potentially correct distant cells too as a result of metabolic cross-correction mechanisms, wherein corrected cells secrete α-galactosidase A that can correct bystander cells. Finally, the inventor's approach entails liver mediated expression of α-galactosidase A following *in-vivo*, AAV mediated gene transfer of hepatocytes, which results in tolerance to the transgenic protein, thereby reducing the risk of developing neutralising antibodies to α-galactosidase A.

Summary

In one aspect, the present invention refers to a nucleic acid molecule comprising a codon optimised nucleotide sequence encoding for a functional α-galactosidase A protein wherein: (a) the nucleotide sequence has at least 95% identity to the sequence of SEQ ID NO. 1, and the codons of the nucleotide sequence have been selected based on the codons used for proteins which are expressed at a high level in the liver; or (b) the nucleotide sequence has at least 98% identity to the sequence of SEQ ID NO. 1. In another aspect, the present invention refers to a vector for expressing α-galactosidase A protein, the vector comprising the nucleic acid molecule of the present invention. In still another aspect, the present invention refers to a host cell comprising the nucleic acid molecule of the present invention or the vector of the present invention. In yet another aspect, the present invention refers to a non-human transgenic animal comprising cells comprising the nucleic acid molecule of the present invention or the vector of the present invention. In one aspect, the present invention refers to a pharmaceutical composition comprising the nucleic acid molecule of the present invention or the vector of the present invention, and one or more pharmaceutically acceptable excipients. In a further aspect, the present invention refers to the vector of the present invention in the manufacture of a medicament for treating Fabry disease in a patient suffering from Fabry disease. In one other aspect, the present invention refers to the nucleic acid molecule of the present invention or the vector of the present invention for use in therapy. In yet another aspect, the present invention refers to the nucleic acid molecule of the present invention or the vector of the present invention for use in the treatment of Fabry disease. In another aspect, the present invention refers to the use of the nucleic acid molecule of the present invention or the vector of the present invention in the manufacture of a medicament for treating Fabry disease. In one other aspect, the present invention refers to the nucleic acid molecule of the present invention or the vector of the present invention for use in delivery of a nucleotide sequence encoding α-galactosidase A protein to a subject. Also described herein is a nucleic acid molecule comprising a nucleotide sequence

encoding for a functional α-galactosidase A protein wherein the nucleotide sequence has at least 85% identity to the sequence of SEQ ID NO. 1.

The inventors have surprisingly found that the novel codon optimised sequence of SEQ ID NO. 1 results in increased expression of the α-galactosidase A protein in hepatocytes transduced with an AAV vector under the control of a liver specific promoter, versus an identical construct containing the wild-type α-galactosidase A cDNA.

The nucleotide sequence has at least 85% identity to the sequence of SEQ ID NO. 1. In some examples, the nucleotide sequence has at least 86% identity to the sequence of SEQ ID NO. 1. In other examples, the nucleotide sequence has at least 87% identity to the sequence of SEQ ID NO. 1. In particular examples, the nucleotide sequence has at least 88% identity to the sequence of SEQ ID NO. 1. In further examples, the nucleotide sequence has at least 89% identity to the sequence of SEQ ID NO. 1. In some examples, the nucleotide sequence has at least 90% identity to the sequence of SEQ ID NO. 1. In other examples, the nucleotide sequence has at least 91% identity to the sequence of SEQ ID NO. 1. In particular examples, the nucleotide sequence has at least 92% identity to the sequence of SEQ ID NO. 1. In further examples, the nucleotide sequence has at least 93% identity to the sequence of SEQ ID NO. 1. In some examples, the nucleotide sequence has at least 94% identity to the sequence of SEQ ID NO. 1. In other examples, the nucleotide sequence has at least 95% identity to the sequence of SEQ ID NO. 1. In particular examples, the nucleotide sequence has at least 96% identity to the sequence of SEQ ID NO. 1. In further examples, the nucleotide sequence has at least 97% identity to the sequence of SEQ ID NO. 1. In some examples, the nucleotide sequence has at least 98% identity to the sequence of SEQ ID NO. 1. In other examples, the nucleotide

sequence has at least 99% identity to the sequence of SEQ ID NO. 1. In particular examples, the nucleotide sequence has the sequence of SEQ ID NO. 1.

