

US Patent & Trademark Office

Patent Public Search | Text View

United States Patent Application Publication

20250255912

Kind Code

A1

Publication Date

August 14, 2025

Inventor(s)

Barbon; Elena et al.

LENTIVIRAL VECTOR

Abstract

The present invention relates to lentiviral vectors encoding a progressive familial intrahepatic cholestasis-associated polypeptide. The present invention also relates to cells and pharmaceutical compositions comprising said lentiviral vectors and to uses of said lentiviral vectors in treating progressive familial intrahepatic cholestasis.

Inventors: Barbon; Elena (Cernusco sul Naviglio (Milano), IT), Cantore; Alessio (Milano, IT)

Applicant: Genespire S.r.l. (Milan, IT)

Family ID: 81753254

Appl. No.: 18/856436

Filed (or PCT Filed): April 14, 2023

PCT No.: PCT/EP2023/059848

Foreign Application Priority Data

GB 2205606.3

Apr. 14, 2022

Publication Classification

Int. Cl.: A61K35/76 (20150101); A61K48/00 (20060101); C07K14/47 (20060101); C12N15/86 (20060101)

U.S. Cl.:

CPC A61K35/76 (20130101); A61K48/005 (20130101); C07K14/47 (20130101); C12N15/86 (20130101); C12N2740/16043 (20130101); C12N2830/008 (20130101)

Background/Summary

FIELD OF THE INVENTION

[0001] The present invention relates to lentiviral vectors encoding a progressive familial intrahepatic cholestasis-associated polypeptide. The present invention also relates to cells and pharmaceutical compositions comprising said lentiviral vectors and to uses of said lentiviral vectors in treating progressive familial intrahepatic cholestasis.

BACKGROUND TO THE INVENTION

[0002] Progressive familial intrahepatic cholestasis (PFIC) is a group of rare autosomal recessive liver disorders of childhood characterized by mutations in genes encoding proteins involved in the hepatocellular bile transport

system (see e.g. Baker, A., et al., 2019. Clinics and research in hepatology and gastroenterology, 43(1), pp. 20-36). [0003] Three main subtypes of PFIC (PFIC-1, PFIC-2, and PFIC-3) have been identified. PFIC-1 is caused by mutations in the ATPase phospholipid transporting 8B1 gene (ATP8B1), which encodes a phospholipid-transporting transmembrane P-type adenosine triphosphatase (FIC1). PFIC-2 is caused by mutations in the ATP binding cassette subfamily B member 11 gene (ABCB11), which encodes the bile salt export pump (BSEP), the main transporter of bile acids from hepatocytes to the canalicular lumen. PFIC-3 is caused by mutations in the ATP binding cassette subfamily B member 4 gene (ABCB4), which encodes multidrug-resistance protein 3 (MDR3).

[0004] Patients with PFIC face debilitating symptoms and poor prognosis. The main clinical features of PFIC include cholestasis, jaundice and pruritus, with symptoms typically appearing in infancy or early childhood. PFIC is associated with a range of potentially fatal complications of the liver, including portal hypertension, liver failure, cirrhosis and hepatocellular carcinoma, as well as extrahepatic manifestations (see e.g. Srivastava, A., 2014. Journal of clinical and experimental hepatology, 4(1), pp. 25-36).

[0005] For example, PFIC-2 patients display cholestasis, jaundice and severe itching and usually develop liver fibrosis and end-stage liver disease before adulthood due to massive bile acid accumulation in the hepatic tissue. Current treatment consists of diet control, pharmacological or surgical treatments such as biliary diversions aiming at interrupting the enterohepatic circulation and decreasing bile acid blood concentration. However, most patients ultimately undergo liver transplantation as the only curative option.

[0006] There is a demand for new approaches for treating or preventing PFIC, for example PFIC-2.

SUMMARY OF THE INVENTION

[0007] The present inventors have developed a gene therapy for treating and/or preventing PFIC, for example PFIC-2. The lentiviral vector mediated gene therapy described herein may allow for a stable gene transfer even in paediatric patients at the first disease stages by virtue of lentiviral vector genomic integration and may therefore alleviate symptoms and limit progressive damage to the hepatocytes.

[0008] The present inventors have surprisingly shown that the lentiviral vector mediated gene therapy described herein is safe and efficacious in a model of PFIC-2. The present inventors produced a lentiviral vector encoding a human codon optimized BSEP-coding sequence under the control of a hepatocyte-specific cassette.

Administration of the lentiviral vector resulted in a significant decrease of serum bile acids, and a significant decrease in serum bilirubin and ALP biomarkers, indicating prevention of progressive cholestatic damage.

[0009] In one aspect, the present invention provides a lentiviral vector comprising a nucleotide sequence encoding a progressive familial intrahepatic cholestasis (PFIC)-associated polypeptide. The lentiviral vector may be an immune-shielded lentiviral vector.

[0010] The PFIC-associated polypeptide may be selected from familial intrahepatic cholestasis type 1 (FIC1), or a fragment thereof; bile salt export pump (BSEP), or a fragment thereof; multiple drug resistance 3 (MDR3), or a fragment thereof; tight junction protein 2 (TJP2), or a fragment thereof; farnesoid X receptor (FXR), or a fragment thereof; and Myosin-Vb (MYO5B), or a fragment thereof. In some embodiments, the PFIC-associated polypeptide is selected from BSEP, or a fragment thereof; FXR, or a fragment thereof; and MYO5B, or a fragment thereof. In preferred embodiments, the PFIC-associated polypeptide is BSEP, or a fragment thereof.

[0011] Suitably, the PFIC-associated polypeptide comprises or consists of an amino acid sequence which is at least 70% identical to one of SEQ ID NOs: 37, 39, 43, 45, 47, 48 or 50, or a fragment thereof. In some embodiments, the PFIC-associated polypeptide comprises or consists of an amino acid sequence which is at least 70% identical to SEQ ID NO: 39 or a fragment thereof.

[0012] Suitably, the nucleotide sequence encoding a PFIC-associated polypeptide comprises or consists of a nucleotide sequence which is at least 70% identical to one of SEQ ID NOs: 38, 40, 41, 42, 44, 46, 49 or 51, or a fragment thereof. In some embodiments, the nucleotide sequence encoding a PFIC-associated polypeptide comprises or consists of a nucleotide sequence which is at least 70% identical to one of SEQ ID NOs: 40-42, or a fragment thereof. Suitably, the nucleotide sequence encoding a PFIC-associated polypeptide is codon-optimised. In some embodiments, the nucleotide sequence encoding a PFIC-associated polypeptide comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 41 or 42, or a fragment thereof. In some embodiments, the nucleotide sequence encoding a PFIC-associated polypeptide comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 41, or a fragment thereof.

[0013] The lentiviral vector may be a CD47.sup.high lentiviral vector. Suitably, the lentiviral vector is obtained from a CD47.sup.high host cell, preferably wherein the host cell is genetically engineered to increase expression of CD47 on the cell surface. Suitably, the lentiviral vector has at least about 2-fold more CD47 on its surface than a lentiviral vector obtained from an unmodified host cell.

[0014] The lentiviral vector may be a MHC-I.sup.free lentiviral vector. Suitably, the lentiviral vector is obtained from a MHC-I.sup.free host cell, preferably wherein the host cell is genetically engineered to disrupt expression

of MHC-I on the cell surface. Suitably, the lentiviral vector has less than about 50% of the number of surface-exposed MHC-I molecules that are displayed on a lentiviral vector obtained from an unmodified host cell. Suitably, MHC-I is not detectable on the surface of the lentiviral vector.

[0015] The lentiviral vector may be a CD47.sup.high/MHC-I.sup.free lentiviral vector. Suitably, the lentiviral vector is obtained from a CD47.sup.high/MHC-I.sup.free host cell. Suitably, the lentiviral vector has at least about 2-fold more CD47 on its surface than a lentiviral vector obtained from an unmodified host cell and MHC-I is not detectable on the surface of the lentiviral vector.

[0016] The nucleotide sequence encoding a PFIC-associated polypeptide may be operably linked to one or more miRNA target sequences. In some embodiments, the one or more miRNA target sequences suppress transgene expression in one or more cells other than hepatocytes. In some embodiments, the one or more miRNA target sequence suppress transgene expression in hematopoietic-lineage cells and/or antigen-presenting cells.

[0017] In some embodiments, the one or more miRNA target sequences are selected from miR-181, miR-142, miR-223, and miR-155 target sequences. In some embodiments, the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to one or more mir-142 target sequence, two or more mir-142 target sequences, three or more mir-142 target sequences, or four or more mir-142 target sequences. In some embodiments, the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to four mir-142 target sequences. In some embodiments, the one or more miRNA target sequence comprise or consist of a nucleotide sequence which is at least 90% identical to SEQ ID NO: 17. In some embodiments, the one or more miRNA target sequence comprise or consist of a nucleotide sequence which is at least 90% identical to SEQ ID NO: 18.

[0018] The nucleotide sequence encoding a PFIC-associated polypeptide may be operably linked to a liver-specific promoter. In some embodiments, the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a hepatocyte-specific promoter. Suitably, the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a transthyretin (TTR) promoter, an alpha-1-antitrypsin (AAT) promoter, a thyroxine-binding globulin (TBG) promoter, an APoE/hAAT promoter, a HCR-hAAT promoter, a LP1 promoter, or a HLP promoter. In some embodiments, the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a transthyretin (TTR) promoter. In preferred embodiments, the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to an Enh1mTTR (ET) promoter. In some embodiments, the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a promoter which comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 19.

[0019] The lentiviral vector may be pseudotyped. Suitably, the lentiviral vector is VSV.G-pseudotyped. The lentiviral vector may be a self-inactivating (SIN) lentiviral vector. Suitably, the lentiviral vector comprises self-inactivating (SIN) LTRs which comprise or consist of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 23, or a fragment thereof. The lentiviral vector may be an integrating lentiviral vector and/or a replication-defective lentiviral vector. The lentiviral vector may be HIV-derived.

[0020] In some embodiments, the lentiviral vector comprises a nucleotide sequence which is at least 70% identical to SEQ ID NO: 36.

[0021] In one aspect, the present invention provides a method of producing a lentiviral vector according to the present invention.

[0022] In one aspect, the present invention provides a kit or system for producing a lentiviral vector of the present invention.

[0023] In one aspect, the present invention provides a cell comprising a lentiviral vector according to the present invention. The cell may be an isolated cell.

[0024] In one aspect, the present invention provides a pharmaceutical composition comprising a lentiviral vector or a cell according to the present invention, in combination with a pharmaceutically acceptable carrier, diluent or excipient.

[0025] In one aspect, the present invention provides a lentiviral vector, a cell, or a pharmaceutical composition according to the present invention, for use as a medicament.

[0026] In one aspect, the present invention provides use of a lentiviral vector, a cell, or a pharmaceutical composition according to the present invention, for the manufacture of a medicament.

[0027] In one aspect, the present invention provides a method comprising administering a therapeutically effective amount of a lentiviral vector, a cell, or a pharmaceutical composition according to the present invention, to a subject in need thereof.

[0028] In one aspect, the present invention provides a lentiviral vector, a cell, or a pharmaceutical composition according to the present invention, for use in preventing or treating progressive familial intrahepatic cholestasis (PFIC).

[0029] In one aspect, the present invention provides use of a lentiviral vector, a cell, or a pharmaceutical

composition according to the present invention, for the manufacture of a medicament for preventing or treating progressive familial intrahepatic cholestasis (PFIC).

[0030] In one aspect, the present invention provides a method of preventing or treating progressive familial intrahepatic cholestasis (PFIC), comprising administering a therapeutically effective amount of a lentiviral vector, a cell, or a pharmaceutical composition according to the present invention, to a subject in need thereof.

[0031] Suitably, the PFIC-associated polypeptide is FIC1, or a fragment thereof, and the PFIC is PFIC type 1 (PFIC-1); the PFIC-associated polypeptide is BSEP, or a fragment thereof, and the PFIC is PFIC type 2 (PFIC-2); the PFIC-associated polypeptide is MDR3, or a fragment thereof, and the PFIC is PFIC type 3 (PFIC-3); the PFIC-associated polypeptide is TJP2, or a fragment thereof, and the PFIC is PFIC type 4 (PFIC-4); the PFIC-associated polypeptide is FXR, or a fragment thereof, and the PFIC is PFIC type 5 (PFIC-5); or the PFIC-associated polypeptide is MYO5B, or a fragment thereof, and the PFIC is PFIC type 6 (PFIC-6). In some embodiments, the PFIC-associated polypeptide is BSEP, or a fragment thereof, and the PFIC is PFIC-2; the PFIC-associated polypeptide is FXR, or a fragment thereof, and the PFIC is PFIC-5; or the PFIC-associated polypeptide is MYO5B, or a fragment thereof, and the PFIC is PFIC-6. In preferred embodiments, the PFIC-associated polypeptide is BSEP, or a fragment thereof, and the PFIC is PFIC-2.

[0032] The lentiviral vector, cell, or pharmaceutical composition may be administered to any subject in need thereof. In some embodiments, the subject is a human subject. In some embodiments, the subject is a juvenile. In some embodiments, the subject is a paediatric patient. In some embodiments, the subject is a neonatal patient or an infantile patient.

[0033] The lentiviral vector, cell, or pharmaceutical composition may be administered by any suitable route. In some embodiments, the lentiviral vector, cell, or pharmaceutical composition is administered systemically (e.g. by intravenous injection or intraperitoneal injection). In some embodiments, the lentiviral vector, cell, or pharmaceutical composition is administered locally (e.g. by direct injection, intra-arterial injection, or intraportal injection). In some embodiments, the lentiviral vector, cell, or pharmaceutical composition is administered locally to the liver (e.g. by intrahepatic injection, intrahepatic arterial injection, or intraportal injection).

[0034] The lentiviral vector, cell, or pharmaceutical composition may be administered in any suitable dose. In some embodiments, the lentiviral vector is administered at a dose of at least about 10^{sup.8} TU/kg, at least about 10^{sup.9} TU/kg, or at least about 10^{sup.10} TU/kg. In some embodiments, the lentiviral vector is administered in a dose of from about 10^{sup.8} to about 10^{sup.11} TU/kg, from about 10^{sup.8} to about 10^{sup.10} TU/kg, or from about 10^{sup.9} to about 10^{sup.10} TU/kg.

[0035] The lentiviral vector may integrate into the genome of liver cells and be maintained as the liver cells duplicate. Suitably, the lentiviral vector integrates into the genome of hepatocytes and is maintained as the hepatocytes duplicate.

[0036] Following administration, the subject's serum bile acid levels may be reduced and/or normalised; the subject's liver function may be improved (e.g. the serum level of bilirubin and/or one or more liver enzymes may be reduced and/or normalised); and/or the formation of liver fibrosis in the subject may be slowed and/or reduced.

[0037] In one aspect, the present invention provides an immune-shielded lentiviral vector for use in a method of therapy, wherein the method comprises administration of the immune-shielded lentiviral vector to a juvenile or paediatric subject. In one aspect, the present invention provides a cell for use in a method of therapy, wherein the cell comprises an immune-shielded lentiviral vector, and wherein the method comprises administration of the cell to a juvenile or paediatric subject.

[0038] In one aspect, the present invention provides use of an immune-shielded lentiviral vector for the manufacture of a medicament for treatment or prevention of a disease, wherein the treatment or prevention comprises administration of the immune-shielded lentiviral vector to a juvenile or paediatric subject. In one aspect, the present invention provides use of a cell for the manufacture of a medicament for treatment or prevention of a disease, wherein the cell comprises an immune-shielded lentiviral vector, and wherein the treatment or prevention comprises administration of the cell to a juvenile or paediatric subject.

[0039] In one aspect, the present invention provides a method of treatment, wherein the method comprises administration of an immune-shielded lentiviral vector to a juvenile or paediatric subject. In one aspect, the present invention provides a method of treatment, wherein the method comprises administration of a cell to a juvenile or paediatric subject, wherein the cell comprises an immune-shielded lentiviral vector.

[0040] In one aspect, the present invention provides an immune-shielded lentiviral vector for use in a method of treatment of a juvenile or paediatric subject. In one aspect, the present invention provides a cell for use in a method of treatment of a juvenile or paediatric subject, wherein the cell comprises an immune-shielded lentiviral vector.

[0041] In one aspect, the present invention provides use of an immune-shielded lentiviral vector for the manufacture of a medicament for treatment or prevention of a disease in a juvenile or paediatric subject. In one

aspect, the present invention provides use of a cell for the manufacture of a medicament for treatment or prevention of a disease in a juvenile or paediatric subject, wherein the cell comprises an immune-shielded lentiviral vector.

[0042] In one aspect, the present invention provides a method of treatment of a juvenile or paediatric subject, wherein the method comprises administration of an immune-shielded lentiviral vector to the juvenile or paediatric subject in need thereof. In one aspect, the present invention provides a method of treatment of a juvenile or paediatric subject, wherein the method comprises administration of a cell to the juvenile or paediatric subject in need thereof, wherein the cell comprises an immune-shielded lentiviral vector.

[0043] In some embodiments, the subject is a neonatal subject or an infantile subject.

[0044] In some embodiments, the method comprises transduction of a liver cell with the immune-shielded lentiviral vector

[0045] In some embodiments, the immune-shielded lentiviral vector comprises a nucleotide sequence encoding a progressive familial intrahepatic cholestasis (PFIC)-associated polypeptide.

Description

DESCRIPTION OF DRAWINGS

[0046] FIG. 1—Design and testing of a lentiviral vector (LV) encoding a BSEP transgene (a) Schematic of the third-generation lentiviral vectors (provirus). LTRs, long terminal repeats: the 5' LTR and the 3' LTR each have an almost complete deletion of the U3 region (Δ U3); ψ , the psi packaging sequence; RRE, Rev response element; cPPT, central polypurine tract; ET, enhanced transthyretin promoter; BSEP wt/co, BSEP coding sequences (wild-type or codon optimised); WPRE, posttranscriptional element from the genome of the woodchuck hepatitis virus; 4 \times mir142T, four copies of miR-142 target sequence in tandem. (b) Flow cytometry analysis on Huh7 cells transduced with LV-ET.BSEPwt, LV-ET.BSEPco_1 or LV-ET.BSEPco_2 at different MOI. UT, untransduced Huh7 cells. MFI, mean fluorescence intensity. (c) Histogram reporting the vector copy number (VCN) per cell (VCN/cell) in Huh7 transduced with LV-ET.BSEPwt, LV-ET.BSEPco_1 or LV-ET.BSEPco_2 at different MOI. (d) Histogram reporting the vector copy number (VCN) per cell (VCN/cell) in Huh7 transduced with LV-ET.BSEPco_1 or LV-hAAT.BSEPco_1 at different MOI. (e) Western Blot analysis of BSEP protein expression on Huh7 cell lysate. UT, untransduced Huh7; hAAT, cells transduced with LV-hAAT.BSEPco; ET, cells transduced with LV-ET.BSEPco. (f) BSEP protein expression as determined by WB densitometry analysis and normalized on GAPDH housekeeping gene in Huh7 cells transduced with LV-ET.BSEPco or LV-hAAT.BSEPco.

[0047] FIG. 2—Immunofluorescence analysis of LV-ET.BSEPco_1 in transduced Huh7 cells Left upper panel, Huh7 cells transduced with LV-ET.BSEPco_1 and stained with Hoechst. Right upper panel, Huh7 cells untransduced (Cntrl-) and stained with Hoechst. Left lower panel, Huh7 cells transduced with LV-ET.BSEPco_1 and stained with an anti-BSEP antibody. Right lower panel, Huh7 cells untransduced (Cntrl-) and stained with an anti-BSEP antibody.

[0048] FIG. 3—Evaluation of LV gene therapy in Abcb11 $^{-/-}$ mice (a) Serum bile acids in Abcb11 $^{-/-}$ animals treated with LV-ET.BSEPco (KO+LV) compared to Abcb11 $^{-/-}$ untreated animals (KO UT) and to WT untreated animals (WT UT). (b) Serum total bilirubin in Abcb11 $^{-/-}$ animals treated with LV-ET.BSEPco (KO+LV) compared to Abcb11 $^{-/-}$ untreated animals (KO UT). Dotted line indicates the upper limit of the bilirubin levels (mg/dL) measured in WT control mice. (c) Serum alkaline phosphatase (ALP) levels (IU/L) in Abcb11 $^{-/-}$ animals treated with LV-ET.BSEPco (KO+LV) compared to Abcb11 $^{-/-}$ untreated animals (KO UT) and to WT untreated animals (WT UT). Dotted line indicates the upper limit of the ALP levels measured in a cohort of WT control mice.

DETAILED DESCRIPTION

[0049] It must be noted that as used herein and in the appended claims, the singular forms “a”, “an”, and “the” include plural referents unless the context clearly dictates otherwise.

[0050] The terms “comprising”, “comprises” and “comprised of” as used herein are synonymous with “including”, “includes” or “containing”, “contains”, and are inclusive or open-ended and do not exclude additional, non-recited members, elements or method steps. The terms “comprising”, “comprises” and “comprised of” also include the term “consisting of”.

[0051] Numeric ranges are inclusive of the numbers defining the range. Unless otherwise indicated, any nucleic acid sequences are written left to right in 5' to 3' orientation; amino acid sequences are written left to right in amino to carboxy orientation, respectively.

[0052] The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that such publications constitute prior art to the claims appended hereto.

[0053] The skilled person will understand that they can combine all features of the invention disclosed herein without departing from the scope of the invention as disclosed.

Lentiviral Vector

[0054] In one aspect, the present invention provides a lentiviral vector comprising a protein-coding sequence, wherein the protein-coding sequence is a nucleotide sequence encoding a progressive familial intrahepatic cholestasis (PFIC)-associated polypeptide.

[0055] A “lentiviral vector” may refer to an enveloped lentiviral genome (i.e. a lentiviral particle). For example, the pharmaceutical composition of the present invention preferably comprises the lentiviral vector in the form of a lentiviral particle and the lentiviral vector is preferably administered in the form of a lentiviral particle. In other embodiments, a “lentiviral vector” may comprise a lentiviral genome, optionally wherein the lentiviral genome is enveloped. As used herein, a “lentiviral genome” may refer to a genome that comprises at least one element derived or derivable from a lentivirus genome. Suitably, a lentiviral genome comprises at least one element that is involved in the mechanisms by which a lentivirus infects cells, expresses genes, and/or is replicated.

[0056] Lentivirus is a genus of retroviruses, which contain an RNA genome that is converted to DNA in the transduced cell by a virally encoded reverse transcriptase. Lentiviral vectors can transduce a wide range of cell types and integrate into the host genome in both dividing and post-mitotic cells, resulting in long-term expression of the protein-coding sequence both in vitro and in vivo (see e.g. Tiscornia, G., et al., 2006. *Nature protocols*, 1(1), pp. 241-245).

[0057] The basic genes required for lentivirus survival and function are the gag, pol, and env genes: gag encodes structural proteins; po/encodes enzymes required for reverse transcription and integration into the host cell genome; and env encodes the viral envelope glycoprotein (see e.g. Milone, M.C. and O'Doherty, U., 2018. *Leukemia*, 32(7), pp. 1529-1541). Lentiviruses may also have additional cis-acting elements, such as a rev response element (RRE), which enables the efficient export of RNA transcripts of the integrated provirus from the nucleus to the cytoplasm of an infected target cell; a retroviral psi packaging element, which is involved in regulating the essential process of packaging the retroviral RNA genome into the viral capsid during replication; a primer binding site (PBS), where reverse transcription is initiated; the TAT activation region (TAR); splice donor and acceptor sites; and central and terminal polypurine tracts, which allow initiation of plus-strand synthesis.

[0058] In a lentivirus genome, these elements are typically flanked at both ends by regions called long terminal repeats (LTRs). The LTRs are responsible for integration and transcription. LTRs may also serve as enhancer-promoter sequences and can control the expression of the lentiviral genes. The LTRs themselves are identical or near-identical sequences that can typically be divided into three regions: U3, R and U5. LTRs may be naturally occurring or may be modified. For example, U3 and U5 modifications are described in Iwakuma, T., et al., 1999. *Virology*, 261(1), pp. 120-132.

[0059] The lentiviral vector of the present invention may comprise a minimal lentiviral genome. As used herein, a “minimal lentiviral genome” may mean that the lentiviral genome has been manipulated so as to remove the non-essential elements and to retain the essential elements in order to provide the required functionality to infect, transduce and deliver a nucleotide sequence of interest to a target host cell (see e.g. Kim, V. N., et al., 1998. *Journal of virology*, 72(1), pp. 811-816 and Sertkaya, H., et al., 2021. *Scientific reports*, 11(1), pp. 1-15).

[0060] A lentiviral genome may comprise from 5' to 3': a 5' LTR, one or more lentiviral-derived cis-acting elements, and a 3' LTR. Suitably, a lentiviral genome may comprise from 5' to 3': a 5' LTR, a RRE, and a 3' LTR. Suitably, a lentiviral genome may comprise from 5' to 3': a 5' LTR, a retroviral psi packaging element, a RRE, and a 3' LTR. Suitably, a lentiviral genome may comprise from 5' to 3': a 5' LTR, a retroviral psi packaging element, a RRE, a cPPT, and a 3' LTR. Suitably, a lentiviral genome may comprise from 5' to 3': a 5' LTR, a PBS, a retroviral psi packaging element, a RRE, a cPPT, and a 3' LTR.

[0061] A lentiviral genome may further comprise a protein-coding sequence and, optionally, one or more regulatory elements (e.g. operably linked to the protein-coding sequence). Suitably, a lentiviral genome may comprise from 5' to 3': a 5' LTR, a RRE, a protein-coding sequence, and a 3' LTR. Suitably, a lentiviral genome may comprise from 5' to 3': a 5' LTR, a retroviral psi packaging element, a RRE, a protein-coding sequence, and a 3' LTR. Suitably, a lentiviral genome may comprise from 5' to 3': a 5' LTR, a retroviral psi packaging element, a RRE, a cPPT, a protein-coding sequence, and a 3' LTR. Suitably, lentiviral genome may comprise from 5' to 3': a 5' LTR, a PBS, a retroviral psi packaging element, a RRE, a cPPT, a protein-coding sequence, and a 3' LTR.

[0062] The lentiviral vector of the present invention may be replication-defective. Typically, at least part of one or more protein coding regions essential for replication may be removed from the lentiviral genome. This makes the lentiviral vector “replication-defective” or “replication-incompetent”. Suitably, one or more of gag, pol, rev, and env genes are deleted (at least partially) in a replication-defective lentiviral vector. Suitably, each of the gag, pol, rev, and env genes are deleted (at least partially) in a replication-defective lentiviral vector. Optionally, the lentiviral vector lacks a functional gag-pol and/or env gene and/or other genes essential for replication.

[0063] The lentiviral vector of the present invention may be derived from any lentivirus. As used herein “lentivirus-derived” or “lentivirus-based” may mean that the lentiviral genome comprises one or more elements from said lentivirus. For example, the coding regions of viral proteins may be deleted, but one or more cis-acting element may be retained from said lentivirus.

[0064] The lentiviral vector may be derived from a primate lentivirus. Examples of “primate” lentiviruses include, but are not limited to, human immunodeficiency virus (HIV) and simian immunodeficiency virus (SIV). The lentiviral vector may be derived from a non-primate lentivirus (i.e. derived from a lentivirus which does not primarily infect primates, especially humans). Examples of “non-primate” lentiviruses include, but are not limited to, the prototype “slow virus” visna/maedi virus (VMV), caprine arthritis-encephalitis virus (CAEV), equine infectious anaemia virus (EIAV), feline immunodeficiency virus (FIV), and bovine immunodeficiency virus (BIV).

[0065] Suitably, the lentiviral vector of the present invention is a HIV-derived lentiviral vector. As used herein “HIV-derived” or “HIV-based” may mean that the lentiviral genome comprises one or more element from HIV. For example, the coding regions of HIV viral proteins may be deleted, and one or more HIV cis-acting element may be retained in the lentiviral genome (see e.g. Johnson, N. M., et al., 2021. *Molecular Therapy-Methods & Clinical Development*, 21, pp. 451-465). A HIV-derived lentiviral genome may comprise from 5' to 3': a 5' LTR, one or more HIV-derived cis-acting elements (e.g. RRE and/or cPPT), and a 3' LTR.

[0066] Suitably, the lentiviral vector of the present invention is a HIV-1-derived lentiviral vector. The prototype lentiviral vector system is based on HIV-1 (see e.g. Merten, O. W., et al., 2016. *Molecular Therapy-Methods & Clinical Development*, 3, p. 16017). It has been shown that sequences that extend into the gag open reading frame may be important for packaging of HIV-1. Therefore, HIV-1 vectors often contain the relevant portion of gag in which the translational initiation codon has been mutated. In addition, HIV-1 vectors often also contain a portion of the env gene that includes the RRE. Rev binds to RRE, which permits the transport of full-length or singly spliced mRNAs from the nucleus to the cytoplasm. In the absence of rev and/or a RRE, full-length HIV-1 RNAs may accumulate in the nucleus. Alternatively, a constitutive transport element from certain simple retroviruses such as Mason-Pfizer monkey virus can be used to relieve the requirement for rev and a RRE. A HIV-1-derived lentiviral genome may comprise from 5' to 3': a 5' LTR, one or more HIV-1-derived cis-acting elements (e.g. a PBS, a retroviral psi packaging element, a RRE and/or a cPPT), and a 3' LTR.

[0067] The lentiviral vector of the present invention may be a self-inactivating lentiviral vector. As used herein, “self-inactivating” or “SIN” lentiviral vectors may comprise lentiviral genomes in which the lentiviral enhancer and promoter sequences have been deleted (see e.g. Zufferey, R., et al., 1998. *Journal of virology*, 72(12), pp. 9873-9880 and Miyoshi, H., et al., 1998. *Journal of virology*, 72(10), pp. 8150-8157). SIN lentiviral vectors can be generated and transduce non-dividing cells in vivo with an efficacy similar to that of wild-type vectors. The transcriptional inactivation of the long terminal repeat (LTR) in the SIN provirus can prevent mobilisation by replication-competent virus. This can also enable the regulated expression of genes from internal promoters by eliminating any cis-acting effects of the LTR.

[0068] The lentiviral vector of the present invention may be integration competent. As used herein, an “integration competent” lentiviral vector is capable of integrating into the genome of a host cell. In contrast to integration competent lentiviral vectors, integration defective lentiviral vectors can be produced, for example, either by packaging the lentiviral vector with catalytically inactive integrase (such as an HIV integrase bearing the D64V mutation in the catalytic site) or by modifying or deleting essential att sequences from the lentiviral genome LTR, or by a combination of the above (see e.g. Wanisch, K. and Yáñez-Muñoz, R. J., 2009. *Molecular Therapy*, 17(8), pp. 1316-1332).

[0069] The lentiviral vector of the present invention may be replication-defective and integrating. The lentiviral vector of the present invention may be replication-defective, integrating, and self-inactivating. The lentiviral vector of the present invention may be replication-defective, integrating, self-inactivating, and HIV-derived.

[0070] The lentiviral vector of the present invention may be a lentiviral particle. A “lentiviral particle” may refer to an enveloped lentiviral genome. Lentiviral particles may be generated by co-transfection of a plasmid containing a lentiviral genome (e.g. a “transfer vector”) with helper plasmids (e.g. “packaging vectors” encoding gag-pol and/or rev and “envelope vectors” encoding env) into host cells and harvesting of the lentivirus-containing supernatant afterwards.

[0071] The lentiviral vector of the present invention may be pseudotyped. Pseudotyping lentiviral vectors with naturally occurring or engineered lentiviral envelopes can allow targeted transduction of specific cell types (see e.g. Joglekar, A. V. and Sandoval, S., 2017. *Human Gene Therapy Methods*, 28(6), pp. 291-301). Suitably, the lentiviral vector is pseudotyped to allow transduction of liver cells (e.g. hepatocytes).

[0072] The lentiviral vector of the present invention may be VSV-G pseudotyped. Vesicular stomatitis virus G protein (VSV-G) is a commonly used envelope protein for pseudotyping. VSV-G is a trimeric protein that binds

phosphatidylserine and low-density lipoprotein receptors on a cell surface to endocytose into the cell. VSV-G pseudotyped lentiviral vectors may efficiently transduce liver cells (e.g. hepatocytes).

[0073] The lentiviral vector of the present invention may be replication-defective, integrating, and VSV-G pseudotyped. The lentiviral vector of the present invention may be replication-defective, integrating, self-inactivating, and VSV-G pseudotyped. The lentiviral vector of the present invention may be replication-defective, integrating self-inactivating, HIV-derived, and VSV-G pseudotyped.

Immune-Shielded Lentiviral Vector

[0074] The lentiviral vector of the present invention may be an immune-shielded lentiviral vector.

[0075] As used herein, an “immune shielded lentiviral vector” may refer to a lentiviral vector which is modified to reduce immune responses following administration. For example, the immune-shielding may reduce activation of acute inflammatory response after administration.

[0076] In some embodiments, the lentiviral vector of the present invention is: (i) a CD47.sup.high lentiviral vector; (ii) a MHC-I.sup.free lentiviral vector; and/or (iii) comprises one or more miRNA target sequences (e.g. which suppress transgene expression in antigen-presenting cells). In preferred embodiments, the lentiviral vector of the present invention is a CD47.sup.high/MHC-I.sup.free lentiviral vector and comprises one or more miRNA target sequences (e.g. which suppress transgene expression in antigen-presenting cells). Each of these modifications may act to reduce immune responses following administration.

[0077] An immune-shielded lentiviral vector may be “phagocytosis-shielded” to reduce uptake by professional phagocytes. In some embodiments, the lentiviral vector is a phagocytosis-shielded lentiviral vector. For example, the lentiviral vector may be a CD47.sup.high lentiviral vector.

CD47.SUP.high .Lentiviral Vectors

[0078] The lentiviral vector of the present invention may be a CD47.sup.high lentiviral vector. As used herein, a “CD47.sup.high lentiviral vector” may refer to a lentiviral vector with increased levels of CD47 (or a fragment thereof) on its surface. A CD47.sup.high lentiviral vector may have reduced uptake by professional phagocytes.

[0079] CD47 (Cluster of Differentiation 47) also known as integrin associated protein (IAP) is a transmembrane protein that in humans is encoded by the CD47 gene. Phagocytosis is physiologically inhibited by CD47, which is a ubiquitously expressed ligand of signal regulatory protein α (SIRP- α) receptor, that is expressed by professional phagocytes. CD47 may be incorporated into lentiviral vectors when they bud from producer cells.

[0080] The lentiviral vector of the present invention may comprise one or more CD47 polypeptides (or a fragment thereof) on its surface. The amount of CD47 (or a fragment thereof) on the surface may be enough to reduce uptake by professional phagocytes. Any suitable assay to quantify the amount of CD47 polypeptides (or fragments thereof) present on the surface of the lentiviral vector may be used.

[0081] In some embodiments, the density of CD47 polypeptides (or fragments thereof) may be determined by immunostaining for CD47 and total internal reflection fluorescence microscopy, e.g. as described in US2010/0316570A1. The CD47 polypeptides (or fragments thereof) may be present in a density of at least about 20 molecules/ μm^2 , at least about 25 molecules/ μm^2 , at least about 30 molecules/ μm^2 , at least about 35 molecules/ μm^2 , at least about 40 molecules/ μm^2 , at least about 45 molecules/ μm^2 , at least about 50 molecules/ μm^2 , at least about 60 molecules/ μm^2 , at least about 70 molecules/ μm^2 , at least about 80 molecules/ μm^2 , at least about 90 molecules/ μm^2 , at least about 100 molecules/ μm^2 , at least about 150 molecules/ μm^2 , at least about 200 molecules/ μm^2 , at least about 250 molecules/ μm^2 , at least about 300 molecules/ μm^2 , at least about 350 molecules/ μm^2 , at least about 400 molecules/ μm^2 , at least about 450 molecules/ μm^2 , at least about 500 molecules/ μm^2 , at least about 600 molecules/ μm^2 , at least about 700 molecules/ μm^2 , at least about 800 molecules/ μm^2 , at least about 900 molecules/ μm^2 , or at least about 1000 molecules/ μm^2 . The CD47 polypeptides (or fragments thereof) may be present in a density of about 1000 molecules/ μm^2 or less, about 500 molecules/ μm^2 or less, or about 250 molecules/ μm^2 or less. The CD47 polypeptides (or fragments thereof) may be present in a density of from about 20 molecules/ μm^2 to about 1000 molecules/ μm^2 , from about 20 molecules/ μm^2 to about 500 molecules/ μm^2 , or from about 20 molecules/ μm^2 to about 250 molecules/ μm^2 .

[0082] In some embodiments, the amount of CD47 polypeptides (or fragments thereof) may be determined by immunostaining for CD47 and electron microscopy, as described in Milani, M., et al., 2019. Science Translational Medicine, 11(493), p.eaav7325. The CD47 polypeptides (or fragments thereof) may be detected in an amount of at least about 10 gold particles/lentiviral particle, at least about 15 gold particles/lentiviral particle, or at least about 20 gold particles/lentiviral particle. The CD47 polypeptides (or fragments thereof) may be detected in an amount of about 100 gold particles/lentiviral particle or less, about 80 gold particles/lentiviral particle or less, or about 60 gold particles/lentiviral particle or less. The CD47 polypeptides (or fragments thereof) may be detected in an amount of from about 10 to about 100 gold particles/lentiviral particle, from about 15 to about 80 gold

particles/lentiviral particle, or from about 20 to about 60 gold particles/lentiviral particle.

[0083] The lentiviral vector of the present invention may be obtained from a CD47.sup.high host cell. As used herein, a "CD47.sup.high host cell" may refer to a host cell with increased levels of CD47 (or a fragment thereof) on its surface.

[0084] A CD47.sup.high host cell may be genetically engineered to increase expression of CD47 (or a fragment thereof) on the cell surface. For example, the host cell may comprise a vector encoding CD47 (or a fragment thereof) or may be edited to introduce a nucleotide sequence encoding CD47 (or a fragment thereof) into its genome. Suitably, the host cell is transduced with a viral vector encoding a CD47 polypeptide (or a fragment thereof).

