(19) DANMARK

(10) **DK/EP 3233130 T3**

Patent- og Varemærkestyrelsen

(12) Oversættelse af europæisk patentskrift

(51) Int.Cl.: ***A 61 K 48/00 (2006.01)*** ***A 61 P 1/16 (2006.01)*** ***C 12 N 9/14 (2006.01)***
C 12 N 15/864 (2006.01)

(45) Oversættelsen bekendtgjort den: **2021-06-21**

(80) Dato for Den Europæiske Patentmyndigheds bekendtgørelse om meddelelse af patentet: **2021-03-31**

(86) Europæisk ansøgning nr.: **15817240.3**

(86) Europæisk indleveringsdag: **2015-12-17**

(87) Den europæiske ansøgnings publiceringsdag: **2017-10-25**

(86) International ansøgning nr.: **EP2015080357**

(87) Internationalt publikationsnr.: **WO2016097219**

(30) Prioritet: **2014-12-17 EP 14382531**

(84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**

(73) Patenthaver: **Fundacion para la Investigacion Medica Aplicada, Avda Pio XII, 55, 31008 Pamplona, Spanien**

(72) Opfinder: **MURILLO SAUCA, Oihana, CIMA - Oficina Propiedad Industrial, Avda Pío XII 55, 31008 Pamplona, Spanien**
GONZÁLEZ ASEGUINOLAZA, Gloria, Cima - Oficina de Propiedad, Industrial, Avda Pio XII 55, 31008 Pamplona, Spanien
HERNÁNDEZ ALCOCEBA, Rubén, CIMA - Oficina de Propiedad Industrial, Avda Pio XII 55, 31008 Pamplona, Spanien

(74) Fuldmægtig i Danmark: **Plougmann Vingtoft A/S, Strandvejen 70, 2900 Hellerup, Danmark**

(54) Benævnelse: **NUKLEINSYREKONSTRUKTIONER OG GENTERAPIVEKTORER TIL ANVENDELSE I BEHANDLINGEN AF WILSON-SYGDOM**

(56) Fremdragne publikationer:
DE-A1- 10 156 121
D. HUSTER ET AL: "The Distinct Roles of the N-terminal Copper-binding Sites in Regulation of Catalytic Activity of the Wilson's Disease Protein", JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 278, no. 34, 22 August 2003 (2003-08-22), pages 32212-32218, XP055254806, US ISSN: 0021-9258, DOI: 10.1074/jbc.M305408200
J L ROYBAL ET AL: "Early gestational gene transfer with targeted ATP7B expression in the liver improves phenotype in a murine model of Wilson's disease", GENE THERAPY, vol. 19, no. 11, 8 December 2011 (2011-12-08), pages 1085-1094, XP055193184, ISSN: 0969-7128, DOI: 10.1038/gt.2011.186 cited in the application
WU ET AL: "Adeno-associated Virus Serotypes: Vector Toolkit for Human Gene Therapy", MOLECULAR THERAPY, NATURE PUBLISHING GROUP, GB, vol. 14, no. 3, 12 August 2006 (2006-08-12), pages 316-327, XP005844824, ISSN: 1525-0016, DOI: 10.1016/J.YMTHE.2006.05.009
HUIMIN HU ET AL: "Preclinical Dose-Finding Study With a Liver-Tropic, Recombinant AAV-2/8 Vector in the Mouse Model of Galactosialidosis", MOLECULAR THERAPY, vol. 20, no. 2, 18 October 2011 (2011-10-18), pages

Fortsættes ...

267-274, XP055193008, ISSN: 1525-0016, DOI: 10.1038/mt.2011.227
RAJENDRA PILANKATTA ET AL: "Functional Effects of N-Metal Binding Domain Deletion and Specific Mutations on the ATP7B (Wilson Disease) Copper ATPase", BIOPHYSICAL JOURNAL, vol. 98, no. 3 suppl., 1 January 2010 (2010-01-01), page 448a, XP055193638,
J. R. FORBES ET AL: "Role of the Copper-binding Domain in the Copper Transport Function of ATP7B, the P-type ATPase Defective in Wilson Disease", JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 274, no. 18, 30 April 1999 (1999-04-30), pages 12408-12413, XP055193630, ISSN: 0021-9258, DOI: 10.1074/jbc.274.18.12408
PORTMANN R ET AL: "Purification and functional reconstitution of the human Wilson copper ATPase, ATP7B", FEBS LETTERS, ELSEVIER, AMSTERDAM, NL, vol. 579, no. 17, 4 July 2005 (2005-07-04) , pages 3589-3595, XP027697055, ISSN: 0014-5793 [retrieved on 2005-07-04]
LORINCZI ET AL: "Delivery of the Cu-transporting ATPase ATP7B to the plasma membrane in Xenopus oocytes", BIOCHIMICA ET BIOPHYSICA ACTA (BBA) - BIOMEMBRANES, ELSEVIER, AMSTERDAM, NL, vol. 1778, no. 4, 3 January 2008 (2008-01-03), pages 896-906, XP022559174, ISSN: 0005-2736
MURILLO OIHANA ET AL: "Long-term metabolic correction of Wilson's disease in a murine model by gene therapy", JOURNAL OF HEPATOLOGY, vol. 64, no. 2, 25 September 2015 (2015-09-25), pages 419-426, XP029389254, ISSN: 0168-8278, DOI: 10.1016/J.JHEP.2015.09.014

Description

TECHNICAL FIELD OF THE DISCLOSURE

[0001] The present disclosure relates to nucleic acid constructs and gene therapy vectors for use in the treatment of Wilson's disease and other conditions.

BACKGROUND ART

[0002] The state of the art regarding gene therapy of Wilson's disease was reviewed by Merle et al. (Current Gene Therapy 2007; 7: 217-220) and is here summarized and completed with later disclosed references.

[0003] Wilson's disease (WD) is an autosomal recessively inherited disorder of copper metabolism with an average prevalence of 1:30,000. WD is caused by mutations of the ATP7B gene coding for a P-type copper transporting ATPase, which is located on chromosome 13. ATP7B is expressed mainly in hepatocytes and functions in the transmembrane transport of copper. Absent or reduced function of ATP7B protein leads to decreased hepatocellular excretion of copper into bile and results in copper accumulation primarily in the liver and subsequently in the neurologic system and other tissues. Failure to incorporate copper into ceruloplasmin is an additional consequence of the loss of functional ATP7B protein.

[0004] WD can present clinically as liver disease, as a progressive neurologic disorder, or as psychiatric illness. Patients with hepatic WD usually present in late childhood or adolescence, and exhibit features of acute hepatitis, fulminant hepatic failure, or progressive chronic liver disease. Neurologic manifestations of WD typically present later than the liver disease, most often in the second or third decade and include extrapyramidal, cerebellar and cerebral-related symptoms.

[0005] The aim of medical treatment of WD is to remove the toxic deposit of copper from the body and to prevent its reaccumulation. Three anti-copper drugs are currently approved for WD: D-penicillamine, trientine, and zinc salts. Medical therapy is effective in most, but not all patients with WD. Liver transplantation is a therapeutic option in WD patients presenting with fulminant liver failure or progressive liver failure. It has been shown to correct the WD phenotype and provides excellent long-term survival.

[0006] However, an interruption of therapy or inadequate treatment can lead to fatalities within few months. Because WD medication has to be taken regularly, adherence to treatment in some patients, especially in adolescent WD patients, is poor.

[0007] Under therapy residual neurological symptoms are relatively common and even progressive symptoms can occur. Because current medical treatment options are not in all WD patients effective and adherence to therapy is a problem, a more comprehensive solution could involve gene therapy.

[0008] Theoretically, expression of wild type ATP7B in hepatocytes would reverse all disease-related abnormalities and rescue the liver and the neurological symptoms. The ultimate goal of an ideal gene therapy for WD would be to deliver ATP7B, in sufficient quantity, specifically to hepatocytes for a lifelong duration.

[0009] All published studies on adenoviral gene transfer for WD have used early-generation adenoviral vectors producing only transient transgene expression. Terada et al. [Terada et al. J. Biol. Chem. 1998; 273:1815-1820; Terada et al. FEBS Lett. 1999; 448: 53-56] demonstrated successful gene transfer by adenovirus mediated gene delivery in the LEC rat model. Restoration of holoceruloplasmin synthesis, of serum ceruloplasmin oxidase activity, and of copper excretion in bile was shown, indicating a therapeutic effect of the gene transfer. These effects were of a very limited duration, with a maximum level at day three and a decline thereafter. Ha-Hao et al. [Z. Gastroenterol. 2002; 40: 209-216] also demonstrated an increased copper content in stool of LEC rats after adenovirus-mediated ATP7B gene transfer, indicating increased copper excretion into the bile. The therapeutic effect was in addition demonstrated by restoration of holoceruloplasmin and of its ferroxidase activity. However, once again the duration of the therapeutic effect in these experiments was only transient with a limited duration of a few days.

[0010] Gutless adenoviral vectors have not been tested for this application so far.

[0011] Other commonly used non-integrating viral vector system, the adeno-associated virus (AAV), has neither been tested for WD so far, mainly because the ATP7B gene (approximately 4.4 kb large) leaves minimum space for allocating the rest of required sequences (e.g. promoter, poly A signal sequence, etc) within the AAV vector, whose packaging capacity is 4.4-4.7 kb. German patent application DE 100156121A1 (published 2003) proposed a recombinant adeno-associated viral vector for the gene therapy of WD that possesses a shortened metal-sensitive promoter (metallothionein-I promoter) to produce copper or zinc inducible expression of ATP7B transgene. Nevertheless, this document does not provide, nor has been later disclosed, any information regarding the therapeutic efficiency and performance of the vector.

[0012] On the other hand, several lentiviral vectors carrying wild type ATP7B have been tested in animal models of WD. Merle et al. [Scan. J. Gastroenterol. 2006; 41: 974-982] reported systemic gene therapy in LEC rats with lentiviral vectors expressing ATP7B under the control of a phosphoglycerokinase promoter. Twenty-four weeks after gene transfer

liver copper content was lowered significantly and liver histology improved in treated rats compared to untreated controls, but the effect was only partial. Serum ceruloplasmin oxidase activity was increased two weeks after gene transfer when compared to controls, however, it declined to lower levels 24 weeks after treatment. More recently, Roybal et al. [Gene Therapy 2012; 19: 1085-1094] have reported early gestational gene transfer in ATP7B$^{-/-}$ mice with a lentivirus carrying human ATP7B under transcriptional control of a liver-specific promoter which contained element of apolipoprotein E and alpha-1 antitrypsin. In utero administration of the vector provided a decrease in liver copper levels, preservation of normal hepatic histology, restoration of copper incorporation into ceruloplasmin and improved cholesterol biosynthesis. However, the efficiency of the treatment was very variable from mice to mice and declined with time and never reached full correction of the different pathologically altered parameters. Huster et al [J Biol Chem vol. 278, no. 34, 22 August 2003, pages 32212-32218] reported studies related to the roles of the N-terminal Copper-binding sites in the in vitro catalytic activity of the Wilson's Disease Protein.

SUMMARY

[0013] The invention is defined in the claims. The inventors have engineered and tested several viral vectors carrying transgenes encoding different truncated forms of the enzyme ATP7B: e.g. vector AAV2/8-AAT-ATP7B(d223-366), encoding ATP7B(d223-366) [ATP7B-T1]; and vector AAV2/8-AAT-ATP7B(d57-486), encoding ATP7B(d57-486) [ATP7B-T2]. When administered to ATP7B knockout mice (a recognized animal model of Wilson's disease), the AAV vector carrying ATP7B-T2 corrected main Wilson's disease pathological characteristics for at least 24 weeks after treatment while the AAV vector carrying ATP7B-T1 had only a partial effect. Cu excretion (Cu urine content), and liver Cu content were significantly reduced in Wilson's disease mice treated with the AAV2/8-AAT-ATP7B(d57-486) vector, and ceruloplasmin activity was significantly restored. On the other hand, the administration of the vector resulted in the normalization of serum transaminases levels and of liver histology, together with a marked reduction of the inflammatory infiltrate, biliary duct proliferation and fibrosis.

[0014] Furthermore, a dose of 1 x 10^{10} vg / mouse of the AAV2/8-AAT-wtATP7B vector was shown to be a "suboptimal dose" for the wt construct both for the obtaining of a normalization of the serum ceruloplasmin activity and a reduction of Cu accumulation in the liver (Figures 10A and 11A); whereas the vector carrying the truncated form was shown to provide statistically significant therapeutic effects (vs untreated) at said suboptimal dose (Figure 10B and 11B). Moreover, the observed difference in activity between the full length ATP7B and T2 constructs at a dose of 1 x 10^{10} vg/ mouse was also shown to be statistically significant for these two therapeutic effects (Fig. 12 and Fig. 14).

[0015] These observations indicated that both a nucleic acid construct encoding the truncated form ATP7B(d57-486) and vectors that carry it, in particular AAV vectors, enable to overcome the most relevant pathological effects of an accumulation of copper linked to a deficiency or dysfunction of ATP7B and thus can be very suitable for gene therapy applied to a condition caused by a deficiency or dysfunction of Copper-transporting ATPase 2, such as Wilson's disease, or a disease and/or condition associated with a decrease of ATP7B-dependent lysosomal exocytosis and copper accumulation. Moreover, unexpectedly the truncated form ATP7B(d57-486) and vectors that carry it, were shown to achieve normalization of some of these pathological manifestations of the disease at dosages where the full length ATP7B protein and vectors encoding the same proved to be less effective.

[0016] Therefore, in a first aspect the disclosure relates to a nucleic acid construct (hereinafter also referred as "nucleic acid construct of the disclosure"), that comprises: a) a nucleotide sequence of an eukaryotic promoter; b) a nucleotide sequence encoding a truncated Copper-transporting ATPase 2 (ATP7B) in which the N-terminal heavy metal associated sites HMA 1, HMA 2, HMA 3, and HMA 4 are totally deleted and HMA 5 and HMA 6 remain undeleted; and c) a polyadenylation signal sequence.

[0017] In another aspect, the disclosure relates to an expression vector (hereinafter also referred as "expression vector of the disclosure"), that comprises a nucleic acid construct of the disclosure.

[0018] In another aspect, the disclosure relates to a host cell comprising a nucleic acid construct or an expression vector of the disclosure.

[0019] In another aspect, the disclosure relates to a viral particle (hereinafter also referred as "viral particle of the disclosure"), that comprises a nucleic construct or an expression vector of the disclosure. Preferably, the nucleic acid construct constitutes the genomic sequence of the viral vector.

[0020] In another aspect, the disclosure relates to a pharmaceutical composition that comprises a product of the disclosure, i.e. a product that comprises a nucleic acid construct of the disclosure, and a pharmaceutically acceptable carrier. The term "product of the disclosure" as used herein refers to and indistinctively covers any of: a) the nucleic acid construct of the disclosure; b) the expression vector of the disclosure, c) the host cell of the disclosure and d) the viral particle of the disclosure.

[0021] In another aspect, the disclosure further relates to a kit comprising a nucleic acid construct, vector, host cell, viral particle or pharmaceutical composition of the disclosure in one or more containers.

[0022] In another aspect, the disclosure relates to a product of the disclosure for use in medicine (as a medicament

or medicinal composition). This use in medicine includes the treatment of a condition caused by a deficiency or dysfunction of Copper-transporting ATPase 2. Said another way, the disclosure relates to: the use of a product of the disclosure in the preparation of a medicament for use in the treatment of a condition caused by a deficiency or dysfunction of Copper-transporting ATPase 2; and to a method for the treatment of a condition caused by a deficiency or dysfunction of Copper-transporting ATPase 2 in a subject or a patient, that comprises administering to the subject or patient a therapeutically effective amount of a product of the disclosure. In a more particular aspect, the product of the disclosure is used for the treatment of Wilson's disease.

[0023] In another aspect, the disclosure further relates to a pharmaceutical composition comprising a product of the disclosure as described above, for the proposed uses in medicine and therapeutic methods herein described.

[0024] In an even further aspect, the disclosure relates to a process of producing viral particles of the disclosure comprising the steps of:

a) culturing a host cell containing a nucleic acid construct or expression vector of the disclosure in a culture medium; and

b) harvesting the viral particles in the cell culture supernatant and/or inside the cells.

[0025] In a related aspect, the present disclosure relates to the use of the nucleic acid construct of the disclosure or the expression vector of the disclosure for the production of viral particles.

BRIEF DESCRIPTION OF THE DRAWINGS

[0026]

Figure 1: Schematic representation of the nucleic acid construct of vector AAV2/8-AAT-wtATP7B which carries human ATP7B; vector AAV2/8-AAT-ATP7B(d223-366) which carries truncated form ATP7B(d223-366) [ATP7B-T1]; and vector AAV2/8-AAT-ATP7B(d57-486) which carries truncated ATP7B(d57-486) [ATP7B-T2]. The elements of the constructs are: a) alpha-1-antitrypsin gene promoter (AAT); b) nucleotide sequence encoding respectively human ATP7B, ATP7B-T1, or ATP7B-T2; c) the polyadenylation signal (pA), and flanking the vector genome d) the inverted terminal repeat sequences of AAV2 (ITRs).

Figure 2: Serum alanine transaminase (ALT) levels in wild type male mice [WT], ATP7B deficient male mice [Wilson's Disease mice, WD], and WD male mice treated with the vectors AAV2/8-AAT-wtATP7B [WD AAV-ATP7B], AAV2/8-AAT-ATP7B(d223-366) [WD AAV-T1]; or AAV2/8-AAT-ATP7B(d57-486) [WD AAV-T2]. A vector dose of 3 $x10^{10}$ vg / mouse was administered when the animals were 6 weeks old. ALT levels were measured 4, 9, 14 and 24 weeks after treatment [Weeks] and is expressed as IU / L (IU: international units). ns: no significant; *: $p<0.05$, **: $p<0.01$; ***: $p<0.001$ [Mann-Whitney unpaired test].

Figure 3: Total urine copper content in wild type male mice [WT], Wilson's Disease male mice [WD], and WD male mice treated with the vectors AAV2/8-AAT-wtATP7B [WD AAV-ATP7B], AAV2/8-AAT-ATP7B(d223-366) [WD AAV-T1]; or AAV2/8-AAT-ATP7B(d57-486) [WD ATP7B-T2]. Vector dose: 3 $x10^{10}$ vg / mouse. Copper content was measured 4, 9, 14 and 24 weeks after treatment [Weeks] in 24 hours urine and expressed as nanograms of Cu (ngr / 24h).

Figure 4: Serum ceruloplasmin activity in wild type male mice [WT], Wilson's Disease male mice [WD], and WD male mice treated with the vectors AAV2/8-AAT-wtATP7B [WD AAV-ATP7B], AAV2/8-AAT-ATP7B(d223-366) [WD AAV-T1]; or AAV2/8-AAT-ATP7B(d57-486) [WD ATP7B-T2]. Vector dose: 3 x 10^{10} vg / mouse. Ceruloplasmin activity was measured 4 weeks after treatment and expressed as the absorbance measured at 570 nm of wave-length [Abs (570 nm)]. ns: no significant; *: $p<0.05$, **: $p<0.01$; ***: $p<0.001$ [Mann-Whitney unpaired test].

Figure 5. Liver Cu content in wild type male mice [WT], Wilson's Disease male mice [WD], and WD male mice treated with the vectors AAV2/8-AAT-wtATP7B [WD AAV-ATP7B], AAV2/8-AAT-ATP7B(d223-366) [WD AAV-T1]; or AAV2/8-AAT-ATP7B(d57-486) [WD ATP7B-T2]. Vector dose: 3 $x10^{10}$ vg / mouse. Cupper content was determined after sacrificing the animals 24 weeks after treatment by atomic absorption spectroscopy; and expressed as μg / g (Cu μg / g of dry liver tissue). ns: no significant; *: $p<0.05$, **: $p<0.01$; ***: $p<0.001$ [Mann-Whitney unpaired test].

Figure 6: Histological images of livers of wild type animals male mice [WT], Wilson's Disease male mice [WD], and WD male mice treated with the vectors AAV2/8-AAT-wtATP7B [WD AAV-ATP7B], AAV2/8-AAT-ATP7B(d223-366) [WD AAV-T1]; or AAV2/8-AAT-ATP7B(d57-486) [WD ATP7B-T2]. Vector dose: 3 $x10^{10}$ vg / mouse. Images were taken after sacrificing the animals (30 weeks of age). A: Images of liver sections stained with hematoxylin and eosin. B: Images of histological samples stained by Timm's sulphide silver technique for detection of copper deposits.

Figure 7: Analysis of liver inflammation, Bile duct proliferation and fibrosis. Images of livers of wild type male mice [WT], Wilson's Disease male mice [WD], and WD male mice treated with the vectors AAV2/8-AAT-wtATP7B [WD AAV-ATP7B], AAV2/8-AAT-ATP7B(d223-366) [WD AAV-T1]; or AAV2/8-AAT-ATP7B(d57-486) [WD ATP7B-T2].

Vector dose: 3 x10^{10} vg / mouse. Analysis was performed after sacrificing the animals (30 weeks of age). CD45: Images of liver sections immunostained with anti-CD45 for detecting liver inflammatory infiltrates. PANCK: Images of liver sections immunostained with anti-PANCK for detecting bile duct proliferation. SR: Images of liver sections stained with Sirius red for detecting fibrosis.

Figure 8: Serum alanine transaminase (ALT) levels in wild type female mice [WT], WD female mice [WD], and WD female mice treated with the vector AAV2/8-AAT-ATP7B(d57-486) [WD AAV-T2]. Different groups of 6 weeks old WD female mice were administered different doses of the vectors (respectively 1 x 10^{10}, 3 x 10^{10}, 1 x 10^{11} vg/ mouse). ALT levels were measured 4, 9, 14 and 24 weeks after treatment [Weeks] and is expressed as IU / L. ns: no significant; *: p<0.05, **: p<0.01; ***: p<0.001 [Mann-Whitney unpaired test].

Figure 9: Urinary Cu Content levels in wild type female mice [WT], WD female mice [WD], and WD female mice treated with the vector AAV2/8-AAT-ATP7B(d57-486) [WD AAV-T2]. Different groups of 6 weeks old WD female mice were administered different doses of the vector (respectively 1 x 10^{10}, 3 x 10^{10}, 1 x 10^{11} vg / mouse). Urinary copper levels were measured 4, 9, 14 and 24 weeks after treatment [Weeks] and is expressed as ngr of Cu in 24 hours urine (ngr / 24h). ns: no significant; *: p<0.05, **: p<0.01; ***: p<0.001 [Mann-Whitney unpaired test].

Figure 10: Ceruloplasmin activity in serum was measured in wild type female mice [WT], WD female mice [WD], and WD female mice treated with the vector AAV2/8-AAT-ATP7B(d57-486) [WD+AAV-T2] or the vector AAV2/8-AAT-wtATP7B [WD+AAV-ATP7B]. For each experimental group, different groups of 6 weeks old WD female mice were administered different doses of the vector (respectively 1 x 10^{10}, 3 x 10^{10}, 1 x 10^{11} vg / mouse). Ceruloplasmin activity was measured 4 weeks after treatment and is expressed as the absorbance measured at 570 nm of wave-length [Abs (570 nm)]. ns: no significant; *: p<0.05, **: p<0.01; ***: p<0.001 [Mann-Whitney unpaired test].

Figure 11: Liver Cu Content was measured in wild type female mice [WT], WD mice [WD], and WD female mice treated with the vector AAV2/8-AAT-wtATP7B [WD AAV ATP7B] or the vector AAV2/8-AAT-ATP7B(d57-486) [WD AAV T2]. For each experimental group, different groups of 6 weeks old WD female mice were administered different doses of the vector (respectively 1 x 10^{10}, 3 x 10^{10}, 1 x 10^{11} vg / mouse). Copper concentration was measured 24 weeks after treatment and is expressed as μg / g of dry tissue ns: no significant; *: p<0.05, **: p<0.01; ***: p<0.001 [Mann-Whitney unpaired test].

Figure 12: Liver Cu Content in wild type male mice [WT, n=15], WD male mice [WD; n=25], and WD male mice treated with the vector AAV2/8-AAT-wtATP7B [WD AAV ATP7B; n=7] or the vector AAV2/8-AAT-ATP7B(d57-486) [WD AAV T2; n=7]. For each experimental group, WD mice were administered a suboptimal dose of the vector (1 x 10^{10} vg / mouse) when the animals were 6 weeks old. Copper concentration was measured 24 weeks after treatment and is expressed as μg / g of dry tissue ns: no significant; *: p<0.05, **: p<0.01; ***: p<0.001 [Mann-Whitney unpaired test].

Figure 13: Liver Cu Content in wild type male mice [WT, n=15], WD male mice [WD; n=25], and WD male mice treated with the vector AAV2/8-AAT-ATP7B(d57-486) [WD AAV T2; n=13] or the vector AAV2/8-AAT-coATP7B(d57-486) [WD AAV coT2; n=4]. For each experimental group, 6 weeks old WD male mice were administered a suboptimal dose of the vector (1 x 10^{10} vg / mouse). Copper concentration was measured 24 weeks after treatment and is expressed as μg / g of dry tissue ns: no significant; *: p<0.05, **: p<0.01; ***: p<0.001 [Mann-Whitney unpaired test].

Figure 14: Ceruloplasmin activity in serum of wild type male mice [WT, n=15], WD male mice [WD; n=25], and WD male mice groups treated with one of the vectors AAV2/8-AAT-wtATP7B [WD AAV ATP7B; n=10], AAV2/8-AAT-coATP7B [WD AAV coATP7B; n=8], AAV2/8-AAT-ATP7B(d57-486) [WD AAV T2; n=13] and AAV2/8-AAT-coATP7B(d57-486) [WD AAV coT2; n=4]. For each experimental group, 6 weeks old WD male mice were administered a suboptimal dose of the vector (1 x 10^{10} vg / mouse). Oxidase activity of ceruloplasmin was measured 4 weeks after treatment and is expressed as the absorbance measured at 570 nm of wave-length [Abs (570 nm)]. ns: no significant; *: p<0.05, **: p<0.01; ***: p<0.001 [Mann-Whitney unpaired test].

DETAILED DESCRIPTION

[0027] The invention is as defined in the claims.

[0028] All terms as used herein in this application, unless otherwise stated, shall be understood in their ordinary meaning as known in the art. Other more specific definitions for certain terms as used in the present application are as set forth below and are intended to apply uniformly through-out the specification and claims unless an otherwise expressly set out definition provides a broader definition.

[0029] The terms "nucleic acid sequence" and "nucleotide sequence" may be used interchangeably to refer to any molecule composed of or comprising monomeric nucleotides. A nucleic acid may be an oligonucleotide or a polynucleotide. A nucleotide sequence may be a DNA or RNA. A nucleotide sequence may be chemically modified or artificial. Nucleotide sequences include peptide nucleic acids (PNA), morpholinos and locked nucleic acids (LNA), as well as glycol nucleic acids (GNA) and threose nucleic acid (TNA). Each of these sequences is distinguished from naturally-

occurring DNA or RNA by changes to the backbone of the molecule. Also, phosphorothioate nucleotides may be used. Other deoxynucleotide analogs include methylphosphonates, phosphoramidates, phosphorodithioates, N3'P5'-phosphoramidates and oligoribonucleotide phosphorothioates and their 2'-0-allyl analogs and 2'-0-methylribonucleotide methylphosphonates which may be used in a nucleotide of the disclosure.

[0030] The term "nucleic acid construct" as used herein refers to a man-made nucleic acid molecule resulting from the use of recombinant DNA technology. A nucleic acid construct is a nucleic acid molecule, either single- or double-stranded, which has been modified to contain segments of nucleic acids sequences, which are combined and juxtaposed in a manner, which would not otherwise exist in nature. A nucleic acid construct usually is a "vector", i.e. a nucleic acid molecule which is used to deliver exogenously created DNA into a host cell.

[0031] The term "expression vector" or "vector" as used herein refers to a recombinant nucleotide sequence that is capable of effecting expression of a gene (transgene) in host cells or host organisms compatible with such sequences. Together with the transgene, expression vectors typically include at least suitable transcription regulatory sequences and optionally, 3' transcription termination signals. Additional factors necessary or helpful in effecting expression may also be present, such as expression enhancer elements able to respond to a precise inductive signal (endogenous or chimeric transcription factors) or specific for certain cells, organs or tissues.

[0032] The term "subject" or "patient" as used herein, refers to mammals. Mammalian species that can benefit from the disclosed methods of treatment include, but are not limited to, humans, non-human primates such as apes; chimpanzees; monkeys, and orangutans, domesticated animals, including dogs and cats, as well as livestock such as horses, cattle, pigs, sheep, and goats, or other mammalian species including, without limitation, mice, rats, guinea pigs, rabbits, hamsters, and the like.

[0033] The term "packaging cells" as used herein, refers to a cell or cell line which may be transfected with a helper vector or virus or a DNA construct, and provides in *trans* all the missing functions which are required for the complete replication and packaging of a viral vector. Typically, the packaging cells express in a constitutive or inducible manner one or more of said missing viral functions.

A NUCLEIC ACID CONSTRUCT OF THE DISCLOSURE

Nucleotide sequence of eukaryotic promoter

[0034] As used herein, the term "eukaryotic promoter" refers to a DNA sequence region that initiates transcription of a particular gene, or one or more coding sequences, in eukaryotic cells. A promoter can work in concert with other regulatory regions or elements to direct the level of transcription of the gene or coding sequence/s. These regulatory elements include, without limitation, transcription factor binding sites, repressor and activator protein binding sites, and any other sequences of nucleotides known to one of skill in the art to act directly or indirectly to regulate the amount of transcription from the promoter, including e.g. attenuators, enhancers, and silencers. The promoter is located near the transcription start site of the gene or coding sequence to which is operably linked, on the same strand and upstream of the DNA sequence (towards the 5' region of the sense strand). A promoter can be about 100-1000 base pairs long. Positions in a promoter are designated relative to the transcriptional start site for a particular gene (i.e., positions upstream are negative numbers counting back from -1, for example -100 is a position 100 base pairs upstream).

[0035] The term "core promoter" or "minimal promoter" refers to the minimal portion of a promoter sequence required to properly initiate transcription. It includes the transcription start site (TSS) and elements directly upstream; a binding site for RNA polymerase (RNA polymerase II); and general transcription factors binding sites. Commonly a promoter also comprises a proximal promoter sequence (upstream of the core promoter), that contains other primary regulatory elements (such as enhancers, silencers, boundary elements/insulators); and a distal promoter sequence (downstream of core promoter), that may contain additional regulatory elements, normally with a weaker influence on the level of transcription of the gene.

[0036] According to the disclosure, the eukaryotic promoter sequence is operably linked to the nucleotide sequence encoding the truncated Copper-transporting ATPase 2. As used herein, the term "operably linked" refers to a linkage of polynucleotide (or polypeptide) elements in a functional relationship. A nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For instance, a promoter or transcription regulatory sequence is operably linked to a coding sequence if it affects the transcription of the coding sequence. Operably linked means that the DNA sequences being linked are typically contiguous; where it is necessary to join two protein encoding regions, they are contiguous and in reading frame.

[0037] According to the disclosure, the eukaryotic promoter sequence of the nucleic acid construct comprises at least the core promoter and, optionally other regulatory regions or elements of the same gene or of different genes (i.e. hybrid or chimeric promoters).

[0038] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the eukaryotic promoter is a constitutive promoter, a tissue specific promoter, or an inducible promoter.

[0039] As used herein, a "constitutive" promoter is a promoter that is active in most eukaryotic tissues under most physiological and developmental conditions.

[0040] A "tissue specific promoter" is a promoter only active in specific types of tissues or cells. That is to say a tissue specific promoter, in the context of this disclosure, is one which is more active in one or several particular tissues (for example two, three or four) than in other tissues (i.e. the promoter is capable of driving a higher expression of the coding sequence to which it is operably linked in the tissue(s) for which it is specific than in the others). Typically, the gene down-stream of a "tissue specific" promoter is active to a much higher degree in the tissue(s) for which the promoter is specific than in any other tissue(s). In this case, there may be little or substantially no activity of the promoter in any tissue other than the one(s) for which it is specific.

