

PROVISIONAL APPLICATION FOR PATENT COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION FOR PATENT under 37 CFR 1.53(c).

Express Mail Label No. ER636323415US

INVENTOR(S)		
Given Name (first and middle [if any])	Family Name or Surname	Residence (City and either State or Foreign Country)
Luk H.	Vandenberghe	Philadelphia, Pennsylvania
Additional inventors are being named on the <u>1</u> separately numbered sheets attached hereto		
TITLE OF THE INVENTION (500 characters max):		
METHOD OF INCREASING THE FUNCTION OF AN AAV VECTOR		
Direct all correspondence to: CORRESPONDENCE ADDRESS		
☒ The address corresponding to Customer Number: 00270		
OR		
☐ Firm or Individual Name		
Address HOWSON AND HOWSON 321 Norristown Road, Suite 200, Box 457		
City Spring House	State PA	Zip 19477
Country USA	Telephone 215-540-9200	Fax 215-540-5818
ENCLOSED APPLICATION PARTS (check all that apply)		
☒ Application Data Sheet. See 37 CFR 1.76		
☐ CD(s), Number of CDs _____		
☒ Specification Number of Pages <u>48</u>		
☒ Other (specify) <u>Sequence Listing- pp. 19</u>		
☒ Drawing(s) Number of Sheets <u>21</u>		
Application Size Fee: If the specification and drawings exceed 100 sheets of paper, the application size fee due is \$250 (\$125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).		
METHOD OF PAYMENT OF FILING FEES AND APPLICATION SIZE FEE FOR THIS PROVISIONAL APPLICATION FOR PATENT		
☐ Applicant claims small entity status. See 37 CFR 1.27.		TOTAL FEE AMOUNT (\$)
☒ A check or money order is enclosed to cover the filing fee and application size fee (if applicable).		
☐ Payment by credit card. Form PTO-2038 is attached		200.00
☒ The Director is hereby authorized to charge the filing fee and application size fee (if applicable) or credit any overpayment to Deposit Account Number: <u>08-3040</u> . A duplicative copy of this form is enclosed for fee processing.		
The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.		
☒ No.		
☐ Yes, the name of the U.S. Government agency and the Government contract number are: _____		

SIGNATURE

Cathy A. Kodroff

Date April 7, 2005

TYPED or PRINTED NAME Cathy A. Kodroff

REGISTRATION NO. 33,980

(if appropriate)

TELEPHONE 215-540-9200

Docket Number: UPN-R3895

USE ONLY FOR FILING A PROVISIONAL APPLICATION FOR PATENT

This collection of information is required by 37 CFR 1.51. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.11 and 1.14. This collection is estimated to take 8 hours to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

PROVISIONAL APPLICATION COVER SHEET
Additional Page

PTO/SB/16 (12-04)

Approved for use through 07/31/2006. OMB 0651-0032

U.S. Patent and Trademark Office; U.S. DEPARTMENT OF COMMERCE

Under the Paperwork Reduction Act of 1995, no persons are required to respond to a collection of information unless it displays a valid OMB control number.

First Named Inventor	Vandenberghe et al	Docket Number	UPN-R3895
INVENTOR(S)/APPLICANT(S)			
Given Name (first and middle [if any])	Family or Surname	Residence (City and either State or Foreign Country)	
Guangping	Gao	Rosemont, Pennsylvania	
James M.	Wilson	Gladwyne, Pennsylvania	

Number 1 of 1

WARNING: Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.

Effective on 12/08/2004.

Fees pursuant to the Consolidated Appropriations Act, 2005 (H.R. 4818).

FEE TRANSMITTAL
For FY 2005**Complete if Known**☐ Applicant claims small entity status. See 37 CFR 1.27

TOTAL AMOUNT OF PAYMENT (\$) 200.00

Application Number	
Filing Date	April 7, 2005
First Named Inventor	Vandenberghe et al
Examiner Name	
Art Unit	
Attorney Docket No.	UPN-R3895

METHOD OF PAYMENT (check all that apply)☒ Check ☐ Credit Card ☐ Money Order ☐ None ☐ Other (please identify): _____☒ Deposit Account Deposit Account Number: 08-3040 Deposit Account Name: HOWSON AND HOWSON

For the above-identified deposit account, the Director is hereby authorized to: (check all that apply)

☐ Charge fee(s) indicated below☐ Charge fee(s) indicated below, except for the filing fee☒ Charge any additional fee(s) or underpayments of fee(s) under 37 CFR 1.16 and 1.17☒ Credit any overpayments**WARNING:** Information on this form may become public. Credit card information should not be included on this form. Provide credit card information and authorization on PTO-2038.**FEE CALCULATION****1. BASIC FILING, SEARCH, AND EXAMINATION FEES**

Application Type	FILING FEES		SEARCH FEES		EXAMINATION FEES		Fees Paid (\$)
	Fee (\$)	Small Entity Fee (\$)	Fee (\$)	Small Entity Fee (\$)	Fee (\$)	Small Entity Fee (\$)	
Utility	300	150	500	250	200	100	
Design	200	100	100	50	130	65	
Plant	200	100	300	150	160	80	
Reissue	300	150	500	250	600	300	
Provisional	200	100	0	0	0	0	200.00

2. EXCESS CLAIM FEES**Fee Description**

Fee (\$)	Small Entity Fee (\$)
Each claim over 20 (including Reissues)	50
Each independent claim over 3 (including Reissues)	200
Multiple dependent claims	360

Total Claims	Extra Claims	Fee (\$)	Fee Paid (\$)
---------------------	---------------------	-----------------	----------------------

_____ - 20 or HP = _____ x _____ = _____

HP = highest number of total claims paid for, if greater than 20.

Indep. Claims	Extra Claims	Fee (\$)	Fee Paid (\$)
----------------------	---------------------	-----------------	----------------------

_____ - 3 or HP = _____ x _____ = _____

HP = highest number of independent claims paid for, if greater than 3.

3. APPLICATION SIZE FEE

If the specification and drawings exceed 100 sheets of paper (excluding electronically filed sequence or computer listings under 37 CFR 1.52(e)), the application size fee due is \$250 (\$125 for small entity) for each additional 50 sheets or fraction thereof. See 35 U.S.C. 41(a)(1)(G) and 37 CFR 1.16(s).

Total Sheets	Extra Sheets	Number of each additional 50 or fraction thereof	Fee (\$)	Fee Paid (\$)
---------------------	---------------------	---	-----------------	----------------------

88 - 100 = 0 / 50 = _____ (round up to a whole number) x _____ = _____

4. OTHER FEES(S)

Non-English Specification, \$130 fee (no small entity discount)

Other (e.g., late filing surcharge): _____

SUBMITTED BY

Signature		Registration No. (Attorney/Agent) 33,980	Telephone 215-540-9200
Name (Print/Type)	CATHY A. RODROFF	Date APRIL 7, 2005	

This collection of information is required by 37 CFR 1.136. The information is required to obtain or retain a benefit by the public which is to file (and by the USPTO to process) an application. Confidentiality is governed by 35 U.S.C. 122 and 37 CFR 1.14. This collection is estimated to take 30 minutes to complete, including gathering, preparing, and submitting the completed application form to the USPTO. Time will vary depending upon the individual case. Any comments on the amount of time you require to complete this form and/or suggestions for reducing this burden, should be sent to the Chief Information Officer, U.S. Patent and Trademark Office, U.S. Department of Commerce, P.O. Box 1450, Alexandria, VA 22313-1450. DO NOT SEND FEES OR COMPLETED FORMS TO THIS ADDRESS. SEND TO: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450.

If you need assistance in completing the form, call 1-800-PTO-9199 and select option 2.

METHOD OF INCREASING THE FUNCTION OF AN AAV VECTOR

5 BACKGROUND OF THE INVENTION

Adeno-associated virus (AAV), a member of the Parvovirus family, is a small nonenveloped, icosahedral virus with single-stranded linear DNA genomes of 4.7 kilobases (kb) to 6 kb. AAV is assigned to the genus, *Dependovirus*, because the virus was discovered as a contaminant in purified adenovirus stocks. AAV's life cycle includes
10 a latent phase at which AAV genomes, after infection, are site specifically integrated into host chromosomes and an infectious phase in which, following either adenovirus or herpes simplex virus infection, the integrated genomes are subsequently rescued, replicated, and packaged into infectious viruses. The properties of non-pathogenicity, broad host range of infectivity, including non-dividing cells, and potential site-specific
15 chromosomal integration make AAV an attractive tool for gene transfer.

AAV vectors have been described for use as delivery vehicles for both therapeutic and immunogenic molecules. To date, there have been several different well-characterized AAVs isolated from human or non-human primates (NHP).

Recently, investigators have described a large number of AAVs of different
20 sequences. [G. Gao, *et al.*, *Proc Natl Acad Sci USA*, **100**(10):6081-6086 (May 13, 2003); US-2003-0138772-A1 (July 24, 2003)] and characterization of these AAVs into different serotypes and clades [G. Gao, *et al.*, *J. Virol.*, **78**(12):6381-6388 (June 2004); International Patent Application No. PCT/US2004/028817]. It has been reported that different AAVs exhibit different transfection efficiencies, and also exhibit tropism for
25 different cells or tissues.

What is desirable are AAV-based constructs for delivery of heterologous molecules to different cell types.

SUMMARY OF THE INVENTION

The present invention provides a method of improving vectors derived from AAV which are non-functional and/or which perform weakly.

5 In one aspect, the method provides a method for correcting singletons in a selected AAV sequence in order to increase the packaging yield, transduction efficiency, and/or gene transfer efficiency of the selected AAV. This method involves altering one or more singletons in the parental AAV capsid to conform the singleton to the amino acid in the corresponding position(s) of the aligned functional AAV capsid sequences.

10 In another aspect, the invention provides modified AAV sequences, *i.e.*, sequences with one or singletons eliminated.

In yet another aspect, the invention provides AAV vectors having modified AAV capsids according to the present invention.

Other aspects and advantages of the invention will be readily apparent from the following detailed description of the invention.

15

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a graph illustrating *in vitro* 293 transduction of singleton corrected AAV vectors. Singleton corrections are indicated after the vector name with – if present – a number to indicate the number of mutations performed.

20 Figs. 2A-2C are line graphs illustrating the titration of AAV vectors on 293 cells at multiplicities of infection ranging from 10^1 to 10^4 , with a comparison between parent rh.8 and singleton-corrected rh.8 (rh.8R) in Fig. 2A, parent rh.37 and modified rh.37 (Fig. 2B), and AAV2 and AAV8 in Fig. 2C. As a control a similar titration of AAV2 and AAV2/8 eGFP expressing vector is presented. % of eGFP positive cells is presented on
25 the Y-axis and was assayed by flow cytometry.

Fig. 3 is a phylogenetic tree of AAV sequences, which indicates their phylogenetic relationship and clades.

30 Figs. 4A-4K are an alignment of the nucleic acid sequences of the capsid protein (vp1) of AAV2, cy.5, rh.10, rh.13, AAV1, AAV3, AAV6, AAV7, AAV8, hu.13, hu.26, hu.37, hu.53, rh.39, rh.43 and rh.46.

Figs. 5A-5D are an alignment of the amino acid sequences of the capsid protein (vp1) of AAV2, cy.5, rh.10, rh.13, AAV1, AAV3, AAV6, AAV7, AAV8, hu.13, hu.26, hu.37, hu.53, rh.39, rh.43 and rh.46.

Fig. 6A-6B are an alignment of the amino acid sequences of the capsid protein (vp1) of rh.13, rh2, rh.8, hu.29, and rh.64.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a method for improving the function of an AAV vector. The present invention is particularly well suited to improve the packaging yield, transduction efficiency, and/or gene transfer efficiency of an AAV vector having a capsid of an AAV which contains one or more singletons. The invention further provides novel AAV capsid sequences identified and prepared according to the method of the invention.

As used throughout this specification and the claims, the terms “comprising” and “including” are inclusive of other components, elements, integers, steps and the like. Conversely, the term “consisting” and its variants are exclusive of other components, elements, integers, steps and the like.

SINGLETON METHOD OF INVENTION

As used herein, the term “singleton” refers to a variable amino acid in a given position in a selected (*i.e.*, parental) AAV capsid sequence. The presence of a variable amino acid is determined by aligning the sequence of the parental AAV capsid with a library of functional AAV capsid sequences. The sequences are then analyzed to determine the presence of any variable amino acid sequences in the parental AAV capsid where the sequences of the AAV in the library of functional AAVs have complete conservation. The parental AAV sequence is then altered to change the singleton to the conserved amino acid identified in that position in the functional AAV capsid sequences. According to the present invention, a parental AAV sequence may have 1 to 6, 1 to 5, 1 to 4, 1 to 3, or 2 singletons. A parental AAV sequence may have more than 6 singletons.

Once modified, the modified AAV capsid can be used to construct an AAV vector having the modified capsid. This vector can be constructed using techniques known to those of skill in the art.

The AAV selected for modification according to the invention method is one for which it is desirable to increase any one or more of the following three functional properties of AAV, *i.e.*, packaging into the viral particle having the capsid of the selected AAV sequence, increasing transduction efficiency, or increasing gene transfer efficiency as compared to the parental AAV. For example, the parental AAV may be characterized by having a lower packaging efficiency than other, closely related AAV. In another example, the parental AAV may have a lower transduction efficiency as compared to closely related AAVs. In another example, the parental AAV may have a lower gene transfer efficiency (*i.e.*, a lower ability to deliver a target molecule *in vivo*) as compared to closely related AAVs. In other examples, the parental AAV is characterized by adequate function in each of these categories, but increased function one or more of these areas is desired.

Thus, the invention provides a library of functional AAVs, the sequences of which are to be compared to the selected (parental) AAV. Suitably, the library contains AAVs which have a desired function which is targeted for improvement in the selected parental AAV. In other words, each of the sequences in the library of functional AAVs is characterized by a desired level of packaging ability, a desired level of *in vitro* transduction efficiency, or a desired level of *in vivo* gene transfer efficiency (*i.e.*, the ability to deliver to a target selected target tissue or cell in a subject). The functional AAVs which compose the library may have individually have one, two or all of these functional characteristics. Other desired functions for the library may be readily determined by one of skill in the art.

In one embodiment, a functional AAV is an AAV characterized by the ability to produce viral particles with equivalent or greater packaging and transduction efficiency as any one of AAV1, AAV2, AAV7, AAV8 or AAV9. Function may be assessed in a pseudotyping setting with AAV2 rep and AAV2 ITRs. Thus, an altered parental AAV can be constructed using conventional techniques and the AAV vector is considered function if virus is produced from the parental AAV at titers of at least 50% when compared to production of AAV2. Further, the ability of AAV to transduce cells can be

readily determined by one of skill in the art. For example, a parental AAV can be constructed such that it contains a marker gene which allows ready detection of virus. For example, the AAV contains eGFP or another gene which allows colorimetric detection. Where the AAV contains CMV-eGFP, when the virus produced from the altered parental AAV capsid is transduced into 293 cells at a multiplicity of infection of 10^4 , function is demonstrated where transduction efficiency is greater than 5% GFP fluorescence of total cells in a context where the cells were pretreated with wild-type human adenovirus type 5 at a multiplicity of infection of 20 for 2 hours.

Suitably, a library is composed of at least three or at least four functional AAV capsid sequences which represent at least two different clades. Preferably, at least two sequences from each of the represented clades is included in the library. In certain embodiments, three, four, five, six, or more clades are represented.

A "clade" is a group of AAV which are phylogenetically related to one another as determined using a Neighbor-Joining algorithm by a bootstrap value of at least 75% (of at least 1000 replicates) and a Poisson correction distance measurement of no more than 0.05, based on alignment of the AAV vp1 amino acid sequence.

The Neighbor-Joining algorithm has been described extensively in the literature. See, e.g., M. Nei and S. Kumar, *Molecular Evolution and Phylogenetics* (Oxford University Press, New York (2000)). Computer programs are available that can be used to implement this algorithm. For example, the MEGA v2.1 program implements the modified Nei-Gojobori method. Using these techniques and computer programs, and the sequence of an AAV vp1 capsid protein, one of skill in the art can readily determine whether a selected AAV is contained in one of the clades identified herein, in another clade, or is outside these clades.

While the AAV clades are based primarily upon naturally occurring AAV vp1 capsids, the clades are not limited to naturally occurring AAV. The clades can encompass non-naturally occurring AAV, including, without limitation, recombinant, modified or altered, chimeric, hybrid, synthetic, artificial, etc., AAV which are phylogenetically related as determined using a Neighbor-Joining algorithm at least 75% (of at least 1000 replicates) and a Poisson correction distance measurement of no more than 0.05, based on alignment of the AAV vp1 amino acid sequence.

The AAV clades which have been described include Clade A (represented by AAV1 and AAV6), Clade B (represented by AAV2) and Clade C (represented by the AAV2-AAV3 hybrid), Clade D (represented by AAV7), Clade E (represented by AAV8), and Clade F (represented by human AAV9). These clades are represented by a member of the clade that is a previously described AAV serotype. Previously described AAV1 and AAV6 are members of a single clade (Clade A) in which 4 isolates were recovered from 3 humans. Previously described AAV3 and AAV5 serotypes are clearly distinct from one another, but were not detected in the screen described herein, and have not been included in any of these clades.

Further discussion of AAV clades is provided in G. Gao, *et al.*, J. Virol., 78(12):6381-6388 (June 2004) and International Patent Application No. PCT/US 2004/028817, filed September 30, 2004.

In one embodiment, the libraries used in the method of the invention exclude AAV5. In another embodiment, the libraries used in the method of the invention exclude AAV4. However, in certain embodiment, *e.g.*, where the parental AAV is similar to AAV5, it may be desirable to include this sequence in the alignment.

Although a library can be constructed that contains the minimal number of sequences, efficiency in identifying singletons may be optimized by utilizing a library containing a larger number of sequences. Suitably, the library contains a minimum of four sequences, with at least two clades being represented. Preferably, the library contains at least two sequences from each of the represented clades. In one embodiment, the library contains more than 100 AAV sequences. In another embodiment, the library contains at least three to 100 AAV sequences. In still another embodiment, the library contains at least six to 50 AAV sequences.

Suitable AAVs for use in the functional libraries of the invention include, *e.g.*, AAV1, AAV2, AAV6, AAV7, AAV8, AAV9, and other sequences which have been described [G. Gao, *et al.*, *Proc Natl. Acad Sci.*, **100**(10):6081-6086 (May 13, 2003); International Patent Application No. PCT/US2004/028817]. One of skill in the art can readily select other AAVs, *e.g.*, those isolated using the methods described in WO 03/093460 A1 (Nov 13, 2003); US-2003-0138772 A1 (July 24, 2003).

According to the present invention, the at least three sequences within the library are least 85% identical over the full-length of their aligned capsid sequences.

The term “percent (%) identity” may be readily determined for amino acid sequences, over the full-length of a protein, or a fragment thereof. Suitably, a fragment is at least about 8 amino acids in length, and may be up to about 700 amino acids.

Generally, when referring to “identity”, “homology”, or “similarity” between two different adeno-associated viruses, “identity”, “homology” or “similarity” is determined in reference to “aligned” sequences. “Aligned” sequences or “alignments” refer to multiple nucleic acid sequences or protein (amino acids) sequences, often containing corrections for missing or additional bases or amino acids as compared to a reference sequence.

Alignments are performed using any of a variety of publicly or commercially available Multiple Sequence Alignment Programs. Sequence alignment programs are available for amino acid sequences, *e.g.*, the “Clustal X”, “MAP”, “PIMA”, “MSA”, “BLOCKMAKER”, “MEME”, and “Match-Box” programs. Generally, any of these programs are used at default settings, although one of skill in the art can alter these settings as needed. Alternatively, one of skill in the art can utilize another algorithm or computer program which provides at least the level of identity or alignment as that provided by the referenced algorithms and programs. *See, e.g.*, J. D. Thomson *et al*, *Nucl. Acids. Res.*, “A comprehensive comparison of multiple sequence alignments”, 27(13):2682-2690 (1999). Multiple sequence alignment programs are also available for nucleic acid sequences. Examples of such programs include, “Clustal W”, “CAP Sequence Assembly”, “MAP”, and “MEME”, which are accessible through Web Servers on the internet. Other sources for such programs are known to those of skill in the art. Alternatively, Vector NTI utilities are also used. There are also a number of algorithms known in the art that can be used to measure nucleotide sequence identity, including those contained in the programs described above. As another example, polynucleotide sequences can be compared using Fasta™, a program in GCG Version 6.1. Fasta™ provides alignments and percent sequence identity of the regions of the best overlap between the query and search sequences. For instance, percent sequence identity between nucleic acid sequences can be determined using Fasta™ with its default parameters (a word size of 6 and the NOPAM factor for the scoring matrix) as provided in GCG Version 6.1, herein incorporated by reference.

