MODIFIED ADENO-ASSOCIATED VIRUS VECTOR COMPOSITIONS

Applicant: UNIVERSITY OF IOWA RESEARCH FOUNDATION, Iowa City, IA (US)

Inventors: Beverly L. Davidson, Iowa City, IA (US); Maria Schneel, Iowa City, IA (US); Ryan Boulbreau, Iowa City, IA (US); Alejandro Mas Montes, Iowa City, IA (US)

Assignee: University of Iowa Research Foundation, Iowa City, IA (US)

Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

Appl. No.: 14/410,983
PCT Filed: Mar. 14, 2013
PCT No.: PCT/US2013/031644
§ 371 (c)(1), (2) Date: Dec. 23, 2014
PCT Pub. No.: WO2014/007858
PCT Pub. Date: Jan. 9, 2014

Prior Publication Data

Related U.S. Application Data
Provisional application No. 61/668,839, filed on Jul. 6, 2012.

Int. Cl.
C07H 21/04 (2006.01)
C12N 15/11 (2006.01)
C12N 15/63 (2006.01)
C12N 15/864 (2006.01)
C12N 15/86 (2006.01)

U.S. CL.
CPC ... C12N 15/86 (2013.01); C12N 2750/14141 (2013.01); C12N 2750/14143 (2013.01); C12N 2830/38 (2013.01)

Field of Classification Search
CPC ..................... C12N 15/86; C12N 2750/14141; C12N 2830/38
USPC ............................... 536/24.2; 435/320.1
See application file for complete search history.

References Cited
U.S. PATENT DOCUMENTS
5,139,941 A 8/1992 Muzychka

FOREIGN PATENT DOCUMENTS
WO 9961034 A1 12/1999
WO 2011133874 A1 10/2011
WO 2012109667 A1 8/2012

OTHER PUBLICATIONS
Dong et al., “Characterization of Genome Integrity for Oversized Recombinant AAV Vector”, Molecular Therapy vol. 18 (1), 87-92 (2010).
Chen et al., “Sialic acid deposition impairs the utility of AAV9, but not peptide-modified AAVs for brain gene therapy in a mouse model of lysosomal storage disease”, Molecular Therapy. 20 (7), 1393-1399 (2012).

* cited by examiner

Primary Examiner — Anne Marie S Wehbe
Attorney, Agent, or Firm — Viksnius Harris & Padoy PLLP

ABSTRACT
An adeno-associated virus filler component comprising a nucleic acid of between 3300 and 4200 nucleotides in length is disclosed.

18 Claims, 32 Drawing Sheets
Figure 2A (SEQ ID NO: 3)

Sequence: 5pFBAAVmU6miHDS1stuffer Assembly Range: 1 to 9110

>5'-_GTVC_-G0202

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTGCTTTCTTTCCCTTCTGCAACGTTCGCGCGCTTTCTCCGTGA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>70</td>
<td>80</td>
<td>90</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>AGCTCTAATCGGGGCTCCCTTTTAGGTTCCGATTTAGTGTCTTACGCG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>110</td>
<td>120</td>
<td>130</td>
<td>140</td>
</tr>
<tr>
<td></td>
<td>TCGAGATTAGCCCCCGAGGGGAATCCCAAGGGCTAATTACGAAATGCGG</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>160</td>
<td>170</td>
<td>180</td>
<td>190</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>TCGCCCTGATAGACGGTTTTTGCCCTTTGACGTTGGAGTCCAGGTCTCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>210</td>
<td>220</td>
<td>230</td>
<td>240</td>
</tr>
<tr>
<td></td>
<td>AGCGGAGACTATCTGCCCAGAAAAGCGGGAAACTGCAACCTCAGGTGCAAGAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>260</td>
<td>270</td>
<td>280</td>
<td>290</td>
<td>300</td>
</tr>
<tr>
<td></td>
<td>CTTTCTTTGATTATAAGGGAATTTGCGGGGATTATTCAGGTTAA</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>310</td>
<td>320</td>
<td>330</td>
<td>340</td>
<td>350</td>
</tr>
<tr>
<td></td>
<td>AAAATGAGCTGTATTTAAACAAAAATTTAAACCGGAATTTTTAACAAAATATT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>360</td>
<td>370</td>
<td>380</td>
<td>390</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>AGCGCTTACATATTAGGTGCGACTTTTCGCGAGGAAATGTGCGCGGAAACCCCT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>410</td>
<td>420</td>
<td>430</td>
<td>440</td>
</tr>
<tr>
<td></td>
<td>ATTTGATTAAAAATGCTCAATTCATATGATGCCATGAGACAGACAAACAAATAAAAAGTTTATGTAAGTTTATCATAGGGAGTACTCTGT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

>Beta-lactamase

<p>| | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>460</td>
<td>470</td>
<td>480</td>
<td>490</td>
<td>500</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ATAACCCCTGATAAATGCTTTCAATAATATAGAAGAAAGAAGGATATGAGTA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
|      |      | TATGCGGACTATTACGAAGTTATATTATAACCTTTTTCCTCATCATCTAT
Figure 2B

510  520  530  540  550
560  570  580  590  600
610  620  630  640  650
660  670  680  690  700
710  720  730  740  750
760  770  780  790  800
810  820  830  840  850
860  870  880  890  900
910  920  930  940  950
960  970  980  990  1000
1010 1020 1030 1040 1050

TTCAACATTTCCGTGTCGCCCTTATCCCCTTTTTTGCGGCGATTTTGCCTT
AAGTTGTAAGAGCAGACGGGAGGAGGAGAGGAGGAGGAGA

CCTGTTTGTGCTCACCAGAAACGCAGTGGTGAAAGTAGATGCTGAAGA
GGACAAAAACGAGTGGGTCTTTTGCGACCACTTTTCTTTTCTACGACTTCT

TCAGTTGGTGCACAGCTGGTTACATCGAAGAGATCTCCACACAGCCTGA
AGTCACCCACGTGCCTCACCAGAATGTAGCTGACCTAGAGTTGCTGGCAT

AGATCCCTGAGATTTTTCGCGGAAAGGATTTTCCAATGATTGAGCAGCT
TCTAGGAACCTCCTCAAAGCGGGGCTTTCTGGACAAAGGTTACTACTCGTGA

TTTAAAGTTCTGCTATGTGGCGCAGTATATTCAGCAGCTTTGCGCGGGA
AAATTTCAAGACGGATACACCGCGCCCTCGATATAGGCATAACTCGGCGG

AGAGCAACCTCGTGCCGCATACACTATATTCGAGAATCTTCGTTGAGTT
TCTCGGGAGCCAGCGGCTATGTGATAAGAGTCTACTAGACCCGCACCCT

ACTCACCACGTCAAGAAACGATCTTTTACCGGATGCGATACGTAAGAGAA
TGAGTTGGTCACTGTCTCCTTTGATGACAAGCTGCTACCGTACTGCTATTCT

TTATGCAGTGCTGCCATAACCATGAGTGATAAACACTGCGCCAACCTACT
AATACGTCACGACGGTTATGGTGACTCACTATTGTGACGGCGGGTTGAA

TCTGACAACGCATCGGAGGAGGAGCTACCGCTTTTTTGCAACAAA
AGACTTGGTGCTAGCCCTCGCCTCCGTAGATGGCATTGGCAGAAAACGTTG

TGGGGAGATCGTAACTCGGCTTTGATGTTGGAACCCGAGCGCTGAATGAA
ACCCCGCTAGTACATGGGCGGAACCTACGAACCTTGCCCTGACTTACTT

GCCCACCACGAGGCTGACACCACAGTAGCTGCTCTAGCAATGGCAAC
CGGTATGGTTGTGCTGTGACAACAGCAGTGGTGGTGCACCGACATCGGTGTT
Figure 2E

2160 2170 2180 2190 2200
TTTGCCTGGGCTTTTGGCTCACATGGTCTTTTCTGGTTATCCCTGATTCT
AAACGACCAGAAGCAGGTAACAAAGAAAAGGACGCAATAGGGGACTAAGA

2210 2220 2230 2240 2250
GTGGATAAACCCTATTTACCCCGCTTTTGGAGTGGAGTCTACCGTGTCGCCGACAG
CACCTATGGCATAATAAGGGCGAAGACAGCAGCTCTGACTATGGCGAGGCGGCGT

2260 2270 2280 2290 2300
CCGAGACACCGGACCCAGGAGCTGACGTGAGGGGAGGAGGAGGAGGAGGACGCG
GGCCTTGCGCTGCCTCGCTCGACTGCTAGCAGCTCGTCTCCTCCTCCGTCC

2310 2320 2330 2340 2350
TGATGGCGGTATTTTTCCTTTACGCATCTGCTGCCGTTTTTACACCAGCATA
ACTACGCCATAAAAGAGGAAATGCGTACGACAGCCCATAAAGAGCTGTCGG

2360 2370 2380 2390 2400
GACCCACCGCGTAAACCTGGAAAAATCGGTACCGTGGAGTAAATAAATGGA
CTGGTCGGGCGTATTGGGCGGTATTTATGCAATGCCACAATCTCATTATACCT

>Tn7R_8-10-11

| 2410 2420 2430 2440 2450 |
TGCCCTGCGTAAACGGGAGCTGGGGCCGAAAATAGCTTTAATTCAGCAGAA
ACGGGACCGATTCGCCACACCCCGCTGTTATTCGGAAGTTGACGGGTT

2460 2470 2480 2490 2500
AATAGATCTAAAATAGCAATAAATGGTCTTTAATAGCGACAGAATAGCTTGT
TTATCTAGATTTTGATACGTTATTTTCGAAATTTGCTGCTTTATCAACA

2510 2520 2530 2540 2550
AAACTGAAAATCAGTCTACCCGTATAGCCTGGGGAAAGCATACTGGGACTTTTG
TTGACTTTTAGTCAGGTCAATACGACACTTTTTCGATGACCTGAAACA

2560 2570 2580 2590 2600
TATGGCTAAAGCAAACCTCTCTCTTTTCTGAGTGGCAAATGTGCCGTCGTA
ATACCGATTTTCGTTTGAGAAGTAAAGGAACCTCAGTTAAAAACCCGACAGCAT

2610 2620 2630 2640 2650
TTAAAGGAGGCCCAGGCACTTTGCTTTTCTGAGTGGCAAATGTGCCGTCGTA
AATTTCTCCGCCACCGGCTTCGCTACATTCTGTGATAAAACGCGCCGAAA

2660 2670 2680 2690 2700
GTGACAAATTATCTGGAGGGCGGCGCCGCGGCGGAGGAGGAGGAGGAGGAGGACGCG
CACGTGTAATGGTGGTTGAGGCCGCGCCTTTCCGCTAGAGGCACGCGAACT
Figure 2F

2710  2720  2730  2740  2750
ACGAAATTGTAGTTGCCGCTACTTGTCGATATCAAAATGATGACATCTTTC
TGCTTAACCAATCCACCGCCATGAACCCAGCTATAGTTTCAACGTAGTGAG

2760  2770  2780  2790  2800
TTCGGCTATGGCCCAACTTTTGTATAGAGACCCACTGCGGGATGGCGTCACCGT
AAGGGCATACGGGTTGAAACATATCTCTCTCGGTGACGCCCTAGCAGTGGCA

2810  2820  2830  2840  2850
AATCTGCTTGCTACAGTACATACAAACCAAGCAAGCCGTTGCGCTCCTATGC
TTAGACGAAACGTGCACTCTAGTGTTCTGTCGCAACCCGGAGTAGACG

2860  2870  2880  2890  2900
TTGAGGAAGATTGATGACCCGCGTGCCATGGCTGGCTGGCTGGCC
AACTCCTCTAACTACTCCGCACCGTACGGAGCGAGCGCCACCGAGG

2910  2920  2930  2940  2950
GGAGAATCGGAGATCATAGATGATATGATCTACATACCGGCGCTGCTCAAACT
CCTGTCAGCGCTCTCAGTATCTATATATAGTGATGCACCCGACGAGCTTGA

2960  2970  2980  2990  3000
TGGGCAGAAGCAGGAAACGCGAGCGAAGCCAAACCCCTCTGTTGTCGAG
ACCCGCTTTGCTACGGGCTCTCGGCCTGTTGCGAGAACCAGCTTC

3010  3020  3030  3040  3050
GCCGGCAAGCGSVGATGATGCTTTACTACGCGAGCAAGTTCCCGAGTGAATC
CGTCTGTCAGCGCTCTACTACGCAAGAATGACCTGTCTCCAGGGCTCCATTAG

3060  3070  3080  3090  3100
GGAGTCGCCGTGTATGTTGGAGATAGTGGCTACGTCTCGGAACCTACCGAC
CTCAGGGCGAGCTACAAACCTCTCATCCACCAGATCGAGGCTTTGAGTGCTG

3110  3120  3130  3140  3150
CGAAAAGATCAAGACAGGCAGCCGCGATGATGCTTTTGACTTGTCGAGGGCAGAGC
GCTTTTCTAGTTCTCGCTGGGCATACCTAAACTGAAACCGAGTCCGGCTCG

3160  3170  3180  3190  3200
CTACATGTGGCAATGATTGCCATACTTGGGACACCATTAACTTTTGGTTAGG
GATGTAACACGGCTTACTACGGGTTGATGAAACTCGGTGGATTGAACAAATCC

<gentamicin
 |

3210  3220  3230  3240  3250
CGGACTGCCCCTGCTGCATGAAATATCGTGGTGCTGCTCGTAACATCGGCTG
CGCTGACGGGACAGCGACATTGTAGCAACGACGACGACATTGTAAGCAACGAC
Figure 2H

>AAV_I TR_ (94bp)
  3810  3820  3830  3840  3850
  TGCAGCCAGCTCCGGCCCTGCTGTCTCACTGAGCCGCCCCGCTGGGG
  ACGTCCGTCCGACGCGGACGAGCGAGTGAACCTCGGCGGGGCCCACC

  3860  3870  3880  3890  3900
  CGACCTTCCTTTGGTGCCCGCCCTCGCTGCTGAGCCGAGCGCGAGAGG
  GCCTGGAAACCAGCCCCGAGTGCACTCGCTGCCGCCTCTCCTCCT

  3910  3920  3930  3940  3950
  GTGGCCAACTCCATCATGAGGGTCTCTTGATGTTATAGTTAATGTAAC
  CCGCCCCGGGTTTATGGTGATCTGCAAGTTCACAAAAATCAC