[0041] An "inducible" promoter is a promoter that is physiologically or developmentally regulated, e.g. by the application of a chemical inducer.

[0042] Many promoters are known in the art [Sambrook and Russell (Molecular Cloning: a Laboratory Manual; Third Edition; 2001 Cold Spring Harbor Laboratory Press); and Green and Sambrook (Molecular Cloning: a Laboratory Manual, cuarta edición, 2012 Cold Spring Harbor Laboratory Press)].

[0043] Suitable tissue specific promoters may be found in the Tissue-Specific Promoter Database, TiProD (Nucleic Acids Research 2006; J4: D104-D107).

[0044] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the eukaryotic promoter is a liver specific promoter. In the context of this disclosure, a "liver specific promoter" is a promoter which is more active in the liver than in any other tissue of the body. Typically, the activity of a liver specific promoter will be considerably greater in the liver than in other tissues. For example, such a promoter may be at least 2, at least 3, at least 4, at least 5 or at least 10 times more active (for example as determined by its ability to drive the expression in a given tissue in comparison to its ability to drive the expression in other cells or tissues). Accordingly, a liver specific promoter allows an active expression in the liver of the gene linked to it and prevents its expression in other cells or tissues.

[0045] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the eukaryotic promoter is a nucleotide sequence of the α1-antitrypsin gene promoter (AAT), or a chimeric promoter sequence EalbPa1AT that comprises an α1-antitrypsin gene promoter sequence (AAT or Pa1AT) combined with an albumin gene enhancer element (Ealb). Both promoter sequences have properties of liver specific promoters.

[0046] In a particular embodiment, optionally in combination with one or more features of the various embodiments described above or below, the eukaryotic promoter sequence is the sequence delimited by bases 156..460 of SEQ.ID.NO.1 (AAT); or SEQ.ID.NO.5 (EalbPa1AT).

Truncated Copper-transporting ATPase 2 (ATP7B)

[0047] Copper-transporting ATPase 2 (ATP7B) is a P-type cation transport ATPase that functions exporting copper out of the cells.

[0048] The gene that encodes human enzyme is located at chromosome 13 (chromosome location 13q14.3; gene name ATP7B). Information on human ATP7B polypeptide (amino acid sequences, structure, domains and other features) is for example available at Uniprot with Accession number: P35670 (http://www.uniprot.org/uniprot/P35670; Entry version 168 (03 Sep 2014), Sequence version 4 (16 Jun 2009)). Information on the ATP7B gene encoding this enzyme is available at Entrez with accession number Gene ID: 540 (http://www.ncbi.nlm.nih.gov/gene/540; updated on 19-Sep-2014). 4 isoforms produced by alternative splicing have been described for ATP7B; isoform 1 (identifier P35670-1, 1465 amino acids long) is chosen as the canonical sequence.

[0049] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the nucleotide acid construct of the disclosure comprises a nucleotide sequence that encodes a truncated form of a human ATP7B, preferably a human ATP7B whose amino acid sequence is the canonical sequence (SEQ.ID.NO.2), herein also referred to as wtATP7B.

[0050] Several conserved motifs are present in ATP7B that are characteristic for the P-type ATPase protein family. These motifs are required for ATP catalysis and include the nucleotide binding domain (N-domain), the phosphorylation domain (P-domain) and the actuator domain (A-domain). Highly conserved signature residues are present in these motifs; SEHPL in the N-domain, DKTG in the P-domain, and TGE in the A-domain. The amino terminal tail of human ATP7B contains "six metal binding sites" (MBS), also indistinctively named as "heavy metal associated (HMA)" sites or domains, each containing the core sequence MxCxxC. These HMA bind Cu(I) in a stoichiometry of one atom of Cu(I) per HMA. These amino-terminal HMAs of ATP7B are required for several aspects of its function, including copper translocation, incorporation of copper in cuproenzymes, ATPase activity, localization and trafficking, and protein-protein interactions. The HMA sites are identified starting at the amino end, as domains HMA 1 (amino acids 59 - 125 in the canonical sequence), HMA 2 (amino acids 144 - 210), HMA 3 (258 - 327), HMA 4 (360 - 426), HMA 5 (489 - 555), and

HMA 6 (565 - 631).

[0051] According to the disclosure, optionally in combination with one or more features of the various aspects described above or below, the nucleic acid construct comprises a nucleotide sequence that encodes a truncated ATP7B in which the N-terminal heavy metal associated sites HMA 1, HMA 2, HMA 3, and HMA 4 are totally or partially deleted.

[0052] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the nucleotide sequence that encodes truncated ATP7B keeps the 56 amino acids of N-terminal signal sequence of ATP7B.

[0053] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the deletion in truncated ATP7B comprises amino acids 57 to 486 of the canonical sequence.

[0054] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the nucleotide sequence encodes a truncated ATP7B whose amino acids sequence is SEQ.ID.NO.7.

[0055] Because of the codons redundancy, there are numerous nucleotide sequences that can be generated encoding ATP7B polypeptides with same amino acids sequence.

[0056] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the nucleotide sequence encoding the truncated Copper-transporting ATPase 2 is the coding sequence CDS of SEQ.ID.NO.6, bases 473..3580.

[0057] In another embodiment, optionally in combination with one or more features of the various embodiments described above or below, the nucleotide sequence encoding the truncated Copper-transporting ATPase 2 is SEQ.ID.NO.8, a sequence with an optimized codon usage bias for the human cells.

[0058] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the nucleotide sequence encoding the truncated Copper-transporting ATPase 2 is a sequence wherein at least 827, at least 879, at least 931, or at least 983 of the codons encoding truncated Copper-transporting ATPase 2 are identical to the codons of coding sequence SEQ.ID.NO.8.

Polyadenylation signal sequence

[0059] As used herein, the term "polyadenylation signal" or "poly(A) signal" refers to a specific recognition sequence within 3' untranslated region (3' UTR) of the gene, which is transcribed into precursor mRNA molecule and guides the termination of the gene transcription. Poly(A) signal acts as a signal for the endonucleolytic cleavage of the newly formed precursor mRNA at its 3'-end, and for the addition to this 3'-end of a RNA stretch consisting only of adenine bases (polyadenylation process; poly(A) tail). Poly(A) tail is important for the nuclear export, translation, and stability of mRNA. In the context of the disclosure, the polyadenylation signal is a recognition sequence that can direct polyadenylation of mammalian genes and/or viral genes, in mammalian cells.

[0060] Poly(A) signals typically consist of a) a consensus sequence AAUAAA, which has been shown to be required for both 3'-end cleavage and polyadenylation of premessenger RNA (pre-mRNA) as well as to promote downstream transcriptional termination, and b) additional elements upstream and downstream of AAUAAA that control the efficiency of utilization of AAUAAA as a poly(A) signal. There is considerable variability in these motifs in mammalian genes.

[0061] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the polyadenylation signal sequence of the nucleic acid construct of the disclosure is a polyadenylation signal sequence of a mammalian gene or a viral gene. Suitable polyadenylation signals include, among others, a SV40 early polyadenylation signal, a SV40 late polyadenylation signal, a HSV thymidine kinase polyadenylation signal, a protamine gene polyadenylation signal, an adenovirus 5 Elb polyadenylation signal, a growth hormone polydenylation signal, a PBGD polyadenylation signal, *in silico* designed polyadenylation signal (synthetic) and the like.

[0062] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the polyadenylation signal sequence of the nucleic acid construct is a synthetic poly(A) signal sequence which is also capable of directing and effecting the endonucleolytic cleavage and polyadenylation of the precursor mRNA resulting from the transcription of nucleotide sequence coding for truncated ATP7B.

[0063] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the polyadenylation signal sequence of the nucleic acid construct is the synthetic poly(A) signal sequence delimited by bases 4877..4932 of SEQ.ID.NO.1.

Other nucleotide elements

[0064] In one aspect, the nucleic acid construct of the disclosure constitutes the recombinant genome of an expression vector for gene therapy, the expression vector of the disclosure; and more particularly of a viral vector for gene therapy.

[0065] Thus, in one aspect, optionally in combination with one or more features of the various aspects described above or below, the nucleic acid construct of the disclosure further comprises a 5'ITR and a 3'ITR of a virus.

[0066] As used herein the term "inverted terminal repeat (ITR)" refers to a nucleotide sequence located at the 5'-end

(5'ITR) and a nucleotide sequence located at the 3'-end (3'ITR) of a virus, that contain palindromic sequences and that can fold over to form T-shaped hairpin structures that function as primers during initiation of DNA replication. They are also needed for viral genome integration into host genome; for the rescue from the host genome; and for the encapsidation of viral nucleic acid into mature virions. The ITRs are required in *cis* for the vector genome replication and its packaging into the viral particles.

[0067] In one aspect, the nucleic acid construct comprises a 5'ITR, a ψ packaging signal, and a 3'ITR of a virus. "ψ packaging signal" is a cis-acting nucleotide sequence of the virus genome, which in some viruses (e.g. adenoviruses, lentiviruses ...) is essential for the process of packaging the virus genome into the viral capsid during replication.

[0068] In one aspect, optionally in combination with one or more features of the various aspects described above or below, the nucleic acid construct comprises a 5'ITR and a 3'ITR of a virus selected from the group consisting of parvoviruses (in particular adeno-associated viruses), adenoviruses, alphaviruses, retroviruses (in particular gamma retroviruses, and lentiviruses), herpesviruses, and SV40; in a preferred aspect the virus is an adeno-associated virus (AAV), an adenovirus (Ad), or a lentivirus.

[0069] In the invention, optionally in combination with one or more features of the various embodiments described above or below, the nucleic acid construct comprises a 5'ITR and a 3'ITR of an AAV, as defined in the claims.

[0070] The AAV genome is composed of a linear, single-stranded DNA molecule which contains 4681 bases (Berns and Bohenzky, (1987) Advances in Virus Research (Academic Press, Inc.) 32:243-307). The genome includes inverted terminal repeats (ITRs) at each end which function in *cis* as origins of DNA replication and as packaging signals for the virus. The ITRs are approximately 145 bp in length. The internal non-repeated portion of the genome includes two large open reading frames, known as the AAV rep and cap genes, respectively. These genes code for the viral proteins involved in replication and packaging of the virion. In particular, at least four viral proteins are synthesized from the AAV rep gene, Rep 78, Rep 68, Rep 52 and Rep 40, named according to their apparent molecular weight. The AAV cap gene encodes at least three proteins, VP1, VP2 and VP3. For a detailed description of the AAV genome, see, e.g., Muzyczka, N. (1992) Current Topics in Microbiol. and Immunol. 158:97-129.

[0071] The construction of recombinant AAV virions is generally known in the art and has been described for instance in US 5,173,414 and US5,139,941; WO 92/01070, WO 93/03769, (Lebkowski et al. (1988) Molec. Cell. Biol. 8:3988-3996; Vincent et al. (1990) Vaccines 90 (Cold Spring Harbor Laboratory Press); Carter, B. J. (1992) Current Opinion in Biotechnology 3:533-539; Muzyczka, N. (1992) Current Topics in Microbiol. and Immunol. 158:97-129; and Kotin, R. M. (1994) Human Gene Therapy 5:793-801.

[0072] The disclosure may be carried out by using ITRs of any AAV serotype, including AAV1, AAV2, AAV3 (including types 3A and 3B), AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, avian AAV, bovine AAV, canine AAV, equine AAV, ovine AAV, and any other AAV serotype now known or later discovered.

[0073] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the nucleic acid construct comprises a 5'ITR and a 3'ITR of an AAV of a serotype selected from the group consisting of an AAV1, an AAV2, and an AAV4. In a preferred embodiment the nucleic acid construct comprises the ITR sequences delimited by bases 1..141, and bases 4968..5107 of SEQ.ID.NO.1, that are the ITRs sequences of an AAV2.

[0074] The ITRs are the only AAV viral elements which are required in *cis* for the AAV genome replication and its packaging into the viral particles.

[0075] In one aspect, optionally in combination with one or more features of the various embodiments described above or below, the nucleic acid construct comprises a 5'ITR, a ψ packaging signal, and a 3'ITR of an adenovirus of any of the serotypes within any of the classification sub-groups (A-F). In a particular aspect, optionally in combination with one or more features of the various aspects described above or below, these 5'ITR, ψ signal, and 3'ITR sequences come from a sub-group C adenovirus, more preferably from an adenovirus of serotype 2 (Ad2) or serotype 5 (Ad5).

[0076] On the other hand, in other aspects the disclosure can be carried out by using synthetic 5'ITR and/or 3'ITR; and also by using a 5'ITR and a 3'ITR which come from viruses of different serotype.

[0077] All other viral genes required for viral vector replication can be provided in *trans* within the virus-producing cells (packaging cells) as described below. Therefore, their inclusion in the nucleic acid construct of a viral vector genome according to the disclosure is optional.

[0078] In the invention, optionally in combination with one or more features of the various embodiments described above or below, the expression vector is an AAV vector, as defined in the claims.

[0079] In a particular embodiment, the nucleic acid construct of the disclosure constitutes an AAV vector selected from the group of combinations consisting of

a) a vector that comprises a 5'ITR and a 3'ITR nucleotide sequences of an AAV2, an AAT promoter sequence, and a nucleotide sequence encoding truncated human ATP7B(d57-486);
b) a vector that comprises a 5'ITR and a 3'ITR nucleotide sequences of an AAV2, an AAT promoter sequence, and the codon optimized nucleotide sequence SEQ.ID.NO.8 encoding truncated human ATP7B(d57-486);

c) a vector that comprises a 5'ITR and a 3'ITR nucleotide sequences of an AAV2, an EalbPa1AT hybrid promoter sequence, and a nucleotide sequence encoding truncated human ATP7B(d57-486); and
d) a vector that comprises a 5'ITR and a 3'ITR nucleotide sequences of an AAV2, an EalbPa1AT hybrid promoter sequence, and a codon optimized nucleotide sequence SEQ.ID.NO.8 encoding truncated human ATP7B(d57-486).

[0080] Each of these AAV vector embodiments also includes a polyadenylation signal sequence, such as synthetic poly(A) signal sequence of SEQ.ID.NO.1 or any other suitable poly(A) signal; together or not with other optional nucleotide elements.

[0081] In another aspect, optionally in combination with one or more features of the various aspects described above or below, the expression vector is an adenoviral vector. This adenoviral vector according to the disclosure can be, in particular, a first-, second-, or third-generation adenovirus [see Adenovirus. Methods and Protocols. Chillon M. and Bosch A. (Eds); third Edition; 2014 Springer], or any other adenoviral vector system already known or later described.

[0082] In a particular aspect, optionally in combination with one or more features of the various aspects described above or below, the viral vector of the disclosure is a "third generation adenovirus", which may also be referred to as "gutless adenovirus", "helper-dependent adenovirus (HD-Ad)", or "high capacity adenovirus (HC-Ad)". A third generation adenovirus has all viral coding regions removed (gutless); it depends on a helper adenovirus to replicate (helper-dependent); and it can carry and deliver into the host cell up to 36 Kbp inserts of foreign genetic material (high-capacity). A gutless adenovirus keeps the inverted terminal repeats ITRs (5' and 3') and the packaging signal (ψ).

[0083] The nucleic acid construct and expression vector of the disclosure herein described can be prepared and obtained by conventional methods known to those skilled in the art: Sambrook and Russell (Molecular Cloning: a Laboratory Manual; Third Edition; 2001 Cold Spring Harbor Laboratory Press); and Green and Sambrook (Molecular Cloning: a Laboratory Manual; Fourth Edition; 2012 Cold Spring Harbor Laboratory Press).

A VIRAL PARTICLE OF THE DISCLOSURE FOR GENE THERAPY

[0084] The terms "viral particle", and "virion" are used herein interchangeably and relate to an infectious and typically replication-defective virus particle comprising the viral genome (i.e. the nucleic acid construct of the expression viral vector) packaged within a capsid and, as the case may be, a lipidic envelope surrounding the capsid.

[0085] In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the virion of the disclosure is a "recombinant AAV virion" or "rAAV virion" obtained by packaging of a nucleic acid construct of an AAV vector according to the disclosure in a protein shell.

[0086] Proteins of the viral capsid of an adeno-associated virus (capsid proteins VP1, VP2, and VP3) are generated from a single viral gene (cap gene). Differences among the capsid protein sequences of the various AAV serotypes result in the use of different cell surface receptors for cell entry. In combination with alternative intracellular processing pathways, this gives rise to distinct tissue tropisms for each AAV serotype.

[0087] In a particular embodiment, a recombinant AAV virion according to the disclosure may be prepared by encapsidating the nucleic acid construct of an AAV vector/genome derived from a particular AAV serotype on a viral particle formed by natural Cap proteins corresponding to an AAV of the same particular serotype. Nevertheless, several techniques have been developed to modify and improve the structural and functional properties of naturally occurring AAV viral particles (Bünning H et al. J Gene Med 2008; 10: 717-733). Thus, in another AAV viral particle according to the disclosure the nucleotide construct of the viral vector flanked by ITR(s) of a given AAV serotype can be packaged, for example, into: a) a viral particle constituted of capsid proteins derived from the same or different AAV serotype [e.g. AAV2 ITRs and AAV5 capsid proteins; AAV2 ITRs and AAV8 capsid proteins; etc]; b) a mosaic viral particle constituted of a mixture of capsid proteins from different AAV serotypes or mutants [e.g. AAV2 ITRs with AAV1 and AAV5 capsid proteins]; c) a chimeric viral particle constituted of capsid proteins that have been truncated by domain swapping between different AAV serotypes or variants [e.g. AAV2 ITRs with AAV5 capsid proteins with AAV3 domains]; or d) a targeted viral particle engineered to display selective binding domains, enabling stringent interaction with target cell specific receptors [e.g. AAV4 ITRs with AAV2 capsid proteins genetically truncated by insertion of a peptide ligand; or AAV2 capsid proteins non-genetically modified by coupling of a peptide ligand to the capsid surface].

[0088] The skilled person will appreciate that the AAV virion according to the disclosure may comprise capsid proteins from any AAV serotype. In one embodiment, optionally in combination with one or more features of the various embodiments described above or below, the viral particle comprises capsid proteins of an AAV. In a particular embodiment, optionally in combination with one or more features of the various embodiments described above or below, the AAV viral particle comprises capsid proteins from a serotype selected from the group consisting of an AAV1, an AAV5, an AAV7, an AAV8, and an AAV9 which are more suitable for delivery to the liver cells (Nathwani et al. Blood 2007; 109: 1414-1421; Kitajima et al. Atherosclerosis 2006; 186:65-73). In a particular embodiment, optionally in combination with one or more features of the various embodiments described above or below, the viral particle comprises a nucleic acid construct of disclosure wherein the 5'ITR and 3'ITR sequences of the nucleic acid construct are of an AAV2 serotype

and the capsid proteins are of an AAV8 serotype.

[0089] In a particular embodiment, optionally in combination with one or more features of the various embodiments described above or below, the AAV viral particle comprises capsid proteins from Anc80, a predicted ancestor of viral AAVs serotypes 1, 2, 8, and 9 that behaves as a highly potent gene therapy vector for targeting liver, muscle and retina (Zinn et al. Cell Reports 2015; 12:1-13). In a more particular embodiment, the viral particle comprises the Anc80L65 VP3 capsid protein (Genbank accession number: KT235804).

[0090] Viral-glycan interactions are critical determinants of host cell invasion. In a particular embodiment, optionally in combination with one or more features of the various embodiments described above or below, the AAV viral particle comprises capsid proteins comprising one or more amino acids substitutions, wherein the substitutions introduce a new glycan binding site into the AAV capsid protein. In a more particular embodiment, the amino acid substitutions are in amino acid 266, amino acids 463-475 and amino acids 499-502 in AAV2 or the corresponding amino acid positions in AAV1, AAV3, AAV4, AAV5, AAV6, AAV7, AAV 8, AAV9, AAV10 or any other AAV serotype, also included Anc80 and Anc80L65.

[0091] The introduced new glycan binding site can be a hexose binding site [e.g. a galactose (Gal), a mannose (Man), a glucose (Glu) or a fucose (fuc) binding site]; a sialic acid (Sia) binding site [e.g. a Sia residue such as is N-acetyl-neuraminic acid (NeuSAc) or N-Glycolylneuraminic acid (NeuSGc)]; or a disaccharide binding site, wherein the disaccharide is a sialic acid linked to galactose, for instance in the form of Sia(alpha2,3)Gal or Sia(alpha2,6)Gal. Detailed guidance to introduce a new binding site from an AAV serotype into a capsid protein of an AAV of another serotype is given on international patent publication WO2014144229 and in Shen et al. (J. Biol. Chem. 2013; 288(40):28814-28823). In a particular embodiment, optionally in combination with one or more features of the various embodiments described above or below, the Gal binding site from AAV9 is introduced into the AAV2 VP3 backbone resulting in a dual glycan-binding AAV strain which is able to use both HS and Gal receptors for cell entry. Preferably, said dual glycan-binding AAV strain is AAV2G9. Shen et al. generated AAV2G9 by substituting amino acid residues directly involved and immediately flanking the Gal recognition site on the AAV9 VP3 capsid protein subunit onto corresponding residues on the AAV2 VP3 subunit coding region (AAV2 VP3 numbering Q464V, A467P, D469N, I470M, R471A, D472V, S474G, Y500F, and S501A).

[0092] In another aspect, optionally in combination with one or more features of the various aspects described above or below, the virion of the disclosure is an adenoviral virion, such as an Ad5 virion. As it is the case for AAV virions, capsid proteins of Ad virions can also be engineered to modify their tropism and cellular targeting properties, alternative adenoviral serotypes can also be employed.

Production of viral particles

[0093] Production of viral particles carrying the nucleic acid construct of the expression viral vector of the disclosure can be performed by means of conventional methods and protocols, which are selected having into account the structural features chosen for the actual embodiment of the nucleic acid construct and viral particle of the vector to be produced.

[0094] Briefly, viral particles can be produced in a specific virus-producing cell (packaging cell) which is transfected with the nucleic acid construct of the vector to be packaged, in the presence of a helper vector or virus or other DNA construct(s).

[0095] Accordingly, in one aspect the disclosure concerns the use of the nucleic acid construct or expression vector of the disclosure for the production of viral particles.

[0096] In a related aspect, the disclosure concerns a process of producing viral particles of the disclosure comprising the steps of:

a) culturing a host cell comprising a nucleic acid construct or expression vector of the disclosure in a culture medium; and

b) harvesting the viral particles from the cell culture supernatant and/or inside the cells.

[0097] Preferably, said host cell is a packaging cell, as described below. Suitable culture media will be known to a person skilled in the art. The ingredients that compose such media may vary depending on the type of cell to be cultured. In addition to nutrient composition, osmolarity and pH are considered important parameters of culture media. The cell growth medium comprises a number of ingredients well known by the person skilled in the art, including amino acids, vitamins, organic and inorganic salts, sources of carbohydrate, lipids, trace elements (CuS04, FeS04, Fe(N03)3, ZnS04...), each ingredient being present in an amount which supports the cultivation of a cell *in vitro* (i.e., survival and growth of cells). Ingredients may also include different auxiliary substances, such as buffer substances (like sodium bicarbonate, Hepes, Tris...), oxidation stabilizers, stabilizers to counteract mechanical stress, protease inhibitors, animal growth factors, plant hydrolyzates, anti-clumping agents, anti-foaming agents. Characteristics and compositions of the cell growth media vary depending on the particular cellular requirements. Examples of commercially available cell growth

media are: MEM (Minimum Essential Medium), BME (Basal Medium Eagle) DMEM (Dulbecco's modified Eagle's Medium), Iscoves DMEM (Iscove's modification of Dulbecco's Medium), GMEM, RPMI 1640, Leibovitz L-15, CHO, McCoy's, Medium 199, HEK293, Ham (Ham's Media) F10 and derivatives, Ham F12, DMEM/F12, etc.

A HOST CELL OF THE DISCLOSURE

[0098] In another aspect, the disclosure relates to a host cell comprising a nucleic acid construct or expression vector of the disclosure.

[0099] The term "host cell" as used herein refers to any cell line that is susceptible to infection by a virus of interest, and amenable to culture *in vitro*.

[0100] The host cell of the disclosure may be used for *ex vivo* gene therapy purposes. In such aspects, the cells are transfected with the nucleic acid construct or viral vector of the disclosure and subsequently transplanted to the patient or subject. Transplanted cells can have an autologous, allogenic or heterologous origin. For clinical use, cell isolation will generally be carried out under Good Manufacturing Practices (GMP) conditions. Before transplantation, cell quality and absence of microbial or other contaminants is typically checked and liver preconditioning, such as with radiation and/or an immunosuppressive treatment, may be carried out. Furthermore, the host cells may be transplanted together with growth factors to stimulate cell proliferation and/or differentiation, such as Hepatocyte Growth Factor (HGF).

[0101] In a particular aspect, the host cell is used for ex vivo gene therapy into the liver. Preferably, said cells are eukaryotic cells such as mammalian cells, these include, but are not limited to, humans, non-human primates such as apes; chimpanzees; monkeys, and orangutans, domesticated animals, including dogs and cats, as well as livestock such as horses, cattle, pigs, sheep, and goats, or other mammalian species including, without limitation, mice, rats, guinea pigs, rabbits, hamsters, and the like. A person skilled in the art will choose the more appropriate cells according to the patient or subject to be transplanted.

[0102] Said host cell may be a cell with self-renewal and pluripotency properties, such as stem cells or induced pluripotent stem cells. Stem cells are preferably mesenchymal stem cells. Mesenchymal stem cells (MSCs) are capable of differentiating into at least one of an osteoblast, a chondrocyte, an adipocyte, or a myocyte and may be isolated from any type of tissue. Generally MSCs will be isolated from bone marrow, adipose tissue, umbilical cord, or peripheral blood. Methods for obtaining thereof are well known to a person skilled in the art. Induced pluripotent stem cells (also known as iPS cells or iPSCs) are a type of pluripotent stem cell that can be generated directly from adult cells. Yamanaka et al. induced iPS cells by transferring the Oct3/4, Sox2, Klf4 and c-Myc genes into mouse and human fibroblasts, and forcing the cells to express the genes (WO 2007/069666). Thomson et al. subsequently produced human iPS cells using Nanog and Lin28 in place of Klf4 and c-Myc (WO 2008/118820).

[0103] Said host cells may also be hepatocytes. Hepatocyte transplantation procedures, including cell isolation and subsequent transplantation into a human or mice recipient is described for instance in Filippi and Dhawan, Ann NY Acad Sci. 2014, 1315 50-55; Yoshida et al., Gastroenterology 1996, 111: 1654-1660; Irani et al. Molecular Therapy 2001, 3:3, 302-309; and Vogel et al. J Inherit Metab Dis 2014, 37:165-176. A method for ex vivo transduction of a viral vector into hepatocytes is described for instance in Merle et al., Scandinavian Journal of Gastroenterology 2006, 41:8, 974-982.

[0104] In another particular aspect, the host cell is a packaging cell. Said cells can be adherent or suspension cells. The packaging cell, and helper vector or DNA constructs provide together in *trans* all the missing functions which are required for the complete replication and packaging of the viral vector.

[0105] Preferably, said packaging cells are eukaryotic cells such as mammalian cells, including simian, human, dog and rodent cells. Examples of human cells are PER.C6 cells (WO01/38362), MRC-5 (ATCC CCL-171), WI-38 (ATCC CCL-75), HEK-293 cells (ATCC CRL-1573), HeLa cells (ATCC CCL2), and fetal rhesus lung cells (ATCC CL- 160). Examples of non-human primate cells are Vero cells (ATCC CCL81), COS-1 cells (ATCC CRL-1650) or COS-7 cells (ATCC CRL-1651). Examples of dog cells are MDCK cells (ATCC CCL-34). Examples of rodent cells are hamster cells, such as BHK21-F, HKCC cells, or CHO cells.

[0106] As an alternative to mammalian sources, cell lines for use in the disclosure may be derived from avian sources such as chicken, duck, goose, quail or pheasant. Examples of avian cell lines include avian embryonic stem cells (WO01/85938 and WO03/076601), immortalized duck retina cells (WO2005/042728), and avian embryonic stem cell derived cells, including chicken cells (WO2006/108846) or duck cells, such as EB66 cell line (WO2008/129058 & WO2008/142124).

[0107] In another aspect, said host cell are insect cells, such as SF9 cells (ATCC CRL-1711), Sf21 cells (IPLB-Sf21), MG1 cells (BTI-TN-MG1) or High Five™ cells (BTI-TN-5B1-4).

[0108] Accordingly, in a particular aspect, optionally in combination with one or more features of the various aspects described above or below, the host cell comprises:

a) a nucleic acid construct or expression vector of the disclosure (i.e., the recombinant AAV genome), generally as a plasmid;

b) a nucleic acid construct, generally a plasmid, encoding AAV rep and/or cap genes which does not carry the ITR sequences; and/or
c) a nucleic acid construct, generally a plasmid or virus, comprising viral helper genes.

[0109] Viral genes necessary for AAV replication are referred herein as viral helper genes. Typically, said genes necessary for AAV replication are adenoviral helper genes, such as E1A, E1B, E2a, E4, or VA RNAs. Preferably, the adenoviral helper genes are of the Ad5 or Ad2 serotype.

[0110] Conventional methods can be used to produce viral particles of the AAV vector, which consist on transient cell co-transfection with nucleic acid construct (e.g. a plasmid) carrying the recombinant AAV vector/genome of the disclosure; a nucleic acid construct (e.g., an AAV helper plasmid) that encodes rep and cap genes, but does not carry ITR sequences; and with a third nucleic acid construct (e.g., a plasmid) providing the adenoviral functions necessary for AAV replication. Thus, in a particular aspect, optionally in combination with one or more of the features of the various aspects described above or below, said host cell is characterized by comprising:

i) a nucleic acid construct or an expression vector of the disclosure (i.e., the recombinant AAV genome);
ii) a nucleic acid construct encoding AAV rep and cap genes which does not carry the ITR sequences; and
iii) a nucleic acid construct comprising adenoviral helper genes.

[0111] Alternatively, the rep, cap, and adenoviral helper genes can be combined on a single plasmid (Blouin Vet al. J Gene Med. 2004; 6(suppl): S223-S228; Grimm D. et al. Hum. Gene Ther. 2003; 7: 839-850). Thus, in another particular aspect, optionally in combination with one or more of the features of the various aspects described above or below, said host cell is characterized by comprising:

i) a nucleic acid construct or an expression vector of the disclosure (i.e., the recombinant AAV genome); and
ii) a plasmid encoding AAV rep and cap genes which does not carry the ITR sequences and further comprising adenoviral helper genes.

[0112] In a further particular aspect, optionally in combination with one or more features of the various aspects described above or below, the host cell comprises:

a) a nucleic acid construct or an expression vector of the disclosure (i.e., the recombinant AAV genome);
b) a plasmid encoding AAV rep and cap genes which does not carry the ITR sequences; and
c) a plasmid comprising adenoviral helper genes E2a, E4, and VA RNAs,

wherein co-transfection is performed in cells, preferably mammalian cells, that constitutively express and transcomplement the adenoviral E1 gene, like HEK-293 cells (ATCC CRL-1573).

[0113] Large-scale production of AAV vectors according to the disclosure can also be carried out for example by infection of insect cells with a combination of recombinant baculoviruses (Urabe et al. Hum. Gene Ther. 2002; 13: 1935-1943). SF9 cells are co-infected with three baculovirus vectors respectively expressing AAV rep, AAV cap and the AAV vector to be packaged. The recombinant baculovirus vectors will provide the viral helper gene functions required for virus replication and/or packaging.

[0114] By using helper plasmids encoding the rep ORF (open reading frame) of an AAV serotype and cap ORF of a different serotype AAV, it is feasible packaging a vector flanked by ITRs of a given AAV serotype into virions assembled from capsid structural proteins of a different serotype. It is also possible by this same procedure packaging mosaic, chimeric or targeted vectors.