According to the invention, the sequences of the target or parental AAV capsid suspected of containing a singleton are compared to the sequences of the AAV capsids within the library. This comparison is performed using an alignment of the full-length vp1 protein of the AAV capsid.

5 A singleton is identified when, for a selected amino acid position when the AAV sequences are aligned, all of the AAVs in the library have the same amino acid residue (*i.e.*, are completely conserved), but the parental AAV has a different amino acid residue.

10 Typically, when an alignment is prepared based upon the AAV capsid vp1 protein, the alignment contains insertions and deletions which are so identified with respect to a reference AAV sequence (*e.g.*, AAV2) and the numbering of the amino acid residues is based upon a reference scale provided for the alignment. However, any given AAV sequence may have fewer amino acid residues than the reference scale. In the present invention, when discussing the parental AAV and the sequences of the reference library, the term “the same position” or the “corresponding position” refers to the amino acid located at the same residue number in each of the sequences, with respect to the reference scale for the aligned sequences. However, when taken out of the alignment, each of the AAV vp1 proteins may have these amino acids located at different residue numbers.

20 Optionally, the method of the invention can be performed using a nucleic acid alignment and identifying as a singleton a codon which encodes a different amino acid (*i.e.*, a non-synonymous codon). Where the nucleic acid sequences of a given codon are not identical in the parental AAV as compared to the sequences of that codon in the library, but encode the same amino acid, they are considered synonymous and are not a singleton.

25 According to the present invention, a parental AAV containing a singleton is altered such that the singleton residue is replaced with the conserved amino acid residue of the AAVs in the library.

30 Conveniently, this replacement can be performed by using conventional site-directed mutagenesis techniques on the codon for the variable amino acid. Typically, the site-directed mutagenesis is performed using as few steps as required to obtain the desired codon for the conserved amino acid residue. Such methods are well known to

those of skill in the art and can be performed using published methods and/or commercially available kits [*e.g.*, available from Stratagene and Promega]. The site-directed mutagenesis may be performed on the AAV genomic sequence. The AAV sequence may be carried by a vector (*e.g.*, a plasmid backbone) for convenience.

5 Alternatively, one of skill in the art can alter the parental AAV using other techniques known to those of skill in the art.

A parental AAV may have more than one singleton, *e.g.*, two, three, four, five, six or more. However, improvement in function may be observed after correction of one singleton. In the embodiment where a parental AAV carries multiple singletons,
10 each singleton may be altered at a time, followed by assessment of the modified AAV for enhancement of the desired function. Alternatively, multiple singletons may be altered prior to assessment for enhancement of the desired function.

Even where a parental AAV contains multiple singletons and functional improvement is observed altered of a first singleton, it may be desirable to optimize
15 function by altering the remaining singleton(s).

Typically, a parental AAV which has had one or more singleton(s) altered according to the method of the invention, is assessed for function by packaging the AAV into an AAV particle. These methods are well known to those of skill in the art. See, *e.g.*, G. Gao *et al*, Proc Natl Acad Sci., cited above; Sambrook *et al*, Molecular
20 Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor NY.

These altered AAVs have novel capsids produced according to the method of the invention and are assessed for function. Suitable methods for assessing AAV function have been described herein and include, *e.g.*, the ability to produce DNase protected particles, *in vitro* cell transduction efficiency, and/or *in vivo* gene transfer.
25 Suitably, the altered AAVs of the invention have a sufficient number of singletons altered to increase function in one or all of these characteristics, as compared to the function of the parent AAV.

II. Novel AAV of the Invention

30 The invention further provides a method of predicting whether a novel AAV will be functional. The method involves using the singleton method of the invention and

identifying the absence of a singleton in the sequence of the selected AAV, *i.e.*, an AAV which lacks a singleton.

Thus, in one embodiment, the invention provides a method of selecting an AAV for use in producing a vector. This method involves selecting a parental AAV capsid sequence for analysis and identifying the absence of any singletons in the parental AAV capsid in an alignment comprising the parental AAV capsid sequence and a library of functional AAV capsid sequences. Once the absence of a singleton in a selected AAV capsid is determined, the AAV can be used to generate a vector according to known techniques.

The term “substantial homology” or “substantial similarity,” when referring to a nucleic acid, or fragment thereof, indicates that, when optimally aligned with appropriate nucleotide insertions or deletions with another nucleic acid (or its complementary strand), there is nucleotide sequence identity in at least about 95 to 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or an open reading frame thereof, or another suitable fragment which is at least 15 nucleotides in length. Examples of suitable fragments are described herein.

The terms “sequence identity” “percent sequence identity” or “percent identical” in the context of nucleic acid sequences refers to the residues in the two sequences which are the same when aligned for maximum correspondence. The length of sequence identity comparison may be over the full-length of the genome, the full-length of a gene coding sequence, or a fragment of at least about 500 to 5000 nucleotides, is desired. However, identity among smaller fragments, *e.g.* of at least about nine nucleotides, usually at least about 20 to 24 nucleotides, at least about 28 to 32 nucleotides, at least about 36 or more nucleotides, may also be desired.

The term “substantial homology” or “substantial similarity,” when referring to amino acids or fragments thereof, indicates that, when optimally aligned with appropriate amino acid insertions or deletions with another amino acid (or its complementary strand), there is amino acid sequence identity in at least about 95 to 99% of the aligned sequences. Preferably, the homology is over full-length sequence, or a protein thereof, *e.g.*, a cap protein, a rep protein, or a fragment thereof which is at least 8 amino acids, or more desirably, at least 15 amino acids in length. Examples of suitable fragments are described herein.

By the term "highly conserved" is meant at least 80% identity, preferably at least 90% identity, and more preferably, over 97% identity. Identity is readily determined by one of skill in the art by resort to algorithms and computer programs known by those of skill in the art.

5 The term "serotype" is a distinction with respect to an AAV having a capsid which is serologically distinct from other AAV serotypes. Serologic distinctiveness is determined on the basis of the lack of cross-reactivity between antibodies to the AAV as compared to other AAV. Cross-reactivity is typically measured in a neutralizing antibody assay. For this assay polyclonal serum is generated against a specific AAV
10 in a rabbit or other suitable animal model using the adeno-associated viruses. In this assay, the serum generated against a specific AAV is then tested in its ability to neutralize either the same (homologous) or a heterologous AAV. The dilution that achieves 50% neutralization is considered the neutralizing antibody titer. If for two AAVs the quotient of the heterologous titer divided by the homologous titer is lower
15 than 16 in a reciprocal manner, those two vectors are considered as the same serotype. Conversely, if the ratio of the heterologous titer over the homologous titer is 16 or more in a reciprocal manner the two AAVs are considered distinct serotypes.

 In a further embodiment, the invention provides AAV having novel capsids, including rh. 20, rh.32/33, rh.39, rh.46, rh.73, and rh.74. The sequences of rh.20 have
20 the amino acid sequence of SEQ ID NO:1, or a sequence 95 to 99% identical over the full-length of SEQ ID NO:1. The capsid of rh.32/33 has an amino acid sequence of SEQ ID NO:2, or sequences 95% to 99% identical thereto over the full-length of SEQ ID NO:2. The capsid of rh.39 has an amino acid sequence of SEQ ID NO:3, or sequences 95% to 99% identical thereto over the full-length of SEQ ID NO:3. The capsid of rh.46
25 has an amino acid sequence of SEQ ID NO:4, or sequences 95% to 99% identical thereto over the full-length of SEQ ID NO:4. The capsid of rh.73 has an amino acid sequence of SEQ ID NO:5, or sequences 95% to 99% identical thereto over the full-length of SEQ ID NO:5. The capsid of rh.74 has an amino acid sequence of SEQ ID NO:6, or sequences 95% to 99% identical thereto over the full-length of SEQ ID NO:6. Preferably, the
30 sequence identity of these novel AAV capsids is such that it lacks any singletons. The sequences of the novel AAV are provided in the Sequence Listing.

Particularly desirable proteins include the AAV capsid proteins, which are encoded by the nucleotide sequences identified above. The AAV capsid is composed of three proteins, vp1, vp2 and vp3, which are alternative splice variants. Other desirable fragments of the capsid protein include the constant and variable regions,
5 located between hypervariable regions (HVR). Other desirable fragments of the capsid protein include the HVR themselves.

An algorithm developed to determine areas of sequence divergence in AAV2 has yielded 12 hypervariable regions (HVR) of which 5 overlap or are part of the four previously described variable regions. [Chiorini *et al*, *J. Virol*, 73:1309-19 (1999);
10 Rutledge *et al*, *J. Virol.*, 72:309-319] Using this algorithm and/or the alignment techniques described herein, the HVR of the novel AAV serotypes are determined. For example, the HVR are located as follows: HVR1, aa 146-152; HVR2, aa 182-186; HVR3, aa 262-264; HVR4, aa 381-383; HVR5, aa 450-474; HVR6, aa 490-495; HVR7, aa 500-504; HVR8, aa 514-522; HVR9, aa 534-555; HVR10, aa 581-594; HVR11, aa
15 658-667; and HVR12, aa 705-719 [the numbering system is based on an alignment which uses the AAV2 vp1 as a point of reference]. Using the alignment provided herein performed using the Clustal X program at default settings, or using other commercially or publicly available alignment programs at default settings such as are described herein, one of skill in the art can readily determine corresponding fragments of the novel AAV
20 capsids of the invention.

Suitably, fragments are at least 8 amino acids in length. However, fragments of other desired lengths may be readily utilized. Such fragments may be produced recombinantly or by other suitable means, *e.g.*, chemical synthesis.

The invention further provides other AAV sequences which are identified using
25 the sequence information provided herein. For example, given the sequences provided herein, infectious may be isolated using genome walking technology (Siebert *et al.*, 1995, *Nucleic Acid Research*, 23:1087-1088, Friezner-Degen *et al.*, 1986, *J. Biol. Chem.* 261:6972-6985, BD Biosciences Clontech, Palo Alto, CA). Genome walking is particularly well suited for identifying and isolating the sequences adjacent to the novel
30 sequences identified according to the method of the invention. This technique is also useful for isolating inverted terminal repeat (ITRs) of the novel AAV, based upon the novel AAV capsid and rep sequences provided herein.

The novel AAV amino acid sequences, peptides and proteins may be expressed from AAV nucleic acid sequences of the invention. Additionally, these amino acid sequences, peptides and proteins can be generated by other methods known in the art, including, *e.g.*, by chemical synthesis, by other synthetic techniques, or by other methods.

5 The sequences of any of the AAV capsids provided herein can be readily generated using a variety of techniques.

Suitable production techniques are well known to those of skill in the art. See, *e.g.*, Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press (Cold Spring Harbor, NY). Alternatively, peptides can also be synthesized by the
10 well-known solid phase peptide synthesis methods (Merrifield, *J. Am. Chem. Soc.*, **85**:2149 (1962); Stewart and Young, *Solid Phase Peptide Synthesis* (Freeman, San Francisco, 1969) pp. 27-62). These and other suitable production methods are within the knowledge of those of skill in the art and are not a limitation of the present invention.

The sequences, proteins, and fragments of the invention may be produced by any
15 suitable means, including recombinant production, chemical synthesis, or other synthetic means. Such production methods are within the knowledge of those of skill in the art and are not a limitation of the present invention.

III. Production of rAAV with Novel AAV Capsids

20 The invention encompasses novel AAV capsid sequences generated by mutation following use of the method of the invention for identifying singletons. The invention further encompasses novel AAV containing AAV rh.20, rh.32/33, rh.39, rh.46, rh.73, and rh.74 capsid sequences [SEQ ID Nos: 1- 6].

In another aspect, the present invention provides molecules that utilize the novel
25 AAV sequences of the invention, including fragments thereof, for production of viral vectors useful in delivery of a heterologous gene or other nucleic acid sequences to a target cell.

The molecules of the invention which contain AAV sequences include any genetic element (vector) which may be delivered to a host cell, *e.g.*, naked DNA, a
30 plasmid, phage, transposon, cosmid, episome, a protein in a non-viral delivery vehicle (*e.g.*, a lipid-based carrier), virus, etc., which transfers the sequences carried thereon.

The selected vector may be delivered by any suitable method, including transfection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion. The methods used to construct any embodiment of this invention are known to those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. See, *e.g.*, Sambrook *et al*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, NY.

In one embodiment, the vectors of the invention contain, *inter alia*, sequences encoding an AAV capsid of the invention or a fragment thereof. In another embodiment, the vectors of the invention contain, at a minimum, sequences encoding an AAV rep protein or a fragment thereof. Optionally, vectors of the invention may contain both AAV cap and rep proteins. In vectors in which both AAV *rep* and *cap* are provided, the AAV *rep* and AAV *cap* sequences can originate from an AAV of the same clade. Alternatively, the present invention provides vectors in which the *rep* sequences are from an AAV source which differs from that which is providing the *cap* sequences. In one embodiment, the *rep* and *cap* sequences are expressed from separate sources (*e.g.*, separate vectors, or a host cell and a vector). In another embodiment, these *rep* sequences are fused in frame to *cap* sequences of a different AAV source to form a chimeric AAV vector. Optionally, the vectors of the invention are vectors packaged in an AAV capsid of the invention. These vectors and other vectors described herein can further contain a minigene comprising a selected transgene which is flanked by AAV 5' ITR and AAV 3' ITR.

Thus, in one embodiment, the vectors described herein contain nucleic acid sequences encoding an intact AAV capsid which may be from a single AAV sequence. Alternatively, these vectors contain sequences encoding artificial capsids which contain one or more fragments of the singleton-corrected AAV capsid fused to heterologous AAV or non-AAV capsid proteins (or fragments thereof). These artificial capsid proteins are selected from non-contiguous portions of the singleton-corrected capsid or from capsids of other AAVs. These modifications may be to increase expression, yield, and/or to improve purification in the selected expression systems, or for another desired purpose (*e.g.*, to change tropism or alter neutralizing antibody epitopes).

The vectors described herein, *e.g.*, a plasmid, are useful for a variety of purposes, but are particularly well suited for use in production of a rAAV containing a capsid comprising AAV sequences or a fragment thereof. These vectors, including rAAV, their elements, construction, and uses are described in detail herein.

5 In one aspect, the invention provides a method of generating a recombinant adeno-associated virus (AAV) having a novel AAV capsid of the invention. Such a method involves culturing a host cell which contains a nucleic acid sequence encoding a novel AAV capsid protein of the invention, or fragment thereof, as defined herein; a functional *rep* gene; a minigene composed of, at a minimum, AAV inverted terminal
10 repeats (ITRs) and a transgene; and sufficient helper functions to permit packaging of the minigene into the AAV capsid protein.

The components required to be cultured in the host cell to package an AAV minigene in an AAV capsid may be provided to the host cell in *trans*. Alternatively, any one or more of the required components (*e.g.*, minigene, *rep* sequences, *cap* sequences,
15 and/or helper functions) may be provided by a stable host cell which has been engineered to contain one or more of the required components using methods known to those of skill in the art. Most suitably, such a stable host cell will contain the required component(s) under the control of an inducible promoter. However, the required component(s) may be under the control of a constitutive promoter. Examples of suitable inducible and
20 constitutive promoters are provided herein, in the discussion of regulatory elements suitable for use with the transgene. In still another alternative, a selected stable host cell may contain selected component(s) under the control of a constitutive promoter and other selected component(s) under the control of one or more inducible promoters. For example, a stable host cell may be generated which is derived from 293 cells (which
25 contain E1 helper functions under the control of a constitutive promoter), but which contains the *rep* and/or *cap* proteins under the control of inducible promoters. Still other stable host cells may be generated by one of skill in the art.

The minigene, *rep* sequences, *cap* sequences, and helper functions required for producing the rAAV of the invention may be delivered to the packaging host cell in the
30 form of any genetic element which transfer the sequences carried thereon. The selected genetic element may be delivered by any suitable method, including those described herein. The methods used to construct any embodiment of this invention are known to

those with skill in nucleic acid manipulation and include genetic engineering, recombinant engineering, and synthetic techniques. See, *e.g.*, Sambrook et al, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, NY. Similarly, methods of generating rAAV virions are well known and the selection of a suitable method is not a limitation on the present invention. See, *e.g.*, K. Fisher et al, *J. Virol.*, **70**:520-532 (1993) and US Patent No. 5,478,745.

Unless otherwise specified, the AAV ITRs, and other selected AAV components described herein, may be readily selected from among any AAV, including, without limitation, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV9 and one of the other novel AAV sequences of the invention. These ITRs or other AAV components may be readily isolated using techniques available to those of skill in the art from an AAV sequence. Such AAV may be isolated or obtained from academic, commercial, or public sources (*e.g.*, the American Type Culture Collection, Manassas, VA). Alternatively, the AAV sequences may be obtained through synthetic or other suitable means by reference to published sequences such as are available in the literature or in databases such as, *e.g.*, GenBank®, PubMed®, or the like.

A. The Minigene

The minigene is composed of, at a minimum, a transgene and its regulatory sequences, and 5' and 3' AAV inverted terminal repeats (ITRs). In one desirable embodiment, the ITRs of AAV serotype 2 are used. However, ITRs from other suitable sources may be selected. It is this minigene that is packaged into a capsid protein and delivered to a selected host cell.

1. The transgene

The transgene is a nucleic acid sequence, heterologous to the vector sequences flanking the transgene, which encodes a polypeptide, protein, or other product, of interest. The nucleic acid coding sequence is operatively linked to regulatory components in a manner which permits transgene transcription, translation, and/or expression in a host cell.

The composition of the transgene sequence will depend upon the use to which the resulting vector will be put. For example, one type of transgene sequence includes a reporter sequence, which upon expression produces a detectable signal. Such reporter sequences include, without limitation, DNA sequences encoding β -

lactamase, β -galactosidase (LacZ), alkaline phosphatase, thymidine kinase, green fluorescent protein (GFP), enhanced GFP (EGFP), chloramphenicol acetyltransferase (CAT), luciferase, membrane bound proteins including, for example, CD2, CD4, CD8, the influenza hemagglutinin protein, and others well known in the art, to which high affinity antibodies directed thereto exist or can be produced by conventional means, and fusion proteins comprising a membrane bound protein appropriately fused to an antigen tag domain from, among others, hemagglutinin or Myc.

These coding sequences, when associated with regulatory elements which drive their expression, provide signals detectable by conventional means, including enzymatic, radiographic, colorimetric, fluorescence or other spectrographic assays, fluorescent activating cell sorting assays and immunological assays, including enzyme linked immunosorbent assay (ELISA), radioimmunoassay (RIA) and immunohistochemistry. For example, where the marker sequence is the LacZ gene, the presence of the vector carrying the signal is detected by assays for beta-galactosidase activity. Where the transgene is green fluorescent protein or luciferase, the vector carrying the signal may be measured visually by color or light production in a luminometer.

However, desirably, the transgene is a non-marker sequence encoding a product which is useful in biology and medicine, such as proteins, peptides, RNA, enzymes, dominant negative mutants, or catalytic RNAs. Desirable RNA molecules include tRNA, dsRNA, ribosomal RNA, catalytic RNAs, siRNA, small hairpin RNA, trans-splicing RNA, and antisense RNAs. One example of a useful RNA sequence is a sequence which inhibits or extinguishes expression of a targeted nucleic acid sequence in the treated animal. Typically, suitable target sequences include oncologic targets and viral diseases. See, for examples of such targets the oncologic targets and viruses identified below in the section relating to immunogens.

The transgene may be used to correct or ameliorate gene deficiencies, which may include deficiencies in which normal genes are expressed at less than normal levels or deficiencies in which the functional gene product is not expressed. Alternatively, the transgene may provide a product to a cell which is not natively expressed in the cell type or in the host. A preferred type of transgene sequence encodes a therapeutic protein or polypeptide which is expressed in a host cell. The invention

further includes using multiple transgenes. In certain situations, a different transgene may be used to encode each subunit of a protein, or to encode different peptides or proteins. This is desirable when the size of the DNA encoding the protein subunit is large, *e.g.*, for an immunoglobulin, the platelet-derived growth factor, or a dystrophin protein. In order for the cell to produce the multi-subunit protein, a cell is infected with the recombinant virus containing each of the different subunits. Alternatively, different subunits of a protein may be encoded by the same transgene. In this case, a single transgene includes the DNA encoding each of the subunits, with the DNA for each subunit separated by an internal ribozyme entry site (IRES). This is desirable when the size of the DNA encoding each of the subunits is small, *e.g.*, the total size of the DNA encoding the subunits and the IRES is less than five kilobases. As an alternative to an IRES, the DNA may be separated by sequences encoding a 2A peptide, which self-cleaves in a post-translational event. See, *e.g.*, M.L. Donnelly, *et al.*, *J. Gen. Virol.*, **78**(Pt 1):13-21 (Jan 1997); Furler, S., *et al.*, *Gene Ther.*, **8**(11):864-873 (June 2001); Klump H., *et al.*, *Gene Ther.*, **8**(10):811-817 (May 2001). This 2A peptide is significantly smaller than an IRES, making it well suited for use when space is a limiting factor. More often, when the transgene is large, consists of multi-subunits, or two transgenes are co-delivered, rAAV carrying the desired transgene(s) or subunits are co-administered to allow them to concatamerize *in vivo* to form a single vector genome. In such an embodiment, a first AAV may carry an expression cassette which expresses a single transgene and a second AAV may carry an expression cassette which expresses a different transgene for co-expression in the host cell. However, the selected transgene may encode any biologically active product or other product, *e.g.*, a product desirable for study.