>mU6 promoter
  3960  3970  3980  3990  4000
  ATGCTACTTTATCTACGGTACGCATGCTCTGTGAAATGCGGCGCTGGCC
  TACGATGAATAGATGCATCGTGCAGATACCTAGTCGGGCGGTGAG

  4010  4020  4030  4040  4050
  GCGGGCCCGCCGCCGGCCGGCTCGCAGGACAGGAATGTGGGGAGGC
  GTCCGCGCGGCGGGGAGCCTGGTGTGCTGACAACCCTCTCTCGGCTG

  4060  4070  4080  4090  4100
  CTCCCGGCGGCTGGGTTAATTTTGCGATATAATATATGTTCCCGCTAC
  TAGAATGCAGAAGAGAGATGGTGTTTTGCTGATTTTTATGGTATGAT

  4110  4120  4130  4140  4150
  GCCCTTAATGTGCGATAAAAAGCACAGATAATCTGTTCTTTTTATTTATCTG
  GCTCCGGCCGGCTGGGTTAATTTTGCGATATAATATATGTTCCCGCTAC
  TAGAATGCAGAAGAGAGATGGTGTTTTGCTGATTTTTATGGTATGAT

  4160  4170  4180  4190  4200
  CATTATGATGGCTTTGAGCTTCTATAAGGATGACAAATATACTAAATT
  GTAACATGCTAPATATTACGGCTGCTATGCTATGCTATGCTATGCTATG

  4210  4220  4230  4240  4250
  ATTATTTTAAACAGCACAACAAAGGAAACTCAACCCTAACTGTAAGTA
  TAATAAAAATTTTTGCGGTGTTTTGCTGATTTTTATGGTATGATTTT

  4260  4270  4280  4290  4300
  TTGCTGTGTTTTGAGACTAAAAATATCCTCTTGGAGAAAAGGCCTTGTTGCG
  AACACACAAAACTCTGATATTTTTATAGGGAACCTCTTTTCCGGAACAAACGC
Figure 21

>miHDS1

4310 4320 4330 4340 4350
TTTAGTGAAACGTCAGATGTTACCGTTAAAACCTCGAGTGACCAGCATGCTG
AAATCTACTTGCCAGTCTACCATTGGCAATTTGAGCTACCTGACCTAGGACC

4360 4370 4380 4390 4400
CTCGCATGTTGCTAGATCTGTAAGCAACGATGGTGTCAGCACTAGCGAC
GAGCGTACCAGCTATGACATTTCGTCGGTCTACCCACAGCTGCTGCTGCTG

4410 4420 4430 4440 4450
CAGCACCAGCTACTAGAGACGGCCGCCACAGCGAGGAGATCCAGACATGAT
GTCGTGCGCGATGATCTCGCGCAGCGGTGCCCTCTAGGTTCTGACTA

>stuffer

4460 4470 4480 4490 4500
AAGATACTATTTTTGAAATTCCGCCCTCATCCAGGTTGCTCTGTGTCATGGC
TTCTATGTAAAAACTTAAAGCAGGATAGGCTTACCCAGGAAACAAGGCTCC

4510 4520 4530 4540 4550
AAATGGGAGGTTAAGGACAGAGAAATATGAACAGAAACCTGTTCTAAT
TTTACCCTGCACTTCTCCGGTACTTCTTTGAGTCTTTGACAAAGATTA

4560 4570 4580 4590 4600
ATTTGGCTCATTTTTATGTTAAGTATTTGGTTTCTTTTTTAACCTCCTATTTT
TAACCAGTAATATACATTACAATTTTACAACAGAAGAAAATTTGAGAAAGTAAAA

4610 4620 4630 4640 4650
TTTTTTCCAGGAATTTTGCGGAGACAGGTGTGCTGTTGCTGCTGAGGACTG
AAAAGGTTCTCTAAAAACGACCTGTGTCACCCGCAACACACACAGACTCGAC

4660 4670 4680 4690 4700
TAGGGCATGGCCCTAGGTTGGTTTATGTTCTCAGGTGCTTTCTCTGCG
ATCCGGTACCAGGGATCCACAAACACAATCCAGATCCACAGGAGGGACCG

4710 4720 4730 4740 4750
TGCTCCTCTTGTCTCTTTTTTTTACATGTCCTCCAGGCGGCTTTCTCTCT
ACAGAAGGAACGAAGAAAGGTTACAGGAAGAAACAAAGGTTGGCGGTAAGGA

4760 4770 4780 4790 4800
CCCTTTATGCTTAAAGGTTGGCAGCAGGGTGGTTTGCTGCTCTCGAGATCCCT
GGGAATACGAAATTCCAAACCACGTCGGTCCCAAGCCCGAGGAGAGCTCAGGA


Figure 2K

5360 5370 5380 5390 5400
AGGCTGACCTGACCATCTCTATGCTGCAGGAGGTGAGGCACCATGAGGAGGTCA
TCGACTTGTTGCGTGAATACGACGCCGCTCTCACCCTGGTGTACTCCAGTC

5410 5420 5430 5440 5450
GTCAAGGGTGTTTGTACTTTTTGTTAGAAAAAGGTGCTCTTGAAGACT
CAGTTCCCCAACCATGGAACACAAACCATCTCTTTAAATCCCCGAGAACCTTC

5460 5470 5480 5490 5500
TTGGATGTTGTCAGGAGGTGATCTACTTTAGAAGAGTGCAGCGGTGAGG
AACCTACACAGTCCCCCTCACAATAGTTAAATCTTTTCTACGGGCGACTCC

5510 5520 5530 5540 5550
ACGGGGGTTAGAGGAGAAGACGTTGGAGGGAGTCCAGGATGGGAGTGA
TGCAACCCCATCTCTCTCTGTCCACTCTCCTCCCTCAAGTCCACCCTCACACTCAT

5560 5570 5580 5590 5600
GACCCAGCGAGTGCGAGGGCCCTCGAGCAGGACGTGTGGCAGGCTGCTGA
CTGGGTGCTCCTCAGTCGTCGCCGACGAGTCGCTCTACTACCCGGTGCCGCACCT

5610 5620 5630 5640 5650
GGAGAGGCGAGCCACCTGTGTGCTCTGCGGAAAGCAGGGGAGGAGGAGAG
CCTCTCCGTGGTGAGACACACAGACGGCTTCTGTCGCTCCCCGTCTCCCTTCCT

5660 5670 5680 5690 5700
GCCAGCGCGCTGTGCCACCCGCCAGCGACTGCGTGATGGTGGAGAGAC
CGTTGCGTCGCACGACCGTAGGTGCTGGCGCTGACCGCATCATAACACTCTCTG

5710 5720 5730 5740 5750
CATTCCCTGGCTTCTTAGAGGGGCTGAGTTTTATAGTGTCTCTGTGTTATACA
GTAAGGGACAGGAATCTCCTCCCCGACTCAAAAATCAAAGAAGAAATAATG

5760 5770 5780 5790 5800
ATAGGCTGTGTTTTTTTATTTTGGTAAAGCTAAATCAAGGTTT
TATTCGAACCATAAACAAATGTTTTTGTAACATTTCGATTTAGTTCCAAA

5810 5820 5830 5840 5850
GATAAGCTTTCTAGTGTGTATTAAAGAACTATGTTGAAATAATATGTGTTG
CTATTCGGAAGATCAAAATAATCTCTCTCATTACAACAATTTATTATCAAAC

5860 5870 5880 5890 5900
CCAATTGCCCTTGTCTCATTAAAAGACCTTCTAGTACAAACTCAGCAACACAAG
GGTTAAGCCGAAACGAGTAAATCTCTGTGACGTGTTTGTC
Figure 2M

<table>
<thead>
<tr>
<th>6460</th>
<th>6470</th>
<th>6480</th>
<th>6490</th>
<th>6500</th>
</tr>
</thead>
<tbody>
<tr>
<td>TGGGATGTAAGAATACATACATCTTCATTTGATTTAGTTAGGGAG</td>
<td>ACCCTACAACCTGTATCTTTAGTTGAAAGATGACATCTACATCAAATCCCTC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6510</td>
<td>6520</td>
<td>6530</td>
<td>6540</td>
<td>6550</td>
</tr>
<tr>
<td>AAATTTGATACAGTTCCCTGTCGAGTTAGCCAGAATGTTAGGGTTGCTT</td>
<td>TTGAAATATCAGGTGCACAACCATCTACTACAGCTCTACGCCACCGAAA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6560</td>
<td>6570</td>
<td>6580</td>
<td>6590</td>
<td>6600</td>
</tr>
<tr>
<td>CCTATTTGTTTCAGAACACTTTTTAAATTAATCTCAGAGAAATGAAATTTAA</td>
<td>GGATAAACAAGTCTTGAAGAAATTTATATGAGCTTGTATACCTAAATATT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6610</td>
<td>6620</td>
<td>6630</td>
<td>6640</td>
<td>6650</td>
</tr>
<tr>
<td>AGGATTTAAAAAAAAAATCTTTGGAATGATTATTTTTCACACATAGGCCTT</td>
<td>TCCTAAAAATTAAAAATTAATGAATGTTACCTGAAATAGTGATGATCTAAA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6660</td>
<td>6670</td>
<td>6680</td>
<td>6690</td>
<td>6700</td>
</tr>
<tr>
<td>CTTGATAAATGAAATTCCTCAGGTATTCTCTCTGTGTGTTACTCAAATAGTTA</td>
<td>GAACTATTTTTATTTAGCTTGAAAGAGGAGAACAAATGATAGTATCAAT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6710</td>
<td>6720</td>
<td>6730</td>
<td>6740</td>
<td>6750</td>
</tr>
<tr>
<td>CCCTTTATGATTTAAAAATTTTTTTTTTTTTTTTCTGAAAATTTTATGCAATCAGACTT</td>
<td>GAAAGAATACCAAAAAAAGGGGACTTTTTAGTAAATAGTTGGCATGCAACC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6760</td>
<td>6770</td>
<td>6780</td>
<td>6790</td>
<td>6800</td>
</tr>
<tr>
<td>CTATTTTTCTGGAAGATGTTTGTGATAATTTTGGAAAGATGGATGGACCTCAGA</td>
<td>GAATAAAAGACTTCTACTACAAACTATTTAACCCTTCTATTTTACGAAAGT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6810</td>
<td>6820</td>
<td>6830</td>
<td>6840</td>
<td>6850</td>
</tr>
<tr>
<td>TATTTTTACAGGTTTGGGATCTTCTTSGCTTCTGTGGTCGTGTCAGCACTTTTTTTGATGTTAGGGAGCAGACACAATCTCCAGAAGAAAAATGTAGAGTACCTCAGTC</td>
<td>AGGGTTCTCCTCTTGATCCAACACTTTTTTTAACGGAATCTCTCTTGAAAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6860</td>
<td>6870</td>
<td>6880</td>
<td>6890</td>
<td>6900</td>
</tr>
<tr>
<td>TCCCAAAAGCAGAAAGACGGCATGTTTGGGAAATGCGTACAGAGATACCTC</td>
<td>AGGGTTCTCCTCTTGATCCAACACTTTTTTTAACGGAATCTCTCTTGAAAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6910</td>
<td>6920</td>
<td>6930</td>
<td>6940</td>
<td>6950</td>
</tr>
<tr>
<td>TTTTCCACTTCTTAAAACTCATTACATCTCACTTGAAATTTTCCAGGACACTTTTC</td>
<td>AAAAGGTTGCCAAAAAAGTTAGTAGTGATGAGAAACTTTAAGTACGCTCCGTGGAAAAG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6960</td>
<td>6970</td>
<td>6980</td>
<td>6990</td>
<td>7000</td>
</tr>
<tr>
<td>CATGCTCTCTGCTTTGCTATCTGTGTTTAGATTTTTTTCCTTCAATACC</td>
<td>GTACGAGGATCAGGAAAGATAGCAGCAGACAAAAATAAAAGGAAAGGACCTTTATGGG</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2Q

8610 8620 8630 8640 8650
ACTTGTATTATGCAGCTTATAATGGTTACAAATAAAAAGCAATAGCATCACA
TGAACAAATAACGTCAATATAATCTTATATTTTCTGTTATCGTAGTGT

8660 8670 8680 8690 8700
AATTTTCACAATAAAGCATTGGTTTTCACTGCAATTCTAGTGTGTTTTGCT
TTAAAGTGTTTTTCTGATAAAAAAGTGCAGATACACAACACCAACAG

8710 8720 8730 8740 8750
CAAACTCACTCAATGTAATCTTATATCTGATGCTGATGATCAGTATGC
GTGTTGAGTAGTTACATAGAATATAGTACAGACCTAGCTAGTGAATAGCG

>Tb7L

8760 | 8770 8780 8790 8800
CTAGGAGATCCGAGACAAGATAATCTAGTCCCAAAACTATTTTGTCC
GATCCTCTAGGCTTGGCTATTCATTTTCTAGATCAAGGTGGTATAAACAG

8810 8820 8830 8840 8850
ATTTTTAAATTTCGTATTTAGCTTACGAGCTACACCCAGTTCCCATCTAT
TAAAAATTTAAAGCATAATCCGAGTGTGGTGGTCAAGGTTGAGTTA

8860 8870 8880 8890 8900
TTTGTACACTCTTCTCCTAATAATCTCCTTTAAAAACCTCCATTCCACCCCTCC
AAACAGTGAGAAGGGATTTTATTACGAATTTTTGAGTTAAAGGTGGGAGG

8910 8920 8930 8940 8950
CAGTTCCCACATATTGCTTCGCCACACCGGGCATTCCCCCTTCCTTCTCCGTT
GTCAAGGGTTGATAAAAAACAGCGGGGTGTCCCGCGCAAAAAGAAGGACAA

8960 8970 8980 8990 9000
ATGGTTTTAACTAAAAACTCCTGCAACTTCATGTGACAAAACGTGATCTTT
TACAAAAATATTGGTTAGGACGTTGGAGGTACACTGGTTGCGACTAGAA

9010 9020 9030 9040 9050
CGGCTACTTTTTTCTCCTGTCACAGAGAATGAAAAATTTTTCTGTCATCTCTTCCG
GCCGATGAAAAAAGGACAGTGTGCTTTACTTTTAAAAAGACAGTAGAGAAAGC

9060 9070 9080 9090 9100
TTATATTAGTTTGTAAATTGACTGAATATCAACGCCTTTTTTGTGCAGCCTGAA
AATAATTACAAACATTTACTGACTTATATTGTTGCGAATAAACGTCGGACTT
Figure 2R