[0115] On the other hand, the production of HC-Ad vectors according to the disclosure can be carried out by means of mammalian cells that constitutively express and transcomplement the adenoviral E1 gene, and also Cre recombinase (e.g. 293Cre cells). These cells are transfected with the HC-Ad vector genome and infected with a first-generation adenoviral helper virus (E1-deleted) in which the packaging signal is flanked by loxP sequences. [Parks RJ et al. Proc. Natl. Acad. Sci. USA 1996; 13565-13570; for 293Cre cells, see Palmer and Engel. Mol. Ther. 2003; 8:846-852]. Several Cre/loxP-based helper virus systems have been described that can be used for packaging HC-Ad vectors, such as AdAdLC8cluc, or the optimized self-inactivating AdTetCre helper virus (EP2295591; Gonzalez-Aparicio et al. Gene Therapy 2011; 18: 1025-1033).

[0116] Further guidance for the construction and production of viral vectors for gene therapy according to the disclosure can be found in:
Viral Vectors for Gene Therapy, Methods and Protocols. Series: Methods in Molecular Biology, Vol. 737. Merten and Al-Rubeai (Eds.); 2011 Humana Press (Springer).

[0117] Gene Therapy. M. Giacca. 2010 Springer-Verlag.

[0118] Heilbronn R. and Weger S. Viral Vectors for Gene Transfer: Current Status of Gene Therapeutics. In: Drug Delivery, Handbook of Experimental Pharmacology 197; M. Schäfer-Korting (Ed.). 2010 Springer-Verlag; pp. 143-170.
[0119] Adeno-Associated Virus: Methods and Protocols. R.O. Snyder and P. Moulllier (Eds). 2011 Humana Press (Springer).
[0120] Bünning H. et al. Recent developments in adeno-associated virus technology. J. Gene Med. 2008; 10:717-733.
[0121] Adenovirus: Methods and Protocols. M. Chillon and A. Bosch (Eds.); Third Edition. 2014 Humana Press (Springer).

THERAPEUTIC USES

[0122] In a further aspect, the disclosure relates to the product of the disclosure as defined within the Summary for use as a medicament.

[0123] In an additional aspect, the disclosure relates to the product of the disclosure as defined within the Summary for use in the treatment of a condition caused by a deficiency or dysfunction of Copper-transporting ATPase 2, and of any other conditions and illnesses in which an upregulation of Copper-transporting ATPase 2 expression and activity may produce a therapeutic benefit or improvement, in particular a disease or condition associated with a decrease of ATP7B-dependent lysosomal exocytosis and accumulation of copper in lysosomes, such as choleostatic disorders, Alzheimer disease and/or cancer (Polishchuck et al. Dev Cell. 2014, 29(6), 686-700; Gupta and Lutsenko, Future Med. Chem. 2009, 1, 1125-1142)..

[0124] The subject to be treated can be a mammal, and in particular a human patient.

[0125] In a particular embodiment, optionally in combination with one or more features of the various embodiments described above or below, the condition caused by a deficiency or dysfunction of Copper-transporting ATPase is Wilson's disease (WD, Online Mendelian Inheritance in Man catalog accession number OMIN 277900; http://www.omim.org/entry/277900).

[0126] In a related aspect, the disclosure pertains to the use of the product of the disclosure, as defined within the Summary, in the preparation of a medicament for use in the treatment of a condition caused by a deficiency or dysfunction of Copper-transporting ATPase 2, and of any other conditions and illnesses in which an upregulation of Copper-transporting ATPase 2 expression and activity may produce a therapeutic benefit or improvement, preferably for use in the treatment of Wilson's disease.

[0127] In a further aspect, the disclosure relates to the treatment of a condition caused by a deficiency or dysfunction of Copper-transporting ATPase 2, and of any other conditions and illnesses in which an upregulation of Copper-transporting ATPase 2 expression and activity may produce a therapeutic benefit or improvement, preferably for use in the treatment of Wilson's disease, in a patient that comprises administering to the patient a therapeutically effective amount of a nucleic acid construct, an expression vector, a host cell, a viral particle or a pharmaceutical composition of the disclosure.

[0128] The treatment with a product of the disclosure may alleviate, ameliorate, or reduce the severity of one or more symptoms of WD. For example, treatment may increase and/or restore holoceruplasmin synthesis, ceruloplasmin oxidase activity, and /or copper excretion in the bile (thus reducing copper accumulation in serum, liver, brain and urine); and as a consequence may alleviate, ameliorate, or reduce the severity of abdominal pain, fatigue, jaundice, frequency of uncontrolled movements, muscle stiffness, problems with speech, swallowing or physical coordination.

[0129] The product of the disclosure will be typically included in a pharmaceutical composition or medicament, optionally in combination with a pharmaceutical carrier, diluent and/or adjuvant. Such composition or medicinal product comprises the product of the disclosure in an effective amount, sufficient to provide a desired therapeutic effect, and a pharmaceutically acceptable carrier or excipient.

[0130] Accordingly, in a further aspect, the disclosure relates to a pharmaceutical composition that comprises a nucleic acid construct, an expression vector, a host cell or a viral particle of the disclosure, and a pharmaceutically acceptable carrier.

[0131] Any suitable pharmaceutically acceptable carrier or excipient can be used in the preparation of a pharmaceutical composition according to the disclosure (See e.g., Remington: The Science and Practice of Pharmacy, Alfonso R. Gennaro (Editor) Mack Publishing Company, April 1997). Pharmaceutical compositions are typically sterile and stable under the conditions of manufacture and storage. Pharmaceutical compositions may be formulated as solutions (e.g. saline, dextrose solution, or buffered solution, or other pharmaceutically acceptable sterile fluids), microemulsions, liposomes, or other ordered structure suitable to accommodate a high product concentration (e.g. microparticles or nanoparticles). The carrier may be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, or sodium chloride in the composition. Prolonged

absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, monostearate salts and gelatin. The product of the disclosure may be administered in a controlled release formulation, for example in a composition which includes a slow release polymer or other carriers that protect the product against rapid release, including implants and microencapsulated delivery systems. Biodegradable and biocompatible polymers may for example be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, polylactic acid and polylactic / polyglycolic copolymers (PLG). Preferably, said pharmaceutical composition is formulated as a solution, more preferably as an optionally buffered saline solution.

[0132] Supplementary active compounds can also be incorporated into the pharmaceutical compositions of the disclosure. Guidance on co-administration of additional therapeutics can for example be found in the Compendium of Pharmaceutical and Specialties (CPS) of the Canadian Pharmacists Association.

[0133] In one aspect, optionally in combination with one or more features of the various aspects described above or below, the pharmaceutical composition of the disclosure is a parenteral pharmaceutical composition, including a composition suitable for intravenous, intraarterial, subcutaneous, intraperitoneal or intramuscular administration. These pharmaceutical compositions are exemplary only and do not limit the pharmaceutical compositions suitable for other parenteral and non-parenteral administration routes.

[0134] In the context of the disclosure, an "effective amount" means a therapeutically effective amount.

[0135] As used herein 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 an elevation of copper translocation activity, thus increasing copper in bile and reducing copper in serum, liver, brain and urine. The therapeutically effective amount of the product of the disclosure, or pharmaceutical composition that comprises it may vary according to factors such as the disease state, age, sex, and weight of the individual, and the ability of the product or pharmaceutical composition to elicit a desired response in the individual. Dosage regimens may be adjusted to provide the optimum therapeutic response. A therapeutically effective amount is also typically one in which any toxic or detrimental effect of the product or pharmaceutical composition is outweighed by the therapeutically beneficial effects.

[0136] In one aspect, optionally in combination with one or more features of the various aspects described above or below, the pharmaceutical composition carrying the product of the disclosure is administered to the subject or patient by a parenteral route.

[0137] In one aspect, optionally in combination with one or more features of the various aspects described above or below, the pharmaceutical composition is administered by intravenous, intraarterial, subcutaneous, intraperitoneal, or intramuscular route.

[0138] In one aspect, optionally in combination with one or more features of the various aspects described above or below, the pharmaceutical composition comprising a product of the disclosure is administered by interstitial route, i.e. by injection to or into the interstices of a tissue. The tissue target may be specific, for example the liver tissue, or it may be a combination of several tissues, for example the muscle and liver tissues. Exemplary tissue targets may include liver, skeletal muscle, heart muscle, adipose deposits, kidney, lung, vascular endothelium, epithelial and/or hematopoietic cells. In a preferred aspect, optionally in combination with one or more features of the various aspects described above or below, it is administered by intrahepatic injection, i.e. injection into the interstitial space of hepatic tissue.

[0139] The amount of product of the disclosure that is administered to the subject or patient may vary depending on the particular circumstances of the individual subject or patient including, age, sex, and weight of the individual; the nature and stage of the disease, the aggressiveness of the disease; the route of administration; and/or concomitant medication that has been prescribed to the subject or patient. Dosage regimens may be adjusted to provide the optimum therapeutic response.

[0140] For any particular subject, specific dosage regimens may be adjusted over time according to the individual needs and the professional judgment of the person administering or supervising the administration of the compositions. Dosage ranges set forth herein are exemplary only and do not limit the dosage ranges that may be selected by medical practitioners.

[0141] In one embodiment, an AAV vector according to the disclosure can be administered to the subject or patient for the treatment of Wilson's disease in an amount or dose comprised within a range of 5×10^{11} to 1×10^{14} vg / kg (vg: viral genomes; kg: subject's or patient's body weight). In a more particular embodiment, the AAV vector is administered in an amount comprised within a range of 1×10^{12} to 1×10^{13} vg / kg.

[0142] In another aspect, a HC-Ad vector according to the disclosure can be administered to the subject or patient for the treatment of Wilson's disease in an amount or dose comprised within a range of 1×10^{9} to 1×10^{11} iu / kg (iu: infective units of the vector).

[0143] In another aspect, the disclosure further relates to a kit comprising a nucleic acid construct, vector, host cell, viral particle or pharmaceutical composition of the disclosure in one or more containers. A kit of the disclosure may include instructions or packaging materials that describe how to administer a nucleic acid construct, vector, host cell or viral particle of the disclosure contained within the kit to a patient. Containers of the kit can be of any suitable material, e.g., glass, plastic, metal, etc., and of any suitable size, shape, or configuration. In certain embodiments, the kits may

include one or more ampoules or syringes that contain a nucleic acid construct, vector, host cell, viral particle or pharmaceutical composition of the disclosure in a suitable liquid or solution form.

[0144] Throughout the description and claims the word "comprise" and variations of thereof, are not intended to exclude other technical features, additives, components, or steps. Furthermore, the word "comprise" encompasses the case of "consisting of". Additional objects, advantages and features of the disclosure will become apparent to those skilled in the art upon examination of the description or may be learned by practice of the disclosure. Furthermore, all possible combinations of particular and preferred embodiments are described herein.

EXAMPLES

Example 1. Construction of recombinant expression vectors

[0145] Five different AAV vectors that carry and express human ATP7B, or a truncated form of human ATP7B, were designed and produced for conducting gene therapy of Wilson's Disease (WD): AAV2/8-AAT-wtATP7B, AAV2/8-AAT-coATP7B, AAV2/8-AAT-ATP7B(d223-366), AAV2/8-AAT-ATP7B(d57-486), and AAV2/8-AAT-coATP7B(d57-486).

1.1 Vector AAV2/8-AAT-wtATP7B [herein also named as AAV-wtATP7B1

[0146] Genomic sequence of this vector is identified as SEQ.ID.NO.1.

[0147] Firstly, the plasmid pUC-ATP7B was assembled at request (GenScript) by cloning nucleic acid construct into a pUC57 plasmid. Nucleic acid construct contained cDNA sequence encoding human ATP7B (transgene) together with a synthetic polyadenylation signal sequence (Levitt N. et al. Genes & Development 1989; 3(7):1019-1025) downstream of the transgene.

[0148] Next, the minimal promoter of alpha 1 anti-trypsin gene (AAT) was introduced into the plasmid pUC-ATP7B, upstream the ATP7B gene. The minimal promoter consists on the sequence from nucleotide -261 to nucleotide +44 relative to cap site of the AAT promoter (Kramer M.G. et al. Mol. Therapy 2003; 7(3): 375-385) and contains the tissue-specific element (TSE), required for liver function, and the distal region (DRI) required for whole promoter activity. The AAT promoter was obtained by PCR amplification using as template the pEnhAlbAAT-luciferase plasmid (provided by M.G. Kramer) and the following primers

Primer AAT-Forward
5' CTGGTCTAGAACGCGTCGCCACCCCCTCCACCTTGG 3' (SEQ.ID.NO.10); and
Primer AAT-reverse
5' ATCATGATGCGGCCGCTTCACTGTCCCAGGTCAGTG 3' (SEQ.ID.NO.11).

[0149] The AAT-Forward primer has a restriction site for XbaI and MluI and the 3' AAT-reverse primer has a restriction site for NotI.

[0150] Therefore, in order to obtain plasmid pUC-AAT-ATP7B, the plasmid pUC-ATP7B was digested with XbaI and NotI and ligated to AAT promoter previously digested with the same enzymes.

[0151] The expression cassette was subsequently subcloned into the AAV transfer plasmid pAAV-MCS (Agilent technologies) by digestion with restriction enzymes PmlI and MluI, thus producing the plasmid pAAV2-AAT-wtATP7B.

[0152] Once the plasmid had been constructed, the AAV vector was made by double transfection into 293 cells of the plasmid pAAV2-AAT-ATP7B and of the plasmid pDP8 (obtained from PlasmidFactory, Bielefeld, Germany; plasmid pDP8 expresses AAV8 capsid protein, AAV2 rep protein and the adenoviral molecules required for production and packaging of AAV).

[0153] The vector was finally purified by iodixanol gradient and titrated by quantitative PCR.

1.2 Vector AAV2/8-AAT-coATP7B [herein also named as AAV-coATP7B]

[0154] Genomic sequence of this vector is identified as SEQ.ID.NO.3.

[0155] To obtain the AAV vector expressing a codon optimized version of the ATP7B gene (coATP7B), the plasmid pUC-coATP7B was firstly assembled at request (GenScript) by cloning nucleic acid construct into a pUC57 plasmid. Next the coATP7B was excised from the pUC-coATP7B by digestion with the restriction enzymes NotI and KpnI and subcloned into the pAAV2-AAT-wtATP7B plasmid previously digested with the same enzymes, NotI and KpnI, to obtain the plasmid pAAV2-AAT-coATP7B.

[0156] Once the plasmid had been constructed, the production of vector genome and packaging of viral particles was performed as has been described previously for the vector AAV2/8-AAT-wtATP7B: double transfection of previously obtained plasmid pAAV2-AAT-coATP7B with plasmid pDP8, purification (iodixanol gradient) and titration.

1.3 Vector AAV2/8-AAT-ATP7B(d223-366)[herein also named as AAV-T1]

[0157] This vector carries as the transgene a nucleic acid sequence (SEQ.ID.NO.12) encoding ATP7B(d223-366), a truncated form of human ATP7B in which amino acids 223 to 366 have been deleted. The deleted sequence includes HMA 3 domain and seven amino acids of the HMA 4 domain.

[0158] To obtain the vector, the plasmid pUC57-wtATP7B was digested with the restriction enzymes Mfel and Nae I, to obtain the plasmid pUC57-ATP7B-T1. This way, the size of the codifying region was reduced in 432 nucleotides and the size of the protein in 144 amino acids.

[0159] Once the plasmid pUC57-ATP7B-T1 had been constructed, the production of vector genome and packaging of viral particles was performed as described previously for the vector AAV2/8-AAT-wtATP7B: ligation to AAT promoter, subcloning into plasmid pAAV-MCS, double transfection of previously obtained plasmid pAAV2-AAT-T1 with plasmid pDP8, virus purification (iodixanol gradient) and titration.

1.4 Vector AAV2/8-AAT-ATP7B(d57-486)Therein also named as AAV-T21

[0160] Genomic sequence of this vector is identified as SEQ.ID.NO.6.

[0161] This vector carries as the transgene a nucleic acid sequence encoding ATP7B(d57-486) [also named as ATP7B-T2], a truncated form of human ATP7B in which amino acids 57 to 486 have been deleted. This way, the first 4 HMA domains have been eliminated while maintaining the signal sequence that comprises the 56 amino acids of the amino terminal region, reducing the size of the codifying region in 1.29 Kb and the protein in 430 amino acids.

[0162] The nucleotide sequence of ATP7B(d57-486) was obtained by PCR amplification using the pUC57-wtATP7B as template and two sets of primers;
a first set of primers amplifying the amino terminal sequence:

Primer F1:
5' CTAGATGCGGCCGCCACCATGCCTG 3' (SEQ.ID.NO.14), and
Primer R1:
5' CTGAGAAGAAGGGCCCAGGCC 3' (SEQ.ID.NO.15); and

a second set of primers amplifying the carboxy terminal region:

Primer F2:
5' GGCCCTTCTTCTCAGCCGCAGAAGTGCTTCTTACAG 3' (SEQ.ID.NO.16), and
Primer R2:
5' ACCAAAATCGATAAAACCGATTACAATCC 3' (SEQ.ID.NO.17).

[0163] The 5' terminal sequences of primers R1 and F2 are complementary. Using equimolecular amounts of the two PCR purified fragments as template, and primers F1 and R2, PCR was performed to obtain nucleotide sequence encoding ATP7B(d57-486) . The PCR product was then digested with NotI and ClaI and cloned into the pUC57-AAT-wtATP7B plasmid previously digested with both enzymes obtaining the plasmid pUC57-ATP7B-T2.

[0164] Once the plasmid pUC57-ATP7B-T2 had been constructed, the production of vector genome and packaging of viral particles was performed as described previously for the vector AAV2/8-AAT-wtATP7B: ligation to AAT promoter, subcloning into plasmid pAAV-MCS, double transfection of previously obtained plasmid pAAV2-AAT-T2 with plasmid pDP8, purification (iodixanol gradient) and titration.

1.5 Vector AAV2/8-AAT-coATP7B(d57-486)Therein also named as AAV-AAT-coT2]

[0165] This vector carries as transgene a codon optimized nucleic acid sequence [SEQ.ID.NO.8; coATP7B(d57-486) or coATP7B-T2] that also encodes ATP7B(d57-486).

[0166] The nucleotide sequence of coATP7B(d57-486) was obtained by PCR amplification using the pUC57-coATP7B as template and two sets of primers;
a first set of primers amplifying the amino terminal sequence:

Primer F3:
5' ACGCGTGCGGCCGCCACCATGCCAG 3' (SEQ.ID.NO.18), and
Primer R3:
5' CTGGGAGCTAGGTCCCAGTCC 3' (SEQ.ID.NO.19); and

A second set of primers amplifying the carboxy terminal region:

Primer F4:
5' GGACCTAGCTCCCAGCCTCAGAAGTGTTTTCTGCAG 3' (SEQ.ID.NO.20), and
Primer R4:
5' TGTTCCTCGCGAATGATCAGGTTGTCCTC 3' (SEQ.ID.NO.21).

[0167] The 5' terminal sequences of primers R3 and F4 are complementary. Using equimolecular amounts of the two PCR purified fragments as template, and primers F3 and R4, PCR was performed to obtain codon optimized nucleotide sequence encoding ATP7B(d57-486). The PCR product was then digested with NotI and NruI and cloned into the pUC57-AAT-wtATP7B plasmid previously digested with both enzymes obtaining the plasmid pUC57-coATP7B-T2.

[0168] Once the plasmid pUC57-coATP7B-T2 had been constructed, the production of vector genome and packaging of viral particles was performed as described previously for the vector AAV2/8-AAT-wtATP7B: ligation to AAT promoter, subcloning into plasmid pAAV-MCS, double transfection of previously obtained plasmid pAAV2-AAT-coT2 with plasmid pDP8, virus purification (iodixanol gradient) and titration.

Example 2. Wilson's disease animal model: ATP7B KO

[0169] The therapeutic performance of the vectors AAV2/8-AAT-ATP7B-T1 and AAV2/8-AAT-ATP7B-T2 was tested in ATP7B knockout mice (ATP7B KO, ATP7B$^{-/-}$ or WD mice) which are a representative animal model of WD. This animal model was developed by Buiakova et al., by introducing an early termination codon in the mouse ATP7B mRNA by engineering the substitution of a portion of ATP7B exon 2 with a neomycin cassette oriented in the opposite transcriptional frame (Buikova O.I. et al. Human Molecular Genetics 1999; 8(9): 1665-1671). ATP7B knockout mice show no ATP7B expression in the liver and high Cu excretion in the urine, low holoceruloplasmin levels in serum, high transaminase levels, high Cu concentration in the liver and a pathologic liver histology. These mice exhibit the typical biochemical characteristics of human Wilson's disease except for the neurological affectation (Lutsenko S. Biochemical Society Transactions 2008; 36(Pt 6): 1233-1238).

Example 3. Determination of the therapeutic effect of viral vectors AAV2/8-AAT-ATP7B-T1 and AAV2/8-AAT-ATP7B-T2 in Wilson's disease mice

[0170] Six weeks (6w) old male ATP7B-/- mice were divided in 4 groups of 5 mice each: 1 of the groups were treated intravenously with the vector AAV2/8-AAT-wtATP7B at a dose of $3x10^{10}$ vg / mouse (vg: viral genomes); a second group with the same dose of the vector AAV2/8-AAT-ATP7B-T1; a third group with the same dose of the vector AAV2/8-AAT-ATP7B-T2; and a fourth group was left untreated. An additional group of wild type mice was kept untreated as a control group (control). Animals were sacrificed twenty-four weeks after vector administration (w30).

[0171] Four weeks after vector administration and every five weeks after that up to week 30; serum transaminases (ALT) levels and urine Cu content were determined in all the groups. Serum ceruloplasmin activity was measured 4 weeks after treatment.

[0172] Serum transaminases (ALT) levels were determined by the DGKC method (Roche Diagnostics, Mannheim, Germany) using a Hitachi 747 Clinical Analyzer (Hitachi, Tokyo, Japan).

[0173] Serum ceruloplasmin activity was determined with o-dianisidine dihydrochloride (4, 4'-diamino-3,3'-dimethoxy-biphenyl) as substrate (Sigma-Aldrich, San Louis, MO, United States) as described by Schosinsky and cols. (Clinical Chemistry 1974; 20(12): 1556-1563). Absorbance was measured at 540 nm in a spectrophotometer.

[0174] Urine copper content was determined by atomic absorption spectroscopy (SIMAA 6000, from Perkin-Elmer GmbH, Bodenseewerk).

[0175] After the sacrifice the liver was excised for histological analyses.

[0176] Hepatic copper content was determined in dry liver tissue by atomic absorption spectroscopy (SIMAA 6000, from Perkin-Elmer GmbH, Bodenseewerk), and by Timm's sulphide silver staining (Danscher G. and Zimmer J. Histochemistry 1978; 55(1): 27-40).

[0177] Liver structure was assessed in sections stained with hematoxylin and eosin.

[0178] Immunohistochemistry with anti-mouse CD45 antibody (BioLegend, San Diego, USA; Catalog Number 103102) was performed to detect inflammatory infiltration in the liver.

[0179] Immunohistochemistry with anti-mouse PanCk antibody (Invitrogen/Life Technologies, 18-0132, clon AE1/AE3) was also performed to detect biliary cells.

[0180] To determine fibrosis we used conventional Sirius Red staining as a method for collagen determination.

[0181] As shown in Figure 2, transaminase levels were normalized in the mice receiving AAV2/8-AAT-wtATP7B or AAV2/8-AAT-ATP7B-T2 but no in animals treated with AAV2/8-AAT-ATP7B-T1. Furthermore, the concentration of Cu

in urine was significantly lower in the animals that received AAV2/8-AAT-wtATP7B, AAV2/8-AAT-ATP7B-T1, or AAV2/8-AAT-ATP7B-T2; however AAV2/8-AAT-ATP7B-T1 was less efficient in reducing Cu concentration in urine (Figure 3). Ceruloplasmin activity was restored four weeks after treatment in the animals receiving AAV2/8-AAT-wtATP7B or AAV2/8-AAT-ATP7B-T2 but no in animal treated with AAV2/8-AAT-ATP7B-T1 (Figure 4). This result was corroborated by western blot analysis. Holoceruloplasmin was detected in mice treated with AAV2/8-AAT-wtATP7B or AAV2/8-AAT-ATP7B-T2 but no in animals treated with AAV2/8-AAT-ATP7B-T1 where as in untreated WD mice only the apoceruloplasmin form could be detected.

[0182] On the other hand, the administration of the AAV2/8-AAT-wtATP7B, AAV2/8-AAT-ATP7B-T1, or AAV2/8-AAT-ATP7B-T2 significantly reduced Cu content in the liver; however, AAV2/8-AAT-ATP7B-T1 was less efficient in reducing Cu concentration in the liver (Figure 5). The results were confirmed in the image obtained after Timm's staining (Figure 6B). Regarding liver histology, untreated animals showed an abnormal hepatic architecture with huge hepatocytes containing enormous nuclei. The administration of the vectors AAV2/8-AAT-wtATP7B or AAV2/8-AAT-ATP7B-T2 but no AAV2/8-AAT-ATP7B-T1 resulted in the normalization of liver histology (Figure 6A). Furthermore, WD animals presented a strong liver infiltrate mainly composed by CD45 positive cells; infiltration disappeared after treatment with the recombinant viral vectors (Figure 7). Thus, the administration of AAV vector resulted in a marked reduction of the inflammatory infiltrate. Furthermore, biliary duct proliferation and liver fibrosis were also significantly reduced in AAV2/8-AAT-wtATP7B, AAV2/8-AAT-ATP7B-T2, and AAV2/8-AAT-ATP7B-T1-treated WD mice (Figure 7).

Example 4. Therapeutic effect of viral vector AAV2/8-AAT-ATP7B(d57-486) in Wilson's disease female mice.

[0183] Six weeks (6w) old female ATP7B-/- mice were divided in 4 groups of 5 mice each: animals of the groups 1 - 3 were treated intravenously with the viral vector AAV2/8-AAT-ATP7B(d57-486), each group receiving a different dose (respectively 1×10^{10}, 3×10^{10}, and 1×10^{11} vg / mouse); a fourth group were left untreated. An additional group of wild type mice was kept untreated as a control group (WT).

[0184] Four weeks after vector administration and every five weeks after that up to 24 weeks after treatment (when the mice were 30 weeks old), serum transaminases (ALT) levels and urine Cu concentration were determined in all the groups, by the same methods as described in Example 3.

[0185] As shown in Figure 8, AAV2/8-AAT-ATP7B(d57-486) normalized transaminase levels in WD female mice at the two highest doses (3×10^{10}, and 1×10^{11} vg / mouse); the lowest dose 1×10^{10} vg / mouse significantly reduced transaminase levels but failed to eliminate liver damage. However, treatment with the three different doses significantly reduced Cu urinary excretion reaching the levels found in WT mice (Figure 9).

Example 5. Comparison of the therapeutic effect of viral vectors AAV2/8-AAT-wtATP7B and AAV2/8-AAT-ATP7B(d57-486) in Wilson's disease female mice.

[0186] Two experimental groups were established. For each experimental group, six weeks (6w) old female ATP7B-/- mice were divided in 4 groups of 5 mice each: 3 of the groups were treated intravenously with a viral vector to be tested, each group receiving a different dose (respectively 1×10^{10}, 3×10^{10}, and 1×10^{11} vg / mouse; a fourth group were left untreated. An additional group of wild type mice were kept untreated as a control group (WT).

[0187] In first experimental group (experimental group 1), WD mice receiving treatment were administered with the vector AAV2/8-AAT-wtATP7B; in second experimental group (experimental group 2) they were administered with the vector AAV2/8-AAT-ATP7B(d57-486).

[0188] Serum ceruloplasmin activity determined 4 weeks after treatment, and hepatic Cu content determined 24 weeks after treatment, were measured by the same methods as described in example 3.

Serum ceruloplasmin activity

[0189] Serum ceruloplasmin activity was corrected only by the administration of the highest dose of the AAV2/8-AAT-wtATP7B vector (Figure 10A experimental group 1); no effect being observed after the administration of the two lowest doses.

[0190] Conversely, the AAV2/8-AAT-ATP7B(d57-486) vector significantly increased ceruloplasmin levels at the lowest dose of 1×10^{10} vg / mouse; the administration of the medium dose of vector normalized ceruloplasmin levels and the highest dose increased ceruloplasmin activity over the normal levels (Figure 10B experimental group 2).

Cu concentration in the liver

[0191] Besides, Cu concentration in the liver was reduced but not normalized by the administration of the two highest doses of AAV2/8-AAT-wtATP7B; and no effect was observed at the lowest dose (Figure 11A experimental group 1). On

the contrary, Cu concentration was shown to be reduced after administration of the AAV2/8-AAT-ATP7B(d57-486) vector at all the tested doses, and at the highest dose the levels were close to normal (Figure 11B experimental group 2).

[0192] Accordingly, a dose of 1 x 10^{10} vg / mouse of the AAV2/8-AAT-wtATP7B vector was shown to be a "suboptimal dose" for the wt construct both for the obtaining of a normalization of the serum ceruloplasmin activity and a reduction of Cu accumulation in the liver; whereas the vector carrying the truncated form unexpectedly provided statistically significant therapeutic effects at said suboptimal dose.

Example 6. Comparison of the therapeutic effect of viral vectors AAV2/8-AAT-wtATP7B and AAV2/8-AAT-ATP7B(d57-486) in WD mice.

[0193] Six weeks (6w) old male ATP7B-/- mice were divided in 3 groups of mice: 2 groups of animals were respectively treated with a suboptimal intravenous dose (1 x 10^{10} vg / mouse) of the vector AAV2/8-AAT-wtATP7B or the vector AAV2/8-AAT-ATP7B(d57-486); a third group were left untreated. An additional group of wild type mice were kept untreated as a control group (WT).

[0194] Hepatic Cu content was measured by the same method as described in example 3.

[0195] As it is shown in Figure 12, although both AAV2/8-AAT-wtATP7B and AAV2/8-AAT-ATP7B(d57-486) vectors given at a suboptimal dose reduced accumulation of copper in the liver of WD mice, AAV2/8-AAT-ATP7B(d57-486) provided a reduction of hepatic copper content that was significantly greater than the reduction provided by AAV2/8-AAT-wtATP7B.

Example 7. Comparison of the therapeutic effect of viral vectors AAV2/8-AAT-ATP7B(d57-486) and AAV-AAT-coATP7B(d57-486) in WD mice.

[0196] Six weeks (6w) old male ATP7B-/- mice were divided in 3 groups of mice: 2 groups of animals were respectively treated with a suboptimal intravenous dose (1 x 10^{10} vg / mouse) of the vector AAV2/8-AAT-ATP7B(d57-486) and AAV-AAT-coATP7B(d57-486); a third group were left untreated. An additional group of wild type mice were kept untreated as a control group (WT).

[0197] Hepatic Cu content was measured by the same method as described in example 3.

[0198] As it is shown in Figure 13, although both AAV2/8-AAT-ATP7B(d57-486) and AAV2/8-AAT-coATP7B(d57-486)vectors given at a suboptimal dose reduced accumulation of copper in the liver of WD mice, AAV2/8-AAT-coATP7B(d57-486) provided a reduction of hepatic copper content that was significantly greater than the reduction provided by AAV2/8-AAT-ATP7B(d57-486).

Example 8. Therapeutic effect of codon optimized viral vector AAV2/8-AAT-coATP7B(d57-486) in WD mice.

[0199] Six weeks (6w) old male ATP7B-/- mice were divided in 5 groups of mice: 4 groups of animals were respectively treated with a suboptimal intravenous dose (1 x 10^{10} vg / mouse) of the vectors AAV2/8-AAT-wtATP7B, AAV2/8-AAT-coATP7B, AAV2/8-AAT-ATP7B(d57-486) or AAV2/8-AAT-coATP7B(d57-486); a fifth group were left untreated. An additional group of wild type mice were kept untreated as a control group (WT).

[0200] Serum ceruloplasmin activity was measured by the same method as described in example 3.

