



US 20250249053A1

(19) United States

(12) Patent Application Publication

Schulze et al.

(10) Pub. No.: US 2025/0249053 A1

(43) Pub. Date: Aug. 7, 2025

## (54) USING INFECTIOUS NUCLEIC ACID TO TREAT CANCER

(71) Applicant: Mayo Foundation for Medical Education and Research, Rochester, MN (US)

(72) Inventors: Autumn J. Schulze, Rochester, MN (US); Stephen J. Russell, Rochester, MN (US); Noura B. Elsedawy, Sioux Falls, SD (US)

(21) Appl. No.: 19/090,602

(22) Filed: Mar. 26, 2025

## Related U.S. Application Data

(63) Continuation of application No. 16/975,782, filed on Aug. 26, 2020, now abandoned, filed as application No. PCT/US2019/021850 on Mar. 12, 2019.

(60) Provisional application No. 62/641,814, filed on Mar. 12, 2018.

## Publication Classification

(51) Int. Cl.

A61K 35/76 (2015.01)  
C12N 7/00 (2006.01)  
C12N 15/85 (2006.01)

(52) U.S. Cl.

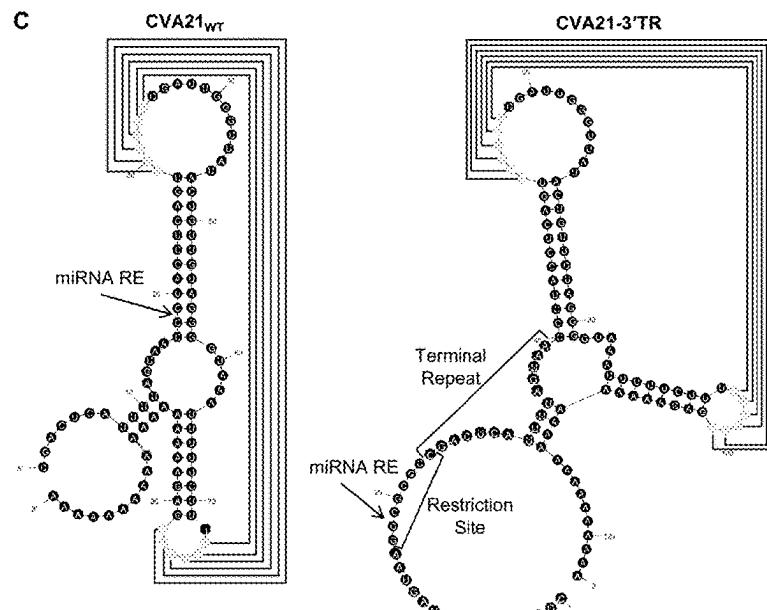
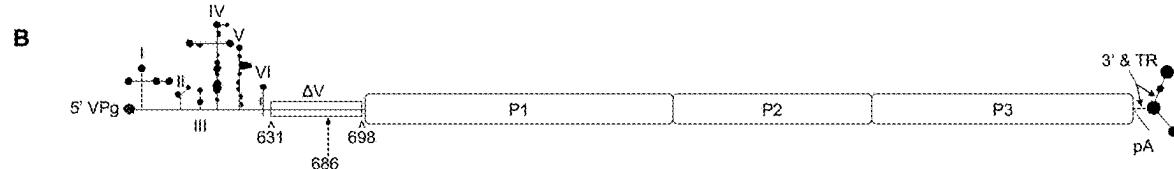
CPC ..... A61K 35/76 (2013.01); C12N 7/00 (2013.01); C12N 15/85 (2013.01); C12N 2310/141 (2013.01)

## (57) ABSTRACT

This document provides methods and materials related to using infectious nucleic acid encoding viruses to reduce the number of viable cancer cells within a mammal. For example, methods for using infectious nucleic acid to treat cancer, engineered viral nucleic acid, and methods for making engineered viral nucleic acid are provided.

Specification includes a Sequence Listing.

A miRT (1x) ACAGCTGGTTGAAGGGGACCAA CTGGAGCCACACACTTCCCTTACATTCCA  
miRT (2x) ACAGCTGGTTGAAGGGGACCAA CGAT ACAGCTGGTTGAAGGGGACCAA CTGGAGCCACACACTTCCCTTACATTCCA TCACCCACACACTTCCCTTACATTCCA



Virus	Insert
CVA21	
CVA21-ΔV(1x)	133-206 Exchanged for 631-698 (5'UTR)
CVA21-ΔV(2x)	133(2x)-206(2x) Exchanged for 631-698 (5'UTR)
CVA21-688(2x)	133(2x)-206(2x) Inserted at position 688 (5'UTR)
CVA21-3'miRT(2x)	133(2x)-206(2x) Inserted at position 7343 (3'UTR)
CVA21-3'TR(1x)	133-206 Between repeats (3'UTR)
CVA21-3'TR(2x)	133(2x)-206(2x) Between repeats (3'UTR)

FIG. 1

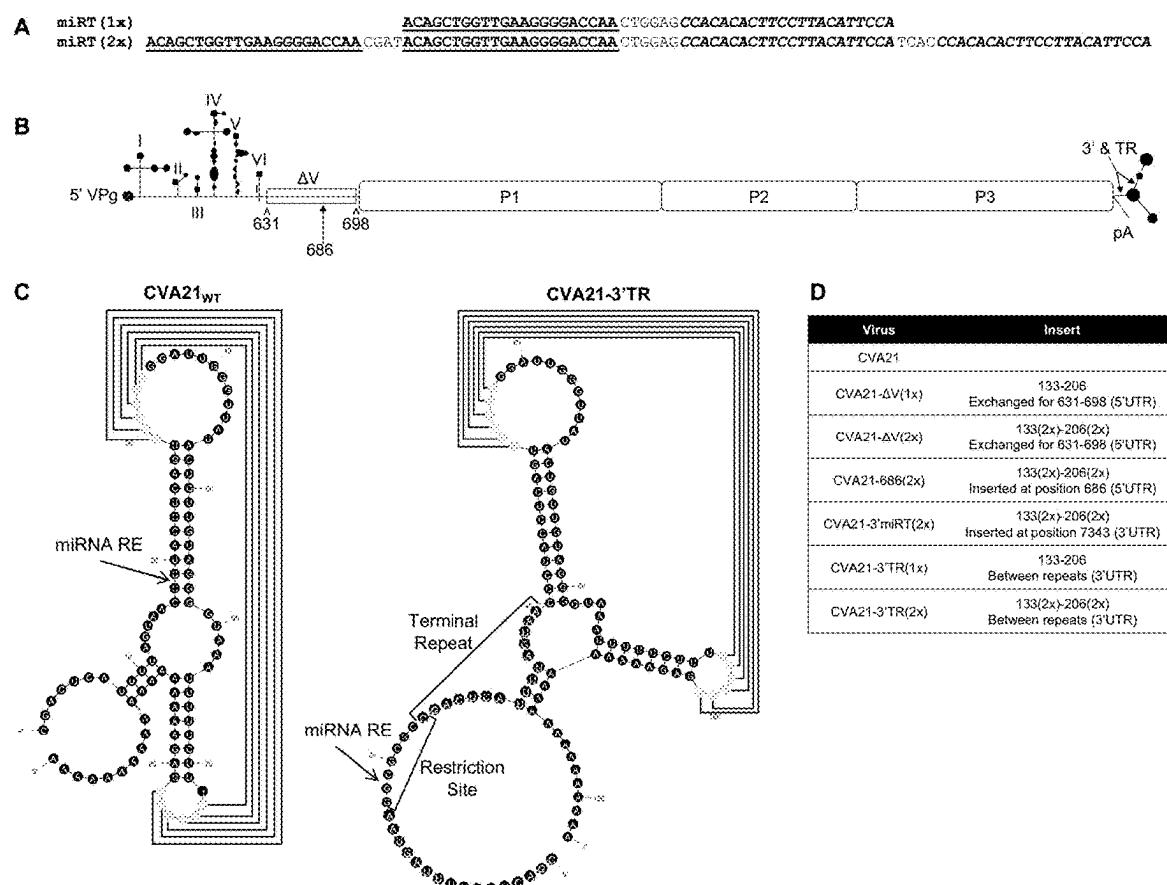


FIG. 2

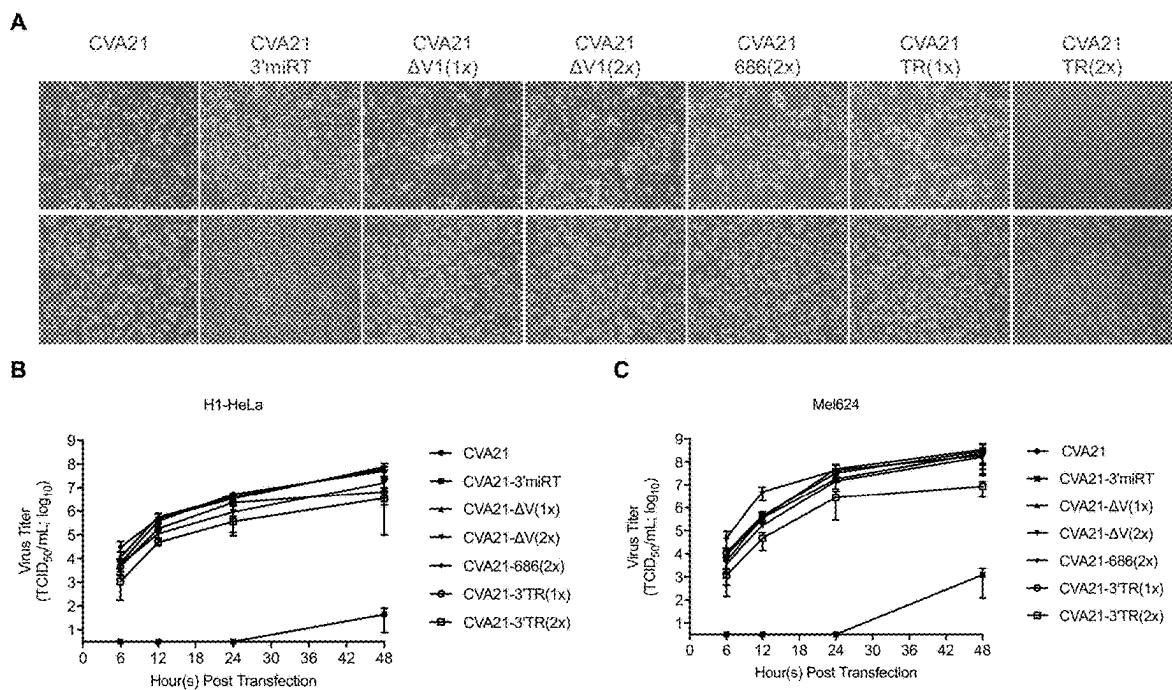
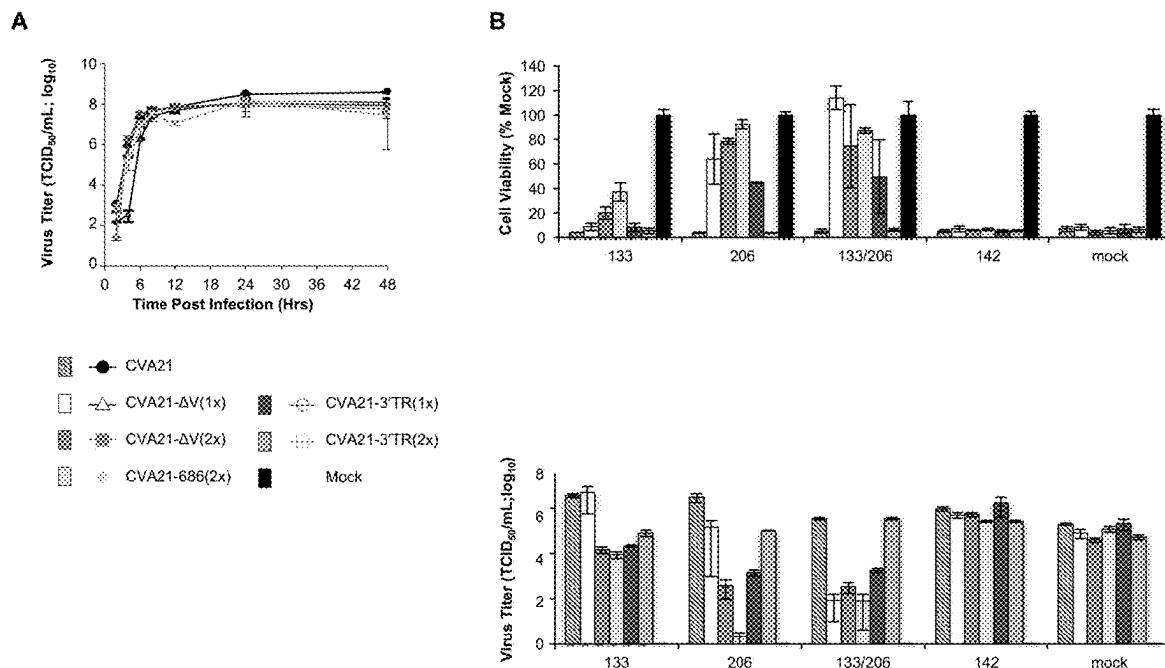
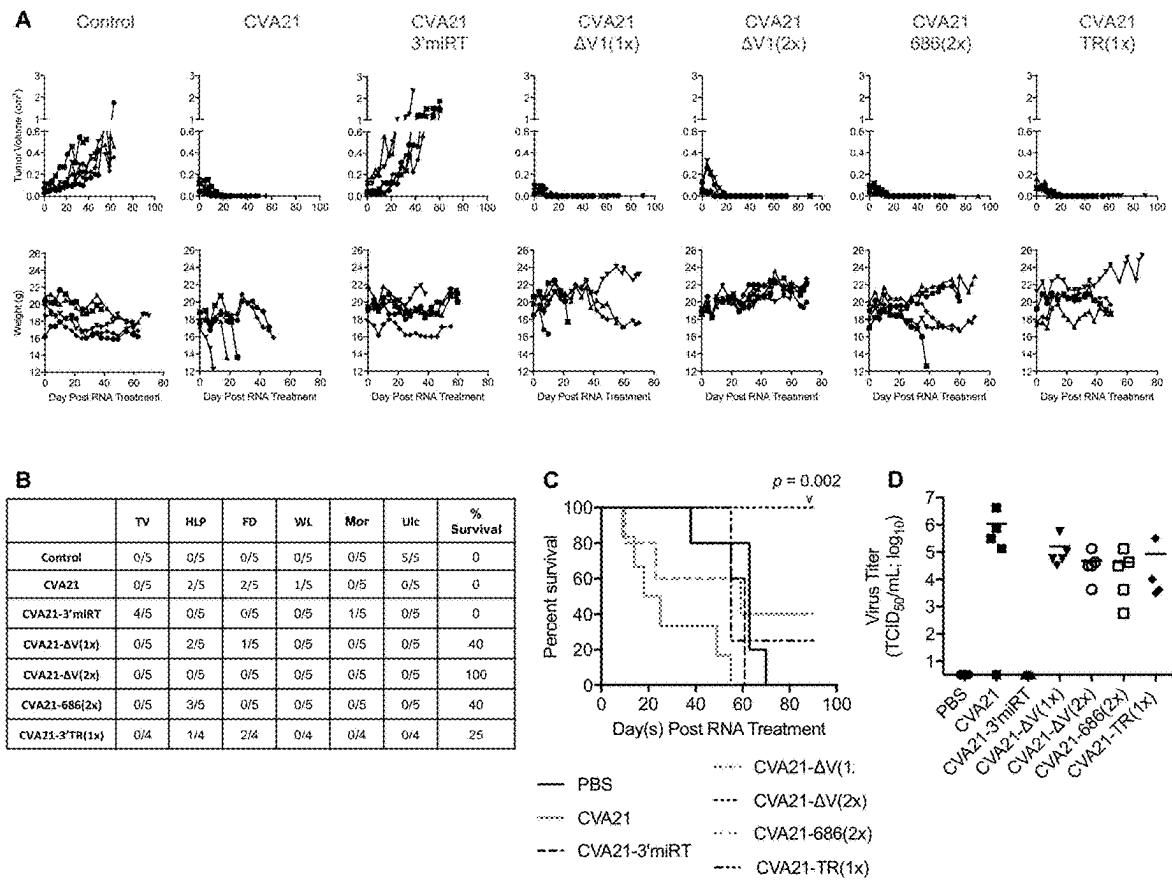


FIG. 3



**FIG. 4**


**FIG. 5**

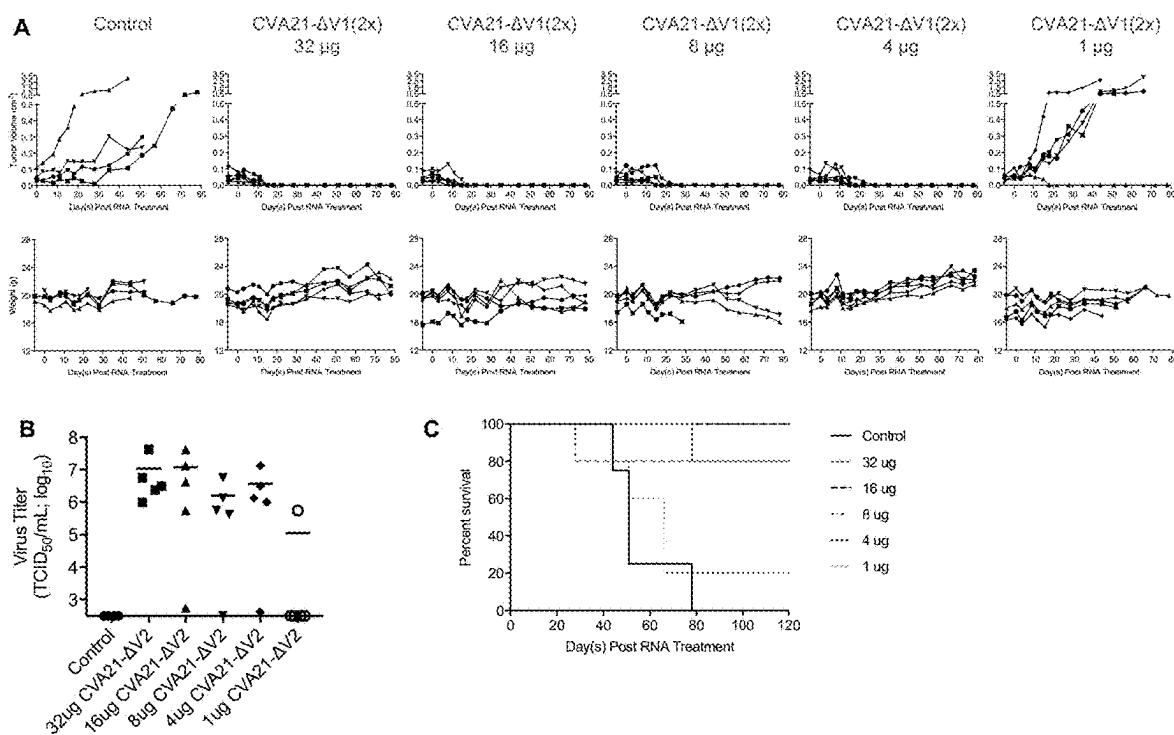


FIG. 6

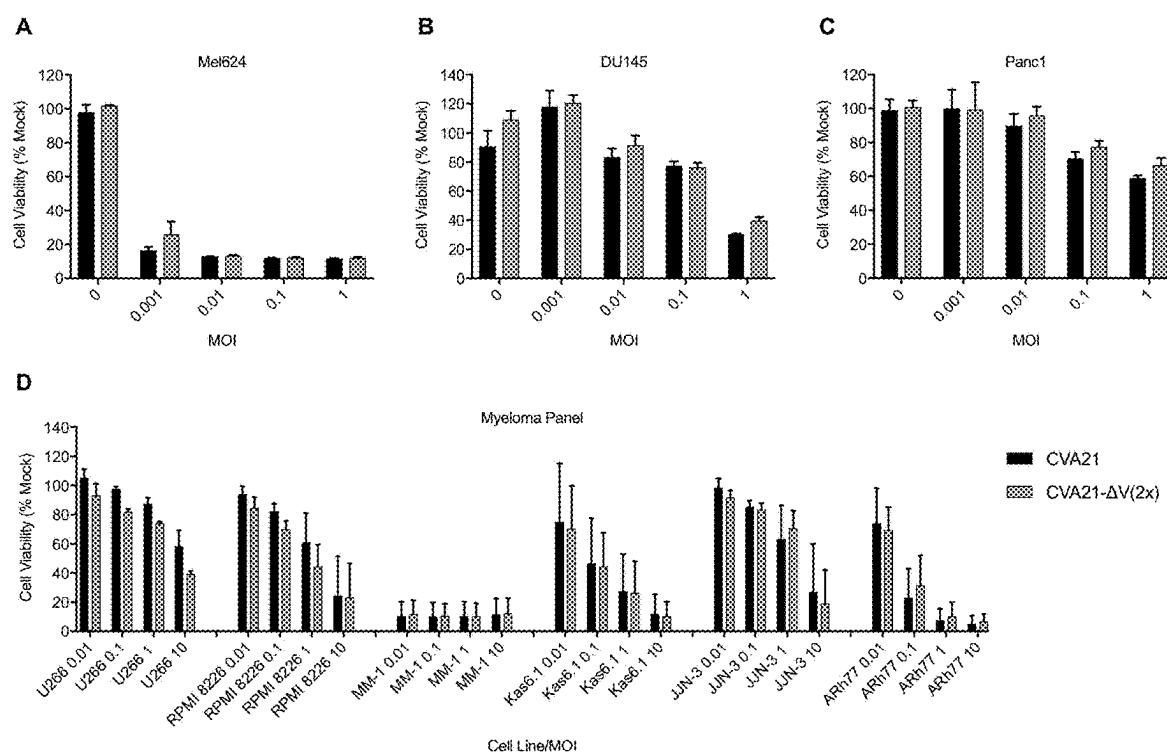


FIG. 7

1 TTAAAACAGC TCTGGGGTTG TTCCCACCCCC AGAGGCCAC GTGGCGGCTA GTACTCTGGT ATTACGGTAC  
 71 CTTTGTCACGC CTGTTTGTGTA TCCCTCCCC CGTAACCTTA GAAGCTTATC AAAAGTCAA TAGCAGGGGT  
 141 ACAAAACAGT ACCTCTACGA ACAAGCACTT CTGTTCCCC GGTGATATCA CATAACTGT ACCCACGGTC  
 211 AAAAGTATT GATCCGGTTAT CCGCTTGAGT ACTTCGAGAA GCCTAGTATC ACCTTGAAT CTTCGATGCG  
 281 TTGCGCTCAA CACTCTGCC CGAGTGTAGC TTAGGCTGAT GAGTCTGGC ACTCCCCACC GGCGACGGTG  
 351 GCCCAGGCTG CGTTGGCGC CTACCCATGG CTGATGCCGT GGGACGCTAG TTGTAACAA GGTGTGAAGA  
 421 GCCTATTGAG CTACTCAAGA GTCTCCGGC CCCTGAATGC GGCTAATCTT AACACGGAG CAACCGCTCA  
 491 CAACCCAGTG AGTAGGTTG CTGAATGCGT AAGTCTGTTG CGGAACGGAC TACTTTGGGT GTCCGTGTT  
 561 CCCTTATAT TCATACTGGC TGCTTATGGT GACAATITAC AAATTGTAC CATATAGCTA TTGGATTGGC  
 631 CACCCAGTAT TGTCAATAT ATITGAGTGT TTCTTCATA AGCCTTATTA ACATCACATT TTIAATCACA  
 701 ATAACAGTG CAAATGGGG CTCAGTTT AACGAAAAG ACCGGTGCACGAGAACATCA AAACGTGGCA  
 771 GCCAATGGAT CCACCATTA TTACACTACT ATCAACTATT ACAAAAGACAG TGCAGAGTAAT TCCGCTACTA  
 841 GACAAGACCT CTCCCAGAT CCATCAAATC ACAGAACCG GGTAAAGGAC TTAATGTTGA AAACAGCACC  
 911 AGCTCTAACAC TCGCTTAACG TGGAAAGCATG TGGGTAGT GACCGTGTGA GGCAAATCAC TTAGGCAAC  
 981 TCGACTTATTA CTACACAAAG AGCAGGAACT GCTATTGTTG CTTACGGTGA ATGGCCACT TACATAATG  
 1051 ATTACAGAAC TAATCCGGTA GATGCACCCA CTGAGCCAGA CGTTAGTACG AACCGGTTT ACACCTAGA  
 1121 ATCGGTGCT TGGAAAGACCA CTTCAGGGG ATGGTGGTGG AAGTTACAG ATTGTTGAA GGACATGGGA  
 1191 ATGTTGGTC AGAATATGTA CTATCACTAC TTGGGGCGCT CTGGTTACAC CATTATGTC CAGTGCAACG  
 1261 CTTCAAAATT TCACCAAGGG GCGTTAGGAG TTTCCTGAT ACCAGAGTT GTCATGGCTT GCAACACTGA  
 1331 GAGTAAAACG TCATACGTTT CATAATCAA TGCAATCTT GGTGAGAGAG GCGGTGAGTT TACGAACACC  
 1401 TACAATCCGT CAAATACAGA CGCCAGTGAG GGCAGAAAGT TTGCAAGCATT GGATTATTG CTGGGTTCTG  
 1471 GTGTTCTAGC AGGAAACGCC TTTGTGTACC CGCACCAAGAT CATCAACTA CGTACCAACA ACAGTGCAAC  
 1541 ATTGTTGGTG CCATACGTAAC ACTCACTGT GATTGATTGT ATGGCAAAAC ACAATAACTG GGGCATTGTC  
 1611 ATATTACCCAC TGGCACCCCTT GGCCTTGCC GCAACATCGT CACCAAGGT GCCTATTACA GTGACCATTG  
 1681 CACCATGTG TACAGAATT AATGGGTTGA GAAACATCAC CGTCCCAGT CATCAAGGGT TGCCGACAAAT  
 1751 GAACACACCT GGTTCACATC AATTCTTAC ATCTGATGAC TTCCAGTCGC CTTGTCCTT ACCTAATT  
 1821 GATGTTACTC CACCAATACA CATAACCGGG GAAAGTAAAGA ATATGATGGA ACTAGCTGAA ATTGACACAT  
 1891 TGATCCAAT GAACGCAGTG GACGGGAAGG TGAACACAAT GGAGATGTAT CAAATACCAT TGAATGACAA  
 1961 TTTGAGCAAG GCACCTATAT TCTGTTATC CCTATCACCT GCTTCTGATA AACGACTGAG CCACACCATG  
 2031 TTGGGTGAAA TCCTAAATT TGACACCCAT TGACGGGGT CCATCAGGTT CACCTTCTA TTTTGTGGCA  
 2101 GTATGATGGC CACTGGTAAA CTGCTCTCA GCTATTCCCC ACCGGGAGCT AAACCAACAA CCAATCGCAA  
 2171 GGATGCAATG CTAGGCACAC ACATCATCTG GGACCTAGGG TTACAATCCA GTTGTCCAT GTTGCACCG  
 2241 TGGATCTCCA ACACAGTGTA CAGACGGTGT GCACGTGATG ACTTCACTGA GGGCGGATT ATAACATTG  
 2311 TCTATCAAAC TAGAATTGTG GTACCTGCTT CAACCCCTAC CAGTATGTT ATGTTAGGCT TTGTTAGTGC  
 2381 GTGTCAGAC TTCAGTGTCA GACTGCTTAG GGACACTCCC CATATTAGTC AATCGAAACT AATAGGACGT  
 2451 ACACAAGGC TTGAAAGACT CATTGACACA GCGATAAAAGA ATGCTTAAG AGTGTCCCAA CCACCCCTCGA  
 2521 CCCAGTCAC TGAAGCAACT AGTGGAGTGA ATAGCAGGAG GGTGCCAGCT CTAATGCTG TGGAAACAGG  
 2591 AGCATCTGGT CAAGCAATCC CCAAGTGATG GGTGGAAACT AGGCACGTGG TAAATTACAA AACCAAGGTCT  
 2661 GAATCGTGTGTTGAGTCATT CTTTGGAGA GCTCGTGTG TCACAACTCTT ATCCTGACCA AACTCTCCA  
 2731 AGAGCGGAGA GGAGAAAAAG CATTCAACA TATGGAATAT TACATACACC GACACTGTCC AGTTACCGAG  
 2801 AAAATTAGAA TTTTCACGT ATTCCAGGTT TGATCTGAA ATGACTTTG TATTACAGA GAACTATCCT  
 2871 AGTACAGCCA GTGGAGAAGT GCGAAACCAAG GTGTACCAAGA TCATGTATAT TCCACCAAGGG GCACCCCGCC  
 2941 CATCATCTG GGATGACTAC ACATGCAAT CCTCTTCAAA CCCTCCATC TTCTACATGT ATGGAAATGC  
 3011 ACCTCCACGG ATGTCAATT CTTACGTAGG GATTGCAAT GCCTATTACAC ACTTCTACGA TGGCTTGTCA  
 3081 CGGGTGCAC TTGAGGGTGA GAAACACCGAT GCTGGCGACA CGTTTACGG TTTAGTGTCC ATAAATGATT  
 3151 TTGGAGTTT AGCAGTTAGA GCAGTAAAGC GCAAGTAATC ACATACAAAT CACACATCTG TGAGAGTGT  
 3221 CATGAAACCA AAACACATTC GGTGTGGTGC CCCCACGACTT CTCAGCTG TATTATACAG GGGAGAGGG  
 3291 GTGGACATGA TATCCAGTGC AATTCTACCT CTGGCCAAGG TAGACTCAAT TACCACTTTT GGGTTTGGTC  
 3361 ATCAGAACAA AGCAGTGTAC GTTGCCTGGT ACAAGATTTG CAACTACAC CTAGCAACCC CAAGTGTACA  
 3431 CTTGAATGCA ATTAGTATGT TATGGGACAG GGATTTAATG GTGGTGAAT CTAGAGCCCG GGGAACTGAT  
 3501 ACCATCGCCA GATGTAGTTG CAGGTGTGGA GTTACTATT GTGAATCTG GAGGAAGTAC TACCCGTCA  
 3571 CTTTACTGG CCCAACGTTT CGATTATGG AAGCAAACGA CTACTATCCA GCAAGATACC AGTCTCACAT  
 3641 GCTGATAGGG TCGGGATTG CAGAACCCGG GGACTGCGGT GGGATACTGA GGTGCACTA TGGGGTAATT  
 3711 GGTATCATT CTGCAGGAGG TGAAGGGTGA GTAGCTTGTG CTGACATTAG AGACCTCTGG GTGTATGAAG  
 3781 AGGAGGCCAT GGAACAGGGA ATAACAAGCT ACATCGAATC TCTCGGCACA GCCTTGGCG CAGGGTTCAC  
 3851 CCACACAATC AGTGAGAAG TGACTGAATT GACAACGATG GTTACCAAGCA CTATCACAGA AAAACTACTG  
 3921 AAAAAACTGG TGAAAATAGT GTCGGCTCTA GTGATTGTTG TGAGAAATT TGAGGACACT ACCACGATCC  
 3991 TTGCAACACT AGCACTACTC GGGTGTGATA TATCTCTTG GCAATGGTTG AAGAAGAAGG CATGTGACTT

FIG. 7 (Continued)

4061 ACTAGAGATT CCTTATGTGA TCGGCCAAGG TGATGGGTGG ATGAAGAAAAT TCACAGAGGC GTGCAATGCA  
 4131 GCTAAAGGCT TAGAGTGGAT TAGCAACAAA ATTCCAAGT TTATAGATTG GTGAAAGTGT AAAATTATCC  
 4201 CAGACGCTAA GGACAAGGTG GAATTCTCA CCAAGTTGAA ACAGCTAGAC ATGTTGGAAA ATCAAATTGC  
 4271 ACCCATCCAC CAATCTGCC CCAGCCAAGA ACAACAAGAG ATTCTTTCA ACAATGTGAG ATGGTTAGCA  
 4341 GTCCAGTCCC GTCGGTTGC ACCATTATAC GCTGTGGAGG CACGCCGAAT TAACAAAATG GAGAGCACAA  
 4411 TAAACAAATT TATACAGTTC AAGAGCAAAC ACCGTATIGA ACCAGTATGT ATGCTCATTG ATGGGTCAAC  
 4481 AGGGACGGGT AAATCTATAG CTACTTCATT AATAGGTAGA GCAATAGCAG AGAAGGAAAG CACATCAGTC  
 4551 TATTCAATGC CACCTGACCC ATCTCACTT GATGGCTATA AACAAACAAGG GGTAGTGATT ATGGACGACC  
 4621 TAAACCAAAA CCCCAGATGGT ATGGACATGA AACTGTTTG CCAAATGTA TCAACAGTGG AGTTTATTCC  
 4691 TCCAATGGCC TCATTAGAGG AGAAGGGCAT TTGTTTACA TCTGATTATG TCCCTGGCTTC TACCAACTCT  
 4761 CATTCATTG TACCACCCAC AGTGGCTCAC AGTGTGATGCCT TAACAGACG ATTTGATCATIT GATGTGGAGG  
 4831 TTTCACCGAT GTCTGAACAT TCACTGAAAG GCAAACTGAA TATGCCACG GCCACTCAAT TGTGTAAGGA  
 4901 TTGTCACCAACA CCTGCAAAATT TTAAAAAGTG TTGCCCTCTC GTTGTGGAA AGGCCTTGCA ATTAATGGAC  
 4971 AGGTACACCA GACAAAGGTT CACTGTAGAT GAGATTACCA CATTATCAT GAATGAGAAA ACAGAAGGG  
 5041 CCAATATCGG CAATTGATG GAAGCCTTGT TTCAAGGACC ACTAAGGTAT AAAGATTGAGA AGATCGATGT  
 5111 GAAGACAGTT CCCCCCCCCTG AGTGCATCG TGATTTGTTA CAAGCAGTGG ATTCTCAAGA GGTTAGGGAT  
 5181 TACTGTGAGA AGAAAGGCTG GATCGTTAAC GTTACTAGCC AGATTCAACT AGAAAGGAAC ATCAATAGGG  
 5251 CCATGACTAT ACTCCAAGCT GTTACCATC TCGCAGCAGT CGCAGGAGTA GTGTATGTA TGACAAACT  
 5321 CTTGCCGGT CAACAGGGTG CATAACTGG CTTGCCAAC AAAAAACCA ATGTCCTAC TATCAGAGTC  
 5391 GCTAAAGTCC AGGGGCCAGG ATTTGACTAC GCAGTGGCAA TGGCAAAAAG AAACATAGTT ACTGCAACCA  
 5461 CCACCAAGGG TGAATTACCG ATGCTAGGGG TGATGATAA TGTAGCAATA TTGCCAACCC ATGCCGCTCC  
 5531 AGGAGAACCC ATTATTATG ATGGGAAAGA AGTAGAGATC CTAGACGCCA GAGCCTTAGA AGATCAAGCG  
 5601 GGAACCAATC TTGAGATCAC CATTATTACT CTAAAAAGAA ATGAGAACCT TAGAGACATC AGATCACATA  
 5671 TTCCCCACCA AATTACTGAA ACTAACGATG GAGTGTGAT CTGAAACACT AGCAAGTACC CCAATATGTA  
 5741 TGTCCCCGT GGTGCTGTGA CCGAACAGGG ATATCTTAAT CTCAGTGGAC GTCAAACACTGC TCGCACTTAA  
 5811 ATGTACAACCT TTCAACAAAG GGCAGGCCAG TGCGGAGGAA TCATCACTTG TACTGGCAAAT GTCATTGGGA  
 5881 TGATGTTGG CGGGAACGGT TCACATGGGT TTGCTGCAGC CCTCAAGCGA TCATACTTCA CTAAAATCA  
 5951 GGGCGAACATC CAGTGGATGA GGTCAAA AGAAGTGGGG TACCCATTA TAAATGCCCA ATCCAAGACA  
 6021 AAGTTAGAAC CCAGTGTCTT CCACTATGTT TTGAAGGTG TTAAGGAACC AGCTGTACTC ACTAAGAATG  
 6091 ACCCCAGACT AAAAACAGAT TTGAAGAAG CCATCTTTC TAAATATGT GGGAAACAAA TTACTGAAGT  
 6161 GGACGAGTAC ATGAAAGAAG CAGTGGATCA CTATGCAGGA CAGITAATGT CACTGGATAT CAACACAGAA  
 6231 CAGATGTGCC TGGAGGATGC CATGTACGGT ACCGATGGTC TTGAGGCCCT GGATCTTAGC ACTAGTGTG  
 6301 GATATCCTTA TGTGCAATG GGGAAAAGA AAAGAGACAT TCTAGATAAA CAGACCAGAG ATACTAAGGA  
 6371 GATGCAGAGA CTTITAGATA CCTATGGAAT CAATCTACCA TTAGTCACGT ACGTGAAAGA TGAACCTAGG  
 6441 TCAAAGACTA AAGTGGAAACA AGGAAAGTC AAGATTGATTG AAGCTTCCAG CTTTAATGAT TCAGTTGCAA  
 6511 TGAGAATGGC CTTTGGCAAT CTTTACGCAG CTTTCCACAA GAATCCAGGT GTGGTGACAG QATCAGCAGT  
 6581 TGGTGTGAC CCAGATTGT TTTGGAGTAAGATACAGTG CTAATGGAAG AAAAACTCTT CGCTTTGAC  
 6651 TACACAGGGT ATGATGCCTC ACTCAGCCCT GCTGGTTG AAGCTTTAA ATGGTGTAA GAAAAAATTG  
 6721 GATTGGCAG TAGAGTAGAC TATATAGACT ACCTGAACCA CTCTCACCAC CTTTACAAA ACAAGACTTA  
 6791 TTGTGTCAAA GGCAGCATGC CATCCGGCTG CTCGGCACC TCAATTTC ACTCAATGAT TAACAACCTG  
 6861 ATCATTAGGA CGCTTTTACT GAGAACCTAG AAGGGCATAG ACTTGGACCA TTAAAAATG ATTGCCTATG  
 6931 GTGATGACGT GATAGCTTCC TACCCCATG AGGTTGACGC TACTCTCTCA GCCCAATCAG GAAAAGACTA  
 7001 TGGACTAACATGACTCCAG CAGATAAATC AGAACCTTT GAAACAGTC CATGGGAGAA TGTAAACATT  
 7071 CTGAAAAGAT TTTCAGAGC AGATGAGAAG TATCCATTCC TGGTCATCC AGTGTGACCA ATGAAAGAAA  
 7141 TTCACGAATC AATCAGATGG ACCAAGGACC CTAGAACAC ACAGGATCAC GTACGCTCGT TGTGCTATT  
 7211 AGCTGGCAC AACGGTGAAG AAGAATACAA TAAATTGTTA GCTAAATCA GAAGTGTGCC AATCGGAAGA  
 7281 GCTTATTGTC TCCAGAGTA CTCTACATTG TACCGCCGAT GGTCGACTC ATTITAGTAA CCCTACCTCA  
 7351 GTCGGATTGG ATTGGTTAT ACTGTTGAG GGGTAAATT TTCTTAAATT CGGAGAAAAA AAAAAAAA  
 7421 AAAAGAGCTC CCAATCACTA GTGAATTGCGC GGCAGCTGC AGGTGACCA TATGGGAGAG CTCCCAACGC  
 7491 GTTGGATGCA TAGCTGAGT ATTCTATAGT GTCACCTAA TAGCTGGCG TAATCATGGT CATAGCTGTT  
 7561 TCCTGTGTGA ATTGTATAC CGCTCACAAT TCCACACAAAC ATACGAGCCG GAAGCATAAA GTGAAAGCC  
 7631 TGGGGTGCCT ATGAGTGAG CTAACCTACA TTAATTGCGT TGCGCTACT GCCCCCTTC CAGTCGGGAA  
 7701 ACCTGTCGTG CGAGCTGCAT TAATGAAATCG GCAAACGCGC GGGGAGAGGC GGTTGCGTA TTGGGCGCTC  
 7771 TTCCGCTTCC TCGCTCACTG ACTCGCTGCCT CGCGTCGTT CGCGCTGCAG GAGCAGTAC AGCTCACTCA  
 7841 AAGGCGGTAA TACGGTTATC CACAGAATCA GGGGATAACCG CAGGAAGAA CATGTGAGCA AAAGGCCAGC  
 7911 AAAAGGCCAG GAACCGTAA AAGGCCAGT TCTGGCGTT TTTCATAGG CTCCGCCCCC CTGACGAGCA  
 7981 TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATACCA GGCCTTCCC  
 8051 CCTGGAAGCT CCCTCGTGCCT CTCTCTGTT CGACCCCTGC CGCTTACCGG ATACCTGTCC GCCTTCTCC  
 8121 CTTCGGAAAG CGTGGCGCTT TCTCATAGCT CACGCTGTAG GTATCTAGT CGGGTGTAGG TCGTTGCGCTC

FIG. 7 (Continued)

8191 CAAGCTGGGC TGTGTGCACG AACCCCCCGT TCAGCCCCGAC CGCTGCCCT TATCCGGTAA CTATCGTCTT  
8261 GAGTCCAACC CGGTAAGACA CGACTTATCG CCACCTGGCAG CAGCCACTGG TAACAGGATT AGCAGAGCGA  
8331 GGTATGTAGG CGGTGCTACA GAGTTCTGA AGTGGTGGCC TAACTACGGC TACACTAGAA GAACAGTATT  
8401 TGGTATCTGC GCTCTGCTGA AGCCAGTTAC CTTCGGAAAA AGAGTTGTA GCTCTGATC CGGCAAACAA  
8471 ACCACCGCTG GTAGCGGTGG TTTTTTGTG TGCAAGCAGC AGATTACCGG CAGAAAAAAA GGATCTCAAG  
8541 AAGATCCTT GATCTTTCT ACGGGGTCTG ACGCTCAGTG GAACGAAAAC TCACGITAAG GGATTITGGT  
8611 CATGAGAGTA TCAAAAAGGA TCTTCACCTA GATCCCTTTA AAITAAAAT GAAGTTTAA ATCAATCTAA  
8681 AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTG GGCACCTATC TCAGCGATCT  
8751 GTCTATTTCG TTCATCCATA GTTGCCTGAC TCCCCTGCTGT GTAGATAACT ACGATAACGGG AGGGCTTACC  
8821 ATCTGGCCCC AGTGTGTCGAA TGATACCGCG AGACCCACGC TCACCGGCTC CAGATTATTC AGCAATAAAC  
8891 CAGCCAGCCG AGAGGGCCGA GCGCAGAAAGT GGCTCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAAATT  
8961 GTTGCCTGGGA AGCTAGAGTA AGTAGTTGCA CAGTTAATAG TTTGCGCAAC GTTGTGCCA TTGCTACAGG  
9031 CATCGTGGTG TCACGCTCGT CGTTGGTAT GGCTTCATIC AGCTCCGGTT CCCAACGATC AAGGGAGTT  
9101 ACATGATCCC CCATGTTGTG CAAAAAAGCG GTTAGCTCTC TCGGTCTCC GATCGTTGTC AGAAGTAAGT  
9171 TGGCCGCAGT GTTATCACT ATGGTTATGG CAGCACTGCA TAATTCTCTT ACTGTATGC CATCCGTAAG  
9241 ATGCTTTCT GTGACTGGTG AGTACTCAC CAAAGTCATTC TGAGAATAGT GTATGCCGCG ACCGAGTTGC  
9311 TCTTGCCCGG CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC ATCATTGGAA  
9381 AACGTTCTTC GGGCGAAAAA CTCTCAAGGA TCTTACCGCT GTTGAGATCC AGTTGATGT AACCCACTCG  
9451 TGCACCCCAAC TGATCTTCAG CATCTTTAC TTTCACCCAGC GTTCTGGGT GAGCAAAAC AGGAAGGCCAA  
9521 AAATCCGCAAA AAAAAAGGAAT AAGGGCACA CGGAAATGTT GAGAATCTCAT ACTCTCCCTT TTCAATATT  
9591 ATTGAAGCAT TTATCAGGGT TATTGCTCTA TGAGCGGATA CATATTGAA TGTATTAGA AAAATAAAC  
9661 AATAAGGGTT CCGCGCACAT TTCCCCGAAA AGTGCCACCT GATGCGGTGT GAAATACCGC ACAGATGCGT  
9731 AAGGAGAAAAA TACCGCATCA GGAAATTGTA AGCGTTAATA TTTTGTAAATT CGCGTTA AATTTTTGTT  
9801 AAATCAGCTC ATTTTTAAC CAATAGGCCG AAATCGGCAA AATCCCTTAAATCAAAG AATAGACCGA  
9871 GATAGGGTTG AGTGTGTTG CAGTTGGAA CAAGAGTCCA CTATTAAAGA ACGTGGACTC CAACGTCAAA  
9941 GGGCGAAAAA CCGTCTATCA GGGCGATGGC CCACTACGTG AACCACATACC CTAATCAAGT TTTTGGGGT  
10011 CGAGGTGCG TAAAGCACTA AATCGGAACC CTAAGGGAG CCCCCGATTT AGAGCTTGAC GGGGAAAGCC  
10081 GGCGAACGTG GCGAGAAAGG AAGGGAAGAA AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG  
10151 GTCACGCTGC GCGTAACCAC CACACCCGCC GCGCTTAATG CGCCGCTACA GGGCGCGTCC ATTCGCCATT  
10221 CAGGCTGCGC AAATGTTGGG AAGGGCGATC GGTGCGGGCC TCTTCGCTAT TACGCCAGCT GCGAAAGGG  
10291 GGATGTGCTG CAAGGCGATT AAGTTGGGTA ACGCCAGGGT TTCCCAAGTC ACGACGTTGT AAAACGACGG  
10361 CCAGTGAATT GTAATACGAC TCACTATAGG GCGAATTGGG CCCGACGTG CATGCTCCCG GCCGCCATGG  
10431 CGGCCGCGGG AATTGATTG AGGCATGCTA ATACGACTCA CTATAGG