The nucleotide sequence encodes for a functional α-galactosidase A protein. A functional α-galactosidase A protein hydrolyses the terminal alpha-galactosyl moieties from glycolipids and glycoproteins. Suitable methods for assaying for α-galactosidase A activity are well known to those skilled in the art. Preferably, the nucleotide sequence encodes for α-galactosidase A protein having the wild type amino acid sequence. This sequence is well known to those skilled in the art. For example, this information can be found on GenBank (http://www.ncbi.nlm.nih.gov/genbank) under accession number CAA29232.1 (GI: 757912). Further information on this protein can be found under NCBI Reference Sequence: NP_000160.1 (GI: 4504009). The wild type protein sequence has 429 amino acids.

Also described herein is a vector for expressing α-galactosidase A protein.

The vector comprises the nucleic acid molecule described above. This means that the vector contains a nucleotide sequence encoding for a functional α-galactosidase A protein so that when this sequence is expressed, a functional α-galactosidase A protein is produced by the cell in which the vector is contained.

The sequence of SEQ ID NO. 1 is a codon optimised α-galactosidase A nucleotide sequence. This sequence has not been codon optimised in a normal way. Instead, the codons have been selected based on the codons used for proteins which are expressed at a high level in the liver. The reason for this is that the vector is normally expressed in the liver. This special codon optimisation process has been found to produce a nucleotide sequence which gives surprisingly high expression.

The nucleotide sequence encoding for a α-galactosidase A protein is preferably between 1265 and 1315 nucleotides in length. In some examples, the nucleotide sequence encoding for a functional α-galactosidase A protein is between 1270 and 1310 nucleotides in length. In other examples, the nucleotide sequence encoding for a functional α-galactosidase A protein is between 1275 and 1305 nucleotides in length. In particular examples, the

nucleotide sequence encoding for a functional α-galactosidase A protein is between 1280 and 1300 nucleotides in length.

Preferably the vector further comprises a promoter. The promoter causes expression of the nucleotide sequence encoding for a functional α-galactosidase A protein. Any appropriate promoter may be used, such as HLP, LP1, HCR-hAAT, ApoE-hAAT, and LSP. These promoters are described in more detail in the following references: HLP: McIntosh J. *et al.*, Blood 2013 Apr 25, 121(17):3335-44; LP1: Nathwani *et al.*, Blood. 2006 April 1, 107(7): 2653–2661; HCR-hAAT: Miao *et al.,* Mol Ther. 2000;1: 522-532; ApoE-hAAT: Okuyama *et al.,* Human Gene Therapy, 7, 637-645 (1996); and LSP: Wang *et al.,* Proc Natl Acad Sci U S A. 1999 March 30, 96(7): 3906–3910. A preferred promoter is also described in WO 2011/005968. Another preferred promoter has the sequence of SEQ ID NO. 2 and is referred to as HLP2. The promoter having SEQ ID NO. 2 is a liver specific promoter which has been found to give particularly good expression in the liver. Whilst giving good expression, this promoter is also relatively small which allows more efficient packaging of the vector. Preferably, the promoter is a liver specific promoter.

SEQ ID NO. 3 is the nucleotide sequence of a vector construct including a promoter and a nucleotide sequence encoding for a functional α-galactosidase A protein. Therefore, the nucleic acid molecule as described herein may comprise a nucleotide sequence having at least 90% identity to the sequence of SEQ ID NO. 3. In other examples, the nucleotide sequence has at least 91% identity to the sequence of SEQ ID NO. 3. In particular examples, the nucleotide sequence has at least 92% identity to the sequence of SEQ ID NO. 3. In further examples, the nucleotide sequence has at least 93% identity to the sequence of SEQ ID NO. 3. In some examples, the nucleotide sequence has at least 94% identity to the sequence of SEQ ID NO. 3. In other examples, the nucleotide sequence has at least 95% identity to the sequence of SEQ ID NO. 3. In particular examples, the nucleotide sequence has at least 96% identity to the sequence of SEQ ID NO. 3. In further examples, the nucleotide sequence has at least 97% identity to the sequence of SEQ ID NO. 3. In some examples, the nucleotide sequence has at least 98% identity to the sequence of SEQ ID NO. 3. In other examples, the nucleotide sequence has at least 99% identity to the sequence of SEQ ID NO. 3. In particular examples, the nucleotide sequence has the sequence of SEQ ID NO. 3.