[0201] As it is shown in Figure 14, the two vectors carrying nucleotide sequence of truncated ATP7B-T2 restored ceruloplasmin oxidase activity when administered to WD mice at the suboptimal dose, while vectors carrying nucleotide sequences encoding complete human ATP7B did not provide any significant improvement of ceruloplasmin activity when administered at the same treatment conditions.

SEQUENCE LISTING

[0202]

<110> FUNDACION PARA LA INVESTIGACION MEDICA APLICADA

<120> NUCLEIC ACID CONSTRUCTS AND GENE THERAPY VECTORS FOR USE IN THE TREATMENT OF WILSON DISEASE AND OTHER CONDITIONS

<130> 2014005EP00

<160> 21

<170> BiSSAP 1.3

<210> 1
<211> 5107
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleic acid construct of expression vector AAV2-AAT-wtATP7B

<220>
<221> repeat_region
<222> 1..141
<223> /note="5' ITR of adeno-associated virus serotype 2"

<220>
<221> promoter
<222> 156..460
<223> /note="alpha 1 antitrypsin"

<220>
<221> CDS
<222> 473..4870
<223> /note="Sequence encoding ATP7B (Copper transporting ATPase 2)" /transl_table=1

<220>
<221> polyA_signal
<222> 4877..4932

<220>
<221> repeat_region
<222> 4968..5107
<223> /standard_name="3' ITR of adeno-associated virus serotype 2"

<400> 1

```
cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc    60

gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca   120

actccatcac taggggttcc tgcggccgca cgcgtcgcca ccccctccac cttggacaca   180

ggacgctgtg gtttctgagc caggtacaat gactcctttc ggtaagtgca gtggaagctg   240

tacactgccc aggcaaagcg tccgggcagc gtaggcgggc gactcagatc ccagccagtg   300

gacttagccc ctgtttgctc ctccgataac tggggtgacc ttggttaata ttcaccagca   360

gcctccccg ttgcccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc    420

tcctcagctt caggcaccac cactgacctg ggacagtgaa gcggccgcca cc atg cct    478
                                                        Met Pro
```

1

gag cag gag aga cag atc aca gcc aga gaa ggg gcc agt cgg aaa atc 526
Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg Lys Ile
5 10 15

tta tct aag ctt tct ttg cct acc cgt gcc tgg gaa cca gca atg aag 574
Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala Met Lys
20 25 30

aag agt ttt gct ttt gac aat gtt ggc tat gaa ggt ggt ctg gat ggc 622
Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu Asp Gly
35 40 45 50

ctg ggc cct tct tct cag gtg gcc acc agc aca gtc agg atc ttg ggc 670
Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile Leu Gly
55 60 65

atg act tgc cag tca tgt gtg aag tcc att gag gac agg att tcc aat 718
Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile Ser Asn
70 75 80

ttg aaa ggc atc atc agc atg aag gtt tcc ctg gaa caa ggc agt gcc 766
Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly Ser Ala
85 90 95

act gtg aaa tat gtg cca tcg gtt gtg tgc ctg caa cag gtt tgc cat 814
Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val Cys His
100 105 110

caa att ggg gac atg ggc ttc gag gcc agc att gca gaa gga aag gca 862
Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly Lys Ala
115 120 125 130

gcc tcc tgg ccc tca agg tcc ttg cct gcc cag gag gct gtg gtc aag 910
Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val Val Lys
135 140 145

ctc cgg gtg gag ggc atg acc tgc cag tcc tgt gtc agc tcc att gaa 958
Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser Ile Glu
150 155 160

ggc aag gtc cgg aaa ctg caa gga gta gtg aga gtc aaa gtc tca ctc 1006
Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val Ser Leu
165 170 175

agc aac caa gag gcc gtc atc act tat cag cct tat ctc att cag ccc 1054
Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile Gln Pro
180 185 190

gaa gac ctc agg gac cat gta aat gac atg gga ttt gaa gct gcc atc 1102
Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala Ala Ile
195 200 205 210

aag agc aaa gtg gct ccc tta agc ctg gga cca att gat att gag cgg 1150
Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Asp Ile Glu Arg
215 220 225

tta caa agc act aac cca aag aga cct tta tct tct gct aac cag aat 1198
Leu Gln Ser Thr Asn Pro Lys Arg Pro Leu Ser Ser Ala Asn Gln Asn
230 235 240

ttt aat aat tct gag acc ttg ggg cac caa gga agc cat gtg gtc acc 1246

```
Phe Asn Asn Ser Glu Thr Leu Gly His Gln Gly Ser His Val Val Thr
        245                 250                 255

ctc caa ctg aga ata gat gga atg cat tgt aag tct tgc gtc ttg aat    1294
Leu Gln Leu Arg Ile Asp Gly Met His Cys Lys Ser Cys Val Leu Asn
    260                 265                 270

att gaa gaa aat att ggc cag ctc cta ggg gtt caa agt att caa gtg    1342
Ile Glu Glu Asn Ile Gly Gln Leu Leu Gly Val Gln Ser Ile Gln Val
275                 280                 285                 290

tcc ttg gag aac aaa act gcc caa gta aag tat gac cct tct tgt acc    1390
Ser Leu Glu Asn Lys Thr Ala Gln Val Lys Tyr Asp Pro Ser Cys Thr
                295                 300                 305

agc cca gtg gct ctg cag agg gct atc gag gca ctt cca cct ggg aat    1438
Ser Pro Val Ala Leu Gln Arg Ala Ile Glu Ala Leu Pro Pro Gly Asn
            310                 315                 320

ttt aaa gtt tct ctt cct gat gga gcc gaa ggg agt ggg aca gat cac    1486
Phe Lys Val Ser Leu Pro Asp Gly Ala Glu Gly Ser Gly Thr Asp His
        325                 330                 335

agg tct tcc agt tct cat tcc cct ggc tcc cca ccg aga aac cag gtc    1534
Arg Ser Ser Ser Ser His Ser Pro Gly Ser Pro Pro Arg Asn Gln Val
    340                 345                 350

cag ggc aca tgc agt acc act ctg att gcc att gcc ggc atg acc tgt    1582
Gln Gly Thr Cys Ser Thr Thr Leu Ile Ala Ile Ala Gly Met Thr Cys
355                 360                 365                 370

gca tcc tgt gtc cat tcc att gaa ggc atg atc tcc caa ctg gaa ggg    1630
Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu Glu Gly
                375                 380                 385

gtg cag caa ata tcg gtg tct ttg gcc gaa ggg act gca aca gtt ctt    1678
Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr Val Leu
            390                 395                 400

tat aat ccc tct gta att agc cca gaa gaa ctc aga gct gct ata gaa    1726
Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala Ile Glu
        405                 410                 415

gac atg gga ttt gag gct tca gtc gtt tct gaa agc tgt tct act aac    1774
Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser Thr Asn
    420                 425                 430

cct ctt gga aac cac agt gct ggg aat tcc atg gtg caa act aca gat    1822
Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr Thr Asp
435                 440                 445                 450

ggt aca cct aca tct gtg cag gaa gtg gct ccc cac act ggg agg ctc    1870
Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly Arg Leu
                455                 460                 465

cct gca aac cat gcc ccg gac atc ttg gca aag tcc cca caa tca acc    1918
Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln Ser Thr
            470                 475                 480

aga gca gtg gca ccg cag aag tgc ttc tta cag atc aaa ggc atg acc    1966
Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly Met Thr
        485                 490                 495
```

```
tgt gca tcc tgt gtg tct aac ata gaa agg aat ctg cag aaa gaa gct    2014
Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys Glu Ala
    500                 505                 510

ggt gtt ctc tcc gtg ttg gtt gcc ttg atg gca gga aag gca gag atc    2062
Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala Glu Ile
515                 520                 525                 530

aag tat gac cca gag gtc atc cag ccc ctc gag ata gct cag ttc atc    2110
Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln Phe Ile
                535                 540                 545

cag gac ctg ggt ttt gag gca gca gtc atg gag gac tac gca ggc tcc    2158
Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala Gly Ser
            550                 555                 560

gat ggc aac att gag ctg aca atc aca ggg atg acc tgc gcg tcc tgt    2206
Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala Ser Cys
        565                 570                 575

gtc cac aac ata gag tcc aaa ctc acg agg aca aat ggc atc act tat    2254
Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile Thr Tyr
    580                 585                 590

gcc tcc gtt gcc ctt gcc acc agc aaa gcc ctt gtt aag ttt gac ccg    2302
Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe Asp Pro
595                 600                 605                 610

gaa att atc ggt cca cgg gat att atc aaa att att gag gaa att ggc    2350
Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu Ile Gly
                615                 620                 625

ttt cat gct tcc ctg gcc cag aga aac ccc aac gct cat cac ttg gac    2398
Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His Leu Asp
            630                 635                 640

cac aag atg gaa ata aag cag tgg aag aag tct ttc ctg tgc agc ctg    2446
His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys Ser Leu
        645                 650                 655

gtg ttt ggc atc cct gtc atg gcc tta atg atc tat atg ctg ata ccc    2494
Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu Ile Pro
    660                 665                 670

agc aac gag ccc cac cag tcc atg gtc ctg gac cac aac atc att cca    2542
Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile Ile Pro
675                 680                 685                 690

gga ctg tcc att cta aat ctc atc ttc ttt atc ttg tgt acc ttt gtc    2590
Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr Phe Val
                695                 700                 705

cag ctc ctc ggt ggg tgg tac ttc tac gtt cag gcc tac aaa tct ctg    2638
Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys Ser Leu
            710                 715                 720

aga cac agg tca gcc aac atg gac gtg ctc atc gtc ctg gcc aca agc    2686
Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala Thr Ser
        725                 730                 735

att gct tat gtt tat tct ctg gtc atc ctg gtg gtt gct gtg gct gag    2734
Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val Ala Glu
    740                 745                 750
```

```
aag gcg gag agg agc cct gtg aca ttc ttc gac acg ccc ccc atg ctc    2782
Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro Met Leu
755                 760                 765                 770

ttt gtg ttc att gcc ctg ggc cgg tgg ctg gaa cac ttg gca aag agc    2830
Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala Lys Ser
                775                 780                 785

aaa acc tca gaa gcc ctg gct aaa ctc atg tct ctc caa gcc aca gaa    2878
Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala Thr Glu
            790                 795                 800

gcc acc gtt gtg acc ctt ggt gag gac aat tta atc atc agg gag gag    2926
Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg Glu Glu
        805                 810                 815

caa gtc ccc atg gag ctg gtg cag cgg ggc gat atc gtc aag gtg gtc    2974
Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys Val Val
    820                 825                 830

cct ggg gga aag ttt cca gtg gat ggg aaa gtc ctg gaa ggc aat acc    3022
Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly Asn Thr
835                 840                 845                 850

atg gct gat gag tcc ctc atc aca gga gaa gcc atg cca gtc act aag    3070
Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val Thr Lys
                855                 860                 865

aaa ccc gga agc act gta att gcg ggg tct ata aat gca cat ggc tct    3118
Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His Gly Ser
            870                 875                 880

gtg ctc att aaa gct acc cac gtg ggc aat gac acc act ttg gct cag    3166
Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu Ala Gln
        885                 890                 895

att gtg aaa ctg gtg gaa gag gct cag atg tca aag gca ccc att cag    3214
Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro Ile Gln
    900                 905                 910

cag ctg gct gac cgg ttt agt gga tat ttt gtc cca ttt atc atc atc    3262
Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile Ile Ile
915                 920                 925                 930

atg tca act ttg acg ttg gtg gta tgg att gta atc ggt ttt atc gat    3310
Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe Ile Asp
                935                 940                 945

ttt ggt gtt gtt cag aga tac ttt cct aac ccc aac aag cac atc tcc    3358
Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His Ile Ser
            950                 955                 960

cag aca gag gtg atc atc cgg ttt gct ttc cag acg tcc atc acg gtg    3406
Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile Thr Val
        965                 970                 975

ctg tgc att gcc tgc ccc tgc tcc ctg ggg ctg gcc acg ccc acg gct    3454
Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro Thr Ala
    980                 985                 990

gtc atg gtg ggc acc ggg gtg gcc gcg cag aac ggc atc ctc atc aag    3502
Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu Ile Lys
```

```
995                 1000                1005                1010

gga ggc aag ccc ctg gag atg gcg cac aag ata aag act gtg atg ttt    3550
Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val Met Phe
                1015                1020                1025

gac aag act ggc acc att acc cat ggc gtc ccc agg gtc atg cgg gtg    3598
Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met Arg Val
            1030                1035                1040

ctc ctg ctg ggg gat gtg gcc aca ctg ccc ctc agg aag gtt ctg gct    3646
Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val Leu Ala
        1045                1050                1055

gtg gtg ggg act gcg gag gcc agc agt gaa cac ccc ttg ggc gtg gca    3694
Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly Val Ala
    1060                1065                1070

gtc acc aaa tac tgt aaa gag gaa ctt gga aca gag acc ttg gga tac    3742
Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu Gly Tyr
1075                1080                1085                1090

tgc acg gac ttc cag gca gtg cca ggc tgt gga att ggg tgc aaa gtc    3790
Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys Lys Val
                1095                1100                1105

agc aac gtg gaa ggc atc ctg gcc cac agt gag cgc cct ttg agt gca    3838
Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu Ser Ala
            1110                1115                1120

ccg gcc agt cac ctg aat gag gct ggc agc ctt ccc gca gaa aaa gat    3886
Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu Lys Asp
        1125                1130                1135

gca gtc ccc cag acc ttc tct gtg ctg att gga aac cgt gag tgg ctg    3934
Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu Trp Leu
    1140                1145                1150

agg cgc aac ggt tta acc att tct agc gat gtc agt gac gct atg aca    3982
Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala Met Thr
1155                1160                1165                1170

gac cac gag atg aaa gga cag aca gcc atc ctg gtg gct att gac ggt    4030
Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile Asp Gly
                1175                1180                1185

gtg ctc tgt ggg atg atc gca atc gca gac gct gtc aag cag gag gct    4078
Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln Glu Ala
            1190                1195                1200

gcc ctg gct gtg cac acg ctg cag agc atg ggt gtg gac gtg gtt ctg    4126
Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val Val Leu
        1205                1210                1215

atc acg ggg gac aac cgg aag aca gcc aga gct att gcc acc cag gtt    4174
Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr Gln Val
    1220                1225                1230

ggc atc aac aaa gtc ttt gca gag gtg ctg cct tcg cac aag gtg gcc    4222
Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys Val Ala
1235                1240                1245                1250

aag gtc cag gag ctc cag aat aaa ggg aag aaa gtc gcc atg gtg ggg    4270
```

```
Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met Val Gly
                1255                1260                1265

gat ggg gtc aat gac tcc ccg gcc ttg gcc cag gca gac atg ggt gtg    4318
Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met Gly Val
            1270                1275                1280

gcc att ggc acc ggc acg gat gtg gcc atc gag gca gcc gac gtc gtc    4366
Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp Val Val
        1285                1290                1295

ctt atc aga aat gat ttg ctg gat gtg gtg gct agc att cac ctt tcc    4414
Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His Leu Ser
    1300                1305                1310

aag agg act gtc cga agg ata cgc atc aac ctg gtc ctg gca ctg att    4462
Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala Leu Ile
1315                1320                1325                1330

tat aac ctg gtt ggg ata ccc att gca gca ggt gtc ttc atg ccc atc    4510
Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met Pro Ile
                1335                1340                1345

ggc att gtg ctg cag ccc tgg atg ggc tca gcg gcc atg gca gcc tcc    4558
Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala Ala Ser
            1350                1355                1360

tct gtg tct gtg gtg ctc tca tcc ctg cag ctc aag tgc tat aag aag    4606
Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr Lys Lys
        1365                1370                1375

cct gac ctg gag agg tat gag gca cag gcg cat ggc cac atg aag ccc    4654
Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met Lys Pro
    1380                1385                1390

ctg acg gca tcc cag gtc agt gtg cac ata ggc atg gat gac agg tgg    4702
Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp Arg Trp
1395                1400                1405                1410

cgg gac tcc ccc agg gcc aca cca tgg gac cag gtc agc tat gtc agc    4750
Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr Val Ser
                1415                1420                1425

cag gtg tcg ctg tcc tcc ctg acg tcc gac aag cca tct cgg cac agc    4798
Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg His Ser
            1430                1435                1440

gct gca gca gac gat gat ggg gac aag tgg tct ctg ctc ctg aat ggc    4846
Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu Asn Gly
        1445                1450                1455

agg gat gag gag cag tac atc tga ggtaccaata aagacctctt attttcattc    4900
Arg Asp Glu Glu Gln Tyr Ile
    1460                1465

atcaggtgtg gttggttttt ttgtgtgggg gcggatccat cggatcccgt gcggaccgag    4960

cggccgcagg aacccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc    5020

actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc ggcctcagtg    5080

agcgagcgag cgcgcagctg cctgcag                                        5107
```

<210> 2
<211> 1465
<212> PRT
<213> Artificial Sequence

<220>
<223> [CDS]:473..4870 from SEQ ID NO 1

<400>

<400> 2

Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1 5 10 15
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
20 25 30
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
35 40 45
Asp Gly Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile
50 55 60
Leu Gly Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile
65 70 75 80
Ser Asn Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly
85 90 95
Ser Ala Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val
100 105 110
Cys His Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly
115 120 125
Lys Ala Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val
130 135 140
Val Lys Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser
145 150 155 160
Ile Glu Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val
165 170 175
Ser Leu Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile
180 185 190
Gln Pro Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala
195 200 205
Ala Ile Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Asp Ile
210 215 220
Glu Arg Leu Gln Ser Thr Asn Pro Lys Arg Pro Leu Ser Ser Ala Asn
225 230 235 240
Gln Asn Phe Asn Asn Ser Glu Thr Leu Gly His Gln Gly Ser His Val
245 250 255
Val Thr Leu Gln Leu Arg Ile Asp Gly Met His Cys Lys Ser Cys Val
260 265 270
Leu Asn Ile Glu Glu Asn Ile Gly Gln Leu Leu Gly Val Gln Ser Ile
275 280 285
Gln Val Ser Leu Glu Asn Lys Thr Ala Gln Val Lys Tyr Asp Pro Ser
290 295 300
Cys Thr Ser Pro Val Ala Leu Gln Arg Ala Ile Glu Ala Leu Pro Pro
305 310 315 320
Gly Asn Phe Lys Val Ser Leu Pro Asp Gly Ala Glu Gly Ser Gly Thr
325 330 335
Asp His Arg Ser Ser Ser Ser His Ser Pro Gly Ser Pro Pro Arg Asn
340 345 350
Gln Val Gln Gly Thr Cys Ser Thr Thr Leu Ile Ala Ile Ala Gly Met
355 360 365
Thr Cys Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu
370 375 380
Glu Gly Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr
385 390 395 400
Val Leu Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala
405 410 415
Ile Glu Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser
420 425 430

```
Thr Asn Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr
        435                 440                 445
Thr Asp Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly
    450                 455                 460
Arg Leu Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln
465                 470                 475                 480
Ser Thr Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly
                485                 490                 495
Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys
            500                 505                 510
Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala
        515                 520                 525
Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln
    530                 535                 540
Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala
545                 550                 555                 560
Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala
                565                 570                 575
Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile
            580                 585                 590
Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe
        595                 600                 605
Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu
    610                 615                 620
Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His
625                 630                 635                 640
Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys
                645                 650                 655
Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu
            660                 665                 670
Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile
        675                 680                 685
Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr
    690                 695                 700
Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys
705                 710                 715                 720
Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala
                725                 730                 735
Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val
            740                 745                 750
Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro
        755                 760                 765
Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala
    770                 775                 780
Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala
785                 790                 795                 800
Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg
                805                 810                 815
Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys
            820                 825                 830
Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly
        835                 840                 845
Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val
    850                 855                 860
Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His
865                 870                 875                 880
Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu
                885                 890                 895
Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro
            900                 905                 910
Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile
        915                 920                 925
Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe
```

```
    930                 935                 940
Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His
945                 950                 955                 960
Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile
                965                 970                 975
Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro
            980                 985                 990
Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu
        995                 1000                1005
Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val
    1010                1015                1020
Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met
1025                1030                1035                1040
Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val
                1045                1050                1055
Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly
            1060                1065                1070
Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu
        1075                1080                1085
Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys
    1090                1095                1100
Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu
1105                1110                1115                1120
Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu
                1125                1130                1135
Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu
            1140                1145                1150
Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala
        1155                1160                1165
Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile
    1170                1175                1180
Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln
1185                1190                1195                1200
Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val
                1205                1210                1215
Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr
            1220                1225                1230
Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys
        1235                1240                1245
Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met
    1250                1255                1260
Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met
1265                1270                1275                1280
Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp
                1285                1290                1295
Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His
            1300                1305                1310
Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala
        1315                1320                1325
Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met
    1330                1335                1340
Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala
1345                1350                1355                1360
Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr
                1365                1370                1375
Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met
            1380                1385                1390
Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
        1395                1400                1405
Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr
    1410                1415                1420
Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg
1425                1430                1435                1440
```

His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu
1445 1450 1455
Asn Gly Arg Asp Glu Glu Gln Tyr Ile
1460 1465

<210> 3
<211> 5107
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleic acid construct of expression vector AAV2-AAT-coATP7B

<220>
<221> repeat_region
<222> 1..141
<223> /note="5' ITR of adeno-associated virus serotype 2"

<220>
<221> promoter
<222> 156..460
<223> /note="alpha 1 antitrypsin"

<220>
<221> CDS
<222> 473..4870
<223> /note="Codon optimized sequence encoding ATP7B" /transl_table=1

<220>
<221> polyA_signal
<222> 4877..4932

<220>
<221> repeat_region
<222> 4968..5107
<223> /note="3' ITR of adeno-associated virus serotype 2"

<400> 3

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120

actccatcac taggggttcc tgcggccgca cgcgtcgcca ccccctccac cttggacaca 180

ggacgctgtg gtttctgagc caggtacaat gactcctttc ggtaagtgca gtggaagctg 240

tacactgccc aggcaaagcg tccgggcagc gtaggcgggc gactcagatc ccagccagtg 300

gacttagccc ctgtttgctc ctccgataac tggggtgacc ttggttaata ttcaccagca 360

gcctccccccg ttgcccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc 420

tcctcagctt caggcaccac cactgacctg ggacagtgaa gcggccgcca cc atg cca 478
Met Pro
1

gaa cag gaa cgc cag atc aca gca aga gag gga gca agt cgg aaa atc 526
Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg Lys Ile
5 10 15

ctg agc aaa ctg agc ctg cca acc aga gca tgg gaa ccc gca atg aag 574
Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala Met Lys

```
    20                  25                  30

aaa agc ttc gcc ttt gac aac gtg gga tac gag gga ggg ctg gat gga      622
Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu Asp Gly
35                  40                  45                  50

ctg gga cct agc tcc cag gtg gcc acc tct aca gtc cga atc ctg ggc      670
Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile Leu Gly
                55                  60                  65

atg act tgc cag agt tgc gtg aaa tca att gaa gac cgg atc agt aat      718
Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile Ser Asn
            70                  75                  80

ctg aag gga atc att agc atg aaa gtg tcc ctg gag cag ggc tca gcc      766
Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly Ser Ala
        85                  90                  95

acc gtg aag tat gtc cct agc gtg gtc tgc ctg cag cag gtg tgc cac      814
Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val Cys His
    100                 105                 110

cag atc ggc gat atg ggg ttc gag gcc tcc att gct gaa ggg aaa gcc      862
Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly Lys Ala
115                 120                 125                 130

gct tct tgg cct agc cgg tcc ctg cca gca cag gaa gca gtg gtc aag      910
Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val Val Lys
                135                 140                 145

ctg aga gtg gag gga atg aca tgc cag agc tgc gtg agc agt atc gaa      958
Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser Ile Glu
            150                 155                 160

gga aag gtc cga aaa ctg cag ggc gtg gtc cgg gtg aag gtc tct ctg     1006
Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val Ser Leu
        165                 170                 175

agt aac cag gag gcc gtg att acc tac cag ccc tat ctg atc cag cct     1054
Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile Gln Pro
    180                 185                 190

gaa gac ctg agg gat cac gtg aat gac atg ggc ttc gag gca gcc atc     1102
Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala Ala Ile
195                 200                 205                 210

aag tcc aaa gtg gcc cca ctg tct ctg ggg ccc att gat atc gaa aga     1150
Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Asp Ile Glu Arg
                215                 220                 225

ctg cag tcc acc aac cca aag agg ccc ctg tca agc gcc aac cag aac     1198
Leu Gln Ser Thr Asn Pro Lys Arg Pro Leu Ser Ser Ala Asn Gln Asn
            230                 235                 240

ttc aac aat agt gag acc ctg gga cac cag ggc tca cat gtg gtc aca     1246
Phe Asn Asn Ser Glu Thr Leu Gly His Gln Gly Ser His Val Val Thr
        245                 250                 255

ctg cag ctg agg att gac ggc atg cac tgc aag tct tgc gtg ctg aac     1294
Leu Gln Leu Arg Ile Asp Gly Met His Cys Lys Ser Cys Val Leu Asn
    260                 265                 270

att gag gaa aat atc ggc cag ctg ctg ggg gtg cag tct atc cag gtc     1342
```

```
Ile Glu Glu Asn Ile Gly Gln Leu Leu Gly Val Gln Ser Ile Gln Val
275                 280                 285                 290

agt ctg gag aac aag act gct cag gtg aaa tac gat cct tca tgc acc    1390
Ser Leu Glu Asn Lys Thr Ala Gln Val Lys Tyr Asp Pro Ser Cys Thr
                295                 300                 305

agc cca gtg gca ctg cag cgc gct atc gaa gca ctg ccc cct gga aat    1438
Ser Pro Val Ala Leu Gln Arg Ala Ile Glu Ala Leu Pro Pro Gly Asn
            310                 315                 320

ttc aag gtg agc ctg cct gac gga gca gag gga tcc gga acc gat cac    1486
Phe Lys Val Ser Leu Pro Asp Gly Ala Glu Gly Ser Gly Thr Asp His
        325                 330                 335

agg tcc tct agt tca cat tcc cca ggg tct cca cca cga aac cag gtg    1534
Arg Ser Ser Ser Ser His Ser Pro Gly Ser Pro Pro Arg Asn Gln Val
    340                 345                 350

cag gga aca tgt tcc acc aca ctg att gca atc gcc ggc atg act tgc    1582
Gln Gly Thr Cys Ser Thr Thr Leu Ile Ala Ile Ala Gly Met Thr Cys
355                 360                 365                 370

gcc tca tgc gtg cac agc att gaa ggg atg atc tct cag ctg gag gga    1630
Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu Glu Gly
                375                 380                 385

gtg cag cag atc tca gtc agc ctg gcc gag ggc act gct acc gtg ctg    1678
Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr Val Leu
            390                 395                 400

tac aat ccc agt gtc atc tca cct gag gaa ctg cgg gct gca att gag    1726
Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala Ile Glu
        405                 410                 415

gac atg ggg ttc gaa gct tcc gtg gtc tcc gaa tct tgc agt acc aac    1774
Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser Thr Asn
    420                 425                 430

ccc ctg ggg aat cat tcc gcc gga aac tct atg gtg cag act acc gac    1822
Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr Thr Asp
435                 440                 445                 450

ggg aca cct act tct gtg cag gag gtc gca cca cac aca gga cgc ctg    1870
Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly Arg Leu
                455                 460                 465

cca gcc aat cat gct ccc gat atc ctg gcc aaa agc ccc cag tcc act    1918
Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln Ser Thr
            470                 475                 480

cga gct gtg gca cct cag aag tgt ttt ctg cag atc aaa ggc atg acc    1966
Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly Met Thr
        485                 490                 495

tgc gcc tct tgc gtg agc aac att gag cgg aat ctg cag aag gaa gct    2014
Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys Glu Ala
    500                 505                 510

ggg gtg ctg agc gtg ctg gtc gca ctg atg gcc gga aag gct gag atc    2062
Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala Glu Ile
515                 520                 525                 530
```

```
aag tac gac cct gaa gtg atc cag cca ctg gag att gcc cag ttc atc    2110
Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln Phe Ile
                535                 540                 545

cag gat ctg ggc ttt gag gcc gct gtg atg gaa gac tat gct ggg agc    2158
Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala Gly Ser
            550                 555                 560

gat gga aac att gaa ctg acc atc acc gga atg act tgt gcc tct tgc    2206
Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala Ser Cys
        565                 570                 575

gtg cac aac atc gag agt aaa ctg act aga acc aat ggg att acc tac    2254
Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile Thr Tyr
    580                 585                 590

gcc agt gtg gcc ctg gct aca tca aag gct ctg gtg aaa ttc gac ccc    2302
Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe Asp Pro
595                 600                 605                 610

gag atc att gga cct agg gat atc att aag atc att gag gaa atc ggc    2350
Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu Ile Gly
                615                 620                 625

ttt cac gca agc ctg gcc cag cgc aac cca aat gcc cac cat ctg gac    2398
Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His Leu Asp
            630                 635                 640

cat aag atg gag atc aag cag tgg aag aaa agt ttc ctg tgc tca ctg    2446
His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys Ser Leu
        645                 650                 655

gtg ttt gga atc ccc gtc atg gcc ctg atg atc tac atg ctg atc cct    2494
Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu Ile Pro
    660                 665                 670

agc aac gag cca cac cag tcc atg gtg ctg gat cat aac atc att cct    2542
Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile Ile Pro
675                 680                 685                 690

ggc ctg tcc atc ctg aat ctg att ttc ttt atc ctg tgc aca ttc gtg    2590
Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr Phe Val
                695                 700                 705

cag ctg ctg gga ggc tgg tac ttt tat gtg cag gca tat aaa tca ctg    2638
Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys Ser Leu
            710                 715                 720

cga cac cgg agc gcc aat atg gac gtg ctg att gtc ctg gca acc tct    2686
Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala Thr Ser
        725                 730                 735

atc gcc tac gtg tat agt ctg gtc atc ctg gtg gtc gca gtg gca gag    2734
Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val Ala Glu
    740                 745                 750

aag gca gaa cgg agc cca gtg act ttc ttt gat acc cct cca atg ctg    2782
Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro Met Leu
755                 760                 765                 770

ttc gtg ttt atc gct ctg ggc aga tgg ctg gaa cat ctg gca aag tca    2830
Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala Lys Ser
                775                 780                 785
```

```
aaa acc agc gag gct ctg gca aag ctg atg agc ctg cag gct acc gaa    2878
Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala Thr Glu
            790                 795                 800

gca aca gtg gtc act ctg gga gag gac aac ctg atc att cgc gag gaa    2926
Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg Glu Glu
        805                 810                 815

cag gtg cct atg gaa ctg gtc cag cga ggc gat atc gtg aag gtg gtc    2974
Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys Val Val
    820                 825                 830

cca ggg gga aaa ttc ccc gtg gac ggc aag gtc ctg gag ggg aat act    3022
Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly Asn Thr
835                 840                 845                 850

atg gcc gat gaa tcc ctg atc acc ggc gag gct atg cct gtg aca aag    3070
Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val Thr Lys
                855                 860                 865

aaa cca gga tca act gtc att gct ggc agc atc aac gca cac ggg tcc    3118
Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His Gly Ser
            870                 875                 880

gtg ctg atc aag gcc aca cat gtc ggg aat gac aca act ctg gct cag    3166
Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu Ala Gln
        885                 890                 895

att gtg aaa ctg gtc gag gaa gcc cag atg tcc aag gct cct atc cag    3214
Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro Ile Gln
    900                 905                 910

cag ctg gcc gat cgg ttc tcc ggc tac ttc gtg ccc ttc atc att atc    3262
Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile Ile Ile
915                 920                 925                 930

atg tct aca ctg act ctg gtg gtc tgg att gtg atc gga ttc att gac    3310
Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe Ile Asp
                935                 940                 945

ttt ggc gtg gtc cag aga tat ttt ccc aac cct aat aag cac atc agc    3358
Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His Ile Ser
            950                 955                 960

cag acc gaa gtg atc atc agg ttc gca ttt cag acc agt att aca gtg    3406
Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile Thr Val
        965                 970                 975

ctg tgc atc gcc tgc cca tgt tca ctg ggg ctg gct acc ccc aca gca    3454
Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro Thr Ala
    980                 985                 990

gtg atg gtc gga aca gga gtg gca gca cag aac gga att ctg atc aag    3502
Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu Ile Lys
995                 1000                1005                1010

ggc ggg aaa ccc ctg gag atg gcc cac aag atc aaa act gtg atg ttt    3550
Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val Met Phe
                1015                1020                1025

gac aaa act ggg acc att aca cat gga gtg ccc cgc gtc atg cga gtg    3598
Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met Arg Val
```

```
                1030                1035                1040

ctg ctg ctg ggc gat gtg gca acc ctg cct ctg aga aag gtc ctg gca    3646
Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val Leu Ala
        1045                1050                1055

gtg gtc gga aca gca gag gct agc tcc gaa cac cca ctg ggg gtg gcc    3694
Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly Val Ala
    1060                1065                1070

gtc aca aag tac tgc aaa gag gaa ctg ggc act gag acc ctg ggg tat    3742
Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu Gly Tyr
1075                1080                1085                1090

tgt act gac ttc cag gca gtg ccc gga tgc gga atc gga tgt aaa gtc    3790
Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys Lys Val
                1095                1100                1105

tct aac gtg gaa ggg att ctg gct cac agt gag cgg ccc ctg agc gca    3838
Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu Ser Ala
            1110                1115                1120

cct gca tcc cat ctg aat gaa gca gga agc ctg cca gca gag aag gac    3886
Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu Lys Asp
        1125                1130                1135

gct gtg cct cag acc ttt tcc gtc ctg atc ggc aac aga gaa tgg ctg    3934
Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu Trp Leu
    1140                1145                1150

cgg aga aat ggg ctg aca att tct agt gac gtg tcc gat gcc atg aca    3982
Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala Met Thr
1155                1160                1165                1170

gat cac gag atg aaa ggc cag act gca att ctg gtg gcc atc gac gga    4030
Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile Asp Gly
                1175                1180                1185

gtc ctg tgc ggc atg att gct atc gca gat gcc gtg aag cag gag gct    4078
Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln Glu Ala
            1190                1195                1200

gca ctg gcc gtc cat acc ctg cag tct atg ggc gtg gac gtg gtc ctg    4126
Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val Val Leu
        1205                1210                1215

atc acc ggg gat aac cgg aaa aca gct aga gca att gcc act caa gtg    4174
Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr Gln Val
    1220                1225                1230

ggc atc aat aag gtg ttc gct gaa gtc ctg cct agc cac aag gtc gca    4222
Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys Val Ala
1235                1240                1245                1250

aaa gtg cag gag ctg cag aac aag ggc aag aaa gtc gcc atg gtg gga    4270
Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met Val Gly
                1255                1260                1265

gac ggc gtg aat gat agc cca gct ctg gca cag gca gac atg gga gtc    4318
Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met Gly Val
            1270                1275                1280

gct att ggg aca gga act gac gtg gca atc gag gcc gct gat gtg gtc    4366
```

```
Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp Val Val
        1285                1290                1295

ctg att agg aat gac ctg ctg gat gtg gtc gct tct att cat ctg agt    4414
Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His Leu Ser
    1300                1305                1310

aag agg aca gtg agg cgc att cgc atc aac ctg gtg ctg gcc ctg atc    4462
Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala Leu Ile
1315                1320                1325                1330

tac aat ctg gtg gga att cca atc gca gcc ggc gtg ttt atg cca att    4510
Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met Pro Ile
                1335                1340                1345

ggg atc gtc ctg cag ccc tgg atg ggc tca gct gca atg gcc gct tca    4558
Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala Ala Ser
            1350                1355                1360

agc gtg agc gtg gtc ctg tcc tct ctg cag ctg aaa tgc tac aag aaa    4606
Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr Lys Lys
        1365                1370                1375

cca gac ctg gag cgg tac gaa gct cag gca cac gga cat atg aag ccc    4654
Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met Lys Pro
    1380                1385                1390

ctg acc gct tcc cag gtg tct gtc cac atc ggc atg gac gat aga tgg    4702
Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp Arg Trp
1395                1400                1405                1410

agg gac agc cca agg gcc act cca tgg gat cag gtc agt tac gtg agc    4750
Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr Val Ser
                1415                1420                1425

cag gtc agc ctg agt tca ctg acc agc gac aag ccc tcc cgc cat tct    4798
Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg His Ser
            1430                1435                1440

gca gcc gct gat gac gac ggg gac aag tgg agc ctg ctg ctg aac gga    4846
Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu Asn Gly
        1445                1450                1455

agg gac gaa gaa cag tat atc taa ggtaccaata aagacctctt attttcattc    4900
Arg Asp Glu Glu Gln Tyr Ile
    1460                1465

atcaggtgtg gttggttttt ttgtgtgggg gcggatccat cggatcccgt gcggaccgag    4960

cggccgcagg aacccctagt gatggagttg gccactccct ctctgcgcgc tcgctcgctc    5020

actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc ggcctcagtg    5080

agcgagcgag cgcgcagctg cctgcag                                        5107
```