FIG. 8

1 TTAAAACAGC TCTGGGGTTG TTCCCACCCCC AGAGGCCAC GTGGCGGCTA GTACTCTGGT ATTACGGTAC  
 71 CTTTGTCACGC CTGTTTGTGTA TCCCTTCCCC CGTAACCTTA GAAGCTTATC AAAAGTCAA TAGCAGGGGT  
 141 ACAAAACAGT ACCTCTACGA ACAAGCACTT CTGTTTCCCC GGTGATATCA CATAGACTGT ACCCACGGTC  
 211 AAAAGTGTATT GATCCCGTTAT CCGCTTGGT ACTTCGAGAA GCCTAGTATC ACCTTGGAT CTTCGATGCG  
 281 TTGCGCTCAA CACTCTGCC CGAGTGTAGC TTAGGCTGAT GAGTCTGGC ACTCCCCACC GGCACGGTG  
 351 GCCCAGGCTG CGTGGCGGC CTACCCATGG CTGATGCCGT GGGACGCTAG TTGTAACAA GGTGTGAAGA  
 421 GCCTATTGAG CTACTCAAGA GTCTCCGGC CCTGAAATGC GGCTAACCTT AACACGGAG CAACCGCTCA  
 491 CAACCCAGT AGTAGGTTGT CGTAATGCGT AAGTCTGTGG CGGAACGGAC TACTTGGGT GTCCGTGTT  
 561 CCCTTATAT TCATACTGGC TGCTTATGGT GACAATTAC AAATTGTTAC CATATAGCTA TTGGATTGGC  
 631 GCGCCGACA GCTGGTTGAA GGGGACCAAC TGAGGCCACA CACTCCCTA CATTCCACGG CGCGCCAATA  
 701 AACAGTGCACAA ATGGGGGCTC AAGTTCAAC GCAAAAGACC GGTGCGCACG AGAATCAAA CGTGGCAGCC  
 771 AATGGATCCA CCAATTAACTA CACTACTATC AACTATTACA AAGACAGTGC GAGTAATTCC GCTACTAGAC  
 841 AAGACCTCTC CCAAGATCCA TCAAAATTCA CAGAACCGGT TAAGGACTTA ATGTTGAAAA CAGCACCCAGC  
 911 TCTAAACTCG CCAACGTGG AAGCATGTGG GTACAGTGAC CGTGTGAGGC AAATCACTT AGGCAACTCG  
 981 ACTATTACTA CACAAGAAC AGCCAATGCT ATTGTTGCTT ACGGTGAATG GCCCACTTAC ATAATGATT  
 1051 CAGAAGCTAA TCCGGTAGAT GCACCCACTG AGCCAGACGT TAGTAGCAAC CGGGTTTACA CCCTAGAAATC  
 1121 GGTGCTTGG AAGACCACTT CAAGGGGATG GTGGTGGAAAG TTACAGATT GTTGAAGGA CATGGGAATG  
 1191 TTTGGTCAGA ATATGTA CCAACTACTG GGCGCTCTG GTTACACCAT TCATGTCAG TGCAACGCTT  
 1261 CAAAATTCA CCAAGGGGGC TTAGGAGTTT TCTGTATACC AGAGTTGTC ATGGCTTGCAC ACAGTGGAG  
 1331 TAAAACGTCA TACGTTTCAAT ACATCAATGC AAATCCTGGT GAGAGAGGCG GTGAGTTAC GAACACCTAC  
 1401 AATCCGTCAA ATACAGACGC CAGTGAGGGC AGAAAGTTT CAGCATTGGA TTATTTGCTG GGTCTGGTG  
 1471 TTCTAGCAGG AAACGCCCTT GTGTACCCGC ACCAGATCAT CAACCTACGT ACCAACAAACA GTGCAACAAAT  
 1541 TGTGGTGCCA TACGTAACACT CACTTGTGAT TGATTGTATG GCAAAACACA ATAACCTGGG CATTGTCTA  
 1611 TTACCACTGG CACCCCTGGC CTTGCGGCAC ACATCGTCAC CACAGGTGCC TATTACAGTG ACCATTGCAC  
 1681 CCATGTGTAC AGAATTCAAT GGGTTGAGAA ACATCACCGT CCCAGTACAT CAAGGGTTGC CGACAATGAA  
 1751 CACACCTGGT TCAATCAAT TCCCTACAT TGATGACTTC CAGTCGCCCT GTGCCCTTACCA TAATTTGAT  
 1821 GTTACTCCAC CAATACACAT ACCGGGGAA GTAAAGAATA TGATGAAACT AGCTGAAATT GACACATTGA  
 1891 TCCCAATGAA CGCAGTGGAC GGGAAAGGTGA ACACAATGGA GATGTATCAA ATACCATGAA ATGACAATT  
 1961 GAGCAAGGCA CCTATATTCT GTTATCCCT ATCACCTGCT TCTGTATAAC GACTGAGGCC CACCATGTT  
 2031 GGTGAAATCC TAAATTATTA CACCCATTGG ACCGGGTCCA TCAGGTTAC CTTCTATTT TGTGGCAGTA  
 2101 TGATGGCCAC TGGTAAACTG CTCCTCAGCT ATTCCCCACC GGGAGCTAA CCACCAACCA ATCGCAAGGA  
 2171 TGCAATGCTA GGCACACACA TCATCTGGG CCTAGGGTTA CAATCCAGTT GTTCCATGGT TGCACCGTGG  
 2241 ATCTCCAACA CAGTGTACAG ACGGTGTGCA CGTGATGACT TCACTGAGGG CGGATTATA ACTTGTCTCT  
 2311 ATCAAACATG AATTGTGGTA CCTGCTTCAA CCCCTACCAAG TATGTTCATG TTAGGCTTIG TTAGTGCCTG  
 2381 TCCAGACITC AGTGTACGAC TGCTTAGGGAA CACTCCCCAT ATTAGTCAAT CGAAACTAAT AGGACAGTACA  
 2451 CAAGGCATTG AAGACCTCAT TGACACAGCG ATAAAGAATG CCTTAAGAGT GTCCAACCA CCCTCGACCC  
 2521 AGTCAACTGA AGCAACTAGT GGAGTGAATA GCCAGGAGGT GCCAGCTTA ACTGCTGTGG AAACAGGAGC  
 2591 ATCTGGTCAA GCAATCCCCA GTGATGTGGT GGGAAACTAGG CACGTGTTAA ATTACAAAAC CAGGTCTGAA  
 2661 TCGTGTCTG AGTCATTCT TGGGAGAGCT GCGTGTGTC CAATCCTATC CTTGACCAAC TCTCTCAAGA  
 2731 GCGGAGAGGA GAAAAGCAT TTCAACATAT GGAATATTAC ATACACCGAC ACTGTCCAGT TACGCAGAAA  
 2801 ATTAGAATTTC TTCACGTATT CCAGGTTGTA TCTGAAATG ACTTTGTTAT TCACAGAGAA CTATCTTAGT  
 2871 ACAGCCAGTG GAGAAGTGC GAAACCAAGGTG TACCAAGATCA TGTATATTCC ACCAGGGCA CCCGCCAT  
 2941 CATCTGGGA TGACTACACA TGGCAATCCT CTCACAAACCC TTCCATCTTC TACATGTATG GAAATGCACC  
 3011 TCCACGGATG TCAATTCCCT ACGTAGGGAT TGCCAATGCC TATTACACACT TCTACGATGG CTTGCACGG  
 3081 GTGCCACTTG AGGGTGAGAA CACCGATGCT GCGACACGT TTACGGTT AGTGTCCATA ATGATTGTTG  
 3151 GAGTTTGTAC AGTTAGAGCA GTAAACCGCA GTAATCCACA TACAATACAC ACATCTGTGA GAGTGTACAT  
 3221 GAAACCAAAA CACATTGGT GTTGGTGCCTT CAGACCTCT CGAGCTGTAT TATACAGGGG AGAGGGAGTG  
 3291 GACATGATAT CCAGTGCAT TCTACCTCTG GCCAAGGTAG ACTCAATTAC CACTTTGGG TTGGTCATC  
 3361 AGAACAAAGC AGTGTACGTT GCGGGTTACA AGATTTGCAA TCAACACCTA GCAACCCCAA GTGATCATT  
 3431 GAATGCAATT AGTATGTTAT GGGACAGGGAA TTAAATGGTG GTGGAATCTA GAGCCCCGGGG AACTGATAC  
 3501 ATCGCCAGAT GTAGTTGCAG GTGTGGAGTT TACTATTGTG AATCTAGGAG GAAGTACTAC CCTGTCACTT  
 3571 TTACTGGCCC AACGTTCCGA TTCAATGAAAG CAAACGACTA CTATCCAGCA AGATACCAAGT CTCACATGCT  
 3641 GATAGGGTGC GGATTGCAAG AACCCGGGAA CTGCGGTGGG ATACTGAGGT GCACTCATGG GGTAAATTGGT  
 3711 ATCATTACTG CAGGAGGTGA AGGGTAGTA GCCTTGTG ACATTAGAGA CCTCTGGGTG TATGAAGAGG  
 3781 AGGCCATGGA ACAGGGAATA ACAAGCTACA TCGAATCTCT CGGCACAGCC TTTGGCGCAG GGTCACCCA  
 3851 CACAATCAGT GAGAAAGTGA CTGAATTGAC AACGATGGTT ACCAGCACTA TCACAGAAAA ACTACTGAAA  
 3921 AACTTGGTGA AAATAGTGTGTC GGCTCTAGTG ATTGTTGTGA GAAATTATGA GGACACTACC ACGATCCTTG  
 3991 CAACACTAGC ACTACTCGGG TGTGATATAT CTCCCTGGCA ATGGTTGAAG AAGAAGGCAT GTGACTTACT

FIG. 8 (Continued)

4061 AGAGATTCT TATGTGATGC GCCAAGGTGA TGGGTGGATG AAGAAATTCA CAGAGGCCGTG CAATGCAGCT  
 4131 AAAGGCTTAG AGTGGATTAG CAACAAAATT TCCAAGTTA TAGATTGGTT GAAGTGTAAA ATTATCCAG  
 4201 ACGCTAAGGA CAAGGTGGAA TTTCTCACCA AGTTGAAACA GCTAGACATG TTGGAAAATC AAATTGCAAC  
 4271 CATCCACCAA TCTTGCCCCA GCCAAGAAC ACAAGAGATT CTTTCAACA ATGTGAGATG GTTAGCAGTC  
 4341 CAGTCCCGTC GGTTTGCACC ATTATACGCT GTGGAGGCAC GCCGAATTAA CAAAATGGAG AGCACAATAA  
 4411 ACAATTATAT ACAGTTCAAG AGCAAAACACC GTATTGAACC AGTATGTATG CTCATTATG GTTCACCCAGG  
 4481 GACGGGTAAA TCTATAGCTA CTTCATTAAT AGGTAGAGCA ATAGCAGAGA AGGAAAGCAC ATCAGTCTAT  
 4551 TCAATGCCAC CTGACCCATC TCACTTTGAT GGCTATAAAC AACAAAGGGT AGTGTGATTATG GACGACCTAA  
 4621 ACCAAAACCC CGATGGTATG GACATGAAAC TGTTTGCCTA AATGGTATCA ACAGTGGAGT TTATTCCCTCC  
 4691 AATGGCCTCA TTAGAGGAGA AGGGCATT TTGTTACATCT GATTATGTCC TGGCTTCTAC CAACTCTCAT  
 4761 TCAATTGTCAC CACCCACAGT GGCTCACAGT GATGCCTTAA CCAGACGATT TGCATTGAT GTGGAGGTTT  
 4831 ACACGATGTC TGAACATTCA GTCAAAGGA AACTGAATAT GGCCACGGCC ACTCAATTGT GTAAGGATTG  
 4901 TCCAACACCT GCAAATTITA AAAAGTGTG CCTCTCGTT TGTGGAAAAG CCTTGCATT AATGGACAGG  
 4971 TACACAGAC AAAGGTTCAC TGAGATGAG ATTACCATAT TAATCATGAA TGAGAAAAAC AGAAGGGCCA  
 5041 ATATCGGCAA TTGATGGAA GCCTTGTTC AAGGACCACT AAGGTATAAA GATTGAGA TCGATGTGAA  
 5111 GACAGTTCCC CCCCCTGAGT GCATCACTGA TTGTTACAA GCAGTGGATT CTCAAGAGGT TAGGGATTAC  
 5181 TGTGAGAAGA AAGGCTGGAT CGTTAACGTT ACTAGCCAGA TTCAACTAGA AAGGAACATC AATAGGGCCA  
 5251 TGACTATAC TCAAGCTGTT ACCACATTG CAGCAGTCGC AGGAGTAGTG TATGTAATGT ACAAACTCTT  
 5321 CGCCGGTCAA CAGGGTGCAT AACTGGCTT GCCAAACAAA AAACCCAATG TCCCTACTAT CAGAGTCGCT  
 5391 AAAGTCCAGG GGCCAGGATT TGACTACGCA GTGGCAATG CAAAAAGAAA CATACTTACT GCAACCCACCA  
 5461 CCAAGGGTGA ATTACCATG CTAGGGTGC ATGATAATGT AGCAATATTG CCAACCCATG CCGCTCCAGG  
 5531 AGAAAACATT ATTATTGATG GGAAAGAAGT AGAGATCCTA GACGCCAGAG CCTTAGAAGA TCAAGCGGGA  
 5601 ACCAATCTTG AGATCACCAT TATTACTTA AAAAGAATG AGAAGTTAG AGACATCAGA TCACATATTIC  
 5671 CCACCCAAAT TACTGAAACT AACGATGGAG TGTGATCGT GAACACTAGC AAGTACCCCA ATATGTATGT  
 5741 CCCCCGGTGT GCTGTGACCG AACAGGGATA TCTTAATCTC AGTGGACGTC AAACGTCTG CACTTAATG  
 5811 TACAACCTTC CAACAAGGGC AGGCCAGTC GGAGGAATCA TCACTTGATC TGGCAAAGTC ATTGGGATG  
 5881 ATGTTGGCGG GAACGGTTCA CATGGGTTG CTGCAGCCCT CAAGCGATCA TACTTCACTC AAAATCAGGG  
 5951 CGAAATCCAG TGGATGAGGT CATCAAAGA AGTGGGGTAC CCCATTATAA ATGCCCATC CAAGACAAAG  
 6021 TTAGAACCCA GTGCTTCCA CTATGTTTT GAAGGTGTTA AGGAACCAGC TGTACTCACT AAGAATGACC  
 6091 CCAAGACTAAA AACAGATTIT GAAGAAGCCA TCTTTCTAA ATATGTGGG AACAAAATTAA CTGAAGTGG  
 6161 CGAGTACATG AAAGAACAGC TGGATCACTA TGCAGGACAG TTAATGTCAC TGGATATCAA CACAGAACAG  
 6231 ATGTCCTGG AGGATGCCAT GTACGGTACG GATGGCTTGG AGGCCCTGGA TCTTAGCCT AGTGTGGAT  
 6301 ATCCITATGT TGCATGGGG AAAAGAAAAA GAGACATCT AGATAAACAG ACCAGAGATA CTAAGGAGAT  
 6371 CGAGAGACTT TTAGATACCT ATGGAATCAA TCTACCATTA GTCACGTACG TGAAAAGATGA ACTCAGGTCA  
 6441 AAGACTAAAG TGGAAACAGG AAAGTCAAGA TTGATTGAAAG CTTCCAGCT TAATGATICA GTTGAATG  
 6511 GAATGGCCTT TGGCAATCTT TACGCAGCTT TCCACAAGAA TCCAGGTGTG GTGACAGGAT CAGCAGTTGG  
 6581 TTGTGACCCA GATTGTTTT GGAGTAAGAT ACCAGTGCTA ATGGAAGAAA AACTCTTCGC TTTTACTAC  
 6651 ACAGGGTATG ATGCCCACT CAGCCCTGCT TGGTTGAAG CTCTTAAAT GGTGTTAGAA AAAATTGGAT  
 6721 TTGGCAGTAG AGTAGACTAT ATAGACTACC TGAACCACTC TCACCACCT TACAAAAACA AGACTTATTG  
 6791 TGTCAAAGGC GGCATGCCAT CCGGCTGCTC TGGCACCTCA ATTTCAACT CAATGATTAA CAACCTGATC  
 6861 ATTAGGACGC TTTACTGAG AACCTACAAG GGCATAGACT TGGACCATT AAAAATGATT GCCTATGGTG  
 6931 ATGACGTGAT AGCTCTCTAC CCCCATGAGG TTGACGTAG TCTCTACCC CAATCAGGAA AGACTATGG  
 7001 ACTAACCATG ACTCCAGCAG ATAATACAGC AACCTTTGAA ACAGTCACAT GGGAGAATGT AACATTCTG  
 7071 AAAAGATTIT TCAGAGCAGA TGAGAAGTAT CCATTCTGG TGACATCCAGT GATGCCATG AAAGAAATTC  
 7141 ACGAATCAAT CAGATGGACC AAGGACCCCTA GAAACACACA GGATCACGTA CGCTCGTTGT GCCTATTAGC  
 7211 TTGCAACAC CGTGAAGAAG AATAACATAA ATTTTAGCT AAAATCAGAA GTGTGCAAT CGGAAGAGCT  
 7281 TTATGCTCC CAGAGTACTC TACATTGATC CGCCGATGGC TCGACTCATT TTAGTAACCC TACCTCAGTC  
 7351 GGATTGGATT GGGTTATACT GTTGTAGGGG TAAATTTTC TTTAATTGCG AGAAAAAAA AAAAAAAA  
 7421 AGAGCTCCA ATCACTAGTG AATTCCGCGC CGCCTGCAGG TCGACCATAT GGGAGAGCTC CCAACCGTT  
 7491 GGATGCATAG CTTGAGTATT CTATAGTGT ACCTAAATAG CTTGGCGTAA TCATGGTCAT AGCTGTTCC  
 7561 TGTGTGAAAT TGTATCCGC TCACAATTCC ACACAACATA CGAGCCGAA GCATAAAAGTG TAAAGCCTGG  
 7631 GGTGCCATAAT GAGTGAGCTA ACTCACATTA ATTGCGTTCG GCTCACTGCC CGCTTCCAG TCGGGAAACC  
 7701 TGTGCGCCA GCTGCATTA TGAATCGGCC AACGCGCGGG GAGAGGCGGT TTGCGTATTG GGCCTCTTC  
 7771 CGCTCTCG CTCACTGACT CGCTCGCTC GGTGCGTTCGG CTGCGCGAG CGGTATCAGC TCACTCAAAG  
 7841 CGGGTAATAC GGTTATCCAC AGAATCAGGG GATAACGAG GAAAGAACAT GTGAGAAAAA GGCCAGCAAA  
 7911 AGGCCAGGAA CCGTAAAAG GCCCGCTTGC TGGCTTTT CCATAGGCTC CGCCCCCTG ACGAGCATCA  
 7981 CAAAAATCGA CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC GTTCCCCCT  
 8051 GGAAGCTCCC TCGTGCCTC TCCGTGTTCCG ACCCTGCCGC TTACCGGATA CCTGTCGCC TTTCTCCCTT  
 8121 CGGGAAAGCGT GGCCTTCT CATAGCTCAC GCTGTAGGTA TCTCAGTTCG GTGTAGGTCG TTCGCTCCAA

FIG. 8 (Continued)

8191 GCTGGGCTGT GTGCACGAAC CCCCCGTTCA GCCCGACCGC TCGCCCTTAT CCGGTAACTA TCGTCTTGAG  
8261 TCCAACCCGG TAAGACACGA CTTATCGCCA CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT  
8331 ATGTAGGCGG TGCTACAGAG TTCTTGAAGT GGTGGCTAA CTACGGCTAC ACTAGAAGAA CAGTATTGG  
8401 TATCTGCGCT CTGCTGAAGC CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG CAAACAAACC  
8471 ACCGCTGGTA GCGGTGGTTT TTTTGTTCG AAGCAGCAGA TTACGCGCAG AAAAAAAAGGA TCTCAAGAAG  
8541 ATCCTTIGAT CTTTCTACG GGGTCTGACG CTCAGTGGAA CGAAAACTCA CGITAAGGGA TTTGGTCAT  
8611 GAGATTATCA AAAAGGATCT TCACCTAGAT CCTTTAAAT TAAAAATGAA GTTTAAATC AATCTAAAGT  
8681 ATATATGAGT AAACCTGGTC TGACAGTTAC CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC  
8751 TATTTCGTTT ATCCATAGTT GCCTGACTCC CGCTCGTGTAA GATAACTACG ATACGGGAGG GCTTACCATC  
8821 TGGCCCCAGT GCTGAATGTA TACCGCGAGA CCCACGCTCA CGCGCTCCAG ATTATCAGC AATAAACCCAG  
8891 CCAGCCGGAA GGCGCAGCG CAGAAGTGGT CCTGCAACTT TATCCGCTC CATCCAGTCT ATTAAATTGTT  
8961 GCGGGGAAGC TAGAGTAAGT AGTTGCCAG TTAATAGTTT GCGCAACGTT GTTGCATTG CTACAGGCAT  
9031 CGTGGTGTCA CGCTCGTCCGTTGGTATGGC TTCAATCAGC TCCGGTCCC AACGATCAAG GCGAGTTACA  
9101 TGATCCCCCA TGTGTGCAA AAAAGCGGTT AGCTCCTCG GTCCTCCGAT CGTGTGAGA AGTAAGTTGG  
9171 CCGCAGTGTG ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CGTAAGATG  
9241 CTTTCTGTG ACTGGTGAGT ACTCAACCAA GTCAATTCTGA GAATAGTGTG TGCAGCGACC GAGTTGCTCT  
9311 TGCCCGCGT CAATACGGGA TAATACCGCG CCACATAGCA GAACTTAAA AGTGTCTCATC ATIGGAAAAC  
9381 GTTCTTCGGG GCGAAAACCTC TCAAGGATCT TACCGCTGTT GAGATCCAGT TCGATGTAAC CCACTCGTGC  
9451 ACCAAACTGA TCTTCAGCAT CTTTACTTT CACCAAGCGTT TCTGGGTGAG CAAAAACAGG AAGGAAAAT  
9521 GCGCAAAAAA AGGGATAAAG GGCAGCACCGG AAATGTTGAA TACTCTACT CTTCTTTT CAATATT  
9591 GAAAGATTAA TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTGAATGT ATTAGAAAA ATAACAAAT  
9661 AGGGGTTCCG CGCACATTTC CCCGAAAAGT GCCACCTGAT GCGGFTGAA ATACCGACA GATGCGTAAG  
9731 GAGAAAATAC CGCATCAGGA AATTGTAAGC GTTAATATIT TGTTAAAATT CGCGTTAAAT TTTGTTAAA  
9801 TCAGCTCATT TTAAACCAA TAGGCCAAA TCGCAAAAT CCCITATAAA TCAAAGAAT AGACCGAGAT  
9871 AGGGTTGAGT GTTGTCCAG TTTGGAAACAA GAGTCCACTA TTAAAGAAGC TGGACTCCAA CGTCAAAGGG  
9941 CGAAAAACCG TCTATCAGGG CGATGCCCA CTACGTGAAC CATCACCTA ATCAAGTTT TTGGGGTCGA  
10011 GGTGCGTAA AGCACTAAAT CGGAACCCCTA AAGGGAGCCC CCGATTAGA GCTTGACCGG GAAAGCCGGC  
10081 GAACGTGGCG AGAAAGGAAG GGAAGAAAGC GAAAGGAGCG GGGGCTAGGG CGCTGGCAAG TGTAGCGGTC  
10151 ACGCTGCGCG TAACCACAC ACCCGCCGCG CTTAATGCGC CGCTACAGGG CGCGTCCATT CGCATTCA  
10221 GCTGCGCAAC TGTGGGAAG GGCAGCTGGT GCGGGCCTCT TCGCTATTAC GCCAGCTGGC GAAAGGGGGA  
10291 TGTGCTGCAA GGCAGTTAAG TTGGGTAACG CCAGGGTTT CCCAGTCACG ACGTTGAAA ACGACGGCCA  
10361 GTGAATTGTA ATACGACTCA CTATAAGGGCG AATTGGGCC GACGTGCGAT GCTCCGGCC GCCATGGCGG  
10431 CGCGGGAAAT TCGATTGAGG CATGCTAATA CGACTCACTA TAGG

FIG. 9

1 TTAAAACAGC TCTGGGGITG TTCCCACCCCC AGAGGCCAC GTGGCGGCTA GTACTCTGGT ATTACGGTAC  
 71 CTITGTACGC CTGTTTGTA TCCCTTCCCC CGTAACTTAA GAAGCTTATC AAAAGTCAA TAGCAGGGGT  
 141 ACAAACCACT ACCTCTACGA ACAAGCACTT CTGTTTCCCC GGTGATATCA CATAGACTGT ACCCACGGTC  
 211 AAAAGTGATT GATCCGTAT CCGCTTGAGT ACTTCGAGAA GCCTAGTATC ACCTTGAAT CTTCGATGCG  
 281 TTGCGCTCAA CACTCTGCC CGAGTGTAGC TTAGGCTGAT GAGTCTGGC ACTCCCCACC GGCGACGGTG  
 351 GCCCAGGCTG CGTGGCGGC CTACCCATGG CTGATGCCGT GGGACGCTAG TTGTAACAA GGTGTGAAGA  
 421 GCCTATTGAG CTACTCAAGA GTCCCTCCGGC CCCTGAATGC GGCTAATCT AACCAACGGAG CAACCGCTCA  
 491 CAACCCAGTG AGTAGGTTGT CGTAATGCGT AAGTCTGTGG CGGAACGGAC TACTTGGGT GTCCGTGTT  
 561 CCCTTATAT TCATACTGGC TGCTTATGGT GACAATTAC AAATTGTAC CATATAGCTA TTGGATTGGC  
 631 GGCGCGACA GCTGGTTGAA GGGGACCAAC GATACAGCTG GTGIAAGGGG ACCAACTGGA GCCACACACT  
 701 TCCTTACATT CCATCACCCA CACACTCCT TACATTCCAC GGCGCGCCAA TAAACAGTGC AAATGGGGC  
 771 TCAAGTTICA ACGCAAAAGA CCGGTGCGCA CGAGAATCAA AACGTGCGAG CCAATGGATC CACCATTAAT  
 841 TACACTACTA TCAACTATTA CAAAGACAGT GCGAGTAATT CCGCTACTAG ACAAGACCTC TCCCAAGATC  
 911 CATCAAAATT CACAGAACCG GTTAAGGACT TAATGTGAA AACAGCACCA GCTCTAACT CGCTAACCGT  
 981 GGAAGCATGT GGGTACAGTG ACCGTGTGAG GCAAATCACT TTAGGCAACT CGACTATTAC TACACAAGAA  
 1051 GCAGCCAATG CTATTGTGC TTACGGTGA TGCCCACCT ACATAATGA TTCAGAAGCT ATCCGGTAG  
 1121 ATGCACCCAC TGAGCCAGAC GTTAGTAGCA ACCGGTTITA CACCCTAGAA TCGGGTGTCTT GGAAGACAC  
 1191 TTCAAGGGGA TGTTGGTGA AGTTACAGA TTGTTGAAAG GACATGGGAA TGTTGGTCA GAATATGTAC  
 1261 TATCACTACT TGGGGCGCTC TGGTTACACC ATTACATGTCC AGTGCAACGC TTCAAAATT CACCAAGGGG  
 1331 CGTITAGGAT TTCTGTATA CCAGAGTTG TCATGGCTG CAACACTGAG AGTAAAACGT CATACTTIC  
 1401 ATACATCAAT GCAAAATCTG GTGAGAGAGG CGGTGAGITT ACGAACACCT ACAATCCGTC AAATACAGAC  
 1471 GCCAGTGAGG GCAGAAAGTT TGCACTTGC GATTATTGC TGGGTTCTGG TGTTCTAGCA GGAACACGCT  
 1541 TTGTTGACCC GCACCAAGATC ATCAACCTAC GTACCAACAA CAGTGCACAA ATTGTGGTGC CATACTAA  
 1611 CTCACCTGTG ATTGATTGTA TGGCAAAACA CAATAACTGG GGCATTGTCA TATTACCACT GGCACCCCTG  
 1681 GCCTTGCGC CAACATCGTC ACCACAGGTG CCTATTACAG TGACCATGTC ACCCATGTGT ACAGAATTCA  
 1751 ATGGGTTGAG AAACATCACC GTCCCAGTAC ATCAAGGGT GCGGACAATG AACACACCTG GITCCAATCA  
 1821 ATTCCCTACA TCTGATGACT TCCAGTCGCC CTGTGCCITA CCTAATTG ATGTACTCC ACCAATACAC  
 1891 ATACCCGGGG AAGTAAAGAA TATGATGGAA CTAGCTGAAA TTGACACATT GATCCAATG AACGCAGTGG  
 1961 ACGGGAAGGT GAACACAATG GAGATGTATC AAATACCATT GAATGACAAT TTGAGCAAGG CACCTATAATT  
 2031 CTGTTATCC CTATCACCTG CTTCTGATAA ACGACTGAGC CACACCATGT TGGTGAAT CCTAAATTAT  
 2101 TACACCCATT GGACGGGGTC CATCAGGTTC ACCTTCTAT TTGTTGGCAG TATGATGGCC ACTGGTAAAC  
 2171 TGCTCTCAG CTATCCCCA CCGGGAGCTA AACACACCAAC CAATCGCAAG GATGCAATGC TAGGCACACA  
 2241 CATCATCTGG GACCTAGGGT TACAATCCAG TTGTCATG GTGACCCGT GGATCTCAA CACAGTGTAC  
 2311 AGACGGTGTG CACGTGATGA CTTCACTGAG GGGGGATTAA TAACCTGCTT CTATCAAAC AGAATTGTGG  
 2381 TACCTGCTTC AACCCTTACCA AGTATGTICA TGTAGGCTT TGTAGTGTG TGTCAGACT TCAGTGTAC  
 2451 ACTGCTTAGG GACACTCCCC ATATTAGTCA ATCGAAACTA ATAGGACGTA CACAAGGCAT TGAAGACCTC  
 2521 ATTGACACAG CGATAAAGAA TGCCTTAAGA GTGCTTAAAC CACCCCGAC CCAGTCACACT GAAGCAACTA  
 2591 GTGGAGTGA TAGCCAGGAG GTGCCAGCTC TAATGCTGT GGAAACAGGA GCATCTGGTC AAGCAATCCC  
 2661 CAGTGTGTTG GTGGAAACTA GGCACGTGGT AAATTACAAA ACCAGGTCTG AATCGTGTCT TGAGTCATT  
 2731 TTGGGAGAG CTGCGTGTGT CACAATCTCA TCCCTGACCA ACTCCTCCAA GAGCGGAGAG GAGAAAAAGC  
 2801 ATTTCACAT ATGGAATATT ACATACACCG ACACTGCTCA GTACGAGA AAATTAGAAT TTTCACGTA  
 2871 TTCCAGGTTT GATCTGAAA TGACTTTGT ATTACACAGAG AACTATCTA GTACAGCCAG TGGAGAAGTG  
 2941 CGAACACAGG TGTAACAGAT CATGTATATT CCACAGGGG CACCCCGCC ATCATCCTGG GATGACTACA  
 3011 CATCGCAATC CTCTTCAAAC CCTTCATCT TCTACATGTA TGGAAATGCA CCTCCACGGA TGTCATTC  
 3081 TTACGTAGGG ATTGCCAATG CCTATTACCA CCTCTACGAT GGCTTGAC GGGTGCCACT TGAGGGTGTG  
 3151 AACACCGATG CTGGCGACAC GTTTACGGT TTAGTGTCCA TAAATGATT TGGAGTTITA GCAGITAGAG  
 3221 CAGTAAACCG CAGTAATCCA CATAACAATAC ACACATCTGT GAGAGTGTAC ATGAAACCAA AACACATTG  
 3291 GTGTGGTGC CCCAGACCTC CTCGAGCTGT ATTACACAGG GGAGAGGGAG TGGACATGAT ATCCAGTGC  
 3361 ATTCTACCTC TGGCCAAGGT AGACTCAATT ACCACTTTG GGTTGGTCA TCAGAACAAA GCAGTGTACG  
 3431 TTGCGGGTTA CAAGAATTGC AACTACCAAC TAGCAACCCC AAGTGTACAC TTGAATGCAA TTAGTATGTT  
 3501 ATGGGACAGG GATTTAATGG TGGTGAATC TAGAGCCCGG GGAACGTATA CCATCGCCAG ATGTAGTTGC  
 3571 AGGTGTGGAG TTACTATTG TGAATCTAGG AGGAAGTACT ACCCTGTCAC TTTCATGGC CCAACGTTTC  
 3641 GATTCAATGGA AGCAACACGAC TACTATCCAG CAAGATACCA GTCTCACATG CTGATAGGGT GCGGATTG  
 3711 AGAACCCGGG GACTGCGGTG GGATACTGAG GTCTCACTCAT GGGGTATTG GTATCATTAC TGCAGGAGGT  
 3781 GAAGGGTAG TAGCCTTGC TGACATTAGA GACCTCTGGG TGTATGAAGA GGAGGCCATG GAACAGGGAA  
 3851 TAACAAGCTA CATCGAATCT CTCGGCACAG CCTTGGCGC AGGGTTCAAC CACACAATCA GTGAGAAAGT

FIG. 9 (Continued)

3921 GACTGAATTG ACAACGATGG TTACCAGCAC TATCACAGAA AAACTACTGA AAAACTTGGT GAAAATAGTG  
 3991 TCGGCTCTAG TGATTGTTGT GAGAAATTAT GAGGACACTA CCACGATCCT TGCAACACTA GCACACTCG  
 4061 GGTGTGATAT ATCTCCTTGG CAATGGTGA AGAAGAAGGC ATGTGACTTA CTAGAGATT CTTATGTGAT  
 4131 GCGCCAAGGT GATGGGTGGA TGAAGAAATT CACAGAGGCG TGCAATGCAG CTAAAGGCTT AGAGTGGATT  
 4201 AGCAACAAAA TTTCCAAGTT TATAGATTGG TTGAAGTGTAA AATTATCCC AGACGCTAAG GACAAGGTGG  
 4271 AATTCTCAC CAAGTGTGAAA CAGCTAGACA TGTGGAAAA TCAAATTGCA ACCATCCACC AATCTTGC  
 4341 CAGCCAAGAA CAACAAGAGA TTCTTTCAA CAATGTGAGA TGGITAGCAG TCCAGTCCCG TCGGTTGCA  
 4411 CCATTACAG CTGCGAGGCC ACAGCGGAATT AACAAAATGG AGAGCACAAAT AAACAATTAT ATACAGTCA  
 4481 AGAGCAAACA CCGTATTGAA CCAGTATGTA TGCTCATTCA TGGGTACCA GGGACGGGTA AATCTATAGC  
 4551 TACTTCITA ATAGGTAGAG CAATAGCAGA GAAGGAAAGC ACATCAGTCT ATTCAATGCC ACCTGACCCA  
 4621 TCTCACITTG ATGGCTATAA ACAACAAGGG GTAGTGTAA TGGACGACCT AAACCAAAAC CCCGATGGTA  
 4691 TGGACATGAA ACTGTTTGC CAAATGGTAT CAACAGTGGA GTTTATTCTT CCAATGGCT CATTAGAGGA  
 4761 GAAGGGCATT TTGTTTACAT CTGATTATGT CCTGGCTTCT ACCAACCTCTC ATTCAATTGT ACCACCCACA  
 4831 GTGGCTCACA GTGATGCCTT AACCAGACGA TTTGCATTG ATGTGGAGGT TTACACGATG TCTGAACATT  
 4901 CAGTCAAAGG CAAACTGAAT ATGGCCACGG CCACTCAATT GTGTAAGGAT TGTCAACAC CTGCAAATT  
 4971 TAAAAAGTGT TGCCCTCTCG TTGTTGAAAGA GGCTTGCAA TTAATGGACA GTGACACCAG ACAAAGGTT  
 5041 ACTGTAGATG AGATTACCACTTAAATCATG AATGAGAAAA ACAGAAGGGC CAATATCGGC AATTGCATGG  
 5111 AAGCCTTGT TCAAGGACCA CTAAGGTATA AAGATTGAA GATCGATGTG AAGACAGTTC CCCCCCTGA  
 5181 GTGCATCAGT GATTGTTAC AAGCAGTGGA TTCTCAAGAG GTTAGGGATT ACTGTGAGAA GAAAGGCTGG  
 5251 ATCGTTAACG TTACTAGCCA GATTCAACTA GAAAGGAACA TCAATAGGGC CATGACTATA CTCCAAGCTG  
 5321 TTACACATT CGCAGCAGTC GCAGGAGTAG TGTATGTAA TGTACAAACTC TTCGCCGGTC AACAGGGTGC  
 5391 ATACACTGGC TTGCCAAACA AAAAACCAA TGTCCCTACT ATCAGAGTCG CTAAAGTCCA GGGGCCAGGA  
 5461 TTGACTACG CAGTGGCAAT GGCAAAAAGA AACATAGTTA CTGCAACCAC CACCAAGGGT GAATTTCACCA  
 5531 TGCTAGGGGT GCATGATAAT GTAGCAATAT TGCCAACCCA TGCGCTCCA GGAGAAACCA TTATTATIGA  
 5601 TGGGAAAGAA GTAGAGATCC TAGACGCCAG AGCCTTAGAA GATCAAGCGG GAACCAATCT TGAGATCACC  
 5671 ATTATTACTC TAAAAAGAAA TGAGAAGTTT AGAGACATCA GATCACATAT TCCCACCCAA ATTACTGAAA  
 5741 CTAACGATGG AGTGTGATC GTGAACACTA GCAAGTACCC CAATATGTAT GTCCCGTTG GTGCTGTGAC  
 5811 CGAACAGGGAA TATCTTAATC TCAGTGGACG TCAAACCTGT CGCACTTAA TGTCACACTT TCCAACAAGG  
 5881 CGAGGCGAGT GCGGGAGGAAT CTCACCTGT ACTGGCAAG TGTTGGGAT GCTATGGC GGGAACGGTT  
 5951 CACATGGTT TGCTGCAGCC CTCAGCGAT CATACTTCAC TCAAATCAG GGCAGAACATCC AGTGGATGAG  
 6021 GTCATCAAAA GAAGTGGGGT ACCCCATTAAATGCCCCA TCCAAGACAA AGTACAAGACC CAGTGCTTTC  
 6091 CACTATGTT TTGAAGGTGT TAAGGAACCA GCTGTACTCA CTAAGAATGA CCCCCAGACTA AAAACAGATI  
 6161 TTGAAGAAGC CATCTTCTTAAATATGTGG GGAACAAAAT TACTGAAGTG GACGAGTACA TGAAAGAAGC  
 6231 AGTGGATCAC TATGCAGGAC AGTTAATGTC ACTGGATATC AACACAGAAC AGATGTGCCT GGAGGATGCC  
 6301 ATGTACGGTA CCGATGGTCT TGAGGCCCTG GATCTTAGCA CTAGTGTGG ATATCCTTAT GTGCAATGG  
 6371 GGAAAAGAA AAGAGACATT CTAGATAAAC AGACCAAGAGA TACTAAGGAG ATGCAGAGAC TTTIAGATAC  
 6441 CTATGGAATC ATCTTACCAT TAGTCACGTA CGTCAAAGAT GAACTCAGGT CAAAGACTAA AGTGGAAACAA  
 6511 GGAAAGTCAA GATTGATTGA AGCTTCCAGC CTTAATGATT CAGTGTCAAT GAGAATGGC TTGGCAATC  
 6581 TTACCGCAGC TTCCACAAAG AATCCAGGTG TGTTGACAGG ATCAGCAGT GGTGGTGC CAGATTGTT  
 6651 TTGGAGTAAG ATACCGATGCA TAATGGAGA AAAACTCTTC GCTTTGACT ACACAGGGTA TGATGCTCA  
 6721 CTCAGCCCTG CTGGTTTGA AGCTCTTAAATGTTGTTAG AAAAATTGG ATTGCGAGT AGAGTAGACT  
 6791 ATATAGACTA CCTGAACCAC TCTCACCACCTTACAAAAAA CAAGACTTAT TGTTGCTAAAG GCGGCATGCC  
 6861 ATCCGGCTGC TCTGGCACCT CAATTITCAA CTCAATGATT AACAAACCTGA TCATTAGGAC GCTTITACTG  
 6931 AGAACCTACA AGGGCATAGA CTTGGACCAT TTAAAAATGA TTGCTTATGG TGATGACGTG ATAGCTTCT  
 7001 ACCCCATGA GTTGACGCT AGTCTCTAG CCCAATCAGG AAAAGACTAT GGACTAACCA TGACTCCAGC  
 7071 AGATAAAATCA GCAACCTTIG AAACAGTCAC ATGGGAGAAAT GTAACATTIC TGAAAAGATT TTTCAGAGCA  
 7141 GATGAGAAATC ATCCATTCTC GGTGCATCCA GTGATGCCAA TGAAAGAAAT TCACGAAATCA ATCAGATGG  
 7211 CCAAGGCCAGGATCAGGATCAGGACAGGATCAGGACAGGATCAGGACAGGATCAGGACAGGATCAGGACAGG  
 7281 AGAATACAAT AAATTAGCTTCTAATGTTAG CTCACCTGAAAGTGTGCAATCAGGAGAG CTTTATTGCT CCCAGAGTAC  
 7351 TCTACATTGT ACCGCCAGT GCTGACTCA TTGTTAGTAACTCCTACCTGAGCAGGATGGTAAAGGTTATA  
 7421 CTGTTGTTAGG GTGAAATTTC TCTTAAATC GGAGAAAAAA AAAAAGAGCTCC CAATCACTAG  
 7491 TGAATTGCGC GCCGCCCTGCA GGTGACCAT ATGGGAGAGC TCCCAACCGCG TTGGATGCAT AGCTTGAGTA  
 7561 TTCTATAGTG TCACCTAAAT AGCTTGGCGT AATCATGGTC ATAGCTGTT CCTGTGTGAA ATTGTTATCC  
 7631 GCTCACAATT CCACACAACA TACGAGCCGG AAGCATAAAG TGAAAGCCT GGGGTGCCTA ATGAGTGCAG  
 7701 TAACTCACAT TAATTGCGTT GCGCTCACTG CCCGCTTCC AGTCGGGAAA CCTGTCGTGC CAGCTGCATT  
 7771 AATGAATCGG CCAACGCGCG GGGAGAGGCG GTTGTGCTAT TGGGCGCTCT TCCGCTCTCG CGCTCACTG  
 7841 CTCGCTGCAGC TCGGTCGTTT GGCTGCCGGCG AGCGGTATCA GCTCACTCAA AGGCAGTAAT ACAGGTTATCC  
 7911 ACAGAATCAG GGGATAACGC AGGAAAGAAC ATGTGAGCAA AAGGCCAGCA AAAGGCCAGG AACCGTAAAA  
 7981 AGGCCCGTGT GCTGGCGTTT TTCCATAGGC TCCGCCCTCC TGACGAGCAT CACAAAATC GACGCTCAAG  
 8051 TCAGAGGTGG CGAAACCCGA CAGGACTATA AAGATACCAAG GCGTTTCCCC CTGGAAGCTC CCTCGTGCAG

FIG. 9 (Continued)

8121 TCTCCTGTT CGACCCTGCC GCTTACCGGA TACCTGTCCG CCTTCTCCC TTCGGGAAGC GTGGCGCTT  
 8191 CTCATAGCTC ACGCTGTAGG TATCTCAGTT CGGTGTAGGT CGTTCGCTCC AAGCTGGGCT GTGTGCACGA  
 8261 ACCCCCCGTT CAGCCCGACC GCTGCGCTT ATCCGTAAC TATCGTCTTG AGTCCAACCC GGTAAAGACAC  
 8331 GACTTATCGC CACTGGCAGC AGCCACTGGT AACAGGATTAA GCAGAGCGAG GTATGTAGGC GGTGCTACAG  
 8401 AGTTCTGAA GTGGTGGCCT AACTACGGCT ACACAGAAG AACAGTATTT GGATCTGCG CTCTGCTGAA  
 8471 GCCAGTTAACCTCGGAAAAA GAGTTGGTAG CTCTTGATCC GGCAAACAAA CCACCGCTGG TAGCGGTGGT  
 8541 TTTCCTGTT GCAAGCAGCA GATTACCGCG AGAAAAAAAG GATCTCAAGA AGATCCITTG ATCTTCTA  
 8611 CGGGGTCTGA CGCTCAGTGG AACGAAAAGT CACGTTAAGG GATTTGGTC ATGAGATTAT CAAAAGGAT  
 8681 CTTCACCTAG ATCCTTAA ATTAAAAATG AAGTTTAA TCAATCTAAA GTATATATGA GTAAACTTGG  
 8751 TCTGACAGTT ACCAATGCTT AATCAGTGAG GCACCTATCT CAGCGATCTG TCTATTCTGT TCATCCATAG  
 8821 TTGCTCTGACT CCCCGTCGTG TAGATAACTA CGATACGGGA GGGCTTACCA TCTGGCCCCA GTGCTGCAAT  
 8891 GATACCGCGA GACCCACGCT CACCGGCTCC AGATTTATCA GCAATAAACCG AGCCAGCCGG AAGGGCCGAG  
 8961 CGCAGAAGTG GTCTGCAAC TTTATCGCC TCCATCCAGT CTATTAATTG TTGCCCCGGAA GCTAGAGTAA  
 9031 GTAGTTGCC AGTTAATAGT TTGCGCAACG TTGTTGCCAT TGCTACAGGC ATCGTGGTGT CACGCTCGTC  
 9101 GTTTGGTATG GCTTCATTCA GCTCCGTTCC CCAACGATCA AGGCGAGTCA CATGATCCCC CATGTTGTGC  
 9171 AAAAAGCGG TTAGCTCTT CGGCTCTCG ATCGTTGTCA GAAGTAAGTT GGCCGAGTG TTATCACTCA  
 9241 TGGTTATGGC AGCACTGCAT AATTCTCTTA CTGTCATGCC ATCCGTAAGA TGCTTTCTG TGACTGGTGA  
 9311 GTACTCAACC AAGTCATTCT GAGAATAGTG TATGCGCGA CCGAGITGCT CTTGCCCCGGC GTCAATACGG  
 9381 GATAATACCG CGCCACATAG CAGAACTTTA AAAGTGCTCA TCATTGGAAA ACGTTCTTCG GGGCGAAAAC  
 9451 TCTCAAGGAT CTTACCGCTG TTGAGATCCA GTTCGATGTA ACCCACTCGT GCACCCAACG GATCTTCAGC  
 9521 ATCTTCTACT TTACCAAGCG TTTCGGGTG AGCAAAAACA GGAAGGCAAA ATGCCGCAAA AAAGGGAAATA  
 9591 AGGGCGACAC GGAAATGTTG AATACTCATA CTCTCTTIT TTCATATTAA TTGAAGCATT TATCAGGGTT  
 9661 ATTGTCTCAT GAGCGGATAC ATATTTGAAT GTATTTAGAA AAATAAACAA ATAGGGGTTTC CGCGCACATT  
 9731 TCCCCGAAAA GTGCCACCTG ATGCGGTGTG AAATACCGCA CAGATGCGTA AGGAGAAAAT ACCGCATCAG  
 9801 GAAATTGTAAGCGTTAATTTTAAATTTGTTAA ATTCGCTTAA ATTTCGTTA AATCAGCTCA TTTTTAAC  
 9871 AATAGGCCGA ATACGGCAAA ATCCCTATA AATCAAAAGA ATAGACCGAG ATAGGGTTGA GTGTTGTTCC  
 9941 AGTTGGAAC AAGAGTCCAC TATTAAGAA CGTGGACTCC AACGTCAAGGGCGAAAAAC CGTCTATCAG  
 10011 GGCGATGGCC CACTACGTGA ACCATCACCC TAATCAAGTT TTTGGGGTC GAGGTGCCGT AAAGCACTAA  
 10081 ATCGGAACCC TAAAGGGAGC CCCCCATTTA GAGCTTGACG GGGAAAGCCG GCGAACGTGG CGAGAAAGGA  
 10151 AGGAAGAAA GCGAAAGAGAG CGGGCGCTAG GGCCTGGCA AGTGTAGCGG TCACGCTGCG CGTAACCACC  
 10221 ACACCCGCG CGCITAATGC GCGCTACAG GGCCTGCTCA ITCGCCATTC AGGCTGCGCA ACTGTTGGGA  
 10291 AGGGCGATCG GTGCGGGCCT CTTCGCTATT ACGCCAGCTG GCGAAAGGGG GATGTGCTGC AAGGCATTA  
 10361 AGTTGGTAA CGCCAGGGTT TTCCCATGTC CGACGTTGTA AAACGACGGC CAGTGAATTG TAATACGACT  
 10431 CACTATAGGG CGAATTGGGC CCGACGTCGC ATGCTCCCGG CCGCCATGGC GGCCGCGGGGA ATTGAGTGA  
 10501 GGCATGCTAA TACGACTCAC TATAGG