Suitable transgenes may be readily selected by one of skill in the art. The selection of the transgene is not considered to be a limitation of this invention.

2. Regulatory Elements

In addition to the major elements identified above for the minigene, the vector also includes conventional control elements which are operably linked to the transgene in a manner which permits its transcription, translation and/or expression in a cell transfected with the plasmid vector or infected with the virus produced by the invention. As used herein, "operably linked" sequences include both expression control sequences that are contiguous with the gene of interest and expression

control sequences that act in *trans* or at a distance to control the gene of interest.

Expression control sequences include appropriate transcription initiation, termination, promoter and enhancer sequences; efficient RNA processing signals such as splicing and polyadenylation (polyA) signals; sequences that stabilize cytoplasmic mRNA; sequences that enhance translation efficiency (*i.e.*, Kozak consensus sequence); sequences that enhance protein stability; and when desired, sequences that enhance secretion of the encoded product. A great number of expression control sequences, including promoters which are native, constitutive, inducible and/or tissue-specific, are known in the art and may be utilized.

Examples of constitutive promoters include, without limitation, the retroviral Rous sarcoma virus (RSV) LTR promoter (optionally with the RSV enhancer), the cytomegalovirus (CMV) promoter (optionally with the CMV enhancer) [see, *e.g.*, Boshart *et al*, *Cell*, **41**:521-530 (1985)], the SV40 promoter, the dihydrofolate reductase promoter, the β -actin promoter, the phosphoglycerol kinase (PGK) promoter, and the EF1 promoter [Invitrogen]. Inducible promoters allow regulation of gene expression and can be regulated by exogenously supplied compounds, environmental factors such as temperature, or the presence of a specific physiological state, *e.g.*, acute phase, a particular differentiation state of the cell, or in replicating cells only. Inducible promoters and inducible systems are available from a variety of commercial sources, including, without limitation, Invitrogen, Clontech and Ariad. Many other systems have been described and can be readily selected by one of skill in the art. Examples of inducible promoters regulated by exogenously supplied compounds, include, the zinc-inducible sheep metallothioneine (MT) promoter, the dexamethasone (Dex)-inducible mouse mammary tumor virus (MMTV) promoter, the T7 polymerase promoter system [International Patent Publication No. WO 98/10088]; the ecdysone insect promoter [No *et al*, *Proc. Natl. Acad. Sci. USA*, **93**:3346-3351 (1996)], the tetracycline-repressible system [Gossen *et al*, *Proc. Natl. Acad. Sci. USA*, **89**:5547-5551 (1992)], the tetracycline-inducible system [Gossen *et al*, *Science*, **268**:1766-1769 (1995), see also Harvey *et al*, *Curr. Opin. Chem. Biol.*, **2**:512-518 (1998)], the RU486-inducible system [Wang *et al*, *Nat. Biotech.*, **15**:239-243 (1997) and Wang *et al*, *Gene Ther.*, **4**:432-441 (1997)] and the rapamycin-inducible system [Magari *et al*, *J. Clin. Invest.*, **100**:2865-2872 (1997)]. Other types of inducible promoters which may be useful in this context are those which are

regulated by a specific physiological state, *e.g.*, temperature, acute phase, a particular differentiation state of the cell, or in replicating cells only.

In another embodiment, the native promoter for the transgene will be used. The native promoter may be preferred when it is desired that expression of the transgene should mimic the native expression. The native promoter may be used when expression of the transgene must be regulated temporally or developmentally, or in a tissue-specific manner, or in response to specific transcriptional stimuli. In a further embodiment, other native expression control elements, such as enhancer elements, polyadenylation sites or Kozak consensus sequences may also be used to mimic the native expression.

Another embodiment of the transgene includes a gene operably linked to a tissue-specific promoter. For instance, if expression in skeletal muscle is desired, a promoter active in muscle should be used. These include the promoters from genes encoding skeletal β -actin, myosin light chain 2A, dystrophin, muscle creatine kinase, as well as synthetic muscle promoters with activities higher than naturally-occurring promoters (see Li *et al.*, *Nat. Biotech.*, **17**:241-245 (1999)). Examples of promoters that are tissue-specific are known for liver (albumin, Miyatake *et al.*, *J. Virol.*, **71**:5124-32 (1997); hepatitis B virus core promoter, Sandig *et al.*, *Gene Ther.*, **3**:1002-9 (1996); alpha-fetoprotein (AFP), Arbuthnot *et al.*, *Hum. Gene Ther.*, **7**:1503-14 (1996)), bone osteocalcin (Stein *et al.*, *Mol. Biol. Rep.*, **24**:185-96 (1997)); bone sialoprotein (Chen *et al.*, *J. Bone Miner. Res.*, **11**:654-64 (1996)), lymphocytes (CD2, Hansal *et al.*, *J. Immunol.*, **161**:1063-8 (1998); immunoglobulin heavy chain; T cell receptor chain), neuronal such as neuron-specific enolase (NSE) promoter (Andersen *et al.*, *Cell. Mol. Neurobiol.*, **13**:503-15 (1993)), neurofilament light-chain gene (Piccioli *et al.*, *Proc. Natl. Acad. Sci. USA*, **88**:5611-5 (1991)), and the neuron-specific *vgf* gene (Piccioli *et al.*, *Neuron*, **15**:373-84 (1995)), among others.

Optionally, plasmids carrying therapeutically useful transgenes may also include selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. Such selectable reporters or marker genes (preferably located outside the viral genome to be rescued by the method of the invention) can be used to signal the presence of the plasmids in bacterial cells, such as ampicillin resistance. Other components of the plasmid may include an origin of

replication. Selection of these and other promoters and vector elements are conventional and many such sequences are available [see, *e.g.*, Sambrook et al, and references cited therein].

5 The combination of the transgene, promoter/enhancer, and 5' and 3' AAV ITRs is referred to as a "minigene" for ease of reference herein. Provided with the teachings of this invention, the design of such a minigene can be made by resort to conventional techniques.

3. Delivery of the Minigene to a Packaging Host Cell

10 The minigene can be carried on any suitable vector, *e.g.*, a plasmid, which is delivered to a host cell. The plasmids useful in this invention may be engineered such that they are suitable for replication and, optionally, integration in prokaryotic cells, mammalian cells, or both. These plasmids (or other vectors carrying the 5' AAV ITR-heterologous molecule-3' AAV ITR) contain sequences permitting replication of the minigene in eukaryotes and/or prokaryotes and selection markers for
15 these systems. Selectable markers or reporter genes may include sequences encoding geneticin, hygromycin or purimycin resistance, among others. The plasmids may also contain certain selectable reporters or marker genes that can be used to signal the presence of the vector in bacterial cells, such as ampicillin resistance. Other components of the plasmid may include an origin of replication and an amplicon, such as the amplicon
20 system employing the Epstein Barr virus nuclear antigen. This amplicon system, or other similar amplicon components permit high copy episomal replication in the cells. Preferably, the molecule carrying the minigene is transfected into the cell, where it may exist transiently. Alternatively, the minigene (carrying the 5' AAV ITR-heterologous molecule-3' ITR) may be stably integrated into the genome of the host cell, either
25 chromosomally or as an episome. In certain embodiments, the minigene may be present in multiple copies, optionally in head-to-head, head-to-tail, or tail-to-tail concatamers. Suitable transfection techniques are known and may readily be utilized to deliver the minigene to the host cell.

30 Generally, when delivering the vector comprising the minigene by transfection, the vector is delivered in an amount from about 5 µg to about 100 µg DNA, about 10 µg to about 50 µg DNA to about 1×10^4 cells to about 1×10^{13} cells, or about

1 x 10⁵ cells. However, the relative amounts of vector DNA to host cells may be adjusted, taking into consideration such factors as the selected vector, the delivery method and the host cells selected.

B. *Rep and Cap Sequences*

5 In addition to the minigene, the host cell contains the sequences which drive expression of a novel AAV capsid protein of the invention (or a capsid protein comprising a fragment thereof) in the host cell and rep sequences of the same source as the source of the AAV ITRs found in the minigene, or a cross-complementing source. The AAV *cap* and *rep* sequences may be independently obtained from an AAV source as described above and may be introduced into the host cell in any manner known to one in the art as described above. Additionally, when pseudotyping an AAV vector, the sequences encoding each of the essential rep proteins may be supplied by different AAV sources (*e.g.*, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9). For example, the *rep78/68* sequences may be from AAV2, whereas the *rep52/40* sequences may be from AAV8.

15 In one embodiment, the host cell stably contains the capsid protein under the control of a suitable promoter, such as those described above. Most desirably, in this embodiment, the capsid protein is expressed under the control of an inducible promoter. In another embodiment, the capsid protein is supplied to the host cell in *trans*. When delivered to the host cell in *trans*, the capsid protein may be delivered via a plasmid which contains the sequences necessary to direct expression of the selected capsid protein in the host cell. Most desirably, when delivered to the host cell in *trans*, the plasmid carrying the capsid protein also carries other sequences required for packaging the rAAV, *e.g.*, the *rep* sequences.

25 In another embodiment, the host cell stably contains the *rep* sequences under the control of a suitable promoter, such as those described above. Most desirably, in this embodiment, the essential rep proteins are expressed under the control of an inducible promoter. In another embodiment, the rep proteins are supplied to the host cell in *trans*. When delivered to the host cell in *trans*, the rep proteins may be delivered via a plasmid which contains the sequences necessary to direct expression of the selected rep proteins in the host cell. Most desirably, when delivered to the host cell in *trans*, the

plasmid carrying the capsid protein also carries other sequences required for packaging the rAAV, *e.g.*, the *rep* and *cap* sequences.

Thus, in one embodiment, the *rep* and *cap* sequences may be transfected into the host cell on a single nucleic acid molecule and exist stably in the cell as an episome. In another embodiment, the *rep* and *cap* sequences are stably integrated into the chromosome of the cell. Another embodiment has the *rep* and *cap* sequences transiently expressed in the host cell. For example, a useful nucleic acid molecule for such transfection comprises, from 5' to 3', a promoter, an optional spacer interposed between the promoter and the start site of the *rep* gene sequence, an AAV *rep* gene sequence, and an AAV *cap* gene sequence.

Optionally, the *rep* and/or *cap* sequences may be supplied on a vector that contains other DNA sequences that are to be introduced into the host cells. For instance, the vector may contain the rAAV construct comprising the minigene. The vector may comprise one or more of the genes encoding the helper functions, *e.g.*, the adenoviral proteins E1, E2a, and E4 ORF6, and the gene for VAI RNA.

Preferably, the promoter used in this construct may be any of the constitutive, inducible or native promoters known to one of skill in the art or as discussed above. In one embodiment, an AAV P5 promoter sequence is employed. The selection of the AAV to provide any of these sequences does not limit the invention.

In another preferred embodiment, the promoter for *rep* is an inducible promoter, such as are discussed above in connection with the transgene regulatory elements. One preferred promoter for *rep* expression is the T7 promoter. The vector comprising the *rep* gene regulated by the T7 promoter and the *cap* gene, is transfected or transformed into a cell which either constitutively or inducibly expresses the T7 polymerase. See International Patent Publication No. WO 98/10088, published March 12, 1998.

The spacer is an optional element in the design of the vector. The spacer is a DNA sequence interposed between the promoter and the *rep* gene ATG start site. The spacer may have any desired design; that is, it may be a random sequence of nucleotides, or alternatively, it may encode a gene product, such as a marker gene. The spacer may contain genes which typically incorporate start/stop and polyA sites. The spacer may be a non-coding DNA sequence from a prokaryote or eukaryote, a repetitive

non-coding sequence, a coding sequence without transcriptional controls or a coding sequence with transcriptional controls. Two exemplary sources of spacer sequences are the phage ladder sequences or yeast ladder sequences, which are available commercially, e.g., from Gibco or Invitrogen, among others. The spacer may be of any size sufficient to
5 reduce expression of the *rep78* and *rep68* gene products, leaving the *rep52*, *rep40* and *cap* gene products expressed at normal levels. The length of the spacer may therefore range from about 10 bp to about 10.0 kbp, preferably in the range of about 100 bp to about 8.0 kbp. To reduce the possibility of recombination, the spacer is preferably less than 2 kbp in length; however, the invention is not so limited.

10 Although the molecule(s) providing *rep* and *cap* may exist in the host cell transiently (*i.e.*, through transfection), it is preferred that one or both of the *rep* and *cap* proteins and the promoter(s) controlling their expression be stably expressed in the host cell, e.g., as an episome or by integration into the chromosome of the host cell. The methods employed for constructing embodiments of this invention are conventional
15 genetic engineering or recombinant engineering techniques such as those described in the references above. While this specification provides illustrative examples of specific constructs, using the information provided herein, one of skill in the art may select and design other suitable constructs, using a choice of spacers, P5 promoters, and other elements, including at least one translational start and stop signal, and the optional
20 addition of polyadenylation sites.

In another embodiment of this invention, the *rep* or *cap* protein may be provided stably by a host cell.

C. The Helper Functions

25 The packaging host cell also requires helper functions in order to package the rAAV of the invention. Optionally, these functions may be supplied by a herpesvirus. Most desirably, the necessary helper functions are each provided from a human or non-human primate adenovirus source, such as those described above and/or are available from a variety of sources, including the American Type Culture Collection (ATCC), Manassas, VA (US). In one currently preferred embodiment, the host cell is provided
30 with and/or contains an E1a gene product, an E1b gene product, an E2a gene product, and/or an E4 ORF6 gene product. The host cell may contain other adenoviral genes such

as VAI RNA, but these genes are not required. In a preferred embodiment, no other adenovirus genes or gene functions are present in the host cell.

By “adenoviral DNA which expresses the E1a gene product”, it is meant any adenovirus sequence encoding E1a or any functional E1a portion. Adenoviral DNA which expresses the E2a gene product and adenoviral DNA which expresses the E4 ORF6 gene products are defined similarly. Also included are any alleles or other modifications of the adenoviral gene or functional portion thereof. Such modifications may be deliberately introduced by resort to conventional genetic engineering or mutagenic techniques to enhance the adenoviral function in some manner, as well as naturally occurring allelic variants thereof. Such modifications and methods for manipulating DNA to achieve these adenovirus gene functions are known to those of skill in the art.

The adenovirus E1a, E1b, E2a, and/or E4ORF6 gene products, as well as any other desired helper functions, can be provided using any means that allows their expression in a cell. Each of the sequences encoding these products may be on a separate vector, or one or more genes may be on the same vector. The vector may be any vector known in the art or disclosed above, including plasmids, cosmids and viruses. Introduction into the host cell of the vector may be achieved by any means known in the art or as disclosed above, including transfection, infection, electroporation, liposome delivery, membrane fusion techniques, high velocity DNA-coated pellets, viral infection and protoplast fusion, among others. One or more of the adenoviral genes may be stably integrated into the genome of the host cell, stably expressed as episomes, or expressed transiently. The gene products may all be expressed transiently, on an episome or stably integrated, or some of the gene products may be expressed stably while others are expressed transiently. Furthermore, the promoters for each of the adenoviral genes may be selected independently from a constitutive promoter, an inducible promoter or a native adenoviral promoter. The promoters may be regulated by a specific physiological state of the organism or cell (*i.e.*, by the differentiation state or in replicating or quiescent cells) or by exogenously added factors, for example.

D. Host Cells And Packaging Cell Lines

The host cell itself may be selected from any biological organism, including prokaryotic (*e.g.*, bacterial) cells, and eukaryotic cells, including, insect cells,

yeast cells and mammalian cells. Particularly desirable host cells are selected from among any mammalian species, including, without limitation, cells such as A549, WEHI, 3T3, 10T1/2, BHK, MDCK, COS 1, COS 7, BSC 1, BSC 40, BMT 10, VERO, WI38, HeLa, 293 cells (which express functional adenoviral E1), Saos, C2C12, L cells, HT1080, HepG2 and primary fibroblast, hepatocyte and myoblast cells derived from mammals including human, monkey, mouse, rat, rabbit, and hamster. The selection of the mammalian species providing the cells is not a limitation of this invention; nor is the type of mammalian cell, *i.e.*, fibroblast, hepatocyte, tumor cell, etc. The requirements for the cell used is that it not carry any adenovirus gene other than E1, E2a and/or E4 ORF6; it not contain any other virus gene which could result in homologous recombination of a contaminating virus during the production of rAAV; and it is capable of infection or transfection of DNA and expression of the transfected DNA. In a preferred embodiment, the host cell is one that has *rep* and *cap* stably transfected in the cell.

One host cell useful in the present invention is a host cell stably transformed with the sequences encoding *rep* and *cap*, and which is transfected with the adenovirus E1, E2a, and E4ORF6 DNA and a construct carrying the minigene as described above. Stable *rep* and/or *cap* expressing cell lines, such as B-50 (International Patent Application Publication No. WO 99/15685), or those described in US Patent No. 5,658,785, may also be similarly employed. Another desirable host cell contains the minimum adenoviral DNA which is sufficient to express E4 ORF6. Yet other cell lines can be constructed using the novel singleton-corrected AAV *cap* sequences of the invention.

The preparation of a host cell according to this invention involves techniques such as assembly of selected DNA sequences. This assembly may be accomplished utilizing conventional techniques. Such techniques include cDNA and genomic cloning, which are well known and are described in Sambrook *et al.*, cited above, use of overlapping oligonucleotide sequences of the adenovirus and AAV genomes, combined with polymerase chain reaction, synthetic methods, and any other suitable methods which provide the desired nucleotide sequence.

Introduction of the molecules (as plasmids or viruses) into the host cell may also be accomplished using techniques known to the skilled artisan and as discussed throughout the specification. In preferred embodiment, standard transfection techniques

are used, *e.g.*, CaPO₄ transfection or electroporation, and/or infection by hybrid adenovirus/AAV vectors into cell lines such as the human embryonic kidney cell line HEK 293 (a human kidney cell line containing functional adenovirus E1 genes which provides *trans*-acting E1 proteins).

5 One of skill in the art will readily understand that the novel AAV sequences of the invention can be readily adapted for use in these and other viral vector systems for *in vitro*, *ex vivo* or *in vivo* gene delivery. Similarly, one of skill in the art can readily select other fragments of the AAV genome of the invention for use in a variety of rAAV and non-rAAV vector systems. Such vectors systems may include, *e.g.*,
10 lentiviruses, retroviruses, poxviruses, vaccinia viruses, and adenoviral systems, among others. Selection of these vector systems is not a limitation of the present invention.

 Thus, the invention further provides vectors generated using the nucleic acid and amino acid sequences of the novel AAV of the invention. Such vectors are useful for a variety of purposes, including for delivery of therapeutic molecules and for
15 use in vaccine regimens. Particularly desirable for delivery of therapeutic molecules are recombinant AAV containing capsids of the novel AAV of the invention. These, or other vector constructs containing novel AAV sequences of the invention may be used in vaccine regimens, *e.g.*, for co-delivery of a cytokine, or for delivery of the immunogen itself.

20 IV. Recombinant Viruses And Uses Therefor

 Using the techniques described herein, one of skill in the art can generate a rAAV having a capsid of an AAV of the invention or having a capsid containing one or more fragments of an AAV of the invention. In one embodiment, a full-length capsid from a
25 singleton-corrected AAV can be utilized.

A. Delivery of Viruses

 In another aspect, the present invention provides a method for delivery of a transgene to a host which involves transfecting or infecting a selected host cell with a recombinant viral vector generated with the singleton-corrected AAV (or functional
30 fragments thereof) of the invention. Methods for delivery are well known to those of skill in the art and are not a limitation of the present invention.

In one desirable embodiment, the invention provides a method for AAV-mediated delivery of a transgene to a host. This method involves transfecting or infecting a selected host cell with a recombinant viral vector containing a selected transgene under the control of sequences that direct expression thereof and the modified capsid proteins of the capsids.