<210> 4
<211> 1465
<212> PRT
<213> Artificial Sequence

<220>

<223> [CDS]:473..4870 from SEQ ID NO 3

<400> 4

```
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                  10                  15
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
        35                  40                  45
Asp Gly Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile
    50                  55                  60
Leu Gly Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile
65                  70                  75                  80
Ser Asn Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly
                85                  90                  95
Ser Ala Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val
            100                 105                 110
Cys His Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly
        115                 120                 125
Lys Ala Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val
    130                 135                 140
Val Lys Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser
145                 150                 155                 160
Ile Glu Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val
                165                 170                 175
Ser Leu Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile
            180                 185                 190
Gln Pro Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala
        195                 200                 205
Ala Ile Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Asp Ile
    210                 215                 220
Glu Arg Leu Gln Ser Thr Asn Pro Lys Arg Pro Leu Ser Ser Ala Asn
225                 230                 235                 240
Gln Asn Phe Asn Asn Ser Glu Thr Leu Gly His Gln Gly Ser His Val
                245                 250                 255
Val Thr Leu Gln Leu Arg Ile Asp Gly Met His Cys Lys Ser Cys Val
            260                 265                 270
Leu Asn Ile Glu Glu Asn Ile Gly Gln Leu Leu Gly Val Gln Ser Ile
        275                 280                 285
Gln Val Ser Leu Glu Asn Lys Thr Ala Gln Val Lys Tyr Asp Pro Ser
    290                 295                 300
Cys Thr Ser Pro Val Ala Leu Gln Arg Ala Ile Glu Ala Leu Pro Pro
305                 310                 315                 320
Gly Asn Phe Lys Val Ser Leu Pro Asp Gly Ala Glu Gly Ser Gly Thr
                325                 330                 335
Asp His Arg Ser Ser Ser Ser His Ser Pro Gly Ser Pro Pro Arg Asn
            340                 345                 350
Gln Val Gln Gly Thr Cys Ser Thr Thr Leu Ile Ala Ile Ala Gly Met
        355                 360                 365
Thr Cys Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu
    370                 375                 380
Glu Gly Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr
385                 390                 395                 400
Val Leu Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala
                405                 410                 415
Ile Glu Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser
            420                 425                 430
Thr Asn Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr
        435                 440                 445
Thr Asp Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly
    450                 455                 460
Arg Leu Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln
465                 470                 475                 480
Ser Thr Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly
                485                 490                 495
```

```
Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys
                500                 505                 510
Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala
            515                 520                 525
Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln
        530                 535                 540
Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala
545                 550                 555                 560
Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala
                565                 570                 575
Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile
                580                 585                 590
Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe
            595                 600                 605
Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu
        610                 615                 620
Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His
625                 630                 635                 640
Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys
                645                 650                 655
Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu
                660                 665                 670
Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile
            675                 680                 685
Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr
        690                 695                 700
Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys
705                 710                 715                 720
Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala
                725                 730                 735
Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val
                740                 745                 750
Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro
            755                 760                 765
Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala
        770                 775                 780
Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala
785                 790                 795                 800
Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg
                805                 810                 815
Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys
                820                 825                 830
Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly
            835                 840                 845
Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val
        850                 855                 860
Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His
865                 870                 875                 880
Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu
                885                 890                 895
Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro
                900                 905                 910
Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile
            915                 920                 925
Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe
        930                 935                 940
Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His
945                 950                 955                 960
Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile
                965                 970                 975
Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro
                980                 985                 990
Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu
```

```
        995                 1000                1005
Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val
    1010                1015                1020
Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met
1025                1030                1035                1040
Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val
                1045                1050                1055
Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly
            1060                1065                1070
Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu
        1075                1080                1085
Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys
    1090                1095                1100
Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu
1105                1110                1115                1120
Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu
                1125                1130                1135
Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu
            1140                1145                1150
Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala
        1155                1160                1165
Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile
    1170                1175                1180
Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln
1185                1190                1195                1200
Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val
                1205                1210                1215
Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr
            1220                1225                1230
Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys
        1235                1240                1245
Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met
    1250                1255                1260
Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met
1265                1270                1275                1280
Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp
                1285                1290                1295
Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His
            1300                1305                1310
Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala
        1315                1320                1325
Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met
    1330                1335                1340
Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala
1345                1350                1355                1360
Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr
                1365                1370                1375
Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met
            1380                1385                1390
Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
        1395                1400                1405
Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr
    1410                1415                1420
Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg
1425                1430                1435                1440
His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu
                1445                1450                1455
Asn Gly Arg Asp Glu Glu Gln Tyr Ile
            1460                1465
```

<210> 5

<211> 695
<212> DNA
<213> Artificial Sequence

<220>
<223> Hybrid promoter EalbPalAT

<220>
<221> enhancer
<222> 1..382
<223> /note="Albumin gene enhancer"

<220>
<221> promoter
<222> 391..695
<223> /note="Alpha 1 antitrypsin promoter"

<400> 5

ctcgaggttc ctagattaca ttacacattc tgcaagcata gcacagagca atgttctact 60

ttaattactt tcattttctt gtatcctcac agcctagaaa ataacctgcg ttacagcatc 120

cactcagtat cccttgagca tgaggtgaca ctacttaaca tagggacgag atggtacttt 180

gtgtctcctg ctctgtcagc agggcacagt acttgctgat accagggaat gtttgttctt 240

aaataccatc attccggacg tgtttgcctt ggccagtttt ccatgtacat gcagaaagaa 300

gtttggactg atcaatacag tcctctgcct ttaaagcaat aggaaaaggc caacttgtct 360

acgtttagta tgtggctgta gatctgtacc cgccaccccc tccaccttgg acacaggacg 420

ctgtggtttc tgagccaggt acaatgactc ctttcggtaa gtgcagtgga agctgtacac 480

tgcccaggca aagcgtccgg gcagcgtagg cgggcgactc agatcccagc cagtggactt 540

agcccctgtt tgctcctccg ataactgggg tgaccttggt taatattcac cagcagcctc 600

ccccgttgcc cctctggatc cactgcttaa atacggacga ggacagggcc ctgtctcctc 660

agcttcaggc accaccactg acctgggaca gtgaa 695

<210> 6
<211> 3817
<212> DNA
<213> Artificial Sequence

<220>
<223> Nucleic acid construct of expression vector AAV2-AAT-ATP7B(d57-486)

<220>
<221> repeat_region
<222> 1..141
<223> /note="5' ITR of adeno-associated virus serotype 2"

<220>
<221> promoter
<222> 156..460

<223> /note="alpha 1 antitrysin"

<220>
<221> CDS
<222> 473..3580
<223> /note="Truncated ATP7B (Copper-transporting ATPase 2)" /note="Sequence encoding a truncated ATP7B (Copper transporting ATPase 2), carrying a deletion of amino acids 57..486" /transl_table=1

<220>
<221> polyA_signal
<222> 3587..3642

<220>
<221> repeat_region
<222> 3678..3817
<223> /note="3' ITR of adeno-associated virus serotype 2"

<400> 6

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgtc 60

gggcgacctt tggtcgcccg gcctcagtga gcgagcgagc gcgcagagag ggagtggcca 120

actccatcac taggggttcc tgcggccgca cgcgtcgcca ccccctccac cttggacaca 180

ggacgctgtg gtttctgagc caggtacaat gactcctttc ggtaagtgca gtggaagctg 240

tacactgccc aggcaaagcg tccgggcagc gtaggcgggc gactcagatc ccagccagtg 300

gacttagccc ctgtttgctc ctccgataac tggggtgacc ttggttaata ttcaccagca 360

gcctcccccg ttgcccctct ggatccactg cttaaatacg gacgaggaca gggccctgtc 420

tcctcagctt caggcaccac cactgacctg gacagtgaa gcggccgcca cc atg cct 478
Met Pro
1

gag cag gag aga cag atc aca gcc aga gaa ggg gcc agt cgg aaa atc 526
Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg Lys Ile
5 10 15

tta tct aag ctt tct ttg cct acc cgt gcc tgg gaa cca gca atg aag 574
Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala Met Lys
20 25 30

aag agt ttt gct ttt gac aat gtt ggc tat gaa ggt ggt ctg gat ggc 622
Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu Asp Gly
35 40 45 50

ctg ggc cct tct tct cag ccg cag aag tgc ttc tta cag atc aaa ggc 670
Leu Gly Pro Ser Ser Gln Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly
55 60 65

atg acc tgt gca tcc tgt gtg tct aac ata gaa agg aat ctg cag aaa 718
Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys
70 75 80

gaa gct ggt gtt ctc tcc gtg ttg gtt gcc ttg atg gca gga aag gca 766
Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala
85 90 95

gag atc aag tat gac cca gag gtc atc cag ccc ctc gag ata gct cag 814
Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln
100 105 110

ttc atc cag gac ctg ggt ttt gag gca gca gtc atg gag gac tac gca 862
Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala

```
115                 120                 125                 130

ggc tcc gat ggc aac att gag ctg aca atc aca ggg atg acc tgc gcg      910
Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala
                135                 140                 145

tcc tgt gtc cac aac ata gag tcc aaa ctc acg agg aca aat ggc atc      958
Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile
            150                 155                 160

act tat gcc tcc gtt gcc ctt gcc acc agc aaa gcc ctt gtt aag ttt     1006
Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe
        165                 170                 175

gac ccg gaa att atc ggt cca cgg gat att atc aaa att att gag gaa     1054
Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu
    180                 185                 190

att ggc ttt cat gct tcc ctg gcc cag aga aac ccc aac gct cat cac     1102
Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His
195                 200                 205                 210

ttg gac cac aag atg gaa ata aag cag tgg aag aag tct ttc ctg tgc     1150
Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys
                215                 220                 225

agc ctg gtg ttt ggc atc cct gtc atg gcc tta atg atc tat atg ctg     1198
Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu
            230                 235                 240

ata ccc agc aac gag ccc cac cag tcc atg gtc ctg gac cac aac atc     1246
Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile
        245                 250                 255

att cca gga ctg tcc att cta aat ctc atc ttc ttt atc ttg tgt acc     1294
Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr
    260                 265                 270

ttt gtc cag ctc ctc ggt ggg tgg tac ttc tac gtt cag gcc tac aaa     1342
Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys
275                 280                 285                 290

tct ctg aga cac agg tca gcc aac atg gac gtg ctc atc gtc ctg gcc     1390
Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala
                295                 300                 305

aca agc att gct tat gtt tat tct ctg gtc atc ctg gtg gtt gct gtg     1438
Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val
            310                 315                 320

gct gag aag gcg gag agg agc cct gtg aca ttc ttc gac acg ccc ccc     1486
Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro
        325                 330                 335

atg ctc ttt gtg ttc att gcc ctg ggc cgg tgg ctg gaa cac ttg gca     1534
Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala
    340                 345                 350

aag agc aaa acc tca gaa gcc ctg gct aaa ctc atg tct ctc caa gcc     1582
Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala
355                 360                 365                 370

aca gaa gcc acc gtt gtg acc ctt ggt gag gac aat tta atc atc agg     1630
```

```
Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg
                375                 380                 385

gag gag caa gtc ccc atg gag ctg gtg cag cgg ggc gat atc gtc aag    1678
Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys
            390                 395                 400

gtg gtc cct ggg gga aag ttt cca gtg gat ggg aaa gtc ctg gaa ggc    1726
Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly
        405                 410                 415

aat acc atg gct gat gag tcc ctc atc aca gga gaa gcc atg cca gtc    1774
Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val
    420                 425                 430

act aag aaa ccc gga agc act gta att gcg ggg tct ata aat gca cat    1822
Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His
435                 440                 445                 450

ggc tct gtg ctc att aaa gct acc cac gtg ggc aat gac acc act ttg    1870
Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu
                455                 460                 465

gct cag att gtg aaa ctg gtg gaa gag gct cag atg tca aag gca ccc    1918
Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro
            470                 475                 480

att cag cag ctg gct gac cgg ttt agt gga tat ttt gtc cca ttt atc    1966
Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile
        485                 490                 495

atc atc atg tca act ttg acg ttg gtg gta tgg att gta atc ggt ttt    2014
Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe
    500                 505                 510

atc gat ttt ggt gtt gtt cag aga tac ttt cct aac ccc aac aag cac    2062
Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His
515                 520                 525                 530

atc tcc cag aca gag gtg atc atc cgg ttt gct ttc cag acg tcc atc    2110
Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile
                535                 540                 545

acg gtg ctg tgc att gcc tgc ccc tgc tcc ctg ggg ctg gcc acg ccc    2158
Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro
            550                 555                 560

acg gct gtc atg gtg ggc acc ggg gtg gcc gcg cag aac ggc atc ctc    2206
Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu
        565                 570                 575

atc aag gga ggc aag ccc ctg gag atg gcg cac aag ata aag act gtg    2254
Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val
    580                 585                 590

atg ttt gac aag act ggc acc att acc cat ggc gtc ccc agg gtc atg    2302
Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met
595                 600                 605                 610

cgg gtg ctc ctg ctg ggg gat gtg gcc aca ctg ccc ctc agg aag gtt    2350
Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val
                615                 620                 625
```

```
ctg gct gtg gtg ggg act gcg gag gcc agc agt gaa cac ccc ttg ggc    2398
Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly
            630                 635                 640

gtg gca gtc acc aaa tac tgt aaa gag gaa ctt gga aca gag acc ttg    2446
Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu
        645                 650                 655

gga tac tgc acg gac ttc cag gca gtg cca ggc tgt gga att ggg tgc    2494
Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys
    660                 665                 670

aaa gtc agc aac gtg gaa ggc atc ctg gcc cac agt gag cgc cct ttg    2542
Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu
675                 680                 685                 690

agt gca ccg gcc agt cac ctg aat gag gct ggc agc ctt ccc gca gaa    2590
Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu
                695                 700                 705

aaa gat gca gtc ccc cag acc ttc tct gtg ctg att gga aac cgt gag    2638
Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu
            710                 715                 720

tgg ctg agg cgc aac ggt tta acc att tct agc gat gtc agt gac gct    2686
Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala
        725                 730                 735

atg aca gac cac gag atg aaa gga cag aca gcc atc ctg gtg gct att    2734
Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile
    740                 745                 750

gac ggt gtg ctc tgt ggg atg atc gca atc gca gac gct gtc aag cag    2782
Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln
755                 760                 765                 770

gag gct gcc ctg gct gtg cac acg ctg cag agc atg ggt gtg gac gtg    2830
Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val
                775                 780                 785

gtt ctg atc acg ggg gac aac cgg aag aca gcc aga gct att gcc acc    2878
Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr
            790                 795                 800

cag gtt ggc atc aac aaa gtc ttt gca gag gtg ctg cct tcg cac aag    2926
Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys
        805                 810                 815

gtg gcc aag gtc cag gag ctc cag aat aaa ggg aag aaa gtc gcc atg    2974
Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met
    820                 825                 830

gtg ggg gat ggg gtc aat gac tcc ccg gcc ttg gcc cag gca gac atg    3022
Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met
835                 840                 845                 850

ggt gtg gcc att ggc acc ggc acg gat gtg gcc atc gag gca gcc gac    3070
Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp
                855                 860                 865

gtc gtc ctt atc aga aat gat ttg ctg gat gtg gtg gct agc att cac    3118
Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His
            870                 875                 880
```

```
ctt tcc aag agg act gtc cga agg ata cgc atc aac ctg gtc ctg gca    3166
Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala
        885                 890                 895

ctg att tat aac ctg gtt ggg ata ccc att gca gca ggt gtc ttc atg    3214
Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met
    900                 905                 910

ccc atc ggc att gtg ctg cag ccc tgg atg ggc tca gcg gcc atg gca    3262
Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala
915                 920                 925                 930

gcc tcc tct gtg tct gtg gtg ctc tca tcc ctg cag ctc aag tgc tat    3310
Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr
                935                 940                 945

aag aag cct gac ctg gag agg tat gag gca cag gcg cat ggc cac atg    3358
Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met
            950                 955                 960

aag ccc ctg acg gca tcc cag gtc agt gtg cac ata ggc atg gat gac    3406
Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
        965                 970                 975

agg tgg cgg gac tcc ccc agg gcc aca cca tgg gac cag gtc agc tat    3454
Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr
    980                 985                 990

gtc agc cag gtg tcg ctg tcc tcc ctg acg tcc gac aag cca tct cgg    3502
Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg
995                 1000                1005                1010

cac agc gct gca gca gac gat gat ggg gac aag tgg tct ctg ctc ctg    3550
His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu
                1015                1020                1025

aat ggc agg gat gag gag cag tac atc tga ggtaccaata aagacctctt      3600
Asn Gly Arg Asp Glu Glu Gln Tyr Ile
            1030                1035

attttcattc atcaggtgtg gttggttttt ttgtgtgggg gcggatccat cggatcccgt  3660

gcggaccgag cggccgcagg aacccctagt gatggagttg gccactccct ctctgcgcgc  3720

tcgctcgctc actgaggccg ggcgaccaaa ggtcgcccga cgcccgggct ttgcccgggc  3780

ggcctcagtg agcgagcgag cgcgcagctg cctgcag                           3817
```

<210> 7
<211> 1035
<212> PRT
<213> Artificial Sequence

<220>
<223> [CDS]:473..3580 from SEQ ID NO 6

<400> 7

```
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                   10                  15
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
```

```
          35                  40                  45
Asp Gly Leu Gly Pro Ser Ser Gln Pro Gln Lys Cys Phe Leu Gln Ile
     50                  55                  60
Lys Gly Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu
65                  70                  75                  80
Gln Lys Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly
                85                  90                  95
Lys Ala Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile
            100                 105                 110
Ala Gln Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp
        115                 120                 125
Tyr Ala Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr
    130                 135                 140
Cys Ala Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn
145                 150                 155                 160
Gly Ile Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val
                165                 170                 175
Lys Phe Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile
            180                 185                 190
Glu Glu Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala
        195                 200                 205
His His Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe
    210                 215                 220
Leu Cys Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr
225                 230                 235                 240
Met Leu Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His
                245                 250                 255
Asn Ile Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu
            260                 265                 270
Cys Thr Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala
        275                 280                 285
Tyr Lys Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val
    290                 295                 300
Leu Ala Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val
305                 310                 315                 320
Ala Val Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr
                325                 330                 335
Pro Pro Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His
            340                 345                 350
Leu Ala Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu
        355                 360                 365
Gln Ala Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile
    370                 375                 380
Ile Arg Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile
385                 390                 395                 400
Val Lys Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu
                405                 410                 415
Glu Gly Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met
            420                 425                 430
Pro Val Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn
        435                 440                 445
Ala His Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr
    450                 455                 460
Thr Leu Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys
465                 470                 475                 480
Ala Pro Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro
                485                 490                 495
Phe Ile Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile
            500                 505                 510
Gly Phe Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn
        515                 520                 525
Lys His Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr
    530                 535                 540
```

```
Ser Ile Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala
545                 550                 555                 560
Thr Pro Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly
                565                 570                 575
Ile Leu Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys
            580                 585                 590
Thr Val Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg
        595                 600                 605
Val Met Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg
    610                 615                 620
Lys Val Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro
625                 630                 635                 640
Leu Gly Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu
                645                 650                 655
Thr Leu Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile
            660                 665                 670
Gly Cys Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg
        675                 680                 685
Pro Leu Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro
    690                 695                 700
Ala Glu Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn
705                 710                 715                 720
Arg Glu Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser
                725                 730                 735
Asp Ala Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val
            740                 745                 750
Ala Ile Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val
        755                 760                 765
Lys Gln Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val
    770                 775                 780
Asp Val Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile
785                 790                 795                 800
Ala Thr Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser
                805                 810                 815
His Lys Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val
            820                 825                 830
Ala Met Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala
        835                 840                 845
Asp Met Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala
    850                 855                 860
Ala Asp Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser
865                 870                 875                 880
Ile His Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val
                885                 890                 895
Leu Ala Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val
            900                 905                 910
Phe Met Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala
        915                 920                 925
Met Ala Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys
    930                 935                 940
Cys Tyr Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly
945                 950                 955                 960
His Met Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met
                965                 970                 975
Asp Asp Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val
            980                 985                 990
Ser Tyr Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro
        995                 1000                1005
Ser Arg His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu
    1010                1015                1020
Leu Leu Asn Gly Arg Asp Glu Glu Gln Tyr Ile
1025                1030                1035
```

<210> 8
<211> 3108
<212> DNA
<213> Artificial Sequence

<220>
<223> Codon optimized sequence coATP7B(d57-486) encoding truncated ATP7B, with deletion of amino acids 57-486

<220>
<221> CDS
<222> 1..3108
<223> /transl_table=1

<400> 8

```
atg cca gaa cag gaa cgc cag atc aca gca aga gag gga gca agt cgg      48
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                  10                  15

aaa atc ctg agc aaa ctg agc ctg cca acc aga gca tgg gaa ccc gca      96
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30

atg aag aaa agc ttc gcc ttt gac aac gtg gga tac gag gga ggg ctg     144
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
        35                  40                  45

gat gga ctg gga cct agc tcc cag cct cag aag tgt ttc ctg cag atc     192
Asp Gly Leu Gly Pro Ser Ser Gln Pro Gln Lys Cys Phe Leu Gln Ile
    50                  55                  60

aaa ggc atg acc tgc gcc tct tgc gtg agc aac att gag cgg aat ctg     240
Lys Gly Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu
65                  70                  75                  80

cag aag gaa gct ggg gtg ctg agc gtg ctg gtc gca ctg atg gcc gga     288
Gln Lys Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly
                85                  90                  95

aag gct gag atc aag tac gac cct gaa gtg atc cag cca ctg gag att     336
Lys Ala Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile
            100                 105                 110

gcc cag ttc atc cag gat ctg ggc ttt gag gcc gct gtg atg gaa gac     384
Ala Gln Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp
        115                 120                 125

tat gct ggg agc gat gga aac att gaa ctg acc atc acc gga atg act     432
Tyr Ala Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr
    130                 135                 140

tgt gcc tct tgc gtg cac aac atc gag agt aaa ctg act aga acc aat     480
Cys Ala Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn
145                 150                 155                 160

ggg att acc tac gcc agt gtg gcc ctg gct aca tca aag gct ctg gtg     528
Gly Ile Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val
                165                 170                 175

aaa ttc gac ccc gag atc att gga cct agg gat atc att aag atc att     576
Lys Phe Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile
            180                 185                 190
```

```
gag gaa atc ggc ttt cac gca agc ctg gcc cag cgc aac cca aat gcc     624
Glu Glu Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala
        195                 200                 205

cac cat ctg gac cat aag atg gag atc aag cag tgg aag aaa agt ttc     672
His His Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe
    210                 215                 220

ctg tgc tca ctg gtg ttt gga atc ccc gtc atg gcc ctg atg atc tac     720
Leu Cys Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr
225                 230                 235                 240

atg ctg atc cct agc aac gag cca cac cag tcc atg gtg ctg gat cat     768
Met Leu Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His
                245                 250                 255

aac atc att cct ggc ctg tcc atc ctg aat ctg att ttc ttt atc ctg     816
Asn Ile Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu
            260                 265                 270

tgc aca ttc gtg cag ctg ctg gga ggc tgg tac ttt tat gtg cag gca     864
Cys Thr Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala
        275                 280                 285

tat aaa tca ctg cga cac cgg agc gcc aat atg gac gtg ctg att gtc     912
Tyr Lys Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val
    290                 295                 300

ctg gca acc tct atc gcc tac gtg tat agt ctg gtc atc ctg gtg gtc     960
Leu Ala Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val
305                 310                 315                 320

gca gtg gca gag aag gca gaa cgg agc cca gtg act ttc ttt gat acc    1008
Ala Val Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr
                325                 330                 335

cct cca atg ctg ttc gtg ttt atc gct ctg ggc aga tgg ctg gaa cat    1056
Pro Pro Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His
            340                 345                 350

ctg gca aag tca aaa acc agc gag gct ctg gca aag ctg atg agc ctg    1104
Leu Ala Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu
        355                 360                 365

cag gct acc gaa gca aca gtg gtc act ctg gga gag gac aac ctg atc    1152
Gln Ala Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile
    370                 375                 380

att cgc gag gaa cag gtg cct atg gaa ctg gtc cag cga ggc gat atc    1200
Ile Arg Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile
385                 390                 395                 400

gtg aag gtg gtc cca ggg gga aaa ttc ccc gtg gac ggc aag gtc ctg    1248
Val Lys Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu
                405                 410                 415

gag ggg aat act atg gcc gat gaa tcc ctg atc acc ggc gag gct atg    1296
Glu Gly Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met
            420                 425                 430

cct gtg aca aag aaa cca gga tca act gtc att gct ggc agc atc aac    1344
Pro Val Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn
        435                 440                 445
```

```
gca cac ggg tcc gtg ctg atc aag gcc aca cat gtc ggg aat gac aca    1392
Ala His Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr
    450                 455                 460

act ctg gct cag att gtg aaa ctg gtc gag gaa gcc cag atg tcc aag    1440
Thr Leu Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys
465                 470                 475                 480

gct cct atc cag cag ctg gcc gat cgg ttc tcc ggc tac ttc gtg ccc    1488
Ala Pro Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro
                485                 490                 495

ttc atc att atc atg tct aca ctg act ctg gtg gtc tgg att gtg atc    1536
Phe Ile Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile
            500                 505                 510

gga ttc att gac ttt ggc gtg gtc cag aga tat ttt ccc aac cct aat    1584
Gly Phe Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn
        515                 520                 525

aag cac atc agc cag acc gaa gtg atc atc agg ttc gca ttt cag acc    1632
Lys His Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr
    530                 535                 540

agt att aca gtg ctg tgc atc gcc tgc cca tgt tca ctg ggg ctg gct    1680
Ser Ile Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala
545                 550                 555                 560

acc ccc aca gca gtg atg gtc gga aca gga gtg gca gca cag aac gga    1728
Thr Pro Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly
                565                 570                 575

att ctg atc aag ggc ggg aaa ccc ctg gag atg gcc cac aag atc aaa    1776
Ile Leu Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys
            580                 585                 590

act gtg atg ttt gac aaa act ggg acc att aca cat gga gtg ccc cgc    1824
Thr Val Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg
        595                 600                 605

gtc atg cga gtg ctg ctg ctg ggc gat gtg gca acc ctg cct ctg aga    1872
Val Met Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg
    610                 615                 620

aag gtc ctg gca gtg gtc gga aca gca gag gct agc tcc gaa cac cca    1920
Lys Val Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro
625                 630                 635                 640

ctg ggg gtg gcc gtc aca aag tac tgc aaa gag gaa ctg ggc act gag    1968
Leu Gly Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu
                645                 650                 655

acc ctg ggg tat tgt act gac ttc cag gca gtg ccc gga tgc gga atc    2016
Thr Leu Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile
            660                 665                 670

gga tgt aaa gtc tct aac gtg gaa ggg att ctg gct cac agt gag cgg    2064
Gly Cys Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg
        675                 680                 685

ccc ctg agc gca cct gca tcc cat ctg aat gaa gca gga agc ctg cca    2112
Pro Leu Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro
```

```
     690                 695                 700

gca gag aag gac gct gtg cct cag acc ttt tcc gtc ctg atc ggc aac    2160
Ala Glu Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn
705                 710                 715                 720

aga gaa tgg ctg cgg aga aat ggg ctg aca att tct agt gac gtg tcc    2208
Arg Glu Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser
                725                 730                 735

gat gcc atg aca gat cac gag atg aaa ggc cag act gca att ctg gtg    2256
Asp Ala Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val
            740                 745                 750

gcc atc gac gga gtc ctg tgc ggc atg att gct atc gca gat gcc gtg    2304
Ala Ile Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val
        755                 760                 765

aag cag gag gct gca ctg gcc gtc cat acc ctg cag tct atg ggc gtg    2352
Lys Gln Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val
    770                 775                 780

gac gtg gtc ctg atc acc ggg gat aac cgg aaa aca gct aga gca att    2400
Asp Val Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile
785                 790                 795                 800

gcc act caa gtg ggc atc aat aag gtg ttc gct gaa gtc ctg cct agc    2448
Ala Thr Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser
                805                 810                 815

cac aag gtc gca aaa gtg cag gag ctg cag aac aag ggc aag aaa gtc    2496
His Lys Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val
            820                 825                 830

gcc atg gtg gga gac ggc gtg aat gat agc cca gct ctg gca cag gca    2544
Ala Met Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala
        835                 840                 845

gac atg gga gtc gct att ggg aca gga act gac gtg gca atc gag gcc    2592
Asp Met Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala
    850                 855                 860

gct gat gtg gtc ctg att agg aat gac ctg ctg gat gtg gtc gct tct    2640
Ala Asp Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser
865                 870                 875                 880

att cat ctg agt aag agg aca gtg agg cgc att cgc atc aac ctg gtg    2688
Ile His Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val
                885                 890                 895

ctg gcc ctg atc tac aat ctg gtg gga att cca atc gca gcc ggc gtg    2736
Leu Ala Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val
            900                 905                 910

ttt atg cca att ggg atc gtc ctg cag ccc tgg atg ggc tca gct gca    2784
Phe Met Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala
        915                 920                 925

atg gcc gct tca agc gtg agc gtg gtc ctg tcc tct ctg cag ctg aaa    2832
Met Ala Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys
    930                 935                 940

tgc tac aag aaa cca gac ctg gag cgg tac gaa gct cag gca cac gga    2880
```

```
Cys Tyr Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly
945                 950                 955                 960

cat atg aag ccc ctg acc gct tcc cag gtg tct gtc cac atc ggc atg    2928
His Met Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met
                965                 970                 975

gac gat aga tgg agg gac agc cca agg gcc act cca tgg gat cag gtc    2976
Asp Asp Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val
            980                 985                 990

agt tac gtg agc cag gtc agc ctg agt tca ctg acc agc gac aag ccc    3024
Ser Tyr Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro
        995                 1000                1005

tcc cgc cat tct gca gcc gct gat gac gac ggg gac aag tgg agc ctg    3072
Ser Arg His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu
    1010                1015                1020

ctg ctg aac gga agg gac gaa gaa cag tat atc taa                    3108
Leu Leu Asn Gly Arg Asp Glu Glu Gln Tyr Ile
1025                1030                1035
```