9110
TGGCGAATGG
ACCGCTTACC
stuffer sequence
GAAATTCGGCGCTACCCAGGTTGCTGCGATGCAATGGGACGTAAAGAGGCGA
GAGAGATATGAAACAGAAACACGTAGTTCAATATGTTGTCAATTAAATGTGTGTAAGATATTGTTT
CTTTTTAATACCCCTTCTTCTTTTTTTACGAGGATTTGCTGACACTGTTGCTGAT
GTGCTGAGGGACTGAGTACCCGCAGTTGCTGAGTTTGTGGTTTATGCTGGTCTTCT
TGGCGTCTGCTCTGCTCCTCTTTCTCCAGCTGCTCTCTTTTTTCTCCACGCGTTCTC
ATTATGCTTTAGTTTGGGCTCGAGGGTTTTTGGGCTGCTCTCAATCTCTGCTTTCCAG
ATGCTGATGTTGTACGAGCCAGCCGCGGCAGAGGGATACGAGATCTGCTGCTAGTTT
GCTGCTCTAGTGGAGCTGAGTTAGTTGCTGAGGGCGTGAGTTGAGGACAGGCTTG
GCAGGATATGAGTTGCTGAGGGCGTGAGTTGAGGACAGGCTTG
AATTCTCATATGTTGCGATAAAAGTCGCTGCTGCTGAGGTTGCTGCGAAATCTACACCAGAC
GTGCAAGAGCGCCGAGTATTTTGTCTAGGGAGGACAGAGACTTTTCTGCTGAGGAGC
CAGCCAATAGTGAAAGATTGCGGACCTACTTACATTCACGAGGACAGGCTTG
TCAGGAGGCTGACAAGTTGAGGTCGACAGGCTTG
GACACCAGAGCTGAGTTAGTTGCTGAGGGCGTGAGTTGAGGACAGGCTTG
TCGAGGAGACTTTTGGGAGTTGCTGAGGGAGGAGTTAATTGAGGAGAATCTGCTG
AGTGTATTCTTTCTTCTTTATACTAAATAGGCTGTTGATATTGTTGTTATTAAACGACATTG
AAAGCTAAAATCGAGGGTTGTAAGATTTACTTACCTTCTATTAAAGCTAAGTATGTTGAA
TAAAGTTTTTCGCAAATTCGCTTCTTTTCGCTCTTTATTTAAAGACTTTCTAGCTGCAAGG
ACGGATTAGGTTAAACAGTTTCTGAGATGTTTTTACTTCTCAGAAATTTCAGCAG
TGTGATCTGTGTTTTTGATTTTTCAACGTGCTGACAAAAATGTTTACCCACAGGTTT
ACGAGAGCTACGATCAGGTCCGACATCAACAGTGCAAGCTGAGTTAGGAGAGAG
TCGGCTGAGGAATGTTGGGCCAAGCTGCGCAGAGCTTGATCTGCGAGGGGCGA
GGACACGGCAGATCCGCTGAGGAGCTTCTAGGAGGATCTTCCCTGCTGTT
GTCTTATGAAATTTTCTGAGATTCTCCTTTTATTGAGTTTTGGAGATGAAACATAGA
ATCAACTCTACTTTGAGATTTTGGAGGAAACTTATTACCCAGATTTGTTAGC
CCCTGTTCAAGAGTTTTCTTCTTCTTTATTTATTACAGTTTAAATATTTTTTACCTTACG
AAGACGATGAAATTTAAAGAGTTTTTAAAAAAACTTTAAAGATTATTTTACATAGCTC
TTGGCAATTTTCTGATAAATGATCTGATATTCTCCTCTTCTTGGTTATTAAGTAA
TATCTTCTATTGCTGTTGAAATTACATTTAAAGCTAAGGATGTTT
TTTTCTGAGGAGATTGGTATTTTGGGAAGATGAAAGTGAAGCTTACTATTTAAGC
GTTTATGAGGTCTCTTTTATTTTGCTCTCAGTTTTTGCAAGCAGAAAAGG
CATGATGAAATTTTCTGAGGAGAATATTCTTCTTTTCCACGCGTTTTTCAACTTCAAT
CATCTTGAATTTCAGGGCACACTTCTCCATGCTCTTAGTGCTTGTCTATTCTGATATTCT
Figure 3B

GGCCTCCACGAGCTTCGGACCTGCTCTGGGTCTGTAAGCCCGCCTGCGTCACTG
CTTGCTGCACGACCTCTCAGAGTGGGTGCTGAGTGGCTGCCGCTGGTGCTCTAG
TCTTGAGGATTATCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
ACTCCCACTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
TTTTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
AGGTTTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
CTTCCATGGTGAACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
CTGTTGACGGGAGAGTCAACAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG
GGGCTGGCTCTGATACATGACATCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG
GTACATCTCTAGTTTCTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC
CTTCTGACTTTGCCCCAGAGAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT
ACTGACACCTTCTTGAGGCGATGGTTTTATAAATAATTTATAATTTATAATTTAT
TTGGATATAAAATATGAACTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT
ATTATGAAAGAAATTTTCAGTTATTATTATATATTATTATATTATTATTATTATT
CTTTGAGAAATGAGAGTATTATTATTATTATTATTATTATTATTATTATTATTATT
TTCCAGTGCTGACAAACCAACACAGCAACACCCTGCAAGGTGACTGAGTACGCG

mHDS1 sequence
CTCGAGTGAGCGATGCTGGCTCGCATGTCATACGTAAGGCCACAGATGGGTGTC
GACCATGCGACCCAGCACCGCTACTAGA

mU6 promoter
CGAGCGGCCGCTTCTGACTGAGAGGACCTCTCTGACAGACTTTGTGGGAGAACGT
CGCTCTCCCTCGGCCGTTTAAATTGGATATAAATATTTCTCAGTAAACTATAGAAC
TTTAAGTGGCATAAAGACAGATAACTCTCTGCTTCTCTCTTAAATCTAGCTCATATTT
TGTAGGCTTGGATTCTATAAGAGATACAAATACTAAATTATTATTATTAAAAACA
GCAAAAAAGAAACTACACCTAATGTAAGTAAAGTAGTATTGTGTGTTTTGAGACTTAAAT
ATCCCGAATGAAAGACCTTTGTTT

AAV2 ITR (94bp)
CTGGCAGCTCCTGCTCAGCTCACTGAGGCGCCCGCGCGGCGGCGGAGCCCTTTGTCGCCC
GCCTCAGTGAGCGAGCGAGCGAGCGAGAGAGGAGTG

AAV2 ITR (128bp)
AAGGAAACCCCTAGTGATGGAGTTGGCGACCTCCTCTCTCGCGCGCCTCGCCTGACTG
AGGAGGGCGACAAAGGCTCGCAGGCGGCGGCGGCGGCTTTGCGCCCCGGGCGCGCTTCCAGTGAG
CGAGCGAGCGCG
Figure 3C

Gentamicin
TTAGGTGGCGGTACTTGGGTGCTGATATCAAAGTGCATCACTTCTTCCGATATGCCCAAC
TTTGTATAGAGAGCCACCTGGGGATCGTCACCGTAATCTGGTGCAGTGATGCA
TAAGCACCAAGGCCGTGGTGCCTCTATGCTTGAAGGATAATGAGACGGGCTGGAATG
CCTGCTCCTCGGGTTCTTCGAGGACTCAATGATATAGATATGGTACTACG
GCTGCTCAAACTTTGGCGGAGAACGTAAGCCCGAGAGGCCAACAACGCCGCTCTTTG
GAAGGCAAGAACGCCGCTGAATATCTTTACTACGGAAGAAATTTCCGAGTAAATCGGA
GTCCGCTGATATTTGGGAGATGTTGCTAGTCTCCGAACTACGAGCAGGAAAGGATC
AAGACAGACCGCGATCTGGATTTCGCTCAGGCGAGGCTACATGTTGGAATATG
GCCCATACTTGTAGGCGCCACTAACTTCTTGAAGGGCGACTGCCCCTGCTCGGTAACATCG
TTGCTGCTGCGTAAACAT

Beta-lactamase (Ampicillin)
ATGAGTTATTCAACATTCGTCGGCCTTTGATTCCCTTTTTTTGGCGGATTTTGCTTC
CTGTTTTTTGACCGAAGCGCTGTGAAAGTTAAAAAGATGCTGGAATACGAGTTCTG
GTGCAGGTGCGTTGCTCAATCACAAGCAGTTAGATCCTTGGAGAGTT
TTTGCCCCCGAAAGCAGTTTCCCAATGATAGACGGCCTATTAAAGTCTGCTATGTCG
GTTATTATCCCGCTTATGACCGCGGCAAGAACGAAACTCGTGCTCGCGCATACACTTCTTCT
CAGAACTGACCTTGGTATGATACGCACTACGCAAGAAAGCTCTTACGCGGATGGAATG
ACAGAAGAGAAATTATGCAATGCGTCGAGTCAATAACATGAGTGAATAACACTGCGCCCAAC
TTACTCTGACAAACGATGGGAGAACAGAGCTAACCCTGTCTTGGCCAAACATG
GGGAGCTATGTAACCTTCGCTTTTGATGCTTTGGAAACGGAAGCTGGAATAAGGCATACCA
AAGACGGAGGTGACGACCAAGCGTGCGATAGAATGCAGCTAAGCACTGCGCAACAT
TTAACTGGCGAACTACTTACTCTAGTCTCCGCGAAAACATATTAAATAGACTGGATGGAG
GCAGTAAATCTTCGAGACACTTCTCTCGGCGTCGGCGGGCTCGTGCGTTATTGG
CTGTATAAATCTGGAGCGGTTGAGCTGCTGCTGCTGCTGATACATTGACGACTGGGCGC
CAGATGGTGAACCGCCCGTCTACTTTGCGTATCTTACGCGACGGGAGTGCAGGCAACTA
TGGATGAAACGAAATAGACGATCCGTTGAGATAGTGCTCCTCAGTATGAAAGCATATTGC
AA

Tn7R (Transposable element)
TGTTGCGCGGACAAATAGCTTAAAATCTGCAAACAAATAGATTAACTATGCAAAATA
AGCTTTAATCAAGACAGAACATTTGAAACTGAAATCAGTGTTATATGCTGTAAGAA
AAGGATACGTCGACTTTTGTATAGGCTAAAAGCAAACATTCTTATATTCTGTAAGAAA
TTGCCGTCGATTAAAGGGCGGTGGCGCAAGGCGATCGTGTAAGAAC

Tn7 (Transposable element)
AACCAGATAAGTGAATAATCTAGCTTTCCCAATTATTATTTGTGATTTTTATTATTTTCTGATTAG
GCTTTACGACGTACACCCGATTCCTCCTATATTATTGTGACTCTTTCCCTAAATAATTCTT
AAAAACTCTCATTMTCCACCCGCTCAGTTCCCATAGTATTTTGTCCGGCAACA
Figure 5

5pFBAAVmU6miHDS1-stuffer

Map Features:

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>494 - 1354</td>
<td>Beta-lactamase</td>
</tr>
<tr>
<td>2418 - 2642</td>
<td>Tn7R</td>
</tr>
<tr>
<td>27009 - 3242</td>
<td>Gentamicin</td>
</tr>
<tr>
<td>3810 - 3928</td>
<td>AAV ITR (119bp)</td>
</tr>
<tr>
<td>3967 - 4249</td>
<td>mU6 promoter</td>
</tr>
<tr>
<td>4332 - 4417</td>
<td>miHDS1</td>
</tr>
<tr>
<td>4465 - 8239</td>
<td>Stuffer sequence</td>
</tr>
<tr>
<td>8293 - 8423</td>
<td>AAV ITR (130bp)</td>
</tr>
<tr>
<td>8764 - 8929</td>
<td>Tb7L</td>
</tr>
</tbody>
</table>

5pFBAAVmU6miHDS1-stuffer
9111bp
Figure 6A

Plasmid sequence (SEQ ID NO:12):

```
TTGCGTTTTTCTCCCTTCTTTCTCGCAGTCTTGGCGCTTCCCGTCAAGCTTCTAA
TCGAGGGGTCGCCTTTCTAGTGTTGCTCGATAAGCTCGTAAATAGGATCTCCGACCCCTGAA
CTTCTGATAGGGGATGGATTTCCGAGTACTGATGGAAGAGTTTCTTGACCCTGACAA
CTTCAACCCGCTATCTCGGCTATTTATCTGTTATAGGAAAGTCGTCGTTTCCG
ATTGTTAAAAAATAGCAGCTATTTAACAAAAAATTTAACGCGAATTTTAAACCAAAATA
TTAACGCTTACATTAATTTAGGATGCTGCTCCCTTTCTGCGGAAATAAGCGGAGAAGCTACTT
GTATTTTTTCTAAATACATTTCAAATATGTACCGTCAGACAAATTACCTGTAT
AAATTCGTTCAAATATTGAAAAAGGGAAGATGATGATATCTCAACATATTTTGGTGTCG
CCTCTATCCCTTTTTTATTCGAGATTTTGCCTCTTGTTTCTTCTGACCAAGAAGGCT
GGTGAAAATGAAAGTGCTGAAAGATCGATTGGGATGCGACTGGTGGTACATCGAAACT
GAGACCTCAACAGCGTTAAAGATGCTGAGAACAGCTGACTCTCGCCTGCTGAC
AGATCGTAGCTTTAAAATAGTCTGAGATCTGACGAGTGCGACACAGATGGTCTCCTAG
AGCAGATCGCCTCGGGGCCCTCGGCTTRCTTTTTATTCTGCTGTAATTGTCGAGGCGCTC
CTCAGGGGATCAATGCTACCACTGTGAGGTACGCTGCACAGCTGACAGCTTATAG
ATTATTTTTTCTTTTTTTTCGGAACCCAGCTGACGCTGAGCTGAGAAGGCGGAGGCGGCA
GTAGCTTCCGAGCTTAAGATGGACGAGTGCTGATCTGACAGCTGACAGCTGACAGCTGAC
GAAATGGCTTCCCTTTTCTGCTGACGCTGAGCTGAGAAGGCGGAGGCGGCA
```