The vector may be any appropriate vector for expressing the α-galactosidase A protein, including viral and non-viral vectors. Viral vectors include a parvovirus, an adenovirus, a retrovirus, a lentivirus or a herpes simplex virus. The parvovirus may be an adenovirus-associated virus (AAV). The vector is preferably a recombinant adeno-associated viral (rAAV) vector or a lentiviral vector. More preferably, the vector is a rAAV vector.

A vector as described herein may be a gene delivery vector. Such a gene delivery vector may be a viral gene delivery vector or a non-viral gene delivery vector.

Accordingly, described herein are gene delivery vectors based on animal parvoviruses, in particular dependoviruses such as infectious human or simian AAV, and the components thereof (e.g., an animal parvovirus genome) for use as vectors for introduction and/or expression of an α-galactosidase A protein in a mammalian cell. The term "parvoviral" as used herein thus encompasses dependoviruses such as any type of AAV.

Viruses of the Parvoviridae family are small DNA animal viruses. The family Parvoviridae may be divided between two subfamilies: the Parvovirinae, which infect vertebrates, and the Densovirinae, which infect insects. Members of the subfamily Parvovirinae are herein referred to as the parvoviruses and include the genus Dependovirus. As may be deduced from the name of their genus, members of the Dependovirus are unique in that they usually require coinfection with a helper virus such as adenovirus or herpes virus for productive infection in cell culture. The genus Dependovirus includes AAV, which normally infects humans (e.g., serotypes 1, 2, 3A, 3B, 4, 5, and 6) or primates (e.g., serotypes 1 and 4), and related viruses that infect other warm-blooded animals (e.g., bovine, canine, equine, and ovine adeno-associated viruses). Further information on parvoviruses and other members of the Parvoviridae is described in Kenneth I. Berns, "Parvoviridae: The Viruses and Their Replication," Chapter 69 in Fields Virology (3d Ed. 1996). For convenience, the present disclosure is further exemplified and described herein by reference to AAV. It is, however, understood that what is described herein is not limited to AAV but may equally be applied to other parvoviruses.

The genomic organization of all known AAV serotypes is very similar. The genome of AAV is a linear, single-stranded DNA molecule that is less than about 5,000 nucleotides (nt) in length. Inverted terminal repeats (ITRs) flank the unique coding nucleotide sequences for the non-structural replication (Rep) proteins and the structural (VP) proteins. The VP proteins (VP1, -2 and -3) form the capsid. The terminal 145 nt are self-complementary and are organized so that an energetically stable intramolecular duplex forming a T-shaped hairpin may be formed. These hairpin structures function as an origin for viral DNA replication, serving as primers for the cellular DNA polymerase complex. Following wild type (wt) AAV infection in mammalian cells the Rep genes (i.e. encoding Rep78 and Rep52 proteins) are expressed from the P5 promoter and the P19 promoter, respectively, and both Rep proteins have a function in the replication of the viral genome. A splicing event in the Rep ORF results in the expression of actually four Rep proteins (i.e. Rep78, Rep68, Rep52 and Rep40). However, it has been shown that the unspliced mRNA, encoding Rep78 and Rep52 proteins, in mammalian cells are sufficient for AAV vector production. Also in insect cells the Rep78 and Rep52 proteins suffice for AAV vector production.

In an AAV suitable for use as a gene therapy vector, the vector genome typically comprises a nucleic acid to be packaged for delivery to a target cell. According to this particular example, the heterologous nucleotide sequence is located between the viral ITRs at either end of the vector genome. In further preferred examples, the parvovirus (e.g. AAV) cap genes and parvovirus (e.g. AAV) rep genes are deleted from the template genome (and thus from the virion DNA produced therefrom). This configuration maximizes the size of the nucleic acid sequence(s) that can be carried by the parvovirus capsid.