<210> 9
<211> 1035
<212> PRT
<213> Artificial Sequence

<220>
<223> [CDS] :1..3108 from SEQ ID NO 8

<400> 9

```
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                  10                  15
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
        35                  40                  45
Asp Gly Leu Gly Pro Ser Ser Gln Pro Gln Lys Cys Phe Leu Gln Ile
    50                  55                  60
Lys Gly Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu
65                  70                  75                  80
Gln Lys Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly
                85                  90                  95
Lys Ala Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile
            100                 105                 110
Ala Gln Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp
        115                 120                 125
Tyr Ala Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr
    130                 135                 140
Cys Ala Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn
145                 150                 155                 160
Gly Ile Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val
                165                 170                 175
Lys Phe Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile
            180                 185                 190
Glu Glu Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala
        195                 200                 205
His His Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe
    210                 215                 220
Leu Cys Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr
225                 230                 235                 240
Met Leu Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His
```

245 250 255
Asn Ile Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu
260 265 270
Cys Thr Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala
275 280 285
Tyr Lys Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val
290 295 300
Leu Ala Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val
305 310 315 320
Ala Val Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr
325 330 335
Pro Pro Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His
340 345 350
Leu Ala Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu
355 360 365
Gln Ala Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile
370 375 380
Ile Arg Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile
385 390 395 400
Val Lys Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu
405 410 415
Glu Gly Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met
420 425 430
Pro Val Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn
435 440 445
Ala His Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr
450 455 460
Thr Leu Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys
465 470 475 480
Ala Pro Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro
485 490 495
Phe Ile Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile
500 505 510
Gly Phe Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn
515 520 525
Lys His Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr
530 535 540
Ser Ile Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala
545 550 555 560
Thr Pro Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly
565 570 575
Ile Leu Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys
580 585 590
Thr Val Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg
595 600 605
Val Met Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg
610 615 620
Lys Val Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro
625 630 635 640
Leu Gly Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu
645 650 655
Thr Leu Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile
660 665 670
Gly Cys Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg
675 680 685
Pro Leu Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro
690 695 700
Ala Glu Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn
705 710 715 720
Arg Glu Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser
725 730 735
Asp Ala Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val
740 745 750

```
Ala Ile Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val
        755                 760                 765
Lys Gln Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val
    770                 775                 780
Asp Val Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile
785                 790                 795                 800
Ala Thr Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser
                805                 810                 815
His Lys Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val
            820                 825                 830
Ala Met Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala
        835                 840                 845
Asp Met Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala
    850                 855                 860
Ala Asp Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser
865                 870                 875                 880
Ile His Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val
                885                 890                 895
Leu Ala Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val
            900                 905                 910
Phe Met Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala
        915                 920                 925
Met Ala Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys
    930                 935                 940
Cys Tyr Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly
945                 950                 955                 960
His Met Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met
                965                 970                 975
Asp Asp Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val
            980                 985                 990
Ser Tyr Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro
        995                 1000                1005
Ser Arg His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu
    1010                1015                1020
Leu Leu Asn Gly Arg Asp Glu Glu Gln Tyr Ile
1025                1030                1035
```

<210> 10
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer AAT-Forward

<400> 10
ctggtctaga acgcgtcgcc acccccctcca ccttgg 36

<210> 11
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer AAT-Reverse

<400> 11
atcatgatgc ggccgcttca ctgtcccagg tcagtg 36

<210> 12

<211> 3966
<212> DNA
<213> Artificial Sequence

<220>
<223> Sequence encoding truncated ATP7B(d223-366) that encodes ATP7B with a deletion of amino acids 223..366

<220>
<221> CDS
<222> 1..3966
<223> /transl_table=1

<400> 12

```
atg cct gag cag gag aga cag atc aca gcc aga gaa ggg gcc agt cgg      48
Met Pro Glu Gln Glu Arg Gln Ile Thr Ala Arg Glu Gly Ala Ser Arg
1               5                  10                  15

aaa atc tta tct aag ctt tct ttg cct acc cgt gcc tgg gaa cca gca      96
Lys Ile Leu Ser Lys Leu Ser Leu Pro Thr Arg Ala Trp Glu Pro Ala
            20                  25                  30

atg aag aag agt ttt gct ttt gac aat gtt ggc tat gaa ggt ggt ctg     144
Met Lys Lys Ser Phe Ala Phe Asp Asn Val Gly Tyr Glu Gly Gly Leu
        35                  40                  45

gat ggc ctg ggc cct tct tct cag gtg gcc acc agc aca gtc agg atc     192
Asp Gly Leu Gly Pro Ser Ser Gln Val Ala Thr Ser Thr Val Arg Ile
    50                  55                  60

ttg ggc atg act tgc cag tca tgt gtg aag tcc att gag gac agg att     240
Leu Gly Met Thr Cys Gln Ser Cys Val Lys Ser Ile Glu Asp Arg Ile
65                  70                  75                  80

tcc aat ttg aaa ggc atc atc agc atg aag gtt tcc ctg gaa caa ggc     288
Ser Asn Leu Lys Gly Ile Ile Ser Met Lys Val Ser Leu Glu Gln Gly
                85                  90                  95

agt gcc act gtg aaa tat gtg cca tcg gtt gtg tgc ctg caa cag gtt     336
Ser Ala Thr Val Lys Tyr Val Pro Ser Val Val Cys Leu Gln Gln Val
            100                 105                 110

tgc cat caa att ggg gac atg ggc ttc gag gcc agc att gca gaa gga     384
Cys His Gln Ile Gly Asp Met Gly Phe Glu Ala Ser Ile Ala Glu Gly
        115                 120                 125

aag gca gcc tcc tgg ccc tca agg tcc ttg cct gcc cag gag gct gtg     432
Lys Ala Ala Ser Trp Pro Ser Arg Ser Leu Pro Ala Gln Glu Ala Val
    130                 135                 140

gtc aag ctc cgg gtg gag ggc atg acc tgc cag tcc tgt gtc agc tcc     480
Val Lys Leu Arg Val Glu Gly Met Thr Cys Gln Ser Cys Val Ser Ser
145                 150                 155                 160

att gaa ggc aag gtc cgg aaa ctg caa gga gta gtg aga gtc aaa gtc     528
Ile Glu Gly Lys Val Arg Lys Leu Gln Gly Val Val Arg Val Lys Val
                165                 170                 175

tca ctc agc aac caa gag gcc gtc atc act tat cag cct tat ctc att     576
Ser Leu Ser Asn Gln Glu Ala Val Ile Thr Tyr Gln Pro Tyr Leu Ile
            180                 185                 190

cag ccc gaa gac ctc agg gac cat gta aat gac atg gga ttt gaa gct     624
Gln Pro Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala
```

```
                195                 200                 205

gcc atc aag agc aaa gtg gct ccc tta agc ctg gga cca att ggc atg      672
Ala Ile Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Gly Met
    210                 215                 220

acc tgt gca tcc tgt gtc cat tcc att gaa ggc atg atc tcc caa ctg      720
Thr Cys Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu
225                 230                 235                 240

gaa ggg gtg cag caa ata tcg gtg tct ttg gcc gaa ggg act gca aca      768
Glu Gly Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr
                245                 250                 255

gtt ctt tat aat ccc tct gta att agc cca gaa gaa ctc aga gct gct      816
Val Leu Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala
            260                 265                 270

ata gaa gac atg gga ttt gag gct tca gtc gtt tct gaa agc tgt tct      864
Ile Glu Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser
        275                 280                 285

act aac cct ctt gga aac cac agt gct ggg aat tcc atg gtg caa act      912
Thr Asn Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr
    290                 295                 300

aca gat ggt aca cct aca tct gtg cag gaa gtg gct ccc cac act ggg      960
Thr Asp Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly
305                 310                 315                 320

agg ctc cct gca aac cat gcc ccg gac atc ttg gca aag tcc cca caa     1008
Arg Leu Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln
                325                 330                 335

tca acc aga gca gtg gca ccg cag aag tgc ttc tta cag atc aaa ggc     1056
Ser Thr Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly
            340                 345                 350

atg acc tgt gca tcc tgt gtg tct aac ata gaa agg aat ctg cag aaa     1104
Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys
        355                 360                 365

gaa gct ggt gtt ctc tcc gtg ttg gtt gcc ttg atg gca gga aag gca     1152
Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala
    370                 375                 380

gag atc aag tat gac cca gag gtc atc cag ccc ctc gag ata gct cag     1200
Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln
385                 390                 395                 400

ttc atc cag gac ctg ggt ttt gag gca gca gtc atg gag gac tac gca     1248
Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala
                405                 410                 415

ggc tcc gat ggc aac att gag ctg aca atc aca ggg atg acc tgc gcg     1296
Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala
            420                 425                 430

tcc tgt gtc cac aac ata gag tcc aaa ctc acg agg aca aat ggc atc     1344
Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile
        435                 440                 445

act tat gcc tcc gtt gcc ctt gcc acc agc aaa gcc ctt gtt aag ttt     1392
```

```
Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe
    450                 455                 460

gac ccg gaa att atc ggt cca cgg gat att atc aaa att att gag gaa    1440
Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu
465                 470                 475                 480

att ggc ttt cat gct tcc ctg gcc cag aga aac ccc aac gct cat cac    1488
Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His
                485                 490                 495

ttg gac cac aag atg gaa ata aag cag tgg aag aag tct ttc ctg tgc    1536
Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys
            500                 505                 510

agc ctg gtg ttt ggc atc cct gtc atg gcc tta atg atc tat atg ctg    1584
Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu
        515                 520                 525

ata ccc agc aac gag ccc cac cag tcc atg gtc ctg gac cac aac atc    1632
Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile
    530                 535                 540

att cca gga ctg tcc att cta aat ctc atc ttc ttt atc ttg tgt acc    1680
Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr
545                 550                 555                 560

ttt gtc cag ctc ctc ggt ggg tgg tac ttc tac gtt cag gcc tac aaa    1728
Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys
                565                 570                 575

tct ctg aga cac agg tca gcc aac atg gac gtg ctc atc gtc ctg gcc    1776
Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala
            580                 585                 590

aca agc att gct tat gtt tat tct ctg gtc atc ctg gtg gtt gct gtg    1824
Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val
        595                 600                 605

gct gag aag gcg gag agg agc cct gtg aca ttc ttc gac acg ccc ccc    1872
Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro
    610                 615                 620

atg ctc ttt gtg ttc att gcc ctg ggc cgg tgg ctg gaa cac ttg gca    1920
Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala
625                 630                 635                 640

aag agc aaa acc tca gaa gcc ctg gct aaa ctc atg tct ctc caa gcc    1968
Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala
                645                 650                 655

aca gaa gcc acc gtt gtg acc ctt ggt gag gac aat tta atc atc agg    2016
Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg
            660                 665                 670

gag gag caa gtc ccc atg gag ctg gtg cag cgg ggc gat atc gtc aag    2064
Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys
        675                 680                 685

gtg gtc cct ggg gga aag ttt cca gtg gat ggg aaa gtc ctg gaa ggc    2112
Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly
    690                 695                 700
```

```
aat acc atg gct gat gag tcc ctc atc aca gga gaa gcc atg cca gtc    2160
Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val
705                 710                 715                 720

act aag aaa ccc gga agc act gta att gcg ggg tct ata aat gca cat    2208
Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His
                725                 730                 735

ggc tct gtg ctc att aaa gct acc cac gtg ggc aat gac acc act ttg    2256
Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu
            740                 745                 750

gct cag att gtg aaa ctg gtg gaa gag gct cag atg tca aag gca ccc    2304
Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro
        755                 760                 765

att cag cag ctg gct gac cgg ttt agt gga tat ttt gtc cca ttt atc    2352
Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile
    770                 775                 780

atc atc atg tca act ttg acg ttg gtg gta tgg att gta atc ggt ttt    2400
Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe
785                 790                 795                 800

atc gat ttt ggt gtt gtt cag aga tac ttt cct aac ccc aac aag cac    2448
Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His
                805                 810                 815

atc tcc cag aca gag gtg atc atc cgg ttt gct ttc cag acg tcc atc    2496
Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile
            820                 825                 830

acg gtg ctg tgc att gcc tgc ccc tgc tcc ctg ggg ctg gcc acg ccc    2544
Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro
        835                 840                 845

acg gct gtc atg gtg ggc acc ggg gtg gcc gcg cag aac ggc atc ctc    2592
Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu
    850                 855                 860

atc aag gga ggc aag ccc ctg gag atg gcg cac aag ata aag act gtg    2640
Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val
865                 870                 875                 880

atg ttt gac aag act ggc acc att acc cat ggc gtc ccc agg gtc atg    2688
Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met
                885                 890                 895

cgg gtg ctc ctg ctg ggg gat gtg gcc aca ctg ccc ctc agg aag gtt    2736
Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val
            900                 905                 910

ctg gct gtg gtg ggg act gcg gag gcc agc agt gaa cac ccc ttg ggc    2784
Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly
        915                 920                 925

gtg gca gtc acc aaa tac tgt aaa gag gaa ctt gga aca gag acc ttg    2832
Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu
    930                 935                 940

gga tac tgc acg gac ttc cag gca gtg cca ggc tgt gga att ggg tgc    2880
Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys
945                 950                 955                 960
```

```
aaa gtc agc aac gtg gaa ggc atc ctg gcc cac agt gag cgc cct ttg    2928
Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu
                965                 970                 975

agt gca ccg gcc agt cac ctg aat gag gct ggc agc ctt ccc gca gaa    2976
Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu
            980                 985                 990

aaa gat gca gtc ccc cag acc ttc tct gtg ctg att gga aac cgt gag    3024
Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu
        995                 1000                1005

tgg ctg agg cgc aac ggt tta acc att tct agc gat gtc agt gac gct    3072
Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala
    1010                1015                1020

atg aca gac cac gag atg aaa gga cag aca gcc atc ctg gtg gct att    3120
Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile
1025                1030                1035                1040

gac ggt gtg ctc tgt ggg atg atc gca atc gca gac gct gtc aag cag    3168
Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln
                1045                1050                1055

gag gct gcc ctg gct gtg cac acg ctg cag agc atg ggt gtg gac gtg    3216
Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val
            1060                1065                1070

gtt ctg atc acg ggg gac aac cgg aag aca gcc aga gct att gcc acc    3264
Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr
        1075                1080                1085

cag gtt ggc atc aac aaa gtc ttt gca gag gtg ctg cct tcg cac aag    3312
Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys
    1090                1095                1100

gtg gcc aag gtc cag gag ctc cag aat aaa ggg aag aaa gtc gcc atg    3360
Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met
1105                1110                1115                1120

gtg ggg gat ggg gtc aat gac tcc ccg gcc ttg gcc cag gca gac atg    3408
Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met
                1125                1130                1135

ggt gtg gcc att ggc acc ggc acg gat gtg gcc atc gag gca gcc gac    3456
Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp
            1140                1145                1150

gtc gtc ctt atc aga aat gat ttg ctg gat gtg gtg gct agc att cac    3504
Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His
        1155                1160                1165

ctt tcc aag agg act gtc cga agg ata cgc atc aac ctg gtc ctg gca    3552
Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala
    1170                1175                1180

ctg att tat aac ctg gtt ggg ata ccc att gca gca ggt gtc ttc atg    3600
Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met
1185                1190                1195                1200

ccc atc ggc att gtg ctg cag ccc tgg atg ggc tca gcg gcc atg gca    3648
Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala
```

```
                    1205                1210                1215

gcc tcc tct gtg tct gtg gtg ctc tca tcc ctg cag ctc aag tgc tat    3696
Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr
            1220                1225                1230

aag aag cct gac ctg gag agg tat gag gca cag gcg cat ggc cac atg    3744
Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met
        1235                1240                1245

aag ccc ctg acg gca tcc cag gtc agt gtg cac ata ggc atg gat gac    3792
Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
    1250                1255                1260

agg tgg cgg gac tcc ccc agg gcc aca cca tgg gac cag gtc agc tat    3840
Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr
1265                1270                1275                1280

gtc agc cag gtg tcg ctg tcc tcc ctg acg tcc gac aag cca tct cgg    3888
Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg
                1285                1290                1295

cac agc gct gca gca gac gat gat ggg gac aag tgg tct ctg ctc ctg    3936
His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu
            1300                1305                1310

aat ggc agg gat gag gag cag tac atc tga                           3966
Asn Gly Arg Asp Glu Glu Gln Tyr Ile
        1315                1320
```

<210> 13
<211> 1321
<212> PRT
<213> Artificial Sequence

<220>
<223> [CDS]:1..3966 from SEQ ID NO 12

<400> 13

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

```
Gln Pro Glu Asp Leu Arg Asp His Val Asn Asp Met Gly Phe Glu Ala
        195                 200                 205
Ala Ile Lys Ser Lys Val Ala Pro Leu Ser Leu Gly Pro Ile Gly Met
    210                 215                 220
Thr Cys Ala Ser Cys Val His Ser Ile Glu Gly Met Ile Ser Gln Leu
225                 230                 235                 240
Glu Gly Val Gln Gln Ile Ser Val Ser Leu Ala Glu Gly Thr Ala Thr
                245                 250                 255
Val Leu Tyr Asn Pro Ser Val Ile Ser Pro Glu Glu Leu Arg Ala Ala
            260                 265                 270
Ile Glu Asp Met Gly Phe Glu Ala Ser Val Val Ser Glu Ser Cys Ser
        275                 280                 285
Thr Asn Pro Leu Gly Asn His Ser Ala Gly Asn Ser Met Val Gln Thr
    290                 295                 300
Thr Asp Gly Thr Pro Thr Ser Val Gln Glu Val Ala Pro His Thr Gly
305                 310                 315                 320
Arg Leu Pro Ala Asn His Ala Pro Asp Ile Leu Ala Lys Ser Pro Gln
                325                 330                 335
Ser Thr Arg Ala Val Ala Pro Gln Lys Cys Phe Leu Gln Ile Lys Gly
            340                 345                 350
Met Thr Cys Ala Ser Cys Val Ser Asn Ile Glu Arg Asn Leu Gln Lys
        355                 360                 365
Glu Ala Gly Val Leu Ser Val Leu Val Ala Leu Met Ala Gly Lys Ala
    370                 375                 380
Glu Ile Lys Tyr Asp Pro Glu Val Ile Gln Pro Leu Glu Ile Ala Gln
385                 390                 395                 400
Phe Ile Gln Asp Leu Gly Phe Glu Ala Ala Val Met Glu Asp Tyr Ala
                405                 410                 415
Gly Ser Asp Gly Asn Ile Glu Leu Thr Ile Thr Gly Met Thr Cys Ala
            420                 425                 430
Ser Cys Val His Asn Ile Glu Ser Lys Leu Thr Arg Thr Asn Gly Ile
        435                 440                 445
Thr Tyr Ala Ser Val Ala Leu Ala Thr Ser Lys Ala Leu Val Lys Phe
    450                 455                 460
Asp Pro Glu Ile Ile Gly Pro Arg Asp Ile Ile Lys Ile Ile Glu Glu
465                 470                 475                 480
Ile Gly Phe His Ala Ser Leu Ala Gln Arg Asn Pro Asn Ala His His
                485                 490                 495
Leu Asp His Lys Met Glu Ile Lys Gln Trp Lys Lys Ser Phe Leu Cys
            500                 505                 510
Ser Leu Val Phe Gly Ile Pro Val Met Ala Leu Met Ile Tyr Met Leu
        515                 520                 525
Ile Pro Ser Asn Glu Pro His Gln Ser Met Val Leu Asp His Asn Ile
    530                 535                 540
Ile Pro Gly Leu Ser Ile Leu Asn Leu Ile Phe Phe Ile Leu Cys Thr
545                 550                 555                 560
Phe Val Gln Leu Leu Gly Gly Trp Tyr Phe Tyr Val Gln Ala Tyr Lys
                565                 570                 575
Ser Leu Arg His Arg Ser Ala Asn Met Asp Val Leu Ile Val Leu Ala
            580                 585                 590
Thr Ser Ile Ala Tyr Val Tyr Ser Leu Val Ile Leu Val Val Ala Val
        595                 600                 605
Ala Glu Lys Ala Glu Arg Ser Pro Val Thr Phe Phe Asp Thr Pro Pro
    610                 615                 620
Met Leu Phe Val Phe Ile Ala Leu Gly Arg Trp Leu Glu His Leu Ala
625                 630                 635                 640
Lys Ser Lys Thr Ser Glu Ala Leu Ala Lys Leu Met Ser Leu Gln Ala
                645                 650                 655
Thr Glu Ala Thr Val Val Thr Leu Gly Glu Asp Asn Leu Ile Ile Arg
            660                 665                 670
Glu Glu Gln Val Pro Met Glu Leu Val Gln Arg Gly Asp Ile Val Lys
        675                 680                 685
Val Val Pro Gly Gly Lys Phe Pro Val Asp Gly Lys Val Leu Glu Gly
```

```
    690                 695                 700
Asn Thr Met Ala Asp Glu Ser Leu Ile Thr Gly Glu Ala Met Pro Val
705                 710                 715                 720
Thr Lys Lys Pro Gly Ser Thr Val Ile Ala Gly Ser Ile Asn Ala His
                725                 730                 735
Gly Ser Val Leu Ile Lys Ala Thr His Val Gly Asn Asp Thr Thr Leu
            740                 745                 750
Ala Gln Ile Val Lys Leu Val Glu Glu Ala Gln Met Ser Lys Ala Pro
        755                 760                 765
Ile Gln Gln Leu Ala Asp Arg Phe Ser Gly Tyr Phe Val Pro Phe Ile
    770                 775                 780
Ile Ile Met Ser Thr Leu Thr Leu Val Val Trp Ile Val Ile Gly Phe
785                 790                 795                 800
Ile Asp Phe Gly Val Val Gln Arg Tyr Phe Pro Asn Pro Asn Lys His
                805                 810                 815
Ile Ser Gln Thr Glu Val Ile Ile Arg Phe Ala Phe Gln Thr Ser Ile
            820                 825                 830
Thr Val Leu Cys Ile Ala Cys Pro Cys Ser Leu Gly Leu Ala Thr Pro
        835                 840                 845
Thr Ala Val Met Val Gly Thr Gly Val Ala Ala Gln Asn Gly Ile Leu
    850                 855                 860
Ile Lys Gly Gly Lys Pro Leu Glu Met Ala His Lys Ile Lys Thr Val
865                 870                 875                 880
Met Phe Asp Lys Thr Gly Thr Ile Thr His Gly Val Pro Arg Val Met
                885                 890                 895
Arg Val Leu Leu Leu Gly Asp Val Ala Thr Leu Pro Leu Arg Lys Val
            900                 905                 910
Leu Ala Val Val Gly Thr Ala Glu Ala Ser Ser Glu His Pro Leu Gly
        915                 920                 925
Val Ala Val Thr Lys Tyr Cys Lys Glu Glu Leu Gly Thr Glu Thr Leu
    930                 935                 940
Gly Tyr Cys Thr Asp Phe Gln Ala Val Pro Gly Cys Gly Ile Gly Cys
945                 950                 955                 960
Lys Val Ser Asn Val Glu Gly Ile Leu Ala His Ser Glu Arg Pro Leu
                965                 970                 975
Ser Ala Pro Ala Ser His Leu Asn Glu Ala Gly Ser Leu Pro Ala Glu
            980                 985                 990
Lys Asp Ala Val Pro Gln Thr Phe Ser Val Leu Ile Gly Asn Arg Glu
        995                 1000                1005
Trp Leu Arg Arg Asn Gly Leu Thr Ile Ser Ser Asp Val Ser Asp Ala
    1010                1015                1020
Met Thr Asp His Glu Met Lys Gly Gln Thr Ala Ile Leu Val Ala Ile
1025                1030                1035                1040
Asp Gly Val Leu Cys Gly Met Ile Ala Ile Ala Asp Ala Val Lys Gln
                1045                1050                1055
Glu Ala Ala Leu Ala Val His Thr Leu Gln Ser Met Gly Val Asp Val
            1060                1065                1070
Val Leu Ile Thr Gly Asp Asn Arg Lys Thr Ala Arg Ala Ile Ala Thr
        1075                1080                1085
Gln Val Gly Ile Asn Lys Val Phe Ala Glu Val Leu Pro Ser His Lys
    1090                1095                1100
Val Ala Lys Val Gln Glu Leu Gln Asn Lys Gly Lys Lys Val Ala Met
1105                1110                1115                1120
Val Gly Asp Gly Val Asn Asp Ser Pro Ala Leu Ala Gln Ala Asp Met
                1125                1130                1135
Gly Val Ala Ile Gly Thr Gly Thr Asp Val Ala Ile Glu Ala Ala Asp
            1140                1145                1150
Val Val Leu Ile Arg Asn Asp Leu Leu Asp Val Val Ala Ser Ile His
        1155                1160                1165
Leu Ser Lys Arg Thr Val Arg Arg Ile Arg Ile Asn Leu Val Leu Ala
    1170                1175                1180
Leu Ile Tyr Asn Leu Val Gly Ile Pro Ile Ala Ala Gly Val Phe Met
1185                1190                1195                1200
```

```
Pro Ile Gly Ile Val Leu Gln Pro Trp Met Gly Ser Ala Ala Met Ala
                1205                1210                1215
Ala Ser Ser Val Ser Val Val Leu Ser Ser Leu Gln Leu Lys Cys Tyr
            1220                1225                1230
Lys Lys Pro Asp Leu Glu Arg Tyr Glu Ala Gln Ala His Gly His Met
        1235                1240                1245
Lys Pro Leu Thr Ala Ser Gln Val Ser Val His Ile Gly Met Asp Asp
    1250                1255                1260
Arg Trp Arg Asp Ser Pro Arg Ala Thr Pro Trp Asp Gln Val Ser Tyr
1265                1270                1275                1280
Val Ser Gln Val Ser Leu Ser Ser Leu Thr Ser Asp Lys Pro Ser Arg
                1285                1290                1295
His Ser Ala Ala Ala Asp Asp Asp Gly Asp Lys Trp Ser Leu Leu Leu
            1300                1305                1310
Asn Gly Arg Asp Glu Glu Gln Tyr Ile
        1315                1320
```

<210> 14
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer F1

<400> 14
ctagatgcgg ccgccaccat gcctg 25

<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer R1

<400> 15
ctgagaagaa gggcccaggc c 21

<210> 16
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer F2

<400> 16
ggcccttctt ctcagccgca gaagtgcttc ttacag 36

<210> 17
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer R2

<400> 17

accaaaatcg ataaaaccga ttacaatcc 29

<210> 18
<211> 25
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer F3

<400> 18
acgcgtgcgg ccgccaccat gccag 25

<210> 19
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer R3

<400> 19
ctgggagcta ggtcccagtc c 21

<210> 20
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer F4

<400> 20
ggacctagct cccagcctca gaagtgtttt ctgcag 36

<210> 21
<211> 29
<212> DNA
<213> Artificial Sequence

<220>
<223> Primer R4

<400> 21
tgttcctcgc gaatgatcag gttgtcctc 29

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- DE 100156121 A1 **[0011]**
- DE 2003 A1 **[0011]**
- US 5173414 A **[0071]**
- US 5139941 A **[0071]**
- WO 9201070 A **[0071]**
- WO 9303769 A **[0071]**
- WO 2014144229 A **[0091]**
- WO 2007069666 A **[0102]**
- WO 2008118820 A **[0102]**
- WO 0138362 A **[0105]**
- WO 0185938 A **[0106]**
- WO 03076601 A **[0106]**
- WO 2005042728 A **[0106]**
- WO 2006108846 A **[0106]**
- WO 2008129058 A **[0106]**
- WO 2008142124 A **[0106]**
- EP 2295591 A **[0115]**

Non-patent literature cited in the description

- **MERLE et al.** *Current Gene Therapy,* 2007, vol. 7, 217-220 **[0002]**
- **TERADA et al.** *J. Biol. Chem.,* 1998, vol. 273, 1815-1820 **[0009]**
- **TERADA et al.** *FEBS Lett.,* 1999, vol. 448, 53-56 **[0009]**
- **HA-HAO et al.** *Z. Gastroenterol.,* 2002, vol. 40, 209-216 **[0009]**
- **MERLE et al.** *Scan. J. Gastroenterol.,* 2006, vol. 41, 974-982 **[0012]**
- **ROYBAL et al.** *Gene Therapy,* 2012, vol. 19, 1085-1094 **[0012]**
- **HUSTER et al.** *J Biol Chem,* 22 August 2003, vol. 278 (34), 32212-32218 **[0012]**
- **SAMBROOK ; RUSSELL.** Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory Press, 2001 **[0042] [0083]**
- **GREEN ; SAMBROOK.** Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory Press, 2012 **[0042] [0083]**
- Tissue-Specific Promoter Database, TiProD. *Nucleic Acids Research,* 2006, vol. J4, D104-D107 **[0043]**
- **BERNS ; BOHENZKY.** Advances in Virus Research. Academic Press, Inc, 1987, vol. 32, 243-307 **[0070]**
- **MUZYCZKA, N.** *Current Topics in Microbiol. and Immunol.,* 1992, vol. 158, 97-129 **[0070] [0071]**
- **LEBKOWSKI et al.** *Molec. Cell. Biol.,* 1988, vol. 8, 3988-3996 **[0071]**
- **VINCENT et al.** Vaccines 90. Cold Spring Harbor Laboratory Press, 1990 **[0071]**
- **CARTER, B. J.** *Current Opinion in Biotechnology,* 1992, vol. 3, 533-539 **[0071]**
- **KOTIN, R. M.** *Human Gene Therapy,* 1994, vol. 5, 793-801 **[0071]**
- Adenovirus. Methods and Protocols. Springer, 2014 **[0081]**
- **BÜNNING H et al.** *J Gene Med,* 2008, vol. 10, 717-733 **[0087]**
- **NATHWANI et al.** *Blood,* 2007, vol. 109, 1414-1421 **[0088]**
- **KITAJIMA et al.** *Atherosclerosis,* 2006, vol. 186, 65-73 **[0088]**
- **ZINN et al.** *Cell Reports,* 2015, vol. 12, 1-13 **[0089]**
- **SHEN et al.** *J. Biol. Chem.,* 2013, vol. 288 (40), 28814-28823 **[0091]**
- **FILIPPI ; DHAWAN.** *Ann NY Acad Sci.,* 2014, vol. 1315, 50-55 **[0103]**
- **YOSHIDA et al.** *Gastroenterology,* 1996, vol. 111, 1654-1660 **[0103]**
- **IRANI et al.** *Molecular Therapy,* 2001, vol. 3 (3), 302-309 **[0103]**
- **VOGEL et al.** *J Inherit Metab Dis,* 2014, vol. 37, 165-176 **[0103]**
- **MERLE et al.** *Scandinavian Journal of Gastroenterology,* 2006, vol. 41 (8), 974-982 **[0103]**
- **BLOUIN VET.** *J Gene Med.,* 2004, S223-S228 **[0111]**
- **GRIMM D. et al.** *Hum. Gene Ther.,* 2003, vol. 7, 839-850 **[0111]**
- **URABE et al.** *Hum. Gene Ther.,* 2002, vol. 13, 1935-1943 **[0113]**
- **PARKS RJ et al.** *Proc. Natl. Acad. Sci. USA,* 1996, 13565-13570 **[0115]**
- **PALMER.** *Engel. Mol. Ther.,* 2003, vol. 8, 846-852 **[0115]**
- **GONZALEZ-APARICIO et al.** *Gene Therapy,* 2011, vol. 18, 1025-1033 **[0115]**
- Viral Vectors for Gene Therapy, Methods and Protocols. Series: Methods in Molecular Biology. Humana Press (Springer), 2011, vol. 737 **[0116]**
- **M. GIACCA.** Gene Therapy. Springer-Verlag, 2010 **[0117]**

- Viral Vectors for Gene Transfer: Current Status of Gene Therapeutics. **HEILBRONN R. ; WEGER S.** Drug Delivery, Handbook of Experimental Pharmacology. Springer-Verlag, 2010, vol. 197, 143-170 **[0118]**
- Adeno-Associated Virus: Methods and Protocols. Humana Press (Springer), 2011 **[0119]**
- **BÜNNING H. et al.** Recent developments in adeno-associated virus technology. *J. Gene Med.,* 2008, vol. 10, 717-733 **[0120]**
- Adenovirus: Methods and Protocols. Humana Press (Springer), 2014 **[0121]**
- **POLISHCHUCK et al.** *Dev Cell.,* 2014, vol. 29 (6), 686-700 **[0123]**
- **GUPTA ; LUTSENKO.** *Future Med. Chem.,* 2009, vol. 1, 1125-1142 **[0123]**
- **REMINGTON.** The Science and Practice of Pharmacy. Mack Publishing Company, April 1997 **[0131]**
- **LEVITT N. et al.** *Genes & Development,* 1989, vol. 3 (7), 1019-1025 **[0147]**
- **KRAMER M.G. et al.** *Mol. Therapy,* 2003, vol. 7 (3), 375-385 **[0148]**
- **BUIKOVA O.I. et al.** *Human Molecular Genetics,* 1999, vol. 8 (9), 1665-1671 **[0169]**
- **LUTSENKO S.** *Biochemical Society Transactions,* 2008, vol. 36 (6), 1233-1238 **[0169]**
- **SCHOSINSKY ; COLS.** *Clinical Chemistry,* 1974, vol. 20 (12), 1556-1563 **[0173]**
- **DANSCHER G. ; ZIMMER J.** *Histochemistry,* 1978, vol. 55 (1), 27-40 **[0176]**

Patentkrav

1. Nukleinsyrekonstruktion, som omfatter

a) en nukleotidsekvens af en eukaryotisk promoter;

b) en nukleotidsekvens der koder for trunkeret kobber-transporterende ATPase 2 (ATP7B) i hvilken de N-terminale tungmetalassocierede steder HMA 1, HMA 2, HMA 3 og HMA 4 er helt slettet og HMA 5 og HMA 6 forbliver ikke-slettet; og

c) en polyadenyleringssignalsekvens,

d) flankeret af en 5' inverteret terminal repetition (ITR) og 3'ITR af en adenoassocieret virus (AAV).