**Figure 7A**

**Stuffer sequence (Stuffer #2) (SEQ ID NO:2)**

```
GGGCTATCCAGGTTGCTTCTGTTACTGGCAAAATGGGACGTATAGAGGCGAGGAGA
ATATGAAACAGAAAATCTGCTTAATATTGGCTTTATATATGTGTAAGTGATTGCT
TAAAGCTCTCTTTTCTACAGAGATCTTGCCAGACAGTCGGCCTTTGCTTGTCCT
CAGACTGTAGGCTGAGCCCTGTGCTTCTCTCTTTTTTCCTGAGTCATGTAAG
TGTTGTCAGCCAGGCTGAGGTGGTGGCAGATCATGCTGCTGGTCTCTACAT
TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT
CCCTCTCGGCAGGTAGTCTACGCTACAGGCGTACCAGACAGCGAGAAACAGA
TATATTGTGCGATTACGAGCGAGGACGTGCTGCTGCTGCTGCTGCTGCTGCT
AGGCTGAGTTCTGATGGAGGACGAGAGAAACATGCTGCTGCTGCTGCTGCT
GATTGCAAGCGAGTTCTGATGGAGGACGAGAGAAACATGCTGCTGCTGCTG
GGCTGAGTTCTGATGGAGGACGAGAGAAACATGCTGCTGCTGCTGCTGCT
TGGTGTTCAGCCAGGCTGAGGGCTGAGGACGTGCTGCTGCTGCTGCTGCTGCT
CCCTCTCGGCAGGTAGTCTACGCTACAGGCGTACCAGACAGCGAGAAACAGA
TATATTGTGCGATTACGAGCGAGGACGTGCTGCTGCTGCTGCTGCTGCTGCT
AGGCTGAGTTCTGATGGAGGACGAGAGAAACATGCTGCTGCTGCTGCTGCTG
```

Figure 8

AAV2/1 mU6miHDS25Intron/I/II

96.6% full virions
1.07E+13 vg/mL

Figure 9
Table 1

<table>
<thead>
<tr>
<th></th>
<th>Cap/rAAV</th>
<th>Amp/rAAV</th>
<th>Gent/rAAV</th>
<th>Avg. Empty %</th>
<th>QPCR Titer (vg/ml)</th>
<th>Total vg/ml</th>
<th>Total # of (pt/ml)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAV2/1mU6miSafeIntron1/II</td>
<td>0.00%</td>
<td>0.06%</td>
<td>0.15%</td>
<td>1.30%</td>
<td>2.75E+13</td>
<td>2.76E+13</td>
<td>2.79E+13</td>
</tr>
<tr>
<td>AAV2/1mU6miHDS26Intron1/II</td>
<td>0.15%</td>
<td>1.81%</td>
<td>1.29%</td>
<td>2.00%</td>
<td>3.23E+12</td>
<td>3.34E+12</td>
<td>5.33E+12</td>
</tr>
<tr>
<td>AAV2/1mU6miHDS26Intron1/II</td>
<td>0.80%</td>
<td>2.14%</td>
<td>7.87%</td>
<td>3.90%</td>
<td>1.09E+13</td>
<td>1.22E+13</td>
<td>1.27E+13</td>
</tr>
<tr>
<td>AAV2/1mU6miHDS25Intron1/II</td>
<td>0.19%</td>
<td>1.34%</td>
<td>1.02%</td>
<td>0.90%</td>
<td>2.74E+12</td>
<td>2.81E+12</td>
<td>3.73E+12</td>
</tr>
<tr>
<td>AAV2/1mU6miHDS25Intron1/II</td>
<td>0.08%</td>
<td>0.28%</td>
<td>1.98%</td>
<td>2.70%</td>
<td>1.07E+13</td>
<td>1.09E+13</td>
<td>1.12E+13</td>
</tr>
<tr>
<td>AAV2/1mU6miHDS10Intron1/II</td>
<td>0.12%</td>
<td>1.40%</td>
<td>0.87%</td>
<td>5.60%</td>
<td>3.52E+12</td>
<td>3.60E+12</td>
<td>3.80E+12</td>
</tr>
<tr>
<td>AAV2/1mU6miHDS1Intron1/II</td>
<td>0.01%</td>
<td>0.15%</td>
<td>0.15%</td>
<td>0.70%</td>
<td>1.81E+13</td>
<td>1.82E+13</td>
<td>2.08E+13</td>
</tr>
</tbody>
</table>
MODIFIED ADENO-ASSOCIATED VIRUS VECTOR COMPOSITIONS

RELATED APPLICATIONS

This patent application claims the benefit of priority of U.S. Application Ser. No. 61/668,839, filed Jul. 6, 2012, which application is incorporated by reference herein.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted in ASCII format via EFS-Web and is hereby incorporated by reference in its entirety. Said ASCII copy, created on Mar. 14, 2013, is named 17023.126WO1_SL.txt and is 39,125 bytes in size.

BACKGROUND

Adeno associated virus (AAV) is a small nonpathogenic virus of the parvoviridae family. AAV is distinct from the other members of this family by its dependence upon a helper virus for replication. The approximately 5 kb genome of AAV consists of one segment of single stranded DNA of either plus or minus polarity. The ends of the genome are short inverted terminal repeats which can fold into hairpin structures and serve as the origin of viral DNA replication. Physically, the parvovirus virion is non-enveloped and its icosohedral capsid is approximately 20 nm in diameter.

To-date many serologically distinct AAVs have been identified and have been isolated from humans or primates. Govindasamy et al., “Structurally Mapping the Diverse Phenotype of Adeno-Associated Virus Serotype 4,” J. Vir., 80 (23):11556-11570 (2006). For example, the genome of AAV2 is 4680 nucleotides in length and contains two open reading frames (ORFs). The left ORF encodes the non-structural Rep proteins, Rep 40, Rep 52, Rep 68 and Rep 78, which are involved in regulation of replication and transcription in addition to the production of single-stranded progeny genomes. Rep68/78 has also been shown to possess NTP binding activity as well as DNA and RNA helicase activities. The Rep proteins possess a nuclear localization signal as well as several potential phosphorylation sites. Mutation of one of these kinase sites resulted in a loss of replication activity.

The ends of the genome are short inverted terminal repeats (ITR) which have the potential to fold into T-shaped hairpin structures that serve as the origin of viral DNA replication. Within the ITR region two elements have been described which are central to the function of the ITR, a GAGC repeat motif and the terminal resolution site (trs). The repeat motif has been shown to bind Rep when the ITR is in either a linear or hairpin conformation. This binding serves to position Rep68/78 for cleavage at the trs which occurs in a site- and strand-specific manner.

The following features of AAV have made it an attractive vector for gene transfer. AAV vectors possess a broad host range; transduce both dividing and non-dividing cells in vitro and in vivo and maintain high levels of expression of the transduced genes. Viral particles are heat stable, resistant to solvents, detergents, changes in pH, temperature, and can be concentrated on CsCl gradients. AAV is not associated with any pathogenic event, and transduction with AAV vectors has not been found to induce any lasting negative effects on cell growth or differentiation. The ITRs have been shown to be the only cis elements required for packaging allowing for complete gutting of viral genes to create vector systems.

There is a current need for AAV vectors that have improved packaging features.

SUMMARY

In certain embodiments, the present invention provides an adeno-associated virus (AAV) filler component (also called a “stuffer sequence”) comprising a nucleic acid of between 3300 and 4200 nucleotides in length having at least 90% identity to SEQ ID NO:1 or SEQ ID NO:2.

In certain embodiments, the present invention provides an adeno-associated virus (AAV) filler component consisting of a nucleic acid of between 3300 and 4200 nucleotides in length having at least 90% identity to SEQ ID NO:1 or SEQ ID NO:2.

In certain embodiments, the present invention provides an AAV vector comprising the filler component described above.

BRIEF DESCRIPTION OF THE DRAWINGS AND TABLE

FIG. 1 is a plasmid map of 5pFBAAVmU6mHDS1stuffer (9110 bp).

FIGS. 2A-2R collectively provide the sequence of 5pFBAAVmU6mHDS1stuffer (Stuffer #1) (SEQ ID NO:3).

FIGS. 3A-3C provide the sequences of the various individual components of 5pFBAAVmU6mHDS1stuffer (SEQ ID NO:1, 4-11).

FIG. 4 is a graph showing relative Htt expression.

FIG. 5 is a plasmid map of 5pFBAAVmU6mHDS1stuffer.

FIGS. 6A-6D collectively provide the plasmid sequence for 5pFBAAVmU6mHDS1stuffer (SEQ ID NO:12).

FIGS. 7A-7B collectively provide a stuffer sequence (Stuffer #2) (SEQ ID NO:2).

FIG. 8. EM evaluation of full virions vs. empty virions. Two examples of empty virions are highlighted by the arrows. This prep had only ~4% empty virions, which is quite low.

FIG. 9. Silver stain to examine the capsid integrity of the purified virions. Several different miRNA-expressing constructs were engineered into the shuttle vector along with the intron I/II stuffer to generate near wild type genome size. The purified viruses show optimal VP1, VP2 and VP3 protein ratios.

Table 1. % Packaging efficiencies of miR-intron I/II virions and % contaminants.

DETAILED DESCRIPTION

AAV Vectors and Expression Cassettes

The viral vectors of the invention utilize an AAV vector. An “AAV” vector refers to an adeno-associated virus, and may be used to refer to the naturally occurring wild-type virus itself or derivatives thereof. The term covers all subtypes, serotypes and pseudotypes, and both naturally occurring and recombinant forms, except where required otherwise. As used herein, the term “serotype” refers to an AAV which is identified by and distinguished from other AAVs based on capsid protein reactivity with defined antisera, e.g., there are eight known serotypes of primate AAVs, AAV-1 to AAV-8. For example, serotype AAV-2 is used to
refer to an AAV which contains capsid proteins encoded from the cap gene of AAV-2 and a genome containing 5' and 3' ITR sequences from the same AAV-2 serotype. Pseudotyped rAAV refers to an AAV that contains capsid proteins from one serotype and a viral genome including 5'-3' ITRs of a second serotype. Pseudotyped rAAV would be expected to have cell surface binding properties of the capsid serotype and genetic properties consistent with the ITR serotype. Pseudotyped rAAV are produced using standard techniques described in the art. As used herein, for example, rAAV1 may be used to refer an AAV having both capsid proteins and 5'-3' ITRs from the same serotype or it may refer to an AAV having capsid proteins from serotype 1 and 5'-3' ITRs from a different AAV serotype, e.g., AAV serotype 2.

The abbreviation “rAAV” refers to recombinant adenoadenovirus, which is also referred to as a recombinant AAV vector (or “rAAV vector”). In one embodiment, the AAV expression vectors are constructed using known techniques to at least provide as operatively linked components in the direction of transcription, control elements including a transcriptional initiation region, a DNA of interest and a transcriptional termination region. The control elements are selected to be functional in a mammalian cell. The resulting construct which contains the operatively linked components is flanked (5' and 3') with functional AAV ITR sequences.

By “adenoadenovirus vector inverted terminal repeats” or “AAV ITRs” is meant the art-recognized regions found at each end of the AAV genome which function together in cis as origins of DNA replication and as packaging signals for the virus.

The nucleotide sequences of AAV ITR regions are known. As used herein, an “AAV ITR” need not have the wild-type nucleotide sequence depicted, but may be altered, e.g., by the insertion, deletion or substitution of nucleotides. Additionally, the AAV ITR may be derived from any of several AAV serotypes, including without limitation, AAV-1, AAV-2, AAV-3, AAV-4, AAV-5, AAV7, etc. Furthermore, 5' and 3' ITRs which flank a selected nucleotide sequence in an AAV vector need not necessarily be identical or derived from the same AAV serotype or isolate, so long as they function as intended, i.e., to allow for excision and rescue of the sequence of interest from a host cell genome or vector.

AAV ITRs can be excised from an AAV vector plasmid containing the same and fused 5' and 3' of a selected nucleic acid construct that is present in another vector using standard ligation techniques, such as those described in Sabbrook and Russell, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press Cold Spring Harbor, N.Y. (2001). For example, ligation can be accomplished in 20 mM Tris-Cl pH 7.5, 10 mM MgCl2, 10 mM DTT, 33 μg/ml BSA, 10 mM-50 mM NaCl, and either 40 μM ATP, 0.01-0.02 (Weiss) units T4 DNA ligase at 0° C. (for "sticky end" ligation) or 1 mM ATP, 0.3-0.6 (Weiss) units T4 DNA ligase at 14° C. (for "blunt end" ligation). Intermolecular “sticky end” ligations are usually performed at 30-100 μg/ml total DNA concentrations (5-100 nM total end concentration). AAV vectors which contain ITRs have been described in, e.g., U.S. Pat. No. 5,139,941. In particular, several AAV vectors are described therein which are available from the American Type Culture Collection (“ATCC”) under Accession Numbers 53222, 53223, 53224, 53225 and 53226.

The adenoadenovirus vector preferentially packages a full-length genome, i.e., one that is approximately the same size as the native genome, and is not too big or too small. Many target nucleic acid sequences, or expression cassettes encoding target nucleic acid sequences, are very small. To avoid packaging of fragmented genomes, the present inventors designed and tested a nucleic acid sequence when linked to an expression cassette, resulted in a genome whose size was near-normal in length between the ITRs. The starting sequence was of mammalian origin, but was significantly modified to ensure that this “filler component” (also called a “stuffer sequence”) was devoid of enhancers, promoters, splicing regulators, noncoding RNAs or antisense sequences, among others things. In other words, the stuffer sequences are “silent” and confer no activity to the expression cassette.

In the present invention, suitable DNA molecules for use in AAV vectors will include, for example, a stuffer sequence and an expression cassette encoding a siRNA molecule of the invention. Many expression cassettes are very small, for example, those expressing inhibitory RNAs (siRNAs and shRNAs). Thus, there is a need to add sequences to the cassette such that it makes up a full-length or near full-length AAV genome. If only the small genome was used in the AAV production, the recombinant viruses would be heterogeneous and contain various size genomes. This is because the virus likes to package full length genomes so it will pick up other DNA fragments to fill that space. The stuffer cannot be too big, as AAV genomes above 105% of the wild-type genome size will generally not be packaged.

In certain embodiments, the present invention provides an adenoadenovirus (AAV) stuffer component (also called a “stuffer sequence”) comprising a nucleic acid of between 3300 and 4200 nucleotides in length having at least 90% identity to SEQ ID NO:1 or SEQ ID NO:2.
-continued

TAGATAACTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AGCGTGAACTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

GCTGCAACTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TCCTGAACTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

GTCTCAACTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

CATTGAACTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGCTCAACTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

CATTGAACTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGCTCAACTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TAGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

AAGCTGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT

TGGCTGCACTGGAATGCTGAAGCTCACTCGAACTCGTCAACTGAACTCTACGTTTCTGACATCTTGCTACCTCTCTACGATGCTTCT
In certain embodiments, the present invention provides an adeno-associated virus (AAV) filler component consisting of a nucleic acid of between 3300 and 4200 nucleotides in length having at least 90% identity to SEQ ID NO: 1 or SEQ ID NO: 2. In certain embodiments, the filler component consists of at least 90% identity with SEQ ID NO: 1 or SEQ ID NO: 2. In certain embodiments, the filler component has 95% identity, 98% identity, 99% identity, or even 100% identity with SEQ ID NO: 1 or SEQ ID NO: 2. In certain embodiments, the filler component has a length of about 3500-4000 nucleotides, or of about 3700-3850 nucleotides. In the present invention, the filler component is “silent” in terms of biological activity, in that it is devoid of enhancers, promoters, splicing regulators, noncoding RNAs, antisense sequences, or coding sequences.