[0085] A CD47.sup.high host cell may have a higher concentration of CD47 (or a fragment thereof) on its surface than an unmodified host cell (e.g. an unmodified producer cell or packaging cell, as described herein). Suitably, the host cell has at least about 2-fold, at least about 3-fold, at least about 4-fold, at least about 5-fold, at least about 10-fold, at least about 20-fold, or at least about 30-fold more CD47 on its cell surface than an unmodified host cell. Suitably, the host cell has from about 5-fold to about 30-fold more CD47 (or a fragment thereof) on its cell surface than an unmodified host cell.

[0086] Suitably, the lentiviral vector of the present has a higher concentration of CD47 (or a fragment thereof) on its surface than a lentiviral vector obtained from an unmodified host cell (e.g. an unmodified producer cell or packaging cell, as described herein). Suitably, the lentiviral vector has at least about 2-fold, at least about 3-fold, at least about 4-fold, at least about 5-fold, at least about 10-fold, at least about 20-fold, at least about 30-fold, at least about 40-fold, or at least about 50-fold more CD47 (or a fragment thereof) on its surface than a lentiviral vector obtained from an unmodified host cell. Suitably, the lentiviral vector has from about 5-fold to about 30-fold more CD47 (or a fragment thereof) on its surface than a lentiviral vector obtained from an unmodified host cell.

[0087] CD47 is a member of the immunoglobulin (Ig) superfamily of membrane proteins, with a single IgV-like domain at its N-terminus, a highly hydrophobic stretch with five membrane-spanning segments and an alternatively spliced cytoplasmic C-terminus ranging in length from 3 to 36 amino acids. Mouse, rat, bovine and human CD47 molecules have been cloned and show about 70% overall amino acid identity (see e.g. Brown, E. J. and Frazier, W. A., 2001. Trends in cell biology, 11(3), pp. 130-135).

[0088] The CD47 polypeptide (or a fragment thereof) may be a human CD47 polypeptide (or a fragment thereof). A CD47 polypeptide may have an amino acid sequence of UniProtKB Q08722.

[0089] Exemplary CD47 polypeptides are provided by SEQ ID NOs: 1-4. Suitably, a CD47 polypeptide comprises an amino acid sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity to any of SEQ ID NOs: 1-4. Suitably, a CD47 polypeptide comprises or consists of the amino acid sequence of any of SEQ ID NOs: 1-4.

TABLE-US-00001 Exemplary CD47 polypeptide (SEQ ID NO: 1)

MWPLVAALLLGSACCGSAQLLFNKTKSVEFTFCNDTVVIPCFVTNMEAQN
TTEVYVKWKFKGRDIYTFDGALNKSTVPTDFSSAKIEVSQLLKGDASLKM
DKSDAVSHTGNYTCEVTELTREGETIIEELKYRVVSWFSPNENILIVIFPI
FAILLFWGQFGIKTLKYRSGGMDEKTIALLVAGLVITVIVIVGAILFVPG
EYSLKNATGLGLIVTSTGILILLHYVVFSTAIGLTSFVIAILVIQVIAYI

LAVVGLSLCIAACIPMHGPLLISGLSILALAQLLGLVYMKFVASNQKTIQ PPRNN Exemplary CD47 polypeptide (SEQ ID NO: 2)

MWPLVAALLLGSACCGSAQLLFNKTKSVEFTFCNDTVVIPCFVTNMEAQN
TTEVYVKWKFKGRDIYTFDGALNKSTVPTDFSSAKIEVSQLLKGDASLKM
DKSDAVSHTGNYTCEVTELTREGETIIEELKYRVVSWFSPNENILIVIFPI
FAILLFWGQFGIKTLKYRSGGMDEKTIALLVAGLVITVIVIVGAILFVPG
EYSLKNATGLGLIVTSTGILILLHYVVFSTAIGLTSFVIAILVIQVIAYI

LAVVGLSLCIAACIPMHGPLLISGLSILALAQLLGLVYMKFVASNQKTIQ PPRKAVEEPLNAFKESKGMMNDE Exemplary CD47 polypeptide (SEQ ID NO: 3)

MWPLVAALLLGSACCGSAQLLFNKTKSVEFTFCNDTVVIPCFVTNMEAQN
TTEVYVKWKFKGRDIYTFDGALNKSTVPTDFSSAKIEVSQLLKGDASLKM
DKSDAVSHTGNYTCEVTELTREGETIIEELKYRVVSWFSPNENILIVIFPI
FAILLFWGQFGIKTLKYRSGGMDEKTIALLVAGLVITVIVIVGAILFVPG
EYSLKNATGLGLIVTSTGILILLHYVVFSTAIGLTSFVIAILVIQVIAYI

LAVVGLSLCIAACIPMHGPLLISGLSILALAQLLGLVYMKFV Exemplary CD47 polypeptide (SEQ ID NO: 4)

MWPLVAALLLGSACCGSAQLLFNKTKSVEFTFCNDTVVIPCFVTNMEAQN
TTEVYVKWKFKGRDIYTFDGALNKSTVPTDFSSAKIEVSQLLKGDASLKM

DKSDAVSHTGNYTCEVTELTREGETIIEELKYYRVVSWFSPNENILIVIFPI
FAILLFWGQFGIKTLKYRSGGMDEKTIALLVAGLVITVIVIVGAILFVPG
EYSLKNATGLGLIVTSTGILILLHYYVFSTAIGLTSFVIAILVIQVIAYI
LAVVGLSLCIAACIPMHGPLLISGLSILALAQLLGLVYMKFVASNQKTIQ PPRKAVEEPLN
[0090] Exemplary CD47 polypeptides excluding the signal peptide are provided by SEQ ID NOs: 5-8. Suitably, a CD47 polypeptide comprises an amino acid sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity to any of SEQ ID NOs: 5-8. Suitably, a CD47 polypeptide comprises or consists of the amino acid sequence of any of SEQ ID NOs: 5-8.

TABLE-US-00002 Exemplary CD47 polypeptide excluding signal peptide (SEQ ID NO: 5)
QLLFNKTKSVEFTFCNDTVPVPCFVTNMEAQNTTEVYVKWKFKGRDIYTF
DGALNKSTVPTDFSSAKIEVSQLLKGDASLKMDKSDAVSHTGNYTCEVTE
LTREGETIIEELKYYRVVSWFSPNENILIVIFPIFAILLFWGQFGIKTLKYR
SGGMDEKTIALLVAGLVITVIVIVGAILFVPGEYSLKNATGLGLIVTSTG
ILILLHYYVFSTAIGLTSFVIAILVIQVIAYILAVVGLSLCIAACIPMHG
PLLISGLSILALAQLLGLVYMKFVASNQKTIQPPRNN Exemplary CD47 polypeptide excluding
signal peptide (SEQ ID NO: 6)
QLLFNKTKSVEFTFCNDTVPVPCFVTNMEAQNTTEVYVKWKFKGRDIYTF
DGALNKSTVPTDFSSAKIEVSQLLKGDASLKMDKSDAVSHTGNYTCEVTE
LTREGETIIEELKYYRVVSWFSPNENILIVIFPIFAILLFWGQFGIKTLKYR
SGGMDEKTIALLVAGLVITVIVIVGAILFVPGEYSLKNATGLGLIVTSTG
ILILLHYYVFSTAIGLTSFVIAILVIQVIAYILAVVGLSLCIAACIPMHG
PLLISGLSILALAQLLGLVYMKFVASNQKTIQPPRKAVEEPLNAFKESKG MMNDE Exemplary CD47
polypeptide excluding signal peptide (SEQ ID NO: 7)
QLLFNKTKSVEFTFCNDTVPVPCFVTNMEAQNTTEVYVKWKFKGRDIYTF
DGALNKSTVPTDFSSAKIEVSQLLKGDASLKMDKSDAVSHTGNYTCEVTE
LTREGETIIEELKYYRVVSWFSPNENILIVIFPIFAILLFWGQFGIKTLKYR
SGGMDEKTIALLVAGLVITVIVIVGAILFVPGEYSLKNATGLGLIVTSTG
ILILLHYYVFSTAIGLTSFVIAILVIQVIAYILAVVGLSLCIAACIPMHG PLLISGLSILALAQLLGLVYMKFV
Exemplary CD47 polypeptide excluding signal peptide (SEQ ID NO: 8)
QLLFNKTKSVEFTFCNDTVPVPCFVTNMEAQNTTEVYVKWKFKGRDIYTF
DGALNKSTVPTDFSSAKIEVSQLLKGDASLKMDKSDAVSHTGNYTCEVTE
LTREGETIIEELKYYRVVSWFSPNENILIVIFPIFAILLFWGQFGIKTLKYR
SGGMDEKTIALLVAGLVITVIVIVGAILFVPGEYSLKNATGLGLIVTSTG
ILILLHYYVFSTAIGLTSFVIAILVIQVIAYILAVVGLSLCIAACIPMHG
PLLISGLSILALAQLLGLVYMKFVASNQKTIQPPRKAVEEPLN

[0091] A person skilled in the art would be able to generate variants and/or fragments based on conservative substitutions and/or the known structural and functional features of CD47. These are described, for instance in Fenalti, G., et al., 2021. Nature communications, 12(1), pp. 1-14.

[0092] Suitably, a fragment of CD47 and/or CD47 variant retains the ability to inhibit phagocytosis. Suitably, a CD47 fragment and/or CD47 variant may comprise the extracellular domain of CD47. The extracellular domain of human CD47 may interact with SIRP- α and inhibit phagocytosis. Optionally, a CD47 fragment and/or CD47 variant comprises the transmembrane domain of CD47. The domains may be linked by inter-domain linker(s). The fragment and/or variant may have at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the activity of a full-length CD47 polypeptide.

[0093] Suitably, a variant of SEQ ID NO: 1 may comprise one or more variation selected from V5I, C14W, C15R, F22L, S27F, F30L, F32Y, T36S, V38L, V38I, F42V, T44A, N50S, T51A, T52S, T52A, V56I, R63K, A71T, S75Y, T76A, P78L, P78S, P78A, S82R, S82N, S83T, K85N, K85E, V88A, V88L, V88I, Q90R, L91F, K93N, M100I, M100V, D101G, K102R, K102T, S107L, I126F, I127V, K130Q, R132H, S138F, V146I, I150V, I153V, S169A, G170R, G170S, G171S, D173Y, I177V, A178G, V181I, V185A, I186V, V188A, I189T, I191V, V198I, A207S, T215I, I219M, Y226C, A231S, T235A, S236F, A240V, A240T, I241V, V243I, I244T, V246L, Y249F, A252S, A252T, V254A, S257T, I264M, I264L, M266I, M266T, M266V, V287I, V292A, N295S, N295D, Q296L, P302S, N304S, and N304D. These are considered to be tolerated, benign, and/or likely benign variations as predicted by SIFT, PolyPhen, CADD, REVEL, and MetaLR.

[0094] Suitably, a variant of SEQ ID NO: 2 may comprise one or more variation selected from P3L, A6P, F22L, S27F, F30L, F32Y, T36S, V38L, V38I, F42V, N50S, T51A, T52S, T52A, V56I, R63K, A71T, S75Y, T76A, P78L, P78S, P78A, S82R, S82N, S83T, K85N, K85E, V88A, V88L, V88I, Q90R, L91F, K93N, M100I, M100V,

D101G, K102R, K102T, I126F, I127V, K130Q, R132H, S138F, V146I, I150V, I153V, S169A, G170R, G170S, G171S, I177V, A178G, V181I, I186V, V188A, I189T, I191V, V198I, A207S, T215I, I219M, Y226C, A231S, T235A, A240V, A240T, I241V, V243I, I244T, V246L, Y249F, A252S, A252T, V254A, I264M, I264L, M266I, M266T, M266V, V287I, V292A, N295S, N295D, and Q296L. These are considered to be tolerated, benign, and/or likely benign variations as predicted by SIFT, PolyPhen, CADD, REVEL, and MetaLR.

[0095] An exemplary CD47 fragment is provided by SEQ ID NO: 9. Suitably, a CD47 fragment comprises an amino acid sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity to SEQ ID NO: 9. Suitably, a CD47 fragment comprises or consists of the amino acid sequence of SEQ ID NO: 9.

TABLE-US-00003 Exemplary CD47 fragment (SEQ ID NO: 9)
MWPLVAALLLGSACCGSAQLLFNKTGSVEFTFCNDTVVIPCFVTNMEAQN
TTEVYVKWKFKGRDIYTFD GALNKSTVPTDFSSAKIEVSQLLKGDASLKM
DKSDAVSHTGNYTCEVTELTREGETIIEELKYRVVSWFSPN

[0096] An exemplary CD47 fragment excluding the signal peptide is provided by SEQ ID NO: 10. Suitably, a CD47 fragment comprises an amino acid sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity to SEQ ID NO: 10. Suitably, a CD47 fragment comprises or consists of the amino acid sequence of SEQ ID NO: 10.

TABLE-US-00004 Exemplary CD47 fragment excluding signal peptide (SEQ ID NO: 10)
QLLFNKTGSVEFTFCNDTVVIPCFVTNMEAQNTTEVYVKWKFKGRDIYTF
D GALNKSTVPTDFSSAKIEVSQLLKGDASLKM DKSDAVSHTGNYTCEVTE
LTREGETIIEELKYRVVSWFSPN

MHC-I.sup.low or MHC-I.sup.free Lentiviral Vectors

[0097] The lentiviral vector of the present invention may be a MHC-I.sup.low lentiviral vector or a MHC-I.sup.free lentiviral vector. In preferred embodiments, the lentiviral vector of the present invention is a MHC-I.sup.free lentiviral vector.

[0098] As used herein, a “MHC-I.sup.low lentiviral vector” may refer to a lentiviral vector with reduced levels of one or more MHC-I molecules on its surface (i.e. reduced levels of surface-exposed MHC-I molecules). The number of surface-exposed MHC-I molecules may be reduced such that the immune response to the MHC-I is decreased to a therapeutically relevant degree.

[0099] As used herein, a “MHC-I.sup.free lentiviral vector” may refer to a lentiviral vector which is substantially devoid of (or free of) one or more MHC-I molecules on its surface (i.e. substantially devoid of (or free of) surface-exposed MHC-I molecules).

[0100] The major histocompatibility complex class I (MHC-I) is a heterodimeric membrane protein that is displayed on the outer leaflet of the cell membrane (see e.g. Penn, D. J. and Ilmonen, P., 2005. Major histocompatibility complex (MHC). eLS). MHC-I functions to bind and display peptide fragments of proteins to the extracellular environment where they may be recognised by CD8⁺ cytotoxic T cells. Peptide fragments generated from normal cellular proteins will not activate cytotoxic T cells due to central and peripheral tolerance mechanisms. However, foreign peptides (e.g. those originating from viral proteins) will cause activation of an immune response to destroy the cell. An allogeneic MHC-I protein itself may be recognised by the immune system. For example, antibodies may bind MHC-I epitopes directly. As a result, lentiviral vectors that comprise MHC-I molecules originating from an allogeneic source may be targeted and neutralised by the immune system.

[0101] The term “MHC-I molecules” may refer to human MHC-I molecules. Human MHC-I is also referred to as human leukocyte antigen class I (HLA-I) and is expressed on almost all nucleated cells. HLA-I consists of two polypeptide chains, an HLA-I heavy chain (α chain) and β 2 microglobulin (β 2M or β chain). The HLA-I α chain and β 2M are linked non-covalently. The HLA-I α chain is polymorphic. Six HLA-I α chains have been identified to date, including three classical, highly polymorphic α chains (HLA-A, HLA-B and HLA-C) and three non-classical, less polymorphic (HLA-E, HLA-F and HLA-G) α chains. The MHC-I molecules may comprise or consist of HLA-A, HLA-B, and HLA-C molecules, which comprise an invariant β 2M sequence.

[0102] The term “MHC-I molecules” may also include variant MHC-I sequences, such as polymorphisms of HLA-I α chain sequences and/or β 2M sequences. For example, variant MHC-I sequences may include HLA-I α chain sequences and/or β 2M sequences with single nucleotide polymorphisms (SNPs) or multiple SNPs.

[0103] Any suitable assay to quantify the amount of MHC-I molecules present on the surface of the lentiviral vector may be used.

[0104] In some embodiments, the amount of MHC-I molecules may be determined by immunostaining for MHC-I and electron microscopy, as described in Milani, M., et al., 2017. EMBO molecular medicine, 9(11), pp. 1558-

1573. The MHC-I molecules may be detected in an amount of less than about 10 gold particles/lentiviral particle, less than about 9 gold particles/lentiviral particle, less than about 8 gold particles/lentiviral particle, less than about 7 gold particles/lentiviral particle, less than about 6 gold particles/lentiviral particle, less than about 5 gold particles/lentiviral particle, less than about 4 gold particles/lentiviral particle, less than about 3 gold particles/lentiviral particle, less than about 2 gold particles/lentiviral particle, less than about 1 gold particle/lentiviral particle, or about 0 gold particles/lentiviral particle. The MHC-I molecules may be undetectable (e.g. the amount of gold particles detected may not be significantly higher than background levels).

[0105] The lentiviral vector of the present invention may be obtained from a MHC-I^{sup}.low host cell or a MHC-I^{sup}.free host cell. In preferred embodiments, the lentiviral vector of the present invention is obtained from a MHC-I^{sup}.free host cell. As used herein, a “MHC-I^{sup}.low host cell” may refer to a host cell with reduced levels of one or more MHC-I molecule on its surface. As used herein, a “MHC-I^{sup}.free host cell” may refer to a host cell which is substantially devoid of or free of one or more MHC-I molecule on its surface.

[0106] A MHC-I^{sup}.low or MHC-I^{sup}.free host cell may be genetically engineered to decrease expression of MHC-I on the cell surface. For example, the cell may comprise a genetically engineered disruption of a gene encoding β 2-microglobulin and/or a genetically engineered disruption of a gene encoding an MHC-I α chain.

[0107] Methods for genetic engineering to decrease protein expression are known in the art. For example, this may be achieved by targeted gene knockout. To decrease protein expression, the gene encoding the protein itself or its regulatory sequence (e.g. its promoter) may be knocked out. Knockout may be achieved by deletion of a section of the coding nucleic acid sequence, which may delete a section of the protein essential for expression or stability, or alter the reading frame of the coding sequence or by base-editing. Suitable methods for targeted gene knockout include use of zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and CRISPR/Cas-based RNA-guided nucleases (see e.g. Gaj, T. et al., 2013. Trends Biotechnol, 31, pp. 397-405). For example, the CRISPR/Cas9 RNA-guided nuclease may be used to catalyse a double strand break at a specific locus in the genome if provided with appropriate RNA guides designed to bind that locus. Cas9 and the guide RNA may be delivered to a target cell by transfection of vectors encoding the protein and RNA. Cells attempt to repair any double strand breaks in their DNA using the non-homologous end joining (NHEJ) pathway. This is an error-prone mechanism which inserts random nucleotides and often disrupts the reading frame of the targeted gene. Alternatively, the genetic engineering to decrease protein expression may be accomplished using RNAi techniques, microRNA or antisense RNA to suppress expression of the target gene.

[0108] Once the targeted gene knockout or suppression of expression approach has been carried out, the resulting population of cells may be screened to select and enrich for those cells exhibiting the phenotype of interest, for example decreased expression of surface-exposed MHC-I. Suitable techniques for screening and enrichment are known in the art and include flow cytometry and fluorescence-activated cell sorting (FACS).

[0109] In some embodiments, the host cell comprises a genetically engineered disruption of a gene encoding β 2-microglobulin. β 2-microglobulin stabilises MHC-I, thus cells deficient in β 2-microglobulin will exhibit decreased expression of MHC-I on the surface of the cell. The cell may comprise genetically engineered disruptions in all copies of the gene encoding β 2-microglobulin.

[0110] In another embodiment, the cell comprises a genetically engineered disruption of one or more gene encoding an MHC-I α chain. The cell may comprise genetically engineered disruptions in all copies of the gene encoding an MHC-I α chain.

[0111] The cell may comprise both genetically engineered disruptions of genes encoding β 2-microglobulin and genetically engineered disruptions of genes encoding an MHC-I α chain.

[0112] Decreased expression of MHC-I on the surface of the cell may refer to a decrease in the number of MHC-I molecules that are expressed on the surface of the cell that has been genetically engineered, in comparison to the number of MHC-I molecules that are expressed on the surface of a cell lacking the genetic engineering, but under otherwise substantially identical conditions. The expression of MHC-I on the surface of the cell may be decreased such that the number of surface-exposed MHC-I molecules is, for example, less than about 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 4%, 3%, 2% or 1% of the number of surface-exposed MHC-I molecules that are displayed in the absence of the genetic engineering. In some embodiments, the expression of MHC-I on the surface of the cell is decreased such that the number of surface-exposed MHC-I molecules is 0% of the number of surface-exposed MHC-I molecules that are displayed in the absence of the genetic engineering.

[0113] The expression of MHC-I on the surface of the cell is preferably decreased such that the cell is substantially devoid of surface-exposed MHC-I molecules. In this context, “substantially devoid” may mean that there is a substantial decrease in the number of MHC-I molecules that are expressed on the surface of the cell that has been genetically engineered, in comparison to the number of MHC-I molecules that are expressed on the surface of a cell lacking the genetic engineering, such that the immune response to MHC-I on lentiviral vectors produced by the cell is decreased to a therapeutically useful degree.

[0114] Suitably, the lentiviral vector of the present invention has a lower concentration of MHC-I molecules on its surface than a lentiviral vector obtained from an unmodified host cell (e.g. an unmodified producer cell or packaging cell, as described herein). Suitably, the lentiviral vector has less than about 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 4%, 3%, 2% or 1% of the number of surface-exposed MHC-I molecules that are displayed on a lentiviral vector obtained from an unmodified host cell. In some embodiments, the lentiviral vector has less than about 20% of the number of surface-exposed MHC-I molecules that are displayed on a lentiviral vector obtained from an unmodified host cell.

[0115] In some embodiments, the lentiviral vector of the present invention is substantially devoid of MHC-I molecules on its surface. In this context, “substantially devoid” may mean that there is no detectable immune response due to the molecules on the surface of the lentiviral vector.

[0116] In some embodiments, the lentiviral vector of the present invention is free of MHC-I molecules on its surface. In this context, “free” may mean that there are no detectable molecules (e.g. by immunostaining and electron microscopy) on the surface of the lentiviral vector. As used herein, “not detectable” may refer to levels which are not statistically significantly different compared to background levels.

[0117] In some embodiments, the lentiviral vector of the present invention has decreased HLA-A, HLA-B, and/or HLA-C molecules on its surface. Suitably, the lentiviral vector has less than about 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 4%, 3%, 2% or 1% of the number of surface-exposed HLA-A molecules that are displayed on a lentiviral vector obtained from an unmodified host cell. Suitably, the lentiviral vector has less than about 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 4%, 3%, 2% or 1% of the number of surface-exposed HLA-B molecules that are displayed on a lentiviral vector obtained from an unmodified host cell. Suitably, the lentiviral vector has less than about 50%, 40%, 30%, 25%, 20%, 15%, 10%, 5%, 4%, 3%, 2% or 1% of the number of surface-exposed HLA-C molecules that are displayed on a lentiviral vector obtained from an unmodified host cell.

[0118] In some embodiments, the lentiviral vector of the present invention is substantially devoid of HLA-A, HLA-B, and/or HLA-C molecules on its surface. In some embodiments, the lentiviral vector of the present invention is substantially devoid of HLA-A, HLA-B, and HLA-C molecules on its surface. In some embodiments, the lentiviral vector of the present invention is free of HLA-A, HLA-B, and/or HLA-C molecules on its surface. In some embodiments, the lentiviral vector of the present invention is free of HLA-A, HLA-B, and HLA-C molecules on its surface.

[0119] As described above, an HLA-I molecule consists of two polypeptide chains, an HLA-I heavy chain (α chain) and β 2 microglobulin (β 2M or β chain). The HLA-I α chain and β 2M are linked non-covalently.

[0120] The skilled person would readily be able to determine amino acid and nucleic acid sequences of HLA-I α chains. For example, the HLA-I α chains may be identified in a genome sequence using their location within the major histocompatibility complex region of the chromosome (see e.g. Penn, D. J. and Ilmonen, P., 2005. Major histocompatibility complex (MHC). eLS).

[0121] HLA-A alpha chains may have an amino acid sequence of UniProtKB P04439. Exemplary HLA-A alpha chains are provided by SEQ ID NOs: 11 and 12. Suitably, an HLA-A alpha chain comprises an amino acid sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity to SEQ ID NO: 11 or 12. Suitably, a HLA-A alpha chain comprises or consists of the amino acid sequence of SEQ ID NO: 11 or 12.

TABLE-US-00005 Exemplary HLA-A alpha chain (SEQ ID NO: 11)

MAVMAPRTL L L L L L S G A L A L T Q T W A G S H S M R Y F F T S V S R P G R G E P R F I A V G
Y V D D T Q F V R F D S D A A S Q R M E P R A P W I E Q E G P E Y W D Q E T R N V K A Q S Q T D R V
D L G T L R G Y Y N Q S E A G S H T I Q I M Y G C D V G S D G R F L R G Y R Q D A Y D G K D Y I A L
N E D L R S W T A A D M A A Q I T K R K W E A A H E A E Q L R A Y L D G T C V E W L R R Y L E N G K
E T L Q R T D P P K T H M T H H P I S D H E A T L R C W A L G F Y P A E I T L T W Q R D G E D Q T Q
D T E L V E T R P A G D G T F Q K W A A V V V P S G E E Q R Y T C H V Q H E G L P K P L T L R W E L
S S Q P T I P V G I I A G L V L L G A V I T G A V V A A V M W R R K S S D R K G G S Y T Q A A S S D S A Q G S D V S L T A C K V

Exemplary HLA-A alpha chain (SEQ ID NO: 12)

MAVMAPRTL L L L L L S G A L A L T Q T W A G S H S M R Y F F T S V S R P G R G E P R F I A V G
Y V D D T Q F V R F D S D A A S Q R M E P R A P W I E Q E G P E Y W D Q E T R N V K A Q S Q T D R V
D L G T L R G Y Y N Q S E A G S H T I Q I M Y G C D V G S D G R F L R G Y R Q D A Y D G K D Y I A L
N E D L R S W T A A D M A A Q I T K R K W E A A H A A E Q Q R A Y L E G R C V E W L R R Y L E N G K
E T L Q R T D P P K T H M T H H P I S D H E A T L R C W A L G F Y P A E I T L T W Q R D G E D Q T Q
D T E L V E T R P A G D G T F Q K W A A V V V P S G E E Q R Y T C H V Q H E G L P K P L T L R W E L
S S Q P T I P V G I I A G L V L L G A V I T G A V V A A V M W R R K S S G G E G V K D R K G G S Y
T Q A A S S D S A Q G S D V S L T A C K V

[0122] HLA-B alpha chains may have an amino acid sequence of UniProtKB P01889. An exemplary HLA-B alpha chain is provided by SEQ ID NO: 13. Suitably, an HLA-B alpha chain comprises an amino acid sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity to SEQ ID NO: 13. Suitably, a HLA-B alpha chain comprises or consists of the amino acid sequence of SEQ ID NO: 13.

TABLE-US-00006 Exemplary HLA-B alpha chain (SEQ ID NO: 13)
MLVMAPRTVLLLLSAALALTETWAGSHSMRYFYTSVSRPGRGEPFRFISVG
YVDDTQFVRFDSDAASPREPRAPWIEQEGPEYWDRNTQIYKAQAQTDRE
SLRNLRGYYNQSEAGSHTLQSMYGC DVGPDGRLLRGHDQYAYDGKDYIAL
NEDLRSWTAADTAAQITQRKWEAAREAEQRRAYLEGECEVWLRRLYLENGK
DKLERADPPKTHVTHHPISDHEATLRCWALGFYPAEITLTWQRDGEDQTQ
DTELVE TRPAGDRTFQKWA AVVVPSGEEQRYTCHVQHEGLPKPLTLRWE P
SSQSTVPIVGIVAGLAVLAVVVIGAVVA VMCR RKSSGGKGGSY SQAACS DSAQGS DVSLTA

[0123] HLA-C alpha chains may have an amino acid sequence of UniProtKB P10321. Exemplary HLA-C alpha chains are provided by SEQ ID NOs: 14 and 15. Suitably, an HLA-C alpha chain comprises an amino acid sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity to SEQ ID NO: 14 or 15. Suitably, a HLA-C alpha chain comprises or consists of the amino acid sequence of SEQ ID NO: 14 or 15.

TABLE-US-00007 Exemplary HLA-C alpha chain (SEQ ID NO: 14)
MRVMAPRALLLLLSGGLALTETWACSHSMRYFDTAVSRPGRGEPFRFISVG
YVDDTQFVRFDSDAASPRGEPRAPWVEQEGPEYWDRETQKYKRQAQADRV
SLRNLRGYYNQSE DG SHTLQRM SC DLGPDGRLLRGYDQSAYDGKDYIAL
NEDLRSWTAADTAAQITQRKLEAARAAEQRLAYLEGTCVEWLRRLYLENGK
ETLQRAEPPKTHVTHHPLSDHEATLRCWALGFYPAEITLTWQRDGEDQTQ
DTELVE TRPAGDGT FQKWA AVVVPSGQE QRYTCHMQHEGLQEPLT LSWEP
SSQPTIPIMGIVAGLAVLVVLAVLGAVVTAMMCR RKSSGGKGGSCSQAAC SNSAQGSDESLITCKA

Exemplary HLA-C alpha chain (SEQ ID NO: 15)
MRVMAPRALLLLLSGGLALTETWACSHSMRYFDTAVSRPGRGEPFRFISVG
YVDDTQFVRFDSDAASPRGEPRAPWVEQEGPEYWDRETQKYKRQAQADRV
SLRNLRGYYNQSE DG SHTLQRM SC DLGPDGRLLRGYDQSAYDGKDYIAL
NEHLRSC TAADTAAQITQRKLEAARAAEQRLAYLEGTCVEWLRRLYLENGK
ETLQRAEPPKTHVTHHPLSDHEATLRCWALGFYPAEITLTWQRDGEDQTQ
DTELVE TRPAGDGT FQKWA AVVVPSGQE QRYTCHMQHEGLQEPLT LRWGG
KGGSCSQAACSN SQAQGSDESLITCKA

[0124] Amino acid and nucleic acid sequences encoding β 2M are also known in the art. For example, a nucleic acid sequence of a human β 2M is deposited as GenBank Accession No. NM_004048.

[0125] An HLA β chain may be that of UniProtKB P61769. An exemplary HLA β chain is provided by SEQ ID NO: 16. Suitably, a HLA β chain comprises an amino acid sequence having at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or about 100% identity to SEQ ID NO: 16. Suitably, a HLA β chain comprises or consists of the amino acid sequence of SEQ ID NO: 16.

TABLE-US-00008 Exemplary HLA beta chain (SEQ ID NO: 16)
MSRSVALAVLALLSLSGLEAIQRTPKIQVYSRHPAENGKSNFLNCYVSGF
HPSDIEVDLLKNGERIEKVEHSDLSFSKDWSFYLLYYTEFTPTTEKDEYAC RVNHVTL SQPKIVKWDRDM

[0126] The lentiviral vector of the present invention may be a CD47.sup.high/MHC-I.sup.free lentiviral vector or a CD47.sup.high/MHC-I.sup.low lentiviral vector. In preferred embodiments, the lentiviral vector of the present invention is a CD47.sup.high/MHC-I.sup.free lentiviral vector.

[0127] The lentiviral vector of the present invention may be obtained from a CD47.sup.high/MHC-I.sup.free host cell or a CD47.sup.high/MHC-I.sup.low host cell. In preferred embodiments, the lentiviral vector of the present invention is obtained from a CD47.sup.high/MHC-I.sup.free host cell.

miRNA Target Sequence

[0128] The lentiviral vector of the present invention may comprise one or more miRNA target sequences. The one or more miRNA target sequences may be operably linked to the protein-coding sequence. The term “operably linked” may mean that the components described are in a relationship permitting them to function in their intended manner.

[0129] MicroRNA (miRNA) genes are scattered across all human chromosomes, except for the Y chromosome. Similar to protein-coding genes, miRNAs are usually transcribed from polymerase-II promoters, generating a so-called primary miRNA transcript (pri-miRNA). From the pri-miRNA, a stem loop of about 60 nucleotides in length, called miRNA precursor (pre-miRNA), is excised leaving a 5' phosphate and a 2 bp long, 3' overhang. The pre-miRNA is then actively transported from the nucleus to the cytoplasm. Then, Dicer performs a double strand cut at the other end of the stem loop, generating a 19-24 bp duplex, which is composed of the mature miRNA and the opposite strand of the duplex, called miRNA*. One strand of the duplex is selectively loaded into the RNA-induced silencing complex (RISC), and accumulates as the mature microRNA. This strand is usually the one whose 5' end is less tightly paired to its complement. However, there are some miRNAs that support accumulation of both duplex strands to similar extent.

[0130] Once loaded into RISC, the guide strand of the mature microRNA interacts with mRNA target sequences preferentially found in the 3' untranslated region (3'UTR) of protein-coding genes. If the whole guide strand sequence is perfectly complementary to the mRNA target, the mRNA is endonucleolytically cleaved. If only the seed sequence (i.e. nucleotides 2-8 counted from the 5' end of the miRNA) is perfectly complementary to the target mRNA, RNAi may act through alternative mechanisms leading to translational repression.

[0131] Expression of the protein from the protein-coding sequence (i.e. "transgene expression") may be regulated by one or more endogenous miRNAs using one or more corresponding miRNA target sequences. Using this method, one or more miRNAs endogenously expressed in a cell prevent or reduce transgene expression in that cell by interacting with its corresponding miRNA target sequence positioned in the lentiviral genome (see e.g. Brown, B. D. et al. (2007) *Nat Biotechnol* 25:1457-1467).

[0132] Suitable miRNA target sequences which suppress transgene expression in specific cells will be known to the skilled person. Determining a miRNA with the desired expression profile may be achieved using techniques known to those skilled in the art. For example, a mammalian microRNA expression atlas is described in Landgraf, P., et al., 2007. *Cell*, 129(7), pp. 1401-1414 and the distribution of miRNA expression across human tissues is described in Ludwig, N., et al., 2016. *Nucleic acids research*, 44(8), pp. 3865-3877. Once a miRNA has been identified, the corresponding target sequence can readily be determined using, for example, a microRNA database, such as miRBase (Griffiths-Jones, S., et al., 2007. *Nucleic acids research*, 36 (suppl_1), pp. D154-D158).

[0133] A miRNA target sequence may be fully or partially complementary to the corresponding miRNA. The term "fully complementary", as used herein, may mean that the target sequence has a nucleic acid sequence which is 100% complementary to the sequence of the miRNA which recognises it. The term "partially complementary", as used herein, may mean that the target sequence is only in part complementary to the sequence of the miRNA which recognises it, whereby the partially complementary sequence is still recognised by the miRNA. In other words, a partially complementary target sequence in the context of the present invention is effective in recognising the corresponding miRNA and effecting prevention or reduction of transgene expression in cells expressing that miRNA. Suitably, a partially complementary miRNA target sequence may be fully complementary to the miRNA seed sequence.

[0134] Including more than one copy of a miRNA target sequence in a lentiviral vector may increase the effectiveness of the system. Also, different miRNA target sequences can be included. For example, the protein-coding sequence may be operably linked to more than one miRNA target sequence, which may or may not be different. The miRNA target sequences may be in tandem, but other arrangements are envisaged. The lentiviral vector may, for example, comprise 1, 2, 3, 4, 5, 6, 7 or 8 copies of the same or different miRNA target sequences. Suitably, the lentiviral vector comprises 4 miRNA target sequences of each miRNA target sequence.

[0135] Copies of miRNA target sequences may be separated by a spacer sequence. A spacer sequence may comprise, for example, at least one, at least two, at least three, at least four or at least five nucleotide bases.

[0136] Suitably, the lentiviral vector comprises one or more miRNA target sequence, two or more miRNA target sequences, three or more miRNA target sequences, or four or more miRNA target sequences. Suitably, the protein-coding sequence is operably linked to one or more miRNA target sequence, two or more miRNA target sequences, three or more miRNA target sequences, or four or more miRNA target sequences. In some embodiments, the protein-coding sequence is operably linked to four miRNA target sequences.

[0137] The miRNA target sequence may be a human miRNA target sequence. Suitably, the miRNA target sequence is a -5p or -3p miRNA target sequence.

[0138] The one or more miRNA target sequence may suppress transgene expression in one or more cells other than liver cells (e.g. hepatocytes).

[0139] The one or more miRNA target sequence may suppress transgene expression in hematopoietic-lineage cells. Hematopoietic stem cells give rise to different types of blood cells, in lines called myeloid and lymphoid. As used herein, "hematopoietic-lineage cells" may include myeloid cells and lymphoid cells. Myeloid cells may include monocytes, macrophages, neutrophils, basophils, and eosinophils. Lymphoid cells may include T cells, B

cells, natural killer cells, and innate lymphoid cells.

[0140] The one or more miRNA target sequences may suppress transgene expression in antigen-presenting cells. As used herein, an “antigen presenting cell” (APC) may refer to a cell that displays antigen bound by major histocompatibility complex (MHC) proteins on its surface. APCs may be hematopoietic-lineage cells. The antigen-presenting cells may be professional antigen-presenting cells. Professional APCs specialize in presenting antigens to T cells and may include macrophages, B cells and dendritic cells. Suitably, the APCs are splenic and/or hepatic APCs.

[0141] The one or more miRNA target sequences may suppress transgene expression in hematopoietic-lineage antigen-presenting cells.

[0142] By preventing transgene expression in antigen-presenting cells, while permitting high levels of expression in other cells, miRNA regulation may enable strong and stable gene transfer in the absence of an immune response.

[0143] As used herein, the term “suppress expression” may refer to a reduction of expression in the relevant cell type(s) of a transgene to which the one or more miRNA target sequence is operably linked as compared to transgene expression in the absence of the one or more miRNA target sequence, but under otherwise substantially identical conditions. In some embodiments, transgene expression is suppressed by at least 50%. In some embodiments, transgene expression is suppressed by at least 60%, at least 70%, at least 80%, at least 90%, or at least 95%. In some embodiments, transgene expression is substantially prevented, e.g. not detectable.