According to this particular example, the nucleic acid is located between the viral ITRs at either end of the substrate. It is possible for a parvoviral genome to function with only one ITR. Thus, in a gene therapy vector as described herein based on a parvovirus, the vector genome is flanked by at least one ITR, but, more typically, by two AAV ITRs (generally with one either side of the vector genome, i.e. one at the 5' end and one at the 3' end). There may be intervening sequences between the nucleic acid in the vector genome and one or more of the ITRs.

Preferably, the nucleotide sequence encoding a functional α-galactosidase A protein (for expression in the mammalian cell) will be incorporated into a parvoviral genome located between two regular ITRs or located on either side of an ITR engineered with two D regions.

AAV sequences that may be used herein for the production of AAV gene therapy vectors can be derived from the genome of any AAV serotype. Generally, the AAV serotypes have genomic sequences of significant homology at the amino acid and the nucleic acid levels, provide an identical set of genetic functions, produce virions which are essentially physically and functionally equivalent, and replicate and assemble by practically identical mechanisms. For the genomic sequence of the various AAV serotypes and an overview of the genomic similarities see e.g. GenBank Accession number U89790; GenBank Accession number J01901; GenBank Accession number AF043303; GenBank Accession number AF085716; Chiorini *et al*, 1997; Srivastava *et al*, 1983; Chiorini *et al*, 1999; Rutledge *et al*, 1998; and Wu *et al*, 2000. AAV serotype 1, 2, 3, 4, 5, 6, 7, 8 or 9 may be used herein. However, AAV serotypes 1, 5 or 8 are preferred sources of AAV sequences for use in the context of the present disclosure. The sequences from the AAV serotypes may be mutated or engineered when being used in the production of gene therapy vectors.

Preferably, the AAV ITR sequences for use in the context of the present disclosure are derived from AAV1, AAV2, AAV4 and/or AAV6. Likewise, the Rep (Rep78 and Rep52) coding sequences are preferably derived from AAV1, AAV2, AAV4 and/or AAV6. The sequences coding for the VP1, VP2, and VP3 capsid proteins for use in the context of the present disclosure may however be taken from any of the known 42 serotypes, more preferably from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 or AAV9 or newly developed AAV-like particles obtained by e.g. capsid shuffling techniques and AAV capsid libraries.

AAV Rep and ITR sequences are particularly conserved among most serotypes. The Rep78 proteins of various AAV serotypes are e.g. more than 89% identical and the total nucleotide sequence identity at the genome level between AAV2, AAV3A, AAV3B, and AAV6 is around 82% (Bantel-Schaal *et al*, 1999). Moreover, the Rep sequences and ITRs of many AAV serotypes are known to efficiently cross-complement (i.e., functionally substitute) corresponding sequences from other serotypes in production of AAV particles in mammalian

cells. US 2003148506 reports that AAV Rep and ITR sequences also efficiently cross-complement other AAV Rep and ITR sequences in insect cells.

The AAV VP proteins are known to determine the cellular tropicity of the AAV virion. The VP protein-encoding sequences are significantly less conserved than Rep proteins and genes among different AAV serotypes. The ability of Rep and ITR sequences to cross-complement corresponding sequences of other serotypes allows for the production of pseudotyped AAV particles comprising the capsid proteins of a serotype (e.g., AAV1, 5 or 8) and the Rep and/or ITR sequences of another AAV serotype (e.g., AAV2). Such pseudotyped rAAV particles are a part of the present invention.

Modified "AAV" sequences also can be used in the context of the present disclosure, e.g. for the production of AAV gene therapy vectors. Such modified sequences e.g. include sequences 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%, or more nucleotide and/or amino acid sequence identity (e.g., a sequence having about 75-99% nucleotide sequence identity) to an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8 or AAV9 ITR, Rep, or VP can be used in place of wild-type AAV ITR, Rep, or VP sequences.