2. Nukleinsyrekonstruktionen ifølge krav 1, hvor sletningen i det trunkerede ATP7B omfatter aminosyrer 57 til 486 af SEQ.ID.NO.2-sekvensen.

3. Nukleinsyrekonstruktionen ifølge et hvilket som helst af kravene 1 eller 2, hvor aminosyresekvensen af trunkeret ATP7B er SEQ.ID.NO.7.

4. Nukleinsyrekonstruktionen ifølge krav 3, hvor nukleotidsekvensen der koder for det trunkerede ATP7B er valgt fra gruppen bestående af

a) kodningssekvensen CDS af SEQ.ID.NO.6, baser 473..3580;

b) sekvensen SEQ.ID.NO.8; og

c) en sekvens, hvor mindst 827, mindst 879, mindst 931 eller mindst 983 af kodonerne der koder for trunkeret ATP7B er identiske med kodonerne af kodningssekvens SEQ.ID.NO.8.

5. Nukleinsyrekonstruktionen ifølge et hvilket som helst af kravene 1-4, hvor nukleotidsekvensen af den eukaryotiske promoter er en nukleotidsekvens af α1-antitrypsin-genpromoter, eller en kimær promotersekvens der omfatter en α1-antitrypsin-genpromotersekvens kombineret med et albumingenfremmerelement.

6. Nukleinsyrekonstruktionen ifølge et hvilket som helst af kravene 1-5, hvor nukleotidsekvensen af den eukaryotiske promoter er sekvensen afgrænset af baserne 156..460 af SEQ.ID.NO.1 (AAT) eller SEQ.ID.NO.5 (EalbPa1AT).

7. Nukleinsyrekonstruktionen ifølge krav 1, hvor 5'ITR- og 3'ITR-sekvenserne af et AAV er af en serotype valgt fra gruppen bestående af AAV1, AAV2, og AAV4, fortrinsvis er af AAV2-serotypen.

8. Ekspressionsvektor, der omfatter en nukleinsyrekonstruktion ifølge et hvilket som helst af kravene 1-7.

9. Ekspressionsvektoren ifølge krav 9, hvor vektoren er en AAV-vektor.

10. Værtscelle omfattende en nukleinsyrekonstruktion ifølge et hvilket som helst af kravene 1-7, eller en ekspressionsvektor ifølge et hvilket som helst af kravene 8-9.

11. Viruspartikel, der omfatter en nukleinsyrekonstruktion ifølge et hvilket som helst af kravene 1-7 eller en ekspressionsvektor ifølge et hvilket som helst af kravene 8-9.

12. Viruspartiklen ifølge krav 11, hvor viruspartiklen omfatter capsidproteiner af et AAV.

13. Farmaceutisk sammensætning, der omfatter en nukleinsyrekonstruktion ifølge et hvilket som helst af kravene 1-7, en ekspressionsvektor ifølge et hvilket som helst af kravene 8-9, en værtscelle ifølge krav 10 eller en viruspartikel ifølge et hvilket som helst af kravene 11-12 og en farmaceutisk acceptabel bærer.

14. Nukleinsyrekonstruktion, der omfatter

a) en nukleotidsekvens af en eukaryotisk promoter;

b) en nukleotidsekvens der koder for et trunkeret ATP7B i hvilken de N-terminale tungmetalassocierede steder HMA 1, HMA 2, HMA 3 og HMA 4 er helt slettet og HMA 5 og HMA 6 forbliver ikke-slettet; og

c) en polyadenyleringssignalsekvens;

d) flankeret af en 5' inverteret terminal repetition (ITR) og 3'ITR af en adenoassocieret virus (AAV);

en ekspressionsvektor ifølge et hvilket som helst af kravene 8-9, en værtscelle ifølge krav 10 eller en viruspartikel ifølge et hvilket som helst af kravene 11-12, til anvendelse som et medikament.

15. Nukleinsyrekonstruktion, der omfatter

a) en nukleotidsekvens af en eukaryotisk promoter;

b) en nukleotidsekvens der koder for et trunkeret ATP7B i hvilken de N-terminale tungmetalassocierede steder HMA 1, HMA 2, HMA 3 og HMA 4 er helt slettet og HMA 5 og HMA 6 forbliver ikke-slettet; og

c) en polyadenyleringssignalsekvens;

d) flankeret af en 5' inverteret terminal repetition (ITR) og 3'ITR af en adenoassocieret virus (AAV);

en ekspressionsvektor ifølge et hvilket som helst af kravene 8-9, en værtscelle ifølge krav 10 eller en viruspartikel ifølge et hvilket som helst af kravene 11-12 eller en farmaceutisk sammensætning ifølge krav 13, til anvendelse i behandlingen af en tilstand forårsaget af en defekt eller dysfunktion af ATP7B hos mennesker, valgt fra de følgende tilstande: choleostatiske lidelser, Alzheimers sygdom, kræft og Wilsons sygdom.

16. Fremgangsmåde til fremstilling af viruspartikler ifølge et hvilket som helst af kravene 11-12, omfattende trinnene:

a) at dyrke en værtscelle ifølge krav 10 i et dyrkningsmedium; og

b) at høste viruspartiklerne fra cellekultursupernatanten og/eller inden i cellerne.

DRAWINGS

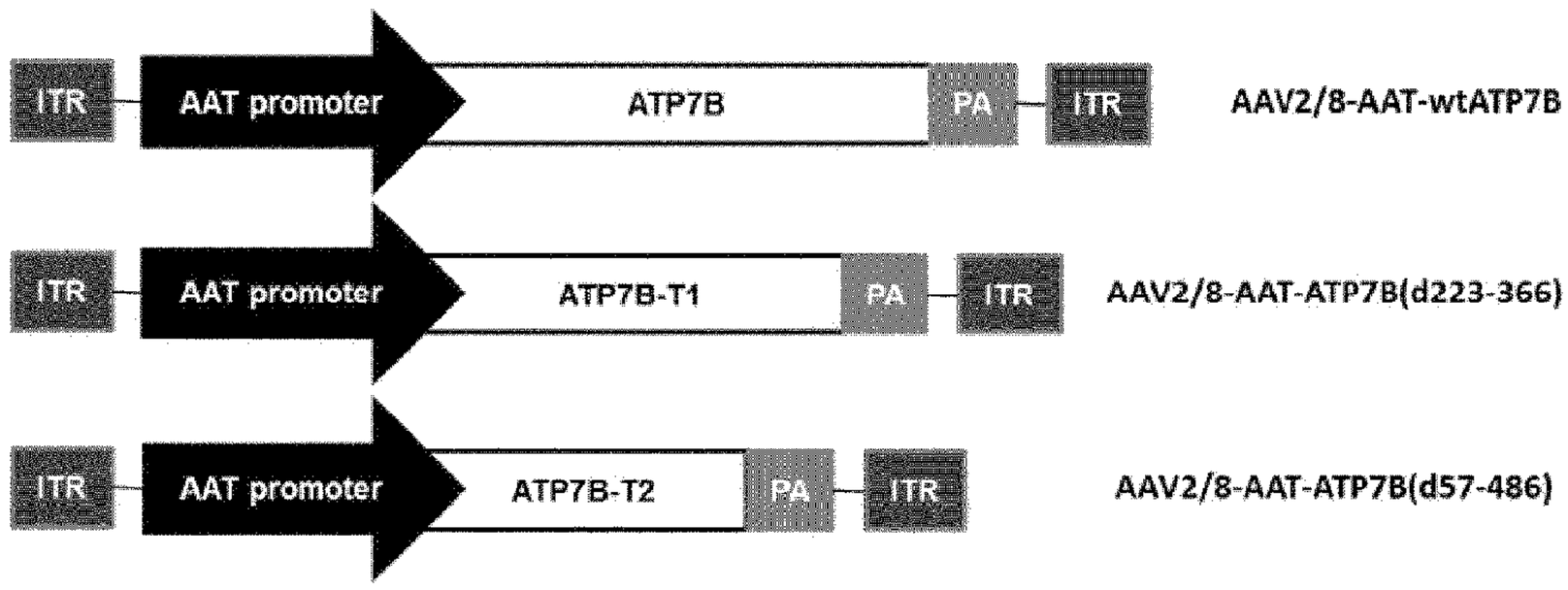


Figure 1

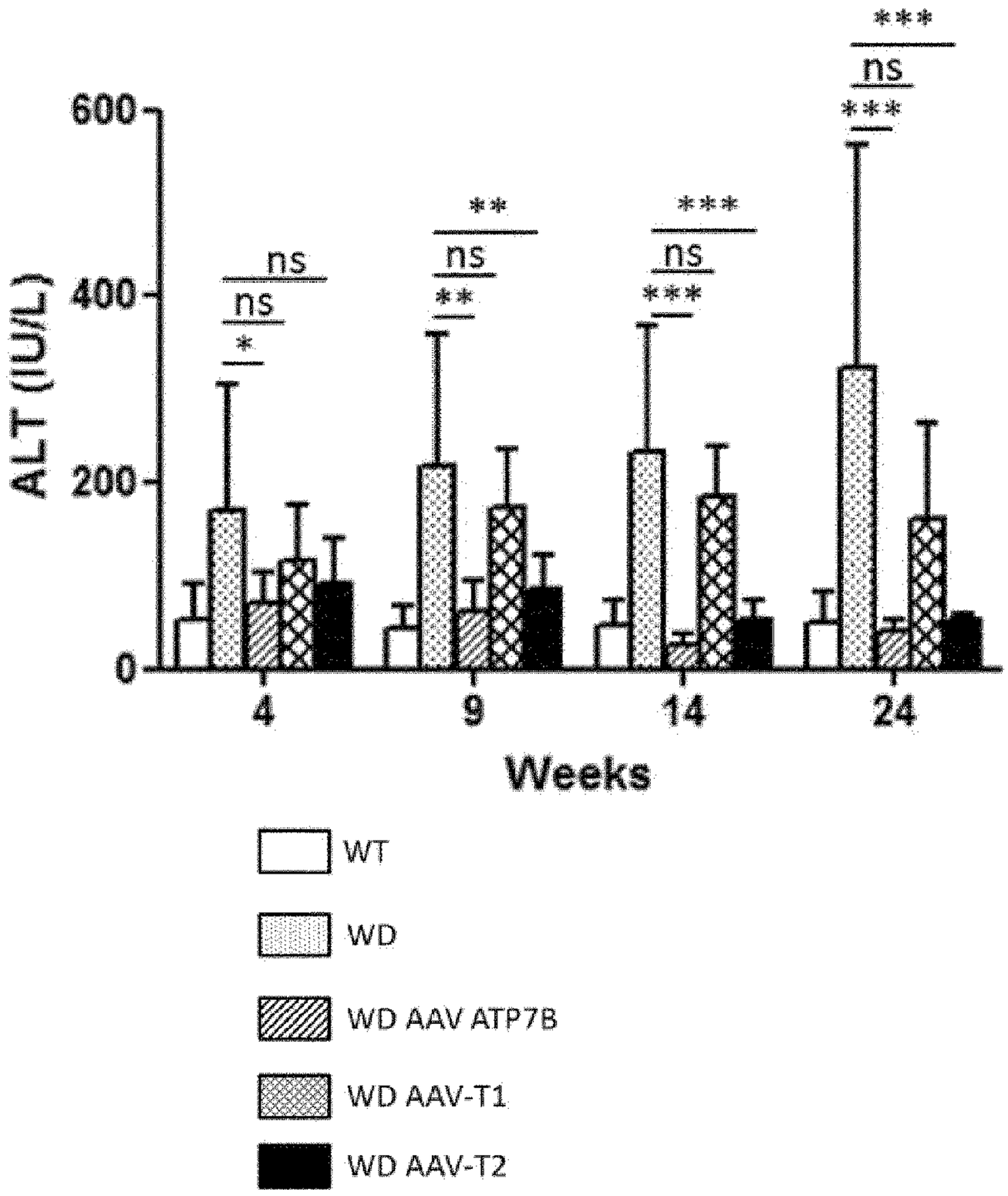

ALT (IU/L)
600
400
200
0
ns
ns
*
**
ns
**
***
ns
***
***
ns
4
9
14
24
Weeks
WT
WD
WD AAV ATP7B
WD AAV-T1
WD AAV-T2

Figure 2

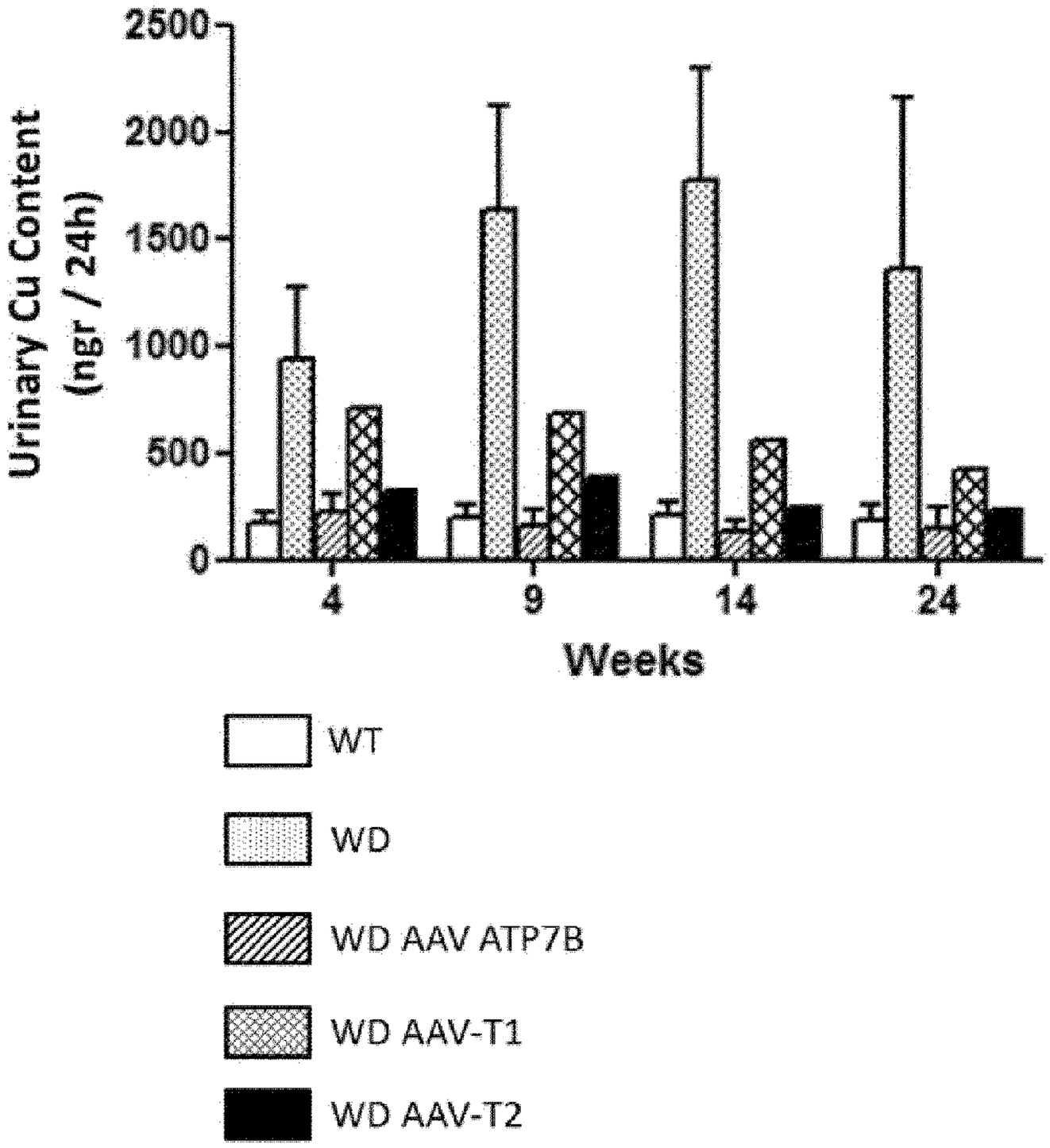

Urinary Cu Content
(ngr / 24h)
2500
2000
1500
1000
500
0
4
9
14
24
Weeks
WT
WD
WD AAV ATP7B
WD AAV-T1
WD AAV-T2

Figure 3

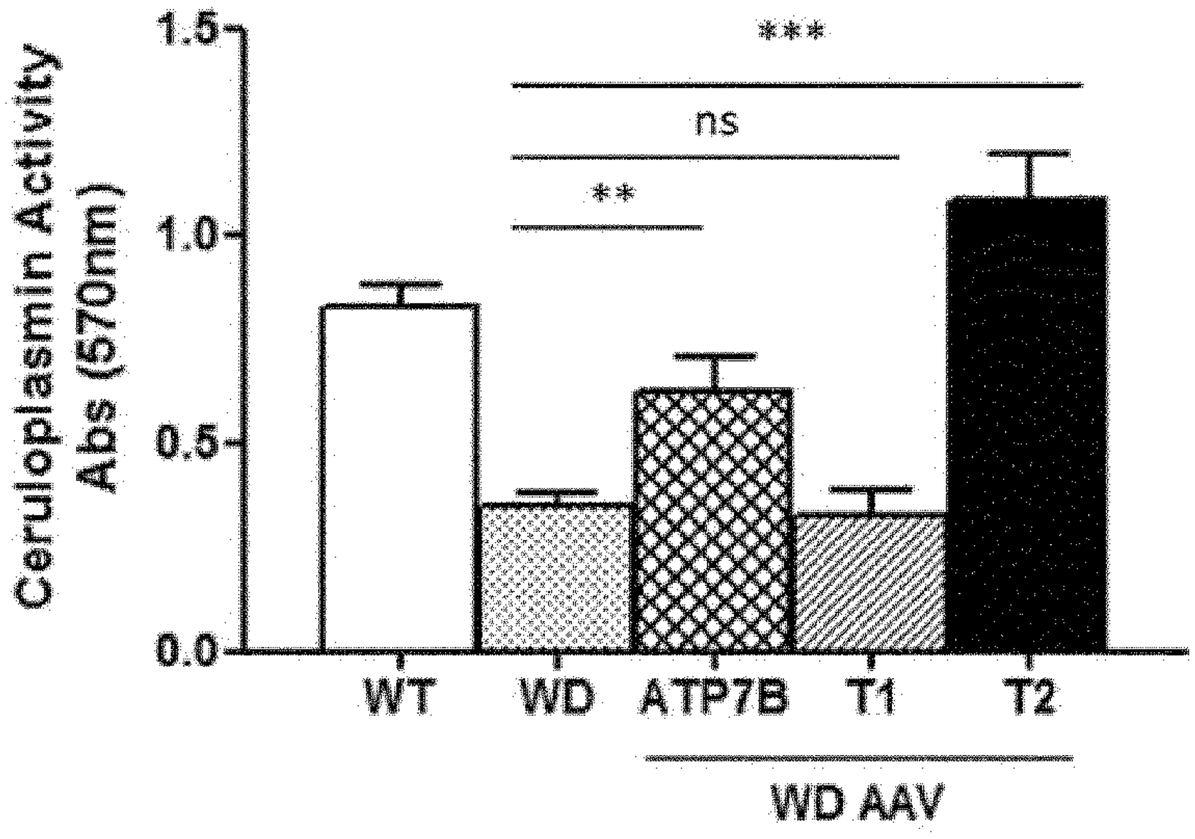

Ceruloplasmin Activity
Abs (570nm)
1.5
1.0
0.5
0.0
***
ns
**
WT
WD
ATP7B
T1
T2
WD AAV