[0144] The miRNA-mediated approach for restricting gene expression has several advantages over other strategies of regulating transgenes. Although using tissue-specific promoters can successfully limit expression to target cells, leaky expression in a fraction of non-target cells is observed. This occurs because the reconstituted promoter, modified for inclusion into a vector system, often loses some of its cell specificity and also because vector integration near active promoters and enhancers can activate the tissue-specific promoter and drive transgene expression. In contrast, because miRNA-mediated silencing occurs at the post-transcriptional level, promoter and enhancer trapping is irrelevant. As such, miRNA-regulation can be used to effectively de-target transgene expression from a particular cell type, while still allowing for broad tissue expression. miRNA regulation may also be used as in combination with tissue-specific promoter/enhancers. By including the miRNA target sequence in expression cassettes already under the control of a tissue-specific promoter, an additional layer of regulation is added which may eliminate off-target expression.

[0145] Exemplary miRNA target sequences which suppress transgene expression in hematopoietic-lineage cells and/or antigen-presenting cells, include, but are not limited to, miR-181, miR-142, miR-223, and miR-155 target sequences. Other miRNA target sequences which suppress transgene expression in hematopoietic-lineage cells and/or antigen-presenting cells are known in the art (see e.g. Ghafouri-Fard, S., et al., 2021. *Non-coding RNA research*, 6(1), pp. 8-14). miRNAs which are expressed in hematopoietic-lineage cells and/or antigen-presenting cells interact with the corresponding miRNA target sequence and reduce the expression of the target gene (see e.g. Brown, B. D., et al., 2006. *Nature medicine*, 12(5), pp. 585-591 and Brown, B. D., et al., 2007. *Nature biotechnology*, 25(12), pp. 1457-1467).

[0146] Further miRNA target sequences that suppress transgene expression in hematopoietic-lineage cells and/or antigen-presenting cells can be identified by any suitable method, for example miRNA expression analysis as described in Monticelli, S., et al., 2005. *Genome biology*, 6(8), pp. 1-15.

[0147] Suitably, the one or more miRNA target sequence comprise or consist of: (i) one or more (e.g. 1, 2, 3, 4, 5, 6, 7, 8) miR-142 target sequence; (ii) one or more (e.g. 1, 2, 3, 4, 5, 6, 7, 8) miR-181 target sequence; (iii) one or more (e.g. 1, 2, 3, 4, 5, 6, 7, 8) miR-223 target sequence; and/or (iv) one or more (e.g. 1, 2, 3, 4, 5, 6, 7, 8) miR-155 target sequence.

[0148] In some embodiments, the one or more miRNA target sequence comprise or consist of: (i) two or more miR-142 target sequences; (ii) two or more miR-181 target sequences; (iii) two or more miR-223 target sequences; and/or (iv) two or more miR-155 target sequences. In some embodiments, the one or more miRNA target sequence comprise or consist of: (i) at least four miR-142 target sequences; (ii) at least four miR-181 target sequences; (iii) at least four miR-223 target sequences; and/or (iv) at least four miR-155 target sequences. In some embodiments, the one or more miRNA target sequence comprise or consist of: (i) four miR-142 target sequences; (ii) four miR-181 target sequences; (iii) four miR-223 target sequences; and/or (iv) four miR-155 target sequences. Suitably, the target sequences are separated by spacer sequences.

[0149] In some embodiments, the one or more miRNA target sequence comprise or consist of one or more (e.g. 1, 2, 3, 4, 5, 6, 7, 8) miR-142 target sequence. In some embodiments, the one or more miRNA target sequence comprise or consist of two or more miR-142 target sequences. In some embodiments, the one or more miRNA target sequence comprise or consist of three or more miR-142 target sequences. In some embodiments, the one or more miRNA target sequence comprise or consist of four or more miR-142 target sequences. In some

embodiments, the one or more miRNA target sequence comprises or consist of four miR-142 target sequences. Suitably, the target sequences are separated by spacer sequences.

[0150] The miR-142 target sequence may be a human miRNA target sequence. Suitably, the miR-142 target sequence is a miR-142-5p or miR-142-3p miRNA target sequence. In some embodiments, the miR-142 target sequence is a miR-142-3p miRNA target sequence.

[0151] In some embodiments, the miR-142 target sequence comprises or consists of a nucleotide sequence which is at least 80% identical to SEQ ID NO: 17 or a fragment thereof. Suitably, the miR-142 target sequence comprises or consists of a nucleotide sequence which is at least 85%, at least 90%, or at least 95% identical to SEQ ID NO: 17 or a fragment thereof.

[0152] In some embodiments, the miR-142 target sequence comprises or consists of the nucleotide sequence SEQ ID NO: 17 or a fragment thereof.

TABLE-US-00009 Exemplary miR-142 target sequence (SEQ ID NO: 17)

TCCATAAAGTAGGAAACACTACA

[0153] In some embodiments, the one or more miRNA target sequence comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 18 or a fragment thereof. Suitably, the one or more miRNA target sequence comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% identical to SEQ ID NO: 18 or a fragment thereof.

[0154] In some embodiments, the one or more miRNA target sequence comprises or consists of the nucleotide sequence of SEQ ID NO: 18 or a fragment thereof.

TABLE-US-00010 Exemplary 4x miR-142 target sequence (SEQ ID NO: 18)

TCCATAAAGTAGGAAACACTACACGATTCCATAAAGTAGGAAACACTACA

ACCGGTTCCATAAAGTAGGAAACACTACATCACTCCATAAAGTAGGAAAC ACTACA

PFIC-Associated Polypeptide

[0155] The protein-coding sequence delivered by the lentiviral vector of the present invention is a nucleotide sequence encoding a progressive familial intrahepatic cholestasis (PFIC)-associated polypeptide. Once delivered to a cell, the protein encoded by the protein-coding sequence may be expressed in the cell (i.e. “transgene expression” may occur).

[0156] As used herein, a “PFIC-associated polypeptide” may be any polypeptide associated with progressive familial intrahepatic cholestasis (PFIC). As described above, PFIC is a group of rare autosomal recessive liver disorders characterized by mutations in hepatocellular transport system genes involved in bile formation. A PFIC-associated polypeptide may be any polypeptide associated with bile formation from hepatocytes (e.g. FIC1, BSEP, MDR3, or FXR). A deficiency of the bile acid export pump (BSEP) activity impairing bile salt handling is seen in several forms of PFIC. A PFIC-associated polypeptide may refer to a hepatic bile acid transporter (e.g. BSEP) or a protein essential for expression and/or function of a hepatic bile acid transporter (e.g. FXR, MYO5B).

[0157] Suitably, the PFIC-associated polypeptide is selected from familial intrahepatic cholestasis type 1 (FIC1), or a fragment and/or variant thereof; bile salt export pump (BSEP), or a fragment and/or variant thereof; multiple drug resistance 3 (MDR3), or a fragment and/or variant thereof; tight junction protein 2 (TJP2), or a fragment and/or variant thereof; farnesoid X receptor (FXR), or a fragment and/or variant thereof; and Myosin-Vb (MYO5B), or a fragment and/or variant thereof.

[0158] The protein-coding sequence may be codon-optimised. For example, the protein-coding sequence may be codon-optimised for expression in a mammalian (e.g. human) cell.

[0159] Different cells differ in their usage of particular codons. This codon bias corresponds to a bias in the relative abundance of particular tRNAs in the cell type. By altering the codons in the sequence so that they are tailored to match with the relative abundance of corresponding tRNAs, it is possible to increase expression. By the same token, it is possible to decrease expression by deliberately choosing codons for which the corresponding tRNAs are known to be rare in the particular cell type. Thus, an additional degree of translational control is available. Codon usage tables are known in the art for mammalian cells (e.g. humans), as well as for a variety of other organisms.

FIC1

[0160] In some embodiments, the PFIC-associated polypeptide is FIC1 or a fragment and/or variant thereof.

[0161] “FIC1” is the abbreviated name of the polypeptide encoded by the ATP8B1 gene and is also known as familial intrahepatic cholestasis type 1 protein, phospholipid-transporting ATPase IC, ATPase class I type 8B member 1, and 4-ATPase flippase complex alpha subunit ATP8B1. FIC1 is a member of the P.sub.4-subfamily of P-type adenosine triphosphatases (ATPases) and transports phospholipids (e.g. phosphatidylserine) from the outer to the inner leaflet of the plasma membrane. FIC1 is localized on the apical membrane of epithelial cells, including the canalicular membrane of hepatocytes (see e.g. Amer, S. and Hajira, A., 2014. Gastroenterology

[0162] A fragment and/or variant of FIC1 may retain FIC1 activity (see e.g. EC 7.6.2.1). For example, a fragment and/or variant of FIC1 may localise on the canalicular membrane of hepatocytes and transport phospholipids from the outer to the inner leaflet of the plasma membrane. Suitably, a fragment and/or variant of FIC1 may have the same or similar activity to FIC1, e.g. may have at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the activity of FIC1.

[0163] A person skilled in the art would be able to generate fragments and/or variants using conservative substitutions, based on the known structural and functional features of FIC1 (see e.g. Klomp, L. W., et al., 2004. *Hepatology*, 40(1), pp. 27-38 and Shin, H. W. and Takatsu, H., 2019. *The FASEB Journal*, 33(3), pp. 3087-3096), and/or based on known variants (see e.g. NCBI Gene ID: 5205 and NCBI HomoloGene: 21151).

[0164] The ATP8B1 gene is conserved in chimpanzee, Rhesus monkey, dog, cow, mouse, and rat. The FIC1 may be a human FIC1. Suitably, the FIC1 may comprise or consist of a polypeptide sequence of UniProtKB accession 043520, or a fragment and/or variant thereof.

[0165] In some embodiments, the FIC1 comprises or consists of an amino acid sequence which is at least 70% identical to SEQ ID NO: 37 or a fragment thereof. Suitably, the FIC1 comprises or consists of an amino acid sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 37 or a fragment thereof.

[0166] In some embodiments, the FIC1 comprises or consists of SEQ ID NO: 37 or a fragment thereof.

TABLE-US-00011 Exemplary FIC1 (SEQ ID NO: 37)

MSTERDSETTFDEDSQPNDEVVPYSDDETEDELDDQGS AVEPEQNRVNRE
AEENREPFRKECTWQVKANDRKYHEQPHFMNTKFLCIKESKYANNAIKTY
KYNAFTFIPMNLFEQFKRAANLYFLALLILQAVPQISTLAWYTTTLVPLL
VLGVTAIKDLVDDVARHKMDKEINNRTCEVIKDG RFKVAKWKEIQVGDVI
RLKKNDFVPADILLSSSEPNSLCYVETAELDGETNLKFKMSLEITDQYL
QREDTLATFDGFIECEEPNNRLDKFTGTLFWRNTSFPLDADKILLRGCVI
RNTDFCHGLVIFAGADTKIMKNSGKTRFKRTKIDYLMNYMVYTIFVVLIL
LSAGLAIGHAYWEAQVGNSSWYLYDGEDDTPSYRGFLIFWGYIIVLNTMV
PISLYVSVEVIRLGQSHFINWDLQMYA EKDTPAKARTTTLNEQLGQIH
YIFSDKTGTLTQNIMTFKKCCINGQIYGDHRDASQHNHNKIEQVDFSWNTY
ADGKLAFYDHYLIEQIQSGKEPEVRQFFFLAVCHTVMVDRTDGQLNYQA
ASPDEGALVNAARNFGFAFLARTQNTITISELG TERTYNVLAILDFNSDR
KRMSIIVRTPEGNIKLYCKGADTVIYERLHRMNPTKQETQDALDIFANET
LRTLCLCYKEIEEKEFTEWNKKEMAASVASTNRDEALDKVYEEIEKDLIL
LGATAIEDKLQDGV PETISKLA KADIKIWVLTGDKKETAENIGFACELT
EDTTICYGEDINSL LHARMENQRNRGGVYAKFAPPVQESFFPPGGNRALI
ITGSWLNEILLEKKT KRNKILKLKFPRT EEERRMRTQSKRRLEAKKEQRQ
KNFVDLACECSAVICCRVTPKQKAMVVDLVKRYKKAITLAIGDGANDVNM
IKTAHIGVGISGQEGMQAVMSSDYSFAQFRYLQRLLL VHGRWSYIRMCKF
LRYFFYKNFAFTLVHFWYSFFNGYSAQTAYEDW FITLYNVLYTSLPVLLM
GLLDQDVSDKLSL RFPGLYIVGQRDLLFNYKRFFV SLLHGVLTSMILFFI
PLGAYLQTVGQDGEAPSDYQSFAVTIASALVITVNFQIGLDTSYWTFVNA
FSIFGSIALYFGIMFDFHSAGIHVLFPSAFQFTGTASNALRQPYIWLTI
LAVAVCLLPVVAIRFLSMTIWPSESDKI QKHRKRLKAE EQWQRRQVFR
GVSTRRSAYAFSHQRGYADLISSGRSIRKKRSPLDAIVADGTAEYRRTGDS

[0167] Suitably, a FIC1 variant may comprise one or more variation selected from: N45T, H78Q, I393V, I577V, S580N, M647T, K814N, R833Q, R930Q, R952Q, and A1152T. These are considered to be benign (or likely benign) variations based on clinical data.

[0168] Suitably, a nucleotide sequence encoding FIC1, or a fragment and/or variant thereof, may comprise or consist of a nucleotide sequence of NCBI reference sequence NM_005603, NM_001374385 or NM_001374386, or a fragment and/or variant thereof.

[0169] In some embodiments, the nucleotide sequence encoding FIC1, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 38 or a fragment thereof. Suitably, the nucleotide sequence encoding FIC1, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 38 or a fragment thereof.

[0170] In some embodiments, the nucleotide sequence encoding FIC1, or a fragment and/or variant thereof, comprises or consists of the nucleotide sequence SEQ ID NO: 38 or a fragment thereof.

TABLE-US-00012 Exemplary nucleotide encoding FIC1 (SEQ ID NO: 38)
ATGAGTACAGAAAGAGACTCAGAAACGACATTTGACGAGGATTCTCAGCC
TAATGACGAAGTGGTTCCCTACAGTGATGATGAAACAGAAGATGAACTTG
ATGACCAGGGGTCTGCTGTTGAACCAGAACAAAACCGAGTCAACAGGGAA
GCAGAGGAGAACCGGGAGCCATTACAGAAAAGAATGTACATGGCAAGTCAA
AGCAAACGATCGCAAGTACCACGAACAACCTCACTTTATGAACACAAAAT
TCTTGTGTATTAAGGAGAGTAAATATGCGAATAATGCAATTAAAACATAC
AAGTACAACGCATTTACCTTTATACCAATGAATCTGTTTGAGCAGTTTAA
GAGAGCAGCCAATTTATATTTCTGGCTCTTCTTATCTTACAGGCAGTTC
CTCAAATCTCTACCCTGGCTTGGTACACCACACTAGTGCCCCTGCTTGTG
GTGCTGGGCGTCACTGCAATCAAAGACCTGGTGGACGATGTGGCTCGCCA
TAAAATGGATAAGGAAATCAACAATAGGACGTGTGAAGTCATTAAGGATG
GCAGGTTCAAAGTTGCTAAGTGGAAAGAAATTCAAGTTGGAGACGTCATT
CGTCTGAAAAAAAATGATTTTGTTCAGCTGACATTCTCCTGCTGTCTAG
CTCTGAGCCTAACAGCCTCTGCTATGTGGAAACAGCAGAACTGGATGGAG
AAACCAATTTAAAATTTAAGATGTCACCTGAAATCACAGACCAGTACCTC
CAAAGAGAAGATACATTGGCTACATTTGATGGTTTTATTGAATGTGAAGA
ACCCAATAACAGACTAGATAAGTTTACAGGAACACTATTTTGGAGAAACA
CAAGTTTTCTTTGGATGCTGATAAAATTTTGTACGTGGCTGTGTAATT
AGGAACACCGATTTCTGCCACGGCTTAGTCATTTTTGCAGGTGCTGACAC
TAAAATAATGAAGAATAGTGGGAAAACCAGATTTAAAAGAACTAAAATTG
ATTACTTGATGAACTACATGGTTTACACGATCTTTGTTGTTCTTATTCTG
CTTTCTGCTGGTCTTGCCATCGGCCATGCTTATTGGGAAGCACAGGTGGG
CAATTCCTCTTGGTACCTCTATGATGGAGAAGACGATACACCCTCCTACC
GTGGATTCCTCATTTTCTGGGGCTATATCATTGTTCTCAACACCATGGTA
CCCATCTCTCTATGTACGCGTGGAAGTGATTCGTCTTGGACAGAGTCA
CTTCATCAACTGGGACCTGCAAATGTACTATGCTGAGAAGGACACACCCG
CAAAAGCTAGAACCACCACACTCAATGAACAGCTCGGGCAGATCCATTAT
ATCTTCTCTGATAAGACGGGGACACTCACACAAAATATCATGACCTTTAA
AAAGTGCTGTATCAACGGGGCAGATATATGGGGACCATCGGGATGCCTCTC
AACACAACCACAACAAAATAGAGCAAGTTGATTTTAGCTGGAATACATAT
GCTGATGGGAAGCTTGCATTTTATGACCACTATCTTATTGAGCAAATCCA
GTCAGGGAAAGAGCCAGAAGTACGACAGTTCTTCTTCTTGCTCGCAGTTT
GCCACACAGTCATGGTGGATAGGACTGATGGTCAGCTCAACTACCAGGCA
GCCTCTCCCGATGAAGGTGCCCTGGTAAACGCTGCCAGGAACCTTTGGCTT
TGCTTCTCCTCGCCAGGACCCAGAACACCATCACCATCAGTGAAGTGGGCA
CTGAAAGGACTTACAATGTTCTTGCCATTTTGGACTTCAACAGTGACCGG
AAGCGAATGTCTATCATTGTAAGAACCCAGAAAGGCAATATCAAGCTTTA
CTGTAAAGGTGCTGACACTGTTATTTATGAACGGTTACATCGAATGAATC
CTACTAAGCAAGAAACACAGGATGCCCTGGATATCTTTGCAAATGAAACT
CTTAGAACCTATGCCTTTGCTACAAGGAAATTGAAGAAAAAGAATTTAC
AGAATGGAATAAAAAGTTTATGGCTGCCAGTGTGGCCTCCACCAACCGGG
ACGAAGCTCTGGATAAAGTATATGAGGAGATTGAAAAAGACTTAATTCTC
CTGGGAGCTACAGCTATTGAAGACAAGCTACAGGATGGAGTTCCAGAAAC
CATTTCAAAACTTGCAAAAGCTGACATTAAGATCTGGGTGCTTACTGGAG
ACAAAAAGGAAACTGCTGAAAATATAGGATTTGCTTGTGAACCTTCTGACT
GAAGACACCACCATCTGCTATGGGGAGGATATTAATTCTCTTCTTCATGC
AAGGATGGAAAACCAGAGGAATAGAGGTGGCGTCTACGCAAAGTTTGCAC
CTCCTGTGCAGGAATCTTTTTTTTCCACCCGGTGGAACCGTGCCTTAATC
ATCACTGGTTCTTGGTTGAATGAAATTCTTCTCGAGAAAAAGACCAAGAG
AAATAAGATTCTGAAGCTGAAGTTCCCAAGAACAGAAGAAGAAAGACGGA
TGCGGACCCAAAGTAAAAGGAGGCTAGAAGCTAAGAAAGAGCAGCGGCAG
AAAAACTTTGTGGACCTGGCCTGCGAGTGCAGCGCAGTCATCTGCTGCCG
CGTACCCCCCAAGCAGAAGGCCATGGTGGTGGACCTGGTGAAGAGGTACA
AGAAAGCCATCACGCTGGCCATCGGAGATGGGGCCAATGACGTGAACATG
ATCAAAACTGCCACATTGGCGTTGGAATAAGTGGACAAGAAGGAATGCA
AGCTGTCATGTCGAGTGACTATTCCTTTGCTCAGTTCCGATATCTGCAGA

GGCTACTGCTGGTGGCGGATGGTCTTACATAAGGATGTGCAAGTTC
CTACGATACTTCTTTTACAAAACTTTGCCTTTACTTTGGTTCATTTCTG
GTACTCCTTCTTCAATGGCTACTCTGCGCAGACTGCATACGAGGATTGGT
TCATCACCTCTACAACGTGCTGTACACCAGCCTGCCCGTGCTCCTCATG
GGGCTGCTCGACCAGGATGTGAGTGACAACTGAGCCTCCGATTCCCTGG
GTTATACATAGTGGGACAAAGAGACTTACTATTCAACTATAAGAGATTCT
TTGTAAGCTTGTTGCATGGGGTCCTAACATCGATGATCCTCTTCTTCATA
CCTCTTGGAGCTTATCTGCAAACCGTAGGGCAGGATGGAGAGGCACCTTC
CGACTACCAGTCTTTTGCCGTCACCATTGCCTCTGCTCTTGTAATAACAG
TCAATTTCCAGATTGGCTTGGATACTTCTTATTGGACTTTTGTGAATGCT
TTTTCAATTTTGGGAAGCATTGCACTTTATTTTGGCATCATGTTTGACTT
TCATAGTGCTGGAATACATGTTCTCTTTCCATCTGCATTTCAATTTACAG
GCACAGCTTCAAACGCTCTGAGACAGCCATACATTTGGTTAACTATCATC
CTGGCTGTTGCTGTGTGCTTACTACCCGTCGTTGCCATTTCGATTCTGTC
AATGACCATCTGGCCATCAGAAAGTGATAAGATCCAGAAGCATCGCAAGC
GGTTGAAGGCGGAGGAGCAGTGGCAGCGACGGCAGCAGGTGTTCCGCCGG
GGCGTGTCAACGCGGCGCTCGGCCTACGCCTTCTCGCACCAGCGGGGCTA
CGCGGACCTCATCTCCTCCGGGCGCAGCATCCGCAAGAAGCGCTCGCCGC
TTGATGCCATCGTGCGGATGGCACCGCGGAGTACAGGCGCACCGGGGAC AGCTGA
BSEP

[0171] In some embodiments, the PFIC-associated polypeptide is BSEP or a fragment and/or variant thereof.

[0172] “BSEP” is the abbreviated name of the polypeptide encoded by the ABCB11 gene and is also known as Bile salt export pump, ATP-binding cassette sub-family B member 11, ABCB11, ABC16, BRIC2, PFIC-2, PFIC2, PGY4, and SGP. BSEP catalyses the transport of the major hydrophobic bile salts, such as taurine and glycine-conjugated cholic acid across the canalicular membrane of hepatocytes in an ATP-dependent manner. BSEP therefore participates in hepatic bile acid homeostasis and consequently to lipid homeostasis through regulation of biliary lipid secretion in a bile salts dependent manner.

[0173] A fragment and/or variant of BSEP may retain BSEP activity. For example, a fragment and/or variant of BSEP may transport hydrophobic bile salts (e.g. taurine and glycine-conjugated cholic acid) across the canalicular membrane of hepatocytes. Suitably, a fragment and/or variant of BSEP may have the same or similar activity to BSEP, e.g. may have at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the activity of BSEP. Assays to determine BSEP activity (e.g. bile acid transport function) are described, for instance in Hayashi, H., et al., 2005. *Hepatology*, 41(4), pp. 916-924 and Noé, J., et al., 2002. *Gastroenterology*, 123(5), pp. 1659-1666.

[0174] A “fragment of BSEP” may refer to a portion or region of full-length BSEP that has the same or similar activity as full-length BSEP, i.e. the fragment may be a functional fragment (e.g. the fragment may transport hydrophobic bile salts across the canalicular membrane of hepatocytes). The fragment may have at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the activity of full-length BSEP.

[0175] A “BSEP variant” may include an amino acid sequence or a nucleotide sequence which may be at least 50%, at least 55%, at least 65%, at least 70%, at least 75%, at least 75%, at least 80%, at least 85% or at least 90% identical, optionally at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to a wild-type BSEP. BSEP variants may have the same or similar activity to wild-type BSEP (e.g. BSEP variants may transport hydrophobic bile salts across the canalicular membrane of hepatocytes). BSEP variants may have at least at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the activity of a wild-type BSEP.

[0176] A person skilled in the art would be able to generate a fragment of BSEP and/or BSEP variants using conservative substitutions, based on the known structural and functional features of BSEP (see e.g. Wang, L., et al., 2020. *Cell Research*, 30(7), pp. 623-625 and Pauli-Magnus, C., et al., 2004. *Hepatology*, 39(3), pp. 779-791) and/or based on known variants (see e.g. NCBI Gene ID: 8647 and NCBI HomoloGene: 74509).

[0177] Suitably, a fragment of BSEP and/or a BSEP variant comprises two homologous halves, each comprising a hydrophobic membrane-anchoring domain (e.g. an ABC transporter transmembrane region) and an ATP binding cassette (ABC) domain. The domains/sequences may be linked by inter-domain linker(s).

[0178] The ABCB11 gene is conserved in chimpanzee, Rhesus monkey, dog, cow, mouse, and rat. The BSEP may be a human BSEP. Suitably, the BSEP may comprise or consist of a polypeptide sequence of UniProtKB accession 095342, or a fragment and/or variant thereof.

[0179] In some embodiments, the BSEP comprises or consists of an amino acid sequence which is at least 70%

identical to SEQ ID NO: 39 or a fragment thereof. Suitably, the BSEP comprises or consists of an amino acid sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 39 or a fragment thereof.

[0180] In some embodiments, the BSEP comprises or consists of SEQ ID NO: 39 or a fragment thereof.

TABLE-US-00013 Exemplary BSEP (SEQ ID NO: 39)

MSDSVILRSIKKFGGEENDGFESDKSYNNDDKKSRLQDEKKGDGVRVGVFFQL
FRFSSSTDIWLMFVGLCAFLHGIAQPGVLLIFGTMTDVFIDYDVELQEL
QIPGKACVNNITIVWTNSSLNQNMNTNGTRCGLLNIESEMIKFASYAGIAV
AVLITGYIQICFWVIAAARQIQKMRKFYFRRIMRMEIGWFDCNSVGELNT
RFSDDINKINDAIADQMALFIQRMTSTICGFLLGFFRGWKLTLVIISVSP
LIGIGAATIGLSVSKFTDYELKAYAKAGVVADEVISSMRTVAAFGGEKRE
VERYEKNLVFAQRWGIRKGIVMGFFTGFVWCLIFLCYALAFWYGSTLVLD
EGEYTPGTLVQIFLSVIVGALNLGNASPCLEAFATGRAAATSIFETIDRK
PIIDCMSEDGYKLDRIKGEIEFHNVTFHYPSPREVKILNDLNMVIKPGEM
TALVGPSGAGKSTALQLIQRFYDPCEGMVTVDGHDIRSLNIQWLRDQIGI
VEQEPVLFSTTIAENIRYGREDATMEDIVQAAKEANAYNFIMDLPQQFDT
LVGEGGGQMSGGQKQRVAIARALIRNPKILLDMATSALDNESEAMVQEV
LSKIQHGHHTIISVAHRLSTVRAADTIIGFEHGTAVERGTHEELLERKGVY
FTLVTLQSQGNQALNEEDIKDATEDDMLARTFSRGSYQDSLRSIRQRSK
SQLSYLVHEPPLAVVDHKSTYEEDRKDKDIPVQEEVEPAPVRRILKFSAP
EWPYMLVGSVGA AVNGTVTPLYAFLFSQILGTFSIPDKEEQRSQINGVCL
/LFVAMGCVSLFTQFLQGYAFAKSGELLTKRLRKFGFRAMLGQDIAWFDD
LRNSPGALTTRLATDASQVQGAAGSQIGMIVNSFTNVTVMIIAFSFSWK
LSLVILCFFPFLALSGATQTRMLTGFASRDKQALEMVGQITNEALSNIRT
VAGIGKERRFIEALETELEKPFKTAIQKANIYGFCFAFAQCIMFIANSAS
YRYGGYLISNEGLHFSYVFRVISAVVLSATALGRAFSYTPSYAKAKISAA
RFFQLLDRQPPISVYNTAGEKWDNFQKGIDFVDCKFTYPSRPDSQVLNGL
SVSISPGQTLAFVGSSGCGKSTSIQLLERFYDPDQGKVMIDGHDSKKVNV
QFLRSNIGIVSQEPVLFACSIMDNIKYGDNTKEIPMERVIAAAKQAQLHD
FVMSLPEKYETNVGSQGSQLSRGEKQRIAIARAIVRDPKILLLDEATSAL
DTESEKTVQVALDKAREGRTCIVIAHRLSTIQNADIIVMAQGVVIEKGT
HEELMAQKGAYYKLVTTGSPIS

[0181] Suitably, a BSEP variant may comprise one or more variation selected from R8Q, K12E, F13S, F13Y, E15Q, D18H, G19D, G19R, E21D, E21Q, S22P, D23H, D23N, K24E, Y26H, Y26N, D29G, K31T, G40V, G40D, G40S, V43F, V43I, R44T, V45I, S56L, S56P, T57A, I59L, V64A, I74T, V79M, V95A, V95I, I102F, G104E, V108M, N120H, Q121R, Q121K, M123R, M123T, M123L, M123V, T124K, R128H, N133S, I139T, S143I, A151V, L153F, I160V, I165T, I171T, M174L, N193S, F202L, I206V, N207S, A218S, A218T, M224L, F236S, I245V, I254M, T258A, S262T, K265R, M288I, M288V, R299K, V301A, V301L, V301I, K306E, V309A, R313H, R313S, I320L, V321A, M322L, F324L, V329M, C336S, G352E, G352R, T355A, I362V, L364F, N375S, A384T, I393V, R399S, R399T, I402T, I402V, R415Q, H428R, E434D, V444A, P447S, M450T, M478T, M478K, E521K, D522N, T524S, M525I, V529I, D543G, D543A, D543N, Q546K, G557S, V567I, K578R, M596I, M596V, E599Q, I610T, V613I, I627T, A634V, R637S, R637K, T639N, E641K, E642D, E645G, R646K, T652A, Q659R, Q662L, Q662E, A663D, A663T, L664F, L664V, E666K, I669M, K670N, D675G, M677V, A679V, A679E, R680K, R680G, T681I, T681N, F682I, S686R, Q688R, A693T, R698H, E709K, P710L, P711A, L712S, A713V, A713S, V714F, V715L, V715I, D716G, D716A, H717R, H717N, T720S, T720A, Y721C, E722D, E722G, D724A, R725T, R725G, D727E, D729V, I730S, I730T, I730L, V732A, V732L, V732M, P740S, V741F, V741I, A749V, M755I, M755L, M755L, F774L, F774I, Q778E, Q778K, I785V, P786L, P786S, K788N, E790Q, N796D, A804V, F834L, A838T, G841R, R851K, S853N, I879M, V886I, M890I, F908V, A916V, T917A, N941S, S945G, R958Q, I960V, T965S, T965A, E966K, K969N, K969E, F971L, F971Y, Q976R, I991V, I1007V, N1009K, N1009S, N1009H, P1060S, V1063A, Y1064F, A1067G, E1069D, N1073S, K1077N, K1077E, V1081I, K1084T, K1136R, V1147L, V1147I, K1181R, E1182D, P1184S, M1185K, A1190T, S1203A, N1211D, V1259I, K1263E, M1289I, A1290T, V1293M, T1299S, E1301A, E1302D, Q1306R. These are considered to be tolerated, benign, and/or likely benign variations as predicted by SIFT, PolyPhen, CADD, REVEL, and MetaLR.

[0182] Suitably, a BSEP variant may comprise one or more variation selected from: V95I, I206V, N375S, V444A, Q546K, N591S, R646K, A663T, M677V, R698H, S701P, A865V, R1057Q, and E1186K. These are considered to be benign (or likely benign) variations based on clinical data.

[0183] Suitably, a nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, may comprise or

consist of a nucleotide sequence of NCBI reference sequence NM_003742, or a fragment and/or variant thereof. [0184] In some embodiments, the nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 40 or a fragment thereof. Suitably, the nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 75%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 40 or a fragment thereof.

[0185] In some embodiments, the nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, comprises or consists of the nucleotide sequence SEQ ID NO: 40 or a fragment thereof.

TABLE-US-00014 Exemplary nucleotide encoding BSEP (SEQ ID NO: 40)

ATGTCTGACTCAGTAATTCTTCGAAGTATAAAGAAATTTGGAGAGGAGAATGATGGTTTTGAGTCAGAT
AAATCATATAATAATGATAAGAAATCAAGGTTACAAGATGAGAAGAAAGGTGATGGCGTTAGAGTTGGC
TTCTTTCAATTGTTTTCGGTTTTCTTCATCAACTGACATTTGGCTGATGTTTGTGGGAAGTTTGTGTGCA
TTTCTCCATGGAATAGCCCAGCCAGGCGTGCTACTCATTTTTTGGCACAATGACAGATGTTTTTATTGAC
TACGACGTTGAGTTACAAGAACTCCAGATTCCAGGAAAAGCATGTGTGAATAACACCATTGTATGGACT
AACAGTTCCTCAACCAGAACATGACAAATGGAACACGTTGTGGGTTGCTGAACATCGAGAGCGAAATG
ATCAAATTTGCCAGTTACTATGCTGGAATTGCTGTGCGCAGTACTTATCACAGGATATATTCAAATATGC
TTTTGGGTCATTGCCGCAGCTCGTCAGATACAGAAAATGAGAAAATTTTACTTTAGGAGAATAATGAGA
ATGGAAATAGGGTGGTTTGACTGCAATTCAGTGGGGGAGCTGAATACAAGATTCTCTGATGATATTAAT
AAAATCAATGATGCCATAGCTGACCAAATGGCCCTTTTCATTTCAGCGCATGACCTCGACCATCTGTGGT
TTCCTGTTGGGATTTTTTCAGGGGTTGGAACTGACCTTGTTTATTATTTCTGTTCAGCCCTCTCATTGGG
ATTGGAGCAGCCACCATTGGTCTGAGTGTGTCCAAGTTTACGGACTATGAGCTGAAGGCCTATGCCAAA
GCAGGGGTGGTGGCTGATGAAGTCATTTTCATCAATGAGAACAGTGGCTGCTTTTGGTGGTGAGAAAAGA
GAGGTTGAAAGGTATGAGAAAAATCTTGTGTTTCGCCAGCGTTGGGGAATTAGAAAAGGAATAGTGATG
GGATTCTTTACTGGATTTCGTGTGGTGTCTCATCTTTTTGTGTTATGCACTGGCCTTCTGGTACGGCTCC
ACACTTGTCTGGATGAAGGAGAATATACACCAGGAACCTTGTCCAGATTTTCCTCAGTGTCATAGTA
GGAGCTTTAAATCTTGGCAATGCCTCTCCTTGTTTGGAAGCCTTTGCAACTGGACGTGCAGCAGCCACC
AGCATTTTTGAGACAATAGACAGGAAACCCATCATTGACTGCATGTCAGAAGATGGTTACAAGTTGGAT
CGAATCAAGGGTGAAATTGAGTTCCATAATGTGACCTTCCATTATCCTTCCAGACCAGAGGTGAAGATT
CTAAATGACCTCAACATGGTCATTAAACCAGGGGAAATGACAGCTCTGGTAGGACCCAGTGGAGCTGGA
AAAAGTACAGCACTGCAACTCATTTCAGCGATTCTATGACCCCTGTGAAGGAATGGTGACCGTGGATGGC
CATGACATTTCGCTCTCTTAACATTCAGTGGCTTAGAGATCAGATTGGGATAGTGGAGCAAGAGCCAGTT
CTGTTCTCTACCACCATTGCAGAAAATATTCGCTATGGCAGAGAAGATGCAACAATGGAAGACATAGTC
CAAGCTGCCAAGGAGGCCAATGCCTACAACCTTCATCATGGACCTGCCACAGCAATTTGACACCCTTGTT
GGAGAAGGAGGAGGCCAGATGAGTGGTGGCCAGAAACAAAGGGTAGCTATCGCCAGAGCCCTCATCCGA
AATCCCAAGATTCTGCTTTTTGGACATGGCCACCTCAGCTCTGGACAATGAGAGTGAAGCCATGGTGCAA
GAAGTGCTGAGTAAGATTCAGCATGGGCACACAATCATTTTCAGTTGCTCATCGCTTGTCTACGGTCAGA
GCTGCAGATACCATCATTGGTTTTGAACATGGCACTGCAGTGGAAAGAGGGACCCATGAAGAATTACTG
GAAAGGAAAGGTGTTTACTTCACTCTAGTGACTTTGCAAAGCCAGGGAAATCAAGCTCTTAATGAAGAG
GACATAAAGGATGCAACTGAAGATGACATGCTTGCGAGGACCTTTAGCAGAGGGAGCTACCAGGATAGT
TTAAGGGCTTCCATCCGGCAACGCTCCAAGTCTCAGCTTTCTTACCTGGTGCACGAACCTCCATTAGCT
GTTGTAGATCATAAGTCTACCTATGAAGAAGATAGAAAGGACAAGGACATTCCTGTGCAGGAAGAAGTT
GAACCTGCCCCAGTTAGGAGGATTCTGAAATTCAGTGCTCCAGAATGGCCCTACATGCTGGTAGGGTCT
GTGGGTGCAGCTGTGAACGGGACAGTCACACCCTTGATGCCTTTTTTATTTCAGCCAGATTCTTGGGACT
TTTTCAATTCCTGATAAAGAGGAACAAAGGTCACAGATCAATGGTGTGTGCCTACTTTTTGTAGCAATG
GGCTGTGTATCTCTTTTACCCAATTTCTACAGGGATATGCCTTTGCTAAATCTGGGGAGCTCCTAACA
AAAAGGCTACGTAAATTTGGTTTCAGGGCAATGCTGGGGCAAGATATTGCCTGGTTTGATGACCTCAGA
AATAGCCCTGGAGCATTGACAACAAGACTTGCTACAGATGCTTCCCAAGTTCAAGGGGCTGCCGGCTCT
CAGATCGGGATGATAGTCAATTCCTTCACTAACGTCAGTGTGGCCATGATCATTGCCTTCTCCTTTAGC
TGGAAGCTGAGCCTGGTCATCTTGCTTCTTCCCCTTCTTGGCTTTATCAGGAGCCACACAGACCAGG
ATGTTGACAGGATTTGCCTCTCGAGATAAGCAGGCCCTGGAGATGGTGGGACAGATTACAAATGAAGCC
CTCAGTAACATCCGCACTGTTGCTGGAATTGGAAAGGAGAGGCGGTTTCATTGAAGCACTTGAGACTGAG
CTGGAGAAGCCCTTCAAGACAGCCATTCAGAAAGCCAATATTTACGGATTCTGCTTTGCCTTTGCCAG
TGCATCATGTTTATTGCGAACTCTGCTTCCTACAGATATGGAGGTTACTTAATCTCCAATGAGGGGCTC
CATTTTCAGCTATGTGTTTCAGGGTGATCTCTGCAGTTGTACTGAGTGCAACAGCTCTTGGAAGAGCCTTC
TCTTACACCCCAAGTTATGCAAAAGCTAAAATATCAGCTGCACGCTTTTTTCAACTGCTGGACCGACAA
CCCCCAATCAGTGTATACAATACTGCAGGTGAAAAATGGGACAACCTTCAGGGGAAGATTGATTTTGT

GATTGTAATTACATCTTCTCCTCGACCTCGCAAGTTCTGAATCGTCTCAGTGTCTGCTAGT
CCAGGGCAGACACTGGCGTTTGTGGGAGCAGTGGATGTGGCAAAGCACTAGCATTGCTGTTGGAA
CGTTTCTATGATCCTGATCAAGGGAAGGTGATGATAGATGGTCATGACAGCAAAAAAGTAAATGTCCAG
TTCCTCCGCTCAAACATTGGAATTGTTTCCCAGGAACCAAGTGTGTTTGCCTGTAGCATAATGGACAAT
ATCAAGTATGGAGACAACACCAAAGAAATTCCTCATGGAAAGAGTCATAGCAGCTGCAAAACAGGCTCAG
CTGCATGATTTTGTTCATGTCACTCCCAGAGAAATATGAACTAACGTTGGGTCCCAGGGGTCTCAACTC
TCTAGAGGGGAGAAACAACGCATTGCTATTGCTCGGGCCATTGTACGAGATCCTAAAATCTTGCTACTA
GATGAAGCCACTTCTGCCTTAGACACAGAAAGTGAAAAGACGGTGCAGGTTGCTCTAGACAAAGCCAGA
GAGGGTCGGACCTGCATTGTTCATTGCCCATCGCTTGTCACCATCCAGAACGCGGATATCATTGCTGTC
ATGGCACAGGGGGTGGTGATTGAAAAGGGGACCCATGAAGAACTGATGGCCCAAAAAGGAGCCTACTAC
AACTAGTCACTGATCCCCCATCAGTTGA

[0186] In some embodiments, the nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, is codon optimised. Exemplary codon-optimised sequences are provided in SEQ ID NOs: 41 and 42.