The term “nucleic acid” refers to deoxyribonucleic acid (DNA) or ribonucleic acid (RNA) and polymers thereof in either single- or double-stranded form, consisting of monomers (nucleotides) containing a sugar, phosphate and a base that is either a purine or pyrimidine. Unless specifically limited, the term encompasses nucleic acids containing known analogs of natural nucleotides that have similar binding properties as the reference nucleic acid and are metabolized in a manner similar to naturally occurring nucleotides. Unless otherwise indicated, a particular nucleic acid sequence also encompasses conservatively modified variants thereof (e.g., degenerate codon substitutions) and complementary sequences, as well as the sequence explicitly indicated. Specifically, degenerate codon substitutions may be achieved by generating sequences in which the third position of one or more selected (or all) codons is substituted with mixed-base and/or deoxyinosine residues. A “nucleic acid fragment” is a portion of a given nucleic acid molecule.

A “nucleotide sequence” is a polymer of DNA or RNA that can be single-stranded or double-stranded, optionally containing synthetic, non-natural or altered nucleotide bases capable of incorporation into DNA or RNA polymers. The terms “nucleic acid,” “nucleic acid molecule,” “nucleic acid fragment,” “nucleic acid sequence or segment,” or “polynucleotide” are used interchangeably and may also be used interchangeably with gene, cDNA, DNA and RNA encoded by a gene.

The invention encompasses isolated or substantially purified nucleic acid compositions. In the context of the present invention, an “isolated” or “purified” DNA molecule or RNA molecule is a DNA molecule or RNA molecule that exists apart from its native environment and is therefore not a product of nature. An isolated DNA molecule or RNA molecule may exist in a purified form or may exist in a non-native environment such as, for example, a transgenic host cell. For example, an “isolated” or “purified” nucleic acid molecule or biologically active portion thereof, is
substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized. In one embodiment, an “isolated” nucleic acid is free of sequences that naturally flank the nucleic acid (i.e., sequences located at the 5’ and 3’ ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0.5 kb, or 0.1 kb of nucleotide sequences that naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Fragments and variants of the disclosed nucleotide sequences are also encompassed by the present invention. By “fragment” or “portion” is meant a full length or less than full length of the nucleotide sequence.

“Naturally occurring,” “native,” or “wild-type” is used to describe an object that can be found in nature as distinct from being artificially produced. For example, a protein or nucleic acid sequence present in an organism (including a virus), which can be isolated from a source in nature and that has not been intentionally modified by a person in the laboratory, is naturally occurring.

“Genome” refers to the complete genetic material of an organism.

A “vector” is defined to include, inter alia, any viral vector, as well as any plasmid, cosmid, phage or binary vector in double or single stranded linear or circular form that may or may not be self-transmissible or mobilizable, and that can transform prokaryotic or eukaryotic host.

AAV ITRs

An “AAV virus” or “AAV viral particle” refers to a viral particle composed of at least one AAV capsid protein (preferably all of the capsid proteins of a wild-type AAV) and an encapsidated polynucleotide. If the particle comprises heterologous polynucleotide (i.e., a polynucleotide other than a wild-type AAV genome such as a transgene to be delivered to a mammalian cell), it is typically referred to as “AAV”.

In one embodiment, the AAV expression vectors are constructed using known techniques to at least provide as operatively linked components in the direction of transcription, control elements including a transcriptional initiation region, the DNA of interest and a transcriptional termination region. The control elements are selected to be functional in a mammalian cell. The resulting construct which contains the operatively linked components is flanked (5’ and 3’) with functional AAV ITR sequences.

By “adeno-associated virus inverted terminal repeats” or “AAV ITRs” is meant the art-recognized regions found at each end of the AAV genome which function together to cis as origins of DNA replication and as packaging signals for the virus. AAV ITRs, together with the AAV rep coding region, provide for the efficient excision from plasmids expressing them.

The nucleotide sequences of AAV ITR regions are known. As used herein, an “AAV ITR” need not have the wild-type nucleotide sequence depicted, but may be altered, e.g., by the insertion, deletion or substitution of nucleotides. Additionally, the AAV ITR may be derived from any of several AAV serotypes, including without limitation, AAV1, AAV2, AAV3, AAV4, AAV5, AAV7, etc. Furthermore, 5’ and 3’ ITRs which flank a selected nucleotide sequence in an AAV vector need not necessarily be identical or derived from the same AAV serotype or isolate, so long as they function as intended, i.e., to allow for excision and rescue of the sequence of interest from a vector, and to package the desired genome into the AAV virion.

In one embodiment, AAV ITRs can be derived from any of several AAV serotypes, including without limitation, AAV1, AAV2, AAV3, AAV4, AAV5, AAV7, etc. Furthermore, 5’ and 3’ ITRs which flank a selected nucleotide sequence in an AAV expression vector need not necessarily be identical or derived from the same AAV serotype or isolate, so long as they function as intended, i.e., to allow for excision and rescue of the sequence of interest from a vector, and to package the desired genome into the AAV virion.

In certain embodiments, the present invention provides an adeno-associated virus (AAV) vector comprising the filler component as described above operably linked to an expression cassette. In certain embodiments, the expression cassette comprises a promoter. In certain embodiments, the promoter is a pol III promoter. In certain embodiments, the promoter is a mTS3 promoter. In certain embodiments, the AAV vector further comprising a target sequence. In certain embodiments, the target sequence is an RNAi molecule.

“Expression cassette” as used herein means a nucleic acid sequence capable of directing expression of a particular nucleotide sequence in an appropriate host cell, which may include a promoter operably linked to the nucleotide sequence of interest that may be operably linked to termination signals. The coding region usually codes for a functional RNA of interest, for example an RNAi molecule. The expression cassette including the nucleotide sequence of interest may be chimeric. The expression cassette may also be one that is naturally occurring but has been obtained in a recombinant form useful for heterologous expression.

Double-stranded RNA (dsRNA) can induce sequence-specific posttranscriptional gene silencing in many organisms by a process known as RNA interference (RNAi). RNA fragments are the sequence-specific mediators of RNAi. Interference of gene expression by these RNA interference (RNAi) molecules is now recognized as a naturally occurring strategy for silencing genes in the cells of many organisms.

Certain embodiments of the present invention provide a vector that encodes an isolated RNAi molecule. As used herein the term “encoded by” is used in a broad sense, similar to the term “comprising” in patent terminology. RNAi molecules include siRNAs, shRNAs and other small RNAs that can or are capable of modulating the expression of a target gene, for example via RNA interference. Such small RNAs include without limitation, shRNAs and miRNAs (miRNAs).

“Operably-linked” refers to the association of nucleic acid sequences on single nucleic acid fragment so that the function of one of the sequences is affected by another. For example, a regulatory DNA sequence is said to be “operably linked to” or “associated with” a DNA sequence that codes for an RNA or a polypeptide if the two sequences are situated such that the regulatory DNA sequence affects expression of the coding DNA sequence (i.e., that the coding sequence or functional RNA is under the transcriptional control of the promoter). Coding sequences can be operably linked to regulatory sequences in sense or antisense orientation.

Operably linked nucleic acids are nucleic acids placed in a functional relationship with another nucleic acid sequence. For example, a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate trans-
lation. Generally, operably linked DNA sequences are DNA sequences that are linked are contiguous. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accord with conventional practice.

The invention will now be illustrated by the following non-limiting Examples.

Example 1

A plasmid FBAAVmU6miHDS1 stuffer was generated that included AAV2 ITRs, mU6 promoter, miHDS1 target sequence, filler component stuffer, and an AAV backbone (FIG. 1). The sequence for 5pFBAAVmU6miHDS1 AAVstuffer is provided in FIG. 2, and the sequences for the individual components of the plasmid are provided in FIG. 3. The full-length filler component ("stuffer sequence") consisted of 3776 nucleotides.

Example 2

The in vivo silencing efficiency of a vectors expressing miHDS1 was compared. Four vectors were constructed: (1) a vector expressing a control sequence (miSAFE) and containing a control sequence (eGFP), (2) a vector expressing the target sequence (miHDS1) and containing a control sequence (eGFP), (3) a vector expressing a control sequence (miSAFE) and containing the stuffer sequence described in Example 1, and (4) a vector expressing the target sequence (miHDS1) and containing the stuffer sequence described in Example 1.

(1) AAV2/1 mU6miSAFE-eGFP (4.81E12 µg/ml)
(2) AAV2/1 mU6miHDS1-eGFP (4.81E12 µg/ml)
(3) AAV2/1 mU6miSAFE-stuffer (4.81E12 µg/ml)
(4) AAV2/1 mU6miHDS1-stuffer (4.81E12 µg/ml)

The sequences for miSAFE and miHDS1 have been previously discussed (see, PCT/US2012/024904, which is hereby incorporated by reference herein in its entirety). Wild type mice were injected in the striatum with the four vectors. Mice were sacrificed one month later and htt expression was determined relative to Actb expression levels by QPCR. FIG. 4 shows that there was a 20% decrease in expression between the miSAFE/eGFP and the miHDS1/eGFP expression cassettes, whereas there was a 60% decrease in expression between the miSAFE/stuffer and the miHDS1/stuffer expression cassettes, i.e., a 60% decrease in expression when the stuffer was used.

Example 3

A plasmid 5pFBAAVmU6miHDS1 stuffer was generated that included AAV2 ITRs, mU6 promoter, miHDS1 target sequence, filler component stuffer, and an AAV backbone (FIG. 5). The sequence for the plasmid 5pFBAAVmU6miHDS1AAV-stuffer is provided in FIG. 6. The sequence for the stuffer (Stuffer #2) is provided in FIG. 7.

Example 4

One of the considerations with AAV packaging is maintaining optimal genome size. When this occurs, the ratio of virions that form which are lacking genomes are minimized. Experiments were performed testing the packaging efficiency of the new stuffer sequences and found high efficiency packaging. For example, see Table 1 “Average empty” and FIG. 8. It was also measured if genetic material that was packaged contained non-mRNA:intron stuffer sequences. It was found that the incorporation of unintended genomic material used in virus production was extremely low (Cap/rAAV, Amp/rAAV, Gent/rAAV). Finally, the quality of the viruses were analyzed by Silver Stain after polyacrylamide gel electrophoresis and found to contain the appropriate proportions of the various capsid proteins (VP1, VP2, and VP3; FIG. 9). In summary, the intron I/H stuffer sequence allows optimal packaging of desired transgenes into AAV capsids.

All publications, patents and patent applications are incorporated herein by reference. While in the foregoing specification this invention has been described in relation to certain preferred embodiments thereof, and many details have been set forth for purposes of illustration, it will be apparent to those skilled in the art that the invention is susceptible to additional embodiments and that certain of the details described herein may be varied considerably without departing from the basic principles of the invention.

The use of the terms “a” and “an” and “the” and similar referents in the context of describing the invention are to be construed to cover both the singular and the plural, unless otherwise indicated herein or clearly contradicted by context. The terms “comprising,” “having,” “including,” and “containing” are to be construed as open-ended terms (i.e., meaning “including, but not limited to”) unless otherwise noted. Recitation of ranges of values herein are merely intended to serve as a shorthand method of referring individually to each separate value falling within the range, unless otherwise indicated herein, and each separate value is incorporated into the specification as if it were individually recited herein. All methods described herein can be performed in any suitable order unless otherwise indicated herein or otherwise clearly contradicted by context. The use of any and all examples, or exemplary language (e.g., “such as”) provided herein, is intended merely to better illuminate the invention and does not pose a limitation on the scope of the invention unless otherwise claimed. No language in the specification should be construed as indicating any nonclaimed element as essential to the practice of the invention.

Embodiments of this invention are described herein, including the best mode known to the inventors for carrying out the invention. Variations of those embodiments may become apparent to those of ordinary skill in the art upon reading the foregoing description. The inventors expect skilled artisans to employ such variations as appropriate, and the inventors intend for the invention to be practiced otherwise than as specifically described herein. Accordingly, this invention includes all modifications and equivalents of the subject matter recited in the claims appended hereto as permitted by applicable law. Moreover, any combination of the above-described elements in all possible variations thereof is encompassed by the invention unless otherwise indicated herein or otherwise clearly contradicted by context.
SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 12
<210> SEQ ID NO 1
<211> LENGTH: 3774
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 1