FIG. 10

1 TTAAAACAGC TCTGGGTTT TTCCCACCCC AGAGGCCAC GTGGCGGCTA GTACTCTGGT ATTACGGTAC  
 71 CTTTGTACGC CTGTTTGTA TCCCTCCCC CGTAACTTTA GAAGCTTATC AAAAGTCAA TAGCAGGGT  
 141 ACAAAACAGT ACCTCTACGA ACAAGCACTT CTGTTCCCC CGTGATATCA CATAGACTGT ACCCACGGTC  
 211 AAAAGTATT GATCCGTAT CCGCTTGAGT ACTTCGAGAA GCCTAGTATC ACCTTGAAT CTTCGATGCG  
 281 TTGCGCTCAA CACTCTGCC CGAGTGTAGC TTAGGCTGAT GAGTCTGGC ACTCCCCACC GGCACGGTG  
 351 GCCCAGGCTG CGTGGCGGC CTACCCATGG CTGATGCCGT GGGACGCTAG TTGTAACAA GGTGTGAAGA  
 421 GCCTATTGAG CTACTCAAGA GTCCCTCCGC CCCTGAATGC GGCTAATCCT AACACAGGAG CAACCGCTCA  
 491 CAACCCAGTG AGTAGGTTG CGTAATGCGT AAGTCTGTGG CGGAACCGAC TACTTGGGT GTCCGTGTT  
 561 CCCTTATAT TCATACTGGC TGCTTATGGT GACAATTAC AAATTGTTAC CATATAGCTA TTGGATTGGC  
 631 CACCCAGTAT TGTCAATAT ATTTGAGTGT TTCTTCTATA AGCCTTATTA ACATGGCGC GCCGCACAGC  
 701 TGGTGAAGG GGACCAACGA TACAGCTGGT TGAAGGGGAC CAACTGGAGC CACACACTTC CTTACATTCC  
 771 ATCACCCACA CACTTCTTA CATTCCACGG CGCGCCACAT TTAAATCAC AATAAACAGT GCAATGGGG  
 841 GCTCAAGTTT CAACGCAAAA GACCGGTGCG CACGAGAATC AAAACGTGGC AGCCAATGGA TCCACCATTA  
 911 ATTACACTAC TATCAACTAT TACAAAGACA GTCGAGTAA TTCCGCTACT AGACAAGACCA TCTCCAAAGA  
 981 TCCATCAAA TTCACAGAAC CGGTTAAGGA CTTAATGTTG AAAACAGCAC CAGCTCTAAA CTCGCCTAAC  
 1051 GTGGAAGCAT GTGGGTACAG TGACCGTGTG AGGCAAATCA CTTAGGCAA CTCGACTATT ACTACACAAG  
 1121 AAGCAGCCAA TGCTATTGTT GCTTACGGTG AATGGCCAC TTACATAAT GATTAGAAG CTAATCCGGT  
 1191 AGATGCACCC ACTGAGCCAG ACGTTAGTAG CAACCGGTT TACACCTAG AATCGGTGTC TTGGAAGACC  
 1261 ACTTCAAGGG GATGGTGGTG GAAGTTACCA GATTGTTGA AGGACATGGG AATGTTGGT CAGAATATGT  
 1331 ACTATCACTA CTTGGGCGC TCTGGTACA CCATTATGT CCAGTGCAC GCTTAAAT TTCACCAAGG  
 1401 GGCGTTAGGA GTTTCTGA TACCAGAGTT TGTATGGCT TGCAACACTG AGAGTAAAC GTCATACGGT  
 1471 TCATACATCA ATGCAAATCC TGGTGAGAGA GGCAGGTGAGT TTACGAACAC CTACAATCCG TCAAATACAG  
 1541 ACGCCAGTGA GGGCAGAAAG TTGCAAGCAT TGATTATTIT GCTGGTTCT GGTGTTCTAG CAGGAAACGC  
 1611 CTTGTGTAC CGCACCAGA TCATCAACCT ACCTTACCAAC AACAGTGCAC CAATGTTGGT GCCATACGTA  
 1681 AACTCACTTG TGATTGATG TATGGAAAAA CACAATAACT GGGGCATGTT CATATTACCA CTGGCACCCCT  
 1751 TGGCCTTGC CGCAACATCG TCACACAGG TGCTTACCAT AGTGACCATT GCACCATGTT GTACAGAATT  
 1821 CAATGGGTTG AGAAACATCA CGCTCCAGT ACATCAAGGG TTGGCGACAA TGAACACACC TTGTTCCAAT  
 1891 CAATTCTTA CATCTGATGA CTTCCAGTCG CCTGTGCT TACCTAATT TGATGTTACT CCACCAATAC  
 1961 ACATACCCGG GGAAGTAAAG AATATGATGG AACTAGCTGA AATTGACACA TTGATCCAA TGAACGCAGT  
 2031 GGACGGGAAG GTAACACAA TGGAGATGTA TCAAATACCA TTGAATGACA ATTGAGCAA GGCACCTATA  
 2101 TTCTGTTAT CCCTATCACC TGCTTCTGAT AAACGACTGA GCCACACCAT GTGGGTGAA ATCCTAAATT  
 2171 ATTACACCA TTGGACGGGG TCCATCAGGT TCACCTTCT ATTTGTTGC AGTATGATGG CCACTGGTAA  
 2241 ACTGCTCTC AGCTATTCCC CACCGGGAGC TAAACCACCA ACCAATGCAC AGGATGCAAT GCTAGGCACA  
 2311 CACATCATCT GGGACCTAGG GTTACAATCC AGTTGTTCCA TGGTGCACC GTGGATCTCC AACACAGTGT  
 2381 ACAGACGGTG TGCACGTGAT GACTTCACTG AGGGCGGATT TATAACTTGC TTCTATCAA CTAGAATTGT  
 2451 GGTACCTGCT TCAACCCCTA CCAGTATGTT CATGTTAGGC TTGTTAGTC CGTGTCCAGA CTTCACTGTC  
 2521 AGACTGCTTA GGGACACTCC CCATATTAGT CAATCGAAAC TAATAGGACG TACACAAGGC ATTGAAGACC  
 2591 TCATGACAC AGCGATAAAAG AATGCTTAA GAGTGTCTCA ACCACCTCG ACCAGTCAC CTGAAGCAAC  
 2661 TAGTGGAGTG AATAGCCAGG AGGTGCCAGC TCTAACTGCT GTGAAACAG GAGCATCTGG TCAAGCAATC  
 2731 CCCACTGATG TGGTGGAAAC TAGGCACGTG GTAAATTACA AAACCCAGGTC TGAATGTTGT CTGAGTCAT  
 2801 TCTTGGGAG AGCTGCGTGT GTCAACATCC TATCCTTGAC CAACTCCTCC AAGAGCGGAG AGGAGAAAAA  
 2871 GCATTCAAC ATATGGAATA TTACATACAC CGACACTGTC CAGTTACGCA GAAAATTAGA ATTTTCACG  
 2941 TATTCCAGGT TTGATCTGAT AATGACTTTT GTATTACAG AGAAACTATCC TAGTACAGCC AGTGGAGAAG  
 3011 TGCACCAAC GGTGTACCAAG ATCATGTATA TTCCACCAGG GGCACCCGC CCATCATCTT GGGATGACTA  
 3081 CACATGGCAA TCCCTCTCAA ACCCTTCCAT TTCTCTACATG TATGGAAATG CACCTCCACG GATGTCAATT  
 3151 CCTTACCGTAG GGATTGCAA TGCTTACATCA CACTTACAG ATGGCTTGC ACGGGTGCCA CTTGAGGGTG  
 3221 AGAACAGCA TGCTGGCGAC ACGTTTACG GTTGTAGTC CATAATGAT TTGAGGTTT TAGCAGTTAG  
 3291 AGCAGTAAAC CGCAGTAAAC CACATACAAT ACACACATCT GTGAGAGTGT ACATGAAACC AAAACACATT  
 3361 CGGTGTTGGT GCGACAGACCC TCCTCGAGCT GTATTACCA GGGGAGGG AGTGGACATG ATATCCAGTG  
 3431 CAATTCTACC TCTGGCCAAG GTAGACTCAA TTACCAATT TTGTTGGTGT CATCAGAACAA AAGCAGTGT  
 3501 CGTTGCCGT TACAAGATT GCAACTACCA CCTAGCAACC CCAAGTGTAC ACTTGAATGC ATTAGTATG  
 3571 TTATGGACA GGGATTTAAT GTGGTGGAA TCTAGAGGCC GGGGAACTGA TACCATGCC AGATGTAGTT  
 3641 GCAGGTGTGG AGTTTACTAT TGTGAATCTA GGAGGAAGTA CTACCTGTC ACTTTACTG GCCCAACGTT  
 3711 TCGATTCTATG GAAGCAAACG ACTACTATCC AGCAAGATAC CAGTCTACCA TGCTGATAGG GTGCGGATT  
 3781 GCAGAACCCG GGGACTGCGG TGGGATACTG AGGTGCACTC ATGGGTAAT TGGTATCATT ACTGCAGGAG  
 3851 GTGAAGGGGT AGTAGCCTTT GCTGACATTA GAGACCTCTG GTGAGTGTGAA GAGGAGGCCA TGGAACAGGG  
 3921 AATAACAAGC TACATCGAAT CTCTCGGCAC AGCCTTGGC GCAGGGTTCA CCCACACAAAT CAGTGAAGAAA  
 3991 GTGACTGAAT TGACAAACGAT GGTTACCAAGC ACTATCACAG AAAAACTACT GAAAAACTTG GTGAAAATAG

FIG. 10 (Continued)

4061 TGTCGGCTCT AGTGATTGTT GTGAGAAATT ATGAGGACAC TACCACGATC CTTGCAACAC TAGCACTACT  
 4131 CGGGTGTGAT ATATCTCCTT GGCAATGGTT GAAGAAGAAG GCATGTGACT TACTAGAGAT TCCTATGTG  
 4201 ATGCGCCAAG GTGATGGGTG GATGAAGAAA TTACAGAGG CGTGCATGAGC AGCTAAAGGC TTAGAGTGG  
 4271 TTAGCAACAA AATTCCAAG TTTATAGATT GTTGAAGTG TAAAATTATC CCAGACGCTA AGGACAAGGT  
 4341 GGAATTCTC ACCAAGTTGA AACAGCTAGA CATGTGGAA AATCAAATTG CAACCATCCA CCAATCTTGC  
 4411 CCCAGCCAA ACAACAAGA GATTCTTTC AACAATGTGA GATGGTAGC AGTCAGTCC CGTCGGTTG  
 4481 CACCAATTATA CGCTGTGGAG GCACGCCAA TTAACAAAT GGAGAGCAC ATAACAAATT ATATAACAGT  
 4551 CAAGAGCAA CACCGTATTG AACCAGTATG TATGCTCATT CATGGGTAC CAGGGACGGG TAAATCTATA  
 4621 GCTACTTCAT TAATAGGTAG AGCAATAGCA GAGAAGGAA GCACATCAGT CTATTCAATG CCACCTGACC  
 4691 CATTCACIT TGATGGCTAT AAACAACAAG GGGTAGTGAT TATGGACGAC CTAAACAAA ACCCCGATGG  
 4761 TATGGACATG AAACGTGTTT GCCAATGGT ATCAACAGTG GAGTTTATIC CTCCAATGGC CTCATTAGAG  
 4831 GAGAAGGGCA TTTGTTTAC ATCTGATTAT GTCTGGCTT CTACCAACTC TCATTCAATT GTACCACCA  
 4901 CAGTGGCTCA CAGTGATGCC TTAACCAGAC GATTGCAIT TGATGTGGAG GTTACACGA TGCTGAACA  
 4971 TTCAGTCAA GGCAAACCTGA ATATGCCAC GGCACTCAA TTGTGTAAGG ATTGCCAAC ACCTGCAAAT  
 5041 TTIAAAAAGT GITGCCCTC CGTTGTGGAA AAGGCCCTGC AATTAATGGA CAGGTACACC AGACAAAGGT  
 5111 TCACTGTAGA TGAGATTACC ACATTATCA TGAATGAGAA AAACAGAAGG GCCAATATCG GCAATTGCA  
 5181 GGAAGCCTTG TTCAAGGAC CACTAAGGTAA AAGATTTG AAGATCAGT TGAGACAGT TCCCCCCCCCT  
 5251 GAGTCATCA GTGATTGTT ACAAGCAGTG GATTCTCAAG AGGTTAGGAA TTACTGTGAG AAGAAAGGCT  
 5321 GGATCGTAA CGTTACTAGC CAGATTCAAC TAGAAAGGAA CATCAATAGG GCCATGACTA TACTCCAAGC  
 5391 TGTTACACCA TTGCAAGCAG TCGCAGGAGT AGTGTATGTA ATGTACAAAC TCTTCGGGG TCAACAGGGT  
 5461 GCATACACTG GCTGCCAAA CAAAAAACCC AATGTCCTA CTATCAGAGT CGCTAAAGTC CAGGGGCCAG  
 5531 GATTGACTA CGCAGTGGCA ATGGCAAAAA GAAACATAGT TACTGCAACC ACCACCAAGG GTGAATTAC  
 5601 CATGCTAGGG GTGCATGATA ATGTAGCAAT ATTGCCAACC CATGCCCTC CAGGAGAAC CATTATTATT  
 5671 GATGGGAAAG AAGTAGAGAT CCTAGACGCC AGAGCCTAG AAGATCAAGC GGGAACCAAT CTTGAGATCA  
 5741 CCATTATTAC TCTAAAAAGA AATGAGAAAGT TTAGAGACAT CAGATCACAT ATTCCCCACCC AAATTACTGA  
 5811 AACTAACGAT GGAGTGTGTA TCGTGAACAC TAGCAAGTAC CCCAATATGT ATGTCCCCGT TGTTGCTGTG  
 5881 ACCGAACAGG GATATCTTAA TCTCAGTGGA CGTCAAACCTG CTGCACTTT AATGTACAAC TTICCAACAA  
 5951 GGGCAGGGCA TGCGGAGGA ATCATCACTT GTACTGGCAA AGTCATTGGG ATGCATGTTG GCGGGAACGG  
 6021 TTCACATGGG TTGCTGCAG CCCTCAAGCG ATCATCACTT ACTCAAACAT AGGGCGAAAT CCAGTGGATG  
 6091 AGGTCAACAA AGAAAGTGGG GTACCCCATT ATAAATGCC CATCCAAGAC AAAGTTAGAA CCCAGTGCT  
 6161 TCCACTATGT TTTGAAGGT GTAAAGGAAC CAGCTGACT CACTAAGAAT GACCCAGAC TAAAACAGA  
 6231 TTTGAAGAA GCCATCTTT CTAAATATGT GGGGAACAAA ATTACTGAAG TGGACGAGTA CATGAAAGAA  
 6301 GCAGTGGATC ACTATGCAGG ACAGTTAATG TCACTGGATA TCAACACAGA ACAGATGTGC CTGGAGGATG  
 6371 CCATGTACGG TACCGATGGT CTTGAGGCC TGGATCTTAG CACTAGTGCT GGATATCCTT ATGTTGCAAT  
 6441 GGGGAAAAAG AAAAGAGACA TTCTAGATAA ACAGACCAGA GATACTAAGG AGATGCAGAG ACTTTAGAT  
 6511 ACCTATGGAA TCAATCTACC ATTAGTCACG TACGTGAAGG ATGAACACTCG GTCAAAAGCT AAAGTGGAAC  
 6581 AAGGAAAGTC AAGATTGATT GAAGCTTCCA GCCATATGA TTCACTGCA ATGAGAATGG CCTTGGC  
 6651 TCTTACGCA GCTTCCACA AGAATCAGG TGTGGTGCAG GGATCAGCAG TTGGTGTGAA CCCAGATITG  
 6721 TTTGGAGTA AGATACCACTG GCTAATGGAA GAAAAACTCT CGCTTTGTA CTACACAGGG TATGATGCC  
 6791 CACTCAGCCC TGCTGGTTT GAAGCTTTA AAATGGTGT AGAAAAAAATT GGATTTGGCA GTAGAGTGA  
 6861 CTATATAGAC TACCTGAACC ACTCTCACCA CCTTTACAAA AACAAGACTT ATTGTGTCAA AGGCCGCATG  
 6931 CCATCCGGCT GCTCTGGCAC CTCAATTTC AACTCAATGA TTAACAACTT GATCATTAGG ACGCTTTAC  
 7001 TGAGAACCTA CAAGGGCATA GACTTGGACC ATTAAAAAT GATTGCCTAT GGTGATGACG TGATAGCTTC  
 7071 CTACCCCAT GAGGGTGCAG CTAGTCTCCT AGCCCAATCA GGGAAAGACT ATGGACTAAC CATGACTCCA  
 7141 GCAGATAAT CAGCAACCTT TGAAACAGTC ACATGGGAGA ATGTAACATT TCTGAAAAGA TTTTCAGAG  
 7211 CAGATGAGAA GTATCATTCT CGGGPCATC CAGTGTGCCC AATGAAAGAA ATTCACGAAT CAATCAGATG  
 7281 GACCAAGGAC CCTAGAAACA CACAGGATCA CGTACCGCTCG TTGTCCTAT TAGCTGGCA CAACGGTGA  
 7351 GAAGAATACA ATAAATTTC AGCTAAACAT AGAAGTGTGC CAATCGGAAG AGCTTATTG CTCCAGAGT  
 7421 ACTCTACATT GTACCGCCGA TGGCTCGACT CATTCTAGTA ACCCTACCTC AGTCGATTG GATGGGTTA  
 7491 TACTGTGTA GGGGTAATT TTTCTTAAT TCGGAGAAAA AAAAAAAA AAAAAGAGCT CCAATCACT  
 7561 AGTGAATTGCG CGGCCGCCTG CAGGTCGACC ATATGGGAGA GCTCCAAAG CGTTGGATGC ATAGCTTGAG  
 7631 TATTCTATAG TGTCACCTAA ATAGCTGGC GTAATCATGG TCATAGCTGT TTCTGTGAA AATTTGTTAT  
 7701 CCGCTCACAA TTCCACACAA CATACTGAGCC GGAAGCATAA AGTGTAAAGC CTGGGGTGCC TAATGAGTGA  
 7771 GCTAACCTAC ATTAAATTGCG TTGCGCTCAC TGCCCGCTT CCAGTCGGGA AACCTGTGCGT GCCAGCTGCA  
 7841 TTAATGAATC GGCCAAACGCG CGGGGAGAGG CGGTTTGCCT ATTGGGCCT CTTCCGCTTC CTCGCTCACT  
 7911 GACTCGCTGC GCTCGCTCGT TCGGCTGCAG CGAGCGGTAT CAGCTCACTC AAAGCGGTAA ATACGGTTAT  
 7981 CCACAGAATC AGGGGATAAC GCAGGAAAGA ACATGTGAGC AAAAGGCCAG CAAAAGGCCA GGAACCGTAA  
 8051 AAAGGCCGCG TTGCTGGCGT TTTCCATAG GCTCCGCCCC CCTGACGAGC ATCACAAAAA TCGACGCTCA

FIG. 10 (Continued)

8121 AGTCAGAGGT GGCGAAACCC GACAGGACTA TAAAGATACC AGGCCTTCC CCCTGGAAGC TCCCTCGTGC  
 8191 GCTCTCTGT TCCGACCCCTG CCGCTTACCG GATACTGTC CGCCTTCTC CCTTCGGGAA GCCTGGCGCT  
 8261 TTCTCATAGC TCACGCTGTA GGTATCTCAG TCCGGTAG GTCGTTCGCT CCAAGCTGGG CTGTGTGCAC  
 8331 GAACCCCCCG TTCAGCCCGA CGCIGGCC TTATCCGGTA ACTATCGCT TGAGTCCAAC CCGTAAGAC  
 8401 ACGACTTATC GCCACTGGCA GCAGCCACTG GTACACAGGAT TAGCAGAGCG AGGTATGTAG GCGGTGCTAC  
 8471 AGAGTTCTTG AAGTGGTGGC CTAACCTACGG CTACACTAGA AGAACAGTAT TTGGTATCTG CGCTCTGCTG  
 8541 AAGCCAGTTA CTTGGGAAA AAGAGTTGGT AGCTCTTGAT CCGGCAAAC ACCACCGCT GTAGCGGTG  
 8611 GTTTTTTGT TTGCAAGCAG CAGATTACGC GCAGAAAAAA AGGATCTCAA GAAGATCCTT TGATCTTT  
 8681 TACGGGGTCT GACGCTCA GTGAACGAAAA CTCACGTTAA GGGATTTGG TCATGAGATT ATCAAAAAGG  
 8751 ATCTCACCT AGATCTTTT AAATAAAAAA TGAAGTTTA AATCAATCTA AAGTATATAT GAGTAAACTT  
 8821 GGTCTGACAG TTACCAATGC TTAATCAGTC AGGCACCTAT CTCAGCGATC TGTCTATTC GTTCATCCAT  
 8891 AGTGCCTGA CTCCCCGTCG TGTAGATAAC TACGATACGG GAGGGCTAC CATCTGGCCC CAGTGTGCA  
 8961 ATGATACCGC GAGACCCACG CTCACCGCT CCAGATTAT CAGCAATAAA CCAGCCAGCC GGAAGGGCG  
 9031 AGCGCAGAAG TGGTCTGCA ACTTTATCCG CCTCCATCCA GTCTATTAAAT TGTGCGGGG AAGCTAGAGT  
 9101 AAGTAGTTCG CCAGTTAATA GTTTGGCAA CGTTGTTGCC ATTGCTACAG GCATCGTGGT GTCACGCTCG  
 9171 TCGTTGGTA TGGCTTCATT CAGCTCCGGT TCCCAACGAT CAAGGGCAGT TACATGATCC CCCATGTTG  
 9241 GCAGAAAAAGC GTTGTCTGC TTGGCTCTC CGATCGTGT CAGAAGTAAAG TTGGCCGAG TGTATTC  
 9311 CATGGTTATG GCAGCACTGC ATAATTCTCT TACTGTCATG CCATCGTAA GATGCTTTTC TGTGACTGGT  
 9381 GAGTACTCAA CCAAGTCAT CTGAGAAATAG TGTATGCCGC GACCGAGTTG CTCTTGCCCG GCGTCAATAC  
 9451 GGGATAATAC CGCGCCACAT AGCAGAACTT TAAAAGTGCT CATCATGGAA AAACGTTCTT CGGGCGAAA  
 9521 ACTCTCAAGG ATCTTACCGC TGTGAGATC CAGTCGATG TAACCCACTC GTGCACCCAA CTGATCTTC  
 9591 GCATCTTTA CTTTACCAAG CGTTCTGGG TGAGCAAAAA CAGGAAGGCA AAATGCCGCA AAAAAGGGAA  
 9661 TAAGGGCGAC ACGGAAATGT TGAATACTCA TACTCTTCCT TTTCAATAT TATTGAAGCA TTTATCAGGG  
 9731 TTATTGTC ATGAGCGGAT ACATATTGA ATGATTTAG AAAAATAAAC AAATAGGGGT TCCGCGCACA  
 9801 TTTCGGGAA AAGTGCACC TGATGCCGTG TGAAATACCG CACAGATCG TAAGGAGAAA ATACCGCATC  
 9871 AGGAAATTGT AAGCGITAAT ATITGTTAA AATTCGCGTT AAATTTITGT TAAATCAGCT CATTTTAA  
 9941 CCAATAGGCC GAAATCGGCA AAATCCTTA TAAATCAAAA GAATAGACCG AGATAGGGIT GAGTGTGTT  
 10011 CCAGTTGGA ACAAGAGTCC ACTATTAAAG AACGTGGACT CCAACGTCAA AGGGCGAAAA ACCGTCTATC  
 10081 AGGGCGATGG CCCACTACGT GAACCATCAC CCTAATCAAG TTTTTGGGG TCGAGGTGCC GTAAAGCACT  
 10151 AAATCGGAAC CCTAAAGGGG GCCCCGATT TAGAGCTTGA CGGGGAAAGC CGCGAACGT GGCAGAGAAAG  
 10221 GAAGGGAAAGA AAGCGAAAGG AGCAGGGCGCT AGGGCGCTGG CAAGTGTAGC GGTACGCTG CGCGTAACCA  
 10291 CCACACCCGC CGCGTTAAT GCGCCGCTAC AGGGCGCGTC CATTGCGCAT TCAGGCTGCG CAACTGTTGG  
 10361 GAAGGGCGAT CGGTGCGGC CTCTTCGCTA TTACGCCAGC TGGCGAAAGG GGGATGTGCT GCAAGGGCGAT  
 10431 TAAGTTGGGT AACGCCAGGG TTTTCCAGT CACGACGTTG TAAAACGACG GCCAGTGAAT TGTAAATACGA  
 10501 CTCACATAG GGCGAATTGG GCCCCACGTC GCATGCTCCC GGCGCCATG CGGGCCGCGG GAATTGCGATT  
 10571 GAGGCATGCT AATACGACTC ACTATAGG

FIG. 11

1 TTAAACAGC TCTGGGTTG TTCCCACCCC AGAGGCCAC GTGGCGCTA GTACTCTGGT ATTACGGTAC  
 71 CTTTGACGC CTGTTTGTATCCCTCCCC CGTAACTTA GAAGCTTATC AAAAGTCAA TAGCAGGGT  
 141 ACAAACAGT ACCTCTACGA ACAAGCACTT CTGTTTCCCC GGTGATATCA CATAGACTGT ACCCACGGTC  
 211 AAAAGTATT GATCCGTAT CCGCTTGGT ACTTCGAGAA GCCTAGTATC ACCTTGGAT CTTCGATGCG  
 281 TTGCGCTCAA CACTCTGCC CGAGTGTAGC TTAGGCTGAT GAGTCTGGC ACTCCCCACC GGCGACGGTG  
 351 GCCCAGGCTG CGTTGGCGGC CTACCATGG CTGATGCCGT GGGACGCTAG TTGTAACAA GGTGTGAAGA  
 421 GCCTATTGAG CTACTCAAGA GTCTCCGGC CCCTGAATGC GGCTAATCCT AACACGGAG CAACCGCTCA  
 491 CAACCCAGTG AGTAGGTTG CGTAATGCGT AAGTCTGTGG CGGAACGGAC TACTTGGGT GTCCGTGTT  
 561 CCCTTATAT TCATACTGGC TGCTTATGGT GACAATTAC AAATTGTAC CATATAGCTA TTGGATTGGC  
 631 CACCCAGTAT TGTCAATAT ATTGAGTGT TTCTTICATA AGCCTTATTA ACATCACATT TTIAATCACA  
 701 ATAAACAGTG CAAATGGGG CTCAAGTTT AACGAAAAG ACCGGTGCACGAGAATCA AAACGTGGCA  
 771 GCCAATGGAT CCACCATTA TTACACTACT ATCAACTATT ACAAAAGACAG TGCAGTAAT TCCGCTACTA  
 841 GACAAGACCT CTCCCAAGAT CCATCAAAAT TCACAGAACCGT TAATGTGAA AAACAGCACC  
 911 AGCTCTAACAC TCGCTAACG TGGAACATG TGTTGACAGT GACCGTGTGA GGCAAACTAC TTAGGCAAC  
 981 TCGACTATTA CTACACAAAGA AGCAGGAACT GCTATTGTTG CTTACGGTGA ATGGCCACT TACATAATG  
 1051 ATTCAAGAGC TAATCCGTTA GATGCACCCA CTGAGCCAGA CGTTAGTAGC AACCGGTTT ACACCTAGA  
 1121 ATCGGTGCT TGGAGAGACCA CTTCAGGGG ATGGTGGTGG AAGTACCAAGG ATTGTTGAA GGACATGGGA  
 1191 ATGTTGGTC AGAATATGTA CTATCACTAC TTGGGGCGCT CTGGTTACAC CATTATGTC CAGTGCACG  
 1261 CTTCAAAATT TCACCAAGGG GCGTTAGGAG TTTCCTGAT ACCAGAGTT GTCATGGCTT GCAACACTGA  
 1331 GAGTAAAACG TCATACGTTT CATAACATCAA TGCAATCCT GGTGAGAGAG GCGGTGAGTT TACGAACACC  
 1401 TACAATCCGT CAAATACAGA CGCCAGTGGAG GGCAGAAAGT TTGCACTGATT GGATTATTG CTGGGTTCTG  
 1471 GTGTCTAGC AGGAAACGCC TTGTTGTTAC CGCACCGAGT CATCAACTA CGTACCAACA ACAGTGCACAC  
 1541 AATTGTTGGTG CCATACGTA ACTCACTGT GATTGATTGT ATGGCAAAAC ACAAAACTG GGGCATTGTC  
 1611 ATATACCCAC TGGCACCCCTT GGCTTGGC GCAACATCGT CACCAAGGGT GCCTATTACA GTGACCATTG  
 1681 CACCATGTG TACAGAATTC AATGGGGTTG GAAACATCAC CGTCCCAGTATCAAGGGT TGCCGACAAT  
 1751 GAACACACCT GGTTCCAATC AATTCTTAC ATCTGATGAC TTCCAGTCGC CCTGTGCTT ACCTAATT  
 1821 GATGTTACTC CACCAATACA CATAACGGGG GAAGTAAAGA ATATGATGGA ACTAGCTGAA ATTGACACAT  
 1891 TGATCCAAT GAACGCAGTG GACGGGAAGG TGAACACAAT GGAGATGTAT CAAATACCAT TGAATGACAA  
 1961 TTTGAGCAAG GCACCTATAT TCTGTTTATC CCTATCACCT GCTTCTGATA AACGACTGAG CCACACCATG  
 2031 TTGGGTGAAA TCCTAAATT TTACACCCAT TGGACGGGGT CCATCAGGTT CACCTTCTA TTTTGTGGCA  
 2101 GTATGATGGC CACTGGTAAA CTGCTCTCA GCTATTCCCC ACCGGGAGCT AAACCAACCA CAAATCGCAA  
 2171 GGATGCAATG CTAGGCACAC ACATCATCTG GGACCTAGGG TTACAATCCA GTTGTCCAT GTTGCACCG  
 2241 TGGATCTCCA ACACAGTGT CAGACGGTGT GCACGTGATG ACTTCACGTG GGGCGGATT ATAACCTGCT  
 2311 TCTATCAAAC TAGAATTGTTG GTACCTGCTT CAACCCCTAC CAGTATGTT ATGTTAGGCT TTGTTAGTGC  
 2381 GTGTCAGAC TTCAGTGTCA GACTGTTAG GGACACTCCC CATATAGTC AATGCAAACCT AATAGGACGT  
 2451 ACACAAAGGCA TTGAGAACCT CATTGACACA GCGATAAAGA ATGCCTTAAG AGTGTCCCAA CCACCCCTCGA  
 2521 CCCAGTCAC TGAAGCAACT AGTGGAGTGA ATAGCAGGA GGTGCCAGCT CTAACGTCTG TGAAACAGG  
 2591 AGCATCTGGT CAAGCAATCC CCAGTGATGT GGTGGAAACT AGGCACGTGG TAAATTACAA ACCAGGTCT  
 2661 GAATCGTGTG TTGAGTCATT CTTTGGGAGA GCTGCGTGTG TCACAATCTT ATCCTGACC AACTCCTCCA  
 2731 AGAGCGGAGA GGAGAAAAAG CATTCAACAA TATGGAATAT TACATACACC GACACTGTCC AGTACCGAG  
 2801 AAAATTAGAA TTTTCACGT ATTCCAGGTT TGATCTTGAATGACTTTG TATTACAGA GAACTATCCT  
 2871 AGTACAGCCA GTGGAGAAAGT GCGAAACCAG GTGTACAGA TCATGTATAT TCCACCAGGG GCACCCGCC  
 2941 CATCATCTG GGATGACTAC ACATGCAAT CCTCTTCAAAC CCCTTCCATC TTCTACATGT ATGGAAATGC  
 3011 ACCTCCACGG ATGTCATT CTTACGTAGG GATTGCAAT GCCTATTACAC ACTTCACGA TGGCTTGTCA  
 3081 CGGGTCCAC TTGAGGGTGA GAAACACCGAT GTGGCGACAA CGTTTACGG TTTAGTGTCC ATAAATGATT  
 3151 TTGGAGTTT AGCAGTGTG CAGTAAACCG GCGATAATCC ACATACAAAT CACACATCTG TGAGAGTGT  
 3221 CATGAAACCA AAACACATTC GGTGTTGGTGC CCCAGACCT CCTCGAGCTG TATTATACAG GGGAGAGGG  
 3291 GTGGACATGA TATCCAGTGC AATTCTACCT CTGGCCAAGG TAGACTCAAT TACCACTTT GGGTTTGGTC  
 3361 ATCAGAACAA AGCAGTGTAC GTTGCAGGTT ACAAGATTG CAACTACAC CTAACGACCC CAAGTGTACA  
 3431 CTTGAATGCA ATTAGTATGT TATGGGACAG GGATTTAATG GTGGTGGAT CTAGAGCCCG GGGAACTGAT  
 3501 ACCATGCCA GATGTAGTT CAGGTGTGGA GTTACTATT GTGAATCTAG GAGGAAGTAC TACCTGTCA  
 3571 CTTTACTGG CCCAACGTT CGATTCAATGG AAGCAAACGA CTAATACCA GCAAGATACC AGTCTCACAT  
 3641 GCTGATAGGG TGCGGATTG CAGAACCCGG GGACTGCGGT GGGATACTGA GGTGCACTCA TGGGGTAATT  
 3711 GGTATCATTATGCAAGGAGG TGAAGGGTGA GTAGCCTTG CTGACATTAG AGACCTCTGG GTGTATGAAAG  
 3781 AGGAGGCCAT GGAACAGGGG ATAACAAAGCT ACATCGAATC TCTCGGCACA GCCTTTGGCG CAGGGTTCAC  
 3851 CCACACAATC AGTGAGAAG TGACTGAATT GACAACGATG GTTACCAAGCA CTATCACAGA AAAACTACTG

FIG. 11 (Continued)

3921 AAAAACTTGG TGAAAATAGT GTCGGCTCTA GTGATTGTIG TGAGAAATT TGAGGACACT ACCACGATCC  
 3991 TTGCAACACT AGCACTACTC GGGTGTGATA TATCTCCITG GCAATGGTTG AAGAAGAAGG CATGTGACTT  
 4061 ACTAGAGATT CCTTATGTGA TGCGCCAAGG TGATGGGTGG ATGAAGAAAT TCACAGAGGC GTGCAATGCA  
 4131 GCTAAAGGCT TAGAGTGGAT TAGCAACAAA ATTCCAAGT TTATAGATTG GTTGAAGTGT AAAATTATCC  
 4201 CAGACGCTAA GGACAAGGTG GAATTCTCA CCAAGTTGAA ACAGCTAGAC ATGTTGGAAA ATCAAATTGC  
 4271 AACCATCCAC CAATCTGGC CCAGCCAAGA ACAACAAGAG ATTCTTCTCA ACAATGTGAG ATGGTTAGCA  
 4341 GTCCAGTCCC GTCGGTTTGC ACCATTATAC GCTGTTGGAGG CACGCCAAT TAACAAAATG GAGAGCACAA  
 4411 TAAACAATTA TATACAGTTC AAGAGCAAAC ACCGTATTGA ACCAGTATGT ATGCTCATTTC ATGGGTCA  
 4481 AGGGACGGGT AAATCTATAG CTACTCATT AATAGGTAGA GCAATAGCAG AGAAGGAAAG CACATCAGTC  
 4551 TATTCATGC CACCTGACCC ATCTCACTT GATGGCTATA ACAACAAGG GGTAGTGTGATT ATGGACGACC  
 4621 TAAACCAAAA CCCCGATGGT ATGGACATGA AACTGTTTG CCAAATGGTA TCAACAGTGG AGTTTATTCC  
 4691 TCCAATGGCC TCATTAGAGG AGAAGGGCAT TTGTTTACA TCTGATTATG TCTGGCTTC TACCAACTCT  
 4761 CATTCAATTG TACCACCCAC AGTGGCTCAC AGTGTGATGCCT TAACCAGACG ATTTGCAATTGATGTG  
 4831 TTTACACGAT GTCTGAACAT TCAGTCAAAG GCAAACGTAA TATGGCCACG GCCACTCAAT TGTGTAAGGA  
 4901 TTGTCACAA CCTGCAAATT TTAAAAGTG TTGCCCCCTC GTTTGTGGAA AGGCCTTGCA ATTAATGGAC  
 4971 AGGTACACCA GACAAAGGTT CACTGTAGAT GAGATTACCA CATTAACTAT GAATGAGAAA AACAGAAGGG  
 5041 CCAATATCGG CAATTGCGATG GAAGCCTTGT TTCAAGGACCA ACTAAGGTT AAAGATTGAG ATGATCGATG  
 5111 GAAGACAGTT CCCCCCCCCTG AGTGCATCG TGATTGTTA CAAGCAGTGG ATTCTCAAGA GTTGTAGGAT  
 5181 TACTGTGAGA AGAAAGGCTG GATCGTTAAC GTTACTAGCC AGATTCAACT AGAAAGGAAC ATCAATAGGG  
 5251 CCATGACTAT ACTCCAAGCT GTTACCAACAT TCGCAGCAGT CGCAGGAGTA GTGTATGTAA TGTACAAACT  
 5321 CTTGCCGGT CAACAGGGTG CATAACTGG CTGCCAAC AAAAAACCCA ATGCCCTAC TATCAGAGTC  
 5391 GCTAAAGTCC AGGGGCCAGG ATTTGACTAC GCAGTGGCAA TGGCAAAAG AAACATAGTT ACTGCAACCA  
 5461 CCACCAAGGG TGAATTTACC ATGCTAGGGG TGATGATAA TGTAGCAATA TTGCCAACCC ATGCCCTCC  
 5531 AGGAGAAACC ATTATTATG ATGGGAAAGA AGTAGAGATC CTAGACGCCA GAGCCTTAGA AGATCAAGCG  
 5601 GGAACCAATC TTGAGATCAC CATTAACTCTAAGA ATGAGAAGTT TAGAGACATC AGATCACATA  
 5671 TTCCCAACCA AATTACTGAA ACTAACGATG GAGTGTGAT CGTGAACACT AGCAAGTACC CCAATATGTA  
 5741 TGTCCCCGTT GGTGCTGTGA CGGAACAGGG ATATCTTAAT CTCAGTGAC GTCAAACCTGC TCGCACTTA  
 5811 ATGTAACACT TTCCAACAAG GGCAGGCCAG TGCGGAGGAA TCATCACTTG TACTGGCAA GTCAATTGGGA  
 5881 TGCAATGTTG CGGGAACGGT TCACATGGGT TTGCTGCAGC CCTCAAGCGA TCATACTTCA CTCAAAATCA  
 5951 GGGGAAATC CAGTGGATGA GGTCAAAAGA AGAAGTGGGG TACCCCATTA TAAATGCCCA ATCCAAGACA  
 6021 AAGTTAGAAC CCAGTGCTT CCACTATGT TTGAGGTTAAGGAAC AGCTGTACTC ACTAAGAATG  
 6091 ACCCCAGACT AAAAACAGAT TTGAGAAGAG CCATCTTTC TAAATATGTG GGGAAACAAA TTACTGAAGT  
 6161 GGACGAGTAC ATGAAAGAAG CAGTGGATCA CTATGCAGGA CAGTTAATGT CACTGGATAT CAACACAGAA  
 6231 CAGATGTGCC TGGAGGATGC CATGTACGGT ACCGATGGTC TTGAGGCCCT GGATCTTAGC ACTAGTGTG  
 6301 GATATCTTA TGTTGCAATG GGGAAAAAGA AAAGAGACAT TCTAGATAAA CAGACAGAG ATACTAAGGA  
 6371 GATGCAGAGA CTTTAGATA CCTATGGAAT CAATCTACCA TTAGTCACGT ACGTGAAAGA TGAACCTCAGG  
 6441 TCAAAGACTA AAGTGGAAACA AGGAAAGTCA AGATTGATG AAGCTTCCAG CCTTAATGAT TCAGTTGCAA  
 6511 TGAGAATGGC CTTGGCAAT CTTACGCAG CTTCCACAA GAATCCAGG GTGGTGACAG GATCAGCAGT  
 6581 TGGTGTGAC CCAGATITGT TTGGAGTAA GATACCAGTG CTAATGGAAAG AAAAATCTT CGCTTTTGAC  
 6651 TACACAGGGT ATGATGCCCTC ACTCAGCCCT GCTGGTTG AAGCTTAA AATGGTGTAA GAAAAATTG  
 6721 GATTGGCG TAGAGTAGAC TATAGACT ACCTGAACCA CTCTCACCA CTTACAAAA ACAAGACITA  
 6791 TTGTTGTCAAA GGCAGCATGC CATCCGGCTG CTCTGGCACC TCAATTTCA ACTCAATGAT TAACAACCTG  
 6861 ATCATTAGGA CGTTTACT GAGAACCTAC AAGGGCATAG ACTTGGACCA TTAAAATG ATGCTTATG  
 6931 GTGATGACGT GATAGCTTCC TACCCCATG AGGTTGACGC TAGTCTCTA GCCCAATCAG GAAAAGACTA  
 7001 TGGACTAACC ATGACTCCAG CAGATAAACATG GCAACACCTT GAAACAGTCA CATGGGAGAA TGTAACATT  
 7071 CTGAAAAGAT TTTCAGAGC AGATGAGAAG TATCCATTCC TTGTCATCC AGTGTGACCA ATGAAAGAAA  
 7141 TTCACGAATC ATCAGATGG ACCAAGGACC CTAGAACAC ACAGGATCAC GTACGCTCGT TGTGCTTATT  
 7211 AGCTTGGCAC AACGGTGAAG AAGAACATCAA TAAATTTTA GCTAAATCA GAAGTGTGCC AATCGGAAGA  
 7281 GCTTATTGTC TCCAGAGTA CTCTACATTG TACCGCCGAT GGCTCGACTC ATTTAGTAA GGCAGCCCG  
 7351 ACACACTGGTT GAAGGGGACCA AACTGGAGAC ACACACTTCC TTACATTCA CGGCAGCCCG ACTCATTAA  
 7421 GTAAACCTAC CTCAGTCGA TTGGATTGGG TTAACTGTG GTAGGGTAA ATTTCCTT AATTGGAGA  
 7491 AAAAAGAAAAAAGA GCTCCAATC ACTAGTGAAT TCGCGCCGC CTGCAGGTCG ACCATATGGG  
 7561 AGAGCTCCA ACGCGTTGGA TGCATAGCTT GAGTATTCTA TAGTGTACCA TAAATAGCTT GCGTAATCA  
 7631 TGGTCATAGC TGTTCTGT GTGAAATTGT TATCCGCTCA CAATTCCACA CAACATACGA GCGGAAGCA  
 7701 TAAAGTGTAA AGCCTGGGGT GCCTAATGAG TGAGCTAACT CACATTAATT GCGTTGCGCT CACTGCCCG  
 7771 TTTCAGTCG GGAAACCTGT CGTGCAGCT GCATTAATGA ATCGGCAAC GCGCGGGGAG AGGCGGTTG  
 7841 CGTATTGGGC GCTCTTCCGC TTCCCTGCTC ACTGACTCGC TGCGCTCGT CGTTCGGCTG CGGCGAGCGG  
 7911 TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA ATCAGGGAT AACGCAGGAA AGAACATGTG

FIG. 11 (Continued)

7981 AGCAAAAGGC CAGCAAAAGG CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTITCCA TAGGCTCCGC  
 8051 CCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA CCCGACAGGA CTATAAAGAT  
 8121 ACCAGGCCTT TCCCCCTGGA AGCTCCCTCG TGCGCTCTCC TGTCGACCC CTGCCGCTTA CGCGATACCT  
 8191 GTCCGCCTT CTCCCTTCGG GAAGCGTGGC GCTTCTCAT AGCTCACGCT GTAGGTATCT CAGTCGGTG  
 8261 TAGTCGTC GCTCCAAGCT GGGCTGTTG CACGAACCCC CGGTTCAAGCC CGACCGCTGC GCCTTATCCG  
 8331 GTAACATCG TCTTGAGTC AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG  
 8401 GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAAGTGGT GGCTTAACTA CGGCTACACT  
 8471 AGAAGAACAG TATTTGGTAT CTGCGCTCTG CTGAAGCCAG TTACCTCGG AAAAAGAGTT GGTAAGCTTT  
 8541 GATCCGCAA ACAAAACCACC GCTGGTAGCG GTGGTTTTTG TGTTGCAAG CAGCAGATT CGCGCAGAAA  
 8611 AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG TCTGACGCTC AGTGGAAACGA AAACTCACGT  
 8681 TAAGGGATTGTT TGGTCATGAG ATTATCAAAA AGGATCTTC CCTAGATCCT TTAAATTAA AAATGAAGTT  
 8751 TTAAATCAAT CTAAAGTATA TATGAGTAAA CTTGGTCTGA CAGTACCAA TGCTTAATCA GTGAGGCACC  
 8821 TATCTCAGCG ATCTGTCTAT TTCGTICATC CATAAGTGCC TGACTCCCCG TCGTGTAGAT AACTACGATA  
 8891 CGGGAGGGCT TACCATCTGG CCCCAGTGCT GCAATGATAC CGCGAGACCC ACGCTCACCG GCCTCAGATT  
 8961 TATCAGCAAT AAACCAGCCA GCCGGAAGGG CCGAGCGCAG AAGTGGTCTC GCAACTTTAT CGCCCTCCAT  
 9031 CCAGTCTATT ATTGGTGC GGGAAGCTAG AGTAAGTGT AGTGGCTTA ATAGTTGCG CAACGTTGTT  
 9101 GCCATTGCTA CAGGCATCGT GGTGTCATCGC TCGTCGTTTG GTATGGCTTC ATTCACTCC GGTTCCAAAC  
 9171 GATCAAGCG AGTACATGA TCCCCATGT TGTGCAAAA AGCGGTIAAGC TCCITCGTC CTCCGATCGT  
 9241 TGTCAGAAGT AAGTGGCCG CAGTGTATC ACTCATGGTT ATGGCAGCAC TGCATAATTCTCTTACTGTGTC  
 9311 ATGCCATCCG TAAGATGCTT TTCTGTGACT GGTGAGTACT CAACCAAGTC ATTCTGAGAA TAGTGTATGC  
 9381 GGCGACCGAG TTGCTCTGC CGCGCTCAA TACGGGATAA TACCGCGCCA CATAGCAGAA CTTAAAGT  
 9451 GCTCATCATT GAAAACGTT CTTCGGGGCG AAAACTCTCA AGGATCTAC CGCTGGTAGAT ATCCAGTTG  
 9521 ATGTAACCCA CTCGTGCACC CAACTGATCT TCAGCATCTT TTACTTCAC CAGCGTTCT GGGTGAGCAA  
 9591 AAACAGGAAG GCAAAATGCC GCAAAAAGG GAATAAGGGC GACACGAAA TGTGAATAC TCATACTCTT  
 9661 CCTTTTCAA TATTGAA GCATTATCA GGGTATTGTT CTCACTGAGCG GATACATATT TGAATGTATT  
 9731 TAGAAAATAA AACAATAGG GGTCCCGCG ACATTTCCCC GAAAAGTGC ACCTGATGCG GTGTGAAATA  
 9801 CGCGACAGAT GCGTAAGGAG AAAATACCGC ATCAGGAAAT TGTAAAGCGT AATATTGTT TAAAATTCG  
 9871 GTTAAATTTT GTTAAATCA GCTCATTTT TAACCAATAG GCGAAATCG GCAAAATCCC TTATAATCA  
 9941 AAAGAATAGA CCGAGATAGG GTTGAGTGTGTT GTCAGGTTT GGAACAAAGAG TCCACTATTA AAGAACGTGG  
 10011 ACTCCAACGT CAAAGGGCGA AAAACCGTCT ATCAGGGCGA TGGCCCACTA CGTGAACCAT CACCCCTAATC  
 10081 AAGTTTTTG GGGTCGAGGT GCCGTAAGAC ACTAAATCGG AACCTAAAG GGAGCCCCCG ATTAGAGCT  
 10151 TGACGGGGAA AGCCGGCGAA CGTGGCGAGA AAGGAAGGGA AGAAAGCAGA AGGAGCGGGCGCTAGGGCGC  
 10221 TGGCAAGTGT AGCGGTCAAG CTGCGCGTAA CCACCAACACC CGCCGCGCTT AATGGCGCCGC TACAGGGCGC  
 10291 GTCCATTGCG CATTCAAGGCT GCGCAACTGT TGGGAAGGGC GATCGGTGCG GGCCCTTTCG CTATTACGCC  
 10361 AGCTGGCGAA AGGGGGATGT GCTGCAAGGC GATTAAGTTG GTAAACGCCA GGTTTTCCC AGTCACGACG  
 10431 TTGAAAACG ACGGCCAGTG AATTGTAATA CGACTCACTA TAGGGCGAAT TGCCCCGAC GTCGCATGCT  
 10501 CCCGGCCGCC ATGGCGGCCG CGGGAAATTG ATTGAGGCAT GCTAATACGA CTCACATAG G