[0187] In some embodiments, the nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 41 or a fragment thereof. Suitably, the nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 75%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 41 or a fragment thereof.

[0188] In some embodiments, the nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, comprises or consists of the nucleotide sequence SEQ ID NO: 41 or a fragment thereof.

TABLE-US-00015 Exemplary codon-optimised nucleotide encoding BSEP (SEQ ID NO: 41)
ATGAGCGACTCCGTGATCCTGCGGAGCATCAAGAAGTTTGGCGAGGAGAACGACGGCTTCGAGTCTGAT
AAGAGCTACAACAATGATAAGAAGTCCAGACTGCAGGACGAGAAGAAGGGCGATGGCGTGCGGGTGGGC
TTCTTTCAGCTGTTCCGGTTCAGCAGCAGCACCGACATCTGGCTGATGTTTGTGGGCTCCCTGTGCGCC
TTCCTGCACGGAATCGCACAGCCAGGCGTGCTGCTGATCTTTGGCACCATGACAGACGTGTTTCATCGAC
TATGATGTGGAGCTGCAGGAGCTGCAGATCCCTGGCAAGGCCTGCGTGAACAATACCATCGTGTGGACA
AATAGCTCCCTGAACCAGAATATGACCAACGGCACAAGATGTGGCCTGCTGAATATCGAGTCTGAGATG
ATCAAGTTTGGCAGCTACTATGCAGGAATCGCAGTGGCCGTGCTGATCACCGGCTACATCCAGATCTGC
TTCTGGGTCATCGCAGCAGCAAGGCAGATCCAGAAGATGAGAAAGTTCTATTTTCGGAGAATCATGCGG
ATGGAGATCGGCTGGTTTGAAGTCTGTGGGCGAGCTGAATACAAGATTCAGCGACGACATCAAC
AAGATCAATGACGCCATCGCCGATCAGATGGCCCTGTTTATCCAGCGGATGACCAGCACAACTCTGTGGC
TTCCTGCTGGGCTTCTTTAGAGGCTGGAAGCTGACCCTGGTCATCATCTCTGTGAGCCCACTGATCGGA
ATCGGAGCAGCAACAATCGGCCTGTCCGTGTCTAAGTTACCGACTACGAGCTGAAGGCATATGCAAAG
GCAGGAGTGGTGGCAGATGAAGTGATCAGCAGCATGAGGACCGTGGCAGCCTTTGGCGGAGAGAAGAGG
GAGGTGGAGCGGTACGAGAAGAACCTGGTGTTCGCCCAGCGGTGGGGCATCAGAAAGGGCATCGTGATG
GGCTTCTTTACAGGCTTCGTGTGGTGCCTGATCTTCCTGTGCTACGCCCTGGCCTTTTGGTATGGCTCC
ACCCTGGTGTGCTGGACGAGGGAGAGTATACCCCTGGCACACTGGTGCAGATCTTCCTGAGCGTGATCGTG
GGCGCCCTGAACCTGGGAAATGCCTCCCCATGCCTGGAGGCCTTTGCCACAGGCAGGGCAGCAGCCACC
TCCATCTTCGAGACAATCGACCGCAAGCCATCATCGACTGTATGTCTGAGGATGGCTACAAGCTGGAC
AGGATCAAGGGCGAGATCGAGTTTCACAATGTGACCTTCCACTATCCCAGCCGCCCTGAGGTGAAGATC
CTGAACGATCTGAATATGGTCATCAAGCCAGGAGAGATGACAGCCCTGGTGGGACCCTCTGGAGCAGGC
AAGAGCACCGCCCTGCAGCTGATCCAGCGGTTTTACGACCCTTGCGAGGGAATGGTGACCGTGGACGGA
CACGACATCAGGTCCCTGAACATCCAGTGGCTGCGCGATCAGATCGGCATCGTGGAGCAGGAGCCAGTG
CTGTTCTCTACCACAATCGCCGAGAATATCAGATACGGCCGCGAGGATGCCACAATGGAGGACATCGTG
CAGGCCGCCAAGGAGGCCAACGCCTATAACTTCATCATGGATCTGCCCCAGCAGTTCGACACCCTGGTG
GGAGAGGGAGGAGGACAGATGTCCGGAGGCCAGAAGCAGAGAGTGGCCATCGCCAGAGCCCTGATCCGC
AACCCTAAGATCCTGCTGCTGGATATGGCCACAAGCGCCCTGGACAATGAGTCCGAGGCCATGGTGCAG
GAGGTGCTGAGCAAGATCCAGCACGGCCACACCATCATCTCTGTGGCACACCGGCTGAGCACAGTGAGA
GCAGCCGACACCATCATCGGCTTTGAGCACGGAACAGCAGTGGAGAGGGGAACCCACGAGGAGCTGCTG
GAGAGGAAGGGCGTGTACTTCACCCTGGTGACACTGCAGAGCCAGGGCAACCAGGCCCTGAATGAGGAG
GACATCAAGGATGCCACAGAGGACGATATGCTGGCCCGGACCTTCTCCAGAGGCTCTTATCAGGATTCT
CTGCGGGCCAGCATCAGGCAGCGCAGCAAGTCCCAGCTGAGCTACCTGGTGCACGAGCCACCTCTGGCC
GTGGTGGACCACAAGTCCACCTATGAGGAGGATAGAAAGGACAAGGACATCCCCGTGCAGGAGGAGGTG
GAGCCTGCCCCAGTGAGGCGCATCCTGAAGTTTTCCGCCCCAGAGTGGCCCTACATGCTGGTGGGATCT
GTGGGAGCAGCAGTGAACGGCACCGTGACACCACTGTATGCCTTCCTGTTTTCCAGATCCTGGGCACC
TTCTCTATCCCCGACAAGGAGGAGCAGAGGTCCCAGATCAATGGCGTGTGCCTGCTGTTTGTGGCCATG

GGTTCGCTGAGCCTGTTTACACAGTTCTGCTTTCGCCAAGAGCGGCGAGCTGCTGACC
AAGCGGCTGAGAAAGTTCGGCTTTTCGCGCAATGCTGGGACAGGACATCGCATGGTTTGACGATCTGCGG
AACAGCCCAGGCGCCCTGACCACAAGACTGGCCACAGATGCCTCTCAGGTGCAGGGAGCAGCAGGCAGC
CAGATCGGCATGATCGTGAACCTCTTACCAATGTGACAGTGGCCATGATCATCGCCTTCTCTTTTAGC
TGGAAGCTGTCTCTGGTCATCCTGTGCTTCTTCCCTTTCTGGCCCTGAGCGGAGCAACCCAGACAAGG
ATGCTGACCGGCTTCGCCTCCAGAGACAAGCAGGCCCTGGAGATGGTGGGCCAGATCACAAACGAGGCC
CTGAGCAATATCAGGACCGTGGCAGGAATCGGCAAGGAGCGGCGGTTTCATCGAGGCCCTGGAGACAGAG
CTGGAGAAGCCTTTCAAGACCGCCATCCAGAAGGCCAACATCTACGGCTTCTGCTTTGCCTTCGCCCAG
TGTATCATGTTTCATCGCCAACTCCGCCTCTTACAGATATGGCGGCTACCTGATCAGCAATGAGGGCCTG
CACTTTTCCTACGTGTTTCAGAGTGATCAGCGCCGTGGTGCTGTCTGCCACAGCCCTGGGAAGGGCCTTC
TCCTACACCCCATCTTATGCCAAGGCCAAGATCTCTGCCGCCAGGTTCTTTTCAGCTGCTGGACCGCCAG
CCACCCATCAGCGTGTACAACACAGCCGGCGAGAAGTGGGATAATTTCCAGGGCAAGATCGACTTTGTG
GATTGCAAGTTCACCTATCCTAGCAGGCCAGACTCCCAGGTGCTGAACGGCCTGAGCGTGTCCATCTCT
CCAGGCCAGACACTGGCCTTTGTGGGCTCCTCTGGCTGTGGCAAGAGCACCTCCATCCAGCTGCTGGAG
AGGTTCTACGACCCCGATCAGGGGCAAAGTGATGATCGACGGCCACGATTCTAAGAAGGTGAACGTGCAG
TTTCTGCGCTCCAATATCGGCATCGTGTCTCAGGAGCCTGTGCTGTTTCGCCTGCAGCATCATGGATAAC
ATCAAGTACGGCGACAATACAAAGGAGATCCCAATGGAGAGAGTGATCGCAGCAGCAAAGCAGGCACAG
CTGCACGATTTTCGTGATGTCCCTGCCCCGAGAAGTATGAGACAAACGTGGGCAGCCAGGGCTCCCAGCTG
TCTAGGGGCGAGAAGCAGAGGATCGCAATCGCCAGGGCCATCGTGCGCGATCCCAAGATCCTGCTGCTG
GACGAGGCCACCAGCGCCCTGGATACAGAGTCCGAGAAGACCGTGCAGGTGGCCCTGGACAAGGCCCGG
GAGGGAAGAACATGTATCGTGATCGCCCACCGCCTGAGCACCATCCAGAATGCCGACATCATCGCCGTG
ATGGCCCAGGGCGTGGTCATCGAGAAGGGCACCCACGAGGAGCTGATGGCCCAGAAGGGCGCCTACTAT
AAGCTGGTGACCACAGGCAGCCCTATCTCCTGA

[0189] In some embodiments, the nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 42 or a fragment thereof. Suitably, the nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 75%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 42 or a fragment thereof.

[0190] In some embodiments, the nucleotide sequence encoding BSEP, or a fragment and/or variant thereof, comprises or consists of the nucleotide sequence SEQ ID NO: 42 or a fragment thereof.

TABLE-US-00016 Exemplary codon-optimised nucleotide encoding BSEP (SEQ ID NO: 42)
ATGAGCGATTCTGTGATCCTGCGGAGCATCAAGAAGTTCGGCGAGGAAAACGACGGCTTCGAGAGCGAC
AAGAGCTACAACAACGACAAGAAGTCCCGGCTGCAGGACGAGAAGAAAGGCGACGGTGTGAGAGTGGGC
TTCTTCCAGCTGTTCCGGTTCAGCAGCAGCACCGACATCTGGCTGATGTTCTGTGGGCAGCCTGTGCGCC
TTTCTGCACGGAATTGCTCAGCCTGGCGTGCTGCTGATCTTCGGCACCATGACCGACGTGTTTCATCGAC
TACGACGTGGAAGTCAAGAGCTGCAGATCCCTGGCAAGGCCTGCGTGAACAACACCATCGTGTGGACC
AACAGCAGCCTGAACCAGAACATGACCAACGGCACCAGATGCGGCCTGCTGAATATCGAGTCCGAGATG
ATCAAGTTCGCCAGCTACTACGCCGGAATCGCCGTGGCCGTTCTGATCACCGGCTACATCCAGATCTGC
TTTTGGGTATCGCCGCTGCCAGACAGATCCAGAAGATGCGGAAGTTCTACTTCCGGCGGATCATGCGG
ATGGAAATCGGATGGTTCGACTGCAACAGCGTGGGCGAGCTGAACACCAGATTTCAGCGACGACATCAAC
AAGATCAACGACGCCATTGCCGACCAGATGGCCCTGTTTCATCCAGCGGATGACCTCTACCATCTGCGGC
TTTCTGCTGGGCTTTTTTCAGAGGCTGGAAGCTGACCCTGGTCATCATCAGCGTGTCCCCACTGATCGGA
ATCGGAGCCGCCACAATCGGCCTGAGCGTGTCCAAGTTCACCGATTACGAGCTGAAGGCCTACGCCAAA
GCTGGCGTGGTGGCCGATGAAGTGATCAGCTCCATGAGAACCGTGGCCGCCTTTGGCGGCGAGAAGAGA
GAGGTGGAACGCTACGAGAAAAACCTGGTGTTTCGCCCAGAGATGGGGCATCAGAAAGGGCATCGTGATG
GGATTCTTACCCGGCTTCGTGTGGTGCTGATCTTCTGTGTTACGCCCTGGCCTTTTGGTACGGCAGC
ACCCTGGTTCTGGATGAGGGCGAGTATACCCCTGGAAGTCTGGTGCAGATCTTTCTGAGCGTGATCGTG
GGCGCCCTGAACCTGGGAAATGCCTCTCCATGCCTGGAAGCCTTTGCCACAGGCAGAGCCGCCGCTACC
AGCATCTTCGAGACAATCGACAGAAAGCCCATCATCGACTGCATGAGCGAGGACGGCTACAAGCTGGAC
AGAATCAAGGGCGAGATCGAGTTCACAAACGTGACCTTTCTACTACCCAGCAGACCCGAAGTGAAGATC
CTGAACGACCTGAACATGGTCATCAAGCCCGGCGAGATGACAGCCCTCGTTGGACCTTCTGGCGCCGGA
AAATCTACAGCCCTGCAGCTGATCCAGCGGTTCTACGATCCTTGCAGAGGCATGGTCAACCGTGGACGGC
CACGACATCAGATCTCTGAACATCCAGTGGCTGCGGGACCAGATCGGCATCGTGGAACAAGAGCCTGTG
CTGTTTCAGCACCACAATCGCCGAGAACATCAGATACGGCCGCGAGGATGCCACCATGGAAGATATTGTG
CAGGCCGCTAAAGAGGCCAACGCCTACAACCTTCATCATGGACCTGCCTCAGCAGTTCGACACCCTCGTC
GGAGAAGGCGGAGGACAAATGTCTGGCGGCCAGAAACAGAGAGTGGCCATTGCTAGAGCCCTGATCAGA

AAACCCAAAGATCCTGTGCTGGACATGGCCACAAAGCGCCCTGGATAATGAGAGCGAGGCCATGGTGCAA
GAGGTGCTGAGCAAGATCCAGCACGGCCACACCATCATCAGTGTGGCCACAGACTGAGCACCGTGCGG
GCTGCCGATACCATCATCGGATTTGAGCACGGCACCGCCGTGGAAAGAGGACACACACGAGGAACTGCTG
GAACGGAAGGGCGTGTACTTCACCCTCGTGACACTGCAGAGCCAGGGCAATCAGGCCCTGAACGAAGAG
GACATCAAGGACGCCACCGAGGACGACATGCTGGCCAGAACCTTTAGCAGAGGCAGCTACCAGGACAGC
CTGCGGGCCTCTATCAGACAGAGAAGCAAGAGCCAGCTGAGCTACCTGGTGCACGAACCTCCACTGGCT
GTGGTGGATCACAAGAGCACCTACGAGGAAGATCGGAAGGACAAGGACATCCCCGTGCAAGAAGAAGTG
GAACCCGCTCCTGTGCGGCGGATCCTGAAATTTTCTGCCCCTGAGTGGCCCTACATGCTCGTGGGATCT
GTGGGAGCCGCCGTGAATGGAACAGTGACCCCTCTGTACGCCTTCTGTCTCCAGATCCTGGGCACC
TTCAGCATCCCCGACAAAGAGGAACAGCGGAGCCAGATCAATGGCGTGTGCCTGCTGTTTGTGGCTATG
GGCTGCGTGTCCCTGTTTACCCAGTTCTCTGCAGGGATACGCCTTCGCCAAGAGCGGAGAGCTGCTGACC
AAGCGGCTGAGAAAGTTCCGGCTTCAGAGCCATGCTCGGCCAGGATATCGCTTGGTTTGACGACCTGAGA
AACAGCCCTGGCGCTCTGACAACCAGACTGGCCACAGATGCTTCTCAGGTGCAAGGCGCTGCCGGCTCT
CAGATCGGAATGATCGTGAACAGCTTCACCAATGTGACCGTGGCCATGATCATTGCCTTTAGCTTCAGC
TGAAACTGAGCCTCGTGATCCTGTGCTTCTTCCCATTCTTGCCCTGTCCGGCGCTACCCAGACCAGA
ATGCTTACAGGCTTCGCCAGCAGAGACAAGCAGGCCCTGGAAATGGTCGGACAGATACCAACGAGGCC
CTGAGCAACATCAGAACAGTGGCCGGCATCGGCAAAGAGCGGCGGTTTATTGAGGCTCTGGAAACCGAG
CTGGAAAACCCCTTCAAGACCGCCATCCAGAAGGCCAATATCTACGGCTTCTGCTTCGCCTTCGCTCAG
TGCATCATGTTTATCGCCAACAGCGCCAGCTATAGATACGGCGGCTACCTGATCAGCAATGAGGGCCTG
CACTTCAGCTACGTGTTCCGCGTGATCTCTGCCGTGGTGTCTGCCACAGCTCTGGGCAGAGCATTC
AGCTACACACCCAGCTACGCCAAGGCCAAGATTAGCGCCGCCAGATTCTTTCAGCTGCTGGATAGACAG
CCTCCTATCTCCGTGTACAACACCGCTGGCGAGAAGTGGGACAACCTTCAGGGCAAGATCGACTTCGTG
GATTGCAAGTTCACATACCCCTCTCGGCCCCGACAGCCAGGTGCTGAATGGACTGTCCGTGTCTATCAGC
CCCGGACAGACACTGGCCTTTGTGGGAAGCTCTGGCTGCGGCAAAAGCACCAGCATCCAGCTGCTCGAG
AGATTCTACGACCCCGACCAGGGCAAAGTGATGATCGACGGACACGACAGCAAGAAAGTGAACGTGCAG
TTCCTGAGATCCAACATCGGGATCGTGTCTCAAGAACCCGTGCTGTTTGCCTGCTCCATCATGGATAAC
ATTAAGTACGGCGACAACACCAAAGAAATCCCCATGGAAAGAGTGATCGCTGCCGCCAAACAGGCCCAG
CTGCACGATTTCTGTGATGAGCCTGCCTGAGAAGTACGAGACAAACGTGGGCTCTCAGGGCAGCCAGCTG
TCTAGAGGCGAAAAGCAGCGGATCGCTATCGCCAGAGCCATCGTGCGGGATCCTAAGATTCTGCTCCTG
GATGAAGCCACCAGCGCACTGGACACCGAGTCCGAAAAGACAGTGCAAGTGGCCCTGGACAAAGCCAGA
GAGGGCAGAACCTGTATCGTGATCGCCCACCGGCTGAGCACAATCCAGAACGCCGATATTATCGCCGTG
ATGGCCCAGGGCGTCGTGATCGAAAAGGGAACCCACGAAGAACTGATGGCTCAGAAGGGCGCCTACTAC
AAGCTCGTGACAACAGGCAGCCCTATCTCCTGA

MDR3

[0191] In some embodiments, the PFIC-associated polypeptide is MDR3 or a fragment and/or variant thereof.

[0192] “MDR3” is the abbreviated name of the polypeptide encoded by the ABCB4 gene and is also known as Multidrug resistance protein 3, Phosphatidylcholine translocator ABCB4, and ATP-binding cassette sub-family B member 4. MDR3 is a floppase that translocates specifically phosphatidylcholine (PC) from the inner to the outer leaflet of the canalicular membrane bilayer. Translocation of PC makes the biliary phospholipids available for extraction into the canaliculi lumen by bile salt mixed micelles and therefore protects the biliary tree from the detergent activity of bile salts.

[0193] A fragment and/or variant of MDR3 may retain MDR3 activity (see e.g. EC 7.6.2.1). For example, a fragment and/or variant of MDR3 may translocate phosphatidylcholine from the inner to the outer leaflet of the canalicular membrane bilayer. Suitably, a fragment and/or variant of MDR3 may have the same or similar activity to MDR3, e.g. may have at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the activity of MDR3.

[0194] A person skilled in the art would be able to generate fragments and/or variants using conservative substitutions, based on the known structural and functional features of MDR3 (see e.g. Oude Elferink, R. P. and Paulusma, C. C., 2007. *Pflügers Archiv-European Journal of Physiology*, 453(5), pp. 601-610; Olsen, J. A., et al., 2020. *Nature structural & molecular biology*, 27(1), pp. 62-70; and Nosol, K., et al., 2021. *PNAS*, 118(33), e2106702118), and/or based on known variants (see e.g. NCBI Gene ID: 5244 and NCBI HomoloGene: 136368).

[0195] Suitably, a fragment of MDR3 and/or a MDR3 variant comprises two homologous halves, each comprising a hydrophobic membrane-anchoring domain (e.g. an ABC transporter transmembrane region) and an ATP binding cassette (ABC) domain. The domains/sequences may be linked by inter-domain linker(s).

[0196] The ABCB4 gene is conserved in chimpanzee, dog, cow, mouse, and rat. The MDR3 may be a human MDR3. Suitably, the MDR3 may comprise or consist of a polypeptide sequence of UniProtKB accession P21439, or a fragment and/or variant thereof.

[0197] In some embodiments, the MDR3 comprises or consists of an amino acid sequence which is at least 70% identical to SEQ ID NO: 43 or a fragment thereof. Suitably, the MDR3 comprises or consists of an amino acid sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 43 or a fragment thereof.

[0198] In some embodiments, the MDR3 comprises or consists of SEQ ID NO: 43 or a fragment thereof.

TABLE-US-00017 Exemplary MDR3 (SEQ ID NO: 43)

MDLEAAKNGTAWRPTSAEGDFELGISSKQKRKTKTVKMIGVLTFLFRYSD
WQDKLFMSLGTIMAIHGSGLPLMMIVFGEMTDKFVDTAGNFSFPVNFSL
SLLNPGKILEEMTRYAYYYSGLGAGVLVAAYIQVSFWTLAAGRQIRKIR
QKFFHAILRQEIGWFDINDTTELNTRLTDDISKISEGIGDKVGMFFQAVA
TFFAGFIVGFIRGWKLTLVIMAISPILGLSAAVWAKILSAFSDKELAAAYA
KAGAVAEALGAIRTVIAFGGQNKELERYQKHLENAKEIGIKKAISANIS
MGIAFLLIYASYALAFWYGSTLVISKEYTIGNAMTVFFSILIGAFSVGQA
APCIDAFANARGAAYVIFDIIDNNPKIDSFSEGRGHKPDSEIKGNLEFNDVH
FSYPSRANVKILKGLNLKVQSGQTVALVGSSGCGKSTTVQLIQRLYDPDE
GTINIDGQDIRNENVNYLREIIGVVSQEPVLESTTIAENICYGRGNVTMD
EIKKAVKEANAYEFIMKLPQKFDTLVGERGAQLSGGQKQRIAIARALVRN
PKILLDEATSALDTESEAEVQAALDKAREGRTTIVIAHRLSTVRNADVI
AGFEDGVIVEQGSHELMKKEGVYFKLVNMQTSGSQIQSEEFELNDEKAA
TRMAPNGWKSRLFRHSTQKNLKNQMCQKSLDVETDGLEANVPPVSFLKV
LKLNKTEWPYFVVGTVCAIANGGLQPAFSVIFSEIIAIFGPGDDAVKQQK
CNIFSLIFLFLGIISFFTFLLQGFTFGKAGEILTRRLRSMAFKAMLRQDM
SWFDDHKNSTGALSTRLATDAAQVQGATGTRLALIAQNIANLGTGIIISF
IYGWQLTLLLLAVVPIIAVSGIVEMKLLAGNAKRDKKELEAAGKIAATEAI
ENIRTVVSLTQERKFESMYVEKLYGPYRNSVQKAHIYGITFSISQAFMYF
SYAGCFRFGAYLIVNGHMRFRDVILVESAIVFGAVALGHASSFAPDYAKA
KLSAAHLFMLFERQPLIDSYSEGLKPKDFEGNITFNEVVFNYPTRANVP
VLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPLAGTVFVDFGFQ
LLDGQEAKKLNQWLRAQLGIVSQEPILFDCSIAENIAYGDNSRVVSQDE
IVSAAKAANIHPFIETLPHKYETRVGDKGTQLSGGQKQRIAIARALIRQP
QILLDEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIV
VFQNGRVKEHGHQQLLAQKGIYFSMVSQAGTQNL

[0199] Suitably, a MDR3 variant may comprise one or more variation selected from: E4D, T175A, L238V, E528D, R652G, and R788Q. These are considered to be benign (or likely benign) variations based on clinical data. Suitably, a fragment of MDR3 and/or a MDR3 variant may comprise a deletion of aa 1094-1100 and/or aa 929-975.

[0200] Suitably, a nucleotide sequence encoding MDR3, or a fragment and/or variant thereof, may comprise or consist of a nucleotide sequence of NCBI reference sequence NM_018849, NM_000443 or NM_018850, or a fragment and/or variant thereof.

[0201] In some embodiments, the nucleotide sequence encoding MDR3, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 44 or a fragment thereof. Suitably, the nucleotide sequence encoding MDR3, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 44 or a fragment thereof.

[0202] In some embodiments, the nucleotide sequence encoding MDR3, or a fragment and/or variant thereof, comprises or consists of the nucleotide sequence SEQ ID NO: 44 or a fragment thereof.

TABLE-US-00018 Exemplary nucleotide encoding MDR3 (SEQ ID NO: 44)

ATGGATCTTGAGGCGGCAAAGAACGGAACAGCCTGGCGCCCCACGAGCGCGGAGGGCGACTTTGAACTG
GGCATCAGCAGCAAACAAAAAAGGAAAAAACGAAGACAGTGAAAATGATTGGAGTATTAACATTGTTT
CGATACTCCGATTGGCAGGATAAATTGTTTATGTCGCTGGGTACCATCATGGCCATAGCTCACGGATCA
GGTCTCCCCCTCATGATGATAGTATTTGGAGAGATGACTGACAAATTTGTTGATACTGCAGGAAACTTC
TCCTTTCCAGTGAACCTTTTCCTTGTCGCTGCTAAATCCAGGCAAAATTCTGGAAGAAGAAATGACTAGA
TATGCATATTACTACTCAGGATTGGGTGCTGGAGTTCTTGTTGCTGCCTATATACAAGTTTCATTTTGG
ACTTTGGCAGCTGGTTCGACAGATCAGGAAAATTAGGCAGAAGTTTTTTCATGCTATTCTACGACAGGAA
ATAGGATGTTTTGACATCAACGACACCACTGAACTCAATACGCGGCTAACAGATGACATCTCCAAAATC
AGTGAAGGAATTGGTGACAAGGTTGGAATGTTCTTTCAAGCAGTAGCCACGTTTTTTTGCAGGATTTCATA
GTGGGATTCATCAGAGGATGGAAGCTCACCTTGTGATAATGGCCATCAGCCCTATTCTAGGACTCTCT

GCAGCCGTTTGGGCCAAAGATACTCTCGGCATTTAGTGTGACAAAGAAGAACTAGCTGCTTATGCCAAAAGCAGGC
 GCCGTGGCAGAAGAGGCTCTGGGGGCCATCAGGACTGTGATAGCTTTTCGGGGGGCCAGAACAAAGAGCTG
 GAAAGGTATCAGAAACATTTAGAAAATGCCAAAGAGATTGGAATTAAGCTATTTTCAGCAAACATT
 TCCATGGGTATTGCCTTCCTGTAAATATATGCATCATATGCACTGGCCTTCTGGTATGGATCCACTCTA
 GTCATATCAAAAGAATATACTATTGGAAATGCAATGACAGTTTTTTTTTTCAATCCTAATTGGAGCTTTC
 AGTGTTGGCCAGGCTGCCCCATGTATTGATGCTTTTGCCAATGCAAGAGGAGCAGCATATGTGATCTTT
 GATATTATTGATAATAATCCTAAAATTGACAGTTTTTTCAGAGAGAGGACACAAACCAGACAGCATCAAA
 GGGAATTTGGAGTTCAATGATGTTCACTTTTCTTACCCTTCTCGAGCTAACGTCAAGATCTTGAAGGGC
 CTCAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAGTAGTGGCTGTGGGAAGAGCACA
 ACGGTCCAGCTGATACAGAGGCTCTATGACCCTGATGAGGGCACAATTAACATTGATGGGCAGGATATT
 AGGAACTTTAATGTAACTATCTGAGGGAAATCATTGGTGTGGTGAGTCAGGAGCCGGTGTCTGTTTTCC
 ACCACAATTGCTGAAAATATTTGTTATGGCCGTGGAAATGTAACCATGGATGAGATAAAGAAAGCTGTC
 AAAGAGGCCAACGCCTATGAGTTTATCATGAAATTACCACAGAAATTTGACACCCTGGTTGGAGAGAGA
 GGGGCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTTCGCAACCCCAAG
 ATCCTTCTGCTGGATGAGGCCACGTCAGCATTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCTCTG
 GATAAGGCCAGAGAAGGCCGGACCACCATTTGTGATAGCACACCGACTGTCTACGGTCCGAAATGCAGAT
 GTCATCGCTGGGTTTGAGGATGGAGTAATTGTGGAGCAAGGAAGCCACAGCGAACTGATGAAGAAGGAA
 GGGGTGTACTTCAAACCTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACTA
 AATGATGAAAAGGCTGCCACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACT
 CAGAAAAACCTTAAAAATTCACAAATGTGTGAGAAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCA
 AATGTGCCACCAGTGTCTTTCTGAAGGTCCTGAACTGAATAAAACAGAATGGCCCTACTTTGTCTGTG
 GGAACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCGGCATTTCAGTCATATTCTCAGAGATCATA
 GCGATTTTTGGACCAGGCGATGATGCAGTGAAGCAGCAGAAGTGAACATATTCTCTTTGATTTTCTTA
 TTTCTGGGAATTATTTCTTTTTTTACTTTCTTCCTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATC
 CTCACCAGAAGACTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGATGAC
 CATAAAAACAGTACTGGTGCATTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGAGCCACA
 GGAACCAGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGAAGCTGGTATTATCATATCATTTATC
 TACGGTTGGCAGTTAACCTTATTGCTATTAGCAGTTGTTCCAATTATTGCTGTGTGTCAGGAATTGTTGAA
 ATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAAAAAGAACTGGAAGCTGCTGGAAAGATTGCAACA
 GAGGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACCCAGGAAAGAAAATTTGAATCAATGTATGTT
 GAAAAATTGTATGGACCTTACAGGAATTCTGTGCAGAAGGCACACATCTATGGAATTACTTTTAGTATC
 TCACAAGCATTTATGTATTTTTCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATTGTGAATGGA
 CATATGCGCTTCAGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACAT
 GCCAGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCCTTATTCATGCTGTTTGAA
 AGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATAACA
 TTTAATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGGGCTGAGCCTGGAG
 GTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTC
 CTGGAGCGGTTCTACGACCCCTTGCGGGGACAGTGTTTGTGGACTTTGGTTTTCAGCTTCTCGATGGT
 CAAGAAGCAAAGAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGGAGCCTATC
 CTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAA
 ATTGTGAGTGCAGCCAAAGCTGCCAACATACATCCTTTCATCGAGACGTTACCCCAACAATATGAAACA
 AGAGTGGGAGATAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCAGGCCCTC
 ATCAGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAAAAGGTT
 GTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACC
 ATCCAGAATGCAGACTTAATAGTGGTGTTCAGAATGGGAGAGTCAAGGAGCATGGCACGCATCAGCAG
 CTGCTGGCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCTGGGACACAGAACTTATGA
 TJP2

[0203] In some embodiments, the PFIC-associated polypeptide is TJP2 or a fragment and/or variant thereof.

[0204] “TJP2” is the abbreviated name of the polypeptide encoded by the TJP2 gene and is also known as Tight junction protein ZO-2, Tight junction protein 2, Zona occludens protein 2, and Zonula occludens protein 2. TJP2 is a multi-domain protein that consists of an SH3 domain, a GK domain and three copies of a PDZ domain that plays a role in tight junctions and adherens junctions. TJP2 is a cytosolic component of several classes of cell-cell junctions (see e.g. Fanning, A. S., et al., 2012. Molecular biology of the cell, 23(4), pp. 577-590).

[0205] A fragment and/or variant of TJP2 may retain TJP2 activity. For example, a fragment and/or variant of TJP2 may bind to the cytoplasmic C termini of junctional transmembrane proteins and link them to the actin cytoskeleton. Suitably, a fragment and/or variant of TJP2 may have the same or similar activity to TJP2, e.g. may have at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least

100% of the activity of TJP2.

[0206] A person skilled in the art would be able to generate fragments and/or variants using conservative substitutions, based on the known structural and functional features of TJP2 (see e.g. in Sambrotta, M., et al., 2014. Nature genetics, 46(4), pp. 326-328 and Chen, H., et al., 2009. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun., 65(4), pp. 327-330), and/or based on known variants (see e.g. NCBI Gene ID: 9414 and NCBI HomoloGene: 3541).

[0207] Suitably, a fragment of TJP2 and/or a TJP2 variant comprises three PDZ domains, a SH3 domain, and a nucleoside/nucleotide kinase domain. The domains/sequences may be linked by inter-domain linker(s).

[0208] The TJP2 gene is conserved in chimpanzee, Rhesus monkey, dog, mouse, and rat. The TJP2 may be a human TJP2. Suitably, the TJP2 may comprise or consist of a polypeptide sequence of UniProtKB accession Q9UDY2, or a fragment and/or variant thereof.

[0209] In some embodiments, the TJP2 comprises or consists of an amino acid sequence which is at least 70% identical to SEQ ID NO: 45 or a fragment thereof. Suitably, the TJP2 comprises or consists of an amino acid sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 45 or a fragment thereof.

[0210] In some embodiments, the TJP2 comprises or consists of SEQ ID NO: 45 or a fragment thereof.

TABLE-US-00019 Exemplary TJP2 (SEQ ID NO: 45)

MPVRGDRGFPPRRELSGWLAPGMEELIWEQYTVTLQKDSKRGFGIAVSG
GRDNPHFENGETSIVISDVLPGGPADGLLQENDRVVMVNGTPMEDVLHSF
AVQQLRKSGKVAAIVVKRPRKVQVAALQASPLDQDDRAFEVMDEFDGRS
FRSGYSERSRLNSHGGRSRSWEDSPERGRPHERARSRELDLSRDRSRGRS
LERGLDQDHARTDRSRGRSLERGLDHDFGSPSRDRDRDRSRGRSIDQDYE
RAYHRAYDPDYERAYSPEYRRGARHDARSRGPRSRSRREHPHSRSPSPEPR
GRPGPIGVLLMKSRANEEYGLRLGSQIFVKEMTRTGLATKDGNLHEGDII
LKINGTVTENMSLTDARKLIEKSRGKLQLVVLRLDSQQT LINIPSLNDS
EIEDISEIESNRSFSPEERRHQYSDYDYHSSSEKLKERPSSREDTPSRLS
RMGATPTPFKSTGDIAGTVVPETNKEPRYQEDPPAPQPKAAPRTFLRSP
EDEAIYGPNTKMMVRFKKGDVGLRLAGGNDVGIFVAGIQEGTSAEQEGLQ
EGDQILKVNTQDFRGLVREDAVLYLLEIPKGEMVTILAQSRADVYRDILA
CGRGDSFFIRSHFECEKETPQSLAFTRGEVFRVVDTLYDGKLG NWLAVRI
GNELEKGLIPNKSRAEQMASVQNAQRDNAGDRADFWRMRGQRSGVKKNL
KSREDLTAVVSVSTKFPAYERVLLREAGFKRPVVLFGPIADIAMEKLANE
LPDWFQTAKTEPKDAGSEKSTGVVRLNTVRQIIEQDKHALLDVTPKAVDL
LNYTQWFPIVIFNPDSRQGVKTMRQRLNPTS NKSSRKLFDQANKLKKTC
AHLFTATINLNSANDSWFGSLKDTIQHQGEAVWVSEGKMEGMDDDPEDR
MSYLTAMGADYLSCD SRLISDFEDTDGEGGAYTDNELDEPAEEPLVSSIT
RSSEPVQHEESIRKPSPEPRAQMRAASSDQLRDNSPPPAFKPEPPKAKT
QNKEESYDESKSYEYKSNPSAVAGNETPGASTKGYPPPVA AKPTFGRSIL
KPSTPIPPQEGEEVGESSEEQDNAPKSVLGKVKIFEKMDHKARLQRMQEL
QEAQNARIEIAQKHPDIYAVPIKTHKPDGTPQHTSSRPPEPQKAPSRPY
QDTRGSYGSDAEEEEYRQQLSEHSKRGYYGQSARYRDTL

[0211] Suitably, a TJP2 variant may comprise one or more variation selected from: L15V, Q128K, R215H, R223Q, S296N, G355R, T364M, D482E, R493K, M668I, S711P, Q879R, L937P, P940L, S1010F, and V1064L. These are considered to be benign (or likely benign) variations based on clinical data. Suitably, a fragment of TJP2 and/or a TJP2 variant may comprise a deletion of aa 1-23 and/or aa 961-1107,

[0212] Suitably, a nucleotide sequence encoding TJP2, or a fragment and/or variant thereof, may comprise or consist of a nucleotide sequence of NCBI reference sequence NM_004817, NM_001170414, NM_001170415, NM_001170416, NM_001369870, NM_001369871, NM_001369872, NM_001369873, NM_001369874, NM_001369875 or NM_201629.3, or a fragment and/or variant thereof.