Optionally, a sample from the host may be first assayed for the presence of antibodies to a selected AAV source (*e.g.*, a serotype). A variety of assay formats for detecting neutralizing antibodies are well known to those of skill in the art. The selection of such an assay is not a limitation of the present invention. See, *e.g.*, Fisher et al, *Nature Med.*, 3(3):306-312 (March 1997) and W. C. Manning et al, *Human Gene Therapy*, 9:477-485 (March 1, 1998). The results of this assay may be used to determine which AAV vector containing capsid proteins of a particular source are preferred for delivery, *e.g.*, by the absence of neutralizing antibodies specific for that capsid source.

In one aspect of this method, the delivery of vector with AAV capsid proteins of the invention may precede or follow delivery of a gene via a vector with a different AAV capsid protein. Thus, gene delivery via rAAV vectors may be used for repeat gene delivery to a selected host cell. Desirably, subsequently administered rAAV vectors carry the same transgene as the first rAAV vector, but the subsequently administered vectors contain capsid proteins of sources (and preferably, different serotypes) which differ from the first vector. For example, if a first vector has a singleton-corrected capsid proteins, subsequently administered vectors may have capsid proteins selected from among the other AAV, optionally, from another serotype or from another clade.

Optionally, multiple rAAV vectors can be used to deliver large transgenes or multiple transgenes by co-administration of rAAV vectors concatamerize *in vivo* to form a single vector genome. In such an embodiment, a first AAV may carry an expression cassette which expresses a single transgene (or a subunit thereof) and a second AAV may carry an expression cassette which expresses a second transgene (or a different subunit) for co-expression in the host cell. A first AAV may carry an expression cassette which is a first piece of a polycistronic construct (*e.g.*, a promoter and transgene, or subunit) and a second AAV may carry an expression cassette which is a second piece of a polycistronic construct (*e.g.*, transgene or subunit and a polyA sequence). These two

pieces of a polycistronic construct concatamerize *in vivo* to form a single vector genome that co-expresses the transgenes delivered by the first and second AAV. In such embodiments, the rAAV vector carrying the first expression cassette and the rAAV vector carrying the second expression cassette can be delivered in a single pharmaceutical composition. In other embodiments, the two or more rAAV vectors are delivered as separate pharmaceutical compositions which can be administered substantially simultaneously, or shortly before or after one another.

The above-described recombinant vectors may be delivered to host cells according to published methods. The rAAV, preferably suspended in a physiologically compatible carrier, may be administered to a human or non-human mammalian patient. Suitable carriers may be readily selected by one of skill in the art in view of the indication for which the transfer virus is directed. For example, one suitable carrier includes saline, which may be formulated with a variety of buffering solutions (*e.g.*, phosphate buffered saline). Other exemplary carriers include sterile saline, lactose, sucrose, calcium phosphate, gelatin, dextran, agar, pectin, peanut oil, sesame oil, and water. The selection of the carrier is not a limitation of the present invention.

Optionally, the compositions of the invention may contain, in addition to the rAAV and carrier(s), other conventional pharmaceutical ingredients, such as preservatives, or chemical stabilizers. Suitable exemplary preservatives include chlorobutanol, potassium sorbate, sorbic acid, sulfur dioxide, propyl gallate, the parabens, ethyl vanillin, glycerin, phenol, and parachlorophenol. Suitable chemical stabilizers include gelatin and albumin.

The vectors are administered in sufficient amounts to transfect the cells and to provide sufficient levels of gene transfer and expression to provide a therapeutic benefit without undue adverse effects, or with medically acceptable physiological effects, which can be determined by those skilled in the medical arts. Conventional and pharmaceutically acceptable routes of administration include, but are not limited to, direct delivery to a desired organ (*e.g.*, the liver (optionally via the hepatic artery) or lung), oral, inhalation, intranasal, intratracheal, intraarterial, intraocular, intravenous, intramuscular, subcutaneous, intradermal, and other parental routes of administration. Routes of administration may be combined, if desired.

Dosages of the viral vector will depend primarily on factors such as the condition being treated, the age, weight and health of the patient, and may thus vary among patients. For example, a therapeutically effective human dosage of the viral vector is generally in the range of from about 0.1 mL to about 100 mL of solution containing concentrations of from about 1×10^9 to 1×10^{16} genomes virus vector. A preferred human dosage for delivery to large organs (*e.g.*, liver, muscle, heart and lung) may be about 5×10^{10} to 5×10^{13} AAV genomes per 1 kg, at a volume of about 1 to 100 mL. A preferred dosage for delivery to eye is about 5×10^9 to 5×10^{12} genome copies, at a volume of about 0.1 mL to 1 mL. The dosage will be adjusted to balance the therapeutic benefit against any side effects and such dosages may vary depending upon the therapeutic application for which the recombinant vector is employed. The levels of expression of the transgene can be monitored to determine the frequency of dosage resulting in viral vectors, preferably AAV vectors containing the minigene. Optionally, dosage regimens similar to those described for therapeutic purposes may be utilized for immunization using the compositions of the invention.

Examples of therapeutic products and immunogenic products for delivery by the AAV-containing vectors of the invention are provided below. These vectors may be used for a variety of therapeutic or vaccinal regimens, as described herein. Additionally, these vectors may be delivered in combination with one or more other vectors or active ingredients in a desired therapeutic and/or vaccinal regimen.

B. Therapeutic Transgenes

Useful therapeutic products encoded by the transgene include hormones and growth and differentiation factors including, without limitation, insulin, glucagon, growth hormone (GH), parathyroid hormone (PTH), growth hormone releasing factor (GRF), follicle stimulating hormone (FSH), luteinizing hormone (LH), human chorionic gonadotropin (hCG), vascular endothelial growth factor (VEGF), angiopoietins, angiostatin, granulocyte colony stimulating factor (GCSF), erythropoietin (EPO), connective tissue growth factor (CTGF), basic fibroblast growth factor (bFGF), acidic fibroblast growth factor (aFGF), epidermal growth factor (EGF), platelet-derived growth factor (PDGF), insulin growth factors I and II (IGF-I and IGF-II), any one of the transforming growth factor α superfamily, including TGF α , activins, inhibins, or any of the bone morphogenic proteins (BMP) BMPs 1-15, any one of the

heregulin/neuregulin/ARIA/neu differentiation factor (NDF) family of growth factors, nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophins NT-3 and NT-4/5, ciliary neurotrophic factor (CNTF), glial cell line derived neurotrophic factor (GDNF), neurturin, agrin, any one of the family of semaphorins/collapsins, netrin-1 and netrin-2, hepatocyte growth factor (HGF), ephrins, noggin, sonic hedgehog and tyrosine hydroxylase.

Other useful transgene products include proteins that regulate the immune system including, without limitation, cytokines and lymphokines such as thrombopoietin (TPO), interleukins (IL) IL-1 through IL-25 (including, *e.g.*, IL-2, IL-4, IL-12 and IL-18), monocyte chemoattractant protein, leukemia inhibitory factor, granulocyte-macrophage colony stimulating factor, Fas ligand, tumor necrosis factors α and β , interferons α , β , and γ , stem cell factor, flk-2/flt3 ligand. Gene products produced by the immune system are also useful in the invention. These include, without limitations, immunoglobulins IgG, IgM, IgA, IgD and IgE, chimeric immunoglobulins, humanized antibodies, single chain antibodies, T cell receptors, chimeric T cell receptors, single chain T cell receptors, class I and class II MHC molecules, as well as engineered immunoglobulins and MHC molecules. Useful gene products also include complement regulatory proteins such as complement regulatory proteins, membrane cofactor protein (MCP), decay accelerating factor (DAF), CR1, CF2 and CD59.

Still other useful gene products include any one of the receptors for the hormones, growth factors, cytokines, lymphokines, regulatory proteins and immune system proteins. The invention encompasses receptors for cholesterol regulation and/or lipid modulation, including the low density lipoprotein (LDL) receptor, high density lipoprotein (HDL) receptor, the very low density lipoprotein (VLDL) receptor, and scavenger receptors. The invention also encompasses gene products such as members of the steroid hormone receptor superfamily including glucocorticoid receptors and estrogen receptors, Vitamin D receptors and other nuclear receptors. In addition, useful gene products include transcription factors such as *jun*, *fos*, max, mad, serum response factor (SRF), AP-1, AP2, *myb*, MyoD and myogenin, ETS-box containing proteins, TFE3, E2F, ATF1, ATF2, ATF3, ATF4, ZF5, NFAT, CREB, HNF-4, C/EBP, SP1, CCAAT-box binding proteins, interferon regulation factor (IRF-1), Wilms tumor protein, ETS-binding

protein, STAT, GATA-box binding proteins, *e.g.*, GATA-3, and the forkhead family of winged helix proteins.

Other useful gene products include, carbamoyl synthetase I, ornithine transcarbamylase, arginosuccinate synthetase, arginosuccinate lyase, arginase, 5 fumarylacetate hydrolase, phenylalanine hydroxylase, alpha-1 antitrypsin, glucose-6-phosphatase, porphobilinogen deaminase, cystathione beta-synthase, branched chain ketoacid decarboxylase, albumin, isovaleryl-coA dehydrogenase, propionyl CoA carboxylase, methyl malonyl CoA mutase, glutaryl CoA dehydrogenase, insulin, beta-glucosidase, pyruvate carboxylate, hepatic phosphorylase, phosphorylase kinase, glycine 10 decarboxylase, H-protein, T-protein, a cystic fibrosis transmembrane regulator (CFTR) sequence, and a dystrophin gene product [*e.g.*, a mini- or micro-dystrophin]. Still other useful gene products include enzymes such as may be useful in enzyme replacement therapy, which is useful in a variety of conditions resulting from deficient activity of enzyme. For example, enzymes that contain mannose-6-phosphate may be utilized in 15 therapies for lysosomal storage diseases (*e.g.*, a suitable gene includes that encoding β -glucuronidase (GUSB)).

Still other useful gene products include those used for treatment of hemophilia, including hemophilia B (including Factor IX) and hemophilia A (including Factor VIII and its variants, such as the light chain and heavy chain of the heterodimer 20 and the B-deleted domain; US Patent No. 6,200,560 and US Patent No. 6,221,349). The Factor VIII gene codes for 2351 amino acids and the protein has six domains, designated from the amino to the terminal carboxy terminus as A1-A2-B-A3-C1-C2 [Wood *et al*, *Nature*, **312**:330 (1984); Vehar *et al.*, *Nature* **312**:337 (1984); and Toole *et al*, *Nature*, **342**:337 (1984)]. Human Factor VIII is processed within the cell to yield a heterodimer 25 primarily comprising a heavy chain containing the A1, A2 and B domains and a light chain containing the A3, C1 and C2 domains. Both the single chain polypeptide and the heterodimer circulate in the plasma as inactive precursors, until activated by thrombin cleavage between the A2 and B domains, which releases the B domain and results in a heavy chain consisting of the A1 and A2 domains. The B domain is deleted in the 30 activated procoagulant form of the protein. Additionally, in the native protein, two polypeptide chains (“a” and “b”), flanking the B domain, are bound to a divalent calcium cation.

In some embodiments, the minigene comprises first 57 base pairs of the Factor VIII heavy chain which encodes the 10 amino acid signal sequence, as well as the human growth hormone (hGH) polyadenylation sequence. In alternative embodiments, the minigene further comprises the A1 and A2 domains, as well as 5 amino acids from the N-terminus of the B domain, and/or 85 amino acids of the C-terminus of the B domain, as well as the A3, C1 and C2 domains. In yet other embodiments, the nucleic acids encoding Factor VIII heavy chain and light chain are provided in a single minigene separated by 42 nucleic acids coding for 14 amino acids of the B domain [US Patent No. 6,200,560].

As used herein, a therapeutically effective amount is an amount of AAV vector that produces sufficient amounts of Factor VIII to decrease the time it takes for a subject's blood to clot. Generally, severe hemophiliacs having less than 1% of normal levels of Factor VIII have a whole blood clotting time of greater than 60 minutes as compared to approximately 10 minutes for non-hemophiliacs.

The present invention is not limited to any specific Factor VIII sequence. Many natural and recombinant forms of Factor VIII have been isolated and generated. Examples of naturally occurring and recombinant forms of Factor VII can be found in the patent and scientific literature including, US Patent No. 5,563,045, US Patent No. 5,451,521, US Patent No. 5,422,260, US Patent No. 5,004,803, US Patent No. 4,757,006, US Patent No. 5,661,008, US Patent No. 5,789,203, US Patent No. 5,681,746, US Patent No. 5,595,886, US Patent No. 5,045,455, US Patent No. 5,668,108, US Patent No. 5,633,150, US Patent No. 5,693,499, US Patent No. 5,587,310, US Patent No. 5,171,844, US Patent No. 5,149,637, US Patent No. 5,112,950, US Patent No. 4,886,876; International Patent Publication Nos. WO 94/11503, WO 87/07144, WO 92/16557, WO 91/09122, WO 97/03195, WO 96/21035, and WO 91/07490; European Patent Application Nos. EP 0 672 138, EP 0 270 618, EP 0 182 448, EP 0 162 067, EP 0 786 474, EP 0 533 862, EP 0 506 757, EP 0 874 057, EP 0 795 021, EP 0 670 332, EP 0 500 734, EP 0 232 112, and EP 0 160 457; Sanberg *et al.*, XXth Int. Congress of the World Fed. Of Hemophilia (1992), and Lind *et al.*, *Eur. J. Biochem.*, **232**:19 (1995).

Nucleic acids sequences coding for the above-described Factor VIII can be obtained using recombinant methods or by deriving the sequence from a vector known to include the same. Furthermore, the desired sequence can be isolated directly from cells

and tissues containing the same, using standard techniques, such as phenol extraction and PCR of cDNA or genomic DNA [See, *e.g.*, Sambrook et al]. Nucleotide sequences can also be produced synthetically, rather than cloned. The complete sequence can be assembled from overlapping oligonucleotides prepared by standard methods and
5 assembled into a complete coding sequence [See, *e.g.*, Edge, *Nature* **292**:757 (1981); Nambari *et al*, *Science*, **223**:1299 (1984); and Jay *et al*, *J. Biol. Chem.* **259**:6311 (1984).

Furthermore, the invention is not limited to human Factor VIII. Indeed, it is intended that the present invention encompass Factor VIII from animals other than humans, including but not limited to companion animals (*e.g.*, canine, felines, and
10 equines), livestock (*e.g.*, bovines, caprines and ovines), laboratory animals, marine mammals, large cats, etc.

The AAV vectors may contain a nucleic acid coding for fragments of Factor VIII which is itself not biologically active, yet when administered into the subject improves or restores the blood clotting time. For example, as discussed above, the Factor
15 VIII protein comprises two polypeptide chains: a heavy chain and a light chain separated by a B-domain which is cleaved during processing. As demonstrated by the present invention, co-transducing recipient cells with the Factor VIII heavy and light chains leads to the expression of biologically active Factor VIII. Because most hemophiliacs contain a mutation or deletion in only one of the chains (*e.g.*, heavy or light chain), it may be
20 possible to administer only the chain defective in the patient to supply the other chain.

Other useful gene products include non-naturally occurring polypeptides, such as chimeric or hybrid polypeptides having a non-naturally occurring amino acid sequence containing insertions, deletions or amino acid substitutions. For example, single-chain engineered immunoglobulins could be useful in certain
25 immunocompromised patients. Other types of non-naturally occurring gene sequences include antisense molecules and catalytic nucleic acids, such as ribozymes, which could be used to reduce overexpression of a target.

Reduction and/or modulation of expression of a gene is particularly desirable for treatment of hyperproliferative conditions characterized by
30 hyperproliferating cells, as are cancers and psoriasis. Target polypeptides include those polypeptides which are produced exclusively or at higher levels in hyperproliferative cells as compared to normal cells. Target antigens include polypeptides encoded by oncogenes

such as myb, myc, fyn, and the translocation gene bcr/abl, ras, src, P53, neu, trk and EGRF. In addition to oncogene products as target antigens, target polypeptides for anti-cancer treatments and protective regimens include variable regions of antibodies made by B cell lymphomas and variable regions of T cell receptors of T cell lymphomas which, in some embodiments, are also used as target antigens for autoimmune disease. Other tumor-associated polypeptides can be used as target polypeptides such as polypeptides which are found at higher levels in tumor cells including the polypeptide recognized by monoclonal antibody 17-1A and folate binding polypeptides.

Other suitable therapeutic polypeptides and proteins include those which may be useful for treating individuals suffering from autoimmune diseases and disorders by conferring a broad based protective immune response against targets that are associated with autoimmunity including cell receptors and cells which produce "self"-directed antibodies. T cell mediated autoimmune diseases include Rheumatoid arthritis (RA), multiple sclerosis (MS), Sjögren's syndrome, sarcoidosis, insulin dependent diabetes mellitus (IDDM), autoimmune thyroiditis, reactive arthritis, ankylosing spondylitis, scleroderma, polymyositis, dermatomyositis, psoriasis, vasculitis, Wegener's granulomatosis, Crohn's disease and ulcerative colitis. Each of these diseases is characterized by T cell receptors (TCRs) that bind to endogenous antigens and initiate the inflammatory cascade associated with autoimmune diseases.

C. Immunogenic Transgenes

Suitably, the AAV vectors of the invention avoid the generation of immune responses to the AAV sequences contained within the vector. However, these vectors may nonetheless be formulated in a manner that permits the expression of a transgene carried by the vectors to induce an immune response to a selected antigen. For example, in order to promote an immune response, the transgene may be expressed from a constitutive promoter, the vector can be adjuvanted as described herein, and/or the vector can be put into degenerating tissue.

Examples of suitable immunogenic transgenes include those selected from a variety of viral families. Examples of desirable viral families against which an immune response would be desirable include, the picornavirus family, which includes the genera rhinoviruses, which are responsible for about 50% of cases of the common cold; the genera enteroviruses, which include polioviruses, coxsackieviruses, echoviruses, and

human enteroviruses such as hepatitis A virus; and the genera aphthoviruses, which are responsible for foot and mouth diseases, primarily in non-human animals. Within the picornavirus family of viruses, target antigens include the VP1, VP2, VP3, VP4, and VPG. Other viral families include the astroviruses and the calicivirus family. The

5 calicivirus family encompasses the Norwalk group of viruses, which are an important causative agent of epidemic gastroenteritis. Still another viral family desirable for use in targeting antigens for inducing immune responses in humans and non-human animals is the togavirus family, which includes the genera alphavirus, which include Sindbis viruses, Ross River virus, and Venezuelan, Eastern & Western Equine encephalitis, and rubivirus,

10 including Rubella virus. The flaviviridae family includes dengue, yellow fever, Japanese encephalitis, St. Louis encephalitis and tick borne encephalitis viruses. Other target antigens may be generated from the Hepatitis C or the coronavirus family, which includes a number of non-human viruses such as infectious bronchitis virus (poultry), porcine transmissible gastroenteric virus (pig), porcine hemagglutinin encephalomyelitis virus

15 (pig), feline infectious peritonitis virus (cat), feline enteric coronavirus (cat), canine coronavirus (dog), and human respiratory coronaviruses, which may cause the common cold and/or non-A, B or C hepatitis, and which include the putative cause of sudden acute respiratory syndrome (SARS). Within the coronavirus family, target antigens include the E1 (also called M or matrix protein), E2 (also called S or Spike protein), E3 (also called

20 HE or hemagglutinin-esterase) glycoprotein (not present in all coronaviruses), or N (nucleocapsid). Still other antigens may be targeted against the arterivirus family and the rhabdovirus family. The rhabdovirus family includes the genera vesiculovirus (*e.g.*, Vesicular Stomatitis Virus), and the general lyssavirus (*e.g.*, rabies). Within the rhabdovirus family, suitable antigens may be derived from the G protein or the N protein.

25 The family filoviridae, which includes hemorrhagic fever viruses such as Marburg and Ebola virus may be a suitable source of antigens. The paramyxovirus family includes parainfluenza Virus Type 1, parainfluenza Virus Type 3, bovine parainfluenza Virus Type 3, rubulavirus (mumps virus, parainfluenza Virus Type 2, parainfluenza virus Type 4, Newcastle disease virus (chickens), rinderpest, morbillivirus, which includes measles and

30 canine distemper, and pneumovirus, which includes respiratory syncytial virus. The influenza virus is classified within the family orthomyxovirus and is a suitable source of antigen (*e.g.*, the HA protein, the N1 protein). The bunyavirus family includes the genera

bunyavirus (California encephalitis, La Crosse), phlebovirus (Rift Valley Fever), hantavirus (puremala is a hemahagin fever virus), nairovirus (Nairobi sheep disease) and various unassigned bungaviruses. The arenavirus family provides a source of antigens against LCM and Lassa fever virus. Another source of antigens is the bornavirus family.

5 The reovirus family includes the genera reovirus, rotavirus (which causes acute gastroenteritis in children), orbiviruses, and cultivirus (Colorado Tick fever, Lebombo (humans), equine encephalosis, blue tongue). The retrovirus family includes the sub-family oncorivirinal which encompasses such human and veterinary diseases as feline leukemia virus, HTLVI and HTLVII, lentivirinal (which includes HIV, simian
10 immunodeficiency virus, feline immunodeficiency virus, equine infectious anemia virus, and spumavirinal).