FIG. 12

1 TTAAACAGC TCTGGGTTG TTCCCACCCC AGAGGCCAC GTGGCGCTA GTACTCTGGT ATTACGGTAC  
 71 CTTTGACGC CTGTTTGTATCCCTCCCC CGTAACTTA GAAGCTTATC AAAAGTCAA TAGCAGGGT  
 141 ACAAACAGT ACCTCTACGA ACAAGCACTT CTGTTTCCCC GGTGATATCA CATAGACTGT ACCCACGGTC  
 211 AAAAGTATT GATCCGTAT CCGCTTGGT ACTTCGAGAA GCCTAGTATC ACCTTGGAT CTTCGATGCG  
 281 TTGCGCTCAA CACTCTGCC CGAGTGTAGC TTAGGCTGAT GAGTCTGGC ACTCCCCACC GGCGACGGTG  
 351 GCCCAGGCTG CGTTGGCGGC CTACCATGG CTGATGCCGT GGGACGCTAG TTGTAACAA GGTGTGAAGA  
 421 GCCTATTGAG CTACTCAAGA GTCTCCGGC CCCTGAATGC GGCTAATCCT AACACGGAG CAACCGCTCA  
 491 CAACCCAGTG AGTAGGTTG CGTAATGCGT AAGTCTGTGG CGGAACGGAC TACTTGGGT GTCCGTGTT  
 561 CCCTTATAT TCATACTGGC TGCTTATGGT GACAATTAC AAATTGTAC CATATAGCTA TTGGATTGGC  
 631 CACCCAGTAT TGTCAATAT ATTGAGTGT TTCTITCATA AGCCTTATTA ACATCACATT TTIAATCACA  
 701 ATAAACAGTG CAAATGGGG CTCAAGTTT AACGAAAAG ACCGGTGCAC ACGAGAAATCA AAACGTGGCA  
 771 GCCAATGGAT CCACCATTA TTACACTACT ATCAACTATT ACAAAAGACAG TGCAGTAAT TCCGCTACTA  
 841 GACAAGACCT CTCCCAAGAT CCATCAAAAT TCACAGAACCG GTTAAGGAC TTAATGTTGA AAACAGCACC  
 911 AGCTCTAACAC TCGCTAACG TGGAAAGCATG TGGGTACAGT GACCGTGTGA GGCAAAATCAC TTAGGCAAC  
 981 TCGACTATTA CTACACAAAGA AGCAGGAACT GCTATTGTTG CTTACGGTGA ATGGCCACT TACATAATG  
 1051 ATTCAAGAGC TAATCCGTTA GATGCACCCA CTGAGCCAGA CGTTAGTAGC AACCGGTTT ACACCTAGA  
 1121 ATCGGTGCT TGGAAAGACCA CTTCAGGGG ATGGTGGTGG AAGTACCAAGG ATTGTTGAA GGACATGGGA  
 1191 ATGTTGGTC AGAATATGTA CTATCACTAC TTGGGGCGCT CTGGTTACAC CATTATGTC CAGTGCACAG  
 1261 CTCAAAATT TCACCAAGGG GCGTGTAGGAG TTTCTGTAT ACCAGAGTT GTCATGGCTT GCAACACTGA  
 1331 GAGTAAAACG TCATACGTTT CATAACATCAA TGCAATCCT GGTGAGAGAG GCGGTGAGTT TACGAACACC  
 1401 TACAATCCGT CAAATACAGA CGCCAGTGGAG GGCAGAAAGT TTGCACTGATT GGATTATTG CTGGGTTCTG  
 1471 GTGTCTAGC AGGAAACGCC TTGTTGTTAC CGCACCGAGT CATCAACTA CGTACCAACA ACAGTGCACAC  
 1541 AATTGTTGGTG CCATACGTA ACTCACTGT GATTGATTGT ATGGCAAAAC ACAAAACTG GGGCATTGTC  
 1611 ATATACCCAC TGGCACCCCTT GGCTTGGC GCAACATCGT CACCAAGGGT GCCTATTACA GTGACCATTTG  
 1681 CACCATGTG TACAGAATTC AATGGGGTTG GAAACATCAC CGTCCCAGTATCAAGGTT TGCCGACAAAT  
 1751 GAACACACCT GGTTCCAATC AATTCTTAC ATCTGATGAC TTCCAGTCGC CCTGTGCTT ACCTAATT  
 1821 GATGTTACTC CACCAATACA CATAACGGGG GAAGTAAAGA ATATGATGGA ACTAGCTGAA ATTGACACAT  
 1891 TGATCCAAT GAACGCAGTG GACGGGAAGG TGAACACAAT GGAGATGTAT CAAATACCAT TGAATGACAA  
 1961 TTTGAGCAAG GCACCTATAT TCTGTTTATC CCTATCACCT GCTTCTGATA AACGACTGAG CCACACCATG  
 2031 TTGGGTGAAA TCCTAAATT TTACACCCAT TGGACGGGGT CCATCAGGTT CACCTTCTA TTTTGTGGCA  
 2101 GTATGATGGC CACTGGTAAA CTGCTCTCA GCTATTCCCC ACCGGGAGCT AAACCAACCA CAAATCGCAA  
 2171 GGATGCAATG CTAGGCACAC ACATCATCTG GGACCTAGGG TTACAATCCA GTTGTCCAT GTTGCACCG  
 2241 TGGATCTCCA ACACAGTGT CAGACGGTGT GCACGTGATG ACTTCACGTG GGGCGGATT ATAACCTGCT  
 2311 TCTATCAAAC TAGAATTGTTG GTACCTGCTT CAACCCCTAC CAGTATGTT ATGTTAGGCT TTGTTAGTGC  
 2381 GTGTCAGAC TTCAGTGTCA GACTGTTAG GGACACTCCC CATATTAGTC AATGCAAACCT AATAGGACGT  
 2451 ACACAAAGGCA TTGAAAGACCT CATTGACACA GCGATAAAGA ATGCCTTAAG AGTGTCCCCA CCACCCCTCGA  
 2521 CCCAGTCAC TGAAGCAACT AGTGGAGTGA ATAGCAGGA GGTGCCAGCT CTAACTGCTG TGGAAACAGG  
 2591 AGCATCTGGT CAAGCAATCC CCAGTGATGT GGTGGAAACT AGGCACGTGG TAAATTACAA ACCAGGTCT  
 2661 GAATCGTGTG TTGAGTCATT CTTTGGGAGA GCTGCGTGTG TCACAATCTT ATCCTGACC AACTCCTCCA  
 2731 AGAGCGGAGA GGAGAAAAAG CATTCAACAA TATGGAATAT TACATACACC GACACTGTCC AGTTACCGAG  
 2801 AAAATTAGAA TTTTCACGT ATTCCAGGTT TGATCTTGAATGACTTTG TATTACAGA GAACTATCCT  
 2871 AGTACAGCCA GTGGAGAAAGT GCGAAACCAG GTGTACAGA TCATGTATAT TCCACCAAGGG GCACCCGCC  
 2941 CATCATCTG GGATGACTAC ACATGCAAT CCTCTTCAAAC CCCTTCCATC TTCTACATGT ATGGAAATGC  
 3011 ACCTCCACGG ATGTCATT CTTACGTAGG GATTGCAAT GCCTATTACAC ACTTCACGA TGGCTTGTCA  
 3081 CGGGTCCAC TTGAGGGTGA GAAACACCGAT GTGGCGACAA CGTTTACGG TTTAGTGTCC ATAAATGATT  
 3151 TTGGAGTTT AGCAGTGTG CAGTAAACCG GCGATAATCC ACATACAAAT CACACATCTG TGAGAGTGT  
 3221 CATGAAACCA AAACACATTC GGTGTTGGTGC CCCAGACCT CCTCGAGCTG TATTATACAG GGGAGAGGG  
 3291 GTGGACATGA TATCCAGTGC AATTCTACCT CTGGCCAAGG TAGACTCAAT TACCACTTT GGGTTTGGTC  
 3361 ATCAGAACAA AGCAGTGTAC GTTGCAGGTT ACAAGATTG CAACTACAC CTAGCAACCC CAAGTGTACA  
 3431 CTTGAATGCA ATTAGTATGT TATGGGACAG GGATTTAATG GTGGTGGAAAT CTAGAGCCCG GGGAACTGAT  
 3501 ACCATGCCA GATGTAGTTG CAGGTGTGGA GTTACTATT GTGAATCTAG GAGGAAGTAC TACCTGTCA  
 3571 CTTTACTGG CCCAACGTTT CGATTCAATGG AAGCAAACGA CTAATACCA GCAAGATACC AGTCTCACAT  
 3641 GCTGATAGGG TGCGGATTG CAGAACCCGG GGACTGCGGT GGGATACTGA GGTGCACTCA TGGGGTAATT  
 3711 GGTATCATTATGCAAGGAGG TGAAGGGTGA GTAGCCTTG CTGACATTAG AGACCTCTGG GTGTATGAAAG  
 3781 AGGAGGCCAT GGAACAGGGG ATAACAAAGCT ACATCGAATC TCTCGGCACA GCCTTTGGCG CAGGGTTCAC  
 3851 CCACACAATC AGTGAGAAAG TGACTGAATT GACAACGATG GTTACCAAGCA CTATCACAGA AAAACTACTG

FIG. 12 (Continued)

3921 AAAAACTTGG TGAAAATAGT GTCGGCTCA GTGATTGTG TGAGAAATT TGAGGACACT ACCACGATCC  
 3991 TTGCAACACT AGCACTACTC GGGTGTGATA TATCTCCTTG GCAATGGTTG AAGAAGAAGG CATGTGACTT  
 4061 ACTAGAGATT CCTTATGTGA TCGCCAAGG TGATGGGTGG ATGAAGAAAT TCACAGAGGC GTGCAATGCA  
 4131 GCTAAAGGCT TAGAGTGGAT TAGCAACAAA ATTCCAAGT TTATAGATTG GTTGAAGTGT AAAATTATCC  
 4201 CAGACGCTAA GGACAAGGTG GAATTCTCA CCAAGTTGAA ACAGCTAGAC ATGTTGGAAA ATCAAATTGCA  
 4271 AACCATCCAC CAATCTGCC CCAGCCAAGA ACAACAAGAG ATTCCTTCA ACAATGTGAG ATGGTAGCA  
 4341 GTCCAGTCCC GTCCGGTTTC ACCATTATAC GCTGTGGAGG CACCGGAAT TAACAAATG GAGAGCACAA  
 4411 TAAACATTA TATACAGTTC AAGAGCAAC ACCGTATTGA ACCAGTATGT ATGCTCATTTC ATGGGTCA  
 4481 AGGGACGGGT AAATCTATAG CTACTTCAATT AATAGGTAGA GCAATAGCAG AGAAGGAAAG CACATCAGTC  
 4551 TATTCAATGC CACCTGACCC ATCTCACTT GATGGCTATA AACACAACAGG GGTAGTGTGATT ATGGACGACC  
 4621 TAAACCAAAA CCCCGATGGT ATGGACATGA AACTGTTTG CCAAATGGTA TCAACAGTGG AGTTTATTCC  
 4691 TCCAATGGCC TCATTAGAGG AGAAGGGCAT TTTGTTTACA TCTGATTATG TCCTGGCTTC TACCAACTCT  
 4761 CATTCAATTG TACCACCCAC AGTGGCTCAC AGTGTGCTT TAACCAGACG ATTGCAATTGATGTGGAGG  
 4831 TTACACGAT GTCTGAACAT TCAGTCAAAG GCAAACGTAA TATGGCCACG GCCACTCAAT TGTGTAAGGA  
 4901 TTGTTCAACA CCTGCAAATT TTAAAAGTGT TTGCGCTCTC GTTTGTGGAA AGGCCTTGCA ATTAATGGAC  
 4971 AGGTACACCA GACAAAGGT CACTGTAGAT GAGATTACCA CATTAACTCAT GAATGAGAAA ACAGAACGGG  
 5041 CCAATATCGG CAATTGCGATG GAAGCCTTGT TTCAAGGACG ACTAAGGTAAAGATTGAGATCGATGT  
 5111 GAAGACAGTT CCCCCCTCTG AGTCGATCAG TGATTTGITA CAAGCAGTGG ATTCTCAAGA GTTGTAGGGAT  
 5181 TACTGTGAGA AGAAAGGCTG GATCGTTAAC GTTACTAGCC AGATTCAACT AGAAAGGAAC ATCAATAGGG  
 5251 CCATGACTAT ACTCCAAGCT GTTACACAT TCGCAGCAGT CGCAGGAGTA GTGTATGAA TGTACAAACT  
 5321 CTTCGGCGGT CAACAGGGTG CATAACTGG CTTGCCAAC AAAAACCCA ATGTCCTAC TATCAGAGTC  
 5391 GCTAAAGTCC AGGGGCCAGG ATTTGACTAC GCAGTGGCAA TGGCAAAAG AAACATAGTT ACTGCAACCA  
 5461 CCACCAAGGG TGAATTTACC ATGCTAGGGG TGATGATAA TGTAGCAATA TTGCCAACCC ATGCCGCTCC  
 5531 AGGAGAAACC ATTATTATTG ATGGGAAAGA AGTAGAGATC CTAGACGCCA GAGCCTTAGA AGATCAAGCG  
 5601 GGAACCAATC TTGAGATCAC CATTATTACT CTAAAAGAA ATGAGAAATT TAGAGACATC AGATCACATA  
 5671 TTCCCAACCA AATTACTGAA ACTAACGATG GAGTGTGAT CGTGAACACT AGCAAGTACC CCAATATGTA  
 5741 TGTCCCCGTT GTGCTGTGA CGAACACAGGG ATATCTTAAT CTCAGTGGAC GTCAAACCTGC TCGCACTTTA  
 5811 ATGTACAACCTT CTCACAAAGG GGCAGGGCAG TGCGGAGGAA TCATCAGTGC TACTGGCAAA GTCATTGGGA  
 5881 TGATGTGG CGGAAACGGT CTCATGGGT TTGCTGCAGC CCTCAAGGCA TCAACTTCA CTCAAAATCA  
 5951 GGGCGAAATC CAGTGGATGA GGTCACTAA AGAAGTGGGG TACCCCTTA TAAATGCCCA ATCCAAGACA  
 6021 AAGTTAGAAC CCAGTGTCTT CCACTATGTT TTGTAAGGTG TTAAGGAACC AGCTGTACTC ACTAAGAATG  
 6091 ACCCCAGACT AAAAACAGAT TTGAGAAGAG CCATCTTTC TAAATATGTG GGGAAACAAA TTACTGAAGT  
 6161 GGACGAGTAC ATGAAAGAAG CAGTGGATCA CTATGCAGGA CAGTTAATGT CACTGGATAT CAACACAGAA  
 6231 CAGATGTGCC TGGAGGATGC CATGTACGGT ACCGATGGTC TTGAGGCCCT GGATCTTAGC ACTAGTGCTG  
 6301 GATATCCCTA TTGTTGCAATG GGGAAAAAGA AAAGAGACAT TCTAGATAAA CAGACCAGAG ATACTAAGGA  
 6371 GATCAGAGA CTTTTAGATA CCTATGGAAT CAATCTACCA TTAGTCACGT ACGTGAAAGA TGAACACTCAGG  
 6441 TCAAAGACTA AAGTGGAAACA AGGAAGTCA AGATTGATTG AAGCTTCAG CCTTAATGAT TCAGTTGCCA  
 6511 TGAGAATGGC TTGGCAAT CTTACGCAG CTTTCCACAA GAATCCAGGT GTGGTGCAG GATCAGCAGT  
 6581 TGTTGTGAC CCAGATTGT TTTGGAGTAA GATACCACTG CTAATGGAAAG AAAACTCTT CGCTTTGAC  
 6651 TACACAGGGT ATGATGCCCTC ACTCAGCCCT GCTTGGTTG AAGCTTTAA AATGGTGTAA GAAAAAAATTG  
 6721 GATTGGCAG TAGAGTAGAC TATATAGACT ACCGAACCA CTCTCACCAC CTTTACAAAA ACAAGACTTA  
 6791 TTGTTGCAAA GGCAGCATGC CATCCGGCTG CTCTGGCACC TCAATTTCATCA ACTCAATGAT TAAACACCTG  
 6861 ATCATTAGGA CGCTTITACT GAGAACCTAC AAGGGCATAG ACTTGGACCA TTAAAAATG ATTGCCTATG  
 6931 GTGATGACGT GATAGCTTCC TACCCCATG AGGTTGACGC TAGTCCTTA GCCCAATCAG GAAAAGACTA  
 7001 TTGACTAACCTG ATGACTCCAG CAGATAAAATC AGCAACCTT GAAACAGTC CATGGGAGAA TGTAACTT  
 7071 CTGAAAAGAT TTTCAGAGC AGATGAGAAG TATCCATTCC TTGTCATCC AGTGTGCA ATGAAAGAAA  
 7141 TTCAACGAAATC AATCAGATGG ACCAAGGACC CTAGAACAC ACAGGATCAC GTACGCTCGT TTGCTTATG  
 7211 AGCTGGCAC AACGGTGAAG AAGAATACAA TAAATTTTAA GCTAAAATCA GAAGTGTGCC ATGCGGAAGA  
 7281 GCTTTATTGC TCCAGAGTA CTCTACATTG TACCGCCGAT GGCTCGACTC ATTGTTAGTAA GGCGCGCCGC  
 7351 ACAGCTGGTT GAAGGGGACC AACGATACAG CTGGTTGAAG GGGACCAACT GGAGCCACAC ACTTCCTTAC  
 7421 ATTCCATCAC CCACACACTT CCTTACATT CACGGCGCGC CGACTCATTT TAGTAACCTT ACCTCAGTCG  
 7491 GATGGATTG GTTGTACTG TTGTTAGGGT AAATTTTCT TTAATTGGA GAAAAAAAGG AAAAAGGAAA  
 7561 GAGCTCCAA TCACTAGTGA ATTCCGGCC GCGTCAGGT CGACCATATG GGAGAGCTCC CAACCGCTTG  
 7631 GATGCATAGC TTGAGTATTC TATAGTGTCA CCTAAATAGC TTGGCGTAAT CATGGTCATA GCTGTTCT  
 7701 GTGTGAAATT GTTATCCGCT CACAATCCA CACACACATAC GAGCCGGAAG CATAAAAGTGT AAAGCCTGG  
 7771 GTGCTTAATG AGTGGACTAA CTACACATTAA TTGCGTTGCG CTCACGTGCC GCTTTCCAGT CGGGAAACCT  
 7841 GTCGTGCAG CTGCATTAAT GAATCCGCCA ACCCGCGGGG AGAGGGCGGTT TGCGTATTGG GCGCTCTTCC  
 7911 GCTTCTCGC TCACTGACTC GCTGCCTCG GTCGTTCGGC TGCGCGAGC GGTATCAGCT CACTCAAAGG  
 7981 CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCGAGG AAAGAACATG TGAGCAAAAG GCCAGCAAAA

FIG. 12 (Continued)

8051 GGCCAGGAAC CGTAAAAAAGC CCGCGTTGCT GGC GTTTTC C ATAGGCTCC GCCCCCTGA CGAGCATCAC  
 8121 AAAAATCGAC GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAAG ATACCAGGCG TTCCCCCTG  
 8191 GAAGCTCCCT CGTGCCTCT CCTGTTCCGA CCCTGCCGT TACCGGATAC CTGTCGCCCT TTCTCCCTTC  
 8261 GGGAAAGCGTG GCGCTTCTC ATAGCTCACG CTGAGGTAT CTCAGTTCGG TGAGGTCTG TCGCTCCAAG  
 8331 CTGGCTGTG TGCACGAACC CCCGTTCAAG CCCGACCGCT GCGCCTTATC CGTAACAT CGTCTTGAGT  
 8401 CCAACCCGGT AAGACACGAC TTATGCCAC TGCGACGAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA  
 8471 TGAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC TACGGCTACA CTAGAAGAAC AGTATTTGGT  
 8541 ATCTCGCTC TGCTGAAGCC AGTTACCTTC GGAAAAAGAG TTGGTAGCTC TTGATCCGGC AAACAAACCA  
 8611 CCGCTGGTAG CGGTGGTTT TTGTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT CTCAGAAAGA  
 8681 TCCTTGTATC TTTCACGG GGTCTGACGC TCAGTGGAAC GAAAATCAC GTTAAGGGAT TTGGTCATG  
 8751 AGATTATCAA AAAGGATCTT CACCTAGATC CTTTAAATT AAAAATGAAG TTAAATCTCA ATCTAAAGTA  
 8821 TATATGAGTA AACTGGTCT GACAGT TACGCTTAAT CAGTGGAGCA CCTATCTAG CGATCTGTCT  
 8891 ATTICGTICA TCCATAGTT CCTGACTCCC CGTGTGAG ATAACCTAGA TACGGGAGGG CTTACCATCT  
 8961 GGCCCCAGTG CTGCAATGAT ACCGGAGAC CCAGCTCAC CGGCTCAGA TTATCAGCA ATAAACCAGC  
 9031 CAGCCGGAAG GGCGAGCGC AGAAGTGGTC CTGCAACTTT ATCCGCTCC ATCCAGTCTA TTAATTGTTG  
 9101 CCGGGAAGCT AGAGTAAGTA GTTCCAGT TAATAGTTG CGAACGTTG TTGCCATTG TACAGGCATC  
 9171 GTGGTGTAC GCTCGTGTG TGGTATGGCT TCATTAGCT CCGGTTCCA ACAGTCAAGG CGAGTTACAT  
 9241 GATCCCCAT GTTGTGCAA AAGCGGTTA GCTCTTCGG TCCCTCGATC GTGTAGAA GTAAGTTGGC  
 9311 CGCAGTGTAA TCACTCATGG TTATGGCAGC ACTGCATAAT TCTCTACTG TCATGCCATC CGTAAGATGC  
 9381 TTTCCTGTGA CTGGTGAGTA CTCAACCAAG TCATCTGAG AATAGTGTAT GCGGGGACCG AGTTGCTCTT  
 9451 GCGCCGCGTC AATACGGGAT AATACCGCGC CACATAGCAG AACTTTAAA GTGCTCATCA TTGGAAAACG  
 9521 TTCTCGGGG CGAAAACCTA CAAGGATCTT ACCCGCTTG AGATCCAGT CGATGTAACC CACTCGTGC  
 9591 CCCAACTGAT CTTCAAGC ATC TTTCAGTTT ACCAGCGTTT CTGGGTGAGC AAAAACAGGA AGCAAAATG  
 9661 CCGCAAAAAA GGGATAAAGG GCGACACGGA ATGTGAAAT ACTCATAC TTTCTTTTC AATATTATTG  
 9731 AAGCAATTAT CAGGGTTATT GTCTCATGAG CGGATACATA TTTGAATGTA TTAGAAAAA TAAACAAATA  
 9801 GGGGTTCCGC GCACATTTCC CCGAAAAGTG CCACCTGATG CGGTGTGAAA TACCGCACAG ATGCGTAAGG  
 9871 AGAAAATACC GCATCAGGAA ATTGTAAGCG ITAATATTTT GTAAAATIC GCGTAAATT TTGTTAAAT  
 9941 CAGCTCATTI TTTAACCAAT AGGCCAAAT CGCAGAACAT CCTTATAAT CAAAAGAATA GACCGAGATA  
 10011 GGGTTGAGTG TTGTTCCAGT TTGGAACAAG AGTCCACTAT TAAAGAACGT GGACTCCAAC GTCAAAGGGC  
 10081 GAAAAACCGT CTATCAGGGC GATGGCCAC TACGTGAACC ATCACCTAA TCAAGTTTTT TGGGGTGCAG  
 10151 GTGCCGTAAA GCACTAAAT GGAACCTAA AGGGAGCCCC CGATTTAGAG CTTGACGGGG AAAGCCGGCG  
 10221 AACGTGGCGA GAAAGGAAGG GAAGAAAAGCG AAAGGAGCGG GCGCTAGGGC GCTGGCAAGT GTAGCGGTCA  
 10291 CGCTCGCGT ACCACCAAC CCCGCCGC TIAATGCGCC GCTACAGGGC GCGTCCATTG GCCATTGAG  
 10361 CTGCGCAACT GTTGGGAAGG GCGATCGGTG CGGGCTCTT CGCTATTACG CCAGCTGGCG AAAGGGGGAT  
 10431 GTGCTGCAAG GCGATTAAGT TGGGTAACGC CAGGGTTTTC CCAGTCACGA CGTTGAAAAA CGACGGCCAG  
 10501 TGAATTGTA TACGACTCAC TATAGGGCGA ATTGGGCCCG ACGTCGCATG CTCCCGGCCG CCATGGCGGC  
 10571 CGCGGAATT CGATTGAGGC ATGCTAATAC GACTCACTAT AGG

FIG. 13

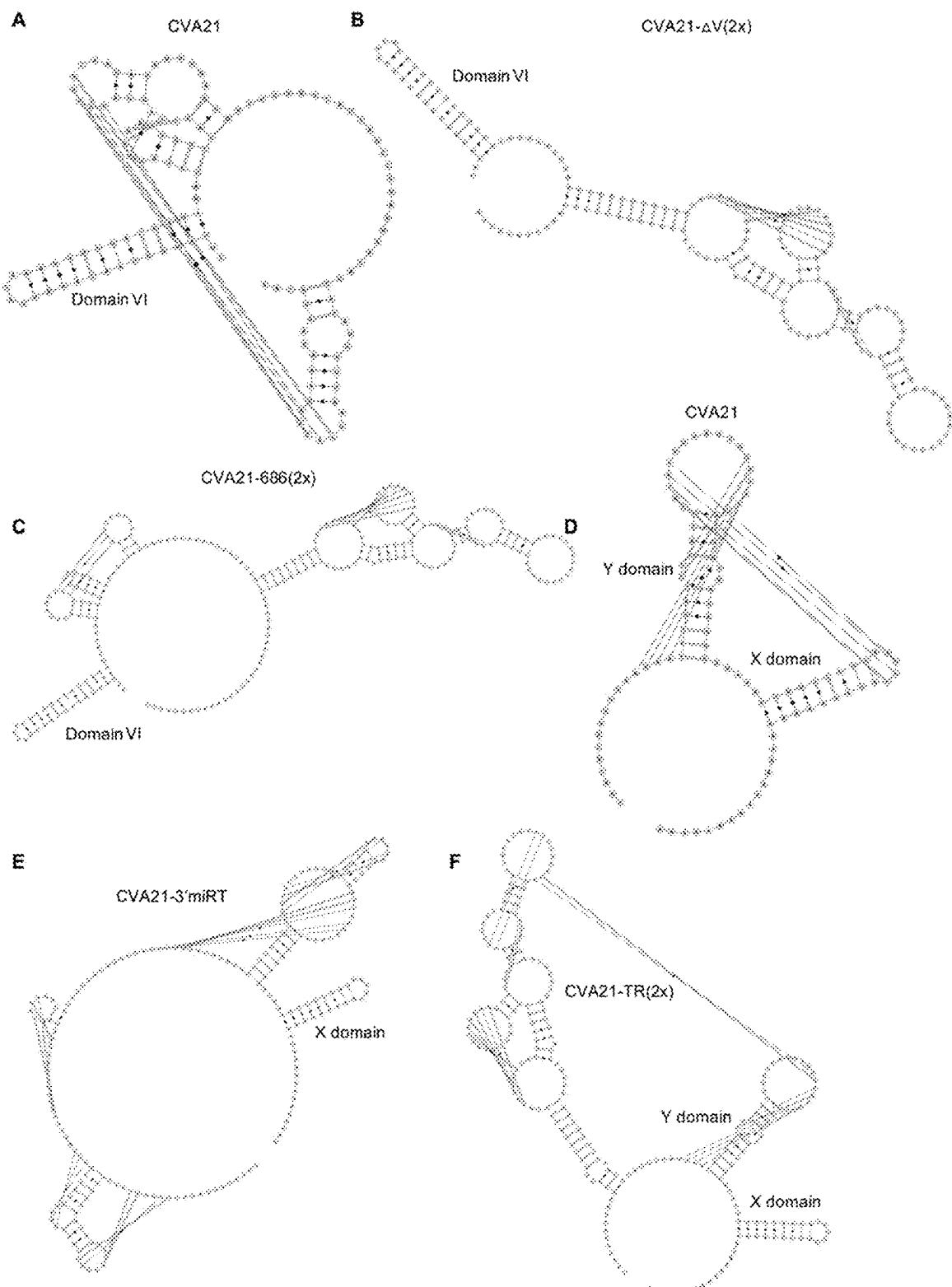


FIG. 14

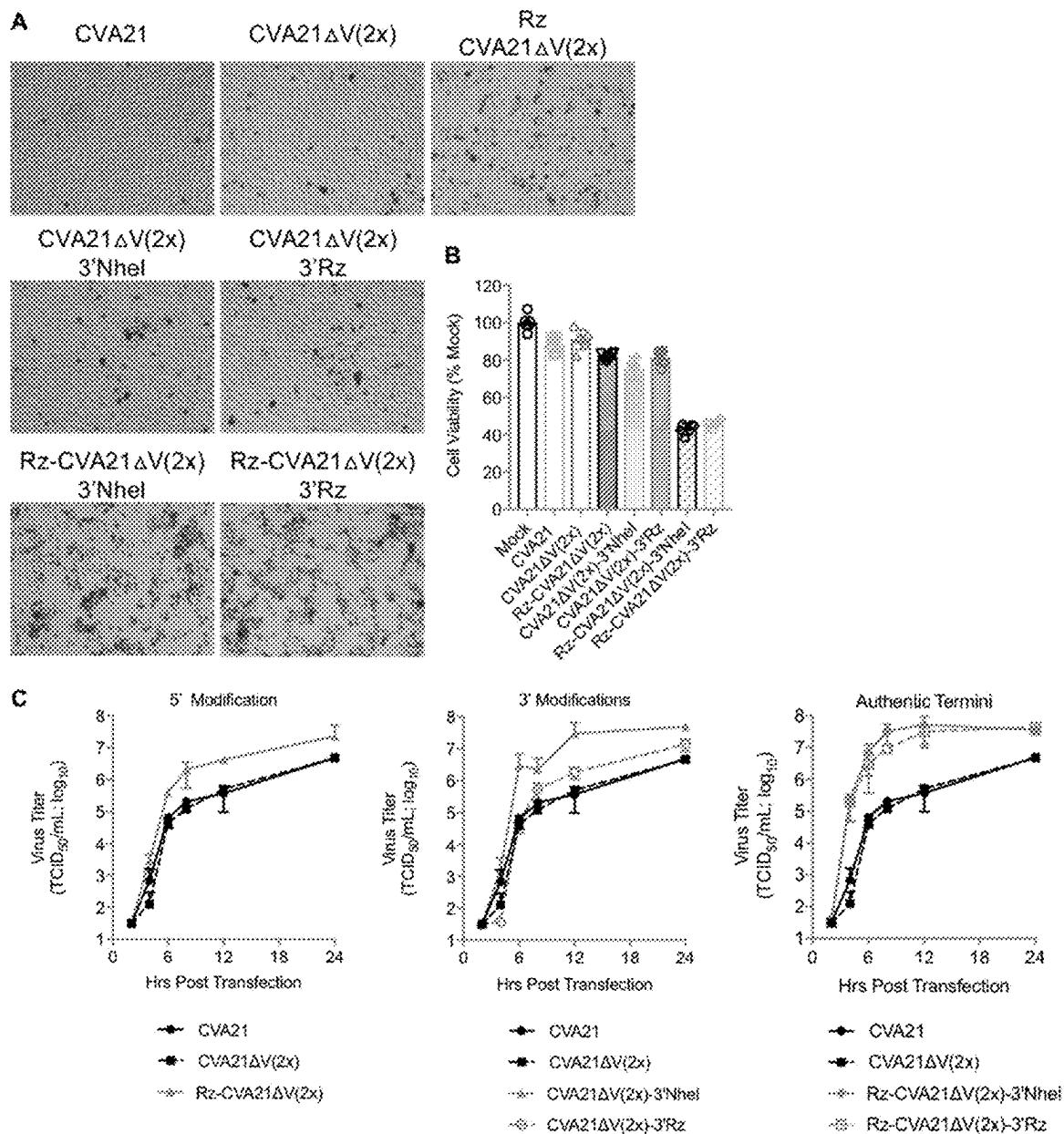


FIG. 15

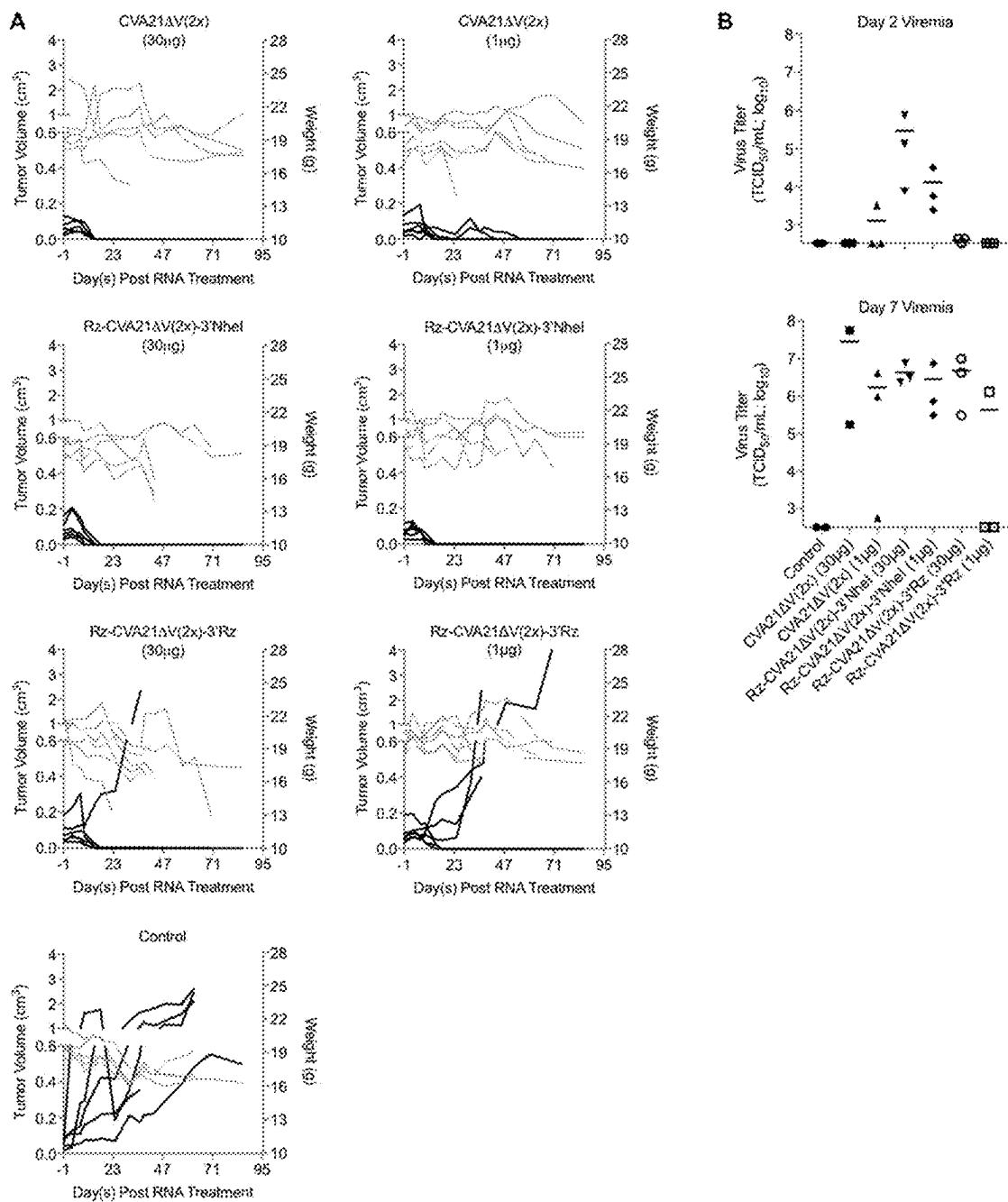


FIG. 16

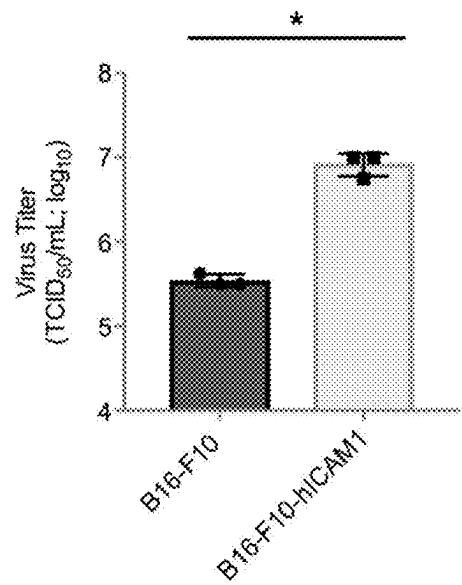
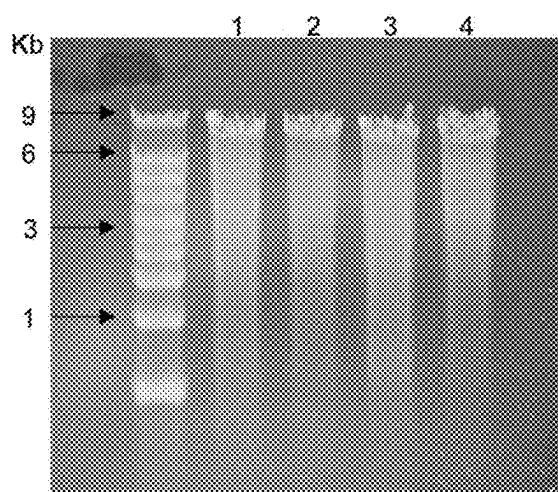


FIG. 17



## USING INFECTIOUS NUCLEIC ACID TO TREAT CANCER

### CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application is a continuation of U.S. application Ser. No. 16/975,782, filed on Aug. 26, 2020, which is a National Stage application under 35 U.S.C. § 371 of International Application No. PCT/US2019/021850, having an International Filing Date of Mar. 12, 2019, which claims priority to U.S. Application Ser. No. 62/641,814, filed on Mar. 12, 2018. The disclosure of the prior applications are considered part of (and are incorporated by reference in) the disclosure of this application.

### SEQUENCE LISTING

[0002] This application contains a Sequence Listing that has been submitted electronically as an XML file named “07039-1756002\_SL\_ST26.XML.” The XML file, created on Mar. 25, 2025, is 125,273 bytes in size. The material in the XML file is hereby incorporated by reference in its entirety.

### BACKGROUND

#### 1. Technical Field

[0003] This document relates to methods and materials involved in treating cancer with viral nucleic acid (e.g., infectious nucleic acid encoding for a picornavirus).

#### 2. Background Information

[0004] The genus enterovirus is a member of the family Picornaviridae and is comprised of four species of human enteroviruses, HEV-A to -D. Enteroviruses are small non-enveloped viruses with positive-sense single-stranded RNA genomes that are around 7500 nucleotides in length. Following cellular internalization, the virion is uncoated, and the genomic RNA is immediately translated into a single large polyprotein, which is processed by the virally-encoded protease (3C) to yield the capsid proteins and non-structural proteins involved in viral replication. A negative-sense strand RNA is synthesized from the positive-sense viral RNA by the virally encoded RNA-dependent RNA polymerase (RdRp-3D). The minus-strand RNA serves as a template for synthesizing more positive-sense RNA molecules that can be translated, replicated or packaged. The virus progeny accumulate in the cytoplasm until the virus induces cell lysis releasing the progeny into the environment.

[0005] Coxsackievirus A21 (CVA21) is a naturally occurring HEV-C picornavirus known to cause upper respiratory infections or in more severe cases myositis in humans. Infants and immunocompromised individuals have a greater chance to develop more severe complications. Intracellular adhesion molecule 1 (ICAM-1) is the main receptor for the virus, and CVA21 has shown potent oncolytic activity against a variety of ICAM-1 expressing cancer cells. However, immunocompromised mice bearing human melanoma and myeloma xenografts develop rapid onset lethal myositis (presented as flaccid limb paralysis) following treatment with CVA21 formulated as virus particles or as infectious nucleic acid.

### SUMMARY

[0006] This document provides methods and materials related to using nucleic acid (e.g., infectious nucleic acid) encoding viruses to reduce the number of viable cancer cells within a mammal. For example, this document provides methods for using infectious nucleic acid to treat cancer, engineered viral nucleic acid, and methods for making engineered viral nucleic acid.

[0007] MicroRNA-targeting can be used to regulate the tropism of positive-sense RNA viruses. This technique exploits tissue-specific microRNAs (miRNA) expressed in host cells where viral replication is undesirable (Kelly et al., *Nat. Med.*, 14:1278-1283 (2008); Barnes et al., *Cell Host Microbe*, 4:239-248 (2008); and Ruiz et al., *J. Virol.*, 90:4078-4092 (2016)). MiRNA target sequences can be inserted into the viral genome such that the genome and mRNAs are recognized by the host miRNAs and subsequently degraded or translationally repressed, preventing viral replication and toxicity. Cancer cells often have dysregulated miRNA levels, allowing permissive virus replication and subsequent tumor cell destruction.

[0008] The shapes of RNA molecules, including viral genomes, are specifically designed to regulate stability, intra- and inter-molecular interactions, and activity. Therefore, it is reasonable to speculate that insertion of additional nucleotides (e.g., miRNA targets) anywhere within the viral genome has the potential to significantly offend a variety of the biological properties of the virus. For example, microRNA target sequences can be inserted into the viral genome in the 3' UTR, and the rescued virus can replicate with kinetics similar to the unmodified virus, maintain oncolytic activity against human xenografts in vivo, and reduce toxicity in normal tissues as described elsewhere (Kelly et al., *Nat. Med.*, 14:1278-1283 (2008)). However, rescue of this virus following transfection of RNA transcripts encoding this modified viral genome was significantly delayed.

[0009] In addition, wild-type viruses can exhibit significant toxicity. For example, injection of RNA transcripts encoding the wild-type viral genome into mice bearing human melanoma or myeloma xenografts nucleated a spreading viral infection resulting in tumor reduction and lethal toxicity as described elsewhere (Hadac et al., *Molecular Therapy*, 19 (6): 1041-1047 (2011)). In contrast, injection of RNA transcripts encoding the Kelly et al. miRNA-targeted viral genome did not mount a spreading infection and thus does not have therapeutic value as infectious nucleic acid.

[0010] As described herein, microRNA-targeted oncolytic viruses were formulated as infectious nucleic acid using a coxsackievirus A21 backbone. Various alternative insertion mechanisms were designed, and the viral genomes were constructed. In addition, the virus rescue kinetics from RNA transcripts, the rescued virus replication kinetics, microRNA-target stability, oncolytic activity, and toxicity were evaluated; each in comparison to the unmodified genome and the Kelly et al. miRNA-targeted genome. Each tested construct exhibited enhanced viral rescue kinetics compared to the Kelly et al. construct and maintained oncolytic activity. In addition, one tested construct (designated CVA21-ΔV<sub>2x</sub> herein) ameliorated toxicity and significantly prolonged overall survival.

[0011] The CVA21-ΔV<sub>2x</sub> construct is an example of a microRNA-targeted viral genome formulated as infectious

nucleic acid that exhibits an acceptable therapeutic index that was experimentally validated. The constructs provided herein can be used for anti-cancer therapy, vaccination with enhanced safety, or segregation of viral growth in various producer and target cells, facilitating manufacturing and experimental evaluation of the virus life cycle and the role of different cells in pathogenesis. Additionally, the techniques for microRNA-targeting provided herein can be used with other type I IRES encoding picornaviruses.

**[0012]** In general, one aspect of this document features a nucleic acid construct comprising (or consisting essentially of or consisting of) an infectious nucleic acid comprising (or consisting essentially of or consisting of) a picornavirus genome comprising (or consisting essentially of or consisting of) one or more heterologous sequence elements of 20 or more bases, where the specific infectivity of the construct is sufficient to initiate a spreading picornavirus infection when administered to a living mammal, where the specific infectivity of the construct is of similar magnitude to the specific infectivity of a comparable construct lacking the one or more heterologous sequence elements. The mammal can be a human. The construct can be formulated as plasmid DNA. The construct can be formulated as an RNA molecule. At least one of the one or more heterologous sequence elements can be a microRNA response element. A microRNA target element of the microRNA response element can comprise (or can consist essentially of or can consist of) at least a region of complementarity to a microRNA present in non-cancer cells. At least one of the one or more heterologous sequence elements can be a tissue-specific microRNA response element. At least one of the one or more heterologous sequence elements can be a muscle-specific, brain-specific, or heart-specific microRNA response element. At least one of the one or more heterologous sequence elements can be inserted into the 5' non-coding region of the picornavirus genome as a substitution for nucleotides within a scanning region. The picornavirus genome can comprise a type I internal ribosome entry site. The picornavirus genome can be a coxsackievirus, poliovirus, echovirus, rhinovirus, or enterovirus genome. The picornavirus genome can be a coxsackievirus A21 genome. The picornavirus genome can comprise a microRNA target element for miR-133. The picornavirus genome can comprise more than one microRNA target element for miR-133. The picornavirus genome can comprise a microRNA target element for miR-206. The picornavirus genome can comprise more than one microRNA target element for miR-206. The picornavirus genome can comprise more than one microRNA target element for miR-133 and more than one microRNA target element for miR-206. The picornavirus genome can be a coxsackievirus A21 genome, and the picornavirus genome can lack at least 20 nucleotides from position 631 to position 698 of a wild-type coxsackievirus A21 genome. The picornavirus genome can be a coxsackievirus A21 genome, and the picornavirus genome can lack nucleotides 631 to 698 of a wild-type coxsackievirus A21 genome. In some cases, the nucleic acid construct can be DNA, and the nucleic acid construct can encode a ribozyme. The ribozyme can be designed to cleave a 5' portion of RNA from an RNA molecule transcribed from the nucleic acid construct, thereby creating a ribozyme-cleaved RNA. The ribozyme-cleaved RNA can have a 5' end with no ribonucleotides that are not present in a picornavirus genome. In some cases, the nucleic acid construct can be RNA, and the nucleic acid

construct can include a ribozyme. The ribozyme can be designed to cleave a 5' portion of RNA from the nucleic acid construct, thereby creating a ribozyme-cleaved RNA. The ribozyme-cleaved RNA can have a 5' end with no ribonucleotides that are not present in a picornavirus genome. In some cases, the nucleic acid construct can be DNA, and the nucleic acid construct can include a restriction endonuclease cut site. The restriction endonuclease cut site can be designed to allow for cleavage, via a restriction endonuclease, of a 3' portion of said nucleic acid construct, thereby creating a restriction endonuclease-cleaved nucleic acid construct. The restriction endonuclease-cleaved nucleic acid construct can encode a virus having a 3' end with less than 10 ribonucleotides that are not present in a picornavirus genome (e.g., a picornavirus genome having a 3' end with no ribonucleotides that are not present in a picornavirus genome). In some cases, the nucleic acid construct can be DNA, the nucleic acid construct can encode a ribozyme, and the nucleic acid construct can include a restriction endonuclease cut site. The ribozyme can be designed to cleave a 5' portion of RNA from an RNA molecule transcribed from the nucleic acid construct, thereby creating a ribozyme-cleaved RNA. The ribozyme-cleaved RNA can have a 5' end with no ribonucleotides that are not present in a picornavirus genome. The restriction endonuclease cut site can be designed to allow for cleavage, via a restriction endonuclease, of a 3' portion of the nucleic acid construct, thereby creating a restriction endonuclease-cleaved nucleic acid construct. The restriction endonuclease-cleaved nucleic acid construct can encode a virus having a 3' end with less than 10 ribonucleotides that are not present in a picornavirus genome (e.g., a virus having a 3' end with no ribonucleotides that are not present in a picornavirus genome).

**[0013]** In another aspect, this document features a method of reducing the number of cancer cells within a living mammal. The method comprises (or consists essentially of or consists of) administering a construct to the mammal. The mammal can be a human. The construct can be formulated as plasmid DNA. The construct can be formulated as an RNA molecule. At least one of the one or more heterologous sequence elements can be a microRNA response element. A microRNA target element of the microRNA response element can comprise (or can consist essentially of or can consist of) at least a region of complementarity to a microRNA present in non-cancer cells. At least one of the one or more heterologous sequence elements can be a tissue-specific microRNA response element. At least one of the one or more heterologous sequence elements can be a muscle-specific, brain-specific, or heart-specific microRNA response element. At least one of the one or more heterologous sequence elements can be inserted into the 5' non-coding region of the picornavirus genome as a substitution for nucleotides within a scanning region. The picornavirus genome can comprise a type I internal ribosome entry site. The picornavirus genome can be a coxsackievirus, poliovirus, echovirus, rhinovirus, or enterovirus genome. The picornavirus genome can be a coxsackievirus A21 genome. The picornavirus genome can comprise a microRNA target element for miR-133. The picornavirus genome can comprise more than one microRNA target element for miR-133. The picornavirus genome can comprise a microRNA target element for miR-206. The picornavirus genome can comprise more than one microRNA target element for miR-206. The picornavirus genome can be a coxsackievirus A21 genome, and the picornavirus genome can lack at least 20 nucleotides from position 631 to position 698 of a wild-type coxsackievirus A21 genome. The picornavirus genome can be a coxsackievirus A21 genome, and the picornavirus genome can lack nucleotides 631 to 698 of a wild-type coxsackievirus A21 genome. In some cases, the nucleic acid construct can be DNA, and the nucleic acid construct can encode a ribozyme. The ribozyme can be designed to cleave a 5' portion of RNA from an RNA molecule transcribed from the nucleic acid construct, thereby creating a ribozyme-cleaved RNA. The ribozyme-cleaved RNA can have a 5' end with no ribonucleotides that are not present in a picornavirus genome. In some cases, the nucleic acid construct can be RNA, and the nucleic acid

microRNA target element for miR-133 and more than one microRNA target element for miR-206. The picornavirus genome can be a coxsackievirus A21 genome, and the picornavirus genome can lack at least 20 nucleotides from position 631 to position 698 of a wild-type coxsackievirus A21 genome. The picornavirus genome can be a coxsackievirus A21 genome, and the picornavirus genome can lack nucleotides 631 to 698 of a wild-type coxsackievirus A21 genome. The cancer cells can be melanoma, myeloma, pancreatic, prostate, bladder, non-small cell lung, or breast cancer cells. The administering step can result in a reduced number of non-cancerous cells present within the mammal undergoing cell lysis following the administering step as compared to the number of non-cancerous cells that undergo lysis when the comparable construct is administered to a comparable mammal. The administering step can result in a similar number or an increased number of cancerous cells present within the living mammal undergoing cell lysis following the administering step as compared to the number of cancerous cells that undergo lysis when the comparable construct is administered to a comparable mammal.

[0014] In another aspect, this document features a method of reducing the number of cancer cells within a living mammal. The method comprises (or consists essentially of or consists of) administering a construct to the mammal, where the cancer cells undergo lysis as a result of unencumbered synthesis of corresponding picornaviruses from the construct, thereby reducing the number of cancer cells within the living mammal. The mammal can be a human. The construct can be formulated as plasmid DNA. The construct can be formulated as an RNA molecule. At least one of the one or more heterologous sequence elements can be a microRNA response element. A microRNA target element of the microRNA response element can comprise (or can consist essentially of or can consist of) at least a region of complementarity to a microRNA present in non-cancer cells. At least one of the one or more heterologous sequence elements can be a tissue-specific microRNA response element. At least one of the one or more heterologous sequence elements can be a muscle-specific, brain-specific, or heart-specific microRNA response element. At least one of the one or more heterologous sequence elements can be inserted into the 5' non-coding region of the picornavirus genome as a substitution for nucleotides within a scanning region. The picornavirus genome can comprise a type I internal ribosome entry site. The picornavirus genome can be a coxsackievirus A21, poliovirus, echovirus, rhinovirus, or enterovirus genome. The picornavirus genome can be a coxsackievirus A21 genome. The picornavirus genome can comprise a microRNA target element for miR-133. The picornavirus genome can comprise more than one microRNA target element for miR-133. The picornavirus genome can comprise a microRNA target element for miR-206. The picornavirus genome can comprise more than one microRNA target element for miR-206. The picornavirus genome can comprise more than one microRNA target element for miR-133 and more than one microRNA target element for miR-206. The picornavirus genome can be a coxsackievirus A21 genome, and the picornavirus genome can lack at least 20 nucleotides from position 631 to position 698 of a wild-type coxsackievirus A21 genome. The picornavirus genome can be a coxsackievirus A21 genome, and the picornavirus genome can lack nucleotides 631 to 698 of a wild-type coxsackievirus A21 genome. The cancer cells

can be melanoma, myeloma, or breast cancer cells. The administering step can result in a reduced number of non-cancerous cells present within the mammal undergoing cell lysis following the administering step as compared to the number of non-cancerous cells that undergo lysis when the comparable construct is administered to a comparable mammal. The administering step can result in a similar number or an increased number of cancerous cells present within the living mammal undergoing cell lysis following the administering step as compared to the number of cancerous cells that undergo lysis when the comparable construct is administered to a comparable mammal.