[0213] In some embodiments, the nucleotide sequence encoding TJP2, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 46 or a fragment thereof. Suitably, the nucleotide sequence encoding TJP2, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 46 or a fragment thereof.

[0214] In some embodiments, the nucleotide sequence encoding TJP2, or a fragment and/or variant thereof, comprises or consists of the nucleotide sequence SEQ ID NO: 46 or a fragment thereof.

TABLE-US-00020 Exemplary nucleotide encoding TJP2 (SEQ ID NO: 46)

ATGCCGTTGCCGAGGAGACCCGCGGTTTCCACCCCGCGGAGCTGTCAAGTTGGCTCCGCGCCCCAGGC
ATGGAAGAGCTGATATGGGAACAGTACACTGTGACCCTACAAAAGGATTCCAAAAGAGGATTTGGAATT
GCAGTGTCCGGAGGCAGAGACAACCCCCACTTTGAAAATGGAGAAACGTCAATTGTCATTTCTGATGTG
CTCCCGGGTGGGCTGCTGATGGGCTGCTCCAAGAAAATGACAGAGTGGTCATGGTCAATGGCACCCCC
ATGGAGGATGTGCTTCATTCTGTTTGCAGTTCAGCAGCTCAGAAAAAGTGGGAAGGTCGCTGCTATTGTG
GTCAAGAGGCCCCGGAAGGTCCAGGTGGCCGCACTTCAGGCCAGCCCTCCCCTGGATCAGGATGACCGG
GCTTTTGAGGTGATGGACGAGTTTGATGGCAGAAGTTTCCGGAGTGGCTACAGCGAGAGGAGCCGGCTG
AACAGCCATGGGGGGCGCAGCCGCAGCTGGGAGGACAGCCCGGAAAGGGGGCGTCCCCATGAGCGGGGCC
CGGAGCCGGGAGCGGGACCTCAGCCGGGACCGGAGCCGTGGCCGGAGCCTGGAGCGGGGCTGGACCAA
GACCATGCGCGCACCCGAGACCGCAGCCGTGGCCGGAGCCTGGAGCGGGGCTGGACCACGACTTTGGG
CCATCCCGGGACCGGGACCGTGACCGCAGCCGCGGCCGGAGCATTGACCAGGACTACGAGCGAGCCTAT
CACCGGGCCTACGACCCAGACTACGAGCGGGCCTACAGCCCGGAGTACAGGCGCGGGGGCCCCGACGAT
GCCCCGCTCTCGGGGACCCCCGAAGCCGCAGCCGCGAGCACCCGCACTCACGGAGCCCCAGCCCCGAGCCT
AGGGGGCGGCCGGGGGCCCATCGGGGTCTCCTGATGAAAAGCAGAGCGAACGAAGAGTATGGTCTCCGG
CTTGGGAGTCAGATCTTCGTAAAGGAAATGACCCGAACGGGTCTGGCAACTAAAGATGGCAACCTTCAC
GAAGGAGACATAATTCTCAAGATCAATGGGACTGTAAGTGAAGAACATGTCTTTAACGGATGCTCGAAAA
TTGATAGAAAAGTCAAGAGGAAAACCTACAGCTAGTGGTGTTGAGAGACAGCCAGCAGACCCTCATCAAC
ATCCCGTCAATTAATGACAGTGAATCAGAAATAGAAGATATTTTCAAGAAATAGAGTCAAACCGATCATTT
TCTCCAGAGGAGAGACGTCAATCAGTATTCTGATTATGATTATCATTCCTCAAGTGAGAAGCTGAAGGAA
AGGCCAAGTTCAGAGAGGACACGCCGAGCAGATTGTCCAGGATGGGTGCGACACCCACTCCCTTTAAG
TCCACAGGGGATATTGCAGGCACAGTTGTCCAGAGACCAACAAGGAACCCAGATACCAAGAGGACCCC
CCAGCTCCTCAACCAAAAAGCAGCCCCGAGAACTTTTCTTCGTCTTAGTCTGAAGATGAAGCAATATAT
GGCCCTAATACCAAAAATGGTAAGGTTCAAGAAGGGAGACAGCGTGGGCCTCCGGTTGGCTGGTGGCAAT
GATGTCGGGATATTTGTTGCTGGCATTCAAGAAGGGACCTCGGCGGAGCAGGAGGGCCTTCAAGAAGGA
GACCAGATTCTGAAGGTGAACACACAGGATTTTCAAGAGGATTAGTGCGGGAGGATGCCGTTCTCTACCTG
TTAGAAATCCCTAAAGGTGAAATGGTGACCATTTTAGCTCAGAGCCGAGCCGATGTGTATAGAGACATC
CTGGCTTGTGGCAGAGGGGATTTCGTTTTTTTATAAGAAGCCACTTTGAATGTGAGAAGGAAACTCCACAG
AGCCTGGCCTTCACCAGAGGGGAGGTCTTCCGAGTGGTAGACACACTGTATGACGGCAAGCTGGGCAAC
TGGCTGGCTGTGAGGATTGGGAACGAGTTGGAGAAAGGCTTAATCCCCAACAAGAGCAGAGCTGAACAA
ATGGCCAGTGTTCAAAATGCCCAGAGAGACAACGCTGGGGACCGGGCAGATTTCTGGAGAATGCGTGGC
CAGAGGTCTGGGGTGAAGAAGAACCTGAGGAAAAGTCGGGAAGACCTCACAGCTGTTGTGTCTGTCAGC
ACCAAGTTCCCAGCTTATGAGAGGGTTTTGCTGCGAGAAGCTGGTTTCAAGAGACCTGTGGTCTTATTC
GGCCCCATAGCTGATATAGCAATGGAAAAATTGGCTAATGAGTTACCTGACTGGTTTCAAAGTCTGCTAAA
ACGGAACCAAAAAGATGCAGGATCTGAGAAATCCACTGGAGTGGTCCGGTTAAATACCGTGAGGCAAATT
ATTGAACAGGATAAGCATGCACTACTGGATGTGACTCCGAAAGCTGTGGACCTGTTGAATTACACCCAG
TGGTTCCCAATTGTGATTTTTTTTCAACCCAGACTCCAGACAAGGTGTCAAAACCATGAGACAAAGGTTA
AATCCAACGTCCAACAAAAGTTCTCGAAAGTTATTTGATCAAGCCAACAAGCTTAAAAAAACGTGTGCA
CACCTTTTTTACAGCTACAATCAACCTAAATTCAGCCAATGATAGCTGGTTTGGCAGCTTAAAGGACACT
ATTGAGCATCAGCAAGGAGAAGCGGTTTTGGGTCTCTGAAGGAAAGATGGAAGGGATGGATGATGACCCC
GAAGACCGCATGTCTTACTTAACCGCCATGGGCGCGGACTATCTGAGTTGCGACAGCCGCCTCATCAGT
GACTTTGAAGACACGGACGGTGAAGGAGGCGCCTACACTGACAATGAGCTGGATGAGCCAGCCGAGGAG
CCGCTGGTGTGCTCCATCACCCGCTCCTCGGAGCCGGTGCAGCACGAGGAGAGCATAAGGAAACCCAGC
CCAGAGCCACGAGCTCAGATGAGGAGGGCTGCTAGCAGCGATCAACTTAGGGACAATAGCCCGCCCCCA
GCATTCAAGCCAGAGCCGCCCAAGGCCAAAACCCAGAACAAAGAAGAATCCTATGACTTCTCCAAATCC
TATGAATATAAGTCAAACCCCTCTGCCGTTGCTGGTAATGAAACTCCTGGGGCATCTACCAAAGGTTAT
CCTCCTCCTGTTGCAGCAAAAACCTACCTTTGGGCGGTCTATACTGAAGCCCTCCACTCCCATCCCTCCT
CAAGAGGGTGAGGAGGTGGGAGAGAGCAGTGAGGAGCAAGATAATGCTCCCAAATCAGTCCTGGGCAAA
GTCAAAATATTTGAGAAGATGGATCACAAGGCCAGGTTACAGAGAATGCAGGAGCTCCAGGAAGCACAG
AATGCAAGGATCGAAATTGCCCAGAAGCATCCTGATATCTATGCAGTTCCAATCAAAACGCACAAGCCA
GACCCTGGCACGCCCCAGCACACGAGTTCCAGACCCCTGAGCCACAGAAAGCTCCTTCCAGACCTTAT
CAGGATACCAGAGGAAGTTATGGCAGTGATGCCGAGGAGGAGGTACCGCCAGCAGCTGTCAGAACAC
TCCAAGCGCGGTTACTATGGCCAGTCTGCCCGATACCGGGACACAGAATTATAG
FXR

[0215] In some embodiments, the PFIC-associated polypeptide is FXR or a fragment and/or variant thereof.

[0216] “FXR” is the abbreviated name of the polypeptide encoded by the NR1H4 gene and is also known as Bile acid receptor, Farnesoid X-activated receptor, Farnesol receptor HRR-1, Nuclear receptor subfamily 1 group H member 4, and Retinoid X receptor-interacting protein 14. FXR is a receptor for bile acids such as

phenodeoxycholic acid, lithocholic acid, deoxycholic acid and alcoholic acid and plays an essential role in bile acid homeostasis.

[0217] A fragment and/or variant of FXR may retain FXR activity. For example, a fragment and/or variant of FXR may sense bile acid levels (e.g. by bile acid binding and subsequent activation). Suitably, a fragment and/or variant of FXR may have the same or similar activity to a FXR, e.g. may have at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the activity of FXR.

[0218] A person skilled in the art would be able to generate fragments and/or variants using conservative substitutions, based on the known structural and functional features of FXR (see e.g. Mi, L. Z., et al., 2003. Molecular cell, 11(4), pp. 1093-1100 and Gomez-Ospina, N., et al., 2016. Nature communications, 7(1), pp. 1-8), and/or based on known variants (see e.g. NCBI Gene ID: 9971 and NCBI HomoloGene: 3760).

[0219] Suitably, a fragment of FXR and/or a FXR variant comprises a nuclear receptor DNA-binding domain and a nuclear receptor ligand-binding domain. The domains/sequences may be linked by inter-domain linker(s).

[0220] The NR1H4 gene is conserved in chimpanzee, Rhesus monkey, dog, cow, mouse, and rat. The FXR may be a human FXR. Suitably, the FXR may comprise or consist of a polypeptide sequence of UniProtKB accession Q96RI1, or a fragment and/or variant thereof.

[0221] In some embodiments, the FXR comprises or consists of an amino acid sequence which is at least 70% identical to SEQ ID NO: 47 or a fragment thereof. Suitably, the FXR comprises or consists of an amino acid sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 47 or a fragment thereof.

[0222] In some embodiments, the FXR comprises or consists of SEQ ID NO: 47 or a fragment thereof.

TABLE-US-00021 Exemplary FXR (SEQ ID NO: 47)

MVMQFQGLENIQISPHCSCTPSGFFMEMMSMKPAKGVLTEQVAGPLGQN
LEVEPYQSQYSNVQFPQVQPQISSSSYYSNLGFYPQQPEEWYSPGIYELRR
MPAETLYQGETEVAEMPVTKKPRMGASAGRIKGDEL CVVCGDRASGYHYN
ALTCEGCKGFFRRSITKNAVYKCKNGGNCVMDMYMRRKCQECRLRKCKEM
GMLAECMYTGLLTEIQCKSKRLRKNVKQHADQTVNEDSEGRDLRQVTSTT
KSCREKTELTPDQQTLLHFIMDSYNKQRMPQEITNKILKEEFSAEENFLI
LTEMATNHVQVLVEFTKKLPGFQTLDHEDQIALLKGSVEAMFLRSAEIF
NKKLPSGHSDLLEERIRNSGISDEYITPMFSFYKSIGELKMTQEEYALLT
AIVILSPDRQYIKDREAVEKLQEPLLDVLQKLCKIHQPENPQHFACLLGR
LTELRFTFNHHHAEMLSWRVNDHKFTPLLCEIWDVQ

[0223] Suitably, a FXR variant may comprise one or more variation selected from V381 and M183T. These are considered to be benign (or likely benign) variations based on clinical data. Suitably, a fragment of FXR and/or a FXR variant may comprise a deletion of aa 207-210 or aa 159-209.

[0224] In some embodiments, the FXR comprises or consists of an amino acid sequence which is at least 70% identical to SEQ ID NO: 48 or a fragment thereof. Suitably, the FXR comprises or consists of an amino acid sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 48 or a fragment thereof.

[0225] In some embodiments, the FXR comprises or consists of SEQ ID NO: 48 or a fragment thereof.

TABLE-US-00022 Exemplary FXR (SEQ ID NO: 48)

MGSKMNLIEHSHLPTTDEFSFSENLFGLVTEQVAGPLGQNLEVEPYQSQYS
NVQFPQVQPQISSSSYYSNLGFYPQQPEEWYSPGIYELRRMPAETLYQGE
TEVAEMPVTKKPRMGASAGRIKGDEL CVVCGDRASGYHYNALTCEGCKGF
FRSITKNAVYKCKNGGNCVMDMYMRRKCQECRLRKCKEMGMLAECMYTG
LLTEIQCKSKRLRKNVKQHADQTVNEDSEGRDLRQVTSTTKSCREKTEL
PDQQTLLHFIMDSYNKQRMPQEITNKILKEEFSAEENFLILTEMATNHVQ
VLVEFTKKLPGFQTLDHEDQIALLKGSVEAMFLRSAEIFNKKLPSGHSD
LLEERIRNSGISDEYITPMFSFYKSIGELKMTQEEYALLTAIVILSPDRQ
YIKDREAVEKLQEPLLDVLQKLCKIHQPENPQHFACLLGRLTELRFTFNHH
HAEMLSWRVNDHKFTPLLCEIWDVQ

[0226] Suitably, a FXR variant may comprise one or more variation selected from V28I and M173T. These are considered to be benign (or likely benign) variations based on clinical data. Suitably, a fragment of FXR and/or a FXR variant may comprise a deletion of aa 197-200 or aa 149-199.

[0227] Suitably, the nucleotide sequence encoding FXR, or a fragment and/or variant thereof, may comprise or consist of a nucleotide sequence of NCBI reference sequence NM_001206993, NM_001206977, NM_001206978, NM_001206979, NM_001206992 or NM_005123.4, or a fragment and/or variant thereof.

[0228] In some embodiments, the nucleotide sequence encoding FXR, or a fragment and/or variant thereof,

comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 49 or a fragment thereof. Suitably, the nucleotide sequence encoding FXR, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 49 or a fragment thereof.

[0229] In some embodiments, the nucleotide sequence encoding FXR, or a fragment and/or variant thereof, comprises or consists of the nucleotide sequence SEQ ID NO: 49 or a fragment thereof.

TABLE-US-00023 Exemplary nucleotide encoding FXR (SEQ ID NO: 49)

```
ATGGTAATGCAGTTTCAGGGGTTAGAAAATCCAATTCAAATTAGTCCTCA
CTGCAGCTGTACGCCGTCAGGATTTTTCATGGAAATGATGAGTATGAAGC
CCGCGAAAGGTGTTTTAACAGAACAAAGTGGCAGGTCCTCTGGGACAGAAC
CTGGAAGTGGAAACCATACTCGCAATACAGCAATGTTTCAGTTTCCCCAAGT
TCAACCACAGATTTCTCTCGTCATCCTATTATTCCAACCTGGGTTTCTACC
CCCAGCAGCCTGAAGAGTGGTACTCTCCTGGAATATATGAACTCAGGCGT
ATGCCAGCTGAGACTCTCTACCAGGGAGAACTGAGGTAGCAGAGATGCC
TGTAACAAAGAAGCCCCGCATGGGCGCGTCAGCAGGGAGGATCAAAGGGG
ATGAGCTGTGTGTTGTTTGTGGAGACAGAGCCTCTGGATACCACTATAAT
GCACTGACCTGTGAGGGGTGTAAAGGTTTCTTCAGGAGAAGCATTACCAA
AAACGCTGTGTACAAGTGTAAAAACGGGGGCAACTGTGTGATGGATATGT
ACATGCGAAGAAAGTGTCAAGAGTGTGCTACTAAGGAAATGCAAAGAGATG
GGAATGTTGGCTGAATGTATGTATACAGGCTTGTTAACTGAAATTCAGTG
TAAATCTAAGCGACTGAGAAAAAATGTGAAGCAGCATGCAGATCAGACCG
TGAATGAAGACAGTGAAGGTCGTGACTTGCGACAAGTGACCTCGACAACA
AAGTCATGCAGGGAGAAAACTGAACTCACCCAGATCAACAGACTCTTCT
ACATTTTATTATGGATTCATATAACAAACAGAGGATGCCTCAGGAAATAA
CAAATAAAATTTTAAAGAAGAATTCAGTGCAGAAGAAAATTTTCTCATT
TTGACGGAAATGGCAACCAATCATGTACAGGTTCTTGTAGAATTCACAAA
AAAGCTACCAGGATTTTCAGACTTTGGACCATGAAGACCAGATTGCTTTGC
TGAAAGGGTCTGCGGTTGAAGCTATGTTCTTCGTTTCAGCTGAGATTTTC
AATAAGAACTTCCGTCTGGGCATTCTGACCTATTGGAAGAAAGAATTCG
AAATAGTGGTATCTCTGATGAATATATAACACCTATGTTTAGTTTTTATA
AAAGTATTGGGGAAGTGAAGAAATGACTCAAGAGGAGTATGCTCTGCTTACA
GCAATTGTTATCCTGTCTCCAGATAGACAATAACATAAAGGATAGAGAGGC
AGTAGAGAAGCTTCAGGAGCCACTTCTTGATGTGCTACAAAAGTTGTGTA
AGATTCACCAGCCTGAAAATCCTCAACACTTTGCCTGTCTCCTGGGTCGC
CTGACTGAATTACGGACATTCAATCATCACCACGCTGAGATGCTGATGTC
ATGGAGAGTAAACGACCACAAGTTTACCCCACTTCTCTGTGAAATCTGGG ACGTGCAGTGA
```

MYO5B

[0230] In some embodiments, the PFIC-associated polypeptide is MYO5B or a fragment and/or variant thereof.

[0231] “MYO5B” is the abbreviated name of the polypeptide encoded by the MYO5B gene and is also known as myosin VB. MYO5B is involved in vesicular trafficking and interacts with rab11. This allows for proper functioning of polarized epithelial cells, including hepatocytes (see e.g. Gonzales, E., et al., 2017. *Hepatology*, 65(1), pp. 164-173).

[0232] A fragment and/or variant of MYO5B may retain MYO5B activity. For example, a fragment and/or variant of MYO5B may interact with rab11 and traffic vesicles. Suitably, a fragment and/or variant of MYO5B may have the same or similar activity to MYO5B, e.g. may have at least 40%, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 100% of the activity of MYO5B.

[0233] A person skilled in the art would be able to generate fragments and/or variants using conservative substitutions, based on the known structural and functional features of MYO5B (see e.g. Müller, T., et al., 2008. *Nature genetics*, 40(10), pp. 1163-1165; Velvarska, H. and Niessing, D., 2013. *PLoS One*, 8(12), p. e82065; and Nascimento, A. F., et al., 2013. *Journal of Biological Chemistry*, 288(47), pp. 34131-34145), and/or based on known variants (see e.g. NCBI Gene ID: 4645 and NCBI HomoloGene: 49481).

[0234] Suitably, a fragment of MYO5B and/or a MYO5B variant comprises a Myosin and Kinesin motor domain, an IQ calmodulin-binding motif, and a DIL domain. The domains/sequences may be linked by inter-domain linker(s).

[0235] The MYO5B gene is conserved in chimpanzee, Rhesus monkey, dog, cow, mouse, and rat. The MYO5B may be a human MYO5B. Suitably, the MYO5B may comprise or consist of a polypeptide sequence of UniProtKB accession Q9ULV0, or a fragment and/or variant thereof.

[0236] In some embodiments, the MYO5B comprises or consists of an amino acid sequence which is at least 70% identical to SEQ ID NO: 50 or a fragment thereof. Suitably, the MYO5B comprises or consists of an amino acid sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 50 or a fragment thereof.

[0237] In some embodiments, the MYO5B comprises or consists of SEQ ID NO: 50 or a fragment thereof.

TABLE-US-00024 Exemplary MYO5B (SEQ ID NO: 50)

MSVGELYSQCTRVWIPDPDEVWRS AELTKDYKEGDKSLQLRLEDETILEY
PIDVQRNQLPFLRNPDLVGENDLTALSYLHEPAVLHNLKVRFLSNHIY
TYCGIVLVAINPYEQLPIYGQDVIYTYSGQNMGMMDPHIFAVAE EAYKQM
ARDEKNQSIIVSGESGAGKTVSAKYAMRYFATVGG SASETNIEEKVLASS
PIMEAIGNAKTTRNDNSSRFGKYIQIGEDKRYHIIIGANMRTYLL EKSRVV
FQADDERNYHIFYQLCAAAGLPEFKELALTS AEDFFYTSQGGDTSIEGVD
DAEDFEKTRQAFTLLGVKESHQMSIFKIIASILHLGSAIQAERDGDSCS
ISPQDVYLSNFCRL LGVEHSQMEHWLCHRKLVT TSETYVKTMSLQQVINA
RNALAKHIYAQLFGWIVEHINKALHTSLKQHSFIGVLDIYGFETFEVNSF
EQFCINYANEKLQQQFN SHVFKLEQEEYMKEQIPWTLIDFYDNQPCIDLI
EAKLGILDLLDEECKVPKGT DQNWAKKLYDRHSSSQHFQKPRMSNTAFII
VHFADKVEYLS DGFLEKNRDTVYEEQINILKASKFPLVADLFHDDKDPVP
ATTPGKGSSSKISVRSARPPMKVSNKEHKKT VGHQFRTSLHLLMETLNAT
TPHYVRCIKPNDEKL PFHFDPKRAVQQLRACGVLETIRISAAGYPSRWAY
HDIFFNRYRVLVKKRELANTDKKAICRSVLENLIKDPDKFQFGRTKIFFRA
GQVAYLEKL RADKFRTATIMIQKTVRGWLQKV KYHRLKGATLTLQRYCRG
HLARRLAEHLRRIRAAVVLQKH YRMQRARQAYQRRRAAVVIQAFTRAMF
VRRTYRQVLM EHKATTIQKHVRGWMARRHFQRLRDAAI VIQCAFRMLKAR
RELKALRIEARS AEHLKRLNVGMENKVVQLQRKIDEQNKEFKTLSEQLSV
TTSTYTMEVERLKKELVHYQQSPGEDTSLRLQEEVESLRTELQRAHSERK
ILED AHSREKDELKRKVADLEQENALLKDEKEQLNNQILCQSKDEFAQNS
VKENLMKKELEEERSRYQNLVKEYS QLEQRYDNLRDEMTHKQTPGHRRN
PSNQSSLES DSNYP SISTSEIGDTEDALQQVEEIGLEKAAMD MTVFLKLQ
KRVRELEQERKKLQVQLEKREQQDSKKVQAEP PQT DIDLDPNADLAYNSL
KRQELESENKKLKN DLNELRKAVADQATQNNSSHGSPDSYSLLLNQLKLA
HEELEVRKEEVLILRTQIVSADQRRLAGRNAEPNINARSSWPNSEKHVDQ
EDAIEAYHGVCQTNSKTEDWGYLNEDGELGLAYQGLKQVARLLEAQLQAQ
SLEHEEEVEHLKAQLEALKEEMDKQQQTFCQTLLLSPEAQVEFGVQQEIS
RLTNENLDL KELVEKLEKNERKLKKQLKIYMKKAQDLEAAQALAQSERKR
HELN RQVTVQRKEKDFQGMLEYHKEDEALLIRNLVTDLKPQMLSGTVPCL
PAYILYMCIRHADYTNDDLKVHSLTSTINGIKKVLKKHND DFEMTSFWL
SNTCRL LHCLKQYSGDEGFMTQNTAKQNEHCLKNEDLTEYRQVLS DLSIQ
IYQQLIKIAEGVLQPMIVSAMLENESI QGLSGVKPTGYRKRSSSMADGDN
SYCLEAIIRQMNAFHTVMCDQGLDPEIILQVFKQLFYMINAVTLNNLLLR
KDVC SWSTGMQLRYNISQLEEWLRGRNLHQSGAVQTMEPLIQAAQLLQLK
KKTQEDAE AICSLCTSLSTQQIVKILNLYTPLNEFEERVTVAFIRTIQAAQ
LQERNDPQQLLLDAKHMFVLPFPNPSSLTMDSIHIPACLNLEFLNEV

[0238] Suitably, a MYO5B variant may comprise one or more variation selected from C10G, E49Q, R56L, L59P, F61L, T126A, I296V, K307N, R344H, S385L, V516I, P598L, R743H, I769N, K781N, R901Q, R911L, R918H, D1011N, P1191L, G1321E, T1382M, S1494L, H1558Q, K1683R, and M1688V. These are considered to be benign (or likely benign) variations based on clinical data. Suitably, a fragment of MYO5B and/or a MYO5B variant may comprise a deletion of aa 1 to 859 and/or aa 1315-1340, or aa 1-1430.

[0239] Suitably, the nucleotide sequence encoding MYO5B, or a fragment and/or variant thereof, may comprise or consist of a nucleotide sequence of NCBI reference sequence NM_001080467, or a fragment and/or variant thereof.

[0240] In some embodiments, the nucleotide sequence encoding MYO5B, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 51 or a fragment thereof. Suitably, the nucleotide sequence encoding MYO5B, or a fragment and/or variant thereof, comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to SEQ ID NO: 51 or a fragment thereof.

[0241] In some embodiments, the nucleotide sequence encoding MYO5B, or a fragment and/or variant thereof,

comprises or consists of the nucleotide sequence SEQ ID NO: 51 or a fragment thereof.

TABLE-US-00025 Exemplary nucleotide encoding MYO5B (SEQ ID NO: 51)

ATGTCGGTGGGCGAGCTCTACAGCCAGTGCACAAGGGTCTGGATCCCTGACCCTGATGAGGTATGGCGC
TCAGCTGAGTTAACCAAGGACTACAAAGAAGGAGACAAGAGCCTACAGCTCAGACTGGAGGATGAAACG
ATTCTGGAATACCCAATTGATGTACAACGCAACCAGCTGCCCTTCTTACGGAATCCAGATATCTTGGTG
GGAGAAAATGACCTGACTGCCCTTAGCTATCTTCATGAGCCTGCAGTTTTGCATAATTTGAAGGTCCGT
TTCCTGGAGTCCAACCATATCTACACTTACTGTGGTATCGTACTTGTGTCATTAATCCTTATGAACAG
TTGCCAATCTATGGACAAGATGTCATCTATACCTACAGTGGCCAAAACATGGGAGACATGGACCCCCAC
ATCTTTGCTGTGGCAGAAGAAGCCTACAAGCAGATGGCCAGAGATGAGAAGAATCAGTCCATCATAGTC
AGTGGGGAGTCTGGAGCCGGGAAGACGGTATCAGCCAAGTATGCCATGCGCTATTTCCGCCACCGTTGGT
GGCTCGGCCAGTGAAACCAACATCGAAGAGAAGGTGCTGGCATCCAGTCCCATCATGGAGGCCATTGGA
AATGCCAAGACCACCCGCAATGACAACAGCAGCCGTTTTTGCCAAGTACATCCAGATTGGCTTTGACAAA
AGGTACCACATCATCGGGGGCCAACATGAGGACTTACCTCTTGAGAGAAGTCCAGAGTGGTCTTCCAGGCA
GATGATGAGAGGAATTACCACATCTTTTACCAGCTCTGTGCTGCTGCCGGTCTTCCAGAATTTAAAGAG
CTTGCACTAACAAGTGCAGAGGACTTTTTCTATACATCACAGGGAGGAGACACTTCCATCGAGGGTGTG
GACGATGCTGAGGACTTTGAGAAGACTCGACAAGCCTTCACACTCCTCGGAGTGAAAGAGTCCCATCAG
ATGAGCATTTTTTAAGATAATTGCTTCTATCTTGCACCTTGAAGTGTGGCGATTCAGGCTGAGCGTGAT
GGTGATTCCTGTAGTATATCACCCAGGATGTATACCTAAGCAACTTCTGCCGACTGCTAGGGGTGGAG
CACAGTCAGATGGAGCACTGGCTGTGTCATCGCAAGCTGGTCAACACCTCGGAGACCTACGTCAAGACC
ATGTCCCTGCAGCAGGTGATCAATGCGCGCAACGCCCTGGCGAAGCACATCTATGCCCAGTTGTTCCGC
TGGATTGTGGAGCACATCAACAAGGCCCTGCACACCTCCCTCAAGCAGCACTCCTTCATCGGGGTCTCTG
GACATCTATGGGTTTTGAGACATTTGAGGTAAACAGCTTTGAGCAGTTCTGTATCAACTATGCAAATGAA
AAGCTCCAGCAGCAGTTCAACTCGCATGTTTTCAAACCTGGAGCAAGAAGAATACATGAAGGAACAGATC
CCTTGGACCCTGATTGATTTTTATGATAACCAACCTTGTATCGACCTCATTGAAGCCAAGCTGGGTATC
TTGGACCTGTTGGATGAAGAATGTAAGGTCCCCAAAGGAACTGACCAGAACTGGGCTCAGAAGCTCTAT
GACCGGCACTCCAGCAGCCAGCACTTCCAGAAGCCCCGCATGTCCAACACGGCCTTCATCATCGTCCAC
TTTGACAGACAAGGTGGAGTACCTCTCTGATGGTTTTCTGGAGAAAAACAGAGACACGGTGTATGAAGAG
CAGATCAATATCCTGAAGGCCAGCAAGTTCCCACTAGTGGCTGACTTGTTTCATGATGACAAGGACCCT
GTTCTGCCACCACCCCTGGGAAGGGGTCATCTTCGAAGATCAGCGTCCGTTCTGCCAGACCCCCCATG
AAAGTCTCCAACAAGGAGCACAAAGAAAACCGTTGGCCACCAGTTCCGTACCTCCCTGCATCTGCTCATG
GAGACCCTGAATGCCACGACACCTCACTATGTCCGCTGCATCAAGCCCAACGATGAGAAGCTCCCCTTT
CACTTTGACCCAAAGAGAGCAGTGCAGCAACTCAGAGCCTGCGGGGTGTTGGAGACGATTCGAATCAGT
GCAGCTGGCTACCCATCCAGGTGGGCCTACCATGACTTTTTCAACCGGTATCGGGTGCTGGTCAAGAAG
AGAGAGCTCGCCAACACAGACAAAAAGGCCATCTGCAGGTCTGTCTTGAGAACTCATCAAGGACCCC
GACAAGTTCCAGTTTGGCCGCACCAAGATCTTCTTTTCGAGCAGGCCAGGTGGCCTACCTGGAGAAGCTG
CGGGCTGACAAGTTCCGGACAGCCACCATCATGATCCAGAAAACCTGTCCGGGGATGGCTGCAGAAGGTG
AAATATCACAGGCTGAAGGGGGCTACCTTAACCCTGCAGAGGTACTGCCGGGGACACCTGGCCCGCAGG
CTGGCTGAGCACCTGCGGAGGATCAGAGCGGCTGTGGTGCTCCAGAAACATTACCGCATGCAGAGGGCC
CGCCAGGCCTACCAGAGGGTCCGCAGAGCTGCCGTTGTTATCCAGGCCTTACCCGGGCCATGTTTGTG
CGGAGAACCTACCGCCAGGTCTCATGGAGCACAAAGGCCACCACCATCCAGAAGCACGTGCGGGGGCTGG
ATGGCACGCAGGCACTTCCAGCGGCTGCGGGATGCAGCCATTGTCATCCAGTGTGCCTTCCGGATGCTC
AAGGCCAGGCGGGAGCTGAAGGCCCTCAGGATTGAGGCCCGCTCAGCAGAGCATCTGAAACGTCTCAAC
GTGGGCATGGAGAACAAGGTGGTCCAGCTGCAGCGGAAGATCGATGAGCAGAACAAAGAGTTCAAGACA
CTTTCAGAGCAGTTGTCCGTGACCACCTCAACATACACCATGGAGGTAGAGCGGCTGAAGAAGGAGCTG
GTGCACTACCAGCAGAGCCCAGGTGAGGACACCAGCCTCAGGCTGCAGGAGGAGGTGGAGAGCCTGCGC
ACAGAGCTGCAGAGGGCCCACTCGGAGCGCAAGATCTTGGAGGACGCCACAGCAGGGAGAAAGATGAG
CTGAGGAAGCGAGTTGCAGACCTGGAGCAAGAAAATGCTCTCTTGAAAGATGAGAAAGAACAGCTCAAC
AACCAAATCCTGTGCCAGTCTAAAGATGAATTTGCCCAGAACTCTGTGAAGGAAAATCTCATGAAGAAA
GAACTGGAGGAGGAGCGATCCCGGTACCAGAACCTTGTGAAGGAATATTCACAGTTGGAGCAGAGATAC
GACAACCTTCGGGATGAAATGACCATCATAAAGCAAACCTCCAGGTCATAGGCGGAACCCATCAAACCAA
AGTAGCTTAGAATCTGACTCCAATTACCCCTCCATCTCCACATCTGAGATCGGAGACACTGAGGATGCC
CTCCAGCAGGTGGAGGAAATTGGCCTGGAGAAGGCAGCCATGGACATGACGGTCTTCCTGAAGCTGCAG
AAGAGAGTACGGGAGCTGGAGCAGGAGAGGAAAAAGCTGCAAGTGCAGCTGGAGAAGAGAGAACAGCA
GACAGCAAGAAAGTCCAGGCGGAACCACCACAGACTGACATAGATTTGGACCCGAATGCAGATCTGGCC
TACAATAGTCTGAAGAGGCAAGAGCTGGAGTCAGAGAACAAAAAGCTGAAGAATGACCTGAATGAGCTG
AGGAAAGCCGTGGCCGACCAAGCCACGCAGAATAACTCCAGCCACGGCTCCCCAGATAGCTACAGCCTC
CTGCTGAACCAGCTCAAGCTGGCCACGAGGAGCTCGAGGTGCGCAAGGAGGAGGTGCTCATCCTCAGG

ACCCAGATCTGTAGCGTCCGACCGGACCGGAGCCGAGCCGAACATTAATGCCAGA
TCAAGTTGGCCTAACAGTGAAAAGCATGTTGACCAGGAGGATGCCATTGAGGCCTATCACGGGGTCTGC
CAGACAAACAGCAAGACTGAGGATTGGGGATATTTAAATGAAGATGGAGAACTCGGCTTGGCCTACCAA
GGCCTAAAGCAAGTTGCCAGGCTGCTGGAGGCTCAGCTGCAGGCCCAGAGCCTGGAGCATGAGGAGGAG
GTGGAGCATCTCAAGGCTCAGCTCGAGGCCCTGAAGGAGGAGATGGACAAACAGCAGCAGACCTTCTGC
CAGACGCTACTGCTCTCCCCAGAGGCCCGAGGTGGAATTCGGCGTTCAGCAGGAAATATCCCGGCTGACC
AACGAGAATCTGGACCTTAAAGAACTGGTAGAAAAGCTGGAAAAGAATGAGAGGAAGCTCAAAAAGCAA
CTGAAGATTTACATGAAGAAAGCCCAGGACCTAGAAGCTGCCCAGGCATTGGCCCAGAGTGAGAGGAAG
CGCCATGAGCTCAACAGGCAGGTCACGGTCCAGCGGAAAGAGAAGGATTTCCAGGGCATGCTGGAGTAC
CACAAAGAGGACGAGGCCCTCCTCATCCGGAACCTGGTGACAGACTTGAAGCCCCAGATGCTGTCTGGGC
ACAGTGGCCTGTCTCCCCGCCTACATCCTCTACATGTGCATCCGGCACGCGGACTACACCAACGACGAT
CTCAAGGTGCACTCCCTGCTGACCTCCACCATCAACGGCATTAAAGAAAGTCCTGAAAAAGCACAAATGAT
GACTTTGAGATGACGTCATTCTGGTTATCCAACACCTGCCGCCTTCTTCACTGTCTGAAGCAGTACAGC
GGGGATGAGGGCTTCATGACTCAGAACACTGCAAAGCAGAATGAACACTGTCTTAAGAATTTTGACCTC
ACCGAATACCGTCAGGTGCTGAGTGACCTTTCCATTAGATCTACCAGCAGCTCATTAAAATTGCCGAG
GGCGTGTTACAGCCGATGATAGTTTCTGCCATGTTGGAAAATGAGAGCATTAGGGGTCTATCTGGTGTG
AAGCCCACCGGCTACCGGAAGCGCTCCTCCAGCATGGCAGATGGGGATAACTCATACTGCCTGGAAGCT
ATCATCCGCCAGATGAATGCCTTTCATACAGTCATGTGTGACCAGGGCTTGGACCCTGAGATCATCCTG
CAGGTATTCAAACAGCTCTTCTACATGATCAACGCAGTGAAGTCTTAACAACCTGCTCTTGCGGAAGGAC
GTCTGCTCTTGAGAGCACAGGCATGCAACTCAGGTACAATATAAGTCAGCTTGAGGAGTGGCTTCGGGGA
AGAAACCTTCACCAGAGTGAGAGCAGTTCAGACCATGGAACCTCTGATCCAAGCAGCCCAGCTCCTGCAA
TTAAAGAAGAAAACCCAGGAGGACGCAGAGGCTATCTGCTCCCTGTGTACCTCCCTCAGCACCCAGCAG
ATTGTCAAATTTTAAACCTTTATACTCCCCTGAATGAATTTGAAGAACGGGTAACAGTGGCCTTTATA
CGAACAATCCAGGCACAACACTACAAGAGCGGAATGACCCTCAGCAACTGCTATTAGATGCCAAGCACATG
TTTCCTGTTTTGTTTCCATTTAATCCATCTTCTCTAACCATGGACTCAATCCACATCCCAGCGTGTCTC
AATCTGGAATTCCTCAATGAAGTCTGA

Regulatory Elements

[0242] The lentiviral vector of the present invention may further comprise one or more regulatory elements which may act pre- or post-transcriptionally. Suitably, the protein-coding sequence is operably linked to one or more regulatory elements which may act pre- or post-transcriptionally. The one or more regulatory elements may facilitate expression of the transgene in liver cells (e.g. hepatocytes).