With respect to HIV and SIV, many suitable antigens have been described and can readily be selected. Examples of suitable HIV and SIV antigens include, without limitation the gag, pol, Vif, Vpx, VPR, Env, Tat and Rev proteins, as well as various
15 fragments thereof. For example, suitable fragments of the envelope (env) protein include, *e.g.*, gp41, gp140, and gp120. In addition, a variety of modifications to these and other HIV and SIV antigens have been described. Suitable antigens for this purpose are known to those of skill in the art. For example, one may select a sequence encoding the gag, pol, Vif, and Vpr, Env, Tat and Rev, amongst other proteins. See, *e.g.*, the modified gag
20 protein which is described in US Patent 5,972,596. See, also, the HIV and SIV proteins described in D.H. Barouch et al, J. Virol., 75(5):2462-2467 (March 2001), and R.R. Amara, et al, Science, 292:69-74 (6 April 2001). These proteins or subunits thereof may be delivered alone, or in combination via separate vectors or from a single vector.

The papovavirus family includes the sub-family polyomaviruses (BKU
25 and JCU viruses) and the sub-family papillomavirus (associated with cancers or malignant progression of papilloma). The adenovirus family includes viruses (EX, AD7, ARD, O.B.) which cause respiratory disease and/or enteritis. The parvovirus family includes feline parvovirus (feline enteritis), feline panleucopeniavirus, canine parvovirus, and porcine parvovirus. The herpesvirus family includes the sub-family
30 alphaherpesvirinae, which encompasses the genera simplexvirus (HSVI, HSVII), varicellovirus (pseudorabies, varicella zoster) and the sub-family betaherpesvirinae, which includes the genera cytomegalovirus (HCMV, muromegalovirus) and the

sub-family gammaherpesvirinae, which includes the genera lymphocryptovirus, EBV (Burkitt's lymphoma), human herpesviruses 6A, 6B and 7, Kaposi's sarcoma-associated herpesvirus and cercopithecine herpesvirus (B virus), infectious rhinotracheitis, Marek's disease virus, and rhadinovirus. The poxvirus family includes the sub-family chordopoxvirinae, which encompasses the genera orthopoxvirus (Variola major (Smallpox) and Vaccinia (Cowpox)), parapoxvirus, avipoxvirus, capripoxvirus, leporipoxvirus, suipoxvirus, and the sub-family entomopoxvirinae. The hepadnavirus family includes the Hepatitis B virus. One unclassified virus which may be suitable source of antigens is the Hepatitis delta virus, Hepatitis E virus, and prions. Another virus which is a source of antigens is Nipah Virus. Still other viral sources may include avian infectious bursal disease virus and porcine respiratory and reproductive syndrome virus. The alphavirus family includes equine arteritis virus and various Encephalitis viruses.

The present invention may also encompass immunogens which are useful to immunize a human or non-human animal against other pathogens including bacteria, fungi, parasitic microorganisms or multicellular parasites which infect human and non-human vertebrates, or from a cancer cell or tumor cell. Examples of bacterial pathogens include pathogenic gram-positive cocci include pneumococci; staphylococci (and the toxins produced thereby, *e.g.*, enterotoxin B); and streptococci. Pathogenic gram-negative cocci include meningococcus; gonococcus. Pathogenic enteric gram-negative bacilli include enterobacteriaceae; pseudomonas, acinetobacteria and eikenella; melioidosis; salmonella; shigella; haemophilus; moraxella; *H. ducreyi* (which causes chancroid); brucella species (brucellosis); *Francisella tularensis* (which causes tularemia); *Yersinia pestis* (plague) and other yersinia (pasteurella); streptobacillus moniliformis and spirillum; Gram-positive bacilli include listeria monocytogenes; erysipelothrix rhusiopathiae; *Corynebacterium diphtheria* (diphtheria); cholera; *B. anthracis* (anthrax); donovanosis (granuloma inguinale); and bartonellosis. Diseases caused by pathogenic anaerobic bacteria include tetanus; botulism (*Clostridium botulinum* and its toxin); *Clostridium perfringens* and its epsilon toxin; other clostridia; tuberculosis; leprosy; and other mycobacteria. Pathogenic spirochetal diseases include syphilis; treponematoses: yaws, pinta and endemic syphilis; and leptospirosis. Other infections caused by higher pathogen bacteria and pathogenic fungi include glanders

(*Burkholderia mallei*); actinomycosis; nocardiosis; cryptococcosis, blastomycosis, histoplasmosis and coccidioidomycosis; candidiasis, aspergillosis, and mucormycosis; sporotrichosis; paracoccidioidomycosis, petriellidiosis, torulopsosis, mycetoma and chromomycosis; and dermatophytosis. Rickettsial infections include Typhus fever,

5 Rocky Mountain spotted fever, Q fever (*Coxiella burnetti*), and Rickettsialpox. Examples of mycoplasma and chlamydial infections include: mycoplasma pneumoniae; lymphogranuloma venereum; psittacosis; and perinatal chlamydial infections.

Pathogenic eukaryotes encompass pathogenic protozoans and helminths and infections produced thereby include: amebiasis; malaria; leishmaniasis; trypanosomiasis;

10 toxoplasmosis; *Pneumocystis carinii*; *Trichans*; *Toxoplasma gondii*; babesiosis; giardiasis; trichinosis; filariasis; schistosomiasis; nematodes; trematodes or flukes; and cestode (tapeworm) infections.

Many of these organisms and/or the toxins produced thereby have been identified by the Centers for Disease Control [(CDC), Department of Health and Human

15 Services, USA], as agents which have potential for use in biological attacks. For example, some of these biological agents, include, *Bacillus anthracis* (anthrax), *Clostridium*

botulinum and its toxin (botulism), *Yersinia pestis* (plague), variola major (smallpox), *Francisella tularensis* (tularemia), and viral hemorrhagic fevers [filoviruses (*e.g.*, Ebola, Marburg), and arenaviruses (*e.g.*, Lassa, Machupo)], all of which are currently classified

20 as Category A agents; *Coxiella burnetti* (Q fever); *Brucella* species (brucellosis), *Burkholderia mallei* (glanders), *Burkholderia pseudomallei* (meloidosis), *Ricinus communis* and its toxin (ricin toxin), *Clostridium perfringens* and its toxin (epsilon toxin), *Staphylococcus* species and their toxins (enterotoxin B), *Chlamydia psittaci* (psittacosis), water safety threats (*e.g.*, *Vibrio cholerae*, *Cryptosporidium parvum*), Typhus fever

25 (*Rickettsia powazekii*), and viral encephalitis (alphaviruses, *e.g.*, Venezuelan equine encephalitis; eastern equine encephalitis; western equine encephalitis); all of which are currently classified as Category B agents; and Nipah virus and hantaviruses, which are currently classified as Category C agents. In addition, other organisms, which are so

30 classified or differently classified, may be identified and/or used for such a purpose in the future. It will be readily understood that the viral vectors and other constructs described herein are useful to deliver antigens from these organisms, viruses, their toxins or other

by-products, which will prevent and/or treat infection or other adverse reactions with these biological agents.

Administration of the vectors of the invention to deliver immunogens against the variable region of the T cells elicit an immune response including CTLs to eliminate those T cells. In rheumatoid arthritis (RA), several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-3, V-14, V-17 and V-17. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in RA. In multiple sclerosis (MS), several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-7 and V-10. Thus, delivery of a nucleic acid sequence that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in MS. In scleroderma, several specific variable regions of TCRs which are involved in the disease have been characterized. These TCRs include V-6, V-8, V-14 and V-16, V-3C, V-7, V-14, V-15, V-16, V-28 and V-12. Thus, delivery of a nucleic acid molecule that encodes at least one of these polypeptides will elicit an immune response that will target T cells involved in scleroderma.

Thus, a rAAV-derived recombinant viral vector of the invention provides an efficient gene transfer vehicle which can deliver a selected transgene to a selected host cell *in vivo* or *ex vivo* even where the organism has neutralizing antibodies to one or more AAV sources. In one embodiment, the rAAV and the cells are mixed *ex vivo*; the infected cells are cultured using conventional methodologies; and the transduced cells are re-infused into the patient.

These compositions are particularly well suited to gene delivery for therapeutic purposes and for immunization, including inducing protective immunity. AAV of the invention and compositions containing same can also be used in immunization regimens such as those described in co-owned US Patent Application No. 60/565,936, filed April 28, 2004 for "Sequential Adenovirus and AAV-Mediated Delivery of Immunogenic Molecules".

Further, the compositions of the invention may also be used for production of a desired gene product *in vitro*. For *in vitro* production, a desired product (*e.g.*, a protein) may be obtained from a desired culture following transfection of host

cells with a rAAV containing the molecule encoding the desired product and culturing the cell culture under conditions which permit expression. The expressed product may then be purified and isolated, as desired. Suitable techniques for transfection, cell culturing, purification, and isolation are known to those of skill in the art.

5

The following examples illustrate several aspects and embodiments of the invention.

EXAMPLE 1

10 According to the method of the invention, AAV sequences have been identified as having singletons, when placed in an alignment with a library of sequences containing representatives of each of clades A, B, C, D, E, and F (represented by AAV9). The following table illustrates the capsid sequences and the singleton to be altered to a conserved sequence. For certain mutations, the singleton is followed by an * and then the amino acid residue which replaces it. For other mutations, the singleton is followed by its amino acid position and the residue which replaced it.

15 The amino acid numbering is based upon the published sequences for each of these AAV capsids. See, *e.g.*, G. Gao, *et al.*, J. Virol., 78(12):6381-6388 (June 2004) [all sequences therein deposited with GenBank] and International Patent Application No. PCT/US 2004/028817, filed September 30, 2004, which are incorporated by reference.

CAPSID SEQUENCE	SINGLETON MUTATION	CLADE REPRESENTED
Cy.5	G13D, D403N, R51K, N158K, P161Q	D
Rh.37	E634K, T207M	D
Rh.54	V404M	D
Rh. 67	F119L	D (AAV7)
Rh.40	VKEVTQN*DEGTKTIAN	E
AAV6	F129L, K531E	A
Hu.48	G277S, E322K, S552N	A

Hu.44	G609D, P446L, E137K	A
-------	---------------------	---

EXAMPLE 2

Five clones were selected to test the singleton method of the invention. The table below provides the phenotype description of the 5 clones. The number of predicted singletons is given with the clade and serotype classification.

Packaging phenotype is considered insufficient when their titer is lower than 1×10^{11} GC, low when lower than 1×10^{12} GC, good when lower 1×10^{13} , excellent when higher.

Gene transfer phenotypes were established by CB.A1AT gene expression and indicated as follows; “+++” better than lead candidate for target tissue, “++”, “+” and “-” respectively better than 50%, between 10-50% or lower than 10% of A1AT serum levels of lead candidates (muscle: AAV1, Liver:AAV8, Lung:AAV9). “n/a” indicated that vector could not be produced at sufficient levels for in vivo gene transfer studies.

Cloning of the singleton corrections went as follows. From the original packaging plasmid, site directed mutagenesis was performed. Subsequent to that, vector backbone integrity was assayed by a PstI digest and correction of the singleton was confirmed by sequencing. EGFP-expressing vector was then produced in triplicates on 12-well format side by side with the parental singleton-containing vector, AAV2 and AAV2/8 positive control and a production without presence of packaging plasmid as a negative control. Equal volume of harvested lysate after a 3x freeze was incubated on 293 cells. eGFP expressing was monitored by flow cytometry 72h post transduction.

Site-directed mutagenesis of the singleton residues in clones rh.37, rh.2, ch.5, rh.13, and rh.8 was performed. These particular sequences were selected to represent a variety of phenotypes that were previously documented.

Clone	Packaging	Gene Transfer Phenotype			# Singleton	Clade (serotype)
	Phenotype	Lung	Liver	Muscle		
rh.37	insufficient	n/a	n/a	n/a	2	D (AAV7)
rh.2	Low	++	+	+++	1	E (AAV8)
ch.5	Good	-	-	-	1	Ch.5
rh.13	Excellent	+	+	+	1	D (AAV7)
rh.8	Good	-	+	++	1	Rh.8

An increase in vector expression was noticed for 4 out of 5 clones. The increase was most dramatic for rh.37 and rh.2, vectors that were previously shown to have a low packaging yield. For these vectors productive particles were produced at levels sufficient
5 for detection. Vectors rh.8 and rh.13 showed an increase in transduction. Clone ch.5 showed a dramatic loss in either production or transduction efficiency due to the singleton mutation.

In order to distinguish the effects of the singleton mutation on transduction versus packaging and assembly, small-scale vector preparations were made and titered for Dnase
10 resistant particles by quantitative PCR. For rh.37, a two-log increase in vector production was observed. rh.8 showed a moderate 5 fold increase in titer whereas rh.13 performed equally. All titers of singleton corrected clones were within acceptable range in

comparison to AAV2 and AAV2/8 production and when extrapolated to large-scale preparations. rh.2 was not assayed for titration.

Subsequently the effect of the singleton change was monitored in vitro in a transduction setting with equal particle number per cell. A titration on 293 cells was performed for rh.8 and rh.13. Moderate increases in transduction efficiency were described at all MOIs.

From this initial subset of 5 clones, 3 productively transduced cells. Two clones were unable to yield any eGFP expression in this setting. This is most likely due to a defect in packaging of the vector that could not be predicted by the singleton approach.

EXAMPLE 3

Using the methods described above, the capsid proteins of rh.64 and hu.29 were altered and then used to construct viral vectors with the altered capsids using pseudotyping as described in Example 2.

Rh.64 was produced over a 100-fold more efficiently after a R697W change. A subsequent V686E mutation yields a 2-fold increase in packaging capability.

Hu.29 was also rescued from a deficient packaging capability by changing G396E. A greater than 1000-fold increase in production was observed.

All publications cited in this specification are incorporated herein by reference. While the invention has been described with reference to particularly preferred embodiments, it will be appreciated that modifications can be made without departing from the spirit of the invention.

CLAIMS:

1. A method of increasing the packaging yield, transduction efficiency, and/or gene transfer efficiency of a parental AAV, said method comprising the steps of:
 - (a) comparing a parental AAV capsid sequence with a library of functional AAV capsid sequences in an alignment;
 - (b) identifying at least one singleton in the parental AAV capsid, said singleton being a variable amino acid in a position in the parental AAV capsid in which the amino acid in the corresponding position of the aligned functional AAV capsid sequences is completely conserved;
 - (c) altering the singleton to the amino acid located in the corresponding position in the functional AAV capsid sequences.
2. The method according to claim 1, wherein the library of aligned functional AAV capsid sequences comprises at least 4 AAV sequences which are least 85% identical over the full-length of the aligned vp1 sequences and contains AAV sequences from at least two different AAV clades.
3. The method according to claim 1, wherein the library comprises at least two AAV sequences from each clade.
4. The method according to claim 1, wherein the library of aligned AAV capsid sequences comprises AAV sequences from at least three different clades.
5. The method according to claim 1, wherein the library of aligned AAV capsid sequences comprises at least three to 100 AAV sequences.
6. The method according to claim 1, wherein the library of aligned AAV capsid sequences comprises at least six to 50 AAV sequences.
7. The method according to claim 6, wherein the library comprises at least 12 AAV sequences.

8. The method according to claim 1, wherein the parental AAV contains multiple singletons.
9. The method according to claim 8, wherein the parental AAV contains 2 to 4 singletons.
10. The method according to claim 8, wherein step (c) is repeated at different singletons.
11. The method according to claim 1, wherein the target AAV is altered by site directed mutagenesis of the codon for the variable amino acid.
12. The method according to claim 11, wherein the site directed mutagenesis is performed on the target AAV carried on a plasmid backbone.
13. The method according to claim 1, further comprising the step of packaging the altered AAV into an AAV particle.
14. The method according to claim 13, wherein the AAV particle has increased packaging efficiency as compared to the target AAV.
15. The method according to claim 13, wherein in the AAV particle has increased transduction efficiency as compared to the target AAV.
16. The method according to claim 13, wherein the AAV particle has increased gene transfer efficiency as compared to the target AAV.
17. A viral vector comprising an AAV capsid modified according to the method of claim 1.

18. A viral vector comprising an AAV capsid selected from the group consisting of rh20, rh32/33, rh39, rh46, rh73 and rh74.

19. The viral vector according to claim 18, wherein the AAV capsid is rh20, which has an amino acid sequence of SEQ ID NO:1.

20. The viral vector according to claim 18, wherein the AAV capsid is rh32/33, which has an amino acid sequence of SEQ ID NO:2.

21. The viral vector according to claim 18, wherein the AAV capsid is rh39, which has an amino acid sequence of SEQ ID NO:3.

22. The viral vector according to claim 18, wherein the AAV capsid is rh46, which has an amino acid sequence of SEQ ID NO:4.

23. The viral vector according to claim 18, wherein the AAV capsid is rh73, which has an amino acid sequence of SEQ ID NO:5.

24. The viral vector according to claim 18, wherein the AAV capsid is rh74, which has an amino acid sequence of SEQ ID NO:6.

25. The viral vector according to claim 18, wherein said viral vector comprises a transgene encoding a gene product under the control of regulatory sequences which direct expression of the product in a host cell.

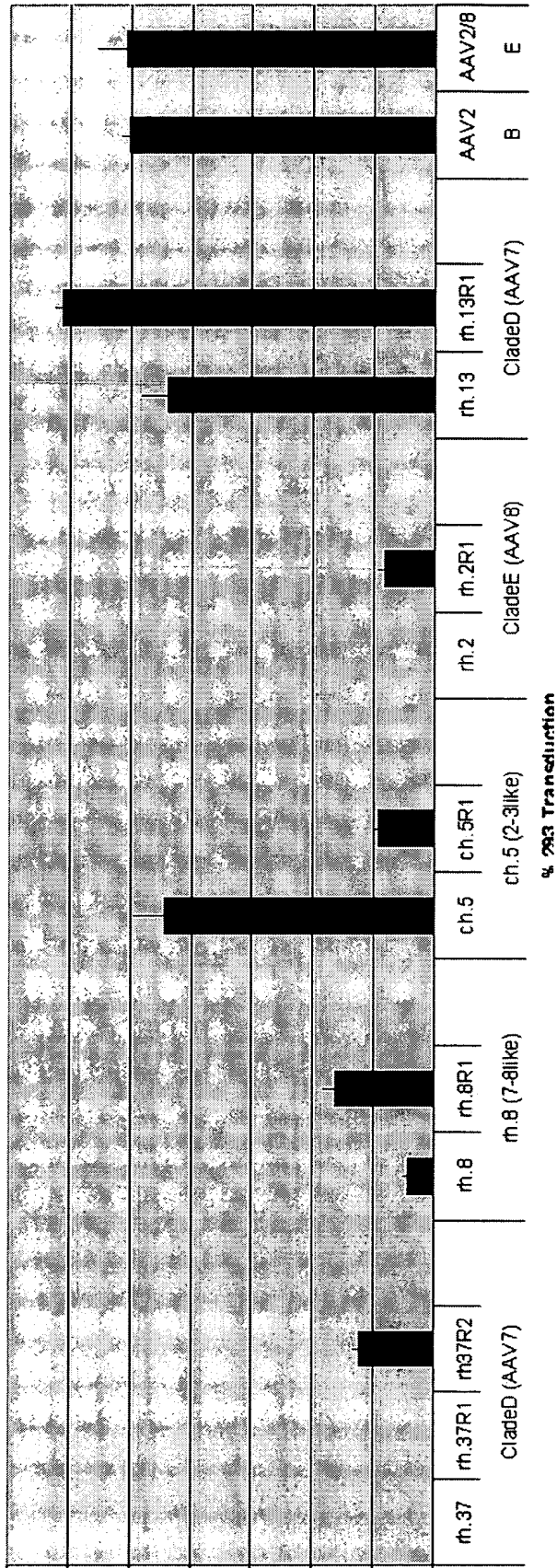
26. Use of a viral vector according to any of claims 17 to 25 in preparing a medicament for delivering a gene product to a host cell.

27. A method of delivering a gene product to a subject, said method comprising delivering the viral vector according to any of claims 17 to 25 to the subject.

ABSTRACT OF THE DISCLOSURE

5 A method of correcting singletons in a selected AAV sequence in order to increasing the packaging yield, transduction efficiency, and/or gene transfer efficiency of the selected AAV is provided. This method involves altering one or more singletons in the parental AAV capsid to conform the singleton to the amino acid in the corresponding position(s) of the aligned functional AAV capsid sequences.