gaatcgggc tatcccaagtt gtccttggtt cattgcaaat gggacgtaaa gagggccagag 60
agaaatagaa cagaaactgt tctaatattg gtcatttaaat gttgtaagttat gttgccctt 120
taaacctccct tccatcttttt tccagggatt ggtgcaccaca gggcgctttg gttgctcctg 180
ggattgtgag gagtgccccc aggtgtgctg tttggtcttt aggtgtcttt cttgctgggt 240
tctgccttt ttcttctcttc ttctgccagc attttcctct cattttaag 300
tttggtgag cagggtggtt gttcgtccag attttgcttt cttcctagtc gtgactgtgc 360
agagccacgag ggctgtcagcc ggattccagaa tttggtcagtt tttggtcagtt cttggtcagga 420
gtattgaggagt ccctgggtg ccacgctgg actcgtgctg gcttgctgtg ccctggctga 480
gtggctcacttg tcctctggttc tcttactcttg ttctctctatt catttactttg tttgctttgt 540
gttacaaaact gttacctctgt acgcccaagga aagggccaca ttttggctgtg ttttggctttgta 600
agagacattgaa tttggtcagtt gttggttttt tttggttatttt cttctgctg gaactggtgttg 660
tccaggggtg aagcagcctgt gttccagtggac actcctctag cattttgctttg 720
tttggtgggt tttggcttctttt tttggttcagtt tttggttatttt tttggttcagtt tttggttatttt 780
agagcagaaag gttctgtcagtt gttcgcaccc agatggtgaa agattactgtt gggacggtcc 840
tttacacatc ccgccctcttg ttctcgaggg cttccgagga ctgctggggc 900
tatggggct acagagtttg aagacagtgg attggctttttt attggagtattt ctttataggaga 960
tggtagacgg gttggtcagtt gttgcgggggg ggggtgtatttt cttctataaa 1020
gacttgagcc ggctgtggatt ggggcgggtt attggacagag gggctggggg cagtttttgag 1080
tttggtcaggg gttggtcagtt gttggtcagtt gttggtcagtt gttggtcagtt gttggtcagtt 1140
agcagcagatg cttcgctcctt gttgcacgctttt gttggtcagtt gttggtcagtt gttggtcagtt 1200
gctcctcatc agagctgggct gttggtcagtt gttggtcagtt gttggtcagtt gttggtcagtt 1260
gtgggggttt ttattcagtt otatttcagtt tttattttttttaa ctttttatttt 1320
aggtgctattc cagtttggatt cttgggctattt cttttattttttaa ctttttatttt 1380
ggtggtcagtt cttgggctattt cttttattttttaa ctttttatttt 1440
aggtgctattc cagtttggatt cttgggctattt cttttattttttaa ctttttatttt 1500
aggtgctattc cagtttggatt cttgggctattt cttttattttttaa ctttttatttt 1560
aggtgctattc cagtttggatt cttgggctattt cttttattttttaa ctttttatttt 1620
aggtgctattc cagtttggatt cttgggctattt cttttattttttaa ctttttatttt 1680
aggtgctattc cagtttggatt cttgggctattt cttttattttttaa ctttttatttt 1740
aggtgctattc cagtttggatt cttgggctattt cttttattttttaa ctttttatttt 1800
aggtgctattc cagtttggatt cttgggctattt cttttattttttaa ctttttatttt 1860
aggtgctattc cagtttggatt cttgggctattt cttttattttttaa ctttttatttt 1920
aggtgctattc cagtttggatt cttgggctattt cttttattttttaa ctttttatttt 1980
tgattttggg atgtgacaa tagaactcaac ttctactgtgt agttgattt aggggaacct
2040
tatacctcg atgtgtaagc aacctgtcga gatgtgagga tgctttctcta ttgttctaga
2100
 accttttaaa ttacccgaa agcacatgaa atatatagaa ttatatataa aactttaaaga
2160
ttatctca tagctctcgc acatcttcg ttaaatgag tatccagttg tccctcttt
2220
ttcttactt ttagtctcttt ttttctccct gaaatctctt tataaaaggt
2280
tagttgctta ttctccagag tgtgctgtctt aatctttgaa gatagtaaag tctctctatt
2340
ttacaagtt ttagttctcttt ttagttctga tgtgttcact gtcacgccca aaaccgaaag
2400
aggggatggt gaaaaatgtcc gtagagaa aactctttt tcaactctata
2460
tcaatcctga atttccaggea cttttccatg gctctctgctg tttatatttt tttatttttt
2520
tttcctctgtg ataccgatga ccttccacatc tctctcaatg ccttgccctc
2580
 ttagaactct gttctcttttg cttctctcctt cccgctcctgtg cttctctgctc ggagctttt
2640
cagctgtaact ggtgtgtagg tgcctctgctg gctctgtagt attctcatgg ctgcctgttt
2700
gttctcttg tttctctctgt ttagttcctctt aagttctctca ctctcccttttttc tgtggtggcc
2760
tctctcagc tctctctcttt tctttctgtt tgtgctttttt gttctctgctc
2820
tctccagac cttttcagcgt cttttctaggct cttttctgcgct atttttcttttt cttttctaggct
2880
ttaacttgctt ttttttttt cccgctccttt ccccttttt cttttcaatagct ttttttttt
2940
gaatcctag cccctcttttt tcttccccctt aacgtgtgctt gcagacgggt gcaagatgctt
3000
tgtcctttg ggcagccaga cttcccccttt ggtctcttcg cccctttcttttt tcccatcaac
3060
tcactctgct tccttcagcct tccttcttgct tccttcttgct tccttcttgct tccttcttgct
3120
cctctctgg cccctctagc cttttttctgt gactttgctt ttaacctgctt tttttctttt
3180
agttctctgt tctttctcttt gctctcttttt ctttccccctt gcagacgggt ttttttttttt
3240
aggggcttct cagattaagtt tctttctccttt cttttcttttttt cttttcttttttt cttttcttttttt
3300
attttttttt ttttttttttt ctttttttttttt ctttttttttttt ctttttttttttt ctttttttttttt
3360
ccattctcct caagttctcctt aatgtgacta aatatctttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttttt
-continued