[0243] As used herein, a “regulatory element” may refer any nucleotide sequence that facilitates expression of a polypeptide, e.g. acts to increase expression of a transcript or to enhance mRNA stability. Suitable regulatory elements include for example promoters, enhancer elements, post-transcriptional regulatory elements, polyadenylation sites, and Kozak sequences.

Promoter

[0244] The lentiviral vector of the present invention may comprise a promoter, preferably a liver-specific (e.g. hepatocyte-specific) promoter. Suitably, the protein-coding sequence is operably linked to a promoter, preferably a liver-specific (e.g. hepatocyte-specific) promoter.

[0245] A “promoter” may refer to a region of DNA that leads to initiation of transcription of a gene. Promoters are located near the transcription start sites of genes, upstream on the DNA (towards the 5' region of the sense strand).

[0246] As used herein, a “tissue-specific promoter” may refer to a promoter which preferentially facilitates expression of a transgene in a specific type of cells or tissue. Suitably, a tissue-specific promoter may facilitate higher expression of a transgene in one cell-type as compared to other cell-types. Higher expression may be measured for example by measuring the expression of a transgene, e.g. green fluorescence protein (GFP), operably linked to the promoter, wherein expression of the transgene correlates with the ability of the promoter to facilitate expression of a gene. For example, a tissue-specific promoter may be a promoter which facilitates transgene expression levels at least 10% higher, at least 20% higher, at least 30% higher, at least 40% higher, at least 50% higher, at least 100% higher, at least 200% higher, at least 300% higher, at least 400% higher, at least 500% higher, or at least 1000% higher in one cell-type as compared to expression levels in other cell-types.

[0247] In some embodiments, the promoter is a liver-specific promoter. In some embodiments, the promoter is a hepatocyte-specific promoter.

[0248] Suitably, the promoter may be (or may be derived from) a promoter associated with a gene with selective expression in human liver cells (e.g. hepatocytes). Suitably, the promoter may be (or may be derived from) a promoter associated with a gene with selective expression in human hepatocyte cells. Methods to identify promoters associated with genes will be well known to those of skill in the art.

[0249] Exemplary liver-specific and/or hepatocyte-specific promoters are described in Kattenhorn, L. M., et al.,

2016. Human gene therapy, 27(12), pp. 947-961 and include transthyretin (TTR) promoters, alpha-1-antitrypsin (AAT) promoters, thyroxine-binding globulin (TBG) promoters, APoE/hAAT promoters, HCR-hAAT promoters, LP1 promoters, and HLP promoters.

[0250] An engineered promoter variant derived from any of these promoters may be used, provided that the variant retains the capacity to drive liver-specific and/or hepatocyte-specific expression of a transgene which is operably coupled to the promoter. A skilled person will be able to arrive at such variants using methods known in the art. The variant may have at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identity to any of the promoters.

[0251] A fragment of any of these promoters (or variants thereof) may be used, provided that the fragment retains the capacity to drive liver-specific and/or hepatocyte-specific expression of a transgene which is operably coupled to the promoter. A skilled person will be able to arrive at such fragments using methods known in the art. The fragment may be, for example, at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, at least 500 nucleotides, or at least 1000 nucleotides in length.

[0252] In some embodiments, the promoter is selected from the group consisting of: a transthyretin (TTR) promoter, an alpha-1-antitrypsin (AAT) promoter, a thyroxine-binding globulin (TBG) promoter, an APoE/hAAT promoter, a HCR-hAAT promoter, a LP1 promoter, and a HLP promoter.

[0253] In some embodiments, the promoter is a TTR promoter, or a variant and/or fragment thereof. In some embodiments, the promoter is an Enh1mTTR (ET) promoter, or a variant and/or fragment thereof.

[0254] An exemplary ET promoter is provided in GenBank accession number AY661265. In some embodiments, the ET promoter comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 19 or a fragment thereof. Suitably, the ET promoter comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 19 or a fragment thereof.

[0255] In some embodiments, the ET promoter comprises or consists of the nucleotide sequence SEQ ID NO: 19 or a fragment thereof.

TABLE-US-00026 Exemplary ET promoter (SEQ ID NO: 19)

```
CGCGAGTTAATAATTACCAGCGCGGGCCAAATAAATAATCCGCGAGGGGC
AGGTGACGTTTGGCCAGCGCGCGCTGGTAATTATTAACCTCGCGAATATT
GATTCGAGGCCGCGATTGCCGCAATCGCGAGGGGCAGGTGACCTTTGCCC
AGCGCGCGTTCCGCCCCGCCCGGACGGTATCGATAAGCTTAGGAGCTTGG
GCTGCAGGTCGAGGGCACTGGGAGGATGTTGAGTAAGATGGAAACTACT
GATGACCCTTGACAGAGACAGAGTATTAGGACATGTTTGAACAGGGGCCGG
GCGATCAGCAGGTAGCTCTAGAGGATCCCCGTCTGTCTGCACATTTCTGTA
GAGCGAGTGTTCCGATACTCTAATCTCCCTAGGCAAGGTTTCATATTTGTG
TAGGTTACTTATTCTCCTTTTGTGACTAAGTCAATAATCAGAATCAGCA
GGTTTGGAGTCAGCTTGGCAGGGATCAGCAGCCTGGGTTGGAAGGAGGGG
GTATAAAAGCCCCCTTCACCAGGAGAAGCCGTCACACAGATCCACAAGCTC CTG
```

[0256] In some embodiments, the promoter is an AAT promoter, or a variant and/or fragment thereof. In some embodiments, the promoter is a human AAT (hAAT) promoter, or a variant and/or fragment thereof.

[0257] In some embodiments, the hAAT promoter comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 20 or a fragment thereof. Suitably, the hAAT promoter comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 20 or a fragment thereof.

[0258] In some embodiments, the hAAT promoter comprises or consists of the nucleotide sequence SEQ ID NO: 20 or a fragment thereof.

TABLE-US-00027 Exemplary hAAT promoter (SEQ ID NO: 20)

```
GATCTTGCTACCAGTGGAACAGCCACTAAGGATTCTGCAGTGAGAGCAGA
GGGCCAGCTAAGTGGTACTCTCCAGAGACTGTCTGACTCACGCCACCCC
CTCCACCTTGACACAGGACGCTGTGGTTTCTGAGCCAGGTACAATGACT
CCTTTCGGTAAGTGCAGTGGAAGCTGTACACTGCCAGGCAAAGCGTCCG
GGCAGCGTAGGCGGGCGACTCAGATCCCAGCCAGTGGACTTAGCCCCCTGT
TTGCTCCTCCGATAACTGGGGTGACCTTGGTTAATATTCACCAGCAGCCT
CCCCCGTTGCCCTCTGGATCCACTGCTTAAATACGGACGAGGACAG
```

[0259] The promoter may be a constitutive promoter. As used herein, a “constitutive promoter” is a promoter which is always active.

[0260] Alternatively, the promoter may be an inducible promoter. As used herein, an “inducible promoter” is a promoter which is only active under specific conditions. For example, expression of the transgene may be induced

by a small molecule or drug (e.g. which binds to a promoter, regulatory sequence or to a transcriptional repressor or activator molecule) or by using an environmental trigger. Types of inducible promoter include chemically-inducible promoters (e.g. a Tet-on system); temperature-inducible promoters (e.g. Hsp70 or Hsp90-derived promoters); and light-inducible promoters. Suitably, the promoter is chemically-inducible. Any suitable method for engineering an inducible promoter may be used.

Enhancer Elements

[0261] The lentiviral vector of the present invention may comprise an enhancer, preferably a liver-specific (e.g. hepatocyte-specific) enhancer. Suitably, the protein-coding sequence is operably linked to an enhancer, preferably a liver-specific (e.g. hepatocyte-specific) enhancer.

[0262] An “enhancer” or “enhancer element” may refer a region of DNA that can be bound by proteins (activators) to increase the likelihood that transcription of a particular gene will occur. Enhancers are cis-acting. They can be located up to 1 Mbp (1,000,000 bp) away from the gene, upstream or downstream from the start site.

[0263] As used herein, a “tissue-specific enhancer” is an enhancer which preferentially facilitates expression of a gene in specific cells or tissues. Suitably, a tissue-specific enhancer may facilitate higher expression of a gene in specific cells-types as compared to other cell-types. Higher expression may be measured for example by measuring the expression of a transgene, e.g. green fluorescence protein (GFP), operably linked to the enhancer, wherein expression of the transgene correlates with the ability of the enhancer to facilitate expression of a gene. For example, a tissue-specific enhancer may be an enhancer which facilitates gene expression levels at least 10% higher, at least 20% higher, at least 30% higher, at least 40% higher, at least 50% higher, at least 100% higher, at least 200% higher, at least 300% higher, at least 400% higher, at least 500% higher, or at least 1000% higher in a specific cell-type compared to expression levels in other cell types.

[0264] Suitable tissue-specific enhancers will be well known to those of skill in the art. The enhancer may be a liver-specific enhancer, preferably a hepatocyte-specific enhancer.

[0265] Suitably, the enhancer may be (or may be derived from) an enhancer associated with a gene with selective expression in human liver cells (e.g. hepatocytes). Suitably, the enhancer may be (or may be derived from) an enhancer associated with a gene with selective expression in human hepatocyte cells. Methods to identify the enhancer regions associated with genes will be well known to those of skill in the art.

[0266] Exemplary liver-specific and/or hepatocyte-specific enhancers are described in Kramer, M. G., et al., 2003. *Molecular therapy*, 7(3), pp. 375-385, and include enhancer regions of the albumin, α 1-antitrypsin, hepatitis B virus core protein, and hemopexin genes. Other liver-specific and/or hepatocyte-specific enhancers include apolipoprotein E (APoE) enhancers, hepatic control region (HCR) enhancers, and alpha-1-antitrypsin (AAT) enhancers.

[0267] An engineered enhancer variant derived from any of these enhancers may be used, provided that the variant retains the capacity to drive liver-specific and/or hepatocyte-specific expression of a transgene which is operably coupled to the enhancer. A skilled person will be arrive at such variants using methods known in the art. The variant may have at least 70%, at least 80%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identity to any of the enhancers.

[0268] A fragment of any of these enhancers (or variants thereof) may be used, provided that the fragment retains the capacity to drive liver-specific and/or hepatocyte-specific expression of a transgene which is operably coupled to the enhancer. A skilled person will be able to arrive at such fragments using methods known in the art. The fragment may be at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, at least 500 nucleotides, or at least 1000 nucleotides in length.

[0269] The vector of the present invention may comprise a liver-specific promoter and/or a liver-specific enhancer, i.e. a liver-specific promoter and/or enhancer. Suitably, the protein-coding sequence is operably linked to a liver-specific promoter and/or enhancer. Suitably, the protein-coding sequence is operably linked to a hepatocyte-specific promoter and/or enhancer. The promoter and enhancer may be a combination of any of the above, for example a hAAT promoter and an ApoE or HCR enhancer.

Post-Transcriptional Regulatory Elements

[0270] The lentiviral vector of the present invention may comprise one or more further post-transcriptional regulatory elements (e.g. in addition to one or more miRNA target sequence). Suitably, the protein-coding sequence is operably linked to one or more further post-transcriptional regulatory elements. The further post-transcriptional regulatory element may improve gene expression.

[0271] The lentiviral vector of the present invention may comprise a Woodchuck Hepatitis Virus Post-transcriptional Regulatory Element (WPRE). Suitably, the protein-coding sequence is operably linked to a WPRE.

[0272] Suitable WPRE sequences will be well known to those of skill in the art (see e.g. Zufferey, R., et al., 1999. *Journal of virology*, 73(4), pp. 2886-2892; and Zanta-Boussif, M. A. et al., 2009. *Gene therapy*, 16(5), pp. 605-619). Suitably, the WPRE is a wild-type WPRE or is a mutant WPRE. For example, the WPRE may be mutated to

abrogate translation of the woodchuck hepatitis virus X protein (WHX) e.g. by mutating the WHX ORF translation start codon.

[0273] In some embodiments, the WPRE comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 21 or a fragment thereof. Suitably, the WPRE comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 21 or a fragment thereof.

[0274] In some embodiments, the WPRE comprises or consists of the nucleotide sequence SEQ ID NO: 21 or a fragment thereof.

TABLE-US-00028 Exemplary WPRE (SEQ ID NO: 21)

```
AATCAACCTCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAA
CTATGTTGCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGT
ATCATGCTATTGCTTCCCGTATGGCTTTCATTTTCTCCTCCTTGTATAAA
TCCTGGTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACG
TGGCGTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCA
TTGCCACCACCTGTCAGCTCCTTTCGGGACTTTCGCTTTCCTCCCTCCCT
ATTGCCACGGCGGAACATCGCCGCCTGCCTTGCCCGCTGCTGGACAGG
GGCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTGTCGGGGAAATCAT
CGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGGG
ACGTCCTTCTGCTACGTCCTTCGGCCCTCAATCCAGCGGACCTTCCTTC CCGC
```

Polyadenylation Sequence

[0275] The lentiviral vector of the present invention may comprise a polyadenylation sequence. Suitably, the protein-coding sequence is operably linked to a polyadenylation sequence. A polyadenylation sequence may be inserted after the protein-coding sequence to improve transgene expression.

[0276] A polyadenylation sequence typically comprises a polyadenylation signal, a polyadenylation site and a downstream element: the polyadenylation signal comprises the sequence motif recognised by the RNA cleavage complex; the polyadenylation site is the site of cleavage at which a poly-A tails is added to the mRNA; the downstream element is a GT-rich region which usually lies just downstream of the polyadenylation site, which is important for efficient processing.

[0277] Suitable polyadenylation sequences will be well known to those of skill in the art (see e.g. Schambach, A., et al., 2007. Molecular Therapy, 15(6), pp. 1167-1173; and Choi, J. H. et al., 2014. Molecular brain, 7(1), pp. 1-10). Exemplary polyadenylation sequences include the bGH poly(A) signal sequence and SV40pA signal sequence.

[0278] Suitably, the polyadenylation sequence may be present in the 3' LTR (i.e. the lentiviral vector does not comprise an additional polyadenylation sequence).

Kozak Sequence

[0279] The lentiviral vector of the present invention may comprise a Kozak sequence. Suitably, the protein-coding sequence is operably linked to a Kozak sequence. A Kozak sequence may be inserted before the start codon to improve the initiation of translation.

[0280] Suitable Kozak sequences will be well known to those of skill in the art (see e.g. Kozak, M., 1987. Nucleic acids research, 15(20), pp. 8125-8148).

[0281] In some embodiments, the Kozak sequence comprises or consists of a nucleotide sequence which is at least 80% identical to SEQ ID NO: 22 or a fragment thereof.

[0282] In some embodiments, the Kozak sequence comprises or consists of the nucleotide sequence SEQ ID NO: 22 or a fragment thereof.

TABLE-US-00029 (SEQ ID NO: 22) GCCACC Exemplary Kozak sequence

Other Cis-Acting Elements

[0283] The lentiviral vector of the present invention may comprise any other suitable cis-acting elements, such as one or more of a rev response element (RRE); a retroviral psi packaging element; a primer binding site (PBS); a TAT activation region (TAR); splice donor and acceptor sites; and central and terminal polypurine tracts.

Long Terminal Repeats (LTRs)

[0284] The lentiviral vector of the present invention may comprise one or more long terminal repeat (LTR). As described above, LTRs are responsible for proviral integration and transcription. Typically, a naturally occurring LTR comprises U3, R, and U5 regions.

[0285] The lentiviral vector may comprise a 5' LTR and/or a 3' LTR. The lentiviral vector may comprise a 5' LTR and a 3' LTR. Suitably, a 5' LTR comprises R and U5 regions, and optionally comprise a U3 region. Suitably, a 3' LTR comprises U3, R, and U5 regions.

[0286] Suitable LTR sequences will be well known to those of skill in the art (see e.g. Frech, K., et al., 1996).

Virology, 224(1), pp. 256-267).

[0287] In some embodiments, a LTR comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 23 or a fragment thereof. Suitably, a LTR comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 23 or a fragment thereof.

[0288] In some embodiments, a LTR comprises or consists of the nucleotide sequence SEQ ID NO: 23 or a fragment thereof.

TABLE-US-00030 (SEQ ID NO: 23)

TGGAAGGGCTAATTCACCTCCCAACGAAGACAAGATCTGCTTTTTGCTTG
TACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGC
TAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGC
TTCAAGTAGTGTGTGCCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATC
CCTCAGACCCCTTTAGTCAGTGTGGAAAATCTCTAGCAG Exemplary LTR

[0289] The lentiviral vector of the present invention may comprise one or more self-inactivating long terminal repeat (SIN-LTR). A "SIN-LTR" may comprise a deletion that abolishes transcription of the full-length virus after it has incorporated into a host cell. For example, a 3' SIN-LTR may comprise a deletion in the U3 region removing the promoter/enhancer elements (see e.g. Zufferey, R., et al., 1998. Journal of virology, 72(12), pp. 9873-9880). This deletion is copied into the 5' LTR after reverse transcription, thereby making the gene expression in target cells dependent on an internal promoter of choice.

[0290] Suitable SIN-LTR sequences will be well known to those of skill in the art (see e.g. Zufferey, R., et al., 1998. Journal of virology, 72(12), pp. 9873-9880 and Miyoshi, H., et al., 1998. Journal of virology, 72(10), pp. 8150-8157).

[0291] In some embodiments, the 5' LTR comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 24 or a fragment thereof. Suitably, the 5' LTR comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 24 or a fragment thereof.

[0292] In some embodiments, the 5' LTR comprises or consists of the nucleotide sequence SEQ ID NO: 24 or a fragment thereof.

TABLE-US-00031 (SEQ ID NO: 24)

GGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAAC
TAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCA
AGTAGTGTGTGCCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTC
AGACCCCTTTAGTCAGTGTGGAAAATCTCTAGCAG Exemplary 5' LTR

[0293] In some embodiments, the 5' LTR and/or the 3' LTR comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 25 or a fragment thereof. Suitably, the 5' LTR and/or the 3' LTR comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 25 or a fragment thereof.

[0294] In some embodiments, the 5' LTR and/or the 3' LTR comprises or consists of the nucleotide sequence SEQ ID NO: 25 or a fragment thereof.

[0295] In some embodiments, the 5' LTR and the 3' LTR comprise or consist of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 25 or a fragment thereof. Suitably, the 5' LTR and the 3' LTR comprise or consist of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 25 or a fragment thereof.

[0296] In some embodiments, the 5' LTR and the 3' LTR comprise or consist of the nucleotide sequence SEQ ID NO: 25 or a fragment thereof.

TABLE-US-00032 (SEQ ID NO: 25)

TGGAAGGGCTAATTCACCTCCCAACGAAGACAAGATCTGCTTTTTGCTTG
TACTGGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGC
TAACTAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGC
TTCAAGTAGTGTGTGCCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATC
CCTCAGACCCCTTTAGTCAGTGTGGAAAATCTCTAGCAG Exemplary 3' LTR
Primer Binding Site (PBS)

[0297] The lentiviral vector of the present invention may comprise a primer binding site (PBS). A PBS is a cis-acting element where a primer may bind to initiate reverse transcription of the RNA genome (see e.g. Lanchy, J. M., et al., 1998. Journal of Biological Chemistry, 273(38), pp. 24425-24432).

[0298] Suitable retroviral PBSs will be well known to those of skill in the art.

[0299] In some embodiments, a PBS comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 26 or a fragment thereof. Suitably, a PBS comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 26 or a fragment thereof.

[0300] In some embodiments, a PBS comprises or consists of the nucleotide sequence SEQ ID NO: 26 or a fragment thereof.

TABLE-US-00033 (SEQ ID NO: 26)

TGGCGCCCGAACAGGGACTTGAAAGCGAAAGGGAAACCAGAGGAGCTCT

CTCGACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGG

GCGGCGACTGGTGAGTACGCCAAAAATTTTGA CTAGCGGAGGCTAGAAG GAGAGAG Exemplary primer binding site

[0301] In some embodiments, a PBS comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 27 or a fragment thereof. Suitably, a PBS comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 27 or a fragment thereof.

[0302] In some embodiments, a PBS comprises or consists of the nucleotide sequence SEQ ID NO: 27 or a fragment thereof.

TABLE-US-00034 (SEQ ID NO: 27)

TGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTC

GACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCG

GCGACTGGTGAGTACGCCAAAAATTTTGA CTAGCGGAGGCTAGAAGGAG AGAG Exemplary primer binding site

Retroviral Psi Packaging Element

[0303] The lentiviral vector of the present invention may comprise a retroviral psi packaging element. As described above, a retroviral psi packaging element is a cis-acting element which is involved in regulating the essential process of packaging the retroviral RNA genome into the viral capsid during replication (see e.g. McBride, M. S., et al., 1997. Journal of virology, 71(6), pp. 4544-4554). A retroviral psi packaging element may form part of the 5' region of the gag gene.

[0304] Suitable retroviral psi packaging elements will be well known to those of skill in the art.

[0305] In some embodiments, a retroviral psi packaging element comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 28 or a fragment thereof. Suitably, a retroviral psi packaging element comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 28 or a fragment thereof.

[0306] In some embodiments, a retroviral psi packaging element comprises or consists of the nucleotide sequence SEQ ID NO: 28 or a fragment thereof.

TABLE-US-00035 (SEQ ID NO: 28)

ATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATG

GGAAAAAATTTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAA

CATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTG

GCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACA

ACCATCCCTTCAGACAGGATCAGAAGAACTTAGATCATTATATAATACA

GTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCA

AGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCAC

CGCACAGCAAGCGGCCGCTGAT Exemplary retroviral psi packaging element

Rev Response Element (RRE)

[0307] The lentiviral vector of the present invention may comprise a rev response element (RRE). As described above, a RRE is a cis-acting element that enables the efficient export of RNA transcripts of the integrated provirus from the nucleus to the cytoplasm of an infected target cell (see e.g. Pollard, V. W. and Malim, M. H., 1998. Annual review of microbiology, 52(1), pp. 491-532).

[0308] Suitable RRE sequences will be well known to those of skill in the art.

[0309] In some embodiments, a RRE comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 29 or a fragment thereof. Suitably, a RRE comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 29 or a fragment thereof.

[0310] In some embodiments, a RRE comprises or consists of the nucleotide sequence SEQ ID NO: 29 or a fragment thereof.

TABLE-US-00036 (SEQ ID NO: 29)

GGAGCTTTGTTCCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCG
CAGCGTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTAT
AGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCAT
CTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCC
TGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATT Exemplary rev response
element

[0311] In some embodiments, a RRE comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 30 or a fragment thereof. Suitably, a RRE comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 30 or a fragment thereof.

[0312] In some embodiments, a RRE comprises or consists of the nucleotide sequence SEQ ID NO: 30 or a fragment thereof.

TABLE-US-00037 (SEQ ID NO: 30)

GGAGCTTTGTTCCCTTGGGTTCTTGGGAGCAGCAGGAAGCACTATGGGCG
CAGCCTCAATGACGCTGACGGTACAGGCCAGACAATTATTGTCTGGTAT
AGTGCAGCAGCAGAACAATTTGCTGAGGGCTATTGAGGCGCAACAGCAT
CTGTTGCAACTCACAGTCTGGGGCATCAAGCAGCTCCAGGCAAGAATCC
TGGCTGTGGAAAGATACCTAAAGGATCAACAGCTCCTGGGGATT Exemplary rev response
element

Central Polypurine Tract (cPPT)

[0313] The lentiviral vector of the present invention may comprise a central polypurine tract (cPPT). As described above, a cPPT may allow initiation of plus-strand synthesis (see e.g. Follenzi, A. et al., 2000. Nature genetics, 25(2), pp. 217-222).

[0314] Suitable cPPT sequences will be well known to those of skill in the art.

[0315] In some embodiments, a cPPT comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 31 or a fragment thereof. Suitably, a cPPT comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 31 or a fragment thereof.

[0316] In some embodiments, a cPPT comprises or consists of the nucleotide sequence SEQ ID NO: 31 or a fragment thereof.

TABLE-US-00038 (SEQ ID NO: 31)

AACTTTTAAAAGAAAAGGGGGGATTGGGGGGTACAGTGCAGGGGAAAGA
ATAGTAGACATAATAGCAACAGACATACAACTAAAGAATTACAAAAC
AAATTACAAAATTCAAAATTTTATC Exemplary central polypurine tract

Other Elements

[0317] The lentiviral vector of the present invention may comprise any other suitable elements.

[0318] In some embodiments, the lentiviral vector of the present invention comprises an element comprising or consisting of a nucleotide sequence which is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 32 or a fragment thereof. In some embodiments, the lentiviral vector of the present invention comprises an element comprising or consisting of the nucleotide sequence of SEQ ID NO: 32 or a fragment thereof.

TABLE-US-00039 (SEQ ID NO: 32)

TCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGGAGAAGTGAATT
ATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAGTAGCACCCACC
AAGGCAAAGAGAAGAGTGGTGCAGAGAGAAAAAAGAGCAGTGGGAATA Exemplary delta ENV1

[0319] In some embodiments, the lentiviral vector of the present invention comprises an element comprising or consisting of a nucleotide sequence which is at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 33 or a fragment thereof. In some embodiments, the lentiviral vector of the present invention comprises an element comprising or consisting of the nucleotide sequence of SEQ ID NO: 33 or a fragment thereof.

TABLE-US-00040 (SEQ ID NO: 33)

GGGTTGCTCTGGAAACTCATTTGCACCACTGCTGTGCCTTGAATGCT
AGTTGGAGTAATAAATCTCTGGAACAGATTTGGAATCACACGACCTGGA
TGGAGTGGGACAGAGAAATTAACAATTACACAAGCTTAATACACTCCTT
AATTGAAGAATCGCAAAACCAGCAAGAAAAGAATGAACAAGAATTATTG
GAATTAGATAAATGGGCAAGTTTGTGGAATTGGTTTAACATAACAAATT

GGCTGTGGTATATAAATATTCTATAATGATAGTAGGAGGCTTGCTAGG
TTTAAGAATAGTTTTTGTCTGTACTTTCTATAGTGAATAGAGTTAGGCAG
GGATATTCACCATTATCGTTTCAGACCCACCTCCCAACCCCGAGGGGAC
CCGACAGGCCCGAAGGAATAGAAGAAGAAGGTGGAGAGAGAGACAGAGA
CAGATCCATTTCGATTAGTGAACGGATC Exemplary delta ENV2

Exemplary Cis-Acting Elements

[0320] The lentiviral vector of the present invention may comprise a cis-acting element comprising a PBS, a retroviral psi packaging element, and a rev response element (RRE).

[0321] In some embodiments, a cis-acting element comprising a PBS, a retroviral psi packaging element, and a RRE comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 34 or a fragment thereof. Suitably, a cis-acting element comprising a PBS, a retroviral psi packaging element, and a RRE comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 34 or a fragment thereof.

[0322] In some embodiments, a cis-acting element comprising a PBS, a retroviral psi packaging element, and a RRE comprises or consists of the nucleotide sequence SEQ ID NO: 34 or a fragment thereof.

TABLE-US-00041 (SEQ ID NO: 34)

TGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTC
GACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCG
GCGACTGGTGAGTACGCCAAAAATTTGACTAGCGGAGGCTAGAAGGAG
AGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGC
GATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATT
AAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAAT
CCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGC
TACAACCATCCCTTCAGACAGGATCAGAAGAAGCTTAGATCATTATATAA
TACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGAC
ACCAAGGAAGCTTTAGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGA
CCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATA
TGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAAT
TGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGACAG
AGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTCCCTTGGGTTCTTGG
GAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACA
GGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAAATTTGCTG
AGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCA
TCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGA
TCAACAGCTCCTGGGGATTT Exemplary cis-acting element comprising a PBS, a retroviral psi packaging element, and a RRE

[0323] The lentiviral vector of the present invention may comprise a cis-acting element comprising a PBS, a retroviral psi packaging element, a rev response element (RRE), and a central polypurine tract (cPPT).

[0324] In some embodiments, a cis-acting element comprising a PBS, a retroviral psi packaging element, a RRE, and a cPPT comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 35 or a fragment thereof. Suitably, a cis-acting element comprising a PBS, a retroviral psi packaging element, a RRE, and a cPPT comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 35 or a fragment thereof.

[0325] In some embodiments, a cis-acting element comprising a PBS, a retroviral psi packaging element, a RRE, and a cPPT comprises or consists of the nucleotide sequence SEQ ID NO: 35 or a fragment thereof.

TABLE-US-00042 (SEQ ID NO: 35)

TGGCGCCCGAACAGGGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTC
GACGCAGGACTCGGCTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCG
GCGACTGGTGAGTACGCCAAAAATTTGACTAGCGGAGGCTAGAAGGAG
AGAGATGGGTGCGAGAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGC
GATGGGAAAAAATTCGGTTAAGGCCAGGGGGAAAGAAAAAATATAAATT
AAAACATATAGTATGGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAAT
CCTGGCCTGTTAGAAACATCAGAAGGCTGTAGACAAATACTGGGACAGC
TACAACCATCCCTTCAGACAGGATCAGAAGAAGCTTAGATCATTATATAA
TACAGTAGCAACCCTCTATTGTGTGCATCAAAGGATAGAGATAAAAGAC

ACCAAGGTAAGCTTAAAGAGCAAAACAAAAGTAAGA
CCACCGCACAGCAAGCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATA
TGAGGGACAATTGGAGAAGTGAATTATATAAATATAAAGTAGTAAAAAT
TGAACCATTAGGAGTAGCACCCACCAAGGCAAAGAGAAGAGTGGTGCAG
AGAGAAAAAAGAGCAGTGGGAATAGGAGCTTTGTTTCCTTGGGTTCTTGG
GAGCAGCAGGAAGCACTATGGGCGCAGCCTCAATGACGCTGACGGTACA
GGCCAGACAATTATTGTCTGGTATAGTGCAGCAGCAGAACAATTTGCTG
AGGGCTATTGAGGCGCAACAGCATCTGTTGCAACTCACAGTCTGGGGCA
TCAAGCAGCTCCAGGCAAGAATCCTGGCTGTGGAAAGATACCTAAAGGA
TCAACAGCTCCTGGGGATTGTTGGGGTTGCTCTGGAAAACCTCATTTGCACC
ACTGCTGTGCCTTGAATGCTAGTTGGAGTAATAAATCTCTGGAACAGA
TTTGGAATCACACGACCTGGATGGAGTGGGACAGAGAAATTAACAATTA
CACAAGCTTAATACACTCCTTAATTGAAGAATCGCAAACCCAGCAAGAA
AAGAATGAACAAGAATTATTGGAATTAGATAAATGGGCAAGTTTGTGGA
ATTGGTTTAACATAACAAATTGGCTGTGGTATATAAAATTATTCATAAT
GATAGTAGGAGGCTTGGTAGGTTTAAGAATAGTTTTTGCTGTACTTTCT
ATAGTGAATAGAGTTAGGCAGGGATATTCACCATTATCGTTTCAGACCC
ACCTCCCAACCCCGAGGGGACCCGACAGGCCCGAAGGAATAGAAGAAGA
AGGTGGAGAGAGAGACAGAGACAGATCCATTGATTAGTGAACGGATCT
CGACGGTATCGGTTAACTTTTAAAAGAAAAGGGGGGATTGGGGGGTACA
GTGCAGGGGAAAGAATAGTAGACATAATAGCAACAGACATACAAATAA
AGAATTACAAAAACAAATTACAAAAATTCAAAATTTTATC Exemplary cis-acting element
comprising a PBS, a retroviral psi packaging element, a RRE, and a cPPT
Exemplary Lentiviral Genomes

[0326] The lentiviral genome of the present invention may comprise from 5' to 3': a 5' LTR, one or more cis-acting elements, a protein-coding sequence, optionally one or more post-transcriptional regulatory sequences, and a 3' LTR.

[0327] For example, the lentiviral genome of the present invention may comprise from 5' to 3': a 5' LTR, a PBS, a retroviral psi packaging element, a RRE, a cPPT, a liver-specific promoter, a protein-coding sequence, a WPRE, one or more miRNA target sequence, and a 3' LTR.

[0328] The lentiviral genome of the present invention may further comprise any other suitable elements, such as any other elements described herein or one or more spacer sequence. The spacer sequence(s) may comprise, for example, at least one (e.g. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10), at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, or at least ten nucleotide bases.

[0329] In some embodiments, the lentiviral genome comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 36 or a fragment thereof. Suitably, the lentiviral genome comprises or consists of a nucleotide sequence which is at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% identical to SEQ ID NO: 36 or a fragment thereof.

[0330] In some embodiments, the lentiviral genome comprises or consists of the nucleotide sequence SEQ ID NO: 36 or a fragment thereof.