[0015] In another aspect, this document features an isolated infectious nucleic acid encoding a coxsackievirus, where the infectious nucleic acid lacks at least 20 nucleotides from position 631 to position 698 of a wild-type coxsackievirus A21 genome (e.g., the Kuykendall CVA21 strain), and where the infectious nucleic acid comprises a microRNA target element for a muscle-specific microRNA that is located between a VI domain and an authentic translation start site of the infectious nucleic acid. The infectious nucleic acid can lack the nucleotides from position 631 to position 698. The muscle-specific microRNA can be miR-133 or miR-206. The infectious nucleic acid can comprise the microRNA target element between a first position and a second position, where the first position corresponds to position 631 of the wild-type coxsackievirus A21 genome, and where the second position corresponds to position 699 of the wild-type coxsackievirus A21 genome. In some cases, the infectious nucleic acid can be DNA, and the infectious nucleic acid can include a restriction endonuclease cut site. The restriction endonuclease cut site can be designed to allow for cleavage, via a restriction endonuclease, of a 3' portion of the infectious nucleic acid, thereby creating a restriction endonuclease-cleaved infectious nucleic acid. The restriction endonuclease-cleaved infectious nucleic acid can encode a coxsackievirus having a 3' end with less than 10 ribonucleotides that are not present in a picornavirus genome (e.g., a coxsackievirus having a 3' end with no ribonucleotides that are not present in a picornavirus genome). In some cases, the infectious nucleic acid can be RNA, the infectious nucleic acid can encode a ribozyme, and the infectious nucleic acid can include a restriction endonuclease cut site. The ribozyme can be designed to cleave a 5' portion of RNA from an RNA molecule transcribed from the infectious nucleic acid, thereby creating a ribozyme-cleaved RNA. The ribozyme-cleaved RNA can have a 5' end with no ribonucleotides that are not present in a picornavirus genome. The restriction endonuclease cut site can be designed to allow for cleavage, via a restriction endonuclease, of a 3' portion of RNA from the infectious nucleic acid, thereby creating a restriction endonuclease-cleaved infectious nucleic acid. The restriction endonuclease-cleaved infectious nucleic acid can encode a coxsackievirus having a 3' end with less than 10 ribonucleotides that are not present in a picornavirus genome (e.g., a coxsackievirus comprising a 3' end with no ribonucleotides that are not present in a picornavirus genome).

[0016] In another aspect, this document features a method for treating cancer present in a mammal. The method comprises (or consists essentially of or consists of) administering, to the mammal, an effective amount of infectious nucleic acid encoding a coxsackievirus, where the infectious

nucleic acid lacks at least 20 nucleotides from position 631 to position 698 of a wild-type coxsackievirus A21 genome, and where the infectious nucleic acid comprises a microRNA target element for a muscle-specific microRNA that is located between a VI domain and an authentic translation start site of the infectious nucleic acid. The mammal can be a human. The cancer can be melanoma, myeloma, pancreatic, prostate, bladder, non-small cell lung, or breast cancer. The infectious nucleic acid can lack the nucleotides from position 631 to position 698. The muscle-specific microRNA can be miR-133 or miR-206. The infectious nucleic acid can comprise the microRNA target element between a first position and a second position, where the first position corresponds to position 631 of the wild-type coxsackievirus A21 genome, and where the second position corresponds to position 699 of the wild-type coxsackievirus A21 genome.

[0017] In another aspect, this document features a method for making infectious RNA comprising a picornavirus genome comprising one or more heterologous sequence elements of 20 or more bases, where the specific infectivity of the infectious RNA is sufficient to initiate a spreading picornavirus infection when administered to a living mammal, where the specific infectivity of the infectious RNA is of similar magnitude to the specific infectivity of a comparable infectious RNA lacking the one or more heterologous sequence elements. The method comprises (or consists essentially of or consists of) providing an DNA construct encoding the infectious RNA, where the DNA construct encodes a ribozyme, and where the nucleic acid construct includes a restriction endonuclease cut site; and contacting the DNA construct with a restriction endonuclease under conditions where at least a portion of the DNA construct is removed, thereby producing a restriction endonuclease-cleaved DNA construct; and where the restriction endonuclease-cleaved DNA construct encodes an infectious RNA having non-picornavirus RNA located at a 5' end, and where the ribozyme cleaves the non-picornavirus RNA located at the 5' end to generate the infectious RNA. The infectious RNA encoded by the restriction endonuclease-cleaved DNA construct can include a 3' end with less than 10 ribonucleotides that are not present in a picornavirus genome. The infectious RNA encoded by the restriction endonuclease-cleaved DNA construct can include a 3' end with no ribonucleotides that are not present in a picornavirus genome. The infectious RNA can include a 5' end with no ribonucleotides that are not present in a picornavirus genome.

[0018] In another aspect, this document features an immunocompetent model that can be infected by infectious nucleic acid encoding a virus, where the immunocompetent model includes a mouse cell expressing a human ICAM-1. The mouse cell can be a murine melanoma B16-F10 cell.

[0019] The term "specific infectivity" as used herein refers to the ratio between infectious viruses to total nucleic acid molecules (i.e., the number of plaque forming units obtained from a set copy number of viral genomes).

[0020] The term "scanning region" as used herein refers to the space between silent or cryptic AUG sites in internal ribosome entry sites and the authentic AUG. Ribosomes scan this region prior to translation initiation. This space generally ranges from 35-160 nucleotides in picornaviruses.

[0021] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this

invention pertains. Although methods and materials similar or equivalent to those described herein can be used to practice the invention, suitable methods and materials are described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. In case of conflict, the present specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and not intended to be limiting.

[0022] The details of one or more embodiments of the invention are set forth in the accompanying drawings and the description below. Other features, objects, and advantages of the invention will be apparent from the description and drawings, and from the claims.

#### DESCRIPTION OF THE DRAWINGS

[0023] FIG. 1. Schematic representation of potential UTR-localized insertion sites for microRNA response element (RE). (A) Sequences of muscle-detargeting response elements. Top=SEQ ID NO:1; bottom=SEQ ID NO:2. 1× constructs contain one copy each, and 2× constructs contain two copies each of target sequences complementary to miR-133 (underlined) and miR-206 (italics). (B) Diagram of CVA21 genome with RE insertion sites analyzed. Domains I to VI in the 5' UTR and response element insertion sites are labeled. The cryptic AUG site in domain VI is represented by a double line. (C) Left: RNA secondary structure model for CVA21<sub>WT</sub> 3' UTR and location for RE constructed by Kelly et al. (SEQ ID NO:3). Right: Secondary structure model for CVA21-3'TR constructs (SEQ ID NO:4). Residues involved in forming the "kissing domain" are shown with interconnecting lines. Residues replicated to form the TR are labeled. (D) Nomenclature for all microRNA-targeted CVA21 constructs analyzed.

[0024] FIG. 2. Transfection of miRT-CVA21 infectious RNA produces virus progeny at rates similar to wild-type CVA21 RNA. (A) Upper panel: H1-HeLa cells 72 hours post transfection with 2.5 µg of infectious RNA encoding CVA21 or miRT-CVA21. Lower panel: H1-HeLa cells 24 hours post infection with cleared lysates collected from transfected cells. (B-C) Time-course production of infectious virus in H1-HeLa cells (B) and Mel624 (C) cells transfected with 1 µg of infectious RNA encoding CVA21 or miRT-CVA21. Experiments were repeated at least in triplicate. Data is represented as mean viral titer+/-standard deviation.

[0025] FIG. 3. Localization of RE within the 5' UTR enhances regulation of viral tropism. (A) One-step growth curve of unmodified and new miRT-CVA21. (B) Upper: Viability of cells 24 hours post infection at an MOI of 1 with unmodified or new miRT-CVA21 of H1-HeLa cells pre-treated with 100 nM miRNA mimics (labeled on x-axis). Lower: Viral titers of supernatants collected from the corresponding infections. Experiments were conducted in triplicate and data is represented as mean viability or mean virus titer+/-standard deviation.

[0026] FIG. 4. Treatment with CVA21-ΔV(2x) RNA results in complete tumor regression without causing toxicity. (A) Tumor volume and weight measurements from SCID mice bearing subcutaneous Mel624 tumors following IT treatment with saline (n=5) or 30 µg of RNA transcripts encoding CVA21 (n=5), CVA21-3'miRT (n=5), CVA21-ΔV (1x) (n=5), CVA21-ΔV(2x) (n=5), CVA21-686(2x) (n=5), or CVA21-3'TR (1x) (n=4). (B) Proportion of mice that develop toxicities in each group and overall percent sur-

vival. (C) Kaplan-Meier survival graphs of treated mice. Statistical significance of survival between saline and CVA21- $\Delta$ V(2x) RNA treated mice was compared using a log-rank test. (D) Viral loads in sera collected on day 7 post therapy from treated mice. Horizontal lines represent mean viral titers.

[0027] FIG. 5. Oncolytic activity of CVA21- $\Delta$ V(2x) RNA is dose dependent. (A) Tumor volume and weight measurements from SCID mice bearing subcutaneous Mel624 tumors following a single IT injection of saline (n=4), 1  $\mu$ g (n=5), 4  $\mu$ g (n=5), 8  $\mu$ g (n=5), 16  $\mu$ g (n=5), or 32  $\mu$ g (n=5) of CVA21- $\Delta$ V(2x) RNA. (B) Viral loads in sera collected on day 9 post RNA administration from mice treated in A. (C) Kaplan-Meier survival graphs of treated mice.

[0028] FIG. 6. Cytotoxicity of CVA21 and CVA21- $\Delta$ V(2x) in a panel of tumor cell lines. 1 $\times$ 10<sup>4</sup> cells per well of Mel624 human melanoma (A), DU145 human prostate (B), Panc1 human pancreatic ductal adenocarcinoma (C), U266, RPMI-8226, MM-1, Kas6.1, JJN-3, or ARh77 human myeloma (D) cells were plated in 96-well tissue culture dishes. The cells were infected with CVA21 or CVA21- $\Delta$ V(2x) at an MOI of 0.001, 0.01, 0.1, 1, or 10 for 2 hours at 37° C. in serum-free media. Following infection, the media and unincorporated virus was removed and replaced with 100  $\mu$ L complete growth media, and the cells were incubated at 37° C. 72 hours post infection, the cells were assayed for proliferation using a 3-(4,5-dimethylthiazolyl-2)-2,5-Diphenyltetrazolium bromide (MTT) kit (ATCC, Manassas, VA). Myeloma panel (n=5). All other lines (n=3). Data is represented as mean percent cell viability normalized to mock infected cells $\pm$ standard deviation.

[0029] FIG. 7 is a sequence listing of the DNA encoding infectious nucleic acid of wild type CVA21 and the encoded RNA (SEQ ID NO:5).

[0030] FIG. 8 is a sequence listing of the DNA encoding infectious nucleic acid of CVA21- $\Delta$ V(1x) and the encoded RNA (SEQ ID NO:6). miRT(1x) is underlined.

[0031] FIG. 9 is a sequence listing of the DNA encoding infectious nucleic acid of CVA21- $\Delta$ V(2x) and the encoded RNA (SEQ ID NO:7). miRT(2x) is underlined.

[0032] FIG. 10 is a sequence listing of the DNA encoding infectious nucleic acid of CVA21-686 (2x) and the encoded RNA (SEQ ID NO:8). miRT(2x) is underlined.

[0033] FIG. 11 is a sequence listing of the DNA encoding infectious nucleic acid of CVA21-3' TR (1x) and the encoded RNA (SEQ ID NO:9). miRT(1x) is underlined.

[0034] FIG. 12 is a sequence listing of the DNA encoding infectious nucleic acid of CVA21-3' TR (2x) and the encoded RNA (SEQ ID NO:10). miRT(2x) is underlined.

[0035] FIG. 13. Predicted secondary RNA structures of microRNA-detargeted CVA21 non-coding regions versus unmodified. Predicted pseudoknot interactions are depicted by interloop lines. Secondary RNA structures were generated using the IPknot web server ([rtips.dna.bio.keio.ac.jp/ipknot/](http://rtips.dna.bio.keio.ac.jp/ipknot/)) available through the Graduate School of Information Science, Nara Institute of Science and Technology Japan. 5' non-coding region predictions span domain VI and the scanning region and were predicted using level 2 (nested pseudoknots), CONTRAfold scoring model with refinements. 3' non-coding region predictions span the junction between 3D and the entire 3'non-coding region with a 20 nucleotide long poly A tail and were predicted using level 3 (pseudoknotted with nested pseudoknots) prediction, CONTRAfold scoring model, with refinements. 5' non-coding

region predictions include (A) unmodified CVA21; (B) CVA21- $\Delta$ V(2x); and (C) CVA21-686(2x). 3' non-coding region prediction for (D) CVA21; (E) CVA21-3'miRT; and (F) CVA21-TR (2x).

[0036] FIG. 14. Authentic viral genome termini enhance virus recovery rate from in vitro-derived RNA transcripts encoding CVA21- $\Delta$ V(2x). (A) 1 $\times$ 10<sup>5</sup> cells per well of H1-HeLa were transfected with 0.5  $\mu$ g of in vitro-derived RNA transcripts encoding CVA21- $\Delta$ V(2x) with or without a ribozyme at the 5' end of the genome (Rz-CVA21- $\Delta$ V(2x)), either a restriction enzyme site (CVA21- $\Delta$ V(2x)-3'NheI) or a different ribozyme (CVA21- $\Delta$ V(2x)-3'Rz) at the 3' end directly adjacent to the poly A tail, or a combination of the 5' ribozyme and either the restriction enzyme site (Rz-CVA21- $\Delta$ V(2x)-3'NheI) or ribozyme (Rz-CVA21- $\Delta$ V(2x)-3'Rz) at the 3' end. 24 hours post transfection, the cells were stained with trypan blue and the CPE imaged. (B) The viability of cells treated as described in (A) was determined using a 3-(4,5-dimethylthiazolyl-2)-2,5-Diphenyltetrazolium bromide (MTT) kit (ATCC, Manassas, VA). (C) Time-course production of infectious virus in H1-HeLa cells transfected with 0.5  $\mu$ g of infectious RNA encoding the constructs described in (A). The experiment was run in duplicate and data are represented as mean viral titers $\pm$ standard deviations.

[0037] FIG. 15. In vitro-derived RNA transcripts encoding CVA21- $\Delta$ V(2x) with authentic termini modifications mount a spreading oncolytic infection. CB-17 SCID mice bearing subcutaneous Mel624 xenografts were treated intratumorally with saline (n=5), 30  $\mu$ g (n=5) or 1  $\mu$ g (n=6) CVA21- $\Delta$ V(2x), 30  $\mu$ g (n=6) or 1  $\mu$ g (n=6) Rz-CVA21- $\Delta$ V(2x)-3'NheI, or 30  $\mu$ g (n=6) or 1  $\mu$ g of Rz-CVA21- $\Delta$ V(2x)-3'Rz RNA. (A) Tumor volumes (black) and weights (gray lines) of all treated mice. (B) Viral titers in sera collected on day 2 or day 7 post RNA administration from mice treated in A. All data points are distinct animals.

[0038] FIG. 16. CVA21- $\Delta$ V(2x) can replicate in B16-F10 cells stably expressing human intracellular adhesion molecule 1 (hICAM-1), a receptor for CVA21. B16-F10-hICAM-1 cells or the parental B16-F10 cells were infected with CVA21- $\Delta$ V(2x) at an MOI of 1. 24 hours post infection, total virus titer was determined. CVA21- $\Delta$ V(2x) replication was significantly enhanced in cells expressing hICAM-1 compared to the parental cell line (p=0.029). The experiment was run in triplicate, and data are represented as mean viral titers $\pm$ standard deviations.

[0039] FIG. 17. CVA21- $\Delta$ V(2x) in vitro transcription reactions can be scaled to milligram levels and maintain similar integrity. Small scale (10  $\mu$ L) and large scale (1 mL) reactions were run simultaneously and purified using lithium chloride precipitation. 1  $\mu$ g of purified RNA from each reaction was run on an RNA FlashGel (Lonza). Lane 1. 10  $\mu$ L reaction #1. Lane 2. 1 mL reaction #1. Lane 3. 10  $\mu$ L reaction #2. Lane 4. 1 mL reaction #2. Total yields from each reaction were 150.55  $\mu$ g; 7.1768 mg; 146.55  $\mu$ g; and 7.3192 mg, respectively.

#### DETAILED DESCRIPTION

[0040] This document provides methods and materials for using nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus having a genome that includes a type I internal ribosome entry site) to reduce the number of viable cancer cells within a mammal. For example, this document provides methods for using infectious nucleic acid

encoding a virus to reduce the number of viable cancer cells within a mammal. Nucleic acid provided herein can encode any appropriate virus and be used to reduce the number of viable cancer cells within a mammal. In some cases, a virus encoded by nucleic acid provided herein can replicate within a cancer cell. In some cases, infectious nucleic acid encoding a picornavirus can be used. A picornavirus can be an enterovirus (e.g., bovine enterovirus, human enterovirus A, human enterovirus B, human enterovirus C, human enterovirus D, human enterovirus E, poliovirus, porcine enterovirus A, and porcine enterovirus B), a rhinovirus (e.g., human rhinovirus A and human rhinovirus B), a cardiovirus (e.g., encephalomyocarditis virus and theilovirus), an aphthovirus (e.g., equine rhinitis A virus and foot-and-mouth disease virus), an hepatovirus (e.g., hepatitis A virus), a parechovirus (e.g., human parechovirus and Ijungan virus), an erbovirus (e.g., equine rhinitis B virus), a kobuvirus (e.g., aichi virus), or a teschovirus (e.g., porcine teschovirus 1-7 and porcine teschovirus). In some cases, infectious nucleic acid provided herein can encode a coxsackievirus A21 (Shafren et al., *Clin. Cancer Res.*, 10(1 Pt. 1):53-60 (2004)), coxsackievirus B3 (Suskind et al., *Proc. Soc. Exp. Biol. Med.*, 94(2):309-318 (1957)), poliovirus type III (Pond and Manueldis, *Am. J. Pathol.*, 45:233-249 (1964)), echovirus I (Shafren et al., *Int. J. Cancer*, 115(2):320-328 (2005)), or an encephalomyocarditis virus type E (Adachi et al., *J. Neurooncol.*, 77(3):233-240 (2006)).

[0041] Nucleic acid (e.g., infectious nucleic acid) encoding a virus can be any appropriate nucleic acid (e.g., DNA, RNA, or a combination thereof). In some cases, nucleic acid encoding a virus can be a nucleic acid construct. For example, a nucleic acid construct encoding a virus can be a plasmid DNA. For example, a nucleic acid construct encoding a virus can be an RNA molecule.

[0042] In some cases, nucleic acid (e.g., infectious nucleic acid) provided herein can encode a picornavirus having a genome that includes a type I internal ribosome entry site. Enteroviruses and rhinoviruses have type I IRESs. Examples of picornaviruses having a genome that includes a type I internal ribosome entry site include, without limitation, coxsackieviruses (e.g., coxsackievirus A21), polioviruses, echoviruses, enteroviruses (e.g., enterovirus-70), and rhinoviruses.

[0043] Nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) provided herein can be administered directly to cancer cells (e.g., by intratumoral administration) or can be administered systemically (e.g., by intravenous, intraperitoneal, intrapleural, or intra-arterial administration). The amount of infectious nucleic acid administered to a mammal can range from about 10 ng to about 1 mg (e.g., from 100 ng to 500 µg, from about 250 ng to about 250 µg, from about 500 ng to about 200 µg, or from about 1 µg to about 100 µg) per kg of body weight. In some cases, from about 100 ng to about 500 µg of infectious nucleic acid encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be administered as a single intratumoral dose. In some cases, the amount of infectious nucleic acid administered to a mammal can be equal to a virus genome copy number of between about  $3 \times 10^{10}$  to about  $3 \times 10^{14}$  genome copies (e.g., between about  $3 \times 10^{10}$  to about  $3 \times 10^{13}$ , between about  $3 \times 10^{10}$  to about  $3 \times 10^{12}$ , between about  $3 \times 10^{11}$  to about  $3 \times 10^{14}$ , between about  $3 \times 10^{11}$  to about  $3 \times 10^{13}$ , or between about  $3 \times 10^{11}$  to about  $3 \times 10^{12}$  genome copies). For example, infec-

tious nucleic acid provided herein can be administered in an amount such that about  $3 \times 10^{11}$  virus genome copies are delivered to a mammal. In some cases, the amount of administered infectious nucleic acid can be between about  $3 \times 10^{10}$  to about  $3 \times 10^{14}$  virus genome copies per kg of body weight.

[0044] Nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to lack one or more (e.g., one, two, three, four, or more) contiguous nucleotide sequences of 10 or more nucleotides in length present in a wild-type version of that virus. For example, infectious nucleic acid encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can lack at least 10 (e.g., at least 10, 20, 30, 40, 50, 60, or more) contiguous nucleotides normally found between the VI domain of the 5' UTR and the translation start site (e.g., the AUG start site) for the viral polyprotein. When infectious nucleic acid encodes a coxsackievirus A21, the infectious nucleic acid can lack at least 10 (e.g., at least 10, 20, 30, 40, 50, 60, or more) contiguous nucleotides from position 631 to position 698 as found in the wild type coxsackievirus A21 genome. In some cases, infectious nucleic acid encoding a coxsackievirus A21 can be designed to lack all the nucleotides from position 631 to position 698 as found in the wild type coxsackievirus A21 genome.

[0045] As described herein, nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to contain a microRNA target element such that a corresponding microRNA (miRNA, specific miRNAs denoted as miR-#) present within a non-cancer cell can reduce virus gene expression, virus replication, or virus stability in that non-cancer cell. MicroRNAs are small, 21-23 nucleotide, highly conserved regulatory RNAs that can mediate translational repression or, in some cases, mRNA destruction by RISC-induced cleavage. MicroRNAs are present within many mammalian cells and can have a tissue-specific tissue distribution. As such, microRNAs can be used to modulate the tropism of a replicating virus to provide a targeting approach for any virus. The ability of infectious nucleic acid encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) to result in non-cancer cell lysis can be reduced using a microRNA target element having at least a region that is complementary to a microRNA present in the non-cancer cells. For example, wild-type coxsackievirus A21 can infect muscle cells. Thus, microRNA target elements that are complementary to microRNAs present in muscle cells can be incorporated into coxsackievirus A21 infectious nucleic acid to reduce muscle cell lysis. Similarly, the safety of vaccines can be improved by modulating the tropism of a virus. For example, a neuronal and/or brain microRNA target element can be incorporated into a poliovirus to reduce the incidence of poliomyelitis induced by an oral polio vaccine.

[0046] This same approach can be used to reduce non-cancer cell lysis by other viral nucleic acids. For example, microRNA target elements having at least a region that is complementary to the microRNAs set forth in Table 1 can be used to reduce cell lysis of the indicated tissue for the listed viruses as well as for other viruses. Other examples of microRNA target elements that can be designed to reduce viral-mediated cell lysis include, without limitation, those having at least a region complementary to a tissue-specific microRNA listed in Table 2. In some cases, infectious

nucleic acid provided herein can encode a virus and contain a microRNA target element having at least a region complementary to a classified tissue-specific microRNA. MicroRNA target elements can have complete complementarity to a microRNA. In some cases, a microRNA target element can contain mismatches in its complementarity to a microRNA provided that it contains complete complementarity to a seed sequence (e.g., base pairs 2-7) of the microRNA. See, e.g., Lim et al., *Nature*, 433(7027):769-73 (2005)).

TABLE 1

Silencing via incorporated microRNA target elements.		
Virus	Tissue	microRNA
Coxsackievirus A21	Muscle	miR-1
Coxsackievirus A21	Muscle	miR-133
Coxsackievirus A21	Muscle	miR-206
Coxsackievirus B3,	Brain	miR-101
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-124a, b
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-125
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-128
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-131
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-132
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-134
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		

TABLE 1-continued

Silencing via incorporated microRNA target elements.		
Virus	Tissue	microRNA
Coxsackievirus B3,	Brain	miR-135
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-138
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-153
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-183
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-219
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-9
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Brain	miR-95
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Heart	miR-1
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Heart	miR-133
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		
Coxsackievirus B3,	Heart	miR-206
Encephalomyocarditis-E		
Poliovirus III		
Echovirus I		

TABLE 2

Classified tissue-specific microRNAs.			
miRNA	Tissue	Sequence	Reference
miR-1	Muscle	UGGAUAGUAAGAAGUAUGUA (SEQ ID NO: 11)	Rao et al., <i>Proc. Nat'l. Acad. Sci.</i> , 103:8721-8726 (2006).
miR-101	Brain	UACAGUACUGUGAUACUGAAG (SEQ ID NO: 12)	Lagos-Quintana et al., <i>Curr. Biol.</i> , 12:735-739 (2002).
miR-122a	Liver	UGGAGUGUGACAAUGGUGUUUG U (SEQ ID NO: 13)	Fu et al., <i>FEBS Lett.</i> , 579:3849-3854 (2005).
miR-124a, b	Brain	UUAAGGCACGCCGGAAUGCCTA (SEQ ID NO: 14)	Lagos-Quintana et al., <i>Curr. Biol.</i> , 12:735-739 (2002).
miR-125	Brain	UCCCUUGAGACCCUUUAACCUGUG G (SEQ ID NO: 15)	Liu et al., <i>Proc. Nat'l. Acad. Sci.</i> , 101:9740-9744 (2004).

TABLE 2-continued

Classified tissue-specific microRNAs.			
miRNA	Tissue	Sequence	Reference
miR-126AS	Digestive	UCGUACCCGUGAGUAAAUAUGC (SEQ ID NO: 16)	Shingara et al., RNA, 11:1461-1470 (2005).
miR-127	Spleen	UCGGAUCCGUCUGAGCUUGGCU (SEQ ID NO: 17)	Lagos-Quintana et al., <i>Curr. Biol.</i> , 12:735- 739 (2002).
miR-128	Brain	UCACAGUGAACCGGUCUCUUUC (SEQ ID NO: 18)	Liu et al., <i>Proc. Nat'l. Acad. Sci.</i> , 101:9740-9744 (2004).
miR-130	Lung	CAGUGCAAUGUUAAAAGGGCAU (SEQ ID NO: 19)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-132	Brain	UAACAGUCUACAGCCAUGGUUCG (SEQ ID NO: 20)	Lagos-Quintana et al., <i>Curr. Biol.</i> , 12:735- 739 (2002).
miR-133	Muscle	UUGGUCCCCUUCUAAACCAGCUGU (SEQ ID NO: 21)	Rao et al., <i>Proc. Nat'l. Acad. Sci.</i> , 103:8721-8726 (2006).
miR-134	Brain	UGUGACUGGUUGUGACCAGAGGG (SEQ ID NO: 22)	Schratt et al., <i>Nature</i> , 439:283-289 (2006).
miR-135	Brain	UAUGGCCUUUUUAUCCUAUGUG A (SEQ ID NO: 23)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-138	Brain	AGCUGGUGUUGUGAAUC (SEQ ID NO: 24)	Obernosterer et al., RNA, 12:1161-1167 (2006).
miR-142s 5p, 3p	Hematopoietic	CAUAAAGUAGAAAGCACUAC (SEQ ID NO: 25) UGUAGUGUUUCCUACUUUAUGG A (SEQ ID NO: 26)	Chen et al., <i>Science</i> , 303:83-86 (2004).
miR-143	Digestive	UGAGAUGAAGCACUGUAGCUA (SEQ ID NO: 27)	Shingara et al., RNA, 11:1461-1470 (2005).
miR-145	Digestive	GUCCAGUUUUCCCAGGAAUCCC UU (SEQ ID NO: 28)	Shingara et al., RNA, 11:1461-1470 (2005).
miR-148	Liver, Stomach	UCAGUGCACUACAGAACUUUGU (SEQ ID NO: 29)	Shingara et al., RNA, 11:1461-1470 (2005).
miR-15 (Down- regulated)	B-cell lymphocytic leukemia	UAGCAGCACAUAAUGGUUGUG (SEQ ID NO: 30)	Calin et al., <i>Proc. Nat'l. Acad. Sc.</i> , 99:15524-15529 (2002).
miR-150	Spleen	UCUCCCAACCCUUGUACCAGUG (SEQ ID NO: 31)	Shingara et al., RNA, 11:1461-1470 (2005).
miR-151	Spleen	ACUAGACUGAACGUCCUUGAGG (SEQ ID NO: 32)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-152	Liver	UCAGUGCAUGACAGAACUUGGG (SEQ ID NO: 33)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-153	Brain	UUGCAUAGUCACAAAAGUGA (SEQ ID NO: 34)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).

TABLE 2-continued

Classified tissue-specific microRNAs.			
miRNA	Tissue	Sequence	Reference
miR-155	Burkitt's Lymphoma	UUAAUGCUAAUCCUGAUAGGGG (SEQ ID NO: 35)	Metzler et al., <i>Genes Chromosomes Cancer</i> , 39:167-169 (2004).
miR-16 (Down-regulated)	B-cell lymphocytic leukemia	UAGCAGCACGUAAAUAUUGGCG (SEQ ID NO: 36)	Calin et al., <i>Proc. Nat'l. Acad. Sci.</i> , 99:15524-15529 (2002).
miR-17-5p	Lymphoma	CAAAGUGCUUACAGUGCAGGUA GU (SEQ ID NO: 37)	He et al., <i>Nature</i> , 435:828-833 (2005).
miR-181	Hematopoietic	AACAUCAACGCUGUCGGUGAG U (SEQ ID NO: 38)	Chen et al., <i>Science</i> , 303:83-86 (2004).
miR-183	Brain	UAUGGCACUGGUAGAAUUCACU G (SEQ ID NO: 39)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-18a, b	Lymphoma	UAAGGUGCAUCUAGUGCAGAUA (SEQ ID NO: 40) UAAGGUGCAUCUAGUGCAGUUA (SEQ ID NO: 41)	He et al., <i>Nature</i> , 435:828-833 (2005).
miR-192	Kidney	CUGACCUAUGAAUUGACAGCC (SEQ ID NO: 42)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-194	Kidney	UGUAACAGCAACUCCAUGUGGA (SEQ ID NO: 43)	Sun et al., <i>Nucleic Acids Res.</i> , 32:e188 (2004).
miR-195	Hematopoietic	UAGCAGCACAGAAAUAUUGGC (SEQ ID NO: 44)	Baskerville et al., <i>RNA</i> , 11:241-247 (2005).
miR-199	Liver	CCCAGUGUUUCAGACUACCUGUU C (SEQ ID NO: 45)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-19a, b	Lymphoma	UGUGCAAAUCUAUGCAAAACUG A (SEQ ID NO: 46) UGUGCAAAUCCAUGCAAAACUG A (SEQ ID NO: 47)	He et al., <i>Nature</i> , 435:828-833 (2005).
miR-204	Kidney	UUCCCUUUGUCAUCCUAUGCUC (SEQ ID NO: 48)	Sun et al., <i>Nucleic Acids Res.</i> , 32:e188 (2004).
miR-204	Testis	UUCCCUUUGUCAUCCUAUGCUC (SEQ ID NO: 49)	Baskerville et al., <i>RNA</i> , 11:241-247 (2005).
miR-206	Muscle	UGGAAUGUAAGGAAGUGUGUGG (SEQ ID NO: 50)	Rao et al., <i>Proc. Nat'l. Acad. Sci.</i> , 103:8721-8726 (2006).
miR-208	Heart	AUAAGACGAGCAAAAGCUUGU (SEQ ID NO: 51)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-212	Spleen	UAACAGUCUCCAGUCACGGCC (SEQ ID NO: 52)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-215	Liver	AUGACCUAUGAAUUGACAGAC (SEQ ID NO: 53)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).

TABLE 2-continued

Classified tissue-specific microRNAs.			
miRNA	Tissue	Sequence	Reference
miR-215	Kidney	AUGACCUCAUAGAAUUGACAGAC (SEQ ID NO: 54)	Sun et al., <i>Nucleic Acids Res.</i> , 32:e188 (2004).
miR-216	Pancreas	UAAUCUCAGCUGGCAACUGUG (SEQ ID NO: 55)	Sood et al., <i>Proc. Nat'l. Acad. Sc.</i> , 103:2746-2751 (2006).
miR-219	Brain	UGAUUGGUCCAAACGCAAUCU (SEQ ID NO: 56)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-221	Hematopoietic	AGCUACAUUGUCUGCUGGGUUUC (SEQ ID NO: 57)	Felli et al., <i>Proc. Nat'l. Acad. Sc.</i> , 102:18081 18086 (2005).
miR-222	Hematopoietic	AGCUACAUUCUGGUACUGGGUC (SEQ ID NO: 58)	Felli et al., <i>Proc. Nat'l. Acad. Sc.</i> , 102:18081 18086 (2005).
miR-223	Hematopoietic	UGUCAGUUUGUCAAAUACCCC (SEQ ID NO: 59)	Chen et al., <i>Science</i> , 303:83-86 (2004).
miR-24	Lung	UGGCUCAGUUUCAGCAGGAACAG (SEQ ID NO: 60)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-25	Lymphoma	CAUUGCACUUUGUCUCGGUCUGA (SEQ ID NO: 61)	He et al., <i>Nature</i> , 435:828-833 (2005).
miR-30b, c	Kidney	UGUAAACAUCCUACACUCAGCU (SEQ ID NO: 62) UGUAAACAUCCUACACUCUCAGC (SEQ ID NO: 63)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-32	Lung	UAUUGCACAUUACUAAGUUGC (SEQ ID NO: 64)	Sempere et al., <i>Genome Biol.</i> , 5:R13 (2004).
miR-375	Pancreas	UUUGUUCGUUCGGCUCGCGUGA (SEQ ID NO: 65)	Poy et al., <i>Nature</i> , 432:226-230 (2004).
miR-7	Pituitary	UGGAAGACUAGUGAUUUUGUUG (SEQ ID NO: 66)	He et al., <i>Nature</i> , 435:828-833 (2005).
miR-9	Brain	UCUUUGGUUAUCUAGCUGUAUGA (SEQ ID NO: 67)	Sun et al., <i>Nucleic Acids Res.</i> , 32:e188 (2004).
miR-95	Brain	UUCAACGGGUAUUUAUUGAGCA (SEQ ID NO: 68)	Babak et al., <i>RNA</i> , 10:1813-1819 (2004).
miR-99b	Brain	CACCCGUAGAACCGGACCUUGCG (SEQ ID NO: 69)	Liu et al., <i>Proc. Nat'l. Acad. Sc.</i> , 101:9740-9744 (2004).

[0047] Molecular cloning techniques can be used to insert microRNA target elements into nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21). An infectious nucleic acid provided herein can contain one microRNA target element or multiple microRNA target elements (e.g., two, three, four, five, six, seven, eight, nine, ten, 15, 20, 25, 30, or more microRNA target elements). For example, an infectious nucleic acid provided herein can include two different microRNA target elements such as one that is a target of miR-133 and one that

is a target of miR-206. In some cases, an infectious nucleic acid provided herein can include two or more identical microRNA target elements. For example, an infectious nucleic acid provided herein can include two microRNA target elements that each are a target of miR-133 or two microRNA target elements that each are a target of miR-206. In some cases, an infectious nucleic acid provided herein can include two or more (e.g., two, three, four, or more) microRNA target elements that each are a target of miR-133

and two or more (e.g., two, three, four, or more) microRNA target elements that each are a target of miR-206.

**[0048]** As described herein, one or more (e.g., one, two, three, four, five, six, or more) microRNA target elements can be inserted into nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) between the VI domain of the 5' UTR and the translation start site (e.g., the AUG start site) for the viral polyprotein. When infectious nucleic acid encodes a coxsackievirus A21, the infectious nucleic acid can include one or more (e.g., one, two, three, four, five, six, or more) microRNA target elements inserted between position 631 and position 698 as found in the wild type coxsackievirus A21 genome. In some cases, an infectious nucleic acid encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) provided herein can lack at least 10 (e.g., at least 10, 20, 30, 40, 50, 60, or more) contiguous nucleotides normally found between the VI domain of the 5' UTR and the translation start site (e.g., the AUG start site) for the viral polyprotein and can include within this same location one or more (e.g., one, two, three, four, five, six, or more) microRNA target elements. In some cases, infectious nucleic acid encoding a coxsackievirus A21 can be designed to lack all the nucleotides from position 631 to position 698 as found in a wild type coxsackievirus A21 genome (e.g., the Kuykendall CVA21 strain) and can include, in place of those removed nucleotides, one or more (e.g., one, two, three, four, five, six, or more) microRNA target elements. Examples of such infectious nucleic acid are set forth in FIGS. 8 and 9. Other examples of infectious nucleic acid provided herein are set forth in FIGS. 10 and 11.

**[0049]** In some cases, microRNA target elements that are complementary to microRNAs that are ubiquitously expressed in normal cells with limited expression in cancer cells can be used to direct cell lysis to cancer cells and not non-cancer cells. For example, when using nucleic acid coding for a virus to treat B-cell lymphocytic leukemia, the nucleic acid (e.g., infectious nucleic acid) can be designed to contain microRNA target elements complementary to microRNAs that are ubiquitously expressed in normal tissue while being downregulated in B-cell lymphocytic leukemia cells. Examples of such microRNAs include, without limitation, miR-15 and miR-16.

**[0050]** Nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to include a ribozyme or nucleic acid encoding a ribozyme. For example, in some cases, an infectious RNA encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to include a ribozyme designed to cleave a portion of RNA from itself, and in some cases, an infectious DNA encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to include a DNA sequence encoding a ribozyme designed to cleave a portion of RNA from the RNA molecule transcribed from the infectious DNA. Such a ribozyme can be a hammerhead ribozyme, a hepatitis delta virus ribozyme, a hairpin ribozyme, a Varkud Satellite ribozyme, a glmS ribozyme, a Twister ribozyme, a Twister sister ribozyme, a Hatchet ribozyme, a Pistol ribozyme, or a synthetic ribozyme. A ribozyme can be designed to remove any portion of RNA from an infectious RNA. For example, a ribozyme can be designed to remove a 5' end portion or a 3' end portion of RNA from an infectious RNA encoding a virus. In some case, infectious nucleic acid (DNA or RNA)

encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to include (a) either nucleic acid encoding a first ribozyme in the case of DNA or a first ribozyme in the case of RNA and (b) either nucleic acid encoding a second ribozyme in the case of DNA or a second ribozyme in the case of RNA. In such cases, the first ribozyme can be designed to remove a 5' end portion of RNA from an infectious RNA encoding a virus and the second ribozyme can be designed to remove a 3' end portion. In some cases, after restriction endonuclease cleavage of nucleic acid encoding a virus, a virus encoded by the nucleic acid encoding a virus can include viral-encoding sequences with less than 88 (e.g., less than 63, or less than 10 such as 5, 4, 3, 2, 1, or 0) non-viral nucleotides (e.g., near authentic termini). In some cases, the resulting infectious RNA encoding a virus after ribozyme cleavage can include viral-encoding sequences with no non-viral sequences (e.g., authentic termini).

**[0051]** Nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to include a restriction endonuclease cut site. For example, nucleic acid (e.g., DNA) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to include a restriction endonuclease cut site designed to cleave a portion of nucleic acid from the infectious nucleic acid. In some cases, DNA (e.g., infectious DNA) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to include a DNA sequence such that the infectious DNA includes a restriction endonuclease cut site capable of being cleaved by a restriction endonuclease that cuts RNA at that restriction endonuclease cut site. In some cases, RNA (e.g., infectious RNA) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to include a restriction endonuclease cut site capable of being cleaved by a restriction endonuclease that cuts RNA and/or DNA at that restriction endonuclease cut site. Any appropriate restriction endonuclease cut site and restriction endonuclease capable of cleaving nucleic acid at that restriction endonuclease cut site can be used. Examples of restriction endonuclease cut sites and restriction endonuclease capable of cleaving nucleic acid at those restriction endonuclease cut sites are provided in Table 3.

TABLE 3

Restriction endonuclease/cut site pairing. R is a purine (A or G). Y is a pyrimidine (A or T).	
Restriction endonuclease	Cut site
NsiI	ATGCA^T
BmtI	GCTAG^C
FseI	GGCCGG^CC
AsiSI	GCGAT^CGC
BstBI	TT^CGAA
NheI	G^CTAGC
MluI	A^CGCGT

TABLE 3-continued

Restriction endonuclease/cut site pairing. R is a purine (A or G). Y is a pyrimidine (A or T).	
Restriction endonuclease	Cut site
HaeII	RGC <sup>G</sup> C <sup>Y</sup>
NspI	RCATG <sup>Y</sup>

[0052] A restriction endonuclease cut site can be designed such that cleavage at that cut site removes any portion of nucleic acid from nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21). For example, a restriction endonuclease cut site can be positioned to remove a 5' end portion and/or a 3' end portion of nucleic acid from nucleic acid encoding a virus. In some case, infectious nucleic acid encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to include one, two, three, four, or more restriction endonuclease cut sites. For example, a first restriction endonuclease cut site can be positioned to remove a 5' end portion of RNA from an infectious RNA encoding a virus and a second endonuclease cut site can be positioned to remove a 3' end portion. When including two or more restriction endonuclease cut sites, the cut sites can be the same (e.g., two NheI cut sites) such that the same restriction endonuclease (e.g., NheI) cuts those sites, or the cut sites can be different (e.g., one NheI cut site and one cut site that is not an NheI cut site) such that different restriction endonucleases (e.g., NheI and one cut site that is not an NheI cut site) cut the different sites. In some cases, after restriction endonuclease cleavage of nucleic acid encoding a virus, a virus encoded by the nucleic acid encoding a virus can include viral-encoding sequences with less than 88 (e.g., less than 63, or less than 10 such as 5, 4, 3, 2, 1, or 0) non-viral nucleotides (e.g., near authentic termini). In some cases, the resulting infectious RNA encoding a virus after restriction endonuclease cleavage can include viral-encoding sequences with no non-viral sequences (e.g., authentic termini).

[0053] In some cases, nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) can be designed to include a combination of one or more ribozymes in the case of infectious RNA (or a combination of sequences encoding one or more ribozymes and one or more restriction endonuclease cut sites in the case of infectious DNA). For example, nucleic acid encoding a virus can include a DNA sequence encoding a ribozyme designed to remove a 5' end portion of the encoded virus, and can include a restriction endonuclease cut site designed to be cleaved by a restriction endonuclease to remove a 3' end portion of the nucleic acid such that encoded virus contains no non-viral sequences (e.g. authentic termini).

[0054] In some cases, when using restriction endonuclease cut site(s), the restriction endonuclease(s) capable of cutting that restriction endonuclease cut site(s) can be exogenously added to a solution containing the nucleic acid (e.g., infectious nucleic acid such as infectious RNA) to be cleaved. For example, an exogenously added restriction endonuclease can be added to a solution containing the infectious nucleic

acid to be cleaved under conditions that allow the restriction endonuclease to cleave the infectious nucleic acid. In some cases, a solution that includes a restriction endonuclease and infectious nucleic acid to be cleaved can be incubated for about 60 minutes to about 300 minutes (e.g., for about 60 minutes to about 240 minutes, for about 60 minutes to about 200 minutes, for about 60 minutes to about 180 minutes, for about 60 minutes to about 150 minutes, for about 60 minutes to about 120 minutes, for about 60 minutes to about 100 minutes, for about 60 minutes to about 90 minutes, for about 90 minutes to about 300 minutes, for about 100 minutes to about 300 minutes, for about 120 minutes to about 300 minutes, for about 150 minutes to about 300 minutes, for about 180 minutes to about 300 minutes, for about 200 minutes to about 300 minutes, for about 240 minutes to about 300 minutes, for about 90 minutes to about 240 minutes, for about 120 minutes to about 210 minutes, for about 150 minutes to about 210 minutes, for about 180 minutes to about 210 minutes, for about 60 minutes to about 90 minutes, for about 90 minutes to about 120 minutes, for about 120 minutes to about 150 minutes, for about 150 minutes to about 180 minutes, for about 180 minutes to about 210 minutes, for about 210 minutes to about 240 minutes, or for about 240 minutes to about 270 minutes) to allow the restriction endonuclease to cleave the infectious nucleic acid. In some cases, a solution that includes a restriction endonuclease and infectious nucleic acid to be cleaved can be incubated at about 37° C. to about 60° C. (e.g., at about 37° C., at about 42° C., at about 45° C., at about 50° C., or at about 55° C.) to allow the restriction endonuclease to cleave the infectious nucleic acid. For example, when a restriction endonuclease is NheI, a solution that includes NheI and infectious nucleic acid to be cleaved can be incubated for about 60 minutes to about 180 minutes at about 37° C. to allow the NheI to cleave the infectious nucleic acid. In some cases, after incubating a solution containing an exogenously added restriction endonuclease and an infectious nucleic acid to be cleaved under conditions that allow the restriction endonuclease to cleave the infectious nucleic acid, the cleaved infectious nucleic acid can be isolated from the solution. Any appropriate technique can be used to isolate cleaved infectious nucleic acid from the solution. For example, ethanol precipitation or column purification can be used to isolate cleaved infectious nucleic acid from the solution.

[0055] When nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21) are administered to a mammal to treat cancer (e.g., B-cell lymphocytic leukemia), the mammal also can be administered one or more additional cancer treatments. The one or more additional cancer treatments can include any appropriate cancer treatment(s). In some cases, a cancer treatment can include surgery. In some cases, a cancer treatment can include radiation therapy. In some cases, a cancer treatment can include administration of one or more anti-cancer agents such as a chemotherapeutics, checkpoint inhibitors (e.g., CTLA-4 inhibitors, PD-1 inhibitors, and PD-L1 inhibitors), oncolytic viruses, gene therapies, histone deacetylase (HDAC) inhibitors, antimicrobials, immunotherapies, vaccines, protein kinase inhibitors, second mitochondrial-derived activator of caspases (SMAC) mimetics, and/or holistic therapies. An anti-cancer agent can be any appropriate type of molecule (e.g., a polypeptide such as an antibody or a small molecule). Examples of anti-cancer agents include, without limitation, ipilimumab, nivolumab,

pembrolizumab, atezolizumab, avelumab, durvalumab, cemiplimab, spartalizuma, talimogene laherparepvec (T-vec), trichostatin A (TSA), panobinostat, entinostat, romidepsin, vorinostat, givinostat, adavosertib, afatinib, axitinib, bosutinib, cetuximab, cobimetinib, crizotinib, cabozantinib, dasatinib, entrectinib, erdafitinib, erlotinib, fostamatinib, gefitinib, ibrutinib, imatinib, lapatinib, lenvatinib, mubritinib, nilotinib, pazopanib, pegaptanib, ruxolitinib, sorafenib, sumitinib, vandetanib, vemurafenib, and combinations thereof. For example, a mammal having cancer (e.g., B-cell lymphocytic leukemia) can be treated by administering nucleic acid encoding a picornavirus such as a coxsackievirus A21 and administering one or more checkpoint inhibitors (e.g., CTLA-4 inhibitors, PD-1 inhibitors, and/or PD-L1 inhibitors) to the mammal.

[0056] In cases where a mammal having cancer is treated with nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21), and is treated with one or more cancer treatments (e.g., is administered one or more anti-cancer agents), the cancer treatment(s) can be administered at the same time or independently. For example, the nucleic acid encoding a virus can be administered first, and the one or more cancer treatments administered second, or vice versa.

[0057] Also provided herein are immunocompetent models that can be infected by nucleic acid (e.g., infectious nucleic acid) encoding a virus (e.g., a picornavirus such as a coxsackievirus A21). For example, cells expressing a virus receptor (e.g., ICAM-1) can be infected by infectious nucleic acid encoding a virus (e.g., a picornavirus such as a coxsackievirus A21). When a virus receptor is ICAM-1, the ICAM-1 can be from any source. For example, an ICAM-1 can be a human ICAM-1. In some cases, an immunocompetent model does not endogenously express a virus receptor. An immunocompetent model expressing a virus receptor can stably express the virus receptor or can transiently express the virus receptor. In some cases, an immunocompetent model described herein can be used to replicate the virus encoded by nucleic acid encoding a virus. In some cases, an immunocompetent model described herein can be used as a model to evaluate and/or monitor the specificity of infection (e.g., following virus production).

[0058] In some cases, an immunocompetent model that can be infected by nucleic acid (e.g., infectious nucleic acid) encoding a virus can be a cell (e.g., a cell line) that can express (e.g., are designed to express) a virus receptor (e.g., ICAM-1 such as human ICAM-1). Any appropriate cell can be used to make an immunocompetent model described herein. In some cases, a cell used to make an immunocompetent model described herein can be obtained from a mammal (e.g., can be a primary cell). In some cases, a cell used to make an immunocompetent model described herein can be obtained from a cell line (e.g., a mammalian cell line such as murine melanoma B16-F10 cells).

[0059] In some cases, an immunocompetent model that can be infected by nucleic acid (e.g., infectious nucleic acid) encoding a virus can be a non-human animal model (e.g., a mouse model) having one or more cells that can express (e.g., are designed to express) a virus receptor (e.g., ICAM-1). Any appropriate non-human animal can be used to make an immunocompetent model described herein. In some cases, a non-human animal used to make an immunocompetent model described herein can be a mammal (e.g., a mouse or a rat).