tctctttccc atgtctcttt ctctgcttcc acgccatttc cccccatgc taagttgggt 300

gcagcaggtt tctgctttct ctagatctcc gttctcccag atgctctgtg tgcagccccc 360

tagcggtgg aggcctgggt cagcagctgc gcagatgtgg cagatgtgg 420

gcagcaggtt gcagctggag ctgacccccc cagatgtgg cagatggcttgc 480

gtacccctta atgtctgctgg atgctttctt cagagatgtg ctgtgttgg tgcagctgaa 540

tactgtcttt actatccttc cagaaaaatgg cacccctttg atgtgctttt ctagaagaaaa 600

tgcacgccat gcttgcacag cggccttcag tctgtgttta ttttgccttc 660

gttgcagag atcacaacaa cggccttcag tctgtgttta ttttgccttc 720

tgcacgccat gcttgcacag cggccttcag tctgtgttta ttttgccttc 780

tgaccccccct tttggctcgat ccggtgctgc gggtcctgcg gggctgctgc ccggtgctgc 840

tgaccccccct tttggctcgat ccggtgctgc gggtcctgcg gggctgctgc ccggtgctgc 900

tgaccccccct tttggctcgat ccggtgctgc gggtcctgcg gggctgctgc ccggtgctgc 960

ttgcacgcag ttaaagttgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1020

tcagctgtgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1080

tcagctgtgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1140

tcagctgtgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1200

tcagctgtgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1260

ttgcacgcag ttaaagttgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1320

ttgcacgcag ttaaagttgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1380

ttgcacgcag ttaaagttgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1440

ttgcacgcag ttaaagttgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1500

ttgcacgcag ttaaagttgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1560

ttgcacgcag ttaaagttgt gatgacatgc atatctattg tattttttgtt tattttttgtt 1620

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 1680

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 1740

tagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 1800

tagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 1860

tagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 1920

tagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 1980

tagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2040

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2100

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2160

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2220

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2280

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2340

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2400

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2460

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2520

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2580

ttagggggagt cggcctgcctgc ccggtgctgc ccggtgctgc ccggtgctgc ccggtgctgc 2640
gaagtgcctt tggagtcct ggcgtctct gctggatctt tgaatttgc ctcctggttc
cttggtcctt tctcgtcagt tgctcagcgt tccactccc cattctctgt gtggcccttc
ctgcacctct ctgatccttt ttgctcttgc tgggtttttg ttttggttto gaagctccac
agaacttcttg cacgctctctt gaagactcgag gacttttctt aagctttaaact tcatcctag
acccattttt ccccttcttg gacagcgaat gctctcctg gactcctctt tccggaatt
ccacgcttctt tttcctccc cacttaccttg tttgccagga gaggccagat tggctgcgt
ctcgagctgc aaccccttcct tctctggtgct cttcacacta catgacatac ccacacarc
ctgctctcttg gaacctcagt ggtgatctgt ctgattttt ctctctttttgc tggccctta
gggccacacc aggtaatca gctatagcat gctatttga ttcacactaca ttttcagttt
cctgccttg ctcctcgtt cttcgtacct tccggcagaga aagcttcttt ctcacaacgg
ctttgcagtt tgtcctcact gacgccatct tttctgaggg cagttttttta caaatattta
ctttctagct cagctgctgc ttaacctctct tgtttactgt gtttttgcg tgcaccacat
ccttgagctt gtaaaaacta atctgctatt aatctatta catgaaactcc ttgattactc
taaacactat tcaacactgt ttaattatct attgtatcga aattgaacacc acacatattg
aatttgacagt tgtatttcct cagggactct agttgataag gaatgagacct tagttatttc
ttttttttga tataacactat acacagattc agtatgttgc gacatataac ctttttttttt
cattatgaa gaaacttgaag aattactcct tccatgtagc ccagrttacac ttggaagctt
ttgaaagctg agttttggtg taattcatcg aaggtacttt ttttcacaaata tttttttca
ctgcgcacac ccacacaccc gcaaggtgag tgtacggtgc
ttcggctctt toocctcttt ttcgccacaag ttgccogget ttcoccgctca agctotaat 60
cgggggtcct ctttggggtt cgggtattcg gttttacgc acctgaaccc caaacaacct 120
gataggggtc ctggctccag tattgggcct aagcccttg cttccctttc
agttgggtact cccggctattt gttttcacta acctggaaca aatcatccac
ctcttcgtc ctctctcttt gataataaag gatctggccc gatttccgcc tattgtttaa 300
aaactgactt gatattcaaca aatatttaac cgtatattca aaaaatattta aagcttacaa
ctttttgtgg acctttttgg gaaatctgggc ggagacctct attttttatt tttttttaaat 420
acatccaatt atgtacgc acatatgaga atctcactga taaacactgt aataattag 490
aaaaagagag atagtagata ttcacactt cctgtgctgc cttacctcct ttttgccgac
atccttcctt ctttggggtt ctacccaccc aacgctttgt aagtaaaag atgtgaag 600
tccgcttggg ccttgcacaggt actctgcctc acagcctgta agatctttttg 660
gagttgggtct cccccacaga gtttccacta gatgagcact ttaattatct tggagtatgg 720
cgggctatc ctcgctattg acgctggcct acagcactt acgctgcgtc 790
tcactagaga tcttggtgat acctcactca cacagaaaaa ctcctctcag atggtatcag 840
agtaagagaa ttcagcgtct cttgcaataac catagctgat aaactgctgc ccacactact 900
tctgacaacg atcgaggac ecgaaggagct aacogcttttt tgtcacaaca tgggggatca 960
tgtactcgc tctgatcgtt gggaacccga gcgtgaatgaa gcctaccaac acggacagcg 1020
tgacacaaac atcgctagtc caatgcaac aaoctgtgcc aactatataa ctgacgaact 1080
acctacttta gcttcgcggc aacataaat agacctgcatg aggccggata aaggtggcag 1140
acactctcgg ctgctgggcc tcctgcgcgttg ctggttcattg tgtgataat gcctgaggc 1200
tgagcgtggt ggctggcgtta tcctgggagc acttggcgcctc acctggtcag 1260
cgtagtttac tacacgagc ggagtccagc actaataag gaaacaaata gacagtcgc 1320
tgagatagtt gctctacaac gttaagcaatg tgaactgcga gaccaagttg aacatataat 1380
actttagatt gatttaaaaatttcaattttttaa attttaaaaggt atctaggtag aagaccttttt 1440
tgataatcct atgacccaaaat ctccttttaacg tgaatgttgct ctcctcgtag cytcgagccc 1500
cgtagaanaa atcaasagat cttctttgagac ctctcttttt gctgcgttaaa ctcctcgttt 1560
gcggaaaacta aacacagctg taccagcggt ggtgtgtttc cgggttgcag aagctacaac 1620
tcttttcggc aagaattactg gcttcagcac gcggcagata ccacactcttg ctctctattg 1680
ctgagcgtggt ttgactgcacc acttcaagag atcctgtgctc gctcctaatg acctgtcctt 1740
gctaatcctt ttaacggcag cttgctgcag cgggcattag cttgtgcctta cgggttgtagg 1800
tcaagacgca ttagatgcag atagggccac ggcggcggcc tgggaagggg gttgctgaca 1860
acacccggac ttcggagccac caacatcagc cgaactagga taactcagcgc gtgagctatg 1920
aggaagacgc aagcttcggc aagggagaaag ggagaggccag ttcgctgtaaaa cgggaggttg 1980
cggaacagga ggcggcagcgc ggagggcctcc aagggggaaac gcgtggattac tttatagcct 2040
tgtgccgggtt gcgtcctgtct gcgtatatgg cttgctgtcgtg cagggggggc 2100
agcgtctatgg gaaaaagccac gcacacggcgc cttttttacaag ctctcgcgct gccgctttgc 2160
cttctgtccc atgctctttc ctgctgtctact gctgtaacct ctggttcatgc 2220
ccttgagtga gcgtgtcagc cttgcgctcag cgaacgaccgc gacagccagc agtcagtggag 2280
cggagagcc gcagaagccct tggattgctga tttttctctctt acgtctcttg ctggtattcc 2340
acacccggata gcacagcgcgc ctaactcggc aaaaagttgtagcg aatattggtttt gagaaggttt 2400
tgctgcctgct aaggcgggtgc gggccgcaaa taagactttaa aactgaaacc aatatgattc 2460
aacttacgata atatacgctt aaaaacttctag aactgaaacta cagtcgatgt 2520
atgtctgtaaa aacgcaattgt ggaacttttct ttaggtctaaa cgaacactttg cattttctga 2580
agtggaaatt gcgcgcgtga ttaaagaaggg gcgtggcaccct gggcatggta aaagactatag 2640
tggcgagcttt tgcacaattt acggacaccc aacctgggcggc gggaagccagcct ctcgggtgta 2700
acgaaatttttg cggcgccggtt gcgcgcgctgc atcaataaag gcgtctactctt cttccgctag 2760
cccaacctcgg tatacagagc caatacgccgg tgcgtacaagc aactgttcttc cacgtagatc 2820
actacaagcc ctcagcgcgtg ggcgttgctg cttggggcct gcctgtgggct gcgtggcattg 2880
ctgccgcttc gcgtgcgtgcc ggactacgag aagcataaga tatgatctct actacagccgc 2940
tgctcaacgc tggcagcagc gtaacccgag agccgcaaca cccacoctct ttcgggcgaag 3000
gcagcacaagc cgatgcaattct cttactacgg agacagttcc accagataac gcgagcctggc 3060
tgatgtgggg agttgtttggc cagcccttcgg aacacaagcata aagagcgacg 3120
cagcatggatt tcagtctggctc agggcgcggc ctactagcgt gcacatgctgc catacttggag 3180
ccaccaac tatgtttttag gcgtggccgc ttcgctgtaa atcctgtgtg cgtgctatgc 3240
atcgtttgctg  ttcataaca  tcaaacatcg  acccaagggg  taacgcgtct  gtgcgttgga  
3300
tgcccaggg  attagctgta  caaaaaaca  gtcataacaa  gcatgaaac  cgcacatgc  
3360
gccgttaccc  cgcgttggtt  ccggtcaacg  ttcgctggtc  gcatagccta  
3420
ctgccctac  actcctagaa  cccagacgagc  ttacctcaat  tgcgtcgctg  ccttcctcgc  
3480	tttcaaggt  tgcgttcacc  cggggaccttt  gggcagcagc  gaaagtgagc  cattttgctc  
3540
tgcgttgccg  aacgcacgccc  agtttctcctg  cctccagcct  cgctcagcct  tgcgggccttt  
3600
gctgctttcc  tcaacgaagg  tgcgtgtcagc  ggtacctgcc  tgcctccagg  aagctgggaa  
3660
aatcggccgc  tgcggtgctg  tgcagcgcggc  gcggcagccc  gatcaaggcc  aatcctgctct  
3720
cgattttgtg  gaaagcgagc  atcgtttggt  ggcctagggc  ttgcagcatt  gtttgtgtg  
3780
ttgccatcag  tgtctgctcc  tgcagcgcggc  tgcggtgctg  tgcagcgcggc  ggcctagggc  
3840
gggcgtggcg  gacacctctgg  tcccccgccc  tcaagaagcgc  agcagtgagc  cagagagagga  
3900
tgcgcccact  ccaatatcctg  ttgctaatg  tttaatcctg  tcatccctct  
3960	ttaagctgagc  cagatactgg  tggctcttggc  gcgtctgctc  ggcgtataat  gcgatatctt  
4020
tccgagagc  tgtggagagc  agctcgtgta  ccttcsgcgc  cgggttaatc  cgcgtatataat  
4080
atctcctagt  aatcataagc  gttataattg  cgctataggg  tgcataatac  ttctttttttta  
4140
atataacttc  catctatgatc  gatagcgttg  gttcttacta  agagatcata  atataatct  
4200
attatattta  aaaaagccag  aaaaaaggaac  tcaacccatac  tcttaatctt  tttctgtgtt  
4260
tgatgactaca  aatatcctcc  ggaggagaac  ttgctgagcg  tttataatcct  cgcagatggtg  
4320	taacgtttata  actcagctgta  gcagcttgctg  tctcagctgtg  cagatacttgta  aagccagaca  
4380
tggttgctca  cctcgccgtgac  gaccagcgc  tctagatcgcg  gcgcacaga  cgggggagatc  
4440
cagactatgc  aagatactct  ttttaatctc  gggtctatcc  agggtgccct  gttctcctgac  
4500
aaaataggccc  tataagggcc  aagagagaata  tgcataatac  tttctttt  attggtcatt  
4560
aatattgtaa  gttattttttt  ttttataactc  tttctctattt  ttttttccag  aatagcgtgga  
4620
ccagcttggc  ttcagcagctg  tctgagcgtg  tggcagcttg  ttgctgggtg  ccccttgttg  tttgctttttc  
4680	ttcctggtcc  tttctctagg  tgcggcctggc  ctattctcgct  cttgtaggttg  ttgctgttttc  
4740
acgcctttctc  ccattcagctt  tgcagcgttgg  gcgcacaggg  ttgcctgtcgc  tcaagattccc  
4800
gctcctcagc  atctctgtgg  tgcagcggc  agggcgggtg  cagcggctagt  cgggtcctgc  
4860	tagctgtgtc  ttcacagtcg  cagaggtgag  ggggcccggg  tggcaggctg  gcgcagctgg  
4920
gccagctgta  gggagcagatg  gcagtgctgc  gctgctgctgc  gggacacgtg  ttgccccgag  
4980
catgctgtag  gttgctgctg  tattggccaa  actagcttct  actctggcaca  aagaataggg  
5040
ccaccttttc  atctgacaggct  tgccacgcgg  gtcgcacgtgc  gttcgctgtgg  gttctgggaaa  
5100
atcgtttctcc  cagcttaggtg  atctctggcgc  gctgctgctgc  gttgctgctgc  gtcgccacctc  
5160
atagagggcc  ttaaagctcc  gtgtctgtgg  gttcgccagc  atcgcataac  ggcacgctgcc  
5220
gactcagtgag  tttttggtcct  gggagcggag  aagatcttcc  tggagtctgc  gcgcagctgg  
5280
tgagagagag  gttggtgacc  cttcattctc  atcagcgcctc  ttgtctccaa  gggggtgccg  
5340
agagcgttgtg  aagctgcgaccc  cattaatgctg  ttgctgtcagc  gcggccacca  cggggatcctg  
5400
gctagagggtg  tgtttacttg  ttgagagcata  ataggggcct  tttctgaattc  ttgattgtgg  
5460
tcaggagagag  ttaaagctct  ggagagctttc  gggagcggag  aagatcttcc  tggagtctgc  
5520
ggggaggaggg  aagctcgagtt  gaggagtaga  gagcgcagcag  gttggtgacc  cttcattctc  
5580
ggagaggctggc  aagctcgagtt  gaggagtaga  gagcgcagcag  gttggtgacc  cttcattctc  
5640
gagggagaag gccagcgacg tgtgctccatc acctccagcag tcggctgtagat tgtgagagac 5700
catcctgcc tgtgaggg ggtctagtgt tcgttttctt tgtttataca ataagccttg 5760
tattgttta caaacaatc gtaaacatca atcaaggttt gataagcttt ctatttttat 5820	ttttaaagat aagtttgtag ccaattgcct tgcgtcattt aaggaccttc 5880
agtacaaacc gcaccaacac cagtaagctt taaagtttctt gtagagatgt ttaatctctg 5940
gaaattccca gaaattgctcc tgttattttat ttttcaacgt gttgacccaa tatttgtaacc 6000
cacaagtttt aagcaagac ccctcatgctt cttcagcctt cttcagctgt cagttttaaa 6060
gagatcatctg acagttgtt ggccagcatt ctctcctcttt gggtggtgcaac gcatattgaac 6120
gaggaagctc atatgggtta gaaggtggtgca ccaaaagctct ccctggtgcaaa cccctcaggt 6180
tgagaagccg tttctcaagt tgtatcttcg gaagggctgcttg aatggagatctg 6240
accatcggga gttataacct cgcagctgtaga cccagagctag caagagttgg cagtatagt 6300
gcgtggagag gttggtgcgg gcagctgcttt ggtcagggcgg gcagagcagca 6360
cgggctatoc tcgagcagttg gagagtttatg tggggtgctca gttctcggctgt ttgctcaggt 6420
aatatttttg tattttctcc tgtgatgtca gcaatagagacctt catctttatc 6480
tttagatggt atatatccccatcctgcatgcttttaaagcttggatatgt 6540
GGGATGTTTT CAGCAGATT CAAACTCAA AAGGATGGTT CACTGACTT CCAGTCATTT 6600
GAAATACAA AAAAAGCTCA AAAGATTTTT CAGAGCTTGC TCCCCACTT TTTGATTAT 6660
GAAATACAA AAAAAGCTCA AAAGATTTTT CAGAGCTTGC TCCCCACTT TTTGATTAT 6720
CCCTAAACCT CATGAGTTCG CTTATTTTAT ACCAGTGTTT GTAATTTTAT 6780
GGGAGGATT CAAACTCAA AAGGATGGTT CTTTATTTAT ACCAGTGTTT GTAATTTTAT 6840
TCAATGCAGC TCCCCAAAGC TTAGCATACG TTTTGTTTAT GATGTTTAT GCTATTTAT 6900
TTTCTACTG GTTTTCACG CATACTACT CTGAACTCG TGGCCCCCT CACTGTTCTT 6960
GTCCTGCTAT GATATTATC TTATTTTCTTC ACTTTTCTTC ACTTTTCTTC ACTTTTCTTC 7020
TGCTCTGCTA CTAGATGGT CTAAGGCTCC AGGAGGATGG TATAGGTACG TATAGGTACG 7080
GCTACGCTGT GAGGTCAGC TTTACGCTG AAGTGTGCTG TTCTTGCTG TTCTTGCTG 7140
CTGACTGCTG TAACTCAGT CTGTGCTGCTG TCTGACAGT CTGTGCTGCTG TCTGACAGT 7200
TGACTGCTG TAACTCAGT CTGTGCTGCTG TCTGACAGT CTGTGCTGCTG TCTGACAGT 7260
GTTTATCTTG TTTTTTCTTTG AGGTCCACCA GAACTTTTAC AGGTCCACCA GAACTTTTAC 7320
AGGTCCACCA TTTTTTCTTTG AGGTCCACCA GAACTTTTAC AGGTCCACCA GAACTTTTAC 7380
TCTTACTCACT CTTCTGCTCT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT 7440
TCTTACTCACT CTTCTGCTCT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT 7500
TTTACTCACT CTTCTGCTCT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT 7560
TTTACTCACT CTTCTGCTCT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT 7620
GTTTATCTTG TTTTTTCTTTG AGGTCCACCA GAACTTTTAC AGGTCCACCA GAACTTTTAC 7680
GTTTATCTTG TTTTTTCTTTG AGGTCCACCA GAACTTTTAC AGGTCCACCA GAACTTTTAC 7740
TTTACTCACT CTTCTGCTCT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT 7800
TTTACTCACT CTTCTGCTCT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT 7860
TTTACTCACT CTTCTGCTCT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT 7920
TTTACTCACT CTTCTGCTCT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT TCTTTTCTTT 7980
<210>SEQ ID NO 4
<211>LENGTH: 86
<212>TYPE: DNA
<213>ORGANISM: Artificial Sequence
<220>FEATURE:
<223>OTHER INFORMATION: Description of Artificial Sequence: Synthetic oligonucleotide

<400>SEQUENCE: 4

ttgaatagc aatagaccct ctgatatctt ctatttggat ataccacata ccagactctg 8040
attatgagc acatttaacc cttttttctt atatatgaa aagatagga attatatctt 8100
ccagtgcc cagttaacc tgaagactt ctaaagagta gtttttatt agcatactga 8160
aagatatttc ttcccaaat attttttcag tgcggcaac aaccagcag acacacccgt 8220
caaagtggtg gtaagggcga ctggagcatg gtaagactta taagagcaat ggggctttaa 8280
tcataacct caagtagacc ctgtagtggg agtggggcag tcgcctctgct cgccctgcgt 8340
gcggtagtg ggggggggag gaaagagcgt cccacggcct gggggtcgctc gggggcgt 8400
cagagtaggca gggagaggg gctggcctgc ttggatgtaga cggctcaactg gtaaag 8460
tcataactgt aagtcagtcg agaagtaacta gagagacctata atacagccata ccacattttg 8520
agaggttta ctctgcttaa aacactccgg cacaccttcgg ctaaacctgta aacataaat 8580
gaaagctatt gttggtgtaa acgtgtttat tcgagcttat aatggttaca aataaaagca 8640
taggcaca aatattcacc aataacatttt ttttacgtct catttatggt tggatgcttc 8700
caancttc acgtatcttt atcagctgctg gatctgatac ctgatatgctc ctggagatac 8760
cgaaacgatg aagtaaagga tagttccaaa aatgatagct atattttagt ttcgtgattag 8820
cctcaacgcc tccaccaagg ctccatact tctgagcctc ttcctctaat aactctaaa 8880
aactatttt cccacccgct cagttccaaa attttttctg gcggcagcag ggggattttt 8940
tctgctgctg aagttttttta tttaacacct tgcacacctc acctgacaa aatcgagcatt 9000
cagagtaggt tttcttgcct cagatgcaaa atttcttctg ctcgctctgct ttaaatgatg 9060
tttgatagc ctgtaatcaca acgtgatttt ttgagcttarg gatggaattg 9120

<210>SEQ ID NO 5
<211>LENGTH: 311
<212>TYPE: DNA
<213>ORGANISM: Artificial Sequence
<220>FEATURE:
<223>OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400>SEQUENCE: 5

cgaagccgcc atctctaggg cgcgcgcggc cccctgcac agacgttggc gagaagctgg 60
gctctccccc tgcgcgggtt aatagcata taatattttc tagtaactat agagctttaa 120
tgtgcatata aagacagata atctctttct tttataacta gtcacatttt atctatagg 180
tttgagtttct taaaagagat acaaatatct aattattt attatatttt taaaactaa cgcacaacaag 240
aaccctaccct tacacttgaat gtaattgtgt gggttggagc tataaatatct cttggacaa 300
aagccctttg t 311

<210>SEQ ID NO 6
<211>LENGTH: 94
ctgcgtgcctc gtcggtctac tgagccgccc agggagtctgg ggaacttttg gtcggtccgtc
ctcagttgac gcggagcgc gcagagaggg agtg

<210> SEQ ID NO 7
<211> LENGTH: 128
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 7
aagggagaacc cattgaggtga gttggtcat cctctcttgc gcgggtgcgc gtcggtcagt
60
gccggcgac ccaggttgcgc ccagcgcggg gcccttgccc gggcggtctg cagtgaggag
120
cggagcgc
128

<210> SEQ ID NO 8
<211> LENGTH: 534
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 8
ttaggtggcc gcacatggtg gcctatcata atctgtccct tgtctctct ctaattcact
60
tgtatgagc gcgcagctgc gcgggtgtag gccgggtgcg cttgcctct cttctttctg
120
accaacgcgg tgggtcctac gcgggtgagc aggcttggga cgggtcgtgct cggcttctct
180
cgggtcgtg ccagagagct cgggtgccgt cgggtgcgtg cgggtgcgtg cgggtgccgt
240
cgggtgcgtg cgggtgcgtg cgggtgcgtg cgggtgcgtg cgggtgcgtg cgggtgcgtg
300
cgggtgcgtg cgggtgcgtg cgggtgcgtg cgggtgcgtg cgggtgcgtg cgggtgcgtg
360
gaggtaggg gtaggtgcat ccacgctgag acacgacgac gagctgactt ggtgcagtt
420
tttgtactg cgccgctgtt gcgggtgagc agctgcaggg agctgagcgc gcgggtgctt
480
ttttgactct gcggcttgag ccacagctgc gatctgttttc cgccggggtg cttctctctg
534