TABLE-US-00043 (SEQ ID NO: 36)

GGGTCTCTCTGGTTAGACCAGATCTGAGCCTGGGAGCTCTCTGGCTAAC
TAGGGAACCCACTGCTTAAGCCTCAATAAAGCTTGCCTTGAGTGCTTCA
AGTAGTGTGTGCCCGTCTGTTGTGTGACTCTGGTAACTAGAGATCCCTC
AGACCCTTTTAGTCAGTGTGGAAAATCTCTAGCAGTGGCGCCCGAACAG
GGACCTGAAAGCGAAAGGGAAACCAGAGCTCTCTCGACGCAGGACTCGG
CTTGCTGAAGCGCGCACGGCAAGAGGCGAGGGGCGGCGACTGGTGAGTA
CGCCAAAAATTTTGACTAGCGGAGGCTAGAAGGAGAGAGATGGGTGCGA
GAGCGTCAGTATTAAGCGGGGGAGAATTAGATCGCGATGGGAAAAAATT
CGGTAAAGGCCAGGGGGAAAGAAAAAATATAAATTAAAACATATAGTAT
GGGCAAGCAGGGAGCTAGAACGATTCGCAGTTAATCCTGGCCTGTTAGA
AACATCAGAAGGCTGTAGACAAATACTGGGACAGCTACAACCATCCCTT
CAGACAGGATCAGAAGAACTTAGATCATTATATAATACAGTAGCAACCC
TCTATTGTGTGCATCAAAGGATAGAGATAAAAGACACCAAGGAAGCTTT
AGACAAGATAGAGGAAGAGCAAAACAAAAGTAAGACCACCGCACAGCAA
GCGGCCGCTGATCTTCAGACCTGGAGGAGGAGATATGAGGGACAATTGG
AGAAGTGAATTATATAAATATAAAGTAGTAAAAATTGAACCATTAGGAG

[illegible]

CGCCCTGAGCTAAGCTTGAATCTGAATATGGTCATCAAGCCAG
GAGAGATGACAGCCCTGGTGGGACCCTCTGGAGCAGGCAAGAGCACCGC
CCTGCAGCTGATCCAGCGGTTTTACGACCCTTGCGAGGGAATGGTGACC
GTGGACGGACACGACATCAGGTCCCTGAACATCCAGTGGCTGCGCGATC
AGATCGGCATCGTGGAGCAGGAGCCAGTGCTGTTCTCTACCACAATCGC
CGAGAATATCAGATACGGCCGCGAGGATGCCACAATGGAGGACATCGTG
CAGGCCGCCAAGGAGGCCAACGCCTATAACTTCATCATGGATCTGCCCC
AGCAGTTCGACACCCTGGTGGGAGAGGGAGGAGGACAGATGTCCGGAGG
CCAGAAGCAGAGAGTGGCCATCGCCAGAGCCCTGATCCGCAACCCTAAG
ATCCTGCTGCTGGATATGGCCACAAGCGCCCTGGACAATGAGTCCGAGG
CCATGGTGAGGAGGTGCTGAGCAAGATCCAGCACGGCCACACCATCAT
CTCTGTGGCACACCGGCTGAGCACAGTGAGAGCAGCCGACACCATCATC
GGCTTTGAGCACGGAACAGCAGTGAGAGAGGGGAACCCACGAGGAGCTGC
TGGAGAGGAAGGGCGTGACTTCACCCTGGTGACACTGCAGAGCCAGGG
CAACCAGGCCCTGAATGAGGAGGACATCAAGGATGCCACAGAGGACGAT
ATGCTGGCCCCGGACCTTCTCCAGAGGCTCTTATCAGGATTCTCTGCGGG
CCAGCATCAGGCAGCGCAGCAAGTCCCAGCTGAGCTACCTGGTGACGA
GCCACCTCTGGCCGTGGTGGACCACAAGTCCACCTATGAGGAGGATAGA
AAGGACAAGGACATCCCCGTGCAGGAGGAGGTGGAGCCTGCCCCAGTGA
GGCGCATCCTGAAGTTTTCCGCCCCAGAGTGGCCCTACATGCTGGTGGG
ATCTGTGGGAGCAGCAGTGAACGGCACCGTGACACCACTGTATGCCTTC
CTGTTTTCCAGATCCTGGGCACCTTCTCTATCCCCGACAAGGAGGAGC
AGAGGTCCCAGATCAATGGCGTGTGCCTGCTGTTTGTGGCCATGGGCTG
CGTGAGCCTGTTTACACAGTTCCTGCAGGGCTACGCCTTCGCCAAGAGC
GGCGAGCTGCTGACCAAGCGGCTGAGAAAGTTCGGCTTTCGCGCAATGC
TGGGACAGGACATCGCATGGTTTGACGATCTGCGGAACAGCCCAGGCGC
CCTGACCACAAGACTGGCCACAGATGCCTCTCAGGTGCAGGGAGCAGCA
GGCAGCCAGATCGGCATGATCGTGAACCTCCTTCACCAATGTGACAGTGG
CCATGATCATCGCCTTCTCTTTTAGCTGGAAGCTGTCTCTGGTCATCCT
GTGCTTCTTCCCCTTTCTGGCCCTGAGCGGAGCAACCCAGACAAGGATG
CTGACCGGCTTCGCCTCCAGAGACAAGCAGGCCCTGGAGATGGTGGGCC
AGATCACAAACGAGGCCCTGAGCAATATCAGGACCGTGGCAGGAATCGG
CAAGGAGCGGCGGTTTCATCGAGGCCCTGGAGACAGAGCTGGAGAAGCCT
TTCAAGACCGCCATCCAGAAGGCCAACATCTACGGCTTCTGCTTTGCCT
TCGCCCAGTGTATCATGTTTCATCGCCAACTCCGCCTCTTACAGATATGG
CGGCTACCTGATCAGCAATGAGGGCCTGCACTTTTCCTACGTGTTCAGA
GTGATCAGCGCCGTGGTGTCTGTCTGCCACAGCCCTGGGAAGGGCCTTCT
CCTACACCCCATCTTATGCCAAGGCCAAGATCTCTGCCGCCAGGTTCTT
TCAGCTGCTGGACCGCCAGCCACCCATCAGCGTGTACAACACAGCCGGC
GAGAAGTGGGATAATTTCCAGGGCAAGATCGACTTTGTGGATTGCAAGT
TCACCTATCCTAGCAGGCCAGACTCCCAGGTGCTGAACGGCCTGAGCGT
GTCCATCTCTCCAGGCCAGACACTGGCCTTTGTGGGCTCCTCTGGCTGT
GGCAAGAGCACCTCCATCCAGCTGCTGGAGAGGTTCTACGACCCCGATC
AGGGCAAAGTGATGATCGACGGCCACGATTCTAAGAAGGTGAACGTGCA
GTTTCTGCGCTCCAATATCGGCATCGTGTCTCAGGAGCCTGTGCTGTTT
GCCTGCAGCATCATGGATAACATCAAGTACGGCGACAATACAAAGGAGA
TCCCAATGGAGAGAGTGATCGCAGCAGCAAAGCAGGCACAGCTGCACGA
TTTCGTGATGTCCCTGCCCCGAGAAGTATGAGACAAACGTGGGCAGCCAG
GGCTCCCAGCTGTCTAGGGGCGAGAAGCAGAGGATCGCAATCGCCAGGG
CCATCGTGCGCGATCCCAAGATCCTGCTGCTGGACGAGGCCACCAGCGC
CCTGGATACAGAGTCCGAGAAGACCGTGCAGGTGGCCCTGGACAAGGCC
CGGGAGGGAAGAACATGTATCGTGATCGCCCACCGCCTGAGCACCATCC
AGAATGCCGACATCATCGCCGTGATGGCCCAGGGCGTGGTCATCGAGAA
GGGCACCCACGAGGAGCTGATGGCCCAGAAGGGCGCCTACTATAAGCTG
GTGACCACAGGCAGCCCTATCTCCTGATGAGTCGACTCGACAATCAACC
TCTGGATTACAAAATTTGTGAAAGATTGACTGGTATTCTTAACTATGTT
GCTCCTTTTACGCTATGTGGATACGCTGCTTTAATGCCTTTGTATCATG

CTATTGCTTCCGTTATGGCTATGTTTCTCCTCCTTGTATAAATCCTG
 GTTGCTGTCTCTTTATGAGGAGTTGTGGCCCGTTGTCAGGCAACGTGGC
 GTGGTGTGCACTGTGTTTGCTGACGCAACCCCCACTGGTTGGGGCATTG
 CCACCACCTGTCAGCTCCTTTCCGGGACTTTCGCTTTCCCCCTCCCTAT
 TGCCACGGCGGAACATCGCCGCCTGCCTTGCCCGCTGCTGGACAGGG
 GCTCGGCTGTTGGGCACTGACAATTCCGTGGTGTGTCGGGGAAATCAT
 CGTCCTTTCCTTGGCTGCTCGCCTGTGTTGCCACCTGGATTCTGCGCGG
 GACGTCCTTCTGCTACGTCCCTTCGGCCCTCAATCCAGCGGACCTTCCT
 TCCCGCGGCCTGCTGCCGGCTCTAGATAATCCATAAAGTAGGAAACACT
 ACACGATTCCATAAAGTAGGAAACACTACAACCGGTTCCATAAAGTAGG
 AAACACTACATCACTCCATAAAGTAGGAAACACTACACCCGGGTCGAGC
 TCGGTACCTTTAAGACCAATGACTTACAAGGCAGCTGTAGATCTTAGCC
 ACTTTTAAAGAAAAGGGGGGACTGGAAGGGCTAATCACTCCCAACG
 AAGACAAGATCTGCTTTTTGCTTGTACTGGGTCTCTCTGGTTAGACCAG
 ATCTGAGCCTGGGAGCTCTCTGGCTAACTAGGGAACCCACTGCTTAAGC
 CTCAATAAAGCTTGCCCTGAGTGCTTCAAGTAGTGTGTGCCCGTCTGTT
 GTGTGACTCTGGTAACTAGAGATCCCTCAGACCCTTTTAGTCAGTGTGG AAAATCTCTAGCA

Exemplary lentiviral genome

Variants, Derivatives, Analogues, and Fragments

[0331] In addition to the specific polypeptides and polynucleotides mentioned herein, the invention also encompasses variants, derivatives, and fragments thereof.

[0332] In the context of the invention, a “variant” of any given sequence is a sequence in which the specific sequence of residues (whether amino acid or nucleic acid residues) has been modified in such a manner that the polypeptide or polynucleotide in question retains at least one or all of its endogenous functions. A variant sequence can be obtained by addition, deletion, substitution, modification, replacement and/or variation of at least one residue present in the naturally occurring polypeptide or polynucleotide.

[0333] The term “derivative” as used herein in relation to proteins or polypeptides of the invention includes any substitution of, variation of, modification of, replacement of, deletion of and/or addition of one (or more) amino acid residues from or to the sequence, providing that the resultant protein or polypeptide retains at least one or all of its endogenous functions.

[0334] Typically, amino acid substitutions may be made, for example from 1, 2 or 3, to 10 or 20 substitutions, provided that the modified sequence retains the required activity or ability. Amino acid substitutions may include the use of non-naturally occurring analogues.

[0335] Polypeptides used in the invention may also have deletions, insertions or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent polypeptide. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues as long as the endogenous function is retained. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include asparagine, glutamine, serine, threonine and tyrosine.

[0336] Conservative substitutions may be made, for example according to the table below. Amino acids in the same block in the second column and in the same line in the third column may be substituted for each other:
 TABLE-US-00044 ALIPHATIC Non-polar G A P I L V Polar - uncharged C S T M N Q Polar - charged D E K R
 H AROMATIC F W Y

[0337] The effect of additions, deletions, substitutions, modifications, replacements and/or variations may be predicted using any suitable prediction tool e.g. SIFT (Vaser, R., et al., 2016. Nature protocols, 11(1), pp. 1-9), PolyPhen-2 (Adzhubei, I., et al., 2013. Current protocols in human genetics, 76(1), pp. 7-20), CADD (Rentzsch, P., et al., 2021. Genome medicine, 13(1), pp. 1-12), REVEL (Ioannidis, N. M., et al., 2016. The American Journal of Human Genetics, 99(4), pp. 877-885), MetaLR (Dong, C., et al., 2015. Human molecular genetics, 24(8), pp. 2125-2137), and/or MutationAssessor (Reva, B., et al., 2011. Nucleic acids research, 39(17), pp. e118-e118) or based on clinical data e.g. ClinVar (Landrum, M. J., et al., 2016. Nucleic acids research, 44(D1), pp. D862-D868). Suitable additions, deletions, substitutions, modifications, replacements and/or variations may be considered tolerated, benign, and/or likely benign.

[0338] Typically, a variant may have a certain identity with the wild type amino acid sequence or the wild type nucleotide sequence.

[0339] In the present context, a variant sequence is taken to include an amino acid sequence which may be at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85% or at least 90%

identical, suitably at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to the subject sequence. Although a variant can also be considered in terms of similarity (i.e. amino acid residues having similar chemical properties/functions), in the context of the present invention it is preferred to express in terms of sequence identity.

[0340] In the present context, a variant sequence is taken to include a nucleotide sequence which may be at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85% or at least 90% identical, suitably at least 95%, at least 96%, at least 97%, at least 98% or at least 99% identical to the subject sequence. Although a variant can also be considered in terms of similarity, in the context of the present invention it is preferred to express it in terms of sequence identity.

[0341] Suitably, reference to a sequence which has a percent identity to any one of the SEQ ID NOs detailed herein refers to a sequence which has the stated percent identity over the entire length of the SEQ ID NO referred to.

[0342] Sequence identity comparisons can be conducted by eye, or more usually, with the aid of readily available sequence comparison programs. These commercially available computer programs can calculate percent identity between two or more sequences.

[0343] Percent identity may be calculated over contiguous sequences, i.e. one sequence is aligned with the other sequence and each amino acid or nucleotide in one sequence is directly compared with the corresponding amino acid or nucleotide in the other sequence, one residue at a time. This is called an “ungapped” alignment. Typically, such ungapped alignments are performed only over a relatively short number of residues.

[0344] Although this is a very simple and consistent method, it fails to take into consideration that, for example, in an otherwise identical pair of sequences, one insertion or deletion in the amino acid or nucleotide sequence may cause the following residues or codons to be put out of alignment, thus potentially resulting in a large reduction in percent identity when a global alignment is performed. Consequently, most sequence comparison methods are designed to produce optimal alignments that take into consideration possible insertions and deletions without penalizing unduly the overall identity score. This is achieved by inserting “gaps” in the sequence alignment to try to maximise local identity.

[0345] However, these more complex methods assign “gap penalties” to each gap that occurs in the alignment so that, for the same number of identical amino acids or nucleotides, a sequence alignment with as few gaps as possible, reflecting higher relatedness between the two compared sequences, will achieve a higher score than one with many gaps. “Affine gap costs” are typically used that charge a relatively high cost for the existence of a gap and a smaller penalty for each subsequent residue in the gap. This is the most commonly used gap scoring system. High gap penalties will produce optimised alignments with fewer gaps. Most alignment programs allow the gap penalties to be modified. However, it is preferred to use the default values when using such software for sequence comparisons. For example when using the GCG Wisconsin Bestfit package the default gap penalty for amino acid sequences is -12 for a gap and -4 for each extension.

[0346] Calculation of maximum percent identity therefore firstly requires the production of an optimal alignment, taking into consideration gap penalties. A suitable computer program for carrying out such an alignment is the GCG Wisconsin Bestfit package (see e.g. Devereux, J., et al., 1984. *Nucleic acids research*, 12(1), pp. 387-395). Examples of other software that can perform sequence comparisons include, but are not limited to, the BLAST package (see e.g. Altschul, S. F., et al., 1990. *Journal of molecular biology*, 215(3), pp. 403-410), BLAST 2 (see e.g. Tatusova, T. A. and Madden, T. L., 1999. *FEMS microbiology letters*, 174(2), pp. 247-250), FASTA (see e.g. Pearson, W. R. and Lipman, D. J., 1988. *PNAS*, 85(8), pp. 2444-2448.), EMBOSS Needle (Madeira, F., et al., 2019. *Nucleic acids research*, 47(W1), pp. W636-W641) and the GENWORKS suite of comparison tools. For some applications, it is preferred to use EMBOSS Needle.

[0347] Although the final percent identity can be measured, the alignment process itself is typically not based on an all-or-nothing pair comparison. Instead, a scaled similarity score matrix is generally used that assigns scores to each pairwise comparison based on chemical similarity or evolutionary distance. An example of such a matrix commonly used is the BLOSUM62 matrix.

[0348] Once the software has produced an optimal alignment, it is possible to calculate percent sequence identity. The software typically does this as part of the sequence comparison and generates a numerical result. The percent sequence identity may be calculated as the number of identical residues as a percentage of the total residues in the SEQ ID NO referred to.

[0349] “Fragments” are also variants and the term typically refers to a selected region of the polypeptide or polynucleotide that is of interest either functionally or, for example, in an assay. “Fragment” thus refers to an amino acid or nucleic acid sequence that is a portion of a full-length polypeptide or polynucleotide.

[0350] Such variants, derivatives, and fragments may be prepared using standard recombinant DNA techniques such as site-directed mutagenesis. Where insertions are to be made, synthetic DNA encoding the insertion together

with 5' and 3' flanking regions corresponding to the naturally-occurring sequence either side of the insertion site may be made. The flanking regions will contain convenient restriction sites corresponding to sites in the naturally-occurring sequence so that the sequence may be cut with the appropriate enzyme(s) and the synthetic DNA ligated into the cut. The DNA is then expressed in accordance with the invention to make the encoded polypeptide. These methods are only illustrative of the numerous standard techniques known in the art for manipulation of DNA sequences and other known techniques may also be used.

Method of Production

[0351] In one aspect, the present invention provides a method of producing a lentiviral vector according to the present invention.

[0352] Suitable methods to produce lentiviral vectors will be well known to those of skill in the art (see e.g. Merten, O. W., et al., 2016. *Molecular Therapy-Methods & Clinical Development*, 3, p. 16017).

[0353] The method of production may comprise: (a) introducing a transfer vector and one or more helper vectors into a host cell; (b) culturing the host cell under conditions suitable to produce lentiviral vectors according to the present invention; and (c) obtaining the lentiviral vectors from the host cell.

[0354] As used herein, a “transfer vector” may encode the lentiviral genome of the present invention. Suitably, the transfer vector used to produce the lentiviral viral genome within a host cell/packaging cell will have sufficient lentiviral genetic information to allow packaging of an RNA genome, in the presence of packaging components (e.g. gag-pol, rev, env), into a viral particle which is capable of infecting a target cell, but is incapable of independent replication to produce infectious viral particles within the final target cell.

[0355] The transfer vector used to produce the viral genome within a host cell/packaging cell may include transcriptional regulatory control sequences operably linked to the lentiviral genome to direct transcription of the genome in a host cell/packaging cell. These regulatory sequences may be the natural sequences associated with the transcribed viral sequence (i.e. the 5' U3 region), or they may be a heterologous promoter, such as another viral promoter (e.g. the CMV promoter). The transfer vector may be a plasmid.

[0356] As used herein, a “helper vector” may encode one or more packaging components (e.g. gag-pol, rev, env). The nucleotide sequence encoding the packaging component(s) may be operably linked to a promoter (e.g. a CMV promoter or a RSV promoter) and/or a polyadenylation signal. The term “helper vector” may include “packaging vectors” (e.g. encoding gag-pol or rev) and “envelope vectors” (e.g. encoding an env gene, such as VSV-g). The helper vectors, packaging vectors, and/or envelope vectors may be plasmids.

[0357] The transfer vector and one or more helper vectors may be introduced into the host cell by any suitable technique known in the art, such as transfection, transduction and/or transformation. Suitably, the helper vectors may be transiently transfected or transduced into the host cell or may be stably maintained (e.g. stably integrated into the cell genome or episomally maintained) within the host cell. Alternatively, a combination of transient transfection or transduction and stable maintenance may be used to introduce the helper vectors into the host cell.

[0358] Suitably, the transfer vector and/or the helper vectors may be plasmids and introduced by transfection. Suitably, a four plasmid system may be used consisting of a transfer plasmid and three helper plasmids. The three helper plasmids may consist of: a first helper plasmid encoding a gag-pol gene; a second helper plasmid encoding a rev gene; and a third helper plasmid encoding an env gene. Alternatively, a three plasmid system may be used which consists of a transfer plasmid, one helper plasmid encoding a gag-pol gene and a rev gene; and one helper plasmid encoding an env gene. Alternatively, a two plasmid system may be used in which all helper functions (e.g. gag-pol, rev, and env) are encoded by one helper plasmid.

[0359] Any suitable host cell may be used to produce the lentiviral vector. Suitable host cells include producer cells and packaging cells, such as those described below (e.g. HEK 293, or derivatives thereof). Suitable conditions for culturing the host cell will be well known to the skilled person. For example, the host cells may be incubated in chemically defined medium for from about 1 day to about 5 days (e.g. about 48 hours, about 72 hours, or about 96 hours).

[0360] The lentiviral vector may be obtained using in any suitable methods known in the art. For example the culture supernatant may be harvested and lentiviral vector subsequently purified from the culture supernatant (e.g. by centrifugation, membrane filtration and/or chromatography). The method of production may further comprise any other suitable process steps e.g. DNA reduction, concentration, formulation and/or sterilization.

Vectors, Kits and Systems

[0361] In one aspect, the present invention provides a vector encoding the lentiviral genome of the present invention. The vector may be a transfer vector, as described herein. For example, the vector may be a plasmid and/or the lentiviral genome may be operably linked to a promoter (e.g. a viral promoter, such as a CMV promoter).

[0362] In one aspect, the present invention provides a kit or system for producing the lentiviral vector of the present invention.

[0363] The kit or system may be a lentivirus packaging kit or system or a lentivirus production kit or system. As used herein, a “lentivirus packaging kit or system” may comprise one or more components, and optionally instructions, for packaging the lentiviral vector of the present invention. As used herein, a “lentivirus production kit or system” may comprise one or more components, and optionally instructions, for producing the lentiviral vector of the present invention.

[0364] The kit or system may comprise a transfer vector encoding the lentivirus genome of the present invention and optionally one or more helper vectors. The kit or system may further comprise host cells (e.g. packaging cells or producer cells) and/or other reagents (e.g. transfection reagent, culture medium, etc.). The kit or system may further comprise any other suitable components, and optionally instructions for packaging and/or producing the lentiviral vector of the present invention.

Cells

[0365] In one aspect, the present invention provides a cell comprising the lentiviral vector of the present invention. The cell may be an isolated cell. Suitably, the cell is a mammalian cell, for example a human cell. The cell may be an isolated human cell.

[0366] Suitably, the cell may be a producer cell. The term “producer cell” includes a cell that produces viral particles, after transient transfection, stable transfection or vector transduction of all the elements necessary to produce the viral particles or any cell engineered to stably comprise the elements necessary to produce the viral particles. Suitable producer cells will be well known to those of skill in the art and may include HEK293, COS-1, COS-7, CV-1, HeLa, CHO, and A549 cell lines. In some embodiments, the producer cell is a HEK293 cell, or a derivative thereof (e.g. a HEK293T cell, a HEK293T Lenti-X, a HEK293T-Rex cell, a HEK293FT cell, a HEK293SF-3F6 cell, a HEK293SF-3F9 cell, a HEK293-EBNA1 cell, or a SJ293TS cell).

[0367] Suitably, the cell may be a packaging cell. The term “packaging cell” includes a cell which contains some or all of the elements necessary for packaging a recombinant virus genome. Typically, such packaging cells contain one or more vectors which are capable of expressing viral structural proteins (e.g. gag-pol, rev, env) and/or one or more genes encoding the viral structural proteins have been integrated into the genome of the packaging cell. Cells comprising only some of the elements required for the production of enveloped viral particles are useful as intermediate reagents in the generation of viral particle producer cell lines, through subsequent steps of transient transfection, transduction or stable integration of each additional required element. These intermediate reagents are encompassed by the term “packaging cell”. Suitable packaging cells will be well known to those of skill in the art (see e.g. Merten, O. W., et al., 2016. *Molecular Therapy-Methods & Clinical Development*, 3, p. 16017).

[0368] Suitably, the cell may be a liver cell, for example a hepatocyte. Suitably, the cell may be an immortalized liver cell, for example an immortalized hepatocyte. Suitable cell lines will be well known to those of skill in the art, for example HepG2, Hep3B, HBG, and HepaRG cell lines. Methods to generate immortalized liver cells (e.g. immortalized hepatocytes) will be well known to those of skill in the art (see e.g. Ramboer, E., et al., 2015. *Methods Mol Biol*, 1250, pp. 53-76). Suitably, the cell may be a stem cell.

Pharmaceutical Compositions

[0369] In one aspect, the present invention provides pharmaceutical composition comprising the lentiviral vector or cell of the present invention. In preferred embodiments, the pharmaceutical composition comprises the lentiviral vector of the present invention in the form of a lentiviral particle.

[0370] A pharmaceutical composition is a composition that comprises or consists of a therapeutically effective amount of a pharmaceutically active agent e.g. the lentiviral vector. A pharmaceutical composition preferably includes a pharmaceutically acceptable carrier, diluent or excipient (including combinations thereof).

[0371] By “pharmaceutically acceptable” is included that the formulation is sterile and pyrogen free. The carrier, diluent, and/or excipient must be “acceptable” in the sense of being compatible with the lentiviral vector and not deleterious to the recipients thereof. Typically, the carriers, diluents, and excipients will be saline or infusion media which will be sterile and pyrogen free, however, other acceptable carriers, diluents, and excipients may be used.

[0372] Acceptable carriers, diluents, and excipients for therapeutic use are well known in the pharmaceutical art. The choice of pharmaceutical carrier, excipient or diluent can be selected with regard to the intended route of administration and standard pharmaceutical practice. The pharmaceutical compositions may comprise as (or in addition to) the carrier, excipient or diluent any suitable binder(s), lubricant(s), suspending agent(s), coating agent(s) or solubilising agent(s).

[0373] Examples of pharmaceutically acceptable carriers include, for example, water, salt solutions, alcohol, silicone, waxes, petroleum jelly, vegetable oils, polyethylene glycols, propylene glycol, liposomes, sugars, gelatin, lactose, amylose, magnesium stearate, talc, surfactants, silicic acid, viscous paraffin, perfume oil, fatty acid monoglycerides and diglycerides, petroethral fatty acid esters, hydroxymethyl-cellulose, polyvinylpyrrolidone,

and the like.

[0374] The lentiviral vector, cell, or pharmaceutical composition according to the present invention may be administered in a manner appropriate for treating and/or preventing the diseases described herein. Suitable administration routes will be known to the skilled person (see e.g. Fumoto, S., et al., 2013. Novel Gene Therapy Approaches, pp. 3-31). The quantity and frequency of administration may be determined by the skilled person, for example depending by such factors as the condition of the subject, and the type and severity of the subject's disease. The pharmaceutical composition may be formulated accordingly.

[0375] The lentiviral vector, cell or pharmaceutical composition according to the present invention may be administered parenterally, (e.g. intravenous, intra-arterial, intramuscular, intrathecal, subcutaneous), or by infusion techniques. The lentiviral vector, cell or pharmaceutical composition may be administered in the form of a sterile aqueous solution which may contain other substances, for example, enough salts or glucose to make the solution isotonic with blood. The aqueous solution may be suitably buffered (preferably to a pH of from 3 to 9). The pharmaceutical composition may be formulated accordingly. The preparation of suitable parenteral formulations under sterile conditions is readily accomplished by standard pharmaceutical techniques well-known to those skilled in the art.

[0376] The lentiviral vector, cell or pharmaceutical composition according to the present invention may be administered systemically, for example by intravenous injection or intraperitoneal injection. In some embodiments, the lentiviral vector, cell or pharmaceutical composition according to the present invention is administered by intravenous injection. The pharmaceutical composition may be formulated accordingly.

[0377] The lentiviral vector, cell or pharmaceutical composition according to the present invention may be administered locally, for example by direct injection, intra-arterial injection, or intraportal injection. In some embodiments, the lentiviral vector, cell or pharmaceutical composition according to the present invention is administered locally to the liver. In some embodiments, the lentiviral vector, cell or pharmaceutical composition according to the present invention is administered by intrahepatic injection, intrahepatic arterial injection, or intraportal injection. The pharmaceutical composition may be formulated accordingly.

[0378] The pharmaceutical compositions may comprise lentiviral vectors or cells of the invention in infusion media, for example sterile isotonic solution. The pharmaceutical composition may be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

[0379] The lentiviral vector, cell or pharmaceutical composition may be administered in a single or in multiple doses. Suitably, the lentiviral vector, cell or pharmaceutical composition may be administered in a single, one off dose. The pharmaceutical composition may be formulated accordingly.

[0380] The lentiviral vector, cell or pharmaceutical composition may be administered at varying doses (e.g. measured in Transducing Units (TU) per kg). The physician in any event may determine the actual dosage which will be most suitable for any individual subject and the dosage may, for example, vary with the age, weight and response of the particular subject. Suitably, the lentiviral vector of the present invention is administered at a dose of at least about 10⁸ TU/kg, at least about 10⁹ TU/kg, or at least about 10¹⁰ TU/kg. Suitably, the lentiviral vector of the present invention is administered at a dose of about 10¹³ TU/kg or less, about 10¹² TU/kg or less, or about 10¹¹ TU/kg or less. Suitably, the lentiviral vector of the present invention is administered in a dose of from about 10⁸ to about 10¹³ TU/kg, from about 10⁹ to about 10¹³ TU/kg, or from about 10¹⁰ to about 10¹³ TU/kg. Suitably, the lentiviral vector of the present invention is administered in a dose of from about 10⁸ to about 10¹² TU/kg, from about 10⁹ to about 10¹² TU/kg, or from about 10¹⁰ to about 10¹² TU/kg. Suitably, the lentiviral vector of the present invention is administered in a dose of from about 10⁸ to about 10¹¹ TU/kg, from about 10⁹ to about 10¹¹ TU/kg, or from about 10¹⁰ to about 10¹¹ TU/kg. In some embodiments, the lentiviral vector of the present invention is administered in a dose of from about 10⁸ to about 10¹¹ TU/kg, from about 10⁸ to about 10¹⁰ TU/kg, or from about 10⁹ to about 10¹⁰ TU/kg. In some embodiments, the lentiviral vector of the present invention is administered in a dose of from about 10⁹ to about 10¹⁰ TU/kg. The pharmaceutical composition may be formulated accordingly.

[0381] The pharmaceutical composition may further comprise one or more other therapeutic agents.

[0382] The invention further includes kits comprising the lentiviral vector, cell and/or pharmaceutical composition of the present invention. Preferably said kits are for use in the methods and used as described herein, e.g., the therapeutic methods as described herein. Preferably said kits comprise instructions for use of the kit components. Methods for Treating and/or Preventing Disease

[0383] In one aspect, the present invention provides the lentiviral vector, cell or pharmaceutical composition according to the present invention for use as a medicament.

[0384] In one aspect, the present invention provides use of the lentiviral vector, cell or pharmaceutical composition according to the present invention in the manufacture of a medicament.

[0385] In one aspect, the present invention provides a method of administering a therapeutically effective amount of the lentiviral vector, cell or pharmaceutical composition according to the present invention to a subject in need thereof.

[0386] The lentiviral vector mediated gene therapy described herein may allow for a stable gene transfer even in paediatric patients at the first disease stages by virtue of lentiviral vector genomic integration.

[0387] Following administration of the lentiviral vector of the present invention, the lentiviral vector may integrate into the genome of liver cells (e.g. hepatocyte cells). Subsequently, the lentiviral vector may be maintained in the genome of liver cells (e.g. hepatocyte cells) as they duplicate. The integration of the lentiviral vector in the genome of liver cells may be determined by integration site (IS) analysis (e.g. quantitative high-throughput vector IS analysis). Suitable methods are known in the art (see e.g. Cantore, A., et al., 2015. *Science translational medicine*, 7(277), p. 277ra28).

[0388] The lentiviral vector, cell or pharmaceutical composition may be administered to any subject in need thereof. The subject may be a mammal (e.g. a human). In some embodiments, the subject is male. In some embodiments, the subject is female. In preferred embodiments, the lentiviral vector is administered in the form of a lentiviral particle.

[0389] In some embodiments, the subject is a juvenile, an adolescent, or a child. The term “juvenile” may refer to an individual that has not yet reached adulthood. The term “adolescent” may refer to an individual during the period from the onset of puberty to adulthood. The term “child” may refer an individual between the stages of birth and puberty.

[0390] In some embodiments, the subject is a young child, a toddler, or an infant. The term “young child” may refer to a human subject aged from 3 years to 5 years. The term “toddler” may refer to a human subject aged from 1 year to 3 years. The term “infant” may mean refer to a human subject under the age of 12 months.

[0391] In some embodiments, the subject is a paediatric patient. The term “paediatric patient” may refer to a human subject until about 18-21 years of age (see e.g. Sawyer, S. M., et al., 2019. *The Lancet Child & Adolescent Health*, 3(11), pp. 822-830).

[0392] In some embodiments, the subject is a neonatal patient or an infantile patient. The term “neonatal patient” may refer to a human subject who is aged about 4 weeks old or younger. The term “infantile patient” may refer to a human subject who is aged from about 4 weeks to about 1 year.

[0393] In other embodiments, the subject is an adult. Human liver is expected to completely renew every 5 years in humans, so integrating vectors are expected to be more persisting compared to mostly episomal vectors (e.g. AAV).

Progressive Familial Intrahepatic Cholestasis (PFIC)

[0394] The vector, cell or pharmaceutical composition according to the present invention may be used to prevent and/or treat progressive familial intrahepatic cholestasis (PFIC).

[0395] In one aspect, the present invention provides the lentiviral vector, cell or pharmaceutical composition according to the present invention for use in preventing or treating progressive familial intrahepatic cholestasis (PFIC).

[0396] In one aspect, the present invention provides use of the lentiviral vector, cell or pharmaceutical composition according to the present invention in the manufacture of a medicament for preventing or treating progressive familial intrahepatic cholestasis (PFIC).

[0397] In one aspect, the present invention provides a method of preventing or treating progressive familial intrahepatic cholestasis (PFIC), the method comprising administering a therapeutically effective amount of the lentiviral vector, cell or pharmaceutical composition according to the present invention to a subject in need thereof.

[0398] As described above, PFIC is a group of rare autosomal recessive liver disorders characterized by mutations in genes encoding proteins involved in the hepatocellular transport system (e.g. bile secretion from hepatocyte to canaliculi). PFIC is characterised by an early onset of cholestasis (usually during infancy) with pruritus and malabsorption, which rapidly progresses.

[0399] Progressive liver fibrosis, cirrhosis, and end stage liver disease may eventually develop (see e.g. Davit-Spraul, A., et al., 2009. Progressive familial intrahepatic cholestasis. *Orphanet journal of rare diseases*, 4(1), pp. 1-12; Jacquemin, E., 2012. *Clinics and research in hepatology and gastroenterology*, 36, pp. S26-S35; and Gunaydin, M. and Cil, A. T. B., 2018. *Hepatic Medicine: Evidence and Research*, 10, pp. 95-104).

[0400] PFICs may be sub-grouped according to the genetic defect, clinical presentation, laboratory findings, and liver histology. PFIC-1, PFIC-2, and PFIC-3 are the main subtypes of PFIC. Other subtypes may include PFIC-4, PFIC-5, and PFIC-6 (see e.g. Bull, L. N. and Thompson, R. J., 2018. *Clinics in liver disease*, 22(4), pp. 657-669). Suitably, PFIC refers to PFIC-1, PFIC-2, PFIC-3, PFIC-4, PFIC-5, and PFIC-6. Suitably, PFIC refers to PFIC-1, PFIC-2, PFIC-3.

[0401] In some embodiments, the PFIC is selected from PFIC-2, PFIC-5, PFIC-6. In preferred embodiments, the PFIC is PFIC-2.

[0402] Following administration of the lentiviral vector of the present invention to a subject in need thereof, serum bile acid levels may be prevented from increasing or reduced. Suitably, serum bile acid levels may be reduced by at least 10%, at least 20%, at least 30%, at least 40%, or at least 50%. Suitably, following administration, the subject's serum bile acid level may be from about 20 to about 30 $\mu\text{mol/L}$, or from about 10 to about 20 $\mu\text{mol/L}$. Following administration of the lentiviral vector of the present invention to a subject in need thereof, serum bile acid levels may be normalised. A normal serum bile acid level may be from about 4.5 to about 19.2 $\mu\text{mol/L}$ (see e.g. Tibesar, E., et al., 2014. Case Reports in Pediatrics, 2014, 185923). Serum bile acid levels may be determined by any method known in the art (see e.g. Barnes, S., et al., 1975. Journal of clinical pathology, 28(6), pp. 506-509).

[0403] Following administration of the lentiviral vector of the present invention to a subject in need thereof the liver function of the subject may be improved. Liver function may be determined by any suitable method known in the art, for example by measuring biochemical markers such as serum bilirubin, alanine amino transferase, aspartate amino transferase, ratio of aminotransferases, alkaline phosphatase, gamma glutamyl transferase, 5' nucleotidase, ceruloplasmin, and α -fetoprotein (see e.g. Gowda, S., et al., 2009. The Pan African Medical Journal, 3, p. 17; Hoekstra, L. T., et al., 2013. Annals of surgery, 257(1), pp. 27-36; and Limdi, J. K. and Hyde, G. M., 2003. Postgraduate medical journal, 79(932), pp. 307-312).

[0404] Following administration of the lentiviral vector of the present invention to a subject in need thereof, serum bilirubin levels (e.g. total and/or direct) may be prevented from increasing or reduced. Suitably, serum bilirubin levels may be reduced by at least 10%, at least 20%, at least 30%, at least 40%, or at least 50%. Suitably, following administration serum bilirubin levels may be less than about 1 mg/dL. Following administration of the lentiviral vector of the present invention to a subject in need thereof, serum bilirubin levels may be normalised. A normal bilirubin level may be from about 2 to about 17 $\mu\text{mol/L}$ (see e.g. Limdi, J. K. and Hyde, G. M., 2003. Postgraduate medical journal, 79(932), pp. 307-312). Serum bilirubin levels may be determined by any method known in the art (see e.g. Lathe, G. H. and Ruthven, C. R. J., 1958. Journal of clinical pathology, 11(2), p. 155).

[0405] Following administration of the lentiviral vector of the present invention to a subject in need thereof, the serum level of one or more liver enzymes may be prevented from increasing or reduced. Suitable liver enzymes include alanine aminotransferase (ALT), aspartate aminotransferase (AST), and alkaline phosphatase (ALP). Suitably, one or more liver enzymes may be reduced by at least 10%, at least 20%, at least 30%, at least 40%, or at least 50%. Suitably, following administration ALP levels may be less than about 150 IU/L. Following administration of the lentiviral vector of the present invention to a subject in need thereof, the serum level of one or more liver enzymes may be normalised. A normal ALT level may be from about 0 to about 45 IU/L, a normal AST level may be from about 0 to about 35 IU/L, and a normal ALP level may be from about 30 to about 120 IU/L (see e.g. Limdi, J. K. and Hyde, G. M., 2003. Postgraduate medical journal, 79(932), pp. 307-312). The serum level of liver enzymes may be determined by any method known in the art.

[0406] Following administration of the lentiviral vector of the present invention to a subject in need thereof, the formation of liver fibrosis may be slowed and/or reduced. Suitably, the extent of fibrosis is reduced by at least 10%, at least 20%, at least 30%, at least 40%, or at least 50%. The extent of liver fibrosis may be determined by any method known in the art (see e.g. Standish, R. A., et al., 2006. Gut, 55(4), pp. 569-578).

PFIC-1

[0407] Progressive familial intrahepatic cholestasis type 1 (PFIC-1, MCID: CHL132, OMIM: 211600) is also called Fic1 deficiency or Byler disease and is caused by a variety of mutations in ATP8B1, the gene coding for FIC-1.

[0408] In some embodiments, the PFIC is PFIC-1 and the PFIC-associated polypeptide is FIC1 or fragment and/or variant thereof.

[0409] In one aspect, the present invention provides the lentiviral vector, cell or pharmaceutical composition according to the present invention for use in preventing or treating PFIC-1.

[0410] In one aspect, the present invention provides use of the lentiviral vector, cell or pharmaceutical composition according to the present invention in the manufacture of a medicament for preventing or treating PFIC-1.

[0411] In one aspect, the present invention provides a method of preventing or treating PFIC-1, the method comprising administering a therapeutically effective amount of the lentiviral vector, cell or pharmaceutical composition according to the present invention to a subject in need thereof.

PFIC-2

[0412] Progressive familial intrahepatic cholestasis type X (PFIC-2, MCID: CHL136, OMIM: 601847) is also called Bsep Deficiency or Severe Abcb11 Deficiency and is caused by a variety of mutations in ABCB11, the

gene coding for BSEP.

[0413] In some embodiments, the PFIC is PFIC-2 and the PFIC-associated polypeptide is BSEP or fragment and/or variant thereof.

[0414] In one aspect, the present invention provides the lentiviral vector, cell or pharmaceutical composition according to the present invention for use in preventing or treating PFIC-2.

[0415] In one aspect, the present invention provides use of the lentiviral vector, cell or pharmaceutical composition according to the present invention in the manufacture of a medicament for preventing or treating PFIC-2.

[0416] In one aspect, the present invention provides a method of preventing or treating PFIC-2, the method comprising administering a therapeutically effective amount of the lentiviral vector, cell or pharmaceutical composition according to the present invention to a subject in need thereof.

PFIC-3

[0417] Progressive familial intrahepatic cholestasis type 3 (PFIC-3, MCID: CHL137, OMIM: 602347) is also called Mdr3 Deficiency and is caused by a variety of mutations in ABCB4, the gene coding for MDR3.

[0418] In some embodiments, the PFIC is PFIC-3 and the PFIC-associated polypeptide is MDR3 or fragment and/or variant thereof.

[0419] In one aspect, the present invention provides the lentiviral vector, cell or pharmaceutical composition according to the present invention for use in preventing or treating PFIC-3.

[0420] In one aspect, the present invention provides use of the lentiviral vector, cell or pharmaceutical composition according to the present invention in the manufacture of a medicament for preventing or treating PFIC-3.