Fig. 1 : Singleton *In Vitro* lysate Transduction



FIGs. 2A- 2C

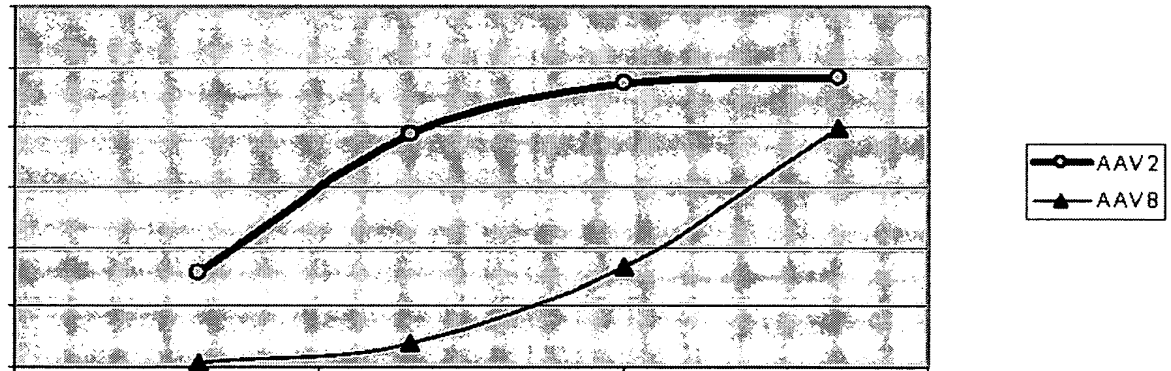
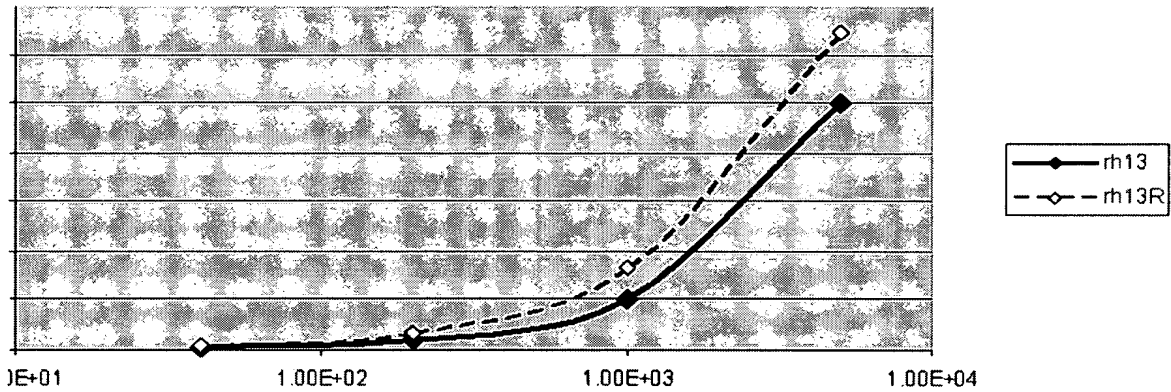
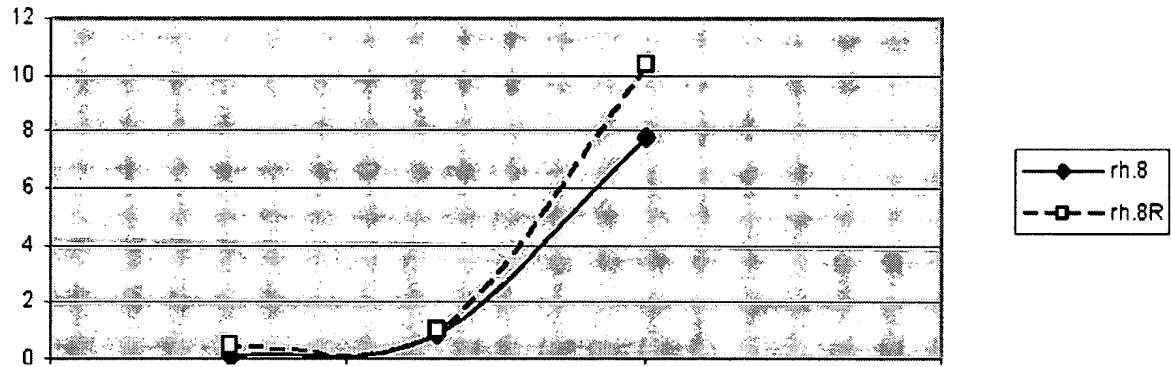
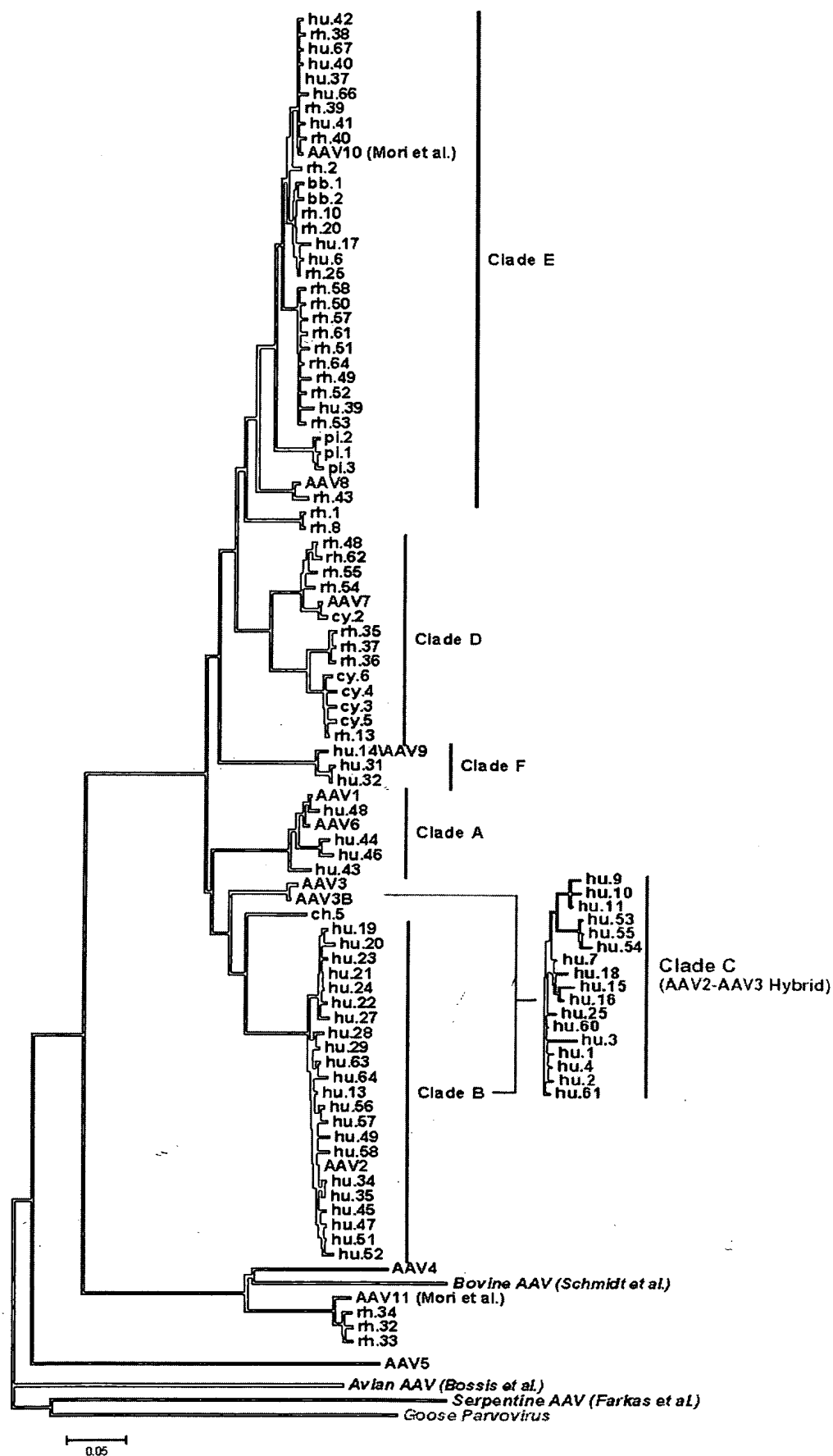


FIG. 3



[illegible][illegible]

FIG. 4A

	210	220	230	240	250	260	270	280	290	300
AAV2	g	g	c	a	c	a	a	a	a	a
cy.5	c	c	c	c	c	c	c	c	c	c
rh.10	a	a	a	a	a	a	a	a	a	a
rh.13	a	a	a	a	a	a	a	a	a	a
AAV1	a	a	a	a	a	a	a	a	a	a
AAV3	a	a	a	a	a	a	a	a	a	a
AAV6	a	a	a	a	a	a	a	a	a	a
AAV7	a	a	a	a	a	a	a	a	a	a
AAV8	a	a	a	a	a	a	a	a	a	a
hu.13	a	a	a	a	a	a	a	a	a	a
hu.26	a	a	a	a	a	a	a	a	a	a
hu.37	a	a	a	a	a	a	a	a	a	a
hu.53	a	a	a	a	a	a	a	a	a	a
rh.39	a	a	a	a	a	a	a	a	a	a
rh.43	a	a	a	a	a	a	a	a	a	a
rh.46	a	a	a	a	a	a	a	a	a	a

	310	320	330	340	350	360	370	380	390	400
AAV2	c	a	a	a	a	a	a	a	a	a
cy.5	c	c	c	c	c	c	c	c	c	c
rh.10	a	a	a	a	a	a	a	a	a	a
rh.13	a	a	a	a	a	a	a	a	a	a
AAV1	a	a	a	a	a	a	a	a	a	a
AAV3	a	a	a	a	a	a	a	a	a	a
AAV6	a	a	a	a	a	a	a	a	a	a
AAV7	a	a	a	a	a	a	a	a	a	a
AAV8	a	a	a	a	a	a	a	a	a	a
hu.13	a	a	a	a	a	a	a	a	a	a
hu.26	a	a	a	a	a	a	a	a	a	a
hu.37	a	a	a	a	a	a	a	a	a	a
hu.53	a	a	a	a	a	a	a	a	a	a
rh.39	a	a	a	a	a	a	a	a	a	a
rh.43	a	a	a	a	a	a	a	a	a	a
rh.46	a	a	a	a	a	a	a	a	a	a

FIG. 4B

AAV2
 cy.5
 rh.10
 rh.13
 AAV1
 AAV3
 AAV6
 AAV7
 AAV8
 hu.13
 hu.26
 hu.37
 hu.53
 rh.39
 rh.43
 rh.46

[illegible]

FIG. 4C

	610	620	630	640	650	660	670	680	690	700
AAV2	actaat	acgctac	agcag	tggtgg	cgccgac	gagtg	ggtaatt	cctcgg	gaaattg	gcgattcca
cy.5	t..gg..a	g	g	a	a	g	g	c	c	
rh.10	t..gg..a	g	t	a	a	g	g	a	a	
rh.13	t..gg..a	g	g	a	a	g	g	c	c	
AAV1	c...c..a	t	g	a	a	g	g	a	a	c
AAV3	t.....a	t	g	t	t	t	a	a	a	c
AAV6	c...c..a	t	g	a	a	g	g	a	a	
AAV7	t..gg..ag	g	g	a	a	t	g	a	a	
AAV8	c.....a	g	g	a	a	g	g	a	a	
hu.13	t	t	t	t	t	t	t	t	t	
hu.26	t	t	t	t	t	t	t	t	t	
hu.37	t..gg..a	g	t	a	a	g	g	a	a	c
hu.53	t..c...a	t	t	t	t	t	a	a	a	c
rh.39	t..gg..a	g	t	a	a	g	g	a	a	
rh.43	c.....a	g	g	a	a	g	g	a	a	
rh.46	t..gg..a	g	g	a	a	t	g	g	g	

	710	720	730	740	750	760	770	780	790	800
AAV2	catggat	ggcgac	agagtc	atcacc	accagc	accgga	acctgg	ggccctg	cccaccta	caaacatttccagccaatca
cy.5	c	c	c	c	c	c	c	g	a	a
rh.10	c	c	c	c	c	c	c	g	a	a
rh.13	c	c	c	c	c	c	c	g	a	a
AAV1	c	c	c	c	t	t	t	g	c	tgct
AAV3	a.....c	c	a	a	t	t	t	g	c	tgct
AAV6	c	c	t	a	t	t	t	g	c	tgct
AAV7	c	t	t	t	t	t	t	g	c	tg
AAV8	c	t	t	t	t	t	t	g	c	a
hu.13	c	t	t	t	t	t	t	g	c	a
hu.26	c	c	c	c	t	t	t	g	c	a
hu.37	c	c	c	c	t	t	t	g	a	atgga
hu.53	a.....c	c	c	c	t	t	t	g	c	tgga
rh.39	c	c	c	c	t	t	t	g	a	atgga
rh.43	c	c	c	c	t	t	t	g	a	atgga
rh.46	c	c	c	c	t	t	t	g	a	atgga

FIG. 4D

[illegible][illegible]

FIG. 4E

[illegible][illegible]

FIG. 4G

	1410	1420	1430	1440	1450	1460	1470	1480	1490	1500
AAV2										
cy.5	tcaggccggagcgagtgacattcgggaccagcttaggaactggcttcctggaccctgttaccgcccagcagcagtgatcaagacatctcgcggataacaac									
rh.10	t..gc.c.acac...ggcc.g.a.a.a.....g.c.....t.g.....a.c.g.....a.acata.acagc.....									
rh.13	gc.t.a.a.....gtc.ct...g.c.aa.....a.c.g.....c.....g.....a.c.c.c.c.c.....ctgt..c.a.t....									
AAV1	t..gc.c.acac...ggcc.g.a.a.a.....g.c.....t.g.....a.c.g.....a.acata.acagc.....									
AAV3	c.gt.ggtctc.agc..g...gtct.tt...c.c.aa.....g.....a.....g.....a.c.t.....tg.aac..c.....									
AAV6	c.....t..gc.tcagtct..gtctttg...g.c.a.t.....g.....c.....g.a.a.c.t.....tg.aac..c.....									
AAV7	c.g..ggtctc.agc..g...gtct.tt...c.c.aa.....a.....g.....c.t.t...t.a.....aaaa.a.c.....									
AAV8	c....g...gc.ttcaact..ggcc.a.ag.c.a..t..t.a.....t.c.t...g.a.aa...c.c.a..gctg.atc.a.....									
hu.13	c..a.gt..gc.t.a.aca..ggcca.t..g.a.a.....g.a.....a.a.c.c.c.....a.c.g.c.a.....									
hu.26	a.....a.....a.....a.....a.....a.....a.....a.....a.....a.....a.....a.....									
hu.37	a..t..gc.tgcaa...gtc.ct...g..a.....a.....t..c.....g.....c.t.c.....ctgt..c.a.....									
hu.53	c..a..t...c.c.caag...gtctctt.ag...aa.....g.....g.....t.....tt.g.....cagg.aaac..c.....									
rh.39	a..t..gc.tgcaa...gtc.ct...g..a.....a.....t..c.....g.....c.t.c.....ctgt..c.a.....									
rh.43	c..a.gt..gc.t.a.aca..ggcca.t..g.a.a.....g.a.....a.a.c.c.c.....a.c.g.c.a.....									
rh.46	gc.t..ca...gtc.ct...g.c.a.....g.....g.....c..a.a.....c.c.c.c.....ctgt..c.a.....									

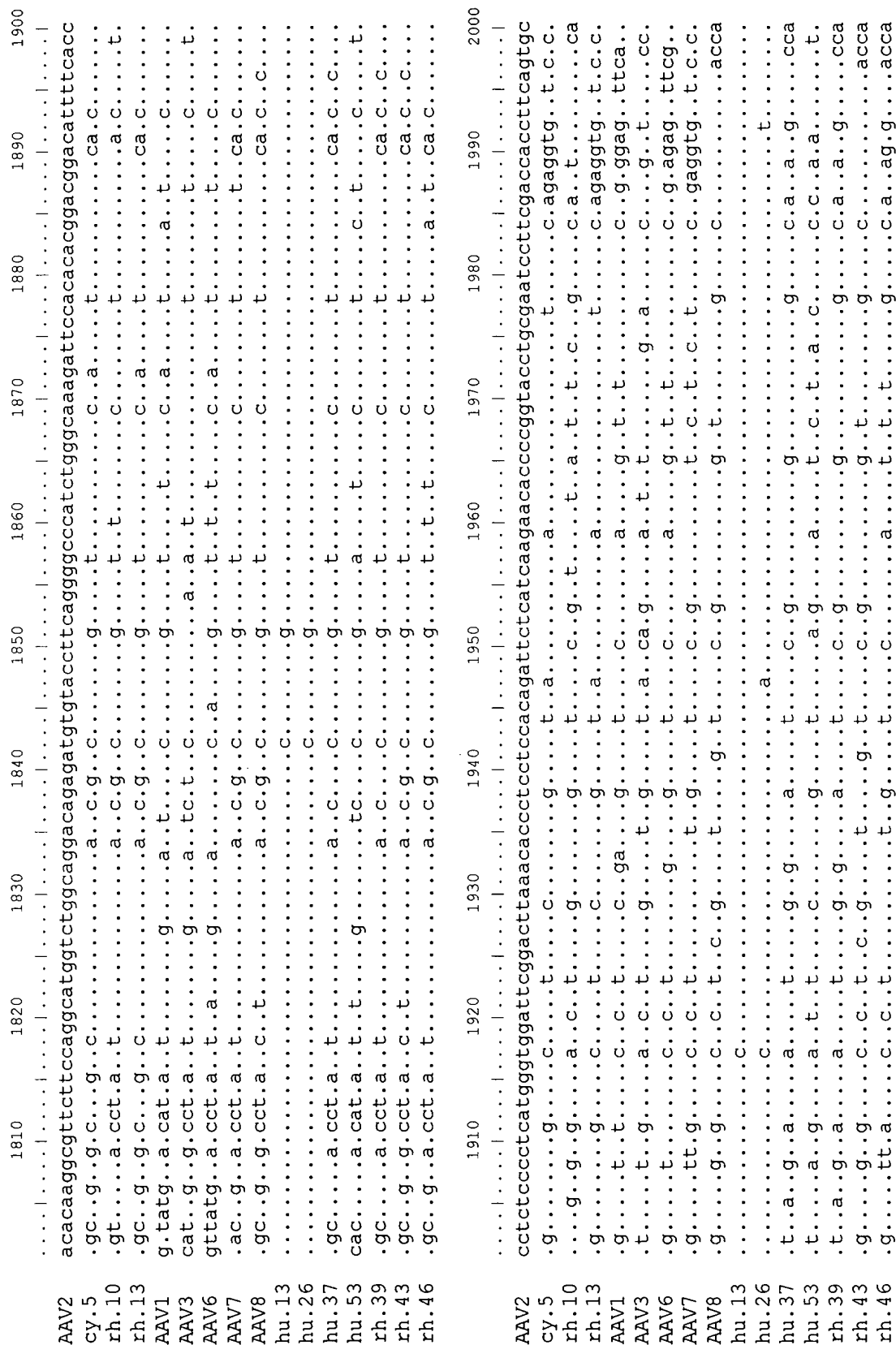
	1510	1520	1530	1540	1550	1560	1570	1580	1590	1600
AAV2										
cy.5	aacagtgaatactcgttgactggagctaccaagctaccacctcaatggcagagactctctgtgaatccggcccgcctatggcaagccacaagcagatg									
rh.10	a.c.ttg.c.....c.g.c.t.a.....t.g.....t..a.t..at.aacc..c.....gta.....c.c.a.....c.c.a.....									
rh.13	ca.c.ttg.c.....c.t.c.....t.t.g.....t.g.....a.....c.tgtc.t.....c.....c.....c.....									
AAV1	a.c.ttg.c.....c.g.c.t.a.....t.g.....t..a.t..at.aacc..c.....gta.....c.c.a.....c.c.a.....									
AAV3	ca.t.tta.c.....t..t.a.a.ta.....gc.t..a.ca.ca.c.c.t..a.t..t.....ctca.....a.....c.....									
AAV6	a.c.ttc.t.....a.cg.c.g..a.t.t.....c.c.....g.....a.a.a.t.....c.t.....c.....									
AAV7	ca.c.tta.c.....t..t.a.a.ta.....t...gc.t..a..a.a.c.c.t..a.t..t.....ctca.....a.....ca									
AAV8	ca.c.ttg.t.....t.c.....a.t.....g.c.....a..gt...t...c..gtc.....ct.....ct.....c.....									
hu.13	t..ca.c.ttg.c.....ct.gg.....a.....t.g.....a.a.t..at..ct...t..atc.t.....ca.....a.....c.....									
hu.26	t.....t.....t.....t.....t.....t.....t.....t.....t.....t.....t.....t.....t.....t.....									
hu.37	ca.c.ttg.t.....t.c.....a.t.....g.c.a.....t...a.....c.tgtc.....c.....c.....c.....									
hu.53	ca.c.ttc.c.....t.....t.....t.....t.....t.....t.....t.....t.....t.....t.....t.....t.....									
rh.39	ca.c.ttg.t.....t.c.....a.t.....g.c.a.....t...a.....c.tgtc.....c.....c.....c.....									
rh.43	t..ca.c.ttg.c.....ct.gg.....a.....t.g.....a.a.t..at..ct...t..atc.t.....ca.....a.....c.....									
rh.46	ca.c.ttg.c.....t.c.....t.t.g.c.....t.t.g.c.....t.t.g.c.....t.t.g.c.....t.t.g.c.....									