<210> SEQ ID NO 9
<211> LENGTH: 8k1
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 9
atgagtac acacacttccc tgtgctttt cttcctcttt tgttgctcct ttgggtgctct
60
gtttttgtc gccacgcaac gctgtgggaat tcggctatca gtttggtgca
120
cgggtgagc cagccttctc ctacgcaggtg acagttaagc ttgttgtagt gttcgccccc
180
gagagcttc tttatagtac gagaacttct aacagttcag ctgattcgct cggattcctc
240
cgattcagc gcgggagcag ccagctctct gcggggcttc gcacggcatc acatattctagt gagaaggtg
300
gtgagatct caacagtcac agaaagcatt cttacggagt gcagcagct aagagaatta 360
tacagtctg ccataaactg catgataaac actgagcgcct acctttctc ccacaaagtc 420
gagagacgga aggagtcatt gcttttttgt cacaacatgg gggatcatgt aactgcctt 480
gatcgctgggg aaccgcagct gtgaagcgg acatcagggc aagacgtga cccacagt 540
cctgtacaa tggcacaacc gtggcgcaaa cttataactg gcaagaact actcttagct 600
tccgcgaacc aataataag cgattagggc cgggataaag tgtcgagggc acctttgccc 660
tgcggccttc cggctgtagt gtttattgtct gataaatctg gacgcgtgga gcttgggtctt 720
cggcgatata tcgacagcact gggccagatg gtaagccctt cccgtagcgt agttatttaac 780
aagaagggag tcacagccaa ttaaggtgaa cgaatagac agatgcgct aataggtgcc 840
tcaagttta agactggtga a 861

<210> SEQ ID NO 10
<211> LENGTH: 225
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 10
tgaggggcga caataagaatg ttaaacagat ccaaatagat ctaaactatg acaataaagt 60
cttaaatag acaagatagt tgaaaactga aatcagcgcct gttargctgt gaaaaagcat 120
aactgagtatt tgtatattct aaseaaactt cttacctttt tgaagttgcaaa attgcccggc 180
gtattaaga gggggcgtggc caagggcatg gtaagaacata tattc 225

<210> SEQ ID NO 11
<211> LENGTH: 166
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 11
aaccagataa gtgaatctaa gttoacaact attttgctat ttttaatatt tgtattagct 60
tagaaccga caccagctac ccaatatttt tgtcaactct ccccaatatt tcctaaaaa 120
ttcatacttc aaccctcaca gttoacaact attttgaccgg ccaca 166

<210> SEQ ID NO 12
<211> LENGTH: 9111
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Description of Artificial Sequence: Synthetic polynucleotide

<400> SEQUENCE: 12
tctgtccctt ttccctcttcctttgcacagt tccccgcacct tccccgcaccttcgctt accc 60
cggggccctt cttccggtct cccagcatttc gtttcgccagcacctgcaccc cccacac 120
gataggggt attgggccag tggccttccatgcgtcagc aacggttttt tgtccctttg 180
aagtgggctt ccccctttgtt aaatgggcacct tctgtgccca aagctgaacaca aacatcaac 240
taatgcggct cttcccctt gattataag gattttgccc gatttcgcgctt aattggcaaa 300
aaaatgcgct gttaattcag ataatttac ccaccttttaca ccaatatttc aacggttcc 360
tttaggtgga cgagcactgg gaaatgtgcc cggaaccctt atttgattat tttgcaaat 420
acatcacaat agttagaccc tctggagaac ataaccctga taatgttctc aataatattg 480
aaaaaggag agctatgagt ttcacacttt cctgctgccg cttgattctc ttttggcgcg 540
attttgcctt cctgttttgg ctcaccacga aaccgttgtg tgaatgaagg agaagagag 600
tcagttgtggt gcacggagtt gttacactca aactgatctt aacagcgtta aagctctgtg 660
gatgacctgc ccgagacac gtttttcaat gtggacgccct tttgaactgc tggattgttg 720
cyccgtatca cccgctattg acgcgggcca agagoacact gcgtgccccga taaacatctc 780
tcgaagatgc ttggttgaag acccaccggc cacaagaag cactctaccg gaggacag 840
agtaagagaa ttagacagt atggccatac caatgtgatg aacatcgggg ccaacctact 900
tcgccaaacg atcggaggct cggaggtttt tttggcacaac tgggggatcta 960
tggactcgcg ctgggtgctt ggcacacgga gcgtgaattg gacatacaca aagacagcag 1020
tgaccaagcg atcgccgcat cagttggaac aaccgttgtc gaactctaat cgggacag 1080
acactctta gttgccggcc acaacttaa agaactgagtc gaggcggata aagttgcaag 1140
acacccctgc gctgcggccc tccccctgtgct gtttttttct gtcgctggttag 1200
tgagcgtggcg ttcggcagtt tccgctgca actgggggca gatggtaggc ctctccgtat 1260
cgtaggcatt cacacagcgg ggagctcggc aacttggtat gaaacgagta gacagcagcg 1320
tgagatagg gcctcactgc ttaaccaatt gtaactgtca gacaaagttt actcaatat 1380
accttagct cagttttaaaac tttttttttt atttttacag atcgctgattg 1440
tgaatctctcg ggcacataa ttttcaagcg tgggttttctg tttcactgag cggctgcacc 1500
cgtaggcgag atcaagatgc atctctttctct cttctctgcgt tggttttctt ggtgctgccg 1560
gcagacaaa aacagcgcgc cacacagcgtt gttttttttg cggctcagac agctaccacc 1620
ttttttttct gggagtaag cttttttttt cgggatggat cgggtctttt gggggttgg 1680
tccagacgag tagtacgcttg ataaggcgcg ccggcggcgcct tggaggggccccgtgcag 1740
gcgcagcccg gacgctgatg cggagggcag cctttttcgc gtctccccgt gttcccgtg 1800
tgtgactgcct ggggagccag tccctcgttgc gatgttctga cggggtttggc 1860
acagccacg ccggtagaag ccagctcact cgcagtcgga caacacgatg ccagctgcctgc tggggttggc 1920
aggaagcgcg ccagtcgcgg cgggagaaaa ggagcgcacgg tttcgctcaag gggcgcggt 1980
cggagccga gcggccggcc cgggagtctc cgggggaaca gctctgtatc ttattatgcc 2040
tcctcggttct gtggctggttag tggctacgctt ggtgttcgaa caggggaggc 2100
cgggcccttt aacagcgcag cggacgctgtt cttttttttt cttcttttct ttttccttttgc 2160
ctctctgagc cttggttcct gttctggtgc cttgttttct gttctgtgtt ccggtttgacg 2220
cggagagcg cggagagcgg tagtttttctc cggagttttt cgggttttcca cgggtatcggc 2280
agacgcagca gacagcagcgg tagtttttctc cggagttttt cgggtattttt cgggtatcggc 2340
acacgcagca gccacagcgg tttttttttt aacagcgttt aagcctgttct ttttttttttgc 2400
tgctggttgt aacggggcgg ggaggggaca ttaaggttata aacggggcagc ttttttttttgc 2460
aaactgagca ataaagttttt aacagcgttt aacggggcagc tttttttttt cgggtatttttgc 2520
agctgtggtg aacggtactt cagtttttttt tattttttcgc aacagcacttt cttttttttttgtg 2580
agtcgcaatt cggagttgtt ttttggggcgt tttttttttt cggagttttt cgggtatttttgc 2640
tcctcggttct gtggctggttag tggctacgctt ggtgttcgaa caggggaggc 2700
aagaattcttg aaggtcggctg atcttttcacg tcaactactc ttctgttggg 2760
cccaaccttg ttatagcagc caactgcccgt tggctcactgt aactctgttg caactgatac 2820
acataagcc caacgctggt ggcctctagc ttgagagatg tgatagacgc ggtggcataat 2880
cctgctcgc ggtgtctccg ggaagagcgc agatacatag cattacagct acaatcgcctg 2940
tgctcaaatc ttgggcagac ttaaaggccag agagccgcaac caacgctgctc ttggtcggag 3000
gcagcagcgc cagatgagatc ttggtctcag cgaagagcgc caagagatcgc aaagagccgct 3060
ctgagcttggg agatggtgttc taaagctcag acacagacgc ccaaagagatc agagccgccg 3120
gcagcttgatc tggactgtgctc agggccgagc ctcacatgtggt gatagtgcgcat ctagcctgag 3180
ccacactaact ttgttttctg gcagactgccc ttcgctgtaa atcgctgtgtt gcctgcttaac 3240
atctggctcg ctcatacaac tcaataacatcg acacgagggc taacgctgtt gcctgcttgaga 3300
tgcttcaggg agatactgtga caaaaaaca gcatacaagc gcctagatta caagcagctc 3360
gcctctcaca cgcgctcggt cggctcaggt tccgagccag ttcggtgagc gcatacaga 3420
ctttgcaatgg agatctcagga ccaagagcgc ttatgctcag tgggctgtggt cctctcagct 3480
cttcacaagc gcctctcaca cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 3540
cggctcttcgc gcctctcaca cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 3600
gtcctcttcgc taagctgcttc gcgtgctctc gcgtgctctc gcgtgctctc agatggaggg 3660
acagtggcgct tcggctgtctg ccggctgtctg gcgtgctctc gcgtgctctc agatggaggg 3720
cggctcttcgc taagctgcttc gcgtgctctc gcgtgctctc gcgtgctctc agatggaggg 3780
ttgagtcgct ccgtgctctc gcggctgtctg gcgtgctctc gcgtgctctc agatggaggg 3840
gcctctcaca cggctcagcta gcggctgtctg gcgtgctctc gcgtgctctc agatggaggg 3900
gggctcttcgc gcctctcaca cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 3960
taaaagccacc ctaacatgctc gttgctgtttc cattctctc gcggctgtctg gcgtgctctc 4020
tctcctgctc ttcgctgttc gcgtgctctc gcgtgctctc gcgtgctctc agatggaggg 4080
atatttcatct actaataggt ccataaagac cgcctctcgc gttttttttttt gttttttttttt 4140
atactagct caacttacag cattagcttt ccatttttata cgttctctata agagataaca atactaaatct 4200
atttttttaa ctaacatgctc gtttctctata cgttctctata agagataaca atactaaatct 4260
tgagcttata ctaacctcttt ccggagagcg ctgctttttgc cattttgtag ctgctaaagct 4320
taagtttaat ctaacatgctc gtttctctata cgttctctata agagataaca atactaaatct 4380
ggctgtcttc gcctctcaca cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 4440
cggctcttcgc gcctctcaca cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 4500
ccggagagcg ctgctttttgc cattttgtag cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 4560
ctattttgta ctttagctag cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 4620
cttccttcttt ctttttttttt ccggctgtctg gcgtgctctc gcgtgctctc agatggaggg 4680
tgcagctgctgc cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 4740
agcagctgctgc cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 4800
gcgtctgtctg ctttctttttt ctttttttttt gcgcctctctgc gcgtgctctc gcgtgctctc 4860
tgctgctgctgc cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 4920
ggcgctgctgc cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 4980
cggtctcagt ccctttttttttt ctttttttttt gcgcctctctgc gcgtgctctc gcgtgctctc 5040
ccggagagcg ctgctttttgc cattttgtag cggctcagctc ggctcaggtt gcctgcaggt cattctctgctc 5100
atgctctcct caggtaggtt aagtcagcgc ggtgaagcac gttgtgccag caaagac tccc 5160
atata gcccaaaatgc tggagct tag gttgtgccag ctggttgcaca atcaacccccc cggctcaagag 5220
gtcgggctg aagagggc acgcggag gagatc aggcttggg aaggtcttgg 5280
tgtaagttgtagc ctgtgtagg aagagggc ctggttgcaca atcaacccc tgggtgccag 5340
gacggttggt ggaggtgctg ggggaagc aagagggc ctggttgcaca atcaacccc tgggtgccag 5400
tgggggcgg ggtggagggg aagagggc ctggttgcaca atcaacccc tgggtgccag 5460
gggcgc ggtggagggg ggtggagggg aagagggc ctggttgcaca atcaacccc tgggtgccag 5520
gggcgc ggtggagggg aagagggc ctggttgcaca atcaacccc tgggtgccag 5580
gggcgc ggtggagggg aagagggc ctggttgcaca atcaacccc tgggtgccag 5640
gggcgc ggtggagggg aagagggc ctggttgcaca atcaacccc tgggtgccag 5700
cattgta tggttccagt gccctagttct aatcagac gtggctagttct 5760
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 5820
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 5880
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 5940
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6000
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6060
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6120
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6180
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6240
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6300
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6360
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6420
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6480
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6540
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6600
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6660
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6720
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6780
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6840
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6900
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 6960
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 7020
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 7080
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 7140
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 7200
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 7260
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 7320
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 7380
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 7440
tggtaaagac aatcagac gtggctagttct aatcagac gtggctagttct 7500
What is claimed is:
1. An adeno-associated virus (AAV) filler component comprising a nucleic acid having at least 90% identity to SEQ ID NO:1 or SEQ ID NO:2, wherein the nucleic acid is between 3500 and 4000 nucleotides.

2. An adeno-associated virus (AAV) filler component consisting of a nucleic acid having at least 90% identity to SEQ ID NO:1 or SEQ ID NO:2.

3. The AAV filler component of claim 2, wherein the nucleic acid is between 3500 and 4000 nucleotides.

4. The AAV filler component of claim 1, wherein the nucleic acid is between 3700 and 3850 nucleotides.

5. A recombinant adeno-associated virus (AAV) vector comprising the filler component of claim 1 operably linked to an expression cassette, wherein the AAV vector is approximately 5 kb in length.

6. The AAV vector of claim 5, wherein the expression cassette comprises a promoter.

7. The AAV vector of claim 6, wherein the promoter is a pol III promoter.

8. The AAV vector of claim 7, wherein the promoter is a mU6 promoter.

9. The AAV vector of claim 5, further comprising a target sequence.

10. The AAV vector of claim 9, wherein the target sequence is an RNAi molecule.

11. The AAV vector of claim 5, wherein the AAV vector is an AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, or AAV8 serotype.

12. The AAV vector of claim 5, further comprising an inverted terminal repeat (ITR) of any one of serotype AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, or AAV8.

13. The AAV filler component of claim 1, wherein the nucleic acid has at least 95% identity to SEQ ID NO:1 or SEQ ID NO:2.

14. The AAV filler component of claim 1, wherein the nucleic acid has at least 98% identity to SEQ ID NO:1 or SEQ ID NO:2.
15. The AAV filler component of claim 1, wherein the nucleic acid has at least 99% identity to SEQ ID NO:1 or SEQ ID NO:2.

16. The AAV filler component of claim 2, wherein the nucleic acid has at least 95% identity to SEQ ID NO:1 or SEQ ID NO:2.

17. The AAV filler component of claim 2, wherein the nucleic acid has at least 98% identity to SEQ ID NO:1 or SEQ ID NO:2.

18. The AAV filler component of claim 2, wherein the nucleic acid has at least 99% identity to SEQ ID NO:1 or SEQ ID NO:2.