[0421] In one aspect, the present invention provides a method of preventing or treating PFIC-3, the method comprising administering a therapeutically effective amount of the lentiviral vector, cell or pharmaceutical composition according to the present invention to a subject in need thereof.

PFIC-4

[0422] Progressive familial intrahepatic cholestasis type 4 (PFIC-4, MCID: CHL143, OMIM: 615878) is also called Tjp2 Deficit and is caused by a variety of mutations in TJP2, the gene coding for TJP2 (see e.g. Sambrotta, M., et al., 2014. Nature genetics, 46(4), pp. 326-328).

[0423] In some embodiments, the PFIC is PFIC-4 and the PFIC-associated polypeptide is TJP2 or fragment and/or variant thereof.

[0424] In one aspect, the present invention provides the lentiviral vector, cell or pharmaceutical composition according to the present invention for use in preventing or treating PFIC-4.

[0425] In one aspect, the present invention provides use of the lentiviral vector, cell or pharmaceutical composition according to the present invention in the manufacture of a medicament for preventing or treating PFIC-4.

[0426] In one aspect, the present invention provides a method of preventing or treating PFIC-4, the method comprising administering a therapeutically effective amount of the lentiviral vector, cell or pharmaceutical composition according to the present invention to a subject in need thereof.

PFIC-5

[0427] Progressive familial intrahepatic cholestasis type X (PFIC-5, MCID: CHL091, OMIM: 617049) is also called Nr1h4 Deficiency and is caused by a variety of mutations in NR1H4, the gene coding for FXR (see e.g. Gomez-Ospina, N., et al., 2016. Nature communications, 7(1), pp. 1-8).

[0428] In some embodiments, the PFIC is PFIC-5 and the PFIC-associated polypeptide is FXR or fragment and/or variant thereof.

[0429] In one aspect, the present invention provides the lentiviral vector, cell or pharmaceutical composition according to the present invention for use in preventing or treating PFIC-5.

[0430] In one aspect, the present invention provides use of the lentiviral vector, cell or pharmaceutical composition according to the present invention in the manufacture of a medicament for preventing or treating PFIC-5.

[0431] In one aspect, the present invention provides a method of preventing or treating PFIC-5, the method comprising administering a therapeutically effective amount of the lentiviral vector, cell or pharmaceutical composition according to the present invention to a subject in need thereof.

PFIC-6

[0432] Progressive familial intrahepatic cholestasis type 6 (PFIC-6, MCID: CHL 186) is caused by a variety of mutations in MYO5B, the gene coding for MYO5B (see e.g. Gonzales, E., et al., 2017. Hepatology, 65(1), pp. 164-173).

[0433] In some embodiments, the PFIC is PFIC-6 and the PFIC-associated polypeptide is MYO5B or fragment

and/or variant thereof.

[0434] In one aspect, the present invention provides the lentiviral vector, cell or pharmaceutical composition according to the present invention for use in preventing or treating PFIC-6.

[0435] In one aspect, the present invention provides use of the lentiviral vector, cell or pharmaceutical composition according to the present invention in the manufacture of a medicament for preventing or treating PFIC-6.

[0436] In one aspect, the present invention provides a method of preventing or treating PFIC-6, the method comprising administering a therapeutically effective amount of the lentiviral vector, cell or pharmaceutical composition according to the present invention to a subject in need thereof.

EXAMPLES

[0437] Preferred features and embodiments of the invention will now be described by way of non-limiting examples.

Example 1—Lentiviral Vector (LV) Encoding a BSEP Transgene

[0438] We generated a lentiviral vector (LV) encoding a BSEP transgene and evaluated in vitro in hepatocyte cell lines and in vivo in a mouse model of PFIC-2, the *Abcb11*^{-/-} mice.

Results

[0439] We generated and produced LV encoding a human wild type or codon optimized BSEP transgenes under the control of a hepatocyte-specific cassette based on the enhanced transthyretin promoter (ET.BSEP) or the human alpha-1 antitrypsin promoter (hAAT.BSEP) (FIG. 1a).

[0440] To compare between the different BSEP transgenes, we first transduced Huh7 cells with LV-ET.BSEPwt, LV-ET.BSEPco_1 or LV-ET.BSEPco_2 at increasing multiplicity of infection (MOI). We performed flow cytometry analysis 10 days after transduction, detecting BSEP overexpression in LV-ET.BSEP transduced cells. Mean fluorescence intensity increased proportionally to the MOI employed and was higher for the BSEPco_1 and BSEPco_2 transgenes compared to BSEPwt at similar vector copy number (VCN) per cell (FIG. 1b,c).

[0441] We then compared the activities of the ET and the hAAT promoters by transducing Huh7 cells with LV-ET.BSEPco_1 or LV-hAAT.BSEP_co1. VCN/cell was comparable in cells transduced with the two vectors at different MOI (FIG. 1d), while BSEP protein expression was higher in LV-ET.BSEPco_1 compared to LV-hAAT.BSEPco_1 transduced cells as detected by western blot analysis (FIG. 1e). By BSEP protein quantification and normalization on the GAPDH housekeeper, we confirmed a higher transgene expression (3-3.5 fold) obtained from the expression cassette driven by the ET promoter compared to the hAAT (FIG. 1f).

[0442] Immunofluorescence analysis on Huh7 cells transduced with LV-ET.BSEPco_1 confirmed proper BSEP membrane and cytoplasmic localization while no signal was detectable in Huh7 untransduced cells (FIG. 2).

[0443] We administered LV-ET.BSEPco_1 to juvenile *Abcb11*^{-/-} mice at a dose of 3×10^{sup.10} transducing units (TU)/kg by i.v. injection. The treatment resulted in a significant decrease of serum BA, bilirubin and ALP biomarkers until 7 months of age, indicating prevention of the progressive cholestatic damage (FIG. 3a-c).

Materials and Methods

Plasmid Construction

[0444] The BSEP coding sequences (WT or codon optimized) used in the study were synthesized by GeneScript and cloned into a third-generation self-inactivating (SIN) LV transfer plasmid (Milani, M., et al., 2019. Science Translational Medicine, 11(493), p.eaav7325) under the control of the enhanced transthyretin promoter (ET) or the human alpha-1 antitrypsin promoter (hAAT).

Vector Production

[0445] Lab-grade VSV.G-pseudotyped third-generation SIN LV were produced by calcium phosphate transient transfection into 293T cells. 293T cells were transfected with a solution containing a mix of the selected LV genome transfer plasmid, the packaging plasmids pMDLg/pRRE and pCMV.REV, pMD2.G and pAdvantage, as previously described (Milani, M., et al., 2017. EMBO molecular medicine, 9(11), pp. 1558-1573; and Milani, M., et al., 2019. Science Translational Medicine, 11(493), p.eaav7325).

[0446] Medium was replaced 14-16 hours post transfection and supernatant was collected around 30 hours after medium change. LV-containing supernatants were sterilized through a 0.22 µm filter (Millipore) and transferred into sterile poliallomer tubes (Beckman) and centrifuged at 20,000 g for 120 min at 20° C. (Beckman Optima XL-100K Ultracentrifuge). LV pellet was resuspended in the appropriate volume of PBS to allow 500-1000× concentration.

LV Titration

[0447] For LV titration, 1×10^{sup.5} 293T cells were transduced with serial LV dilutions in the presence of polybrene (8 µg/ml). Genomic DNA (gDNA) was extracted 10 days after transduction, using Maxwell 16 Cell DNA Purification Kit (Promega), following manufacturer's instructions. VCN was determined by digital droplet PCR (ddPCR) starting from 5-20 ng of template gDNA using primers (HIV fw: 5'-T ACTGACGCTCTCGCACC-

3'; HIV rv: 5'-TCTGCACGCAAGACTCG-3') and a probe (FAM 5'-ATCTCTCTCCTTCTAGCCTC-3') designed on the primer binding site region of LV. The amount of endogenous DNA was quantified by a primers/probe set designed on the human GAPDH gene (Applied Biosystems HS00483111_cm). The PCR reaction was performed with each primer (900 nM) and the probe (250 nM) following manufacturer's instructions (Biorad), read with QX200 reader and analyzed with QuantaSoft software (Biorad). Infectious titer, expressed as TU/mL, was calculated using the formula $TU/mL = (VCN \times 100,000 \times (1/\text{dilution factor}))$. LV physical particles were measured by HIV-1 Gag p24 antigen immunocapture assay (Perkin Elmer) following manufacturer's instructions. LV specific infectivity was calculated as the ratio between infectious titer and physical particles.

Cell Culture and in Vitro Transduction Experiments

[0448] HuH7 cells were maintained under 37° C., 5% CO.sub.2 condition in Iscove's Modified Dulbecco's Medium (IMDM) supplemented with 10% FBS (Thermo Fisher Scientific, Waltham, MA). Cells were seeded into 12-well plates (2×10^{sup.5} cells/well) and transduced with LV variants at different MOI in the presence of polybrene (8 µg/ml). Ten days post-LV transduction, cells were harvested for molecular analyses.

Mice Experiments

[0449] All animal experiments were performed in strict accordance with good animal practices following Italian and European legislation on animal care and experimentation (2010/63/EU). Abcb11^{-/-} mice (Jackson strain n.004125) 2-week old were used in these studies. Animals were administered with LV at a dose of 3×10^{sup.10} TU/kg via the retro-orbital plexus. Blood samples were collected monthly throughout the duration of the study for analysis of the metabolites.

VCN Determination

[0450] DNA was extracted from cells or liver samples using Maxwell 16 Cell or Tissue DNA Purification Kits (Promega). VCN was determined in Huh7 samples as described above (see "LV titration"). VCN in murine DNA was determined by ddPCR, starting from 5-20 ng of template gDNA using a primers/probe set designed on the primer binding site region of LV (see "LV titration" above). The amount of endogenous murine DNA was quantified by a primers/probe set designed on the murine sema3a gene (Sema3A fw: 5'-ACCGATTCCAGATGATTGGC-3'; Sema3A rv: 5'-TCCATATTAATGCAGTGCTTGC-3'; Sema3A probe: HEX 5'-AGAGGCCTGTCCTGCAGCTC

[0451] ATGG-3' BHQ1). The PCR reaction was performed with each primer (900 nM) and the probe (250 nM) following manufacturer's instructions (Biorad), read with QX200 reader and analyzed with QuantaSoft software (Biorad).

Western Blot Analysis

[0452] Protein extracts from HuH7 cell or from mouse liver samples were prepared using RIPA buffer (EMD Millipore) and protease inhibitors (Roche). Protein concentration was determined using the DC Protein Assay (Bio Rad Laboratories). SDS-page electrophoresis was performed in a Bis-Tris 4-12% gradient polyacrylamide gel. After transfer, the membrane was blocked with 1×TBS+5% milk and incubated with an anti-Abcb11 antibody (Sigma-Aldrich HPA019035) or anti-GAPDH (Biolegend 649201). The membrane was incubated with an anti-Mouse and an anti-Rabbit IgG secondary antibodies (Jackson ImmunoResearch 115-035-003 and 111-035-144), then with clarity Western ECL substrate (Bio Rad) and visualized by Uvitec Imaging System (Clever Scientific). For Western blot quantification, we used the Image J software.

BSEP Immunofluorescence Analysis

[0453] Huh7 LV-transduced cells were seeded in an 8-well slide (Nunc® Lab-Tek® Chamber Slide™ system) to reach 40-50% confluency. 24 hours later, cells were fixed in 4% PFA and permeabilized in PBS+Triton X-100 0.1% for 10 minutes at room temperature (RT). Cells were blocked in in PBST+1% BSA 1 hour at RT and stained with a rabbit anti-BSEP (ab255605, dil. 1:100) primary antibody, overnight at 4° C. After an incubation with a goat anti rabbit IgG 488 secondary antibody (Invitrogen, dil. 1:1000), the nuclear staining was performed with Hoechst (diluted at 2 µg/mL in PBS), 5 minutes at RT. The staining slide was acquired at Leica TCS SP8 SMD confocal microscope.

Bile Acids, ALP, Bilirubin Determination

[0454] Total bile acids were determined in serum samples by using the Mouse Total Bile Acid Assay Kit (Crystal Chem, 80471) according to manufacturer's instructions. Serum ALP activity was measured with an International Federation of Clinical Chemistry and Laboratory Medicine optimized kinetic UV method in a SABA chemical analyzer (Seac-Radim), and it is expressed as units per liter (U/L).

EMBODIMENTS

[0455] Various preferred features and embodiments of the present invention will now be described with reference to the following numbered paragraphs (paras).

[0456] 1. A lentiviral vector comprising a nucleotide sequence encoding a progressive familial intrahepatic cholestasis (PFIC)-associated polypeptide, optionally wherein the lentiviral vector is an immune-shielded

lentiviral vector.

[0457] 2. The lentiviral vector according to para 1, wherein the PFIC-associated polypeptide is selected from familial intrahepatic cholestasis type 1 (FIC1), or a fragment thereof; bile salt export pump (BSEP), or a fragment thereof; multiple drug resistance 3 (MDR3), or a fragment thereof; tight junction protein 2 (TJP2), or a fragment thereof; farnesoid X receptor (FXR), or a fragment thereof; and Myosin-Vb (MYO5B), or a fragment thereof.

[0458] 3. The lentiviral vector according to para 1 or 2, wherein the PFIC-associated polypeptide is selected from BSEP, or a fragment thereof; FXR, or a fragment thereof; and MYO5B, or a fragment thereof; preferably wherein the PFIC-associated polypeptide is BSEP, or a fragment thereof.

[0459] 4. The lentiviral vector according to any preceding para, wherein the PFIC-associated polypeptide comprises or consists of an amino acid sequence which is at least 70% identical to one of SEQ ID NOs: 37, 39, 43, 45, 47, 48 or 50, or a fragment thereof, preferably wherein the PFIC-associated polypeptide comprises or consists of an amino acid sequence which is at least 70% identical to SEQ ID NO: 39 or a fragment thereof.

[0460] 5. The lentiviral vector according to any preceding para, wherein the nucleotide sequence encoding a PFIC-associated polypeptide comprises or consists of a nucleotide sequence which is at least 70% identical to one of SEQ ID NOs: 38, 40, 41, 42, 44, 46, 49 or 51, or a fragment thereof, preferably wherein the nucleotide sequence encoding a PFIC-associated polypeptide comprises or consists of a nucleotide sequence which is at least 70% identical to one of SEQ ID NOs: 40-42, or a fragment thereof.

[0461] 6. The lentiviral vector according to any preceding para, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is codon-optimised, preferably wherein the nucleotide sequence encoding a PFIC-associated polypeptide comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 41 or 42, or a fragment thereof.

[0462] 7. The lentiviral vector according to any preceding para, wherein the lentiviral vector is a CD47.sup.high lentiviral vector.

[0463] 8. The lentiviral vector according to any preceding para, wherein the lentiviral vector is obtained from a CD47.sup.high host cell, optionally wherein the host cell is genetically engineered to increase expression of CD47 on the cell surface.

[0464] 9. The lentiviral vector according to any preceding para, wherein the lentiviral vector has at least about 2-fold more CD47 on its surface than a lentiviral vector obtained from an unmodified host cell.

[0465] 10. The lentiviral vector according to any preceding para, wherein the lentiviral vector is a MHC-I.sup.free lentiviral vector.

[0466] 11. The lentiviral vector according to any preceding para, wherein the lentiviral vector is obtained from a MHC-I.sup.free host cell, optionally wherein the host cell is genetically engineered to disrupt expression of MHC-I on the cell surface.

[0467] 12. The lentiviral vector according to any preceding para, wherein MHC-I is not detectable on the surface of the lentiviral vector.

[0468] 13. The lentiviral vector according to any preceding para, wherein the lentiviral vector is a CD47.sup.high/MHC-I.sup.free lentiviral vector.

[0469] 14. The lentiviral vector according to any preceding para, wherein the lentiviral vector is obtained from a CD47.sup.high/MHC-I.sup.free host cell.

[0470] 15. The lentiviral vector according to any preceding para, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to one or more miRNA target sequences.

[0471] 16. The lentiviral vector according to para 15, wherein the one or more miRNA target sequences suppress transgene expression in one or more cells other than hepatocytes, preferably wherein the one or more miRNA target sequence suppress transgene expression in hematopoietic-lineage cells and/or antigen-presenting cells.

[0472] 17. The lentiviral vector according to para 15 or 16, wherein the one or more miRNA target sequences are selected from miR-181, miR-142, miR-223, and miR-155 target sequences.

[0473] 18. The lentiviral vector according to any preceding para, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to one or more mir-142 target sequence, two or more mir-142 target sequences, three or more mir-142 target sequences, or four or more mir-142 target sequences.

[0474] 19. The lentiviral vector according to any preceding para, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to four mir-142 target sequences.

[0475] 20. The lentiviral vector according to any of paras 15-19, wherein the one or more miRNA target sequences comprise or consist of a nucleotide sequence which is at least 90% identical to SEQ ID NO: 17, preferably wherein the one or more miRNA target sequences comprise or consist of a nucleotide sequence which is at least 90% identical to SEQ ID NO: 18.

[0476] 21. The lentiviral vector according to any preceding para, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a liver-specific promoter, preferably wherein the nucleotide

sequence encoding a PFIC-associated polypeptide is operably linked to a hepatocyte-specific promoter.

[0477] 22. The lentiviral vector according to any preceding para, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a transthyretin (TTR) promoter, an alpha-1-antitrypsin (AAT) promoter, a thyroxine-binding globulin (TBG) promoter, a APoE/hAAT promoter, a HCR-hAAT promoter, a LP1 promoter, or a HLP promoter.

[0478] 23. The lentiviral vector according to any preceding para, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a transthyretin (TTR) promoter, preferably wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a Enh1mTTR (ET) promoter.

[0479] 24. The lentiviral vector according to any preceding para, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a promoter which comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 19.

[0480] 25. The lentiviral vector according to any preceding para, wherein the lentiviral vector is pseudotyped, preferably wherein the lentiviral vector is VSV.G-pseudotyped.

[0481] 26. The lentiviral vector according to any preceding para, wherein the lentiviral vector is a self-inactivating (SIN) lentiviral vector, preferably wherein the lentiviral vector comprises self-inactivating (SIN) LTRs which comprise or consist of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 23, or a fragment thereof.

[0482] 27. The lentiviral vector according to any preceding para, wherein the lentiviral vector is an integrating lentiviral vector and/or a replication-defective lentiviral vector.

[0483] 28. The lentiviral vector according to any preceding para, wherein the lentiviral vector comprises a nucleotide sequence which is at least 70% identical to SEQ ID NO: 36.

[0484] 29. An isolated cell comprising a lentiviral vector according to any of paras 1-28.

[0485] 30. A pharmaceutical composition comprising a lentiviral vector according to any of paras 1-28, in combination with a pharmaceutically acceptable carrier, diluent or excipient.

[0486] 31. The lentiviral vector according to any of paras 1-28 or pharmaceutical composition according to para 30, for use as a medicament.

[0487] 32. Use of a lentiviral vector according to any of paras 1-28 or a pharmaceutical composition according to para 30, for the manufacture of a medicament.

[0488] 33. A method comprising administering a therapeutically effective amount of a lentiviral vector according to any of paras 1-28 or a pharmaceutical composition according to para 30, to a subject in need thereof.

[0489] 34. The lentiviral vector according to any of paras 1-28 or pharmaceutical composition according to para 30, for use in preventing or treating progressive familial intrahepatic cholestasis (PFIC).

[0490] 35. Use of a lentiviral vector according to any of paras 1-28 or a pharmaceutical composition according to para 30, for the manufacture of a medicament for preventing or treating progressive familial intrahepatic cholestasis (PFIC).

[0491] 36. A method of preventing or treating progressive familial intrahepatic cholestasis (PFIC), comprising administering a therapeutically effective amount of a lentiviral vector according to any of paras 1-28 or a pharmaceutical composition according to para 30, to a subject in need thereof.

[0492] 37. The lentiviral vector or pharmaceutical composition for use according to para 34, the use according to para 35, or the method according to para 36, wherein the PFIC-associated polypeptide is FIC1, or a fragment thereof, and the PFIC is PFIC type 1 (PFIC-1); the PFIC-associated polypeptide is BSEP, or a fragment thereof, and the PFIC is PFIC type 2 (PFIC-2); the PFIC-associated polypeptide is MDR3, or a fragment thereof, and the PFIC is PFIC type 3 (PFIC-3); the PFIC-associated polypeptide is TJP2, or a fragment thereof, and the PFIC is PFIC type 4 (PFIC-4); the PFIC-associated polypeptide is FXR, or a fragment thereof, and the PFIC is PFIC type 5 (PFIC-5); or the PFIC-associated polypeptide is MYO5B, or a fragment thereof, and the PFIC is PFIC type 6 (PFIC-6).

[0493] 38. The lentiviral vector or pharmaceutical composition for use according to para 34 or 37, the use according to para 35 or 37, or the method according to para 36 or 37, wherein the PFIC-associated polypeptide is BSEP, or a fragment thereof, and the PFIC is PFIC-2; the PFIC-associated polypeptide is FXR, or a fragment thereof, and the PFIC is PFIC-5; or the PFIC-associated polypeptide is MYO5B, or a fragment thereof, and the PFIC is PFIC-6, preferably wherein the PFIC-associated polypeptide is BSEP, or a fragment thereof, and the PFIC is PFIC-2.

[0494] 39. The lentiviral vector or pharmaceutical composition for use according to any of paras para 31, 34, 37-38, the use according to any of paras 32, 35, 37-38, or the method according to any of paras 33, 36-39, wherein the subject is a human subject.

[0495] 40. The lentiviral vector or pharmaceutical composition for use according to any of paras para 31, 34, 37-39, the use according to any of paras 32, 35, 37-39, or the method according to any of paras 33, 36-39, wherein

the subject is a juvenile.

[0496] 41. The lentiviral vector or pharmaceutical composition for use according to any of paras para 31, 34, 37-39, the use according to any of paras 32, 35, 37-39, or the method according to any of paras 33, 36-39, wherein the subject is a paediatric patient, preferably wherein the subject is a neonatal patient or an infantile patient.

[0497] 42. The lentiviral vector or pharmaceutical composition for use according to any of paras 31, 34, 37-41, the use according to any of paras 32, 35, 37-41, or the method according to any of paras 33, 36-41, wherein said lentiviral vector or said pharmaceutical composition is administered systemically, preferably wherein said lentiviral vector or said pharmaceutical composition is administered by intravenous injection or intraperitoneal injection.

[0498] 43. The lentiviral vector or pharmaceutical composition for use according to any of paras 31, 34, 37-41, the use according to any of paras 32, 35, 37-41, or the method according to any of paras 33, 36-41, wherein said lentiviral vector or said pharmaceutical composition is administered locally, preferably wherein said lentiviral vector or said pharmaceutical composition is administered by direct injection, intra-arterial injection, or intraportal injection.

[0499] 44. The lentiviral vector or pharmaceutical composition for use according to para 43, the use according to para 43, or the method according to para 43, wherein said lentiviral vector or said pharmaceutical composition is administered locally to the liver, preferably wherein said lentiviral vector or said pharmaceutical composition is administered by intrahepatic injection, intrahepatic arterial injection, or intraportal injection.

[0500] 45. The lentiviral vector or pharmaceutical composition for use according to any of paras 31, 34, 37-44, the use according to any of paras 32, 35, 37-43, or the method according to any of paras 33, 36-44, wherein the lentiviral vector is administered at a dose of at least about 10.sup.8 TU/kg, at least about 10.sup.9 TU/kg, or at least about 10.sup.10 TU/kg.

[0501] 46. The lentiviral vector or pharmaceutical composition for use according to any of paras 31, 34, 37-45, the use according to any of paras 32, 35, 37-45, or the method according to any of paras 33, 36-45, wherein the lentiviral vector is administered in a dose of from about 10.sup.8 to about 10.sup.11 TU/kg, from about 10.sup.8 to about 10.sup.10 TU/kg, or from about 10.sup.9 to about 10.sup.10 TU/kg.

[0502] 47. The lentiviral vector or pharmaceutical composition for use according to any of paras 31, 34, 37-46, the use according to any of paras 32, 35, 37-46, or the method according to any of paras 33, 36-46, wherein the lentiviral vector integrates into the genome of liver cells and is maintained as the liver cells duplicate, preferably wherein the lentiviral vector integrates into the genome of hepatocytes and is maintained as the hepatocytes duplicate.

[0503] 48. The lentiviral vector or pharmaceutical composition for use according to any of paras 34, 37-47, the use according to any of paras 35, 37-47, or the method according to any of paras 36-47, wherein serum bile acid levels are reduced and/or normalised.

[0504] 49. The lentiviral vector or pharmaceutical composition for use according to any of paras 34, 37-48, the use according to any of paras 35, 37-48, or the method according to any of paras 36-48, wherein the liver function is improved, preferably wherein the serum level of bilirubin and/or one or more liver enzymes is reduced and/or normalised.

[0505] 50. The lentiviral vector or pharmaceutical composition for use according to any of paras 34, 37-49, the use according to any of paras 35, 37-49, or the method according to any of paras 36-49, wherein the formation of liver fibrosis is slowed and/or reduced.

[0506] All publications mentioned in the above specification are herein incorporated by reference. Various modifications and variations of the disclosed vectors, cells, compositions, kits and uses of the invention will be apparent to the skilled person without departing from the scope and spirit of the invention. Although the invention has been disclosed in connection with specific preferred embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the disclosed modes for carrying out the invention, which are obvious to the skilled person are intended to be within the scope of the following claims.

Claims

1. An immune-shielded lentiviral vector comprising a nucleotide sequence encoding a progressive familial intrahepatic cholestasis (PFIC)-associated polypeptide.

2. The lentiviral vector according to claim 1, wherein the PFIC-associated polypeptide is selected from familial intrahepatic cholestasis type 1 (FIC1), or a fragment thereof; bile salt export pump (BSEP), or a fragment thereof; multiple drug resistance 3 (MDR3), or a fragment thereof; tight junction protein 2 (TJP2), or a fragment thereof; farnesoid X receptor (FXR), or a fragment thereof; and Myosin-Vb (MYO5B), or a fragment thereof.

3. The lentiviral vector according to claim 1 or 2, wherein the PFIC-associated polypeptide is selected from BSEP, or a fragment thereof; FXR, or a fragment thereof; and MYO5B, or a fragment thereof; preferably wherein the PFIC-associated polypeptide is BSEP, or a fragment thereof.
4. The lentiviral vector according to any preceding claim, wherein the PFIC-associated polypeptide comprises or consists of an amino acid sequence which is at least 70% identical to one of SEQ ID NOs: 37, 39, 43, 45, 47, 48 or 50, or a fragment thereof, preferably wherein the PFIC-associated polypeptide comprises or consists of an amino acid sequence which is at least 70% identical to SEQ ID NO: 39 or a fragment thereof.
5. The lentiviral vector according to any preceding claim, wherein the nucleotide sequence encoding a PFIC-associated polypeptide comprises or consists of a nucleotide sequence which is at least 70% identical to one of SEQ ID NOs: 38, 40, 41, 42, 44, 46, 49 or 51, or a fragment thereof, preferably wherein the nucleotide sequence encoding a PFIC-associated polypeptide comprises or consists of a nucleotide sequence which is at least 70% identical to one of SEQ ID NOs: 40-42, or a fragment thereof.
6. The lentiviral vector according to any preceding claim, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is codon-optimised, preferably wherein the nucleotide sequence encoding a PFIC-associated polypeptide comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 41 or 42, or a fragment thereof.
7. The lentiviral vector according to any preceding claim, wherein the lentiviral vector is a CD47.sup.high lentiviral vector.
8. The lentiviral vector according to any preceding claim, wherein the lentiviral vector is obtained from a CD47.sup.high host cell, optionally wherein the host cell is genetically engineered to increase expression of CD47 on the cell surface.
9. The lentiviral vector according to any preceding claim, wherein the lentiviral vector has at least about 2-fold more CD47 on its surface than a lentiviral vector obtained from an unmodified host cell.
10. The lentiviral vector according to any preceding claim, wherein the lentiviral vector is a MHC-I.sup.free lentiviral vector.
11. The lentiviral vector according to any preceding claim, wherein the lentiviral vector is obtained from a MHC-I.sup.free host cell, optionally wherein the host cell is genetically engineered to disrupt expression of MHC-I on the cell surface.
12. The lentiviral vector according to any preceding claim, wherein MHC-I is not detectable on the surface of the lentiviral vector.
13. The lentiviral vector according to any preceding claim, wherein the lentiviral vector is a CD47.sup.high/MHC-I.sup.free lentiviral vector.
14. The lentiviral vector according to any preceding claim, wherein the lentiviral vector is obtained from a CD47.sup.high/MHC-I.sup.free host cell.
15. The lentiviral vector according to any preceding claim, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to one or more miRNA target sequences.
16. The lentiviral vector according to claim 15, wherein the one or more miRNA target sequences suppress transgene expression in one or more cells other than hepatocytes, preferably wherein the one or more miRNA target sequence suppress transgene expression in hematopoietic-lineage cells and/or antigen-presenting cells.
17. The lentiviral vector according to claim 15 or 16, wherein the one or more miRNA target sequences are selected from miR-181, miR-142, miR-223, and miR-155 target sequences.
18. The lentiviral vector according to any preceding claim, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to one or more mir-142 target sequence, two or more mir-142 target sequences, three or more mir-142 target sequences, or four or more mir-142 target sequences.
19. The lentiviral vector according to any preceding claim, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to four mir-142 target sequences.
20. The lentiviral vector according to any of claims 15-19, wherein the one or more miRNA target sequences comprise or consist of a nucleotide sequence which is at least 90% identical to SEQ ID NO: 17, preferably wherein the one or more miRNA target sequences comprise or consist of a nucleotide sequence which is at least 90% identical to SEQ ID NO: 18.
21. The lentiviral vector according to any preceding claim, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a liver-specific promoter, preferably wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a hepatocyte-specific promoter.
22. The lentiviral vector according to any preceding claim, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a transthyretin (TTR) promoter, an alpha-1-antitrypsin (AAT) promoter, a thyroxine-binding globulin (TBG) promoter, a APoE/hAAT promoter, a HCR-hAAT promoter, a LP1 promoter, or a HLP promoter.

- 23.** The lentiviral vector according to any preceding claim, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a transthyretin (TTR) promoter, preferably wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a Enh1mTTR (ET) promoter.
- 24.** The lentiviral vector according to any preceding claim, wherein the nucleotide sequence encoding a PFIC-associated polypeptide is operably linked to a promoter which comprises or consists of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 19.
- 25.** The lentiviral vector according to any preceding claim, wherein the lentiviral vector is pseudotyped, preferably wherein the lentiviral vector is VSV.G-pseudotyped.
- 26.** The lentiviral vector according to any preceding claim, wherein the lentiviral vector is a self-inactivating (SIN) lentiviral vector, preferably wherein the lentiviral vector comprises self-inactivating (SIN) LTRs which comprise or consist of a nucleotide sequence which is at least 70% identical to SEQ ID NO: 23, or a fragment thereof.
- 27.** The lentiviral vector according to any preceding claim, wherein the lentiviral vector is an integrating lentiviral vector and/or a replication-defective lentiviral vector.
- 28.** The lentiviral vector according to any preceding claim, wherein the lentiviral vector comprises a nucleotide sequence which is at least 70% identical to SEQ ID NO: 36.
- 29.** An isolated cell comprising a lentiviral vector according to any of claims 1-28.
- 30.** A pharmaceutical composition comprising a lentiviral vector according to any of claims 1-28, in combination with a pharmaceutically acceptable carrier, diluent or excipient.
- 31.** The lentiviral vector according to any of claims 1-28 or pharmaceutical composition according to claim 30, for use as a medicament.
- 32.** Use of a lentiviral vector according to any of claims 1-28 or a pharmaceutical composition according to claim 30, for the manufacture of a medicament.
- 33.** A method comprising administering a therapeutically effective amount of a lentiviral vector according to any of claims 1-28 or a pharmaceutical composition according to claim 30, to a subject in need thereof.
- 34.** The lentiviral vector according to any of claims 1-28 or pharmaceutical composition according to claim 30, for use in preventing or treating progressive familial intrahepatic cholestasis (PFIC).
- 35.** Use of a lentiviral vector according to any of claims 1-28 or a pharmaceutical composition according to claim 30, for the manufacture of a medicament for preventing or treating progressive familial intrahepatic cholestasis (PFIC).
- 36.** A method of preventing or treating progressive familial intrahepatic cholestasis (PFIC), comprising administering a therapeutically effective amount of a lentiviral vector according to any of claims 1-28 or a pharmaceutical composition according to claim 30, to a subject in need thereof.
- 37.** The lentiviral vector or pharmaceutical composition for use according to claim 34, the use according to claim 35, or the method according to claim 36, wherein the PFIC-associated polypeptide is FIC1, or a fragment thereof, and the PFIC is PFIC type 1 (PFIC-1); the PFIC-associated polypeptide is BSEP, or a fragment thereof, and the PFIC is PFIC type 2 (PFIC-2); the PFIC-associated polypeptide is MDR3, or a fragment thereof, and the PFIC is PFIC type 3 (PFIC-3); the PFIC-associated polypeptide is TJP2, or a fragment thereof, and the PFIC is PFIC type 4 (PFIC-4); the PFIC-associated polypeptide is FXR, or a fragment thereof, and the PFIC is PFIC type 5 (PFIC-5); or the PFIC-associated polypeptide is MYO5B, or a fragment thereof, and the PFIC is PFIC type 6 (PFIC-6).
- 38.** The lentiviral vector or pharmaceutical composition for use according to claim 34 or 37, the use according to claim 35 or 37, or the method according to claim 36 or 37, wherein the PFIC-associated polypeptide is BSEP, or a fragment thereof, and the PFIC is PFIC-2; the PFIC-associated polypeptide is FXR, or a fragment thereof, and the PFIC is PFIC-5; or the PFIC-associated polypeptide is MYO5B, or a fragment thereof, and the PFIC is PFIC-6, preferably wherein the PFIC-associated polypeptide is BSEP, or a fragment thereof, and the PFIC is PFIC-2.
- 39.** The lentiviral vector or pharmaceutical composition for use according to any of claims 31, 34, 37-38, the use according to any of claims 32, 35, 37-38, or the method according to any of claims 33, 36-39, wherein the subject is a human subject.
- 40.** The lentiviral vector or pharmaceutical composition for use according to any of claims 31, 34, 37-39, the use according to any of claims 32, 35, 37-39, or the method according to any of claims 33, 36-39, wherein the subject is a juvenile.
- 41.** The lentiviral vector or pharmaceutical composition for use according to any of claims 31, 34, 37-39, the use according to any of claims 32, 35, 37-39, or the method according to any of claims 33, 36-39, wherein the subject is a paediatric patient, preferably wherein the subject is a neonatal patient or an infantile patient.
- 42.** The lentiviral vector or pharmaceutical composition for use according to any of claims 31, 34, 37-41, the use according to any of claims 32, 35, 37-41, or the method according to any of claims 33, 36-41, wherein said lentiviral vector or said pharmaceutical composition is administered systemically, preferably wherein said

lentiviral vector or said pharmaceutical composition is administered by intravenous injection or intraperitoneal injection.

43. The lentiviral vector or pharmaceutical composition for use according to any of claims 31, 34, 37-41, the use according to any of claims 32, 35, 37-41, or the method according to any of claims 33, 36-41, wherein said lentiviral vector or said pharmaceutical composition is administered locally, preferably wherein said lentiviral vector or said pharmaceutical composition is administered by direct injection, intra-arterial injection, or intraportal injection.

44. The lentiviral vector or pharmaceutical composition for use according to claim 43, the use according to claim 43, or the method according to claim 43, wherein said lentiviral vector or said pharmaceutical composition is administered locally to the liver, preferably wherein said lentiviral vector or said pharmaceutical composition is administered by intrahepatic injection, intrahepatic arterial injection, or intraportal injection.

45. The lentiviral vector or pharmaceutical composition for use according to any of claims 31, 34, 37-44, the use according to any of claims 32, 35, 37-43, or the method according to any of claims 33, 36-44, wherein the lentiviral vector is administered at a dose of at least about 10.sup.8 TU/kg, at least about 10.sup.9 TU/kg, or at least about 10.sup.10 TU/kg.

46. The lentiviral vector or pharmaceutical composition for use according to any of claims 31, 34, 37-45, the use according to any of claims 32, 35, 37-45, or the method according to any of claims 33, 36-45, wherein the lentiviral vector is administered in a dose of from about 10.sup.8 to about 10.sup.11 TU/kg, from about 10.sup.8 to about 10.sup.10 TU/kg, or from about 10.sup.9 to about 10.sup.10 TU/kg.

47. The lentiviral vector or pharmaceutical composition for use according to any of claims 31, 34, 37-46, the use according to any of claims 32, 35, 37-46, or the method according to any of claims 33, 36-46, wherein the lentiviral vector integrates into the genome of liver cells and is maintained as the liver cells duplicate, preferably wherein the lentiviral vector integrates into the genome of hepatocytes and is maintained as the hepatocytes duplicate.

48. The lentiviral vector or pharmaceutical composition for use according to any of claims 34, 37-47, the use according to any of claims 35, 37-47, or the method according to any of claims 36-47, wherein serum bile acid levels are reduced and/or normalised.

49. The lentiviral vector or pharmaceutical composition for use according to any of claims 34, 37-48, the use according to any of claims 35, 37-48, or the method according to any of claims 36-48, wherein the liver function is improved, preferably wherein the serum level of bilirubin and/or one or more liver enzymes is reduced and/or normalised.

50. The lentiviral vector or pharmaceutical composition for use according to any of claims 34, 37-49, the use according to any of claims 35, 37-49, or the method according to any of claims 36-49, wherein the formation of liver fibrosis is slowed and/or reduced.

51. An immune-shielded lentiviral vector for use in a method of therapy, wherein the method comprises administration of the immune-shielded lentiviral vector to a juvenile or paediatric subject.

52. A cell for use in a method of therapy, wherein the cell comprises an immune-shielded lentiviral vector, and wherein the method comprises administration of the cell to a juvenile or paediatric subject.

53. The immune-shielded lentiviral vector or cell for use according to claim 51 or 52, wherein the subject is a neonatal subject or an infantile subject.