Although similar to other AAV serotypes in many respects, AAV5 differs from other human and simian AAV serotypes more than other known human and simian serotypes. In view thereof, the production of rAAV5 can differ from production of other serotypes in insect cells. Where methods as described herein are employed to produce rAAV5, it is preferred that one or more constructs comprising, collectively in the case of more than one construct, a nucleotide sequence comprising an AAV5 ITR, a nucleotide sequence comprises an AAV5 Rep coding sequence (i.e. a nucleotide sequence comprises an AAV5 Rep78). Such ITR and Rep sequences can be modified as desired to obtain efficient production of AAV5 or pseudotyped AAV5 vectors. For example, the start codon of the Rep sequences can be modified, VP splice sites can be modified or eliminated, and/or the VP1 start codon and nearby nucleotides can be modified to improve the production of AAV5 vectors.

Thus, the viral capsid used herein may be from any parvovirus, either an autonomous parvovirus or dependovirus, as described above. Preferably, the viral capsid is

an AAV capsid (e. g., AAV1, AAV2, AAV3, AAV4, AAV5 or AAV6 capsid). In general, the AAV1 capsid or AAV6 capsid are preferred. The choice of parvovirus capsid may be based on a number of considerations as known in the art, e.g., the target cell type, the desired level of expression, the nature of the heterologous nucleotide sequence to be expressed, issues related to viral production, and the like. For example, the AAV1 and AAV6 capsid may be advantageously employed for skeletal muscle; AAV1, AAV5 and AAV8 for the liver and cells of the central nervous system (e.g., brain); AAV5 for cells in the airway and lung or brain; AAV3 for bone marrow cells; and AAV4 for particular cells in the brain (e. g., appendable cells).

It is within the technical skills of the skilled person to select the most appropriate virus, virus subtype or virus serotype. Some subtypes or serotypes may be more appropriate than others for a certain type of tissue.

For example, liver-specific expression of a nucleic acid as described herein may advantageously be induced by AAV-mediated transduction of liver cells. Liver is amenable to AAV-mediated transduction, and different serotypes may be used (for example, AAV1, AAV5 or AAV8). Transduction of muscle may be accomplished by administration of an AAV encoding a nucleic acid via the blood stream. Thus, intravenous or intra-arterial administration is applicable.

A parvovirus gene therapy vector prepared according to the disclosure may be a "hybrid" particle in which the viral TRs and viral capsid are from different parvoviruses. Preferably, the viral TRs and capsid are from different serotypes of AAV. Likewise, the parvovirus may have a "chimeric" capsid (e. g., containing sequences from different parvoviruses, preferably different AAV serotypes) or a "targeted" capsid (e. g., a directed tropism).

In the context of the disclosure "at least one parvoviral ITR nucleotide sequence" is understood to mean a palindromic sequence, comprising mostly complementary, symmetrically arranged sequences also referred to as "A," "B," and "C" regions. The ITR functions as an origin of replication, a site having a "cis" role in replication, i.e., being a recognition site for trans-acting replication proteins such as e.g. Rep 78 (or Rep68) which recognize the palindrome and specific sequences internal to the palindrome. One exception to

the symmetry of the ITR sequence is the "D" region of the ITR. It is unique (not having a complement within one ITR). Nicking of single-stranded DNA occurs at the junction between the A and D regions. It is the region where new DNA synthesis initiates. The D region normally sits to one side of the palindrome and provides directionality to the nucleic acid replication step. A parvovirus replicating in a mammalian cell typically has two ITR sequences. It is, however, possible to engineer an ITR so that binding sites are on both strands of the A regions and D regions are located symmetrically, one on each side of the palindrome. On a double-stranded circular DNA template (e.g., a plasmid), the Rep78- or Rep68-assisted nucleic acid replication then proceeds in both directions and a single ITR suffices for parvoviral replication of a circular vector. Thus, one ITR nucleotide sequence can be used in the context of the present disclosure. Preferably, however, two or another even number of regular ITRs are used. Most preferably, two ITR sequences are used. A preferred parvoviral ITR is an AAV ITR. For safety reasons it may be desirable to construct a parvoviral (AAV) vector that is unable to further propagate after initial introduction into a cell. Such a safety mechanism for limiting undesirable vector propagation in a recipient may be provided by using AAV with a chimeric ITR as described in US 2003148506.