[0060] Any appropriate method can be used to make an immunocompetent model described herein. When an immunocompetent model is a cell, nucleic acid encoding a virus receptor can be introduced into a cell such that the virus receptor is expressed by the cell. For example, a lentiviral vector encoding a virus receptor can be transduced into a cell such that the virus receptor is expressed by the cell. When an immunocompetent model is a non-human animal, nucleic acid encoding a virus receptor can be introduced into one or more cells within the non-human animal such that the virus receptor is expressed by one or more cells within the non-human animal. The invention will be further described in the following examples, which do not limit the scope of the invention described in the claims.

## EXAMPLES

### Example 1—Generating Infectious Nucleic Acid that can be Used to Treat Cancer

[0061] Enteroviruses and rhinoviruses use a type I internal ribosome entry site (IRES) for regulating viral translation. The structure of the poliovirus IRES was predicted and validated via digestion/chemical probing and mutagenic analysis (Rivera et al., *Virology*, 165:42-50 (1988); Pilipenko et al., *Virology*, 168:201-209 (1989); Skinner et al., *J. Mol. Biol.*, 207:379-392 (1989); and Burrill et al., *J. Virol.*, 87:11670-11683 (2013)). CVA21 is a polio-like virus and shares relatively 86% homology with the poliovirus IRES. The poliovirus 5' UTR contains six domains (I to VI). Domain I forms a cloverleaf structure that is required for both positive and minus-strand synthesis (Andino et al., *Cell*, 63:369-380 (1990); Andino et al., *EMBO J.*, 12:3587-3598 (1993); Barton et al., *EMBO J.*, 20:1439-1448 (2001); and Vogt et al., *PLoS Pathog.*, 6:e1000936 (2010)). Domains II thru VI are involved in IRES activity. Viral replication and translation require complex RNA-RNA and RNA-protein interactions including recruitment of host IRES-trans acting factors (ITAFs) and initiation factors (Lee et al., *Trends Microbiol.*, 25:546-561 (2017)). Disruption of these domains can result in attenuation or lethal phenotypes for the virus. Similar to poliovirus, CVA21 has a cryptic AUG site in domain VI involved in ribosome loading (Verma et al., *J. Gen. Virol.*, 92:2310-2319 (2011)). Initiation of translation at the primary AUG site depends upon ribosome scanning through a long variable linker referred to as the “scanning region.”

[0062] Infectious nucleic acid was designed to include a microRNA response element (RE) positioned within this scanning region. The RE was designed to lack AUG sites that may initiate translation prior to the authentic AUG.

[0063] Lethal myositis was the primary toxicity that needed to be eliminated to elevate the safety profile of CVA21 to a level sufficient for treatment in immunocompromised patients. The REs provided herein were directed towards eliminating viral replication within muscle tissues. miR-133 and miR-206 were selected. They are highly enriched within muscle tissues and were previously shown to eliminate toxicity when inserted into the 3' UTR of CVA21 as described elsewhere (Kelly et al., *Nat. Med.*, 14:1278-1283 (2008)).

[0064] A single microRNA-target is sufficient to down-regulate viral replication, however, there are a significant number of variables that impact the efficiency of targeting. Additionally, RNA viruses can accumulate mutations rap-

idly due to the low fidelity of the RdRp, which can result in loss of RE functionality shortly after infection. Therefore, REs were designed to encode either one copy each (miRT-1x) or two copies each (miRT-2x) of sequences complementary to miR-133 and miR-206. The sequences for the REs are shown in FIG. 1A. Ascl sites were inserted into pGEM-CVA21 at the desired locations for REs either by overlap-extension PCR, site-directed mutagenesis or by synthesizing fragments of the genome followed by subcloning into the full-length construct (pGEM-CVA21). All REs were inserted into the viral genome by annealing oligonucleotide ultramers encoding the RE flanked by the overhang sequences generated during an Ascl enzymatic digestion followed by ligation into the appropriately digested and purified Ascl-full-length vectors.

[0065] Three different constructs were generated encoding REs in the 5' UTR. The first involved inserting a 2x-RE into the scanning region at nucleotide position 686 (CVA21-686 (2x)). In addition, two constructs were generated where residues 631 thru 698 within the variable scanning region were deleted and the REs were added. The first contained a miRT-1x RE (CVA21-ΔV(1x)), and the other contained a miRT-2x RE (CVA21-ΔV(2x)). FIG. 1B depicts the CVA21 viral genome, the structural elements of the UTRs, and the RE insertion sites tested.

[0066] The majority of previously tested cellular microRNA target sequences are located within the 3' UTR of the targeted mRNAs. Based on experimentally verified modeling of the poliovirus 3' UTR (Pilipenko et al., *EMBO J.*, 15:5428-5436 (1996)), the secondary structure of the 3' UTR of CVA21 (a polio-like virus) was thought to contain two hairpins (Y and X) whose loops form a pseudoknot interaction known as a "kissing domain" (van Ooij et al., *J. Gen. Virol.*, 87:689-695 (2006)). These structures included the origin of replication (oriR) for minus-strand synthesis. Destabilization of the kissing domain in other enteroviruses severely inhibited viral RNA synthesis, resulting in temperature sensitive mutants or lethal phenotypes (van Ooij et al., *J. Gen. Virol.*, 87:689-695 (2006); Melchers et al., *J. Virol.*, 71:686-696 (1997); Wang et al., *Nucleic Acids Res.*, 27:485-490 (1999); and Melchers et al., *RNA*, 6:976-987 (2000)). In poliovirus and coxsackievirus B3, serial passage of the destabilized viruses resulted in revertants wherein the pseudoknot interaction was restored (Pilipenko et al., *EMBO J.*, 15:5428-5436 (1996); and van Ooij et al., *J. Gen. Virol.*, 87:689-695 (2006)). Deletion of the two domains also resulted in truncation of the poly(A) tail (van Ooij et al., *Nucleic Acids Res.*, 34:2953-2965 (2006)). Viruses were recovered from these deletion mutants, however, they acquired poly (AU) stretches corresponding to cellular polyadenylation signals. The lengths of the X and Y domains were highly conserved among enteroviruses, and it was shown that domain Y should be 12 bp in combination with an 8 bp stem for domain X. Additionally, the two most distal base pairs in domain Y relative to the "kissing domain" were required to be Watson-Crick CG pairs (Melchers et al., *J. Virol.*, 71:686-696 (1997)). The length of the linker regions between the two stem loops was important for the correct orientation of the helices and their ability to interact. Finally, the existence of an S domain was also suggested wherein the poly(A) tail base pairs with a tract of four uridine residues upstream of the stop codon in the 3D gene.

[0067] This domain closed off the oriR structure, creating a more rigid 3' UTR structure (Pilipenko et al., *Nucleic Acids*

*Res.*, 20:1739-1745 (1992)). Having a properly folded 3' UTR was therefore involved to ensure correct docking and orientation of the proteins associated with the ribonucleoprotein complex involved in negative-sense strand synthesis.

[0068] The left panel in FIG. 1C shows the secondary structure model of the CVA21 oriR and the location of the RE used in the Kelly et al. construct (CVA21-3'miRT) at nucleotide position 7343. Insertion at this site disrupted the length of domain Y, interfered with the distal CG base pairs, and likely disrupted the linker region between domains X and Y, interfering with the orientation of the helices. All residues in the 3' UTR and several in the coding sequence were involved in maintaining the oriR structure, limiting potential insertion sites for REs. To bypass any structural disruptions, an Ascl restriction site followed by a terminal repeat (TR) element was inserted directly downstream of the stop codon. This TR element included nucleotides 7325 to 7340, containing the four uridine residues thought to be involved in domain S. It was hypothesized that the TR would separate the oriR structure from the RE inserted into the Ascl site while maintaining the coding sequence (FIG. 1C, right). Both miRT-1x (CVA21-3'TR(1x)) and miRT-2x (CVA21-3'TR(2x)) containing constructs were generated.

[0069] MicroRNA target insert integrity was verified in all constructs by sequencing the insert region.

#### Virus Rescue Kinetics from RNA Transcripts

[0070] One objective was to obtain infectious nucleic acid that can nucleate a spreading virus infection from targeted-infectious nucleic acid at a rate similar to the unmodified virus. To analyze the ability of the targeted infectious RNAs to produce virus progeny H1-HeLa cells (ATCC, Manassas, VA) that do not express miR-133 or miR-206 were transfected with infectious RNA encoding each miRT-CVA21. Infectious RNA was prepared by linearizing plasmid DNA encoding unmodified or miRT-CVA21 with restriction endonuclease MluI-HF (New England Biolabs, Ipswich, MA), followed by ethanol precipitation and resuspension in nucleic-acid-free water. 1 µg of linearized DNA was transcribed into RNA transcripts using the Ambion MEGAscript T7 transcription kit, and the RNA was purified using the MEGAclear transcription clean-up kit (Thermo Fisher Scientific Inc., Waltham, MA); both according to the manufacturers' instructions. Transcript size and integrity were verified by running the RNA on an RNA Flash gel (Lonza, Basel, Switzerland). 4×10<sup>5</sup> cells were seeded per well in 6-well tissue culture plates, 24 hours prior to transfection. Each well was transfected with 2.5 µg of purified T7 RNA using TransIT-mRNA transfection kit (Mirus Bio LLC, Madison, WI). Once cytopathic effects (CPE) were observed, the cells were scraped into the supernatant, and the samples collected into individual cryotubes. The samples were subjected to 3 freeze-thaw cycles, cleared by centrifugation at 2500 rpm for 5 minutes at 4°C, and filtered through a 0.22 µm syringe filter. 50 µL of cleared lysate was used to infect fresh H1-HeLa cells in a well of a 6-well tissue culture plate. Cells were infected for 2 hours at 37°C in serum-free media. Media and unincorporated virus were removed, and 2 mL of fresh complete growth media were added per well. At 24 hours post infection, the cells were observed for CPE. All miRT-CVA21 except the Kelly et al. construct, CVA21-3'miRT, generated sufficient virus progeny to produce visible CPE (FIG. 2A).

[0071] To evaluate the rate at which virus progeny were generated from RNA transcripts encoding these targeted

viral genomes, growth curve time courses were conducted in H1-HeLa cells.  $2.5 \times 10^5$  H1-HeLa cells were seeded per well into 12-well tissue culture dishes, 24 hours prior to transfection. Each well was transfected with 1  $\mu$ g of purified T7 RNA using TransIT-mRNA transfection kit (Mirus Bio LLC, Madison, WI). At 6 hours post transfection, the media was removed, and cells were washed once with complete media. 1 mL of complete growth media was added per well, and the cells were incubated at 37° C. until the desired time point. At 6, 12, 24, and 48 hours post transfection, cells were scraped into the supernatant, and the samples were collected into individual cryotubes and stored at -80° C. until all samples were collected. Samples were subjected to 3 freeze-thaw cycles and cleared by centrifugation at 2500 rpm for 5 minutes at 4° C. Cleared lysates were titrated on H1-HeLa cells, and the TCID<sub>50</sub> per mL of each sample was determined using the Spearman-Kärber equation. Virus production from the Kelly et al. construct, CVA21-3'miRT, was severely impaired. The rate of rescue from CVA21-TR(2x) RNA was slightly delayed compared to unmodified CVA21 RNA. All other miRT-CVA21 RNAs rescued virus at rates and levels similar to the unmodified RNA genome (FIG. 2B). Similar results were obtained when the time course was conducted in the human melanoma cell line Mel624 (Imanis Life Sciences, Rochester, MN) (FIG. 2C).

#### Characterization of miRT-CVA21 Viruses

[0072] In order to generate virus stocks, viral RNA were produced as described above. 2.5  $\mu$ g RNA per well was transfected into H1-HeLa cells seeded in 6-well plates as described above. At 48-72 hours post transfection, the cells were scraped into the supernatant, and the samples were collected. The samples were subjected to 3 freeze-thaw cycles, cleared by centrifugation at 2500 rpm for 5 minutes at 4° C., and filtered through a 0.22  $\mu$ m syringe filter. H1-HeLa cells in a T75 flask were infected with the cleared lysates at 37° C. in serum-free media. Two hours post infection, the media and unincorporated virus were removed, and the cells were replenished with complete growth media. Once CPE was observed, the samples were processed in the same manner. The cleared lysate virus stocks were aliquoted and stored at -80° C. Viruses were titrated on H1-HeLa cells.  $1 \times 10^4$  cells per well were plated in 96-well tissue-culture plates, 24 hours prior to infection. Ten-fold serial dilutions ( $1 \times 10^{-2}$  to  $1 \times 10^{-10}$ ) of the virus were made in serum-free media, and 100  $\mu$ L of each dilution was added to each of 8 duplicate wells. The cells were infected for 2 hours at 37° C., and then the media was removed and replaced with 100  $\mu$ L complete growth media. The cells were incubated at 37° C. for 72 hours, and then the wells were visually inspected for CPE. The TCID<sub>50</sub> per mL of each sample was determined using the Spearman-Kärber equation.

[0073] One-step growth curves were performed to compare the replication kinetics of the miRT-CVA21 to the unmodified virus. H1-HeLa cells were infected with unmodified or miRT-CVA21 at a multiplicity of infection of 3.0 in serum-free media. Two hours post infection, the cells were washed, and complete growth media was added. Samples were collected at specific times post transfection (2, 4, 6, 8, 12, 24, and 48 hours) and stored at -80° C. Following the completion of all timepoints, samples were frozen and thawed 3 times, and cellular debris was cleared from the lysates by centrifugation at 2500 rpm for 5 minutes at 4° C. The cleared lysates were then titrated, and virus titers were

determined using the Spearman-Kärber equation. All miRT-CVA21 replicated with kinetics similar to the unmodified virus.

[0074] The efficiency and specificity of microRNA-targeting were analyzed by measuring cell viability and viral replication in H1-HeLa cells transfected with complementary or noncomplementary synthetic miRNA mimics (Dharmacon, Lafayette, CO). Mimics were reverse transfected into H1-HeLa cells using the TransIT-mRNA transfection kit according to the manufacturer's protocol at a concentration of 100 nM each. Briefly, transfection complexes were assembled in a 96-well plate and incubated at room temperature for 5 minutes. H1-HeLa cells in T75 flasks were trypsinized, counted, and resuspended in complete growth media at a concentration of  $1 \times 10^4$  cells per 90  $\mu$ L. 90  $\mu$ L of cells was added per well, and the cells/transfection mixtures were incubated at 37° C. for 12 hours. The cells were infected at an MOI of 1 with each miRT-CVA21 for 2 hours at 37° C. in serum-free media. Following infection, the media and unincorporated virus was removed and replaced with 100  $\mu$ L complete growth media, and the cells were incubated at 37° C. 24 hours post infection, the supernatants were collected and titrated as described above. The cells were assayed for proliferation using a 3-(4,5-dimethylthiazolyl-2)-2,5-Diphenyltetrazolium bromide (MT) kit (ATCC, Manassas, VA). CVA21-3'TR(2x) replication was not regulated by miR-133, miR-206, or a combination of the two mimics. Virus replication was minimally controlled by miR-133 for CVA21-686(2x) and CVA21-ΔV(2x). MicroRNA-206 was much more efficient at controlling virus replication resulting in increased cell viability and decreased virus titers for CVA21-ΔV(1x), CVA21-ΔV(2x), CVA21-686(2x), and CVA21-3'TR(1x). Virus tropism was similarly regulated in H1-HeLa cells transfected with both miR-133 and miR-206. No difference in cell viability and virus titer was observed in H1-HeLa cells transfected with the miR-142 control mimic or non-transfected cells. Of note, viruses with 5' UTR localized REs were more readily controlled than CVA21-3'TR(1x). Based on these results, the analyses did not continue with the CVA21-3'TR(2x) construct.

#### In Vitro Genetic Stability of Response Elements

[0075] The genetic stability of RE at variable locations was evaluated by force passaging miRT-CVA21 in TE671 muscle cells that express miR-133 and miR-206. TE-671 cells were cultured in 2% horse serum for 4 days, which induces the cells to differentiate into myotubes expressing higher levels of miR-133 and miR-206. Differentiated TE-671 cells (dTE-671) in 6-well tissue culture plates were infected at an MOI of 10 with CVA21-ΔV(1x), CVA21-ΔV(2x), CVA21-686(2x), CVA21-3'TR(1x), or CVA21-3'TR(2x) for 2 hours at 37° C. in serum-free media. After 2 hours, the media and unincorporated virus were removed, and the cells were washed with complete media. 1.5 mL of complete media was added per well, and the cells were incubated at 37° C. At 24 hours post infection, the cells were scraped into the supernatant, and the samples were collected. All samples were subjected to 3 freeze-thaw cycles, and the lysates were clarified by centrifugation at 2500 rpm for 5 minutes at 4° C. and filtered through a 0.22  $\mu$ m filter. Virus in clarified lysates was passaged serially in dTE-671 cells seven times, each time using 1 volume of clarified lysate to 2 volumes of fresh media. Viral RNA was isolated from the cleared lysates of each passage with a QIAamp viral RNA mini kit (Qiagen,

CA) according to the manufacturer's instructions. cDNA was synthesized, and regions containing the REs were amplified. Amplicons were sequenced with nested primers. This assay was conducted in duplicate. Escape mutants were observed in CVA21-686(2x) samples as early as passage 2. All other miRT-CVA21 infections gave rise to escape mutants between passages 4 and 7.

#### In Vivo Analysis of Oncolytic Activity

[0076] Based on the in vitro results, the CVA21-3' TR(2x) construct was not assessed in vivo. 4-5 week old female CB17 ICR-SCID mice were purchased from Envigo (Huntingdon, Cambridgeshire, UK). The mice were irradiated and 24 hours later implanted subcutaneously with 5e6 Mel624 cells in the right flank. When the tumors reached an average of 0.5 cm×0.5 cm, the tumors were treated with 30 µg of RNA in 50 µL of saline intratumorally. Tumor volume, measured using a hand-held caliper, weights, and overall health were routinely monitored. Blood was collected from all mice on day 7 post RNA treatment. Mice were anesthetized through the inhalation of isoflurane, and blood was collected from the submandibular vein in a BD microtainer tube with a sera separator gel. Blood was allowed to coagulate for 30 minutes at room temperature, and then sera was separated by centrifugation at 8000 rpm for 5 minutes at 4° C. Sera was stored at -80° C. Virus in sera was titrated as described for virus stocks on H1-HeLa cells. Sera (bled via cardiac puncture) and skeletal muscle tissue were obtained from all mice at the time of euthanasia. Skeletal muscle sections were immediately flash frozen and stored at -80° C. or were fixed in 10% formalin. Total RNA was isolated from flash frozen tissue sections with an RNeasy Plus Universal mini kit (Qiagen, CA) according to the manufacturer's instructions. Viral RNA was isolated from sera using a QIAamp viral RNA mini kit (Qiagen, CA) according to the manufacturer's instructions. cDNA was synthesized, and regions containing the REs were amplified using the Titan One-Tube RT-PCR system (Sigma Aldrich, St. Louis, MO) according to the manufacturer's instructions. Amplicons were sequenced with nested primers.

[0077] Tumor volumes and weights of all mice throughout the duration of the experiment are shown in FIG. 4A. Control treated mice and mice administered CVA21-3'miRT RNA all exhibited progressive tumor growth and were euthanized due to tumor volume exceeding 10% body weight or tumor ulceration. One mouse treated with CVA21-3'miRT was found in a moribund state and was immediately euthanized. Sequence analysis of viral genomes in skeletal muscle tissue from this mouse revealed wild-type reversions. Rapid tumor regression was observed in all mice treated with CVA21, CVA21-ΔV(1x), CVA21-ΔV(2x), CVA21-686(2x), or CVA21-3'TR(1x) RNA. Toxicity in the form of hind-limb paralysis (HLP), sudden death (FD), or excessive weight loss (WL) was observed in all mice treated with CVA21 RNA and a proportion of mice treated with CVA21-ΔV(1x), CVA21-686(2x), or CVA21-3'TR(1x) RNA at 60%, 60%, and 75%, respectively. No toxicity was observed in mice treated with CVA21-ΔV(2x). Toxicities observed and proportions per group are shown in FIG. 4B. All mice treated with CVA21-ΔV(2x) appeared healthy and tumor-free at the end of study day 90. Viral genomes isolated from sera and skeletal muscle from all CVA21-ΔV(2x) treated mice maintained the RE without mutations. Overall

survival for mice treated with CVA21-ΔV(2x) was 100%, significantly improving survival over control treated mice, p=0.002 (FIG. 4C).

[0078] As shown in FIG. 4D, no infectious virus was recovered from sera isolated on day 7 post treatment from mice treated with CVA21-3'miRT RNA. In contrast, viral titers observed in sera from mice treated with CVA21-ΔV (1x), CVA21-ΔV(2x), CVA21-686(2x), or CVA21-3'TR(1x) RNA were at levels similar to those found in mice treated with CVA21 RNA.

#### In Vivo Dose Escalation of CVA21-ΔV(2x) RNA

[0079] 4-5 week old female CB17 ICR-SCID mice from Envigo (Huntingdon, Cambridgeshire, UK) were irradiated and 24 hours later implanted subcutaneously with 5e6 Mel624 cells in the right flank. When the tumors reached an average of 0.5 cm×0.5 cm, the tumors were treated with 1-32 µg of CVA21-ΔV(2x) RNA in 50 µL of saline intratumorally. Tumor volume, measured using a hand-held caliper, weights, and overall health were routinely monitored. Blood was collected from all mice on day 9 post RNA treatment. Tumor size, weight, blood, sera isolation, tissue collection, and genome sequencing were all measured, obtained, processed and analyzed as described for the previous in vivo experiment.

[0080] As shown in FIG. 5A, control treated mice exhibited progressive tumor growth, and all were euthanized due to tumor size or ulceration. 4 of 5 mice treated with 1 µg CVA21-ΔV(2x) RNA also exhibited progressive tumor growth, however, complete tumor regression was observed in one mouse. Complete tumor regression was observed in all other mice treated with CVA21-ΔV(2x) RNA at 4 to 32 µg. Hind-limb paralysis was observed in a single mouse treated with CVA21-ΔV(2x) RNA at 8 µg. Sequence analysis and histological analysis of skeletal muscle did not reveal any mutations in the response element sequence or signs of myositis. Another mouse treated with 32 µg of CVA21-ΔV (2x) RNA was found dead at day 78 post RNA treatment. The majority of mice treated with 4 to 32 µg of CVA21-ΔV (2x) RNA displayed high viral loads in sera on day 9 post RNA treatment, however, a few did not (FIG. 5B). Time course analysis of viral loads in sera following RNA treatment can be used to establish when peak viral loads will be observed. Viremia was only observed in one mouse from the 1 µg group, and this mouse displayed complete tumor regression.

#### Bilateral Tumor Destruction Following CVA21-ΔV(2x) RNA Therapy

[0081] 4-5 week old female CB17 ICR-SCID mice are irradiated and 24 hours later are implanted subcutaneously with 5e6 Mel624 cells in the right flank and 5e6 Mel624 cells in the left flank. When tumors reach an average of 0.5 cm×0.5 cm, the mice are treated. Each mouse is given a single injection of 2, 10, or 30 µg of CVA21-ΔV(2x) RNA in the right flank tumor. Tumor volume, which is measured using a hand-held caliper, weights, and overall health are routinely monitored. Blood is collected from two mice per group on days 2, 4, 6, 8 or 10 post treatment. Tumor size, weight, blood, sera isolation, tissue collection and genome sequencing is measured, is obtained, is processed and analyzed as described for the previous in vivo experiments.

### Potential Infectious Nucleic Acid Formulations for CVA21-ΔV(2x) Therapy

**[0082]** Examples of infectious nucleic acid formulations include, but are not limited to, infectious cDNA clones or RNA transcripts encoding picornavirus genomes. Techniques used to synthesize viral RNA genomes from infectious cDNA can result in additional nucleotides (not part of the viral genome) on the 5' and/or 3' ends. Several positive and negative RNA viruses, including poliovirus, have been shown to require exact termini for efficient replication (Boyer et al., *Virology*, 198:415-426 (1994); Herold and Andino, *J. Virol.*, 74(14):6394-6400 (2000)). Although these residues can be removed during replication generating the authentic viral genomes, their initial presence can have detrimental effects on the specific infectivity of the therapeutic nucleic acid. CVA21-ΔV(2x) infectious nucleic acid therapy may be improved by employing mechanisms to rapidly generate authentic viral genomes. One such mechanism is to encode ribozyme sequences at the 5' and/or 3' termini. Ribozymes are RNA molecules (structures) with catalytic properties. Certain ribozymes are capable of cleaving RNA molecules in cis or in trans at very specific positions. Encoding ribozymes at the 5' and/or 3' end of the viral genome in infectious nucleic acid formulations encoding CVA21-ΔV(2x) may improve the therapeutic efficacy even further.

**[0083]** Three different constructs are made to confirm this. The first includes a ribozyme immediately upstream of the CVA21-ΔV(2x) genome (Rz-CVA21-ΔV(2x)). This ribozyme is modified to allow cleavage in cis at the exact 5' termini of the CVA21-ΔV(2x) genome. The second construct includes a different ribozyme immediately downstream of the CVA21-ΔV(2x) genome (CVA21-ΔV(2x)-Rz) such that it cleaves the RNA in cis directly downstream of the encoded poly A tail of CVA21-ΔV(2x). The third construct includes the CVA21-ΔV(2x) genome flanked by both of these ribozymes (Rz-CVA21-ΔV(2x)-Rz). RNA genomes are synthesized, and their specific infectivity, targeting efficacy/specificity, genetic stability, and therapeutic efficacy/safety are characterized as described above for the unmodified CVA21-ΔV(2x). A mechanism to ensure authentic or near authentic 3' termini is to include sequences encoding a restriction site that can be cleaved by a restriction endonuclease directly adjacent or within 5 nucleotide residues following the encoded 3' end (e.g. poly A tail). An NheI restriction enzyme site was encoding within the DNA encoding the viral genome CVA21-ΔV(2x) directly adjacent to the encoded poly A tail. The DNA was linearized with the NheI restriction endonuclease such that the in-vitro derived RNA transcripts encoding the CVA21-ΔV(2x) contained 5 non-viral nucleotides following the poly A tail (CVA21-ΔV(2x) 3'NheI). Another construct was made with the NheI site at the 3' end of the CVA21-ΔV(2x) genome in conjunction with a ribozyme immediately upstream of the CVA21-ΔV(2x) genome (Rz-CVA21-ΔV(2x)-3'NheI). RNA genomes are synthesized, and their specific infectivity, targeting efficacy/specificity, genetic stability, and therapeutic efficacy/safety are characterized as described above for the unmodified CVA21-ΔV(2x).

### Therapeutic Efficacy in a Variety of Tumor Types

**[0084]** CVA21-ΔV(2x) infectious nucleic acid therapy can be used to treat a variety of cancer types including, but not

limited to, melanoma, myeloma, prostate cancer, breast cancer, lung cancer (e.g., non-small cell lung cancer), and pancreatic cancer.  $1 \times 10^4$  cells per well of representative tumor cell lines for these cancer types were plated in 96-well tissue culture dishes. The cells were infected at an increasing MOI between 0.001 and 1 with CVA21 or CVA21-ΔV2 for 2 hours at 37° C. in serum-free media. Following infection, the media and unincorporated virus was removed and replaced with 100  $\mu$ L complete growth media, and the cells were incubated at 37° C. 72 hours post infection, the cells were assayed for proliferation using a 3-(4,5-dimethylthiazolyl-2)-2,5-Diphenyltetrazolium bromide (MTT) kit (ATCC, Manassas, VA). All cell lines tested were as susceptible to CVA21-ΔV2 as they were to CVA21 (FIG. 6).

**[0085]** This tumor cell panel is analyzed for susceptibility to infectious nucleic acid formulations encoding CVA21-ΔV2.  $6 \times 10^4$  cells per well are plated in 24-well tissue culture dishes. RNA transcripts encoding CVA21, CVA21-ΔV2, Rz-CVA21-ΔV2, CVA21-ΔV2-Rz, or Rz-CVA21-ΔV2-Rz are transfected into the cells. Each well is transfected with 0.5  $\mu$ g of purified T7 RNA using TransIT-mRNA transfection kit (Mirus Bio LLC, Madison, WI). No template transfection controls are used to account for changes in cell viability associated with the transfection protocol. At 24, 48, and 72 hours post transfection, cells are assayed for proliferation using a 3-(4,5-dimethylthiazolyl-2)-2,5-Diphenyltetrazolium bromide (MTT) kit (ATCC, Manassas, VA).

### Example 2—Infectivity and Stability of Nucleic Acid that can be Used to Treat Cancer

#### Insertion of a 2xRE in Domain Y Reduces Specific Infectivity and Eliminates Therapeutic Efficacy of Infectious RNA Encoding the CVA21-3'miRT genome.

**[0086]** Recovery of infectious virus following transfection of 1  $\mu$ g of in vitro-derived infectious RNA encoding CVA21-3'miRT was severely delayed in both H1-HeLa and Mel624 cells compared to RNA encoding unmodified CVA21 (FIG. 2). 4-5 week old female CB17 ICR-SCID mice were purchased from Envigo (Huntingdon, Cambridgeshire, UK). The mice were irradiated and 24 hours later implanted subcutaneously with 5e6 Mel624 cells in the right flank. When the tumors reached an average of 0.5 cm $\times$ 0.5 cm, the tumors were treated with 30  $\mu$ g of RNA encoding either unmodified CVA21 or CVA21-3'miRT in 50  $\mu$ L of saline intratumorally. Tumor volume, measured using a hand-held caliper, weights, and overall health were routinely monitored. Blood was collected from all mice on day 7 post RNA treatment. Mice were anesthetized through the inhalation of isoflurane, and blood was collected from the submandibular vein in a BD microtainer tube with a sera separator gel. Blood was allowed to coagulate for 30 minutes at room temperature, and then sera was separated by centrifugation at 8000 rpm for 5 minutes at 4° C. Sera was stored at -80° C. Virus in sera was titrated as described for virus stocks on H1-HeLa cells. As shown in FIG. 4 A and D, RNA encoding CVA21-3'miRT did not exhibit any oncolytic activity or induce viremia. RNA secondary structural prediction demonstrates the disruption of the Y domain of the oriR and reduced probability of the pseudoknot formation that likely contributes to the reduced specific infectivity of the in vitro-derived RNA and lack of therapeutic efficacy (FIG. 13).

### Scanning Region Replacement Enhances microRNA Response Element Stability.

[0087] At the time of euthanasia, viral genomes were isolated from the sera and skeletal muscle of mice treated with CVA21-ΔV(2x) and mice with clinical signs of toxicity. Viral or total RNA was isolated from the samples, respectively, and cDNA was synthesized. Regions containing the microRNA response elements were amplified and the amplicons sequenced. Reversion mutants were detected in the skeletal muscle of one mouse treated with CVA21-ΔV(1x) that developed hind-limb paralysis and in all three mice treated with CVA21-686(2x) that developed hind-limb paralysis. In contrast, no reversion or escape mutants were detected in any of the mice treated with CVA21-ΔV(2x). Of note, reversion mutants were also detected in both of the evaluable mice treated with CVA21-TR(1x). In an effort to determine the stability of the microRNA response elements *in vitro*, serial passage was performed of each microRNA-detargeted CVA21 in differentiated TE671 (dTE-671) muscle cells that express miR-133 and miR-206. dTE671 cells were initially infected at an MOI of 10. At twenty-four hours post infection, the samples were collected and fresh dTE671 cells infected with 50% of the clarified lysates. Viral RNA was isolated from cleared lysates of seven serial passages, cDNA synthesized and regions containing the response elements amplified for sequencing. In both experiments, reversion mutants were detected in CVA21-686(2x) samples 2-3 passages prior to detection in CVA21-ΔV(2x) and CVA21-ΔV(1x) samples. This data indicates that elongation of the scanning region with direct microRNA response element insertion decreases the genetic stability, an effect that is more pronounced *in vivo*.

Replacement of the Ribosomal Scanning Region with microRNA Response Element Minimizes Potential for Structural Alterations.

[0088] In contrast to the complex structural environment of the 3' NCR, the 5' NCR of CVA21 is predicted to contain a disordered scanning region directly downstream of the IRES. The role of this domain in CVA21 replication is unknown. Although this region has been shown to be dispensable for poliovirus replication *in vitro* and *in vivo*, it has been indicated in binding an IRES trans-acting factor that enhances enterovirus 71 translation. RNA secondary structural analysis predicted that the CVA21-ΔV(2x) configuration resulted in maintenance of domain VI within the IRES and therefore should not significantly impact ribosomal loading and translation (FIG. 13). No pseudoknot formations are impacted by insertion within the 5' NCR as is the case with insertion anywhere within the 3' NCR. As shown in FIG. 13D-F, although the TR(2x) construct separates the microRNA response element from the oriR loops (domains Y and X), the predicted pseudoknot interaction between domain Y and domain X is still lost as depicted by the interloop connecting lines.

### Potential Infectious Nucleic Acid Formulations for CVA21-ΔV(2x) Therapy

[0089] CVA21-ΔV(2x) infectious nucleic acid therapy may be improved by employing mechanisms to rapidly generate authentic viral genomes. One such mechanism is to encode ribozyme sequences at the 5' and 3' termini. Ribozymes are RNA molecules (structures) with catalytic properties. Certain ribozymes are capable of cleaving RNA molecules in *cis* or in *trans* at very specific positions.

Encoding ribozymes at the 5' and/or 3' end of the viral genome in infectious nucleic acid formulations encoding CVA21-ΔV(2x) may improve the therapeutic efficacy even further.

[0090] Five different constructs were made to test this hypothesis. The first included a ribozyme immediately upstream of the CVA21-ΔV(2x) genome (Rz-CVA21-ΔV(2x)). This ribozyme was modified to allow cleavage in *cis* at the exact 5' termini of the CVA21-ΔV(2x) genome. The second construct included a different ribozyme immediately downstream of the CVA21-ΔV(2x) genome (CVA21-ΔV(2x)-Rz) such that it will cleave the RNA in *cis* directly downstream of the encoded poly A tail of CVA21-ΔV(2x). The third construct included an NheI restriction enzyme cut site directly adjacent to the poly A tail of the viral genome encoded in the plasmid DNA. The NheI cut site was used to generate linear transcripts such that only a few residues will follow the poly A tail. The fourth construct included a ribozyme at the 5' termini of the CVA21-ΔV(2x) genome and the NheI cut site following the poly A tail (Rz-CVA21-ΔV(2x)-3'NheI). The fifth construct included the CVA21-ΔV(2x) genome flanked by both of these ribozymes (Rz-CVA21-ΔV(2x)-3'Rz). RNA genomes were synthesized and their specific infectivity, targeting efficacy/specification, genetic stability, and therapeutic efficacy/safety characterized as described above for the unmodified CVA21-ΔV(2x). As shown in FIG. 14, both constructs with dual authentic termini or near-authentic (i.e., Rz-CVA21-ΔV(2x)-3'Rz and Rz-CVA21-ΔV(2x)-3'NheI, respectively) had increased specific infectivity and cytopathic effects observed sooner following transfection of H1-HeLa cells compared to the unmodified CVA21-ΔV(2x) RNA. Both of these constructs exhibit oncolytic activity against subcutaneous Mel624 xenograft tumors in CB-17 SCID mice (FIG. 15A). Both constructs were also able to generate spreading oncolytic infections as exhibited by the development of viremia in the treated mice on days 2 and 7 post therapy (FIG. 15B). Rz-CVA21-ΔV(2x)-3'NheI was more efficient than CVA21-ΔV(2x) and Rz-CVA21-ΔV(2x)-3'Rz and the development of viremia in mice treated with Rz-CVA21-ΔV(2x)-3'NheI was more consistent. Different combinations of ribozyme-encoding sequences and restriction enzyme cut sites (with exogenously added restriction enzymes designed to cut at those cut sites) may be used to modulate the specific infectivity and therapeutic potential of the infectious RNA formulation. This may include but is not limited to the immunogenic potential of the construct in various tumor microenvironments and delivery routes.

Initial Seeding of Heterogenous Tumor Cells is Independent of Virus Receptor Expression, but Spread and Safety is Still Dependent on the Expression of hICAM-1 and/or Decay-Accelerating Factor.

[0091] Murine melanoma B16-F10 cells are not susceptible to CVA21 because they do not express the virus receptors. A B16-F10 cell line stably expressing hICAM-1 was generated via lentiviral transduction, antibiotic selection, single-cell sorting and expansion of a clonal population. Virus recovery from this cell line B16-F10-hICAM1 24 hours post infection with CVA21-ΔV(2x) at an MOI of 1 was significantly enhanced compared to the parental cell line (FIG. 16). This effect can also be applied to other cell lines that do not express the receptors for virus entry to expand the repertoire of *in vivo* models. Infectious nucleic acid can be used to expand initial seeding of tumor cells to include cells

that do not express the receptor for the virus while maintaining the specificity of spread following virus production. Cell type specificity can be further modulated by incorporation of different microRNA target sequences.

#### In Vitro Transcription Reactions can be Scaled to Generate Clinical Preps of CVA21-ΔV(2x) Infectious RNA

**[0092]** Preclinical studies utilized an in vitro RNA transcription kit to generate RNA transcripts. Each 10 microliter reaction would yield ~100 micrograms of RNA that was used in the in vivo studies. To demonstrate the feasibility of scaling a production method without commercially available kits, a similar reaction was set up and scaled to 1 mL (100 $\times$ ), and the resulting transcripts were purified using lithium chloride precipitation. This scaled reaction resulted in >7 milligrams of RNA with integrity similar to 10 microliter reactions as observed by RNA gel electrophoresis (FIG. 17).

#### Modifications to In Vitro-Derived Infectious Nucleic Acid can be Used to Enhance Stability in Different Environments.

**[0093]** Various mechanisms can be used to enhance the stability of RNAs in blood and translatability following internalization within a cell. These techniques include, but are not limited to, using modified bases (e.g., N-methylpseudouridine) in the in vitro transcription to generate RNAs that closely resemble RNAs normally present within the hosts being treated, capping mechanisms, and polyadenylation of transcripts. All of these mechanisms and others to enhance the stability, delivery, uptake, and expression of nucleic acid based therapeutics can be applied to infectious nucleic acid therapy. These mechanisms can also be used to modulate the immunogenicity of the infectious nucleic acid on a per patient basis. Additionally, the delivery of infectious nucleic acid can be enhanced by complexing the nucleic acid with a variety of different carrier molecules, including, but not limited to, lipid-based, polymer-based, nanoparticle-based, and other biosynthetic molecules.

#### Example 3: Reduced Cost and Simplification of Manufacturing

**[0094]** Manufacturing protocols vary for each oncolytic virus, however, the general outline of procedures necessary to produce and purify large quantities of the virus with sufficient titers of infectious particles is similar. A GMP master cell bank and master seed virus are generated and qualified. These cells are seeded and expanded using optimized medium formulations until a sufficient number of cells are established. These cells are often seeded into large bioreactors for further expansion and infection. The virus is harvested, generally requiring lysis and/or clarification of cellular debris followed by enzymatic digestion of residual nucleic acids and purification of the lysate. Downstream purification processes can include various ultrafiltration and diafiltration steps and chromatography (e.g. ion exchange and gel permeation) based purification. This is followed by another filtration step prior to vialing and storing. Production of RNA-based therapeutics requires fewer steps and can be scaled to produce higher yields in lower volumes (i.e., smaller bioreactors). The steps involved in producing infec-

tions RNA are fewer and simpler. A master bank of linearized plasmid DNA encoding the viral genome is made. This is used for in vitro transcription reactions (generally ~1 L per stock) to produce the infectious RNA. DNase digestion is used to remove residual plasmid DNA, and the infectious RNA is then purified via precipitation or column filtration. HPLC or FPLC purification can be used to remove additional contaminants to ensure purity. In contrast to current mRNA-based therapeutics, CVA21-ΔV(2x) transcripts do not require capping or tailing reducing the costs/steps of synthesis and purification using in vitro transcription GMP protocols. Oncolytic CVA21 viruses can be used clinically in combination with immunotherapy.

#### Example 4: Potential for Enhancing Virus Monotherapy

**[0095]** Formulating CVA21 as infectious nucleic acid has the potential to safely enhance its monotherapeutic potency. Nucleic acid is less immunogenic than virus particles providing a mechanism to avoid neutralizing antibodies during repeat dosing. Infectious nucleic acid delivery can infect cells even in the presence of neutralizing antibodies boosting the oncolytic phase of therapy during repeat injections that may enhance the overall efficacy of treatment. Furthermore, infectious nucleic acid is not restricted to tumor cells expressing the virus receptor (e.g., human intracellular adhesion molecule 1). The initial seeding of tumor cells will include those cells normally refractory to virus infection, overcoming the barrier of tumor heterogeneity. This strategy can be applied to other oncolytic picornaviruses.

#### Example 5: Improved Safety Expands Patient Eligibility

**[0096]** The tolerability of CVA21 has been demonstrated in phase I and II clinical trials in patients with several advanced malignancies. However, studies to date have been limited to patients with functional immune systems. Although rare in humans, CVA21 can cause myositis with immunodeficient hosts being particularly vulnerable. Thus, development of a CVA21 that is unable to replicate in muscle cells will allow expansion of clinical analyses to patients with compromised immune systems. CVA21-ΔV(2x) has a microRNA response element that includes sequences recognized by microRNAs enriched within muscle tissues. This element reduces viral replication in cells expressing the cognate microRNAs and ameliorates toxicity observed in immunodeficient mice bearing subcutaneous tumors treated with CVA21. MicroRNA-detargeted viruses including CVA21-ΔV(2x) can be used to improve safety of oncolytic therapies in immunocompromised patients.

#### OTHER EMBODIMENTS

**[0097]** It is to be understood that while the invention has been described in conjunction with the detailed description thereof, the foregoing description is intended to illustrate and not limit the scope of the invention, which is defined by the scope of the appended claims. Other aspects, advantages, and modifications are within the scope of the following claims.