FIG. 4H

[illegible]

	1710	1720	1730	1740	1750	1760	1770	1780	1790	1800
AAV2	ggaatcaggacaaccaatcccg	tgctacggagcagtgatggt	tctgtatctaccaacctccag	agaggaagcaacagacagcag	tgactaccgcagtgatcaccg	cagatgtc	caac			
cy.5	...g...aa...c...c...a...	...a...ag...a...c...gtg...	...c...c...g...g...atcgtct...	...cgccgg...c...ccag...	...acagac...					
rh.10	a...t...aa...c...c...a...	...c...cgtg...gg...cgat...	...g...acagcaa...gcccctc...	...tat...gta...gg...cc...						
rh.13	...g...aa...c...c...c...	...a...ag...a...c...gtg...	...c...c...g...g...atcgtct...	...cgccgg...c...ccag...	...acagac...					
AAV1	...t...aag...c...t...c...t...	...c...c...aaga...t...ga...	...c...gg...agt...tt...	...ca...g...c...g...cc...t...	...g...g...gc...t...					
AAV3	a...g...tc...t...c...c...	...t...a...a...a...aa...	...gg...a...at...t...g...	...ctca...t...c...gctc...	...ca...g...t...g...ac...	...t...				
AAV6	...aag...c...t...c...c...	...c...c...aaga...t...ga...	...gg...agt...t...t...	...ca...g...c...g...cc...t...	...g...g...gc...t...					
AAV7	a...tc...tc...t...t...a...	...c...ag...a...c...gata...	...cagc...g...t...a...	...agcg...ct...t...ctgc...	...ccag...aca...t...					
AAV8	a...aa...c...c...t...c...t...	...t...a...a...c...atc...	...gg...agat...t...g...	...cagcaa...cggtc...	...tcaa...tt...g...ac...					
hu.13	...c...c...c...c...c...	...g...g...c...c...c...								
hu.26	...c...c...c...c...c...	...c...c...c...c...c...								
hu.37	a...t...aa...c...t...c...	...a...c...a...a...a...c...	...gtg...gg...ga...t...g...	...ca...ac...t...c...gggc...	...tat...gtg...g...a...					
hu.53	...c...c...c...c...c...	...a...c...a...a...g...a...	...at...tt...g...a...actca...	...t...ctgct...ag...t...a...	...ac...g...t...					
rh.39	a...t...aa...c...t...c...t...	...a...c...a...a...a...c...	...gtg...gg...gat...t...g...	...ca...ac...t...cggggc...	...tat...gtg...g...a...					
rh.43	a...aa...c...t...c...t...	...a...g...a...c...atc...	...gg...agat...t...g...	...cagcaa...cggtc...	...tcaa...tt...g...ac...					
rh.46	a...a...g...c...c...c...	...c...a...a...c...cgtg...	...gg...gat...t...a...	...cagcaa...ccgctc...	...tat...gtg...gg...cc...					

FIG. 41



	2210	2220
AAV2		
cy.5	acctgactcgtaaatctgtaa	
rh.10	C..C....C.....	
rh.13	C..C....C.....	
AAV1	t..C...CCC.....	----
AAV3	.t..C..a..a..ct..---	----
AAV6	C..C...CCC.....	----
AAV7	C..C....C.....	----
AAV8	C..C....C.....	----
hu.13		
hu.26		
hu.37	C..C....C.....	
hu.53		
rh.39	C..C....C.....	
rh.43	C..C....C.....	
rh.46	C..C....C.....	

FIG. 4L

	10	20	30	40	50	60	70	80	90	100
AAV2	MAADGYLPDWLEDTLSEGIQWVKLKPGPPPPKPAERHKDDSRGLVLPGYKYLGPENGLDKGEPVNEADAAALEHDKAYDRQLDSGDNPYLKYNHADA	EF								
cy.5	GN	E.D.	A.K.	ANQKQ.	G.	R.		K.EQ.		
rh.10	N	E.D.	A.K.	ANQKQ.	G.		A.	Q.KA.	R.	
rh.13	N	E.D.	A.K.	ANQKQ.	G.			K.EQ.		
AAV1	N	E.D.	A.K.	ANQKQ.	G.		A.	Q.KA.	R.	
AAV3	N	E.A.	V.Q.	ANQKQ.	Q.NR.	G.		Q.KA.		
AAV6	N	E.D.	A.K.	ANQKQ.	G.		A.	Q.KA.	R.	
AAV7	N	E.D.	A.K.	ANQKQ.	NG.		A.	Q.KA.	R.	
AAV8	N	E.A.	A.K.	ANQKQ.	G.		A.	Q.QA.	R.	
hu.13										
hu.26										
hu.37	N	E.D.	A.K.	ANQKQ.	G.		A.	Q.KA.	R.	
hu.53										
rh.39	N	E.D.	A.K.	ANQKQ.	G.		A.	Q.KA.	R.	
rh.43	N	E.D.	A.K.	ANQKQ.	G.		A.	Q.EA.	R.	
rh.46	N	E.D.	A.K.	ANQKQ.	G.		A.	Q.KA.	R.	

[illegible]

FIG. 5A

	210	220	230	240	250	260	270	280	290	300
AAV2	TNTMATGSGAPMADNNEGADGVGNSSGNWHCDSTWMDRVITTTSTRTWALPTYNHLYKQISSQS--GASNDNHYFGYSTPWGYFDENRFHCHFSPRDWQ									
cy.5	SG..A.G.....	A.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
rh.10	SG..A.G.....	S.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
rh.13	SG..A.G.....	A.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
AAV1	PT..S.G.....	A.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
AAV3	S...S.G.....	A.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
AAV6	PT..S.G.....	A.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
AAV7	SG.V.A.G.....	A.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
AAV8	P...A.G.....	S.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
hu.13	S.....									
hu.26	S.....									
hu.37	SG..A.G.....	S.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
hu.53	ST.....		Q.L.....	Q.L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
rh.39	SG..A.G.....	S.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
rh.43	P...A.G.....	S.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....
rh.46	SG..A.G.....	S.....	L.....	L.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....	NGTSG.ST..T.....

	310	320	330	340	350	360	370	380	390	400
AAV2	RLINNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPOQGYLTILNNGSQAVGRSSFCLEY									
cy.5	RK.R.....	T.V.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....
rh.10	RK.R.....	T.V.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....
rh.13	RK.R.....	T.V.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....
AAV1	K.S.....	RG.....								
AAV3	K.S.....	RG.....								
AAV6	K.R.....	T.V.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....
AAV7	K.R.....	T.V.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....	I.S.....
AAV8	S.....									
hu.13	S.....									
hu.26	S.....									
hu.37	S.....									
hu.53	S.....									
rh.39	S.....									
rh.43	S.....									
rh.46	S.....									

FIG. 5B

	10	20	30	40	50	60	70	80	90	100																												
rh.13	MAADG	YLPD	WLED	NLSE	GIRE	WDLK	PGAP	KPKAN	QKQD	DDGR	GLVLP	PGYK	YLGPF	NGLD	KGEP	NEAD	AAALE	HDKAY	DKQLE	QGD	NPYL	KYNH	ADAE	F														
rh.2																																						
rh.37																																						
rh.8																																						
hu.29																																						
rh.64																																						
	110	120	130	140	150	160	170	180	190	200																												
rh.13	QERLQ	EDTS	FGNL	GRAV	FQAK	KRVLE	PLGL	VEE	GAKT	APGK	KRPI	E----	SPDS	STGI	GKKG	QQPA	KKLNF	GOTD	SESV	PDQP	PLGE	PPA	AP	SGLG														
rh.2																																						
rh.37																																						
rh.8																																						
hu.29																																						
rh.64																																						
	210	220	230	240	250	260	270	280	290	300																												
rh.13	SGTMA	AGGAP	MADN	NEG	DVG	NAG	SNW	HCD	STWL	GDR	VIT	TSTR	TWAL	PTYN	NHLY	KQIS	-SQS	-GAT	NDN	HFFG	YSTP	WGYF	DFN	RHCH	FS	PR	DW	Q										
rh.2																																						
rh.37																																						
rh.8																																						
hu.29																																						
rh.64																																						
	310	320	330	340	350	360	370	380	390	400																												
rh.13	RLINN	WGF	RPR	KLR	FKL	FN	IQV	KEV	TND	GV	TIAN	NLT	STIQ	VF	SD	SEYQ	LPV	LGSA	HQGL	PPFP	AD	VF	MI	PQY	G	YLT	LN	NG	SQ	SV	GR	SR	SS	F	C	L	E	Y
rh.2																																						
rh.37																																						
rh.8																																						
hu.29																																						
rh.64																																						

FIG. 6A

	410	420	430	440	450	460	470	480	490	500
rh.13		FPSQMLRTGNFFEFSYTFEEV	PFHSSYAHQSQSLDRLMNP	LIQYLYLARTQSTT--GSTRELQFHQAGPNTMAEQSKNWLPGPCYRQQRRLSKNIDSNNN						
rh.2		D.....S.....GGTQG.QQ.L.S.....AN.SA.A.....V.TTISQ....								
rh.37		S.D.....--.....L.F....								
rh.8		Q.....T.G-T.G.QT.A.S.....SS.N.AR.V.....V.TTNQ....								
hu.29		T.....S.NTPSGTTQSR..S..ASDIRD.R.....V.TSAD....								
rh.64		S.....GGTAG.QQ.L.S.....SN.SA.AR.....V.TTISQ....								
	510	520	530	540	550	560	570	580	590	600
rh.13		SNFAWTGATKYHLNGRNSLTNP	GVAMATNKDDEQFFPINGV	IFGETGAAN-KTTLENVIMTSEEEIKTTNPVATEEYGVWSSNLQSSSTAGPQTQTVNS						
rh.2		D.V.....H.....ER..SS..M..KQ..GKDNVDYSS.ML.....Q...AD...QTNGA.IVG....								
rh.37		A.....IP.....K.....-.....S.I....								
rh.8		A.FK...D.M.....SH...D.R..SS..I..KQ..G.DGVDYSQ..I.D....A.....A.AI.N.AANTQA..GL.HN								
hu.29		.EYS.....D.V..P...SH...EK...QS...I..KQ.PEKTNVDI.K.MI.D...R.....Q.S.T...GNTQAA.AD..T								
rh.64		D.V.....R...SS.I.M..KQ..GKDNVDYS..ML.....Q...AD...QONTA.IVGA...								
	610	620	630	640	650	660	670	680	690	700
rh.13		QGALPGMVWQNRDVLQGP	IIWAKIPHTDGNFHPSP	LMGGFGLKHPPPPQILIKNT	PPANPPEVFTPAKEFASFTQYSTGQVSVEIEWELQKENS					
rh.2		E.....								
rh.37		E.....								
rh.8		VI.....D.LT.NQ..LN.....								
hu.29		V.....H.....STT.SA.....								
rh.64		D.TA.NQ..LN.....V.....R....								
	710	720	730							
rh.13		IQYTSNYAKSNVVEFAVNNEG	VYTEPRPIGTRYLTRNL*							
rh.2		Y.T.D...T..T.S.....*								
rh.37		PD.....*								
rh.8		Y.T.D...T...S.....*								
hu.29		N.V..D.T.DTN...S.....*								
rh.64		Y.T.D...T...S.....*								

FIG. 6B

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

UPN-R3895.ST25
SEQUENCE LISTING

<110> vandenbergh, Luk
Gao, Guangping
Wilson, James

<120> Method of Increasing the Function of an AAV Vector

<130> UPN-R3895USA

<160> 6

<170> PatentIn version 3.3

<210> 1

<211> 738

<212> PRT

<213> rhesus adeno-associated virus

<400> 1

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160

Gly Lys Thr Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
165 170 175

UPN-R3895.ST25

Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
 180 185 190
 Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
 195 200 205
 Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
 210 215 220
 Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
 225 230 235 240
 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
 245 250 255
 Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
 260 265 270
 Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
 275 280 285
 Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
 290 295 300
 Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Asn Phe Lys Leu Phe Asn
 305 310 315 320
 Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
 325 330 335
 Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
 340 345 350
 Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
 355 360 365
 Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
 370 375 380
 Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
 385 390 395 400
 Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Glu Phe Ser Tyr
 405 410 415
 Gln Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
 Page 2

420

425

430

Leu Asp Arg₄₃₅ Leu Met Asn Pro Leu₄₄₀ Ile Asp Gln Tyr Leu₄₄₅ Tyr Tyr Leu
 Ser Arg₄₅₀ Thr Gln Ser Thr Gly₄₅₅ Gly Thr Ala Gly Thr₄₆₀ Gln Gln Leu Leu
 Phe Ser Gln Ala Gly Pro₄₇₀ Asn Asn Met Ser Ala₄₇₅ Gln Ala Lys Asn Trp₄₈₀
 Leu Pro Gly Pro Cys₄₈₅ Tyr Arg Gln Gln Arg₄₉₀ Val Ser Thr Thr Leu₄₉₅ Ser
 Gln Asn Asn Asn₅₀₀ Ser Asn Phe Ala Trp₅₀₅ Thr Gly Ala Thr Lys₅₁₀ Tyr His
 Leu Asn Gly₅₁₅ Arg Asp Ser Leu Val₅₂₀ Asn Pro Gly Val Ala₅₂₅ Met Ala Thr
 His Lys₅₃₀ Asp Asp Glu Glu Arg₅₃₅ Phe Phe Pro Ser Ser₅₄₀ Gly Val Leu Met
 Phe Gly Lys Gln Gly Ala₅₅₀ Gly Lys Asp Asn Val₅₅₅ Asp Tyr Ser Ser Val₅₆₀
 Met Leu Thr Ser Glu₅₆₅ Glu Glu Ile Lys Thr₅₇₀ Thr Asn Pro Val Ala₅₇₅ Thr
 Glu Gln Tyr Gly₅₈₀ Val Val Ala Asp Asn₅₈₅ Leu Gln Gln Gln Asn₅₉₀ Ala Ala
 Pro Ile Val₅₉₅ Gly Ala Val Asn Ser₆₀₀ Gln Gly Ala Leu Pro₆₀₅ Gly Met Val
 Trp Gln Asn Arg Asp Val Tyr₆₁₅ Leu Gln Gly Pro Ile₆₂₀ Trp Ala Lys Ile
 Pro His Thr Asp Gly Asn₆₃₀ Phe His Pro Ser Pro₆₃₅ Leu Met Gly Gly Phe₆₄₀
 Gly Leu Lys His Pro₆₄₅ Pro Pro Gln Ile Leu₆₅₀ Ile Lys Asn Thr Pro₆₅₅ Val
 Pro Ala Asp Pro₆₆₀ Pro Thr Thr Phe Ser₆₆₅ Gln Ala Lys Leu Ala₆₇₀ Ser Phe

UPN-R3895.ST25

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
 675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
 690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu
 705 710 715 720

Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
 725 730 735

Asn Leu

<210> 2
 <211> 733
 <212> PRT
 <213> Rhesus Adeno-Associated Virus clone 32/33
 <400> 2

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
 1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
 20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
 35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
 50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
 65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
 85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
 100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
 115 120 125

Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140

UPN-R3895.ST25

Pro Leu Glu Ser Pro Gln Glu Pro Asp Ser Ser Ser Gly Ile Gly Lys
 145 150 155 160
 Lys Gly Lys Gln Pro Ala Lys Lys Arg Leu Asn Phe Glu Glu Asp Thr
 165 170 175
 Gly Ala Gly Asp Gly Pro Pro Glu Gly Ser Asp Thr Ser Ala Met Ser
 180 185 190
 Ser Asp Ile Glu Met Arg Ala Ala Pro Gly Gly Asn Ala Val Asp Ala
 195 200 205
 Gly Gln Gly Ser Asp Gly Val Gly Asn Ala Ser Gly Asp Trp His Cys
 210 215 220
 Asp Ser Thr Trp Ser Glu Gly Lys Val Thr Thr Thr Ser Thr Arg Thr
 225 230 235 240
 Trp Val Leu Pro Thr Tyr Asn Asn His Leu Tyr Leu Arg Leu Gly Thr
 245 250 255
 Thr Ser Asn Ser Asn Thr Tyr Asn Gly Phe Ser Thr Pro Trp Gly Tyr
 260 265 270
 Phe Asp Phe Asn Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln
 275 280 285
 Arg Leu Ile Asn Asn Asn Trp Gly Leu Arg Pro Lys Ala Met Arg Val
 290 295 300
 Lys Ile Phe Asn Ile Gln Val Lys Glu Val Thr Thr Ser Asn Gly Glu
 305 310 315 320
 Thr Thr Val Ala Asn Asn Leu Thr Ser Thr Val Gln Ile Phe Ala Asp
 325 330 335
 Ser Ser Tyr Glu Leu Pro Tyr Val Met Asp Ala Gly Gln Glu Gly Ser
 340 345 350
 Leu Pro Pro Phe Pro Asn Asp Val Phe Met Val Pro Gln Tyr Gly Tyr
 355 360 365
 Cys Gly Ile Val Thr Gly Glu Asn Gln Asn Gln Thr Asp Arg Asn Ala
 370 375 380
 Phe Tyr Cys Leu Glu Tyr Phe Pro Ser Gln Met Leu Arg Thr Gly Asn
 385 390 395 400

UPN-R3895.ST25

Asn Phe Glu Met Ala Tyr Asn Phe Glu Lys Val Pro Phe His Ser Met
 405 410 415
 Tyr Ala His Ser Gln Ser Leu Asp Arg Leu Met Asn Pro Leu Leu Asp
 420 425 430
 Gln Tyr Leu Trp His Leu Gln Ser Thr Thr Ser Gly Glu Thr Leu Asn
 435 440 445
 Gln Gly Asn Ala Ala Thr Thr Phe Gly Lys Ile Arg Ser Gly Asp Phe
 450 455 460
 Ala Phe Tyr Arg Lys Asn Trp Leu Pro Gly Pro Cys Val Lys Gln Gln
 465 470 475 480
 Arg Phe Ser Lys Thr Ala Ser Gln Asn Tyr Lys Ile Pro Ala Ser Gly
 485 490 495
 Gly Asn Ala Leu Leu Lys Tyr Asp Thr His Tyr Thr Leu Asn Asn Arg
 500 505 510
 Trp Ser Asn Ile Ala Pro Gly Pro Pro Met Ala Thr Ala Gly Pro Ser
 515 520 525
 Asp Gly Asp Phe Ser Asn Ala Gln Leu Ile Phe Pro Gly Pro Ser Val
 530 535 540
 Thr Gly Asn Thr Thr Thr Ser Ala Asn Asn Leu Leu Phe Thr Ser Glu
 545 550 555 560
 Glu Glu Ile Ala Ala Thr Asn Pro Arg Asp Thr Asp Met Phe Gly Gln
 565 570 575
 Ile Ala Asp Asn Asn Gln Asn Ala Thr Thr Ala Pro Ile Thr Gly Asn
 580 585 590
 Val Thr Ala Met Gly Val Leu Pro Gly Met Val Trp Gln Asn Arg Asp
 595 600 605
 Ile Tyr Tyr Gln Gly Pro Ile Trp Ala Lys Ile Pro His Ala Asp Gly
 610 615 620
 His Phe His Pro Ser Pro Leu Ile Gly Gly Phe Gly Leu Lys His Pro
 625 630 635 640
 Pro Pro Gln Ile Phe Ile Lys Asn Thr Pro Val Pro Ala Asn Pro Ala
 645 650 655

UPN-R3895.ST25

Thr Thr Phe Thr Ala Ala Arg Val Asp Ser Phe Ile Thr Gln Tyr Ser
660 665 670

Thr Gly Gln Val Ala Val Gln Ile Glu Trp Glu Ile Glu Lys Glu Arg
675 680 685

Ser Lys Arg Trp Asn Pro Glu Val Gln Phe Thr Ser Asn Tyr Gly Asn
690 695 700

Gln Ser Ser Met Leu Trp Ala Pro Asp Thr Thr Gly Lys Tyr Thr Glu
705 710 715 720

Pro Arg Val Ile Gly Ser Arg Tyr Leu Thr Asn His Leu
725 730

<210> 3

<211> 738

<212> PRT

<213> Rhesus Adeno-associated virus clone 39 capsid

<400> 3

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

UPN-R3895.ST25

Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
145 150 155 160

Gly Lys Lys Gly Gln Gln Pro Ala Lys Lys Arg Leu Asn Phe Gly Gln
165 170 175

Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190

Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
210 215 220

Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240

Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255

Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
260 265 270

Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
275 280 285

Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
290 295 300

Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Ser Phe Lys Leu Phe Asn
305 310 315 320

Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
325 330 335

Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
340 345 350

Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
355 360 365

Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
370 375 380

Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
Page 8

385 390 400

Phe Pro Ser Gln Met₄₀₅ Leu Arg Thr Gly Asn₄₁₀ Asn Phe Glu Phe Ser₄₁₅ Tyr

Thr Phe Glu Asp₄₂₀ Val Pro Phe His Ser₄₂₅ Ser Tyr Ala His Ser₄₃₀ Gln Ser

Leu Asp Arg₄₃₅ Leu Met Asn Pro Leu₄₄₀ Ile Asp Gln Tyr Leu₄₄₅ Tyr Tyr Leu

Ser Arg₄₅₀ Thr Gln Ser Thr Gly₄₅₅ Gly Thr Gln Gly Thr₄₆₀ Gln Gln Leu Leu

Phe Ser Gln Ala Gly Pro₄₇₀ Ala Asn Met Ser Ala₄₇₅ Gln Ala Lys Asn Trp₄₈₀

Leu Pro Gly Pro Cys₄₈₅ Tyr Arg Gln Gln Arg₄₉₀ Val Ser Thr Thr Leu₄₉₅ Ser

Gln Asn Asn₅₀₀ Ser Asn Phe Ala Trp₅₀₅ Thr Gly Ala Thr Lys₅₁₀ Tyr His

Leu Asn Gly₅₁₅ Arg Asp Ser Leu Val₅₂₀ Asn Pro Gly Val Ala₅₂₅ Met Ala Thr

His Lys₅₃₀ Asp Asp Glu Glu Arg₅₃₅ Phe Phe Pro Ser Ser₅₄₀ Gly Val Leu Met

Phe Gly Lys Gln Gly Ala₅₅₀ Gly Arg Asp Asn Val₅₅₅ Asp Tyr Ser Ser Val₅₆₀

Met Leu Thr Ser Glu₅₆₅ Glu Glu Ile Lys Thr₅₇₀ Thr Asn Pro Val Ala₅₇₅ Thr

Glu Gln Tyr Gly₅₈₀ Val Val Ala Asp Asn₅₈₅ Leu Gln Gln Thr Asn₅₉₀ Thr Gly

Pro Ile Val₅₉₅ Gly Asn Val Asn Ser₆₀₀ Gln Gly Ala Leu Pro₆₀₅ Gly Met Val

Trp Gln₆₁₀ Asn Arg Asp Val Tyr₆₁₅ Leu Gln Gly Pro Ile₆₂₀ Trp Ala Lys Ile

Pro His Thr Asp Gly Asn₆₃₀ Phe His Pro Ser Pro₆₃₅ Leu Met Gly Gly Phe₆₄₀

UPN-R3895.ST25

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
645 650 655

Pro Ala Asp Pro Pro Thr Thr Phe Ser Gln Ala Lys Leu Ala Ser Phe
660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu
705 710 715 720

Gly Thr Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735

Asn Leu

<210> 4

<211> 737

<212> PRT

<213> Rhesus Adeno-associated virus clone 46 capsid protein

<400> 4

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
100 105 110

UPN-R3895.ST25

Asn Leu Gly Arg Ala Val Leu Gln Ala Lys Lys Arg Val Leu Glu Pro
 115 120 125
 Leu Gly Leu Val Glu Glu Ala Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140
 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
 145 150 155 160
 Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
 165 170 175
 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
 180 185 190
 Pro Ala Gly Pro Ser Gly Leu Gly Ser Gly Thr Met Ala Ala Gly Gly
 195 200 205
 Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn
 210 215 220
 Ala Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
 225 230 235 240
 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
 245 250 255
 Leu Tyr Lys Gln Ile Ser Ser Gln Ser Ala Gly Ser Thr Asn Asp Asn
 260 265 270
 Val Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg
 275 280 285
 Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn
 290 295 300
 Asn Trp Gly Phe Arg Pro Lys Lys Leu Asn Phe Lys Leu Phe Asn Ile
 305 310 315 320
 Gln Val Lys Glu Val Thr Thr Asn Asp Gly Val Thr Thr Ile Ala Asn
 325 330 335
 Asn Leu Thr Ser Thr Val Gln Val Phe Ser Asp Ser Glu Tyr Gln Leu
 340 345 350
 Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro
 355 360 365

UPN-R3895.ST25

Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn
 370 375 380
 Gly Ser Gln Ser Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe
 385 390 395 400
 Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Thr Phe Ser Tyr Thr
 405 410 415
 Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu
 420 425 430
 Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ala
 435 440 445
 Arg Thr Gln Ser Asn Ala Gly Gly Thr Ala Gly Asn Arg Glu Leu Gln
 450 455 460
 Phe Tyr Gln Gly Gly Pro Thr Thr Met Ala Glu Gln Ala Lys Asn Trp
 465 470 475 480
 Leu Pro Gly Pro Cys Phe Arg Gln Gln Arg Val Ser Lys Thr Leu Asp
 485 490 495
 Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
 500 505 510
 Leu Asn Gly Arg Asn Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
 515 520 525
 His Lys Asp Asp Glu Glu Arg Phe Phe Pro Ser Ser Gly Val Leu Ile
 530 535 540
 Phe Gly Lys Thr Gly Ala Ala Asn Lys Thr Thr Leu Glu Asn Val Leu
 545 550 555 560
 Met Thr Asn Glu Glu Glu Ile Arg Pro Thr Asn Pro Val Ala Thr Glu
 565 570 575
 Glu Tyr Gly Thr Val Ser Ser Asn Leu Gln Ala Ala Asn Thr Ala Ala
 580 585 590
 Gln Thr Gln Val Val Asn Asn Gln Gly Ala Leu Pro Gly Met Val Trp
 595 600 605
 Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro
 610 615 620

UPN-R3895.ST25

His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly
625 630 635 640

Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro
645 650 655

Ala Asn Pro Pro Glu Val Phe Thr Pro Ala Lys Phe Ala Ser Phe Ile
660 665 670

Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu Leu
675 680 685

Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser
690 695 700

Asn Phe Asp Lys Gln Thr Gly Val Asp Phe Ala Val Asp Ser Gln Gly
705 710 715 720

Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn
725 730 735

Leu

<210> 5
<211> 738
<212> PRT
<213> Rhesus adeno-associated virus clone 73 capsid
<400> 5

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Trp Asp Leu Lys Pro Gly Ala Pro Lys Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asp Gly Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Glu Pro
50 55 60

Val Asn Ala Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Asp
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Arg Tyr Asn His Ala
85 90 95

UPN-R3895.ST25

Asp Ala Glu Phe Gln Glu Arg Leu Gln Glu Asp Thr Ser Phe Gly Gly
 100 105 110
 Asn Leu Gly Arg Ala Val Phe Gln Ala Lys Lys Arg Val Leu Glu Pro
 115 120 125
 Leu Gly Leu Val Glu Glu Gly Ala Lys Thr Ala Pro Gly Lys Lys Arg
 130 135 140
 Pro Val Glu Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Thr Gly Ile
 145 150 155 160
 Gly Lys Lys Gly Gln Gln Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
 165 170 175
 Thr Gly Asp Ser Glu Ser Val Pro Asp Pro Gln Pro Leu Gly Glu Pro
 180 185 190
 Pro Ala Ala Pro Ser Ser Val Gly Ser Gly Thr Met Ala Ala Gly Gly
 195 200 205
 Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Ser
 210 215 220
 Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
 225 230 235 240
 Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
 245 250 255
 Leu Tyr Lys Gln Ile Ser Asn Gly Thr Ser Gly Gly Ser Thr Asn Asp
 260 265 270
 Asn Thr Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn
 275 280 285
 Arg Phe His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn
 290 295 300
 Asn Asn Trp Gly Phe Arg Pro Lys Arg Leu Ser Phe Lys Leu Phe Asn
 305 310 315 320
 Ile Gln Val Lys Glu Val Thr Gln Asn Glu Gly Thr Lys Thr Ile Ala
 325 330 335
 Asn Asn Leu Thr Ser Thr Ile Gln Val Phe Thr Asp Ser Glu Tyr Gln
 Page 14

340 345 350
 Leu Pro Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe
 355 360 365
 Pro Ala Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn
 370 375 380
 Asn Gly Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr
 385 390 395 400
 Phe Pro Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Ser Phe Ser Tyr
 405 410 415
 Thr Phe Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser
 420 425 430
 Leu Asp Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu
 435 440 445
 Ser Arg Thr Gln Ser Thr Gly Gly Thr Ala Gly Thr Gln Gln Leu Leu
 450 455 460
 Phe Ser Gln Ala Gly Pro Ser Asn Met Ser Ala Gln Ala Arg Asn Trp
 465 470 475 480
 Leu Pro Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Thr Thr Leu Ser
 485 490 495
 Gln Asn Asn Asn Ser Asn Phe Ala Trp Thr Gly Ala Thr Lys Tyr His
 500 505 510
 Leu Asn Gly Arg Asp Ser Leu Val Asn Pro Gly Val Ala Met Ala Thr
 515 520 525
 Asn Lys Asp Asp Glu Asp Arg Phe Phe Pro Ser Ser Gly Ile Leu Met
 530 535 540
 Phe Gly Lys Gln Gly Ala Gly Lys Asp Asn Val Asp Tyr Ser Asn Val
 545 550 555 560
 Met Leu Thr Ser Glu Glu Glu Ile Lys Thr Thr Asn Pro Val Ala Thr
 565 570 575
 Glu Gln Tyr Gly Val Val Ala Asp Asn Leu Gln Arg Gln Asn Thr Ala
 580 585 590

UPN-R3895.ST25

Pro Ile Val Gly Ala Val Asn Ser Gln Gly Ala Leu Pro Gly Met Val
595 600 605

Trp Gln Asn Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile
610 615 620

Pro His Thr Asp Gly Asn Phe His Pro Ser Pro Leu Met Gly Gly Phe
625 630 635 640

Gly Leu Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val
645 650 655

Pro Ala Asp Pro Pro Thr Ala Phe Asn Gln Ala Lys Leu Asn Ser Phe
660 665 670

Ile Thr Gln Tyr Ser Thr Gly Gln Val Ser Val Glu Ile Glu Trp Glu
675 680 685

Leu Gln Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr
690 695 700

Ser Asn Tyr Tyr Lys Ser Thr Asn Val Asp Phe Ala Val Asn Thr Glu
705 710 715 720

Gly Val Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg
725 730 735

Asn Leu

<210> 6

<211> 736

<212> PRT

<213> Rhesus adeno-associated virus clone 74 capsid

<400> 6

Met Ala Ala Asp Gly Tyr Leu Pro Asp Trp Leu Glu Asp Asn Leu Ser
1 5 10 15

Glu Gly Ile Arg Glu Trp Leu Asn Leu Lys Pro Gly Ala Pro Gln Pro
20 25 30

Lys Ala Asn Gln Gln Lys Gln Asp Asn Ala Arg Gly Leu Val Leu Pro
35 40 45

Gly Tyr Lys Tyr Leu Gly Pro Phe Asn Gly Leu Asp Lys Gly Asp Pro
50 55 60

UPN-R3895.ST25

Val Asn Glu Ala Asp Ala Ala Ala Leu Glu His Asp Lys Ala Tyr Gly
65 70 75 80

Gln Gln Leu Lys Ala Gly Asp Asn Pro Tyr Leu Lys Tyr Asn His Ala
85 90 95

Asp Ala Glu Phe Gln Glu Arg Leu Gln Lys Asp Thr Ser Phe Gly Gly
100 105 110

Asn Leu Gly Lys Ala Val Phe Gln Ala Lys Lys Arg Ile Leu Glu Pro
115 120 125

Leu Gly Leu Val Glu Thr Pro Ala Lys Thr Ala Pro Gly Lys Lys Arg
130 135 140

Pro Val Asp Pro Ser Pro Gln Arg Ser Pro Asp Ser Ser Ser Gly Ile
145 150 155 160

Gly Lys Lys Ser Pro His Pro Ala Arg Lys Arg Leu Asn Phe Gly Gln
165 170 175

Thr Gly Glu Ala Glu Ser Val Pro Asp Pro Gln Pro Ile Gly Glu Pro
180 185 190

Pro Ser Asp Pro Ser Ser Val Gly Ser Gly Thr Met Ala Ala Gly Gly
195 200 205

Gly Ala Pro Met Ala Asp Asn Asn Glu Gly Ala Asp Gly Val Gly Asn
210 215 220

Ser Ser Gly Asn Trp His Cys Asp Ser Thr Trp Leu Gly Asp Arg Val
225 230 235 240

Ile Thr Thr Ser Thr Arg Thr Trp Ala Leu Pro Thr Tyr Asn Asn His
245 250 255

Leu Tyr Lys Gln Ile Ser Ser Gln Ser Gly Ala Thr Asn Asp Asn His
260 265 270

Tyr Phe Gly Tyr Ser Thr Pro Trp Gly Tyr Phe Asp Phe Asn Arg Phe
275 280 285

His Cys His Phe Ser Pro Arg Asp Trp Gln Arg Leu Ile Asn Asn Asn
290 295 300

Trp Gly Phe Arg Pro Lys Arg Leu Ser Phe Lys Leu Phe Asn Ile Gln
305 310 315 320

UPN-R3895.ST25

Val Lys Glu Val Thr Gln Asn Asp Gly Thr Thr Thr Ile Ala Asn Asn
325 330 335

Leu Thr Ser Thr Val Gln Val Phe Thr Asp Ser Glu Tyr Gln Leu Pro
340 345 350

Tyr Val Leu Gly Ser Ala His Gln Gly Cys Leu Pro Pro Phe Pro Ala
355 360 365

Asp Val Phe Met Ile Pro Gln Tyr Gly Tyr Leu Thr Leu Asn Asn Gly
370 375 380

Ser Gln Ala Val Gly Arg Ser Ser Phe Tyr Cys Leu Glu Tyr Phe Pro
385 390 395 400

Ser Gln Met Leu Arg Thr Gly Asn Asn Phe Gln Phe Ser Tyr Thr Phe
405 410 415

Glu Asp Val Pro Phe His Ser Ser Tyr Ala His Ser Gln Ser Leu Asp
420 425 430

Arg Leu Met Asn Pro Leu Ile Asp Gln Tyr Leu Tyr Tyr Leu Ser Lys
435 440 445

Thr Gln Gly Thr Asn Ala Thr Val Gln Gly Ala Lys Leu Gln Phe Ser
450 455 460

Gln Ala Gly Pro Ser Asn Met Arg Asp Gln Ala Arg Asn Trp Leu Pro
465 470 475 480

Gly Pro Cys Tyr Arg Gln Gln Arg Val Ser Lys Thr Ala Asn Asp Asn
485 490 495

Asn Asn Ser Glu Tyr Ala Trp Thr Gly Ala Thr Lys Tyr His Leu Asn
500 505 510

Gly Arg Asp Ser Leu Val Asn Pro Gly Pro Ala Met Ala Ser His Lys
515 520 525

Asp Asp Glu Glu Lys Phe Phe Pro Met Asn Gly Thr Leu Val Phe Gly
530 535 540

Lys Asn Gly Ala Gly Asn Ser Asn Val Asp Ile Glu Asn Val Met Ile
545 550 555 560

Thr Asp Glu Glu Glu Ile Arg Thr Thr Asn Pro Val Ala Thr Glu Gln
565 570 575

UPN-R3895.ST25

Tyr Gly Val Val Ser Asp Asn Leu Gln Ser Ser Asn Thr Arg Pro Ile
 580 585 590
 Thr Gly Asp Val Asp Ser Gln Gly Val Leu Pro Gly Met Val Trp Gln
 595 600 605
 Asp Arg Asp Val Tyr Leu Gln Gly Pro Ile Trp Ala Lys Ile Pro His
 610 615 620
 Thr Asp Gly His Phe His Pro Ser Pro Leu Met Gly Gly Phe Gly Leu
 625 630 635 640
 Lys His Pro Pro Pro Gln Ile Leu Ile Lys Asn Thr Pro Val Pro Ala
 645 650 655
 Asn Pro Ala Thr Thr Phe Thr Pro Gly Lys Phe Ala Ser Phe Ile Thr
 660 665 670
 Gln Tyr Ser Thr Gly Gln Val Ser Val Gln Ile Glu Trp Glu Leu Gln
 675 680 685
 Lys Glu Asn Ser Lys Arg Trp Asn Pro Glu Ile Gln Tyr Thr Ser Asn
 690 695 700
 Tyr Asn Lys Ser Val Asn Val Asp Phe Thr Val Asp Thr Asn Gly Val
 705 710 715 720
 Tyr Ser Glu Pro Arg Pro Ile Gly Thr Arg Tyr Leu Thr Arg Asn Leu
 725 730 735

PATENT APPLICATION SERIAL NO _____

U.S. DEPARTMENT OF COMMERCE
PATENT AND TRADEMARK OFFICE
FEE RECORD SHEET

04/11/2005 CNGUYEN2 00000002 60669083

01 FC:1005

200.00 DP

PTO-1556
(5/87)

APPLICATION DATA SHEET

Applicant Information	
Applicant Authority Type::	Inventor
Primary Citizenship Country::	Belgium
Status::	Full Capacity
Given Name::	Luk
Middle Name::	
Family Name::	Vandenberghe
Name Suffix::	
City of Residence::	Philadelphia
State or Province of Residence::	PA
Country of Residence::	US
Street of Mailing Address::	2224 Catharine Street
City of Mailing Address::	Philadelphia
State or Province of Mailing Address::	PA
Country of Mailing Address::	US
Postal or Zip Code of Mailing Address::	19146

Applicant Information	
Applicant Authority Type::	Inventor
Primary Citizenship Country::	US
Status::	Full Capacity
Given Name::	Guangping
Middle Name::	
Family Name::	Gao
Name Suffix::	
City of Residence::	Rosemont
State or Province of Residence::	PA
Country of Residence::	US
Street of Mailing Address::	408 Yorkshire Road
City of Mailing Address::	Rosemont
State or Province of Mailing Address::	PA
Country of Mailing Address::	US
Postal or Zip Code of Mailing Address::	19010

Applicant Information	
Applicant Authority Type::	Inventor
Primary Citizenship Country::	US
Status::	Full Capacity
Given Name::	James
Middle Name::	M.
Family Name::	Wilson
Name Suffix::	
City of Residence::	Gladwyne
State or Province of Residence::	PA
Country of Residence::	US
Street of Mailing Address::	1350 N. Avignon Drive
City of Mailing Address::	Gladwyne
State or Province of Mailing Address::	PA
Country of Mailing Address::	US
Postal or Zip Code of Mailing Address::	19035

Correspondence Information	
Correspondence Customer Number::	00270
Name::	Howson and Howson
Street of Mailing Address	Spring House Corporate Center, Box 457
City of Mailing Address	Spring House
State or Province of Mailing Address	Pennsylvania
Country of Mailing Address	US
Postal or Zip Code of Mailing Address::	19477
Phone Number::	215-540-9200
Fax Number::	215-540-5818
E-Mail Address::	ckodroff@howsonandhowson.com

Application Information	
Application Number::	
Filing Date::	Herewith
Application Type::	Regular
Subject Matter::	Utility
Suggested Classification::	
Suggested Group Art Unit::	
CD-ROM or CD-R::	None
Number of CD disks::	
Number of Copies of CDs::	
Sequence Submission?::	Paper
Computer Readable Form (CRF)?::	
Number of Copies of CRF::	
Title::	METHOD OF INCREASING THE FUNCTION OF AN AAV VECTOR
Attorney Docket Number::	UPN-R3895
Request for Early Publication?	No
Request for Non-Publication?	No
Suggested Drawing Figure::	
Total Drawing Sheets::	21
Small Entity::	No
Latin name::	
Variety denomination name	
Petition Included::	No
Petition Type	
Licensed US Govt. Agency::	
Contract or Grant Number::	
Secrecy Order in Parent Application::	

Representative Information		
Representative Customer No. 00270	Registration Number	Name

Domestic Priority Information			
Application	Continuity Type	Parent Application	Parent Filing Date

Foreign Priority Information			
Country	Application Number	Filing Date	Priority Claimed

Assignee Information	
Assignee Name::	University of Pennsylvania
Street of Mailing Address::	3160 Chestnut Street, Suite 200
City of Mailing Address::	Philadelphia
State or Province of Mailing Address::	PA
Country of Mailing Address::	US
Postal or Zip Code of Mailing Address::	19104-6283