Those skilled in the art will appreciate that the viral Rep protein(s) used for producing an AAV vector as described herein may be selected with consideration for the source of the viral ITRs. For example, the AAV5 ITR typically interacts more efficiently with the AAV5 Rep protein, although it is not necessary that the serotype of ITR and Rep protein(s) are matched.

The ITR(s) used herein are typically functional, i.e. they may be fully resolvable and are preferably AAV sequences, with serotypes 1, 2, 3, 4, 5 or 6 being preferred. Resolvable AAV ITRs according to the present disclosure need not have a wild-type ITR sequence (e. g., a wild-type sequence may be altered by insertion, deletion, truncation or missense mutations), as long as the ITR mediates the desired functions, e. g., virus packaging, integration, and/or provirus rescue, and the like.

Advantageously, by using a gene therapy vector as compared with previous approaches, the restoration of protein synthesis, i.e. α-galactosidase A synthesis, is a characteristic that the transduced cells acquire permanently or for a sustained period of time, thus avoiding the need for continuous administration to achieve a therapeutic effect.

Accordingly, the vectors as described herein therefore represent a tool for the development of strategies for the *in vivo* delivery of an α-galactosidase A nucleotide sequence, by engineering the nucleic acid within a gene therapy vector that efficiently transduces an appropriate cell type, such as a liver cell.

The vector may be a single stranded vector or a self-complementary vector. In some examples, the vector is a single stranded vector. In other examples, the vector is a self-complementary vector.

The vector may further comprise a poly A tail. Preferably, this is positioned downstream of the nucleotide sequence encoding for a functional α-galactosidase A protein. Preferably, the poly A tail is a bovine growth hormone poly A tail. Preferably, this is between 250 and 270 nucleotides in length.

The vector may comprise other elements to allow the functional α-galactosidase A protein to be expressed. Such elements are well known to a person skilled in the art.

Preferably, the nucleic acids described above are isolated.

It would be well with the capabilities of a skilled person to produce the nucleic acid molecules described above. This could be done, for example, using chemical synthesis of a given sequence.

Further, a skilled person would readily be able to determine whether a nucleic acid expresses a functional protein. Suitable methods would be apparent to those skilled in the art. For example, one suitable *in vitro* method involves inserting the nucleic acid into a vector, such as a lentiviral or an AAV vector, transducing host cells, such as 293T or HeLa cells, with the vector, and assaying for α-galactosidase A activity. Alternatively, a suitable *in vivo* method involves transducing a vector containing the nucleic acid into mice with Fabry disease and assaying for functional α-galactosidase A in the plasma of the mice. Suitable methods are described in more detail below.

The nucleic acid can be any type of nucleic acid composed of nucleotides. The nucleic acid should be able to be expressed so that a protein is produced. Preferably, the nucleic acid is DNA or RNA.

The disclosure also describes a host cell comprising any one of the nucleic acid molecules or vectors described above. Preferably, the vector is capable of expressing the α-galactosidase A nucleotide sequence in the host. The host may be any suitable host.

As used herein, the term "host" refers to organisms and/or cells which harbour a nucleic acid molecule or a vector as described herein, as well as organisms and/or cells that are suitable for use in expressing a recombinant gene or protein. It is not intended that the present disclosure be limited to any particular type of cell or organism. Indeed, it is contemplated that any suitable organism and/or cell will find use in the present disclosure as a host. A host cell may be in the form of a single cell, a population of similar or different cells, for example in the form of a culture (such as a liquid culture or a culture on a solid substrate), an organism or part thereof.

A host cell as described herein may permit the expression of a nucleic acid molecule as described herein. Thus, the host cell may be, for example, a bacterial, a yeast, an insect or a mammalian cell.

In addition, also described herein is a transgenic animal comprising cells comprising the nucleic acid molecule encoding for a functional α-galactosidase A protein described above or a vector described above. Preferably the animal is a non-human mammal, especially a primate. Alternatively, the animal may be a rodent, especially a mouse; or may be canine, feline, ovine or porcine.