## SEQUENCE LISTING

Sequence total quantity: 69

SEQ ID NO: 1 moltype = DNA length = 50  
FEATURE Location/Qualifiers  
source 1..50  
mol\_type = other DNA  
organism = synthetic construct

SEQUENCE: 1 acagctgggtt gaaggggacc aactggagcc acacacttcc ttacattcca 50

SEQ ID NO: 2 moltype = DNA length = 102  
FEATURE Location/Qualifiers  
source 1..102  
mol\_type = other DNA  
organism = synthetic construct

SEQUENCE: 2 acagctgggtt gaaggggacc aacgatacag ctgggtgaag gggaccaact ggagccacac 60  
acttccttac attccatcac ccacacattt ctttacattt ca 102

SEQ ID NO: 3 moltype = RNA length = 100  
FEATURE Location/Qualifiers  
source 1..100  
mol\_type = other RNA  
organism = synthetic construct

SEQUENCE: 3 cgactcattt tagtaaccctt acctcagtcg gattggattt gggtataactg ttgttaggg 60  
aaatttttctt ttaattcgga gaaaaaaaaaaa aaaaaaaaaaaa 100

SEQ ID NO: 4 moltype = RNA length = 123  
FEATURE Location/Qualifiers  
source 1..123  
mol\_type = other RNA  
organism = synthetic construct

SEQUENCE: 4 cgactcattt tagtaaggcg cgccgactca ttttagtaac cttacccatcg tcggatttga 60  
ttgggttata ctgttgttgg ggttaattttt tcttttattt ggagaaaaaaaaaaa 120  
aaa 123

SEQ ID NO: 5 moltype = DNA length = 10477  
FEATURE Location/Qualifiers  
source 1..10477  
mol\_type = other DNA  
organism = synthetic construct

SEQUENCE: 5 taaaaacacgc tctgggggtt ttcccaccc agaggccac gtggccgcta gtactctgg 60  
attacggat ctttgatcgc ctgtttgtt tcccttcccc cgttaacttta gaagtttatac 120  
aaaagtccaa tagcagggtt acaaaccatg acctctacga acaagacattt ctgtttcccc 180  
ggtgatcatca catagactgtt accccacggtc aaaagtgtt gatccgttat ccgcttgagt 240  
acttcgagaa ggcatagtt accttggat ctccgtatccg ttgcgttccaa cactctgccc 300  
cgagtgttgc tttagtgggc ggttgcgttgc actcccccacc ggcgacgggtt gcccaggctg 360  
cgttggcgc ctacccatgg ctgtatccgtt gggacgttgc ttgttaacaa ggtgtgaaga 420  
gcctatttagt ctactcaaga gtcctccggc ccctgtatgc ggctaatctt aaccacggag 480  
caaccgttca caacccatgtt agtaggttgcgtt cgtatgttgc aagtctgtgg cggaaaccgac 540  
tactttgggtt tgccgttgc ttcccttataat tcataactggc tggttgcgtt gacaattttac 600  
aaattttgttac catatagctt ttggatggc caccatgtt tgtaatgtt atttgtgtt 660  
tttttttata acgccttattt acatcacattt ttaatcaca ataaacatgtt caaatgggg 720  
ctcaagtttca aacgcacaaag accgggtgcgc acgagaatca aacgtgttca gccaatggat 780  
ccaccatcaa ttatcaactt atcaactt acaaagacatgtt tgccgtatgc ttccgtatcta 840  
gacaagacat cttcccaatgtt ccataaaaaat tcacagaacc ggtttagggc ttatgttgc 900  
aaacagcacc agctctaaac tcgccttcaagc tggaaagcatg tgggttacatgtt gaccgtgttgc 960  
ggcaaaatcac tttaggcaac tcgacttata ctacacaaaga acgagccatgtt gctatttttgc 1020  
cttacggatca tttaggcaactt tacataatgtt acatcggatgtt gatgcacccca 1080  
ctggccatgtt ctttgcgttgc acaccatgtt acatcggatgtt tgccgtatgc 1140  
cttcaagggtt atgggtgtgg aagtttaccatgtt ggttgcgttgc ttgggttgc 1200  
agaatataatgtt ctatcaacttac ttggggcgct ctgttttacatgtt cattcatgtt cagtgttgc 1260  
cttcaaaaatccatcaacatgtt gctttagggat ttttccgtatgtt accatgttgc ttatgttgc 1320  
gcaacactgtt gagtaaaacatgtt tccatgtt acatcggatgtt gatgcacccca 1380  
ggccgtgttgc acaccatgtt acaccatgtt acatcggatgtt tgccgtatgc 1440  
ttgcacgttgc ggttgcgttgc acaccatgtt acatcggatgtt tgccgtatgc 1500  
cgccacatgtt acaccatgtt acaccatgtt acatcggatgtt tgccgtatgc 1560  
actcaactgtt gatgttgcgttgc acaccatgtt acatcggatgtt tgccgtatgc 1620  
tggccacccatgtt gcccttgcgttgc acaccatgtt acaccatgtt acatcggatgtt tgccgtatgc 1680  
caccatgtt acaccatgtt acaccatgtt acaccatgtt acaccatgtt acatcggatgtt tgccgtatgc 1740  
tgcgtgttgc acaccatgtt acaccatgtt acaccatgtt acaccatgtt acaccatgtt acatcggatgtt tgccgtatgc 1800  
cctgtgttgc acaccatgtt acaccatgtt acaccatgtt acaccatgtt acaccatgtt acatcggatgtt tgccgtatgc 1860  
atatgttgcgtt acaccatgtt acaccatgtt acaccatgtt acaccatgtt acaccatgtt acatcggatgtt tgccgtatgc 1920

-continued

---

tgaacacaat	ggagatgtat	caaataccat	tgaatgacaa	ttttagcaag	gcacccatat	1980
tctgtttatc	ccttacacat	gctttctgata	aacgactgag	ccacaccatg	ttgggtgaaa	2040
tccttaattt	ttacacccat	tggacgggggt	ccatcagggt	cacccctcta	ttttgtggca	2100
gtatgatggc	cactggtaaa	ctgctccca	gctattcccc	accgggagct	aaaccaccaa	2160
ccaatcgcaa	ggatgcaatg	ctaggcacac	acatcatctg	ggacctaggg	ttacaatcca	2220
gttggccat	ggttgcaccc	tggatctcca	acacagtgt	cagacgggt	gcacgtgtg	2280
acttcactga	ggccggattt	ataacttgct	tctatcaaac	tagaattgtg	gtacctgttt	2340
caaccctac	cagtagttc	atgttaggt	ttgttagtgc	gtgtccagac	ttcagtgtca	2400
gactgcttag	ggacactccc	catattagtc	aatcgaaact	aataggacgt	acacaaggca	2460
ttgaagacct	cattgacaca	gcgataaaaq	atgccttaaq	agtgtcccaa	ccaccctcga	2520
cccagtcac	tgaagcaact	agtggagt	atagccagg	ggtgcaggt	ctaactgtg	2580
tggaaacagg	agcatcttgt	caagaaatcc	ccagtgtat	ggtggaaact	aggcacgtgg	2640
taaattacaa	aaccaggct	gaatcgtgtc	tttagtcat	cttgggaga	gctgcgtgt	2700
tcacaatct	atcccttgacc	aactccctca	agagcggaga	ggagaaaaag	catttcaaca	2760
tatggatat	tatatacacc	gacactgtc	agttacgcag	aaaatttagaa	tttttcacgt	2820
attccaggtt	tgtatcttgc	atgttttgc	tattcacaga	gaaactatct	agtacagcca	2880
gtggagaagt	gcgaaaccag	gtgtaccaga	tcatgtat	tcaccagg	gcacccgc	2940
catccatctt	ggatgactac	acatggcaat	ccttccaa	ccctccat	ttctacatgt	3000
atggaaatgc	acatccacgg	atgtcaat	cttacgttag	gattgcaat	gcctattcac	3060
acttcttac	tggcttgc	cggtgcac	ttgagggtg	gaaacccgt	gctggcaca	3120
cgttttacgg	tttagtgc	ataatgatt	ttggagttt	agcagttag	gcagtaaacc	3180
gcagtaatcc	acataatca	cacacatgt	tgagagtgta	catgaaacca	aaacacattc	3240
ggtgttgggt	ccccagacct	cctcggat	tattatacg	gggagggaga	gtggacatga	3300
tatccatgtc	aattcttac	ctggccaagg	tagactcaat	taccatctt	gggttggtc	3360
atcagaacaa	agcagtgtac	gttgcgggt	acaagatttgc	caactaccac	ctgcaacccc	3420
caagtgtatc	tttgcgtca	attagtatgt	tatgggacag	ggatttaatg	gtgggtgaat	3480
ctagagcccg	gggaactgt	accatcgcc	gtatgtgt	cagggtgt	gtttactatt	3540
gtgaatcttgc	ggggaaatgt	taccctgtc	cttttactgt	cccaacgtt	cgattcatgg	3600
aagcaaaacca	ctactatcc	gcaagatacc	agtctcacat	gctgataggg	tgccgatgtt	3660
cagaaccccg	ggactgcgg	gggatactga	ggtgcactca	ttgggtattt	gtgtatcatta	3720
ctgcaggagg	tgaagggttgc	gttaggttgc	ctgacattag	agacctctgg	gtgtatgaag	3780
aggaggccat	ggaacaggga	ataacaatgt	acatcgaaatc	tctccgcaca	gccttggcg	3840
cagggttcac	ccacacaatc	agtggaaaag	tgactgaat	gacaacgt	gttaccagca	3900
ctatcacaga	aaaactactg	aaaaacttgg	tgaaaatagt	gtcggctcta	gtgattttgt	3960
tgagaaatgc	tgaggacatt	accacatcc	ttgcacact	agcaactact	gggtgtgata	4020
tatctcccttgc	gcaatgggttgc	agaagaagg	catgtactt	actagaggat	cctttagtgc	4080
tgcccaacgg	tgatgggttgc	atgaaaatgc	tcacagggc	gtgcaatgc	gctaaaggct	4140
tagagttgtat	tagcaacaaa	atttccaatgt	ttatagatttgc	gttgggtgt	aaaattatcc	4200
cacacgttac	ggacaatgttgc	gatattctca	ccaaatgttgc	aaagtagtgc	atgtggaaa	4260
atcaaaatgc	aaccatccac	caatcttgc	ccagccaga	acaacaaag	attctttca	4320
acaaatgttag	atgttttagca	gtccagttcc	gtccgttgc	accattat	gctgtggagg	4380
cacgccgaat	taacaaatgc	gagagcaca	taaacaatta	tatacgttc	aagagcaac	4440
accgtttagtgc	accgtatgt	atgtccat	atgggtcacc	agggacgggt	aaatctatag	4500
ctacttcatt	aataggtat	gcaatagcg	agaaggaaag	cacatcgtc	tatcaatgc	4560
cacctgaccc	atctacttgc	gatggctata	aaacaacgg	ggtagtgcatt	atggacgacc	4620
taaaccaaaa	ccccgtatgttgc	atggacat	aaactgttttgc	ccaaatgttgc	tcaacagtgg	4680
agtttattcc	tccaaatggc	tcattaggg	agaagggttgc	tttgcatttaca	tctgtattatgc	4740
tccctggat	taccaatcttgc	cattcaat	taccacccat	agtggtcacc	agtgtatgc	4800
taaccagacg	atttgcatttgc	gatgtggagg	tttacacgt	gtctgaat	tcagtcaag	4860
gcaaaactgaa	tatggccacg	gccactcaat	tgtgttaggaa	ttgtccaa	cctgcaatt	4920
ttaaaaatgc	ttgccttctgc	gttgggttgc	aggcccttgc	attaatggac	aggtacacca	4980
gacaaagggtt	cactgtatgt	gagatttgc	cattatcat	gaatggaaa	aacagaagg	5040
ccaaatatccg	caatttgc	gaaatgttgc	ttcaaggacc	actaaggat	aaagatttgc	5100
agatcgatgt	gaagacagtt	ccccccctgt	agtgcatcg	tgatgtgttgc	caagcgttgc	5160
attctcaaga	gtttaggtat	tactgttgc	agaaaggctgt	gatcgatgttgc	gttactagcc	5220
agattcaact	agaaaggaaac	atcaataggat	ccatcgactat	actcaatgttgc	gttaccacat	5280
tcgcagcgt	cgccaggatgt	gttgcgttgc	tgtacaaatgttgc	cttcgcgggt	caacagggttgc	5340
catacactgg	cttgcacaaac	aaaaaaacccat	atgtccat	tatcagatgc	gctaaaggat	5400
agggggccagg	atttgcacttgc	gcaatgttgc	ttggcaaaatgc	aaacatgttgc	actgtcaaccc	5460
ccaccaacgg	tgaatttgc	atgttgcgttgc	tgcatgtat	tgtagcaata	ttgccaaccc	5520
atgcgcgttcc	aggagaaacc	atttttgc	atggggaaatgc	agtagatgttgc	ctagacggcc	5580
gagcctttaga	agatcaagcg	ggaaccaatgc	tttgcgttgc	cattattatgc	ctaaaaagaa	5640
atgagaatgtt	tagagacat	agatcacata	tttccacccat	aattactgttgc	actaaccatgc	5700
gagtgttgc	gtgttgcactt	ccaaatgttgc	tgtcccccgtt	ggtgcgttgc	5760	
ccgaacacgg	atattcttgc	ctcgttgc	gtcaatgttgc	tcgcacttgc	atgttacatgc	5820
ttccaaatgc	ggcaggccat	tgcgttgc	ttatgttgc	tttgcgttgc	5880	
tgcgttgc	cgggacacgg	tcacatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	5940
ctcaaaatca	ggggcgaatgc	cagtggatgc	ggtcat	ccaaatgttgc	tttgcgttgc	6000
taaatgc	atccaaatgc	aaatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	6060
ttaaggaaacc	atgttgc	actaaggat	ccatcgacttgc	tttgcgttgc	tttgcgttgc	6120
ccatctttc	taaatatgttgc	gggaaacaaa	ttactgttgc	ggacgttgc	atgttacatgc	6180
cagtggatgc	ctatgc	cgttacatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	6240
tggaggatgc	catgttgc	accgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	6300
gatatccttgc	tgttgc	atgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	6360
atactaagga	gatgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	6420
acgttgc	aaatgttgc	ccaaatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	6480

-continued

---

```

aagttccag ccttaatgtat tcagtgc当地 tgagaatggc ctttggcaat cttaacgc当地 6540
cttc当地 cacaat gaatccaggat gtgggatgac gatcaggc当地 tgggtgtgac ccagatttg 6600
tttgggat当地 gataccgtg ctaatggaa aaaaacttcc ctgttttgac tacacagggt 6660
atgtgc当地 ctcaaggccct gcttgggat当地 aagotcttaa aatgggtgt当地 gaaaaattg 6720
gatttggc当地 tagagttagac tatata>tagact acctgaaccac ctctc当地 accac 6780
acaagactt当地 ttgtgtcaaa ggccggcatgc catccggctg ctctggc当地 tcaatttca 6840
actcaatgat taacaaccctg atcattagga cgcttttact gagaacccatg aaggccatag 6900
acttggacca tttaaaaatg attgocctatg tgatgtgactg gatagettcc tacccccc当地 6960
aggttgc当地 tagtctctt当地 gccaatcg gaaaagacta tggactaacc atgactccag 7020
cagataaaatc agcaacatc gaaacaggatc catgggat当地 tgtaacattt ctgaaaagat 7080
tttc当地 agatgagaag tatcattcc tggtgatcc agtgatgcca atgaaagaaa 7140
ttcacgaatc aatcaatgg accaaaggacc ctagaaacac acaggatcac gtagcctc当地 7200
tgtgc当地 cttt当地 agcttggc当地 aacggatgaa aagaatacaa taaatttt当地 gctaaaatca 7260
gaagtgtgc当地 aatcgga当地 gctttat当地 tccc当地 agatc ctctatcg taccggccat 7320
ggctc当地 gactc attttagat当地 ccctaccatc gtc当地 gattgg attgggat当地 actgttgc当地 7380
gggttaaattt ttctt当地 aatgggat当地 cgggat当地 aaaaaaaaaaaa aaaagagctc ccaatca 7440
gtgaattc当地 gggccgctg aggttgc当地 accatggc当地 tttcccaacgc gttggatgca 7500
tagctt当地 gagtttactatg gtc当地 accatccaa tagtttggc当地 taatcatgtt catacgctgtt 7560
tccctgtgt当地 aattttt当地 tccctaccatc cccacacaaatc acatcgagccg gaaggat当地 7620
gtgttaaacc当地 tggggat当地 aatgatgatc ctaactc当地 ttaatttgc当地 tgc当地 ctact 7680
geccgctt当地 cagtc当地 gggaa acctgtc当地 tggc当地 ccagctgcat taatgaatcg gccaacgc当地 7740
ggggagggc当地 ggttggc当地 ttggggc当地 ttc当地 gcttcc tggctc当地 actcgctg 7800
ctcggtc当地 cggc当地 ggccgctg agtctactc当地 aaggccgatc tacgggat当地 7860
cacaatc当地 gggat当地 aacg cggaaagaa catgtgacccaa aaggccgatc aaaaggccag 7920
gaaccgtaa local aaggccgatc tgctggc当地 tttccatagg ctccgatccc ctgacgagc当地 7980
tcacaaaatc cgacgctc当地 gtc当地 gaggtg gggaaaccgg acaggactat aaagatacc 8040
ggggat当地 cctt当地 ctggatgatc ccccttggc当地 ctggatc当地 cggcttaccgg 8100
atacctgtcc当地 gccc当地 ttccggaaatc cgtggc当地 gcttccatagct cacgctgttag 8160
gtatctc当地 ctgggtgat当地 tc当地 ggcc当地 caagctggc当地 tggatgaccc aacccttccgt 8220
tcagccccatc当地 ctatccggat当地 statcgatc当地 gagtccaaacc cggat当地 aagaca 8280
cgactatc当地 ccaactgatc当地 cagccactatc gagtccatgg taacaggatc agcagagc当地 ggtatgttag 8340
cggtgtc当地 gagtttgc当地 aatggatgatc taatcgatc当地 tacatagaa gaacagtatt 8400
tggtatctc当地 gctctgatc当地 agccatgtatc cttccggaaa agagttggatc gcttccatc 8460
cgccaaacaa local accaccgtc当地 gtggatgatc tttttt当地 ttggatc当地 tgcaagccg agattacgc当地 8520
cagaaaaaaaatc ggtatctc当地 aatggatccatc gatcttccatc acggggatc当地 acgctc当地 8580
gaacgaaaaacatc tcaatggatc当地 ggatggat当地 catgatc当地 tcaatggatc当地 tcttccatc 8640
gatccctttaa local aatggatccatc aatggatccatc aatcatatc当地 agtataatatc agtaaactt当地 8700
gtctgacatc当地 taccatgtatc taatcgatc当地 ggc当地 accatc当地 tc当地 agc当地 gatctt当地 8760
ttcatccatc当地 gtggatc当地 tccctggat当地 gtatgatc当地 acgatc当地 acggggatc当地 8820
atctggccccc当地 atgtatc当地 tgatatc当地 cggccatccgc taccggc当地 cagattt当地 8880
agcaataaaacatc cagccatccgc当地 gaaggccatc当地 ggc当地 cagaatgatc ggtccatc当地 8940
ctccatccatc当地 tctt当地 aatggatccatc gttggccatc当地 agctatc当地 agtagtgc当地 cagttatc 9000
tttgc当地 gacatc当地 ttggatc当地 cttccatc当地 gatccgtggatc当地 tc当地 acgctc当地 cgttggatc 9060
gggttccatc当地 atgtatc当地 cccatccatc当地 aaggccatc当地 acatgatcc当地 ccatgttgc当地 9120
caaaaaaaaatc gttatctc当地 tc当地 ggccatc当地 gatc当地 gtgtatc当地 agaagtaatg tggccatc当地 9180
gttatctc当地 atggatccatc cggccatccatc taatctctt当地 actgtatc当地 catccatc当地 9240
atgtatc当地 ctggatgatc当地 agtactatc当地 caatgtatc当地 tgatgatc当地 gtatgatc当地 9300
accggat当地 ttttggccatc当地 cgtatccatc当地 ggataatccatc ggc当地 acatc当地 gc当地 agatc当地 9360
aaaatgtatc当地 atcatggatc当地 aacgttccatc当地 gggccatc当地 tccatccatc当地 9420
gttgatgatc当地 agtctatc当地 aacccatccatc当地 tgc当地 accatccatc当地 tgc当地 atctt当地 9480
tttccatccatc当地 gtggatc当地 gggatccatc当地 aatggatccatc当地 aatggccatc当地 9540
aaggccatc当地 cggatccatc当地 aatggatccatc当地 actcttccatc当地 tttccatc当地 attgatc当地 9600
ttatcgatc当地 tattgtatc当地 tgatgatc当地 catatggatc当地 tttccatc当地 attgatc当地 9660
aataggggatc当地 cccatccatc当地 ttccatccatc当地 agtccatccatc当地 gatccatc当地 gatccatc当地 9720
acagatc当地 aaggatccatc当地 tccatccatc当地 gggatccatc当地 agtccatccatc当地 tttccatc当地 9780
attcgatc当地 aatggatccatc当地 aatggatccatc当地 atttt当地 aatggatccatc当地 aatggccatc当地 9840
aatccctt当地 aatggatccatc当地 aatggatccatc当地 aatggatccatc当地 aatggatccatc当地 9900
caagatc当地 ctatccatc当地 acgtggatccatc当地 caatgtatc当地 gggccatc当地 cccatccatc当地 9960
gggc当地 gatccatc当地 ccaactacatc当地 aaccatccatc当地 ctatccatc当地 ttttggatc当地 cggatccatc当地 10020
taaaggccatc当地 aatcgatc当地 cttccatc当地 gggatccatc当地 aatggatccatc当地 aatggatccatc当地 10080
ggccatccatc当地 gggatccatc当地 aatggatccatc当地 aatggatccatc当地 aatggatccatc当地 10140
aagtgtatc当地 gtccatccatc当地 gctatccatc当地 cacatccatc当地 ggc当地 cgtatccatc当地 10200
gggc当地 gatccatc当地 cccatccatc当地 aatggatccatc当地 aatggatccatc当地 aatggatccatc当地 10260
tcttccatc当地 tccatccatc当地 gggatccatc当地 aatggatccatc当地 aatggatccatc当地 aatggatccatc当地 10320
acggccatccatc当地 tttccatc当地 acgtggatccatc当地 aatggatccatc当地 aatggatccatc当地 10380
tccatccatc当地 tccatccatc当地 gggatccatc当地 aatggatccatc当地 aatggatccatc当地 10440
aattcgatc当地 aggcatc当地 atacgactc当地 ctatccatc当地 10477

```

SEQ ID NO: 6 moltype = DNA length = 10474  
 FEATURE Location/Qualifiers  
 source 1..10474  
 mol\_type = other DNA  
 organism = synthetic construct

SEQUENCE: 6  
 tttaaaacatc当地 tctggggat当地 ttccatccatc当地 agaggccatc当地 gtggccatc当地 gtacttgc当地 60  
 attcgatc当地 ctttgc当地 acgtggatccatc当地 aatggatccatc当地 aatggatccatc当地 120

-continued

---

aaaaaggtaaa	tagcaggggt	acaaaaccagt	acctctacgaa	acaaggcaccc	ctgtttcccc	180
ggtgatatac	catactact	accccacggtc	aaaagtattt	gatccgttat	ccgcgttgagt	240
acttcggagaa	gccttagtac	accttggaaat	cttcgatgcg	ttgcgtctaa	cactctgccc	300
cggagtgtac	ttaggtgtat	gagttggggc	actccccacc	ggcgacgggt	gcccgaggct	360
cgttggggc	ctacccatgg	ctgtatgcgt	gggacgctag	ttgtgaacaa	ggtgtgaaga	420
gcctatttgg	ctactcaaga	gtccctccgg	ccctgaatgc	ggctaatct	aaccacggag	480
caaccgcgtca	caaccaggta	agtaggtgt	cgtatgcgt	aagtctgtgg	cggaaccggc	540
tactttgggt	gtccgtgttt	ccctttatata	tcatactggc	tgctttaggt	gacaatttac	600
aaatttggta	catatagcta	ttggattggc	gcccgcacaa	gctgggtgaa	ggggaccaac	660
tggagccaca	cacttctta	cattccacgg	cgcgccaata	aacagtgcac	atgggggtc	720
aagtttcaac	gaaaaaggacc	ggtgcgcacg	agaatcaaaa	cgtggcagcc	aatggatcca	780
ccattaatta	cactactata	aactattaca	aagacagtgc	gagtaattcc	gctactagac	840
aagacctctc	ccaagatcca	tcaaattca	cagaaccgg	taaggactt	atgttggaaa	900
cagcaccaggc	tctaaactcg	cctaaacgtgg	aagcatgtgc	tgatgtgag	cgtgtgaggc	960
aaatcacttt	aggcaactcg	actattacta	cacaagaage	agccatgt	atgttgcgtt	1020
acggtgaatg	gcccattac	ataataattt	caagaaatgc	tccggtagat	gcacccactg	1080
agccagacgt	tagtagcaac	cggtttaca	ccctagaatc	gggtgttgg	aagaccactt	1140
caaggggatg	gtggtggaaag	ttttagat	gtttgaagga	catgggaatg	tttggtcaga	1200
atatgtacta	tcaactatcg	ggggcctctg	gttacacatt	tcatgtccag	tgcaacgcgt	1260
caaatttca	ccaaggggcg	ttaggagttt	tctgtatacc	agagtttgc	atggcttgc	1320
acactgagag	taaaacgtca	tacgtttcat	acatcaatgc	aaatcttgc	gagagaggcg	1380
gtgagttac	gaacacccat	aatccgtaa	atacagacgc	cagtggggc	agaaaagttt	1440
cagcatttgg	ttatgtctg	gggttgggtt	ttcttagcagg	aaacgcctt	gtgtaccgc	1500
accagatcat	caacccatcg	accaacaa	gtgcaacaa	tgtgggtc	tacgttaact	1560
cacttgcgt	tgattgtatg	gcaaaacaca	ataactgggg	cattgtcata	ttaccactgg	1620
cacccttgc	ctttggccca	acatgtc	cacaggtgc	tattacagt	accattgcac	1680
ccatgtgtac	agaatcaat	gggttggagaa	acatcacccgt	cccagttacat	caagggttgc	1740
cgacaatgtaa	cacccatgtt	tccatcaat	tccttacat	tgatgactt	cagtgc	1800
gtgccttacc	taatttgtat	gttactccac	caatacacat	acccggggaa	gtaaaagaata	1860
tgatggaaact	agctgaaatt	gacacattga	tcccaatgaa	cgcaagggtac	ggggaaagggt	1920
acacaatgg	gtatgtatca	ataccatgt	atgacaattt	gagcaaggca	ccttatattt	1980
gtttatccct	atcacctgt	tctgtat	gactgagcc	caccatgtt	ggtaatcc	2040
taaatttatt	cacccatgtt	acggggtca	tcaggttac	ctttctattt	tgtggcagta	2100
tgatggccac	tggttaaactg	ctccctcag	atccccacc	gggagctaaa	ccaccaacca	2160
atcgcaagga	tgcaatgtca	ggcacacaca	tcatctgg	cctagggtt	caatccagtt	2220
gttccatgtt	tgcacccgtt	atcttcaaca	tgatgtac	acgggtgtc	cgtgtact	2280
tcactgagg	cggttattata	acttgc	atcaactat	atttgggtt	cctgttca	2340
cccttaccag	tatgttcat	ttaggcttgc	ttagtgcgt	tccagactt	agtgtcagac	2400
tgtttaggg	cactccat	attatgtt	cgaaaactat	aggacgtaca	caaggcattt	2460
aagacctct	tgacacatcg	ataaaatgt	ccttaagat	gtcccaacca	ccctcgaccc	2520
agtcaactga	agcaactatgt	ggatgttata	gcccagggtt	gccagctct	actgtgttgc	2580
aaacaggaggc	atctgttca	gcaatcccc	gtgtatgtt	ggaaacttgg	cacgtgttta	2640
attaccaaac	cagggttca	tcgtgtctt	tgcttctt	ggggagact	gcgtgtgtca	2700
caatecttac	cttacccat	tccttcaaga	gcccgggg	gaaaaggcat	ttcaacat	2760
ggaatattac	atacccgac	actgtccat	tgccatgtt	attagaattt	ttcacgtt	2820
ccagggttga	tcttggat	actttgtat	tcacagagaa	statccat	acagccatgt	2880
gagaatgtcg	aaacccatgt	taccatgtt	tgatattatc	accaggggca	ccccggccat	2940
catcttgg	tgacttacaca	tggcaatct	tttcaaccc	ttccatctt	tatgtat	3000
gaaatgcacc	ttccacggat	tcaatcc	acgttggat	tgccatgtt	tattcact	3060
tctacgttgg	cttgcacgg	gttgcactt	agggttgg	caccatgt	ggcgacacgt	3120
tttacgggtt	atgttgcata	aatgttgc	tgatgttgc	agtttagat	gtaaaaccgc	3180
gtatccatca	tacaatcac	acatctgtt	tgatgtat	gaaacaaa	cacatccgt	3240
gttgggtcc	cgacccctt	cgagctgtat	tatacagggg	agagggttgc	gacatgtat	3300
ccagtgtcaat	tctacccct	gccaaggtag	actcaattt	cacttttgg	tttggtc	3360
agaacaaacg	atgttgcac	ggccgttaca	agatgttca	tttcaaccc	gcaaccccaa	3420
gtgtatccat	gaatgtca	atgttgc	ggatgttgc	tttacgtt	gtggaaatct	3480
gagcccccggg	aactgtatcc	atccgttca	gtatgttgc	gtgtggat	tacttgcgt	3540
aatcttagg	gaagtactac	cctgttca	ttactggcc	aacgtttc	ttcatggaa	3600
caaaccgtact	ctatccat	agataccat	ctcacatgt	gatgggtgc	ggatgttgc	3660
aacccgggg	tcgcgttgg	atacttgc	gactctat	ggtaatttgc	atcttact	3720
cagggttgc	agggttgc	gtcttgcgt	acatgttgc	cctctgggt	tatgttgc	3780
aggccatgtt	acagggata	acaagctaca	tcaatgtt	cgccacagcc	tttggcgc	3840
ggttcaccca	cacaatcg	gagaatgtt	ctgatgtt	aaacgttgc	tttggcgc	3900
tcacagaaaa	actatgtt	aaacttgcgt	tttgcgttgc	tttgcgttgc	tttgcgttgc	3960
gaaattatgtt	ggacat	acgttccat	caactatgc	actatgttgc	tttgcgttgc	4020
ctcttggca	atgttgc	agaaggat	gtgtactt	agatgttgc	tatgttgc	4080
gccaagggtt	tgggttgc	aaatgttca	tgatgttgc	tttgcgttgc	tttgcgttgc	4140
agtggattt	caacaaat	tccatgtt	tgatgttgc	aaatgttgc	tttgcgttgc	4200
acgcttacgg	caagggttgc	tttgcgttgc	tgatgttgc	tttgcgttgc	tttgcgttgc	4260
aaatttgcac	catccat	tcttgc	gccaaggat	acaaggat	tttgcgttgc	4320
atgttgc	atgttgc	gtgtactt	tttgcgttgc	tttgcgttgc	tttgcgttgc	4380
gcccgttta	caaaatgg	tgatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4440
gtattgttgc	atgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4500
cttcattat	aggtatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4560
ctgaccat	tcacttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4620
acaaaaccc	cgatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	4680

-continued

---

ttattccctcc	aatggcctca	ttagaggaga	agggcatttt	gtttacatct	gattatgtcc	4740
tggcttcact	caacttcat	tcaattgtac	caccacagt	ggctcacagt	gtgecttaa	4800
ccagacgatt	tgcatttgat	gtggaggttt	acacgatgtc	tgaacattca	gtcaaaggca	4860
aactgaatat	ggccacggcc	actaattgt	gttaaggattt	tccaacact	gcaaatttt	4920
aaaaggtgtt	ccctetcgtt	tgtggaaagg	ccttgcattt	aatggacagg	tacaccagac	4980
aaagggtcac	tgttagatgag	attaccatc	taatcatgaa	tgagaaaaac	agaaggccc	5040
atatcgccaa	ttgcattggaa	gecttggttc	aaggaccat	aaggatataaa	gatttgaaga	5100
tcgatgtgaa	gacagttccc	ccccctgagt	gcatcagtga	tttggatcaa	gcagtgatt	5160
ctcaagagg	tagggattac	tgtgagaaga	aaggctggat	cgtaacgtt	actagccaga	5220
ttcaactaga	aaggaacatc	aataggccca	tgactatact	ccaagcttt	accacatcg	5280
cagcagtcg	aggagtagt	tatgtatgt	acaacttct	cgccgtcaa	cagggtgc	5340
acactggott	gccaaacaaa	aaacccatg	tccctactat	cagagtgc	aaagtccagg	5400
ggccaggatt	tgactacgca	gtggcaatgg	caaaaagaaa	catagttact	gcaaccacca	5460
caaaagggtga	atttacatc	ctagggtgc	atgataatgt	agcaatattt	ccaacccat	5520
ccgctccagg	agaaacattt	attatgtat	ggaaagaagg	agagatccta	gacgccagag	5580
ccttagaaga	tcaagcggga	accatcttg	agatccat	tattactcta	aaaqaatg	5640
agaagtttag	agacatcaga	tcacatattc	ccacccaaat	tactgaaact	aacgatggag	5700
tgttgatctgt	gaacactagc	aagtacccc	atatgtatgt	ccccgttgg	gctgtgaccc	5760
aacagggtata	tcttaatctc	agtggacgtc	aaactgctcg	cactttaatg	tacaacttcc	5820
caacaaggc	aggccagtg	ggagaaatca	tcacttgatc	tggcaaaatc	attggatgc	5880
atgttggccc	gaacggttca	catgggttg	ctgcagcc	caagcgatca	tacttactc	5940
aaaatccagg	cgaaatccag	tggatgggt	catcaaaaga	agtgggttac	cccattataa	6000
atgcctccatc	caagacaaag	tttagaacc	ctatgtttt	ctatgtttt	gaaggtgtt	6060
aggaaccagg	tgtactact	aagaatgacc	ccagactaa	aacagattt	gaagaaggcc	6120
tctttctaa	atatgtgggg	aacaaaatta	ctgaagtgg	cgagtcatg	aaagaaggag	6180
tggatcaacta	tgcaggacag	ttaatgtac	tggatataca	cacagaaacag	atgtgcctgg	6240
aggatgcctat	gtacggta	gatggcttgc	agggccctgg	tcttagoact	agtgtgttat	6300
atcccttattgt	tgcataatgggg	aaaaaaa	gagacatttc	agataaacag	accagagata	6360
ctaaggagat	gcagagactt	ttagatact	atgaaatcaa	tcttacattt	gtcacgtac	6420
tgaagatgt	actcaggta	aagactaaag	tggacaagg	aaagtaaaga	ttgattgaag	6480
cttccagcctt	taatcaggta	gttgcattga	gaatggctt	tggcaatctt	tacgcactt	6540
tccacaagaa	tccagggtgt	gtgacaggat	cagcagttg	ttgtgacc	gattttttt	6600
ggagtaagat	accagtgc	atgaaagaaa	aactcttcgc	ttttgactac	acagggtatg	6660
atgcctact	cagccctgt	tggttgttgc	ctcttaaaat	ggtgtttag	aaaattggat	6720
ttggcgttag	agtagactat	atagacttac	tgaaaccat	tcaccac	tacaaaaaaca	6780
agacttattt	tgtcaaggc	ggcatccat	ccggctgtc	tggcacctca	attttcaact	6840
caatgattta	caacctgtat	attagggatc	ttttacttgc	aacctacaag	ggcatagact	6900
tggaccattt	aaaaatgatt	gcctatgg	atgacgtat	agtttctac	ccccatgagg	6960
ttgacgctag	tctccatgg	caatcgagg	aaqactatgg	actaaccat	actccagcag	7020
ataaatcage	aaccttggaa	acagtccat	gggagaatgt	aacatttct	aaaagattt	7080
tcagagcaga	tgagaatgt	ccatctctgg	tgcatccat	gatgcaat	aaagaaattc	7140
acgaatcaat	catgggacc	aaggacc	taaaacacaca	ggatcagc	cgctcg	7200
gcctatttgc	ttggcacaac	ggtaaagaag	aatacaataa	atttttatgt	aaaatcagaa	7260
gtgtgccaat	cgaaagat	ttatgttgc	cgagatct	tacatgtac	cgccgtatgc	7320
tcgacttatt	ttagtacacc	taccttcgt	ggatggatt	gggttataact	tttgcattgg	7380
taaatttttc	ttaatttgc	agaaaaaaaa	aaaaaaaaaa	agagctccca	atcactatgt	7440
aattcggcc	ccgcctcg	tcgacat	ggggagatgt	ccaaacgggt	ggatgcat	7500
cttgagattt	ctatagtgc	accttaatag	cttggctgt	tcatgtat	agctgtttcc	7560
tgtgtgaaat	tgttattccgc	tcacaaatcc	acacaatca	cgagccggaa	gcataaaatg	7620
taaaggcttgg	ggtgcctaa	gagtgcata	actcacat	attgcgttgc	gtcaactgc	7680
cgcttccat	tccggaaacc	tgtgcgtca	gtgtcattaa	tgaatcgccc	aacgcgcgg	7740
gagggcgggt	ttgcgtatttgc	ggcgttcttgc	cggttctcg	ctcaactgt	cgctgcgtc	7800
ggtcgttccgg	ttccggccgg	cggtatcgc	tcaactcaaa	cggttaat	gggttac	7860
agaatcagg	gataacgc	gaaagaacat	gtgagcaaa	ggccagcaaa	aggccaggaa	7920
ccgtaaaaag	ccgcgtgtc	tggcgtttt	ccataggctc	ccccccctgc	acgagcat	7980
caaaaactca	cgctcaatgc	agagggtgg	aaacccgc	ggatataaa	gataccaggc	8040
gtttccccc	ggaaatgtcc	tcgtcg	tctgttcc	accctccgc	ttacccgata	8100
cctgtccgc	tttctccctt	cgggaaatgc	ggcgcttct	catagtcac	gctgttagta	8160
tccagttcg	gtgttaggtc	tgcgtccaa	gttgggtgt	gtgcacgaa	cccccg	8220
ggccgacgc	ttgcgcctt	ccggtaacta	tgcgttgc	tcaacccgg	taagacacga	8280
tttccatcoca	tcggcagcag	ccactgttgc	caggatttgc	agagcggat	atgtaggc	8340
tgcatacag	ttcttgcatt	ggtggcttac	ctacggctac	actagaagaa	cagtatttgc	8400
tatctgcgt	tcgtgc	cgttac	cgaaaaaaa	gttggtagt	tttgcattcc	8460
caaaacaaacc	accgtgttgc	ggcggttgc	ttttgttgc	aaggcagc	ttacccgc	8520
aaaaaaagg	tctcaagaag	atccttgcatt	tttttctac	gggttgc	actcagtgg	8580
cgaaaaactca	cggttgcatt	ttttgttgcatt	gagattatca	aaaaggatct	tcaccttagt	8640
ccttttaat	taaaaatgaa	gttttaatc	aatctaaatg	atatatgtat	aaacttgg	8700
tgacagttac	caatgttac	tcagtgc	acccat	tcgcgtat	tatccgtt	8760
atccatattt	gtctgtact	ccgtgtgt	gataactac	atacgggg	gttaccat	8820
tggccccagt	gtgtcaatgc	taccgc	ccacgc	ccggctcc	atttac	8880
aataaaaccc	ccagccggaa	ggccgcgc	cagaatgtt	ctgtcaat	tatccgc	8940
catccgtct	attaattgtt	ggccggaa	taggtatgt	agttcgcc	ttaatgttt	9000
gcccacacgtt	gttgcatttgc	tcacggat	cggtgttgc	cgctcg	tttgcatt	9060
ttcattcgc	tccggttccc	aacgtatca	ggcgat	tgcattccca	ttttgtgca	9120
aaaagcgggtt	agctccttcg	gtcctccat	cggttgc	agaatgttgg	ccgcgtgtt	9180
atcactatcg	tttatggc	cactgcata	ttctcttact	gtcatgc	ccgtat	9240

-continued

---

cttttctgtg	actggtgagt	actcaaccaa	gtcattctga	gaatagtgt	tgcggcgacc	9300
gagttgtctc	tgcccgcgct	caatacgga	taataccgcg	ccacatagca	gaactttaaa	9360
agtgtctatc	atggaaaac	gttctcggg	cgaaaaactc	tcaaggatct	taccgtgtt	9420
gagatccagt	tcgatgtaac	ccactcgto	acccaactga	tcttcagcat	cttttacttt	9480
caccagcggtt	tctgggtgag	caaaaacagg	aaggcaaaaat	gcccacaaa	agggaataag	9540
ggcgacacgg	aaatgttgaa	tactctact	cttcctttt	caatatttt	gaagcattt	9600
tcaggggttat	tgtctcatga	gcccatacat	atttgaatgt	atttagaaaa	ataaacaat	9660
aggggttccg	cgccacattt	ccggaaaatg	gccacctgt	cggtgtgaa	ataccgcaca	9720
gatgcgtaa	gagaaaatac	cgcacatgg	aattgtaaac	gttaatattt	tgtttaaaatt	9780
cgcgtaat	ttttgttaat	tcagtcatt	ttttacccaa	taggcggaaa	tcggcaaaaat	9840
ccttataaa	tcaaaaagaat	agacccgat	agggttgagt	gttggccag	tttggacaaa	9900
gagtccacta	ttaaaaacgc	tggactccaa	cgtccaaagg	cgaaaaacccg	tctatcagg	9960
cgatggccca	ctacgtgaac	catcccccta	atcaagttt	ttggggctga	ggtgcgtaa	10020
agcaactaat	cggaacctca	aaggagggcc	ccgattttga	gttgcacggg	gaaagccggc	10080
gaacgtggcg	agaaaaggaa	ggaagaaag	gaaaggagggc	ggcgttaggg	cgctggcaag	10140
tgtagcggtc	acgctcgccg	taaccacca	acccgcggc	cttaatcgcc	cgctacagg	10200
cgegtccatt	cgccattcag	gctgcgcaac	tgttggaaag	ggcgtatgg	ggggccctct	10260
tcgcttattac	ccgcgttggc	gaaaggggga	tgtgctcaaa	ggcgtatgg	ttgggttaacg	10320
ccagggtttt	cccgatcag	acgttggaa	acgacggcc	gtgaattgt	atacgactca	10380
ctataggccg	aattggggcc	gacgtccat	gttccggcc	gcccattgg	ccggggaaat	10440
tcgatttgg	catgtctata	cgactcacta	tagg			10474

SEQ ID NO: 7                    moltype = DNA length = 10526  
 FEATURE                        Location/Qualifiers  
 source                        1..10526  
 mol\_type = other DNA  
 organism = synthetic construct

SEQUENCE: 7

ttaaaacacg	tctgggggtt	ttcccacccc	agaggccac	gtggcggt	gtactcttgt	60
attaccggtac	ctttgtacgc	ctgtttgt	tcccttcccc	cgtaacttta	gaagcttata	120
aaaagggtca	tagcagggtt	acaaaggat	acctctacga	acaaggac	ctgtttcccc	180
gtgtatca	catagactgt	acccacgg	aaaatgtatt	gtatcggtt	ccgcttgagt	240
acttcgagaa	gctctgtatc	accttggaa	cttcgatcg	ttgcgttcaa	cactctggcc	300
cgagtgttagc	ttaggtgtat	gagttgggg	actccccacc	ggcgcacgg	gcccagg	360
cggtggccgc	ctacccatgg	ctgtatcg	ggggacgtgt	ttgtgaacaa	gttggaa	420
gccttattgt	ctactcaaga	gtcctccgg	ccctgtatgc	ggctaatctt	aaccacgg	480
caaccgtctc	caacccagg	tgatgttgt	cgtatcg	aagtctgtt	cggaacc	540
tactttgggt	gtccgtgtt	ccctttat	tcatactgg	tgttattgtt	gacaatttac	600
aaattgttac	catatgtca	ttggattgg	ggccgcgcaca	tgttggaa	ggggaccaac	660
gatacagctg	ttgttgggg	accaatgg	gccacac	tccttacatt	ccatcaccca	720
cacacttct	taacattccac	ggcgcgc	taaaatgtc	aatggggc	tcaagttca	780
acgaaaaga	ccgggtcgca	cgagaatcaa	aacgtgg	ccaatgg	caccattaa	840
taactacta	tcaacttata	caaaggat	ggcgtat	ccgtactag	acaagac	900
tcccaaggat	catcaaaat	cacagaacc	gtttaagg	actatgtt	aaacgcacca	960
gtctttaact	ccgccttacgt	ggaaggat	gggtatcg	accgtgtt	gcaaatcact	1020
ttaggcaact	cgacttattac	tacacaagaa	gcaggcaat	ctattgtt	ttacgg	1080
tggcccaactt	acataatgtt	ttcaggat	aatccgtt	atgcaccc	tgagcc	1140
gttagtagca	acccgttta	cacccatgg	tcgggttctt	ggaaggac	ttcaagg	1200
tgggtgttga	agtttccat	ttgttggaa	gacatgg	ttttgttca	gaatatgt	1260
tatcaact	tggggcgctc	tggtttacacc	attcatgtt	atgtcaac	ttccaaat	1320
caccaagggg	cggttgggt	tttgcgtat	ccagatgtt	tcatgtt	caacactg	1380
agtaaaacgt	catacggtt	atacatcaat	gcaaatctt	gttgg	gggtt	1440
acgaaacac	acatccgtc	aaatcac	gcccgtt	gcaaaaat	tgcaagg	1500
gattatttgc	tgggttctgg	tgttctat	ggaaacgc	ttgtgtt	gcaccat	1560
atcaacccat	gttccaaac	cagtgc	attgtgg	catacg	ctcactt	1620
attgtgtt	tggcaaaaca	caataat	ggcattgtt	tattaccat	ggcacc	1680
gccttgcgc	caacatcg	accacagg	ccattat	tgaccat	acccatgt	1740
acagaattca	atgggttgg	aaatcacc	gtccc	gttgc	atcaagg	1800
acacacccgt	gttccaaat	attccat	tctgtat	ccactgt	ttccat	1860
cctaatttt	atgttactcc	accaat	ataccgg	ttatgtt	ggaa	1920
ctagctaaa	ttggacat	gttccaaat	aacggcgtt	acgggg	gacacaat	1980
gagatgtatc	aaataccatt	gaatgacaat	tttgcgtt	cacctat	ttgtt	2040
ctatcac	tttgcgtat	acgactg	cacaccat	tgggt	aaat	2100
taacccat	ggacgggg	catcggtt	acccgtt	tttgcgtt	ttccat	2160
actggtaac	tgtctt	cttgcgtt	ccgg	tttgcgtt	ttccat	2220
gtatgtatc	taggcacaca	catatgtt	ggcgtt	tacaat	tttgcgtt	2280
gttgcacgt	ggatcttca	cacagtgt	agacgg	tacgtgt	tttgcgtt	2340
ggcggttta	taacttgc	ctatcaa	agaatttgc	tacgtgt	tttgcgtt	2400
agtatgttca	tgttaggtt	tttgcgtt	tgccatgt	tcagtgt	tttgcgtt	2460
gacactccc	atattatgt	atcgaaact	ataggac	gttgcgtt	tttgcgtt	2520
attgacac	cgatataa	tgcctt	tttgcgtt	ccaccc	tttgcgtt	2580
gaagcaacta	gtggagtt	tagecagg	gttgcgtt	tttgcgtt	tttgcgtt	2640
gcatctggc	aagcaatccc	cagtgt	gttgcgtt	tttgcgtt	tttgcgtt	2700
accagggtct	aatcggtt	tttgcgtt	tttgcgtt	tttgcgtt	tttgcgtt	2760
tccttgc	actcttccaa	gagcgg	tttgcgtt	tttgcgtt	tttgcgtt	2820
acatacacc	acactgttca	tttgcgtt	tttgcgtt	tttgcgtt	tttgcgtt	2880

-continued

---

gatcttggaa	tgactttgt	attcacagag	aactatccta	gtacagccag	tggagaagt	2940
cgaaccagg	tgtaccatg	catgtatatt	ccaccagggg	caccggccc	atcatccctgg	3000
gatgactaca	catggcaatc	ctcttcaaa	ccttcacatct	tctacatgt	tggaaatgca	3060
cctccacgga	tgtcaattcc	ttacgttaggg	attgccaatg	cctattcaca	cttctacatg	3120
ggctttgcac	gggtgcact	tgagggttag	aacaccgtat	tgggcgacac	gttttacgg	3180
ttagtgtcca	taaatgattt	tgggttttgc	gcagtttagag	cagtaaaaccc	cagtaatcca	3240
cataacaatac	acacatctgt	gagagtgtac	atgaaaccaa	aacacattcg	gtgttgggtc	3300
ccccagaccc	ctcgagctgt	attatacagg	ggagagggg	tggacatgtat	atccagtgc	3360
attctacctc	tggccaaggt	agactcaatt	accacttttgc	gttttggta	tcagaacaaa	3420
gcagtgtacg	ttgcccgtta	caagatgttgc	aactaccac	tagcaacccc	aagtgtatcac	3480
ttgaatgcaa	ttagtatgtt	atgggacagg	gattnaatgg	tgggtggaa	tagagccgg	3540
ggaactgata	ccatcgccag	atgtatgttgc	agggtgtggag	tttactatgt	tgaatctagg	3600
aggaagtact	accctgtcac	ttttactggc	ccaacgtttc	gattcatgga	agcaaacac	3660
taactatccg	caagatccaca	gtctcacatct	ctgataggg	ggggatttgc	agaacccggg	3720
gactgcggtg	ggatactgt	gtgcactat	ggggtaatttgc	gtatcattac	tgcaaggaggt	3780
gaagggttag	tagccttgc	tgacattaga	gacccctggg	tgtataga	ggaggccatg	3840
gaacaggggaa	taacaagct	catgaatct	ctcgccacag	cctttggcgc	agggttca	3900
cacacaatca	gtggagaaatg	gactgttgc	acaacgtatgg	ttaccacac	tatcacagaa	3960
aaactactgt	aaaacttggt	gaaaatagtgt	teggcttgc	tgattgttgc	gagaatatt	4020
gaggacacta	ccacgtatct	tgcaacacta	gcaactactcg	gggtgtat	atcccttgc	4080
caatgggtt	agaagaaggc	atgtacttca	ctagagattc	cttatgtat	gcccacatg	4140
gatgggttga	tgaagaaattt	cacagaggcg	tgcattgtcg	ctaaaggctt	agagtggat	4200
agoaacaat	tttccaaatgtt	tatagattgtt	ttgaatgttgc	aaatatcc	agacgctaa	4260
gacaagggttgg	aatttctc	caatgttgc	caatgttgc	tggtaaaaa	tcaaaatttgca	4320
accatccacc	aatcttgc	cagccaa	caacaagaga	tttttttca	caatgtgaga	4380
tggtagcag	tccagttcc	tcggtttca	ccattatacg	ctgtggaggc	acgcggatt	4440
aaaaaaatgg	agagcaat	aaacaaat	atacagttca	agagcaacaa	ccgtatttgc	4500
ccagtatgt	tgctcattca	tgggttcc	gggacgggttgc	aatctatagc	tacttcatttgc	4560
ataggtagag	caatagcaga	gaaggaaatgc	acatcgttgc	attcaatgcc	acctgaccca	4620
tccacttgc	atggctataa	acaacaagg	gtatgttgc	tggacgac	aaaccaaaac	4680
cccgatgttgc	ttggatgttgc	actgtttgc	caatgttgc	caacagtgg	gtttatttgc	4740
ccaaatggc	cattatgg	gaagggttgc	tgggtttat	ctgattatgt	cctggcttgc	4800
accaacttcc	attcaatgtt	accacccaca	gttgcttca	gtgtatgc	tttccacat	4860
tttgcat	atgtggatgt	tttacatgt	tctgttgc	caatgttgc	caatgttgc	4920
atggccac	ccacttcaat	gtgttgc	tttacatgt	tttacatgt	tttttttgc	4980
tgccctctcg	tttgtggaa	ggcccttgc	ttatgttgc	ttatgttgc	ttatgttgc	5040
actgttagat	agatttacc	attatcatgt	aatgagaaaa	acagaaggc	caatatccgc	5100
aatttgcatt	aaacccctgtt	tcaaggatca	ctaaggat	aaagatttgc	gatgtatgt	5160
aagacagtgc	ccccccat	gtgtatgt	gttgcattgt	aagcgttgc	tttctca	5220
gtttaggttgc	actgttgc	gaaaggctgg	atcgatgttgc	tttactatgt	tttactatgt	5280
gaaaggaaaca	tcaatagg	catgtatca	cttcaatgttgc	tttacatgttgc	tttacatgttgc	5340
gcaggaggtag	tgtatgttgc	gttacaaatc	ttccggcgttgc	tttacatgttgc	tttacatgttgc	5400
tttccaaatca	aaaacccat	tgttccatct	atcagatgt	cttcaatgttgc	tttacatgttgc	5460
tttgactatcg	cgttgc	ggccaaat	aatatgttgc	tttcaatgttgc	tttcaatgttgc	5520
gaatttacca	tgcttaggg	gtatgttgc	gtatgttgc	tttcaatgttgc	tttcaatgttgc	5580
ggagaaacca	tttattatgt	tgggttgc	gtatgttgc	tttcaatgttgc	tttcaatgttgc	5640
gtatcaatgc	gaaccaat	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	5700
agagacatca	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	5760
gttacatgttgc	caatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	5820
tatcttcaatc	tcgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	5880
gcaggccatgt	ggggatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	5940
ggggacgggtt	catacttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6000
ggcgaaatcc	atgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6060
tccaaatgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6120
gtctgtacttca	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6180
aaatatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6240
atgtcaggat	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6300
atgtacgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6360
gttgcattatgg	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6420
atgcagatc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6480
gaactcgatgt	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6540
cttaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6600
aatccaggatgt	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6660
ataccatgtc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6720
ctcagccctg	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6780
agagatgtat	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6840
tgtgtcaat	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6900
aacaacatgt	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	6960
ttaaaaatgt	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	7020
agtctccatgt	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	7080
gcaacccttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	7140
gatgagaatgt	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	7200
atccatttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	7260
atcagatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	7320
gcttggcaca	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	7380
atcggaaatgt	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	tttcaatgttgc	7440