Figure 4

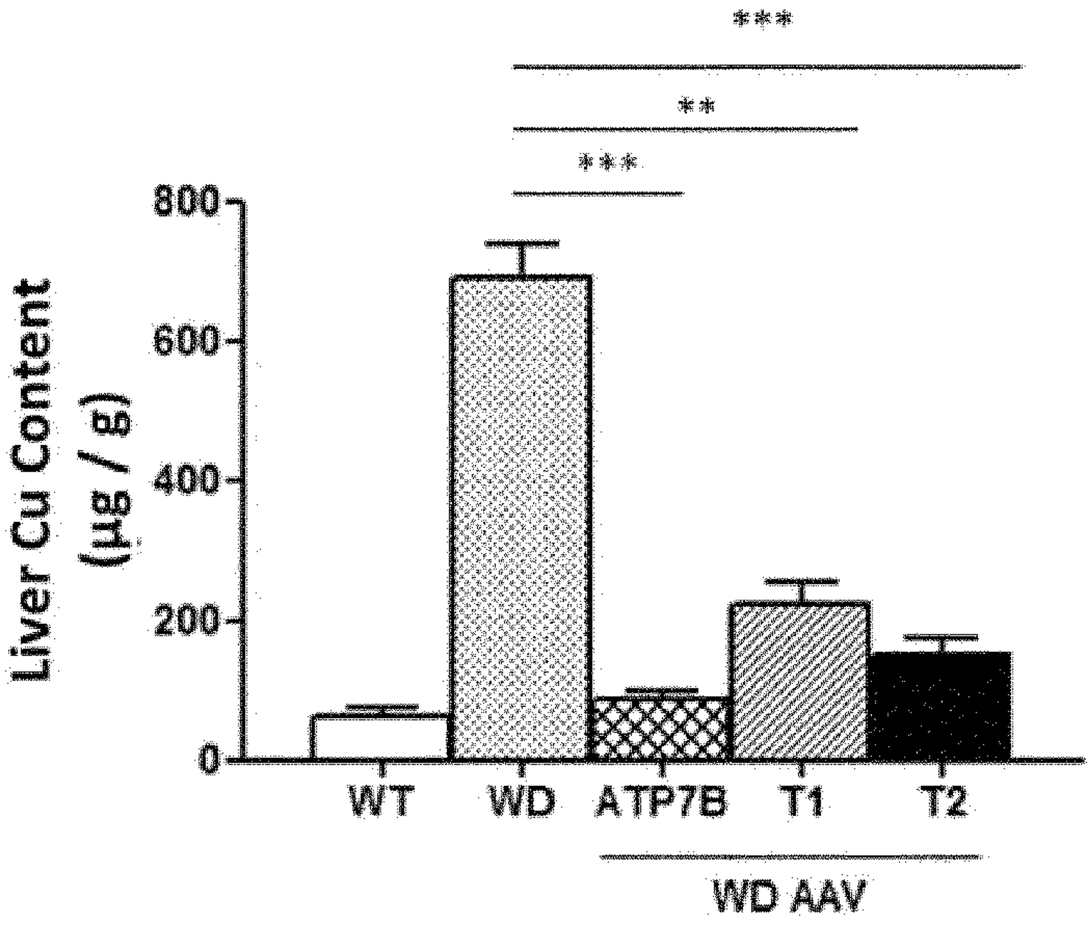


Figure 5

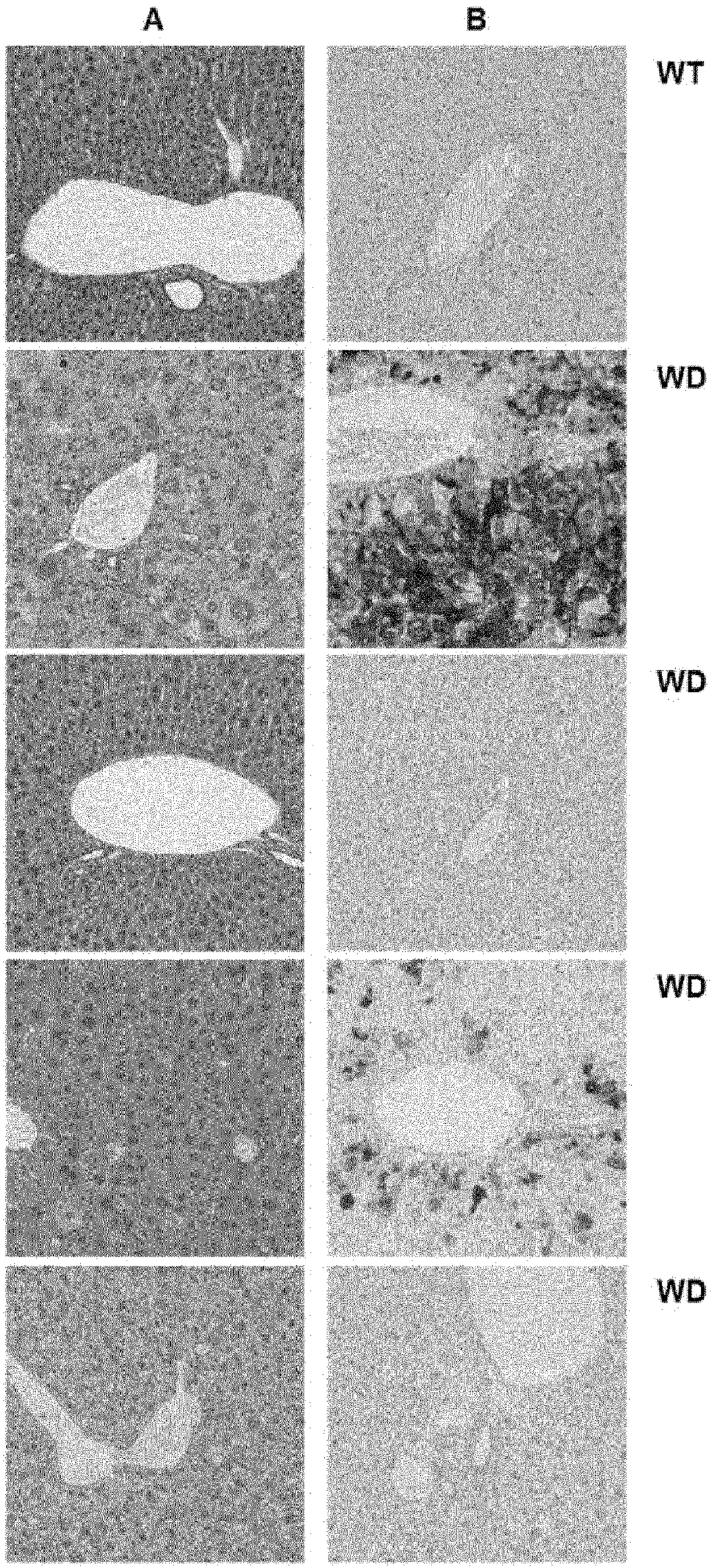


Figure 6

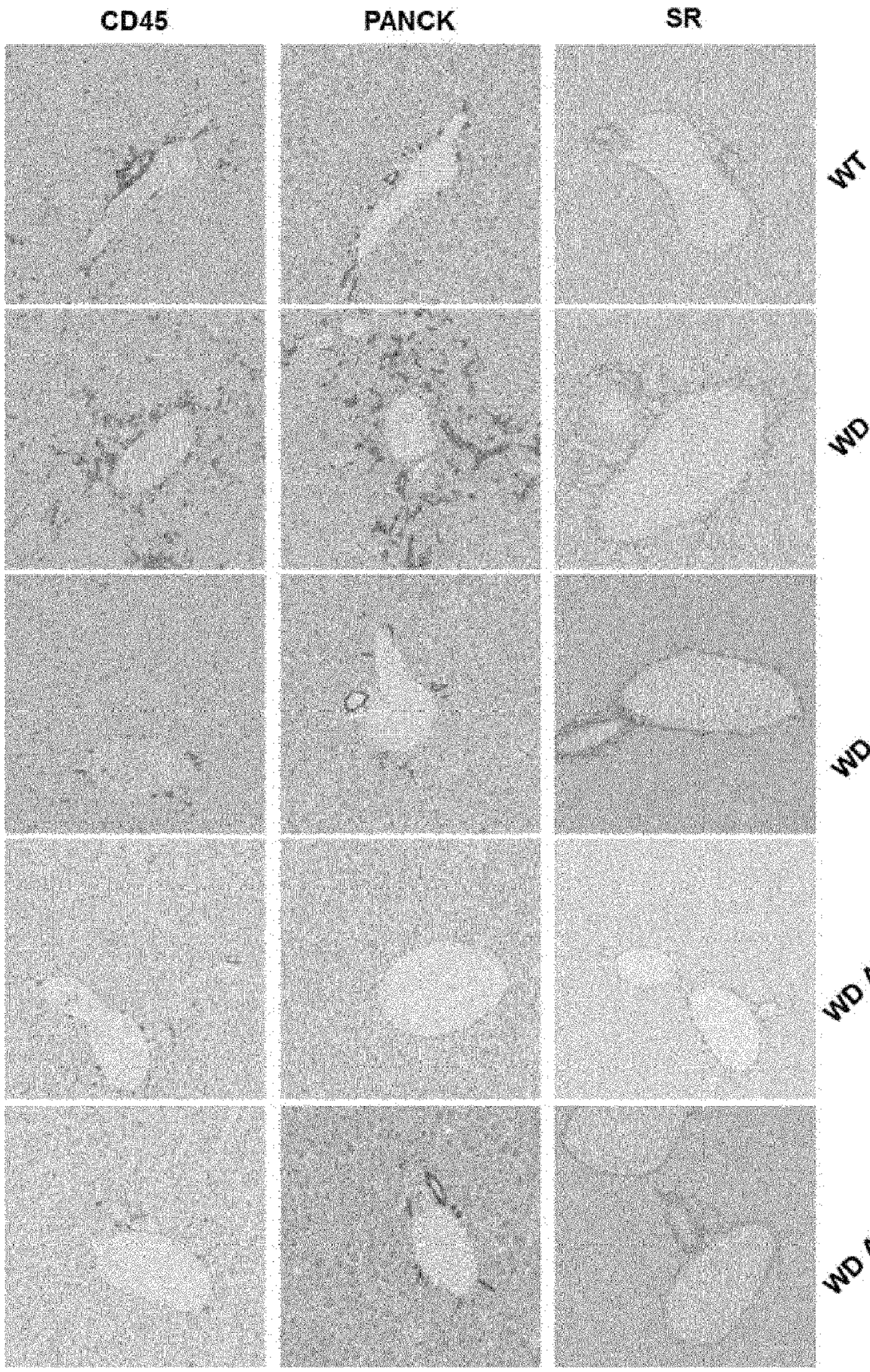


Figure 7

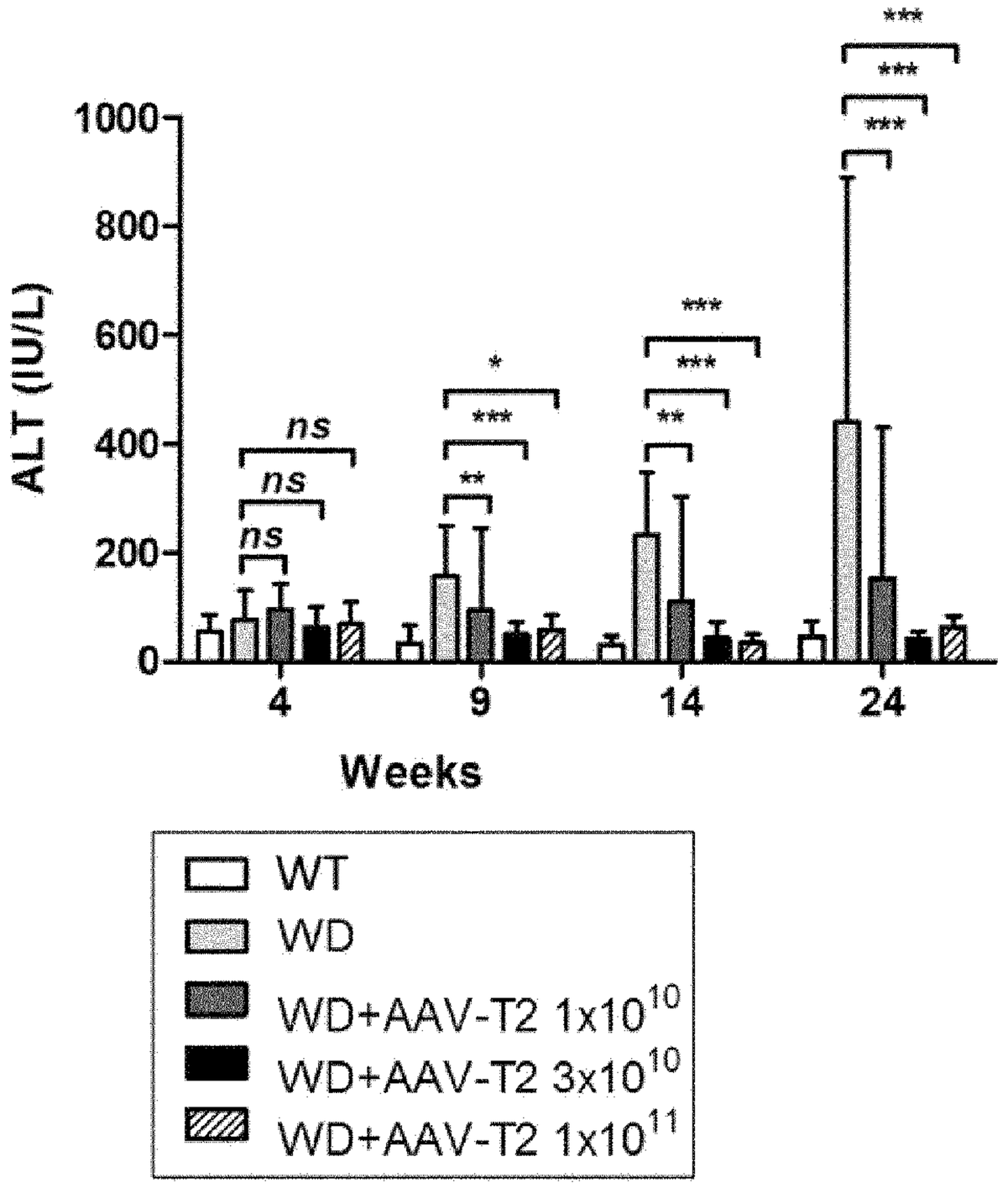

ALT (IU/L)
1000
800
600
400
200
0
4
9
14
24
Weeks
ns
ns
ns
*
**
***
***
***
***
**
***
***
WT
WD
WD+AAV-T2 1x10^10
WD+AAV-T2 3x10^10
WD+AAV-T2 1x10^11

Figure 8

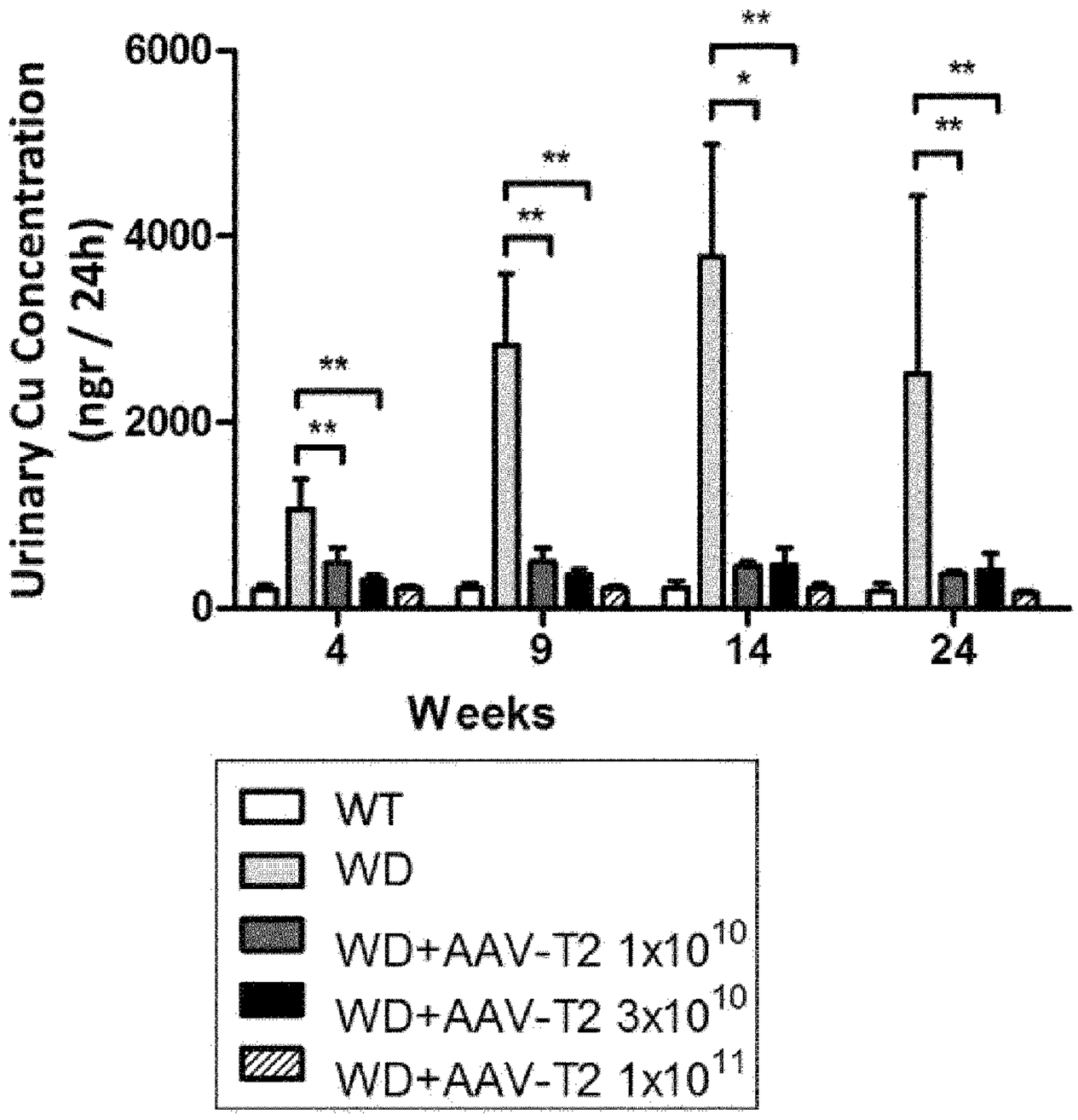
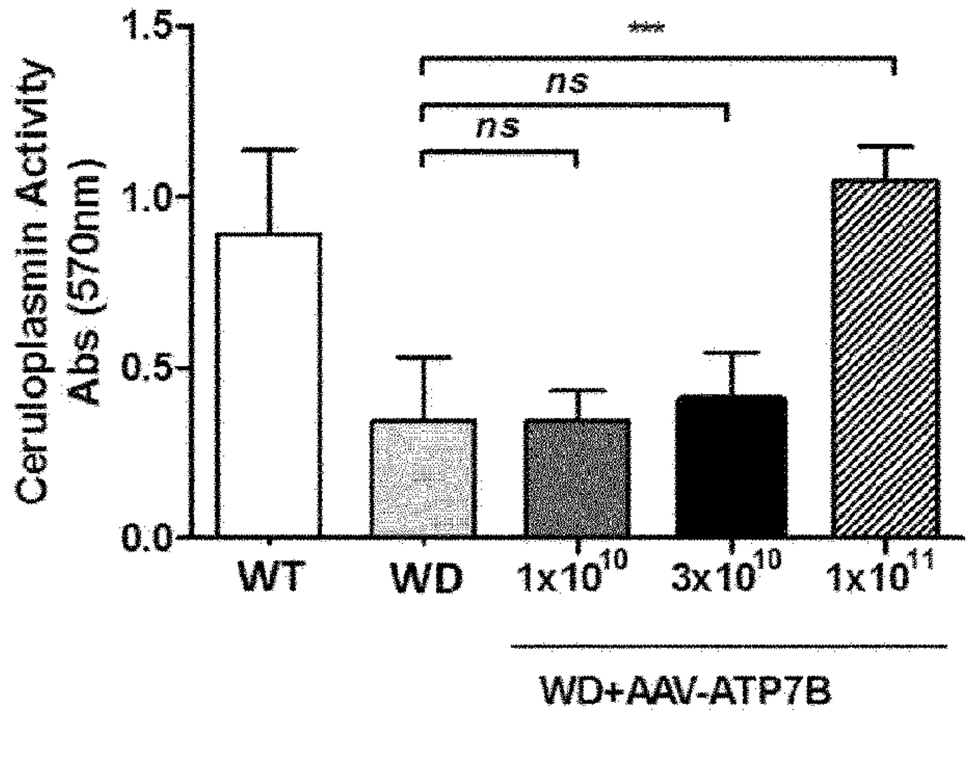

Urinary Cu Concentration (ngr / 24h)
6000
4000
2000
0
**
**
**
**
*
**
**
4
9
14
24
Weeks
WT
WD
WD+AAV-T2 1x10^10
WD+AAV-T2 3x10^10
WD+AAV-T2 1x10^11

Figure 9

A Experimental group 1



B Experimental group 2

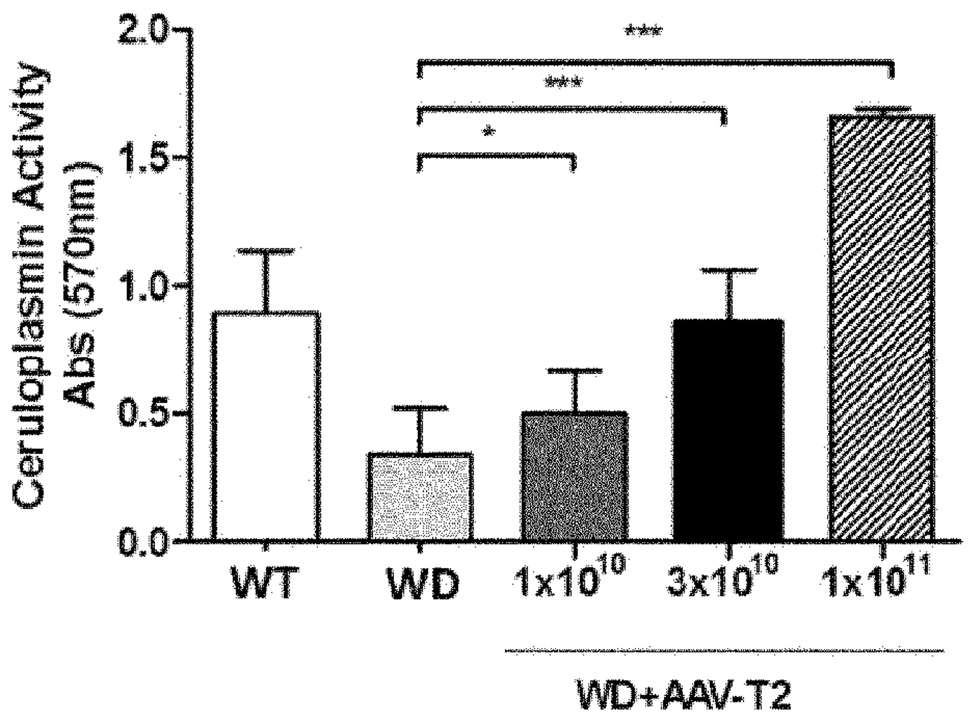
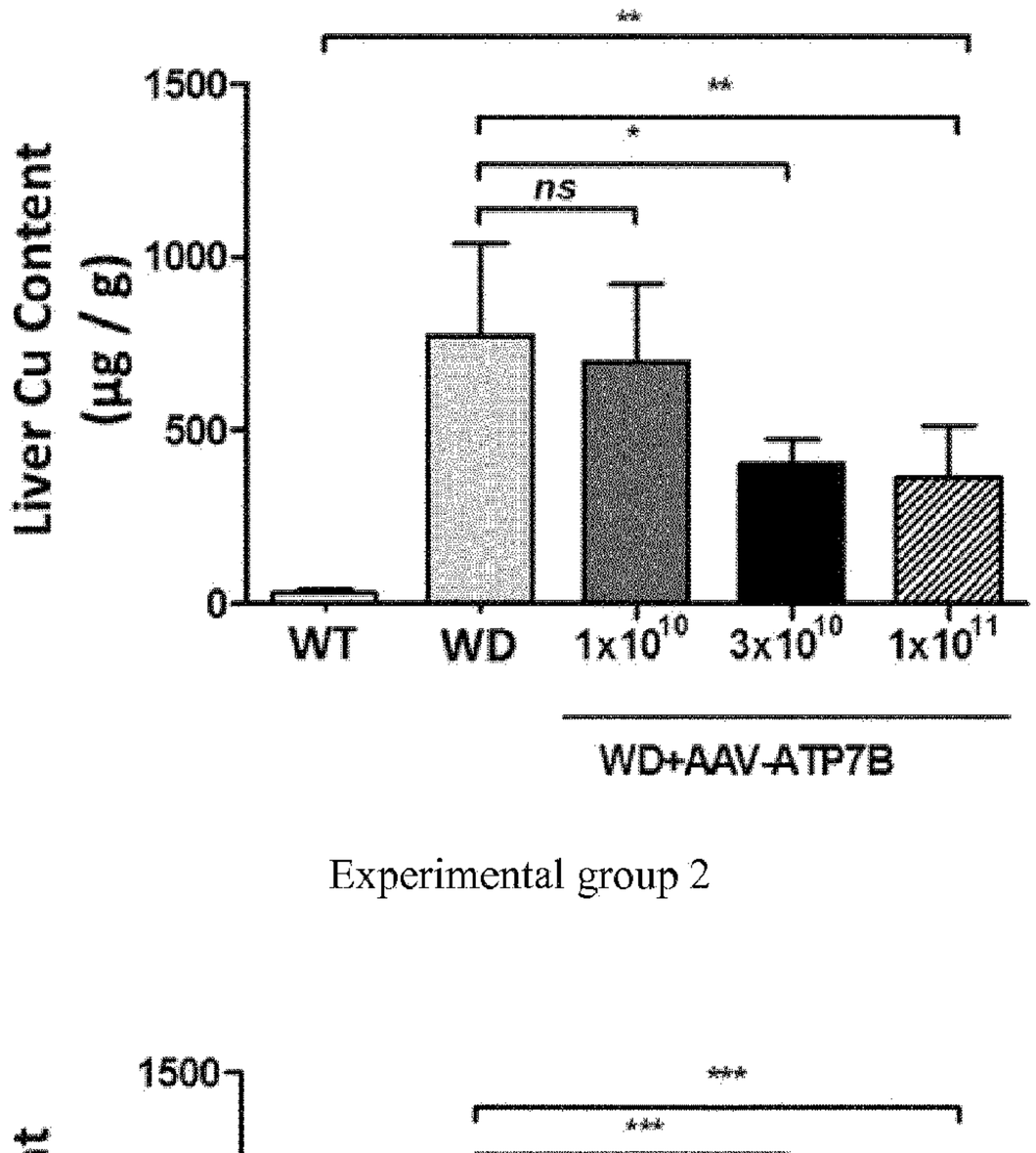


Figure 10

A Experimental group 1



B Experimental group 2

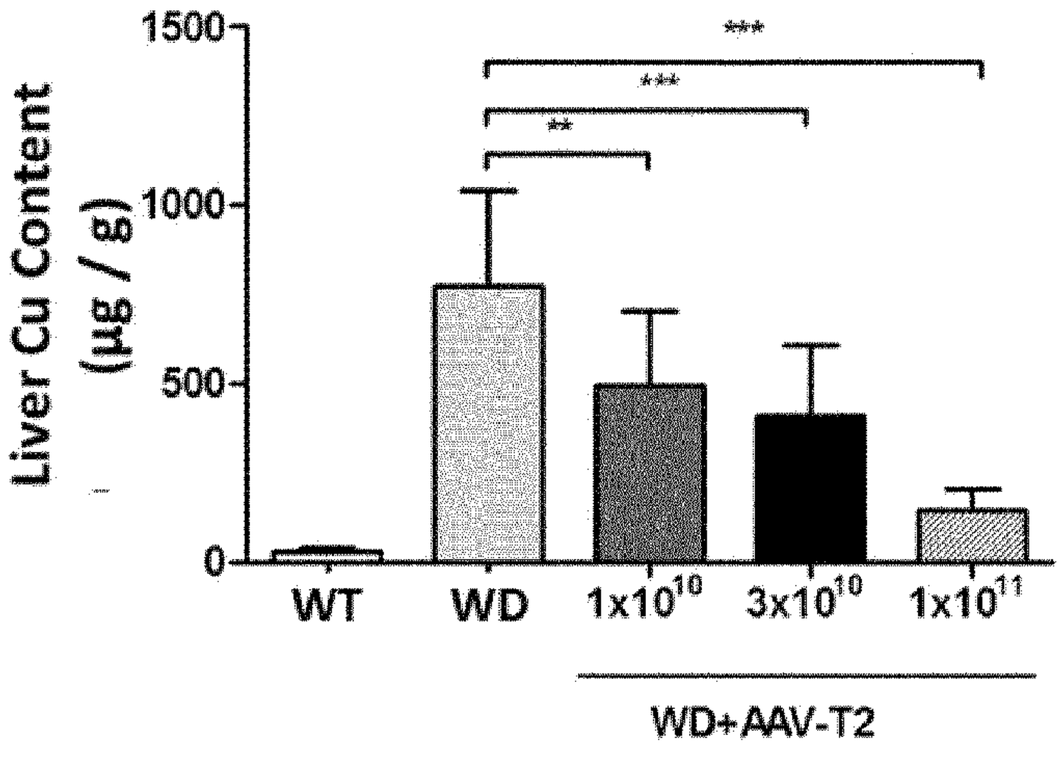


Figure 11

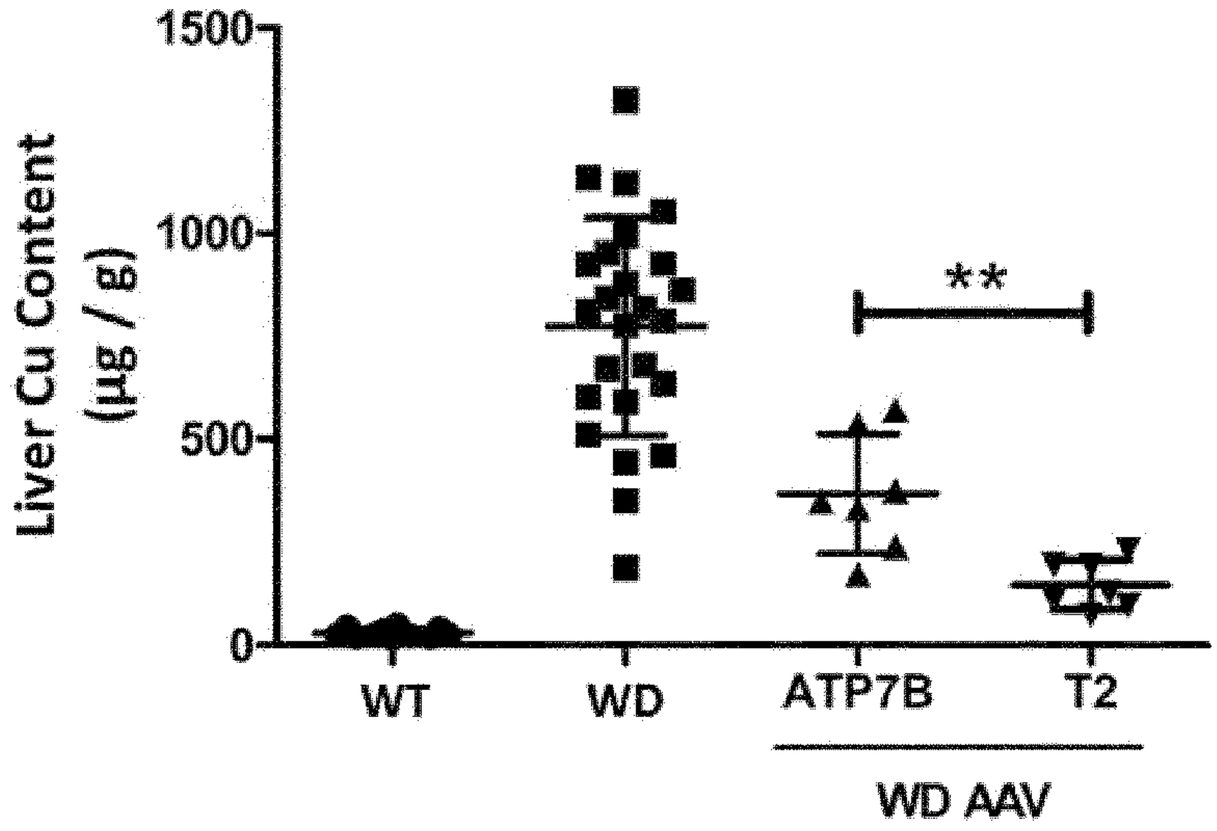


Figure 12

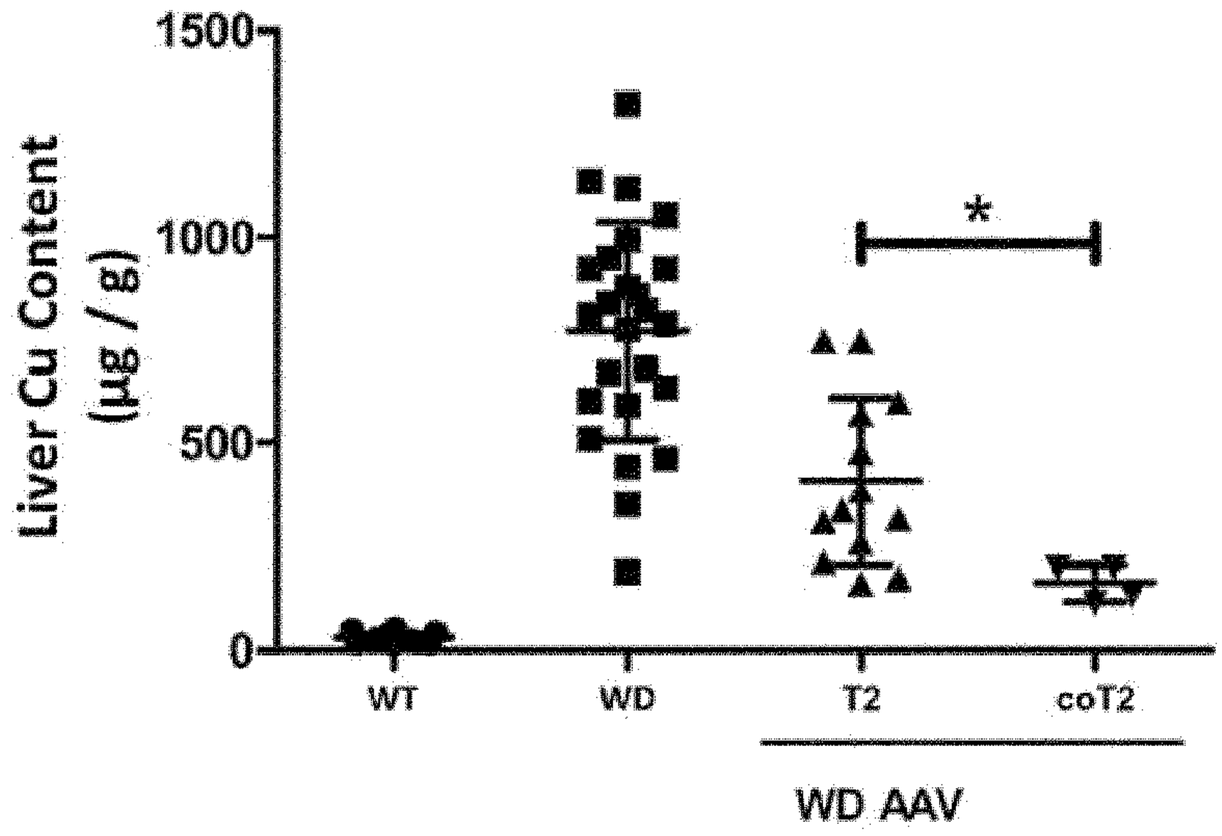


Figure 13

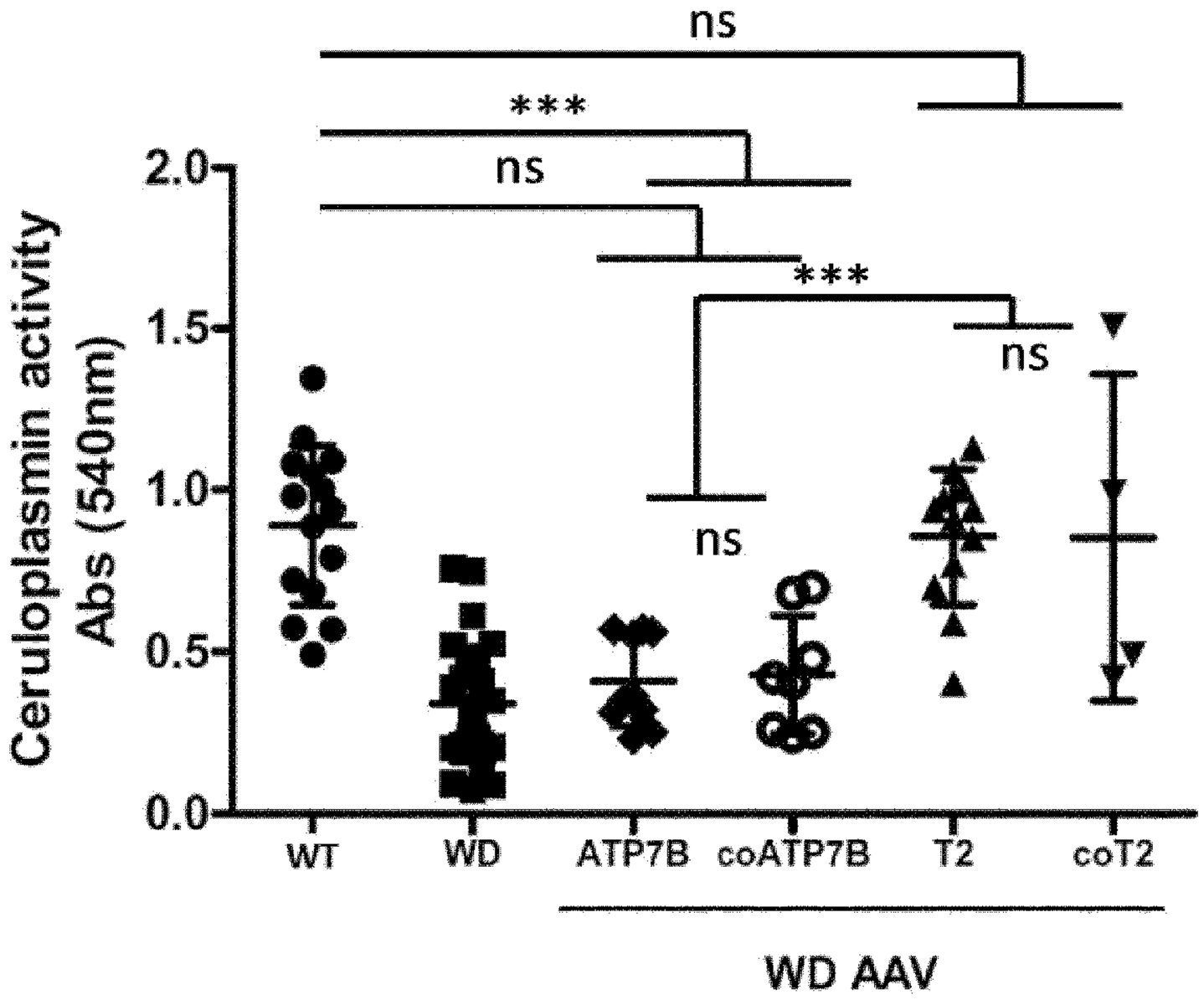

ns
***
ns
***
ns
ns
Ceruloplasmin activity
Abs (540nm)
2.0
1.5
1.0
0.5
0.0
WT
WD
ATP7B
coATP7B
T2
coT2
WD AAV

Figure 14