Also described herein is a pharmaceutical composition comprising a nucleic acid molecule or a vector as described herein and one or more pharmaceutically acceptable excipients. The one or more excipients include carriers, diluents and/or other medicinal agents, pharmaceutical agents or adjuvants, etc.

Also described herein is a method of treating Fabry disease comprising administering a therapeutically effective amount of a vector as described above to a patient suffering from Fabry disease. Preferably, the patient is human.

When Fabry disease is "treated" in the above method, this means that one or more symptoms of Fabry disease are ameliorated. It does not mean that the symptoms of Fabry disease are completely remedied so that they are no longer present in the patient, although in some methods, this may be the case. The method of treating results in one or more of the symptoms of Fabry disease being less severe than before treatment.

A "therapeutically effective amount" refers to an amount effective, at dosages and for periods of time necessary, to achieve the desired therapeutic result, such as raising the level of functional α-galactosidase A in a subject (so as to lead to functional α-galactosidase A production to a level sufficient to ameliorate the symptoms of Fabry disease).

Delivery of a nucleic acid or vector as described herein to a host cell *in vivo* may result in an increase of functional α-galactosidase A in the host, for example to a level that ameliorates one or more symptoms of Fabry disease.

The level of naturally occurring α-galactosidase A in a subject suffering from Fabry disease varies depending on the severity of the Fabry disease. Patients with a severe form of the disease have α-galactosidase A levels of less than about 1% of the level found in a normal healthy subject (referred to herein as "a normal level"). It has been found that when the method of treatment as described herein is used, it can cause an increase in the level of functional α-galactosidase A to at least about 1% of normal levels. In some examples, the method of treatment as described herein causes an increase in the level of functional α-galactosidase A to at least about 2%, at least about 3%, at least about 4%, at least about 10%, at least about 15%, at least about 20% or at least about 25% of normal levels. In a particular example, the method of treatment as described herein causes an increase in the level of functional α-galactosidase A to at least about 30% of normal levels.

In one example, the method of treatment as described herein causes an increase in the level of functional α-galactosidase A to, at most, normal levels.

The activity of functional α-galactosidase A can be measured relatively easily and methods for measuring α-galactosidase A activity are well known to those skilled in the art. The activity of α-galactosidase can conveniently be measured in blood using a blood spot as described in Clin. Biochem. 45(15):1233-8 (2012). The principle of this method is that at acidic pH, α-galactosidase hydrolyses the substrate, 4-methylumbelliferyl-a-D-galactopyranoside, to 4-methylumbelliferone and galactose. Adding alkaline buffer stops the enzyme reaction and causes the 4-methylumbelliferone to fluoresce at a different wavelength from the unhydrolysed substrate, thereby permitting its measurement in the presence of a vast excess of unhydrolysed substrate. In leucocytes usually over 95% of the total α-galactosidase activity is α-galactosidase A whereas in plasma and cultured cells, the isoenzyme, α-galactosidase B can significantly contribute to the total α-galactosidase activity. α-galactosidase A can be measured in the presence of α-galactosidase B by making use of the increased heat liability of the A isoenzyme and in plasma the α-galactosidase B can be inhibited by the addition of a-NAc galactosamine. A key advantage of this method is that only 5ul spot of dried whole blood on filter paper is required. This offers the advantage of measuring α-galactosidase levels in real time as the experiment in Fabry Ko mouse proceeds following vector administration.

Alternatively, α-galactosidase A activity in plasma can be assessed on a terminal bleed in mice following gene transfer. This method is based on the fact that at acidic pH, α-galactosidase hydrolyses the substrate, 4-methylumbelliferyl-a-D-galactopyranoside, to 4-methylumbelliferone and galactose, as above. In addition, α-galactosidase can also be measured using standard Western blot assay or standard (ELISA type) immunoassays which show antigen levels.

Further, also described herein is a nucleic acid molecule encoding for a functional α-galactosidase A protein as described above, or a vector as described above for use in therapy, for example, in the treatment of Fabry disease.

In addition, also described herein is the use of the nucleic acid molecule encoding for a functional α-galactosidase A protein as described above or a vector as described above in the manufacture of a medicament for treating Fabry disease.

Also described herein is a method for delivery of a nucleotide sequence encoding a functional α-galactosidase A protein to a subject, which method comprises administering to the said subject a nucleic acid molecule encoding a functional α-galactosidase A protein as described above or a vector as described above.

In the description above, the term "identity" is used to refer to the similarity of two sequences. For the purpose of this disclosure, it is defined here that in order to determine the percent identity of two nucleotide sequences, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first nucleic acid for optimal alignment with a second amino or nucleic acid sequence). The nucleotide residues at nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid or nucleotide residue as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = number of identical positions/total number of positions (i.e. overlapping positions) x 100). Preferably, the two sequences are the same length.

A sequence comparison may be carried out over the entire lengths of the two sequences being compared or over fragment of the two sequences. Typically, the comparison will be carried out over the full length of the two sequences being compared. However, sequence identity may be carried out over a region of, for example, about twenty, about fifty, about one hundred, about two hundred, about five hundred, about 1000 or about 2000 or more contiguous nucleic acid residues.

The skilled person will be aware of the fact that several different computer programs are available to determine the homology or identity between two sequences. In preferred examples, the identity between two sequences is analysed using the software package Clone Manager Professional version 9 (preferably, version 9.4). This analysis tool is produced by Sci-Ed Software (Scientific & Educational Software, 11010 Lake Grove Blvd, Ste 100, PMB 122, Morrisville, NC 27560, USA - http://www.scied.com/index.htm). The settings used to compare the sequences are preferably as follows: alignment: Global DNA alignment; parameters: both strands; scoring matrix: linear (mismatch 2, OpenGap 4, ExtGap

CLAIMS

1. A nucleic acid molecule comprising a codon optimised nucleotide sequence encoding for a functional α-galactosidase A protein wherein:

(a) the nucleotide sequence has at least 95% identity to the sequence of SEQ ID NO. 1, and the codons of the nucleotide sequence have been selected based on the codons used for proteins which are expressed at a high level in the liver; or

(b) the nucleotide sequence has at least 98% identity to the sequence of SEQ ID NO. 1.

2. The nucleic acid molecule of claim 1, wherein the nucleotide sequence has at least 98% or 99% identity to the sequence of SEQ ID NO. 1.

3. The nucleic acid molecule of claim 1 or claim 2, wherein the nucleotide sequence has the sequence of SEQ ID NO. 1.

4. A vector for expressing α-galactosidase A protein, the vector comprising the nucleic acid molecule of any of claims 1 to 3.

5. The vector of claim 4 further comprising a liver specific promoter.

6. The vector of claim 4 or claim 5, wherein the vector is an AAV vector.

7. The vector of any of claims 4-6, wherein the vector is a single stranded vector.

8. The vector of any of claims 4-7, wherein the vector comprises a nucleotide sequence having at least 90% identity to the sequence of SEQ ID NO. 3.

9. The vector of any of claims 4-8, wherein the vector comprises a nucleotide sequence having the sequence of SEQ ID NO. 3.

10. A host cell comprising the nucleic acid molecule of any one of claims 1-3 or the vector of any one of claims 4-9.

11. A non-human transgenic animal comprising cells comprising the nucleic acid molecule of any one of claims 1-3 or the vector of any one of claims 4-9.

12. A pharmaceutical composition comprising the nucleic acid molecule of any one of claims 1-3 or the vector of any one of claims 4-9, and one or more pharmaceutically acceptable excipients.

13. Use of the vector of any one of claims 4-9 in the manufacture of a medicament for treating Fabry disease in a patient suffering from Fabry disease.

14. The nucleic acid molecule of any one of claims 1-3 or the vector of any one of claims 4-9 for use in therapy.

15. The nucleic acid molecule of any one of claims 1-3 or the vector of any one of claims 4-9 for use in the treatment of Fabry disease.

16. The use of the nucleic acid molecule of any one of claims 1-3 or the vector of any one of claims 4-9 in the manufacture of a medicament for treating Fabry disease.

17. The nucleic acid molecule of any one of claims 1-3 or the vector of any one of claims 4-9 for use in delivery of a nucleotide sequence encoding α-galactosidase A protein to a subject.