-continued

---

tctttaattc	ggagaaaaaaaa	aaaaaaaaaa	aaagagctc	caatcaactag	tgaattcgcg	7500
gcccgcgtca	ggtcgaccat	atggagago	tcccaacccg	ttggatgcat	agcttgagta	7560
ttctatagtg	tcacctaataat	agcttggcgt	aatcatggtc	atagctgtt	cctgtgtgaa	7620
atttgttatcc	gocacaaatt	ccacacaaca	tacgagccgg	aagcataaag	tgtaaaggct	7680
gggggtgecta	atgagtggc	taactcacat	taattgcgtt	gcgctactg	cccgatttcc	7740
agtccggaaa	cctgtcggtc	cagctgcatt	aatgaatccg	ccaaacgcgcg	gggagaggcg	7800
gtttgcgttat	tgggcgctct	tcggcttct	cgctcaactg	ctcgctcgcc	tcggctcggtc	7860
ggotgcggcg	agccgtatca	gctcaactaa	aggccgttaat	acggttatcc	acagaacatcg	7920
gggataacgc	aggaaaagaac	atgtgagcaa	aggccagca	aaaggccagg	aaccgtaaaa	7980
aggccgcegtt	gttgcgtt	ttccataggc	tcggcccccc	tgacgagcat	cacaaaaatc	8040
gacgctcaag	tcagagggtgg	cgaaaccggc	caggactata	aaagatccag	gcgtttcccc	8100
cttgaaggtc	cctctgtcgc	tctctgttc	cgaccctgc	gcttacccgga	tacctgtccg	8160
cctttctccc	ttcggaaagc	gtggcgctt	ctcatagctc	acgctgttagg	tatctcaagtt	8220
cggtgttagt	cggtcgctcc	aagctgggg	gtgtgcacca	accccccgtt	cagcccgacc	8280
gctgcgectt	atccggtaac	tatctgttt	agtccaaacc	ggtaaagacac	gacttatacg	8340
cactggcage	agccactgtt	aacaggatta	gcacagcgag	gtatgttagc	ggtgtacag	8400
agttcttga	gtggggcct	aactacggct	acactagaag	aacagtattt	ggtatctcg	8460
ctctgtgaa	gcccgttacc	ttcggaaaaa	gagttggtag	ctcttgcatt	ggcaaaacaaa	8520
ccaccgctgg	tagccgtgg	ttttttgtt	gcaaggcagca	gattacgcgc	agaaaaaaaag	8580
gatctcaaga	agatcttttgc	atcttttcta	cggggtctg	cgctcaagtgg	aacgaaaact	8640
cacgttaagg	gattttggtc	atagagattat	caaaaaggat	cttcacactag	atccttttaa	8700
attaaaaatg	aaaggtaataa	tcaatctaaa	gtatatatga	gtaaacttgg	tctgacagtt	8760
accatgtt	aatcagttag	gcacatct	cacgcacatcg	totatcttc	tcatccatag	8820
ttgcctgact	ccccgtcg	tagataacta	cgataccgg	gggcttacca	tctggcccca	8880
gtgctgcaat	gataccgcg	gacccacgt	cacccgtcc	agatttatca	gcaataaaacc	8940
agccagccgg	aaaggcccgag	cgccagaatgt	gtcctgcaac	tttateccgg	tccatccagt	9000
ctattaattt	ttggggggaa	gctagatgaa	gtatgtcg	agttatagt	ttgcgcacag	9060
tttggccat	tgctacaggc	atcggtgtt	cacgctcg	gtttgttagt	gcttcattca	9120
gtcccggttc	ccaacgatca	aggcgagtt	catgatccc	catgtgtgc	aaaaaagcg	9180
ttagctctt	cggtctccg	atcgttgc	gaagtaatgt	ggccgcagtg	ttatcactca	9240
tggttatggc	agactgtcat	aatttcttca	ctgtcatgc	atccgtaa	tgctttctg	9300
tgactgttg	gtactcaacc	aagtattctt	gagaatagtgg	tatgcggcga	ccgagtgtct	9360
cttgcggcgc	gtcaataacgg	gataataccg	cgccacatag	cagaacttta	aaagtgc	9420
tcattggaaa	acgttttcg	ggggcggaaac	tctcaaggat	cttaccgc	ttgagatcca	9480
gttcgatgt	accacactgt	gcacccaa	gatcttc	acatctttact	ttcaccagcg	9540
tttctgggt	agccaaaaca	ggaaggcggaa	atgccc	aaagggata	aggccgcacac	9600
ggaatgttg	ataactcata	cttccctt	ttaatatta	ttgaagcatt	tatcagggtt	9660
attgtctcat	gagcgatata	atattgaat	gtatttagaa	aaataaaacaa	ataggggtt	9720
cggcgacatt	ccccggaaa	gtggcaccc	atgcgggttg	aaataccgca	cagatgcgt	9780
aggagaaaaat	acccgtatcg	gaaattgtt	gctgttat	tttgtttaaa	ttcgcgtttaa	9840
atttttgta	atcaatcg	tttttttacc	aataggccg	aatcggaaa	atcccttata	9900
aatcaaaaaga	atagaccgag	atagggttga	gtgttgc	atgttggaa	aagagtccac	9960
tataaaagaa	cggtggactt	aacgcacaa	ggcgaaaaaa	cgctctatcg	ggcgatggcc	10020
cactacgtg	accatcaccc	taatcaatgtt	ttttggggtc	gagggtgcgt	aaagacttaa	10080
atcggaacc	taaaggggc	ccccgttta	gagotttgc	ggggaaacccg	gcgaacgtgg	10140
cgagaaaagg	agggagaaaa	gccaaaggag	cgggcgctag	ggcgctggca	agtgtacgc	10200
tcacgtcg	cgtaaccacc	acacccggc	cgcttaatgc	ggcgctacag	ggcgctggcc	10260
ttccgcattc	aggctggc	actgttgg	agggcgtatc	gtgcggcc	cttcgtatt	10320
acgcccacgt	gcaaaaa	gatgttgc	aaaggcgat	atgttggtaa	cgccagggtt	10380
ttcccagtc	cgacgttg	aaacgacgg	cagtgtt	taatacgt	cactataagg	10440
cgaatttggc	ccgacgtcg	atgtcccg	ccgccc	atgcgttgc	ggccgcggga	10500
ggcatgtt	tacgactc	tatagg				10526

SEQ ID NO: 8                  moltype = DNA length = 10598  
 FEATURE                  Location/Qualifiers  
 source                  1..10598  
 mol\_type = other DNA  
 organism = synthetic construct  
 SEQUENCE: 8  

ttaaaaacagc	tctggggttt	ttcccaccc	agagggccac	gtggcggcta	gtactctgg	60
attacgttac	ctttgtacgc	ctgtttgt	tcccttcccc	cgtaacttta	gaagtttac	120
aaaagttca	tagcagggtt	acaaaaccgt	acctctacga	acaaggacatt	ctgtttcccc	180
ggtgtatca	catagactgt	acccacggtc	aaaagtattt	gatccgttat	ccgcttgcgt	240
acttcgagaa	gctactgtt	accttggat	cttcgtatgc	ttgcgtccaa	cactctggcc	300
cgagtgttgc	ttaggtgtat	gatgttgg	actccccacc	ggcgcacggtg	ggccagggtt	360
cggtggcgcc	ctacccatgg	ctgtatcg	gggacgtat	ttgttacaaa	ggtgtgaaga	420
gcctattgtt	ctactcaaga	gtcctccgg	ccctgtatgc	ggctaatctt	aaccacggag	480
caaccgcgtca	caacccatgt	agttaggtt	cgtaatcg	aagtctgtgg	cggaaccgac	540
tacttgggt	gtccgtt	ccctttat	tcatactgg	tgcttatgtt	gacaatttac	600
aaatttgttac	catatgttca	ttggatttgc	caccatgtt	tgtgttat	atttgatgtt	660
ttcttcata	acccttata	acatcgccgc	gccgcacacg	ttgttgaagg	ggaccaacga	720
taacgtgtt	tgaagggac	caactggac	cacacactt	tttacatcc	atcaccacaca	780
cacttcctt	cattccacgg	cgccgcacat	tttaatc	aataaaacgt	gcaaatgggg	840
gctcaagttt	caacgc	gaccgggtcg	cacgagaatc	aaaacgttgc	agccaatggg	900
tccaccat	attacactac	tatcaactat	tacaaagaca	gtgcgttac	ttccgtact	960
agacaagacc	tctcccaaga	ttcacagaac	cggttaagg	cttaatgtt		1020

-continued

---

aaaacagcac	cagctctaaa	ctcgcttaac	gtggaaagcat	gtgggtacag	tgaccgtgt	1080
aggccaaatca	ctttaggcaa	ctcgactatt	actacacaag	aagcagocaa	tgctatttgt	1140
gcttacggtg	aatggcccac	tatataaat	gattcagaag	ctaatccgg	agatgcaccc	1200
actgagccg	acgttagtag	caacccgtt	tacaccctag	aatcggtgtc	ttggaaagacc	1260
acttcagggg	gatgggtgt	gaaggttaca	gattgttga	aggacatgg	aatgtttgt	1320
cagaatatgt	actatcacta	cttggggcg	tctggttaca	ccattcatgt	ccagtgcac	1380
gettc当地	ttcacaagg	ggcggttag	gttttctgt	taccagagg	tgtcatggct	1440
tgaacactg	agagtaaaac	gtcatacgtt	tcatacatca	atgcaaatcc	tggtgagaga	1500
ggcggtgagt	ttacgaacac	ctacaatccg	tcaaatacag	acgccagtga	gggcagaaag	1560
tttgcagcat	ttggatttt	gctgggtt	gtgttcttag	caggaaacgc	ctttgtgtac	1620
ccgcacccaga	tcatacaact	acgtaccaac	aacagtgcac	caattgtgt	gccccatcta	1680
aactcacttg	tgattgattt	tatggcaaaa	cacaataact	ggggcattgt	catattacca	1740
ctggcaccc	ttgccttgc	cgcaacatcg	tcaccacagg	tgccattttac	agtgaccatt	1800
gcacccatgt	gtacagaatt	caatgggtt	agaaacatca	ccgtcccaat	acatcaagg	1860
tttgcgacaa	ttgaacaccc	tggttccaat	caattcctt	catctgtat	cttcagtcg	1920
cctctgtgc	tacctaattt	tgatgttact	ccaccaatac	acatacccc	ggaagtaaag	1980
aatatgtgg	aactagctga	aattgacaca	ttgateccaa	tgaacgcagt	ggacggaaag	2040
gtgaacacaa	ttggatgt	tcaataccaa	ttgaatgaca	atttgacca	ggcacccata	2100
ttctgtttat	ccctatcacc	tgcttctgt	aaacgactga	gcccacccat	gttgggtgaa	2160
atcttaattt	attacacca	ttggacgggg	tccatcgat	tcacccctt	attttgtgc	2220
agtatgtgg	ccactggtaa	actgtccct	agtatattcc	caccgggagc	taaaccacca	2280
accaatcgca	aggatgcata	gtctggcaca	cacatcatct	gggacccat	gttacaatcc	2340
agttgttcc	ttggtcacc	gtggatctc	aacacagt	acagacgg	tgcacgtgt	2400
gacttcactg	agggccgatt	tataatctg	ttctatcaaa	ctagaatttg	ggtagctgt	2460
tcaaccccta	ccagttatgtt	catgttagg	tttggtagt	ctgttccaga	cttcagtc	2520
agactgttta	gggacactcc	ccatattatg	caatcgaaac	taataggacg	tacacaagg	2580
attgaagacc	tcatttgcac	agcgatataa	aatgccttaa	gagtgttcca	accacccctg	2640
accccgatctaa	ctgaagcaac	tagtgggtt	aatagccagg	agggtccagc	tctaactgt	2700
gtggaaacac	gagcatctgg	tcaagcaatc	cccaagtatgt	ttgtggaaac	taggcacgt	2760
gtaaattaca	aaaccagg	tgaatctgt	ctttagtcat	tctttgggag	agctgcgt	2820
gtcacaatcc	tatcttgc	caacttccct	aagagccgg	aggagaaaaa	gcatttcaac	2880
atatggaaa	tttacatcac	cgacatgtc	cagttacgca	gaaaattaa	attttca	2940
tattccagg	tttagtctt	aatgacttt	gtattcac	agaactatcc	tagtacagcc	3000
agtggagaag	tgcgaaacca	gggttaccag	atcatgtata	ttccaccagg	ggcaccccg	3060
ccatcatct	gggtatgtact	cacatggca	tcctcttca	acccttccat	cttctacatg	3120
tatggaaatg	caccccttac	gtatgtca	ccttacgt	ggatttgc	ttccatttca	3180
cacttctacg	atggcttgc	acgggtgca	tttgggtt	agaacaccg	tgctggcgac	3240
acgttttacg	gttttagtgc	cataaatgt	tttgggtt	tagcagttag	agcagtaaac	3300
cgcgatatac	cacatacaat	acacatgt	gtttagtgc	acatggaaac	aaaacacat	3360
cggtgttgg	ccccccagg	tcctctgt	gtttagtata	ggggagagg	agtggacat	3420
atatccatgt	caatttctacc	tctggcaag	gttagactca	ttaccactt	ttgggttgg	3480
catcagaaca	aaggatgt	cggttccgg	tacaagatt	gcaactacca	cctagcaacc	3540
ccaaatgtac	acttgaatgc	aattatgt	ttatgggaca	gggatttat	gttgggtgaa	3600
tcttagagcc	ggggaaatgt	tccatcgcc	agatgttagt	ggagggttgg	agtttactat	3660
tgtgaatcta	ggggaaatgt	ctaccctgt	atctttatgt	ggcccaatgt	tcgattcatg	3720
gaagcaaaacg	actactatcc	agcaagatac	cagtctcaca	tgctgtatgg	tgccgtt	3780
gcagaaacccg	gggatgtccg	ttgggtactg	gggtgcact	atgggtat	ttgtatcat	3840
actgcaggag	gttgggggt	agtgcctt	gtgtacat	gagacccctg	gggtatgaa	3900
gaggaggcoca	ttggaaacggg	ataaaacago	tatcatcgat	ctctccgg	agcccttggc	3960
gcagggttca	ccccacacaa	cagtggaaa	gtgactgaat	tgacaacgt	ggttaccagg	4020
actatcacag	aaaactact	aaaaacttgc	gtgaaaatgt	tgtcggtct	agtgtattgt	4080
gtgagaaat	atggaggacac	taccacgt	tttgcacac	tagactact	cggggtgt	4140
atatcttcc	ggcaatgg	gaagaagaa	gcatgtact	tactagat	tcctttagtg	4200
atgcgccaag	gtgtgggt	gatgaagaaa	tttacagagg	cggtcaatgc	agctaaaggc	4260
tttagtggta	tttagacaaa	aatttccaa	ttttagatgt	gggtgaatgt	taaaattatc	4320
ccagacgtca	aggacaagg	ggaatttctc	accaatgtt	aacatgt	catgttggaa	4380
atatcaaatt	caacatccca	ccaaatctgc	cccaatgtt	aaacaacaa	gatttttc	4440
aacaatgt	gtatgttgc	agtccatgt	cgccgttgc	caccattata	cgctgtgg	4500
gcacccggaa	ttaaacaaa	gggatgtca	ataaaacaa	atatacagg	caagacaaa	4560
caccgttgg	aaccgtatgt	tatgttgc	catgggtc	cagggtgg	taaaatctata	4620
gtctacttcat	ttatgtgtat	agcaatgt	gagaaggaaa	gcacatgt	ctattcaat	4680
ccacccgtacc	catctactt	tgatgttgc	aaacaacaa	gggtgtgt	tatggacac	4740
ctaaacccaa	accccgatgg	tatggatgt	aaactgttt	gccaatgtt	atcaacatgt	4800
gagtttattt	ctccaaatggc	ctcatttag	ttttgttac	atctgtat	4860	
gtctctgg	cttaccaact	tcattcaat	gttacccaca	cagtggctca	cagtgtatgc	4920
ttaaccacag	gatgttgc	ttatgttgc	gtttacacga	tgctgttca	ttcgttcaaa	4980
ggcoaaactga	atatggccac	ggccactca	ttgtgttgc	attgttcaac	acctgtcaat	5040
tttaaaaatgt	gttggccct	cggttgc	aaggccgt	ggatgtttaa	cggttactgc	5100
agacaaaatgt	tcactgt	tgatgttac	acatataatca	tgaatgt	aaacagaagg	5160
gccaatatacg	gcaatttgc	ggaaggctt	tttcaaggac	cactaagg	taaaggattt	5220
aagatcgat	tgaagacgt	ttccccccct	gagtgcata	gtgattttt	acaaggatgt	5280
gattctcaag	agggttgg	ttactgttgc	aagaaagg	ggatgtttaa	cggttactgc	5340
cagattcaac	tagaaaggaa	catcaatagg	gccatgacta	tacttcaac	tggttaccaca	5400
ttcgcaggcag	tcgcaggagt	agtgtatgt	atgtacaaa	tcttcgg	tcaacagggt	5460
gcatacactg	gttgcacaa	caaaaaaccc	aatgtccct	ctatcaga	cgctaaatgt	5520
caggggccac	gatgttgc	cgcaatgt	atggcaaaaa	gaaacatgt	tactgcac	5580

-continued

---

accaccaagg	gtgaattttac	catgttaggg	gtgcatgata	atgttagcaat	attgccaacc	5640
catgccgcotc	caggagaaac	cattattatt	gatggaaaag	aagttagagat	ccttagacgccc	5700
agagcccttag	aagatcaagc	gggaaccaat	cttgagatca	ccattattac	tctaaaaaga	5760
aatggagaagt	ttagagacat	cagatcacat	atccccacc	aaattactga	aactaacat	5820
ggagtgttga	tcgtgaacac	tagcaagtac	cccaatatgt	atgtccccgt	tggtgctgt	5880
accgaacagg	gatatcttaa	tctcagtgg	cgtcaaaactt	ctcgcaactt	aatgtacaac	5940
tttccaacaa	gggcaggcca	gtgcgggaga	atcatcactt	gtactggcaa	agtcatggg	6000
atgcatgttgc	ggggaaacgg	ttcacatggg	tttgcgtcag	ccctcaacgg	atcataactt	6060
actcaaaatc	agggcgaaat	ccagtggtat	aggtcatca	aagaagtggg	gtaccccatt	6120
ataaatgcoc	catccaaagac	aaagttagaa	cccagtgtt	tccactatgt	ttttqaaggt	6180
gttaaggaaac	cagctgtat	cactaactt	gaccggacag	taaaaaacaga	ttttqaagaa	6240
gccccatctttt	ctaaatatgt	ggggaaacaaa	attacttgaag	ttggacgagta	catgaaagaa	6300
gcagtggatc	actatgcagg	acagttaatg	tcactggata	tcaacacaga	acagatgtgc	6360
ctggaggatc	ccatgtacgg	taccgtatgt	cttgaggccc	ttggatcttag	cactagtgt	6420
ggatatactt	atgttgcatt	ggggaaagaa	aaaagagaca	ttcttagataa	acagaccaga	6480
gataactaaagg	agatcgacag	acttttagat	acctatggaa	tcaatctacc	attagtca	6540
taatgttgc	cactcagccc	tgcttggttt	gaagcttta	aatgtgtt	agaaaaaatt	6840
ggataggcga	gttagagtga	ctatatacgt	tacctgaacc	actctcacca	ccctttacaaa	6900
aacaagactt	atttgttca	aggcgccat	ccatccggct	getctggcac	ctcaattttc	6960
aactcaatgt	ttaacaaacct	gatcattagg	acgtctttac	tgagaaactt	caagggcata	7020
gacttggacc	ataaaaaaat	gattgcctat	ggtgatgacg	tgatagottc	ctacccccat	7080
gaggttgcacg	ctagttccat	agccaaatca	ggaaaagact	atggactaac	catgactcca	7140
gcagataaaat	cagcaacccctt	tgaaaaacgtt	acatggggaa	atgttaacatt	tctgaaaaaga	7200
tttttcagat	cagatgagaa	gtatccatc	ctggtgcatc	cagtatgtcc	aatgaaagaa	7260
attcacaatgt	caatcagatg	gccaaggac	cctagaaaaac	cacaggatca	cgtacgc	7320
tttgtccat	tagcttggca	caacggtaa	gaagaataca	ataaaattttt	agctaaaatc	7380
agaagtgtgc	caatcggaa	agctttttat	ctcccagat	actctacatt	gtacccgcga	7440
tggtctcgact	catttttagt	acccctaccc	agtccgatgg	gattgggtt	tactgttga	7500
ggggtaaattt	tttctttaat	tcggagaaaa	aaaaaaaaaa	aaaaagactt	cccaatca	7560
agtgaattcg	cgccgcctg	caggtcgacc	atatgggaga	gtcccaacg	cggtggatgc	7620
atagcttgat	tattctatag	tgtcacatca	atagcttgcc	gtatcatgg	tcatagctt	7680
ttccctgtgt	aaatttgcatt	ccgcctacaa	ttccacacaa	catacgagcc	ggaagcataa	7740
agtgtaaaacg	ctgggggtgccc	taatgagtga	gctaactcac	attaatttgc	ttgcgctcac	7800
tgccccgttt	ccagtcggga	aacctgtcgt	gcccagctca	ttaatgaatc	ggccaacgcg	7860
cggggagagg	cggtttcgct	atttggcgct	cttcccgctt	ctcgctca	gactcgctgc	7920
gctcgctgt	tcggctcgcc	cgagcgat	cagtcactt	aaaggcggtt	atacggttat	7980
ccacagaaatc	aggggataac	gcaggaaaga	acatgtgacg	aaaaggccag	caaaaggcca	8040
ggaaccgtaa	aaaggccgcg	ttgctggcgt	tttccatag	gctccgcccc	cctgacgagc	8100
atcacaaaaat	tcgacgctca	agtccagat	ggcgaaaccc	gacaggacta	taaaqatacc	8160
aggcgtttcc	cccttggaa	tcctctgt	gtcttcctgt	tcggacccct	ccgcttaccg	8220
gataccgttc	ccgccttctc	ccttcggaa	cgctggcgct	ttctcatagc	tcacgctgt	8280
ggtatcttag	ttcgggtgt	gtcgttcgt	ccaaagctggg	ttgtgtgcac	gaaccccccg	8340
ttcagcccgaa	ccgcgtcgcc	ttatccgtt	actatgtct	tgatccaaac	ccggtaagac	8400
acgacttatac	cccaacttgc	cgacggactt	gtaacaggat	tagcagagc	aggatgttag	8460
gegggtgtac	agagtcttc	aagtgggtt	ctaaactacgg	ctacactaga	agaacatgt	8520
tttgttatctg	cgctctgt	aagccagtt	ccttcggaa	aagagtgtt	agctcttgat	8580
cccgccaaacaa	aaccacccgt	ggtagccgtt	gtttttttgt	ttgcaagcag	cagattacgc	8640
gcagaaaaaa	aggatctaa	gaagatctt	tgatctttt	tacgggggtt	tgacgttcgt	8700
ggaacgaaaa	ttcacgttta	gggtttttgg	tcatagat	atcaaaaaaa	atcttcac	8760
agatcctttt	aaattaaaaaa	tgaattttt	aatcaatcta	aagtataat	gagtaaactt	8820
ggtgtctgc	tttacatgtc	ttatctgt	aggcacccat	ctcagcgatc	tgtcttattt	8880
gttcatccat	agttgcctga	ctccccgtc	tgttagataa	tacatgtcc	gagggtttac	8940
catctggccc	cagtgtcga	atgataccgc	gagacccacg	ctcaccggct	ccagatttt	9000
cagcaataaa	ccagccagcc	ggaaggccgc	agcgacaa	tggtcttgc	actttatccg	9060
cctccatcca	gtcttataat	ttgttccggg	aagcttagat	aagtgttcg	ccagttataa	9120
gtttgcgc	cggtttcgcc	atttgcatac	gcatctgtt	gtcagcgtcg	tcgtttgtt	9180
tggcttattt	cagtcgttgc	ttccaaacgt	caaggcgat	tatcatgtcc	cccatgtt	9240
gcaaaaaaa	ggtttagctt	ttcgggttct	cgatcggtt	cagaagtaa	ttggccgcag	9300
tgttatctact	catgggtat	gcagacttgc	ataatttctt	tactgtat	ccatccgtaa	9360
gtatgttttc	tgtgtactgtt	gagttacttca	ccaaatgtt	ctgagaatag	tgtatgcggc	9420
gaccggatgt	cttttccccc	cgcttataac	gggataatac	ccgcgcacat	agcagaactt	9480
taaaatgtct	cattatgg	aaacgtttt	cgggcgaa	actctca	aggatcc	9540
tgttgatgtc	cagttcgat	taacccactt	gtgcacccaa	ctgatcttca	gcatcttta	9600
cttcaccagg	cgtttctggg	tgagaaaaaa	caggaaggca	aaatggccca	aaaaagggaa	9660
taaaggccgc	acggaaat	tgaatactt	tactcttcc	ttttcaat	tattgaagca	9720
tttatcagg	tttattgtct	atagcgat	acatattttg	aatgtatttt	aaaaataaaac	9780
aaatagggtt	tccgcgcaca	tttccccaa	aagtgcacc	tgtatgcgtt	tgaaaatcc	9840
cacagatgc	taaggagaaa	ataccgcato	aggaaatttt	aagcgttta	attttgtt	9900
aatttcgcgtt	aaatttttgt	taaatcagct	cattttttaa	ccaataggcc	gaaatcgca	9960
aaatccctta	taatcaaaaa	gaatagacc	agatagggtt	gagtgtt	ccagtttgg	10020
acaagagtc	actattaaag	aacgtggat	ccaaacgtca	agggcgaaa	accgtctatc	10080
aggcgatgg	cccaactatc	ctaatca	ttttttttgg	tcgagggt	tgcc	10140

---

-continued

---

gtaaagact	aaatcggaac	cctaaaggga	gcccccgatt	tagagttga	cggggaaagc	10200
cggcgaacgt	ggcgagaaag	gaagggaaga	aagcgaaaagg	agcggggcgt	agggcgcgtgg	10260
caagtgtacg	ggtcacgctg	cgcgtAACCA	ccacaccggc	cgcgttaat	gcgcgcgtac	10320
aggcgcgctc	cattcgccat	tcaggctcg	caactgttg	gaagggcgat	cggtgccggc	10380
ctcttcgcta	ttacgcgcagc	tggcgaaagg	gggatgtgt	gcaaggcgat	taagtgggt	10440
aacgccagg	ttttcccagt	cacgacgttg	taaaacgcag	gccagtgat	tgtaatacga	10500
ctcaactata	ggcgaatttg	gcccgcgtc	gcatgcgtcc	ggccgcgtat	gcggccgcgg	10560
gaattcgatt	gaggcatgct	aatatcgact	actatagg			10598

---

SEQ ID NO: 9	moltype = DNA	length = 10561			
FEATURE	Location/Qualifiers				
source	1..10561				
	mol_type = other DNA				
	organism = synthetic construct				
SEQUENCE: 9					
ttaaacacgc tctggggttt	ttcccacccc	agaggcccac	gtggcggtca	gtactcttgt	60
attacggtag ctttgtacgc	ctgtttgtt	tccttcccc	cgtaacttta	gaagcttata	120
aaaagttcaa tagcagggtt	acaaaccagt	acctctacgc	acaagcaccc	ctgtttcccc	180
gggtgatata catagactgt	acccacgggt	aaaagtgtat	gatccgttat	ccgcgttgagt	240
acttcgagaa gctctgtatc	accttggat	cttcgtatcg	ttgcgttcaa	cactctggcc	300
cgagtgtagc tttaggtgtat	gagttttggc	actccccacc	ggcgacgggt	gcccaggctg	360
cgttggccgc taatccatgg	ctgtatcggt	gggacgtat	ttgtgaacaa	ggtgtgaaga	420
gcctatttgat	ctactcaaga	gtcctccggc	ccctgaatgc	ggctaatctt	480
caaccgtctca acacccatgt	atgtttgtt	ctgtatcgat	aagtctgtgg	cggaacccgac	540
tactttgggt gtccgtgttt	ccctttatat	tcataactggc	tgctttatgtt	gacaattttac	600
aaattgttac catatagcta	ttggatttgg	cacccatgtat	tgtgcaatat	atttgagttt	660
ttctttatata agccatata	acatcacat	tttaatcaca	ataaaacagt	caaattggggg	720
ctcaagtttc aacgaaaaag	acgggtgcgc	acgagaatca	aaacgtggca	gccaatggat	780
ccaccattaa ttacactact	atcaactt	acaaagacag	tgcgagtaat	tccgtacta	840
gacaagacct ctcccaagat	ccatcaaaaat	tcacagaaacc	ggtaaggac	ttaatgttga	900
aaacacgacc agctctaaac	tcgcttacaa	tggaagcat	tgggtacagt	gaccgtgtga	960
ggcaaaatcc tttagggac	tcgtatata	tcacacaaga	agcagccat	gttatttttgc	1020
cttacggtag atggccact	tacataaaatg	attcagaac	taatccgtt	gatgcaccca	1080
ctgagccaga cggttagtagc	aaccgggttt	acaccctaga	atcggtgtt	tggagacca	1140
cttcaaggggg atgggtgtgg	aagttagcc	atgtttgtt	ggacatggg	atgtttggc	1200
agaatatgtt	ctatcaact	ttggggcgct	ctggttacac	cattcatgtc	1260
cttcaaaaattt tcaccaaggg	gctgttaggg	ttttctgtat	accagatgtt	gtcatggct	1320
gcaacactga gagtttttttt	tcatacgttt	catacatca	tgcaatctt	ggtgagagag	1380
ggcgtgagg	tcaccaaccc	tacaatccgt	caaatacaga	cgccagtgag	1440
ttgcagccat	ggatttttttgc	ctgggttctgt	gtgttcttgc	aggaaacggc	1500
cgaccaggat	catcaactca	cgtagcaaca	aattgtgttgc	ccatcatgtt	1560
actcaactgtt gattgtttt	atggcaaaac	acaataactt	gggcattttgc	atattaccac	1620
tggcacccctt	ggcccttgc	gcaacatcg	caccacatgt	gtgaccatttgc	1680
cacccatgtg tacagaatcc	atagggttgc	gaaacatcac	cgccccatgt	catcaagggt	1740
tgccgacaaat	gacacacac	gggttcaat	aattcccttac	atctgtatgc	1800
cctgtgcctt acctaattttt	gatgttactt	caccaataca	cataccggg	gaagtaaga	1860
atatgttgcg actatgtttt	atttgcacat	tgatcccaat	gacgcgtat	gacgggaagg	1920
tgaacacaaat	ggagatgtat	ttttatccat	tgaatgacaa	tttgagcaag	1980
tctgtttatc	cctatccat	gtcttgcata	aacgactgtt	ccacccatgt	2040
tcctaaatata	ttacacccat	tggacgggg	ccatcagggtt	cacccatca	2100
gtatgttgc	cactgttttt	ctgttgcata	gttattttcc	accggggact	2160
ccaaatcgca	ggatgttgc	ctaggcacac	acatcatctg	ggacatgggg	2220
gttggccat	gttgcaccc	tggatcttca	acacatgttgc	cgacgtgttgc	2280
acttcactga	ggccggattt	ataacttgc	tctatcaaaat	tagaattttgttgc	2340
caacccctac	cgtatgttgc	atgttaggtt	ttgttagtgc	gtgtccagac	2400
gactgttttag	ggacactccc	catattttgtc	aatcgaaactt	aataggact	2460
ttaaagactt	cattgttacca	tggtatgttgc	atgccttaaa	ccacccatgttgc	2520
cccaatgttac	tgaagactt	atgttttttttgc	atagccaggat	ggtgcacgtt	2580
tggaaacagg	ggcatatgttgc	caagatgttgc	ccgtgtatgtt	gggtggaaact	2640
taaattttatata	aaccagggttgc	gaatgttgc	tttgcgttgc	gttgcgtgttgc	2700
tcacaaatctt	atcccttgc	aaactcttca	agacggggaa	ggagaaaaaa	2760
tatggatata	tacatatacc	gacactgttgc	agttacgtttgc	aaaatttttttttgc	2820
attccaggtt	tgatcttgc	atgacttttttgc	tatttcacaga	gaactatctt	2880
gttggagaaat	ggcgaaaccc	gttgcgttgc	tcacccatgttgc	ggaccccccgc	2940
catcatatcg	ggatgttgc	acatggcaat	ccttttcaat	cccttccat	3000
atggaaatgc	acctccaccc	atgttgc	cttacgttgc	gatgttgc	3060
acttctacg	tggcttgc	cggttgc	tttgcgttgc	gaacacatgt	3120
cgttttatcg	tttagtgc	ataaaatgtt	tttgcgttgc	gcttgcgaca	3180
gcagtaatcc	acataatcc	cacacatgttgc	ttagatgttgc	catgaaacca	3240
ggttgggttgc	ccccagaccc	cctcgatgttgc	tattataccat	ggggaggggg	3300
tatccatgtc	aatttccat	ctggccaaagg	ttactgtat	taccacttttgc	3360
atcagaacaa	agcgtgttgc	tttgcgttgc	acaagatgttgc	caactaccac	3420
caagtgtatca	cttgcgttgc	attatgttgc	tatggagac	ggatgttgc	3480
ctagagcccg	gggaaacttgc	accatcgca	gttgcgttgc	gttactt	3540
gttgcgttgc	ggggaaatgt	accatcgca	gttgcgttgc	catttgc	3600
aagcaaaacgc	ctactatcc	gcaagatacc	gttgcgttgc	tgccggat	3660

-continued

cagaaccggg	ggactgcggt	gggatactga	ggtgtcaactca	tggggtaatt	ggtatcatta	3720
ctgcaggagg	tgaagggtt	gtagcccttg	ctgacattag	agacctctgg	gtgtatgaag	3780
aggggccat	ggAACAGGGA	ataacaagct	acatcaac	tctcgccaca	gcccTTGGCG	3840
cagggttcac	ccacaaacat	agtggaaa	tgatgttt	tgatgttt	gacaacatgt	3900
ctatcacaga	aaaactacty	aaaaacttgg	tgaaaatagt	gtcggtctt	gtgttgggtt	3960
tgagaaatta	tgaggacact	accacatcc	ttgcaacact	agcactactc	gggtgtgtata	4020
tatctcttgc	gcaatgttg	aagaagaagg	catgtgactt	actagagat	ccttatagtgt	4080
tgcccaagg	tgatgggtt	atagaaata	tcacagggc	gtgcata	gctaagggtt	4140
tagatgtgt	tagcaacaa	atttccaaat	ttatagattt	tttggatgt	aaaaattatcc	4200
cagacgtca	ggacaagggt	gaatttctca	ccaaatgttta	acagcttagac	atgttggaaa	4260
atcaaattgc	aaccatcccc	caatcttgc	ccacggcaaga	acaacaaagag	attctttca	4320
acaatgtgt	atgggttag	gtcccgatcc	gtccgttgc	accattatac	gtctgggg	4380
cacggcaat	taacaaatgt	gagagcacaat	taacaaatata	tatatacttc	aaagacaaac	4440
accgtattga	accaggatgt	atgtcattt	atgggttcc	agggacgggt	aaatctata	4500
ctacttcat	aataggtaga	gcaatagcag	agaaggaaag	ccccatcgt	tattcaatgc	4560
cacccgttcc	atctcaattt	gtatgtttt	aaacaacagg	ggtagtgatt	atggacggcc	4620
taaaaaaaa	ccccggatgt	atggacatgt	aactgtttt	ccaaatgtt	tcacaaatgtt	4680
agtttattcc	tccaaatggcc	tcattagagg	agaagggtt	tttggtttaca	tctgtattat	4740
tcctgggttc	taccaactt	cattcaattt	taccacccac	agtggcttac	agtgtatgcct	4800
taaacacagg	atttgttca	gtatgtgggg	tttacaaatgt	gtgttgcacat	tcacaaatgtt	4860
gcaaaatgg	ttatggccac	gcccacttac	tgtgttgc	tttgcacaa	cttcggaaat	4920
ttaaaaaatgt	ttgccttc	gtttgtggaa	aggccttgc	attaaatggac	aggttacacca	4980
gacaaagggt	cactgttagt	gagattacca	cattaatcat	gaatgagaaa	aacagaagggt	5040
ccaatatccg	caatgttgc	gaagccctgt	tttcaagggcc	actaaatgtt	aaagatgtt	5100
agatcgatgt	gaagacagg	cccccccttc	atgtgcata	tgattttta	caacgcgtt	5160
attctcaaga	ggtttaggtt	tactgtgaga	agaagggtt	gtatgtttt	tttgcacaaat	5220
agattcaact	agaaaaggaa	atcaataggg	ccatgactat	actccaaatgt	gttaccacat	5280
tcgcagcgt	cgcaggatgt	gtgtatgtt	tgtacaaatct	cttcggccgt	caacagggtt	5340
catacaactg	cttgcacaaat	aaaaacccaa	atgtccctat	tatcagatgc	gtcaaaatgtt	5400
agggggccagg	atttgtactt	gcagtggca	tttgcacaaat	aaacatgtt	atgcacca	5460
ccaccaagg	tgaatttacc	atgttgggg	tgcata	tgttgcacata	tttgcacaaat	5520
atgcgcgttc	aggagaaatcc	atttatttt	atggggaaatgt	atgttgcacat	tttgcacaaat	5580
gaggccttgc	agatcaagg	ggaaacaaatc	tttgcacat	catttactt	ctaaaaaaatgt	5640
atagaaatgt	tagagacat	agatcacata	tttccacccat	tttacttgc	actaaatgtt	5700
gagttgttgc	cgtgtacact	agcaagtttac	ccaaatatgtt	tttgcacaaat	tttgcacaaat	5760
ccgaacacgg	atataatccat	ctcagtggac	gtcaaaatgtt	tttgcacaaat	tttgcacaaat	5820
ttccaaacaa	gggcaggccat	tgccggggaa	tcatcaatgtt	tactgttgc	tttgcacaaat	5880
tcatgtttgc	cgggggacatgt	tcatgttttgc	tttgcacaaat	tttgcacaaat	tttgcacaaat	5940
ctcaaaatca	ggggggacat	cagtgtatgt	gggttgcata	aaatgttgc	tacccat	6000
taaatggccc	atccaaagaca	aagttagaaat	ccagtgtttt	ccactatgtt	tttgcacaaat	6060
ttaaaatggcc	actgtgtactt	actaaatggat	tttgcacaaat	tttgcacaaat	tttgcacaaat	6120
ccatctttc	taataatgttgc	ggggacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	6180
cagtgttgc	ctatcgagg	cagttaatgt	cactgttgc	caacacacaa	cagatgttgc	6240
tggaggatgc	catgtacgtt	accgtatgtt	tttgcacaaat	tttgcacaaat	tttgcacaaat	6300
gatattctta	ttgttgcata	ggggaaaatgt	aaatgttgc	tttgcacaaat	tttgcacaaat	6360
atactaaatgg	gatgtcagaga	cttttagata	cctatgttgc	caatcttacca	tttgcacaaat	6420
acgttgcagg	tgaatccat	tcaaaatgtt	aaatgttgc	tttgcacaaat	tttgcacaaat	6480
aagcttccat	ccttaatgtt	tcagtgc	tttgcacaaat	tttgcacaaat	tttgcacaaat	6540
ctttccacaa	gaatccatgt	gttgcacat	tttgcacaaat	tttgcacaaat	tttgcacaaat	6600
tttgggatgt	gataccatgt	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	6660
atgtgcgttc	acttgcctt	gtttgttttgc	aaatgttgc	tttgcacaaat	tttgcacaaat	6720
gatttggcag	taggttgc	ttatataatgt	tttgcacaaat	tttgcacaaat	tttgcacaaat	6780
acaagactt	tttgttgc	ggccggatgt	catccggct	tttgcacaaat	tttgcacaaat	6840
actcaatgtt	taacaaatcc	atcattatgtt	tttgcacaaat	tttgcacaaat	tttgcacaaat	6900
atttggacca	ttttaaaatgt	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	6960
aggttgcac	tagtgcctt	ggccaaatgtt	ggaaatgtt	tttgcacaaat	tttgcacaaat	7020
cagataaaatc	agcaacatcc	gaaacatgtt	catggggaaat	tttgcacaaat	tttgcacaaat	7080
ttttcaatgc	agatgagaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	7140
tttcaatgc	aatcagatgtt	accaggatcc	tttgcacaaat	tttgcacaaat	tttgcacaaat	7200
tgtgttgcatt	atgttgcac	aaatgttgc	tttgcacaaat	tttgcacaaat	tttgcacaaat	7260
gaagtgtgc	aatcggaaat	gttttattgt	tttgcacaaat	tttgcacaaat	tttgcacaaat	7320
ggctcgactt	attttagata	ggcgcgcgc	acatgttgc	tttgcacaaat	tttgcacaaat	7380
acacacttcc	tttacatcc	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	7440
tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	7500
aaaaaaaaaa	gctcccaatc	actatgttgc	tttgcacaaat	tttgcacaaat	tttgcacaaat	7560
agagcttcc	acgcgttgc	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	7620
ggcgttaatca	ttgttgcata	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	7680
caacatataca	ggccggaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	7740
cacattatgt	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	7800
gttgcataat	atccggccat	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	7860
tttgcacaaat	acttgcactt	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	7920
tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	7980
tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	8040
taggttgc	ccccctgc	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	8100
cccgacagg	ctataaaatgt	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	8160
tgttgcac	ctgcgcgtt	tttgcacaaat	tttgcacaaat	tttgcacaaat	tttgcacaaat	8220

-continued

---

gctttctcat	agtcacgct	gttaggtatct	cagttcggtg	taggtcggtc	gctccaagct	8280
gggctgtgt	cacgacccc	ccgttccagcc	cgaccgcgtc	gccttatccg	gtaaactatcg	8340
tcttgatgtcc	aacccggtaa	gacacgactt	atcgccactg	gcagcagcga	ctggtaacag	8400
gattagcaga	gogaggatgt	taggggtgc	tacagagttc	ttgaagtgtt	ggcctaacta	8460
cggtctacact	agaagaacag	tatgggtat	ctgcgctctg	ctgaagccag	ttaccccg	8520
aaaaagagtt	ggtagtctt	gatccggcaa	acaaaccacc	gctggtagcg	gtggttttt	8580
tgttgcag	cacgagatta	cgcgcagaaa	aaaaggatct	caagaagatc	cttgatctt	8640
tcttacgggg	tctgacgctc	agtggaaacga	aaactcactg	taaggattt	ttgtcatgag	8700
attatcaaaa	aggatcttca	cctagatct	tttaaattaa	aatgaagtt	ttaatcaat	8760
ctaaatgtata	tatgatcaa	cttgcgtct	cagttaacca	tgcttaatca	gtgaggcacc	8820
tatctcagcg	tatgtatctat	ttctgtatct	catagttgc	tgactcccc	tcgtgtatag	8880
aactacgata	cgggagggtc	taccatctgg	cccccaagt	gcaatgatc	cgcgagaccc	8940
acgctcaccc	gctccagatt	tatcagcaat	aaaccagcca	gccggaaggg	ccgagcgcag	9000
aagtgggtct	gcaactttat	ccgcctccat	ccagttat	aatttgtcc	ggaaagctag	9060
agtaagtagt	tcgcgttta	atagtttgc	caacgttgc	gccccatcg	9120	
gggtgtcacgc	tcgtcggtt	gtatggctt	attcagttcc	ggttcccaac	gataaaggcg	9180
agttatcatgt	tccccatgt	tgtgcaaaaa	agcggttagc	tccttcggtc	ctccgatctg	9240
tgtcagaatgt	aatgtggccg	cagtgttac	actcatgtt	atggcagcac	tgccataattc	9300
tcttactgtc	atgcctatcc	tgatgttct	ttctgtatgc	gggtgtatct	caaccaagtc	9360
attctcgagaa	tagtgtatgc	ggcgcacgg	ttgcgttgc	ccggcgatca	tacggatata	9420
taccgcgc	catagcagaa	ctttaaaagt	gctcatatt	ggaaaacgtt	cttccggcg	9480
aaaactctca	aggatcttac	cgctgttgc	atcgttgc	atgtaaacca	ctcgatcacc	9540
caactgtatc	ttagcatctt	ttactttcac	cagcgttcc	gggtgagcaa	aaacaggaaag	9600
gcaaaaatgcc	gaaaaaaaaagg	gataaaggdc	gacacggaaa	tgttgaatac	tcatacttt	9660
ccttttca	tattattgaa	gcattttca	gggttattgt	ctcatgagcg	gatacatatt	9720
tgaatgtatt	tagaaaaata	aacaaatagg	ggttccgc	acattttccc	gaaaagtgc	9780
acctgatgc	gtgtgaaata	ccgcacatgt	gctgttgc	atcgttgc	atcggaaat	9840
tgttaagcggtt	aatattttgt	taaaatgtc	gttaaatttt	tgtttaatca	gcttatttt	9900
taaccaatag	gccccaaatcg	gcaaaatccc	tttaaatca	aaagaataga	ccgagatagg	9960
gtttaggtt	gttcaggat	ggaaacaagag	tccactatt	agaacgtgg	actccaaacgt	10020
caaaggccgca	aaacccgtct	atcggccgc	ttggccacta	cgtgaacat	caccctatac	10080
aagttttttgc	gggtcgagg	gctgttgc	actaaatcg	aaccctaaag	ggagcccccg	10140
atttagact	tgacggggaa	agccggcga	ctggcgaga	aaggaaaggga	agaaacgcga	10200
aggagcggc	gtttagggcgc	tggcaagtgt	agcggtcacg	ctgcgcgtaa	ccaccacacc	10260
cgcgcgcgtt	aatgcggccgc	tacaggccgc	gtccatgc	cattcaatgc	ggcaactgt	10320
tggaaaggcc	gatcggtcg	ggcccttc	cttacacgc	agtcgttgc	aggggatgt	10380
gctgcaaggc	gatataatgtt	ggtaacgc	gggttttcc	agtcacgc	ttgttaaacag	10440
acggccatgt	aattgtata	cgactacta	tagggcga	ttggccgcac	gtcgcatgt	10500
cccgccgc	atggccgc	cggttac	atggaggat	gctaatacga	ctcaactata	10560
g						10561

---

SEQ ID NO: 10	moltype = DNA	length = 10613				
FEATURE	Location/Qualifiers					
source	1..10613					
	mol_type = other DNA					
	organism = synthetic construct					
SEQUENCE: 10						
ttaaaacagc	tctgggttgc	ttcccaccc	agaggccac	gtggccgta	gtactctgg	60
attacggtagc	ctttgtacgc	ctgtttgt	tcccttcccc	cgtaacttta	gaagtttac	120
aaaagttcaa	tagcagggtt	acaaaaccgt	acctctacga	acaaggactt	ctgtttcccc	180
ggtgatgtca	catatgtact	acccacggtc	aaaatgtt	gatccgttat	ccgttgcgt	240
acttcgtatc	gccttagtac	accttggat	ttcgtatgc	ttgcgttca	cactctgc	300
cgagtgtac	ttaggtgtat	gatgttggc	actccccacc	ggcgacgggt	ggccaggctg	360
cgttggggc	ctacccatgg	ctgtatgcgt	gggcgttag	ttgtgaacaa	gggtgtgaaga	420
gcctattgt	ctactcaaga	gtccctccgc	ccctgtatgc	ggctaatct	aaccacggag	480
caaccgtatca	aaacccgt	agtaggtgt	cgtaatgcgt	aatgttgc	cggaaccggac	540
tactttgtt	gtccgtgtt	ccctttat	tcatactgc	tgcttatgtt	gacaatttac	600
aaattgttac	catatgtca	ttggattgc	cacccgtat	tgtgcaat	atttgatgt	660
ttttttata	acccttata	acatcacat	ttaatcaca	ataaaacgt	caaatgggg	720
ctcaaggat	aaacccaaag	accgggtgc	acgagaatca	aaacgtggc	gccaatggat	780
ccaccattaa	tttactactat	accaatctt	aaacaaagac	tgccgtat	tcgcgtacta	840
gacaagac	ctcccaagat	ccatcaaaat	tcacagaacc	ggtttggac	ttaatgttga	900
aaacacgacc	agctctaaac	tcgcctaaacg	ttggaaatgt	ttgggttacgt	gaccgtgt	960
ggcaaaatcac	tttaggttgc	tcgttactt	ctacacaaga	agcggccat	gttattgtt	1020
cttacgggt	atggccact	tatataatgt	attcgttgc	ttatccgtt	gatgttgc	1080
ctgagccaga	cggttagtc	aaacggttt	acacccttgc	atccgttgc	ttggaaatgt	1140
cttcaagggg	atgggtgttgc	aaatgttac	ttgttgc	ttgttgc	atgttgc	1200
agaatatgt	ctatctact	ttggggcgc	ctgttacac	cattcatgtc	cagtgcac	1260
cttcaaaaat	tcaccaagg	cggttaggt	ttttccgtat	accaggat	gtctatggct	1320
gcaacactga	gatggaaaac	tcatacgtt	catacatca	tgcaatct	gggtgagag	1380
gcgggtgat	tacgaacacc	tacaatccgt	caaatacga	cgccgttgc	ggcagaaatgt	1440
ttgcagcatt	ggattatgt	ctgggttgc	gtgttgc	aggaaacgc	tttgcgttac	1500
cgcaccatgt	cataccaaca	acgtgcac	aattgttgc	ccatcatgt	catcaagggt	1560
actcacttgt	gatggattgt	atggcaaaac	acaataactg	gggcattgtc	atattacac	1620
tggcaccctt	ggcccttgc	gcaacatgt	caccacgt	gccttataca	gtgaccatt	1680
cacccatgt	tacagaattc	aatgggttga	gaaacatc	cggtcccgat	catcaagggt	1740

-continued

tcggccgacaaat gaacacacactt ggttccaatc aattcccttac atctgtatgc ttccagtcg 1800  
 cctgtgcctt acctaatttt gatgttactc caccataca catacccgaa gaagtaaaga 1860  
 atatgtatggaa actatgtgaa attgacacat tgatcccaat gaacgcgtg gacggaaagg 1920  
 tgaacacaat ggagatgtat caaatccat tgaatgacaa ttggacgaag gcacttat 1980  
 tctgttatac cctatcacct gcttcgtata aacgactgtg ccacacatg ttgggtgaaa 2040  
 tcttaaatta ttacacccat tggacggggt ccatcagggtt caccttcta ttttgtgca 2100  
 gtatgtggc cactgtgaaa ctgtccctca gctattcccc accggaggtt aaaccaccaa 2160  
 ccaatcgcaaa ggatgtcaat ctggacacac acatcatgtg ggacatgggg ttacaatcc 2220  
 gtgttccat ggttgacccat tggatctcaaa acacagtgtg caacggctgt gacgtgtat 2280  
 acttcaactg gggcgattt ataacttgc tctatcaaa tagaaatgtg gtacctgtt 2340  
 caacccctac cagtagtgc atgttaggt ttgttagtgc ttgtccagac ttcaactgtca 2400  
 gactgttggg gggacactcc catattagtc atacgaaact aataggacgtt acacaaaggc 2460  
 ttaaagacact cattgacaca gcatgtaaaatg atgccttaag atgtttccaa ccacccctca 2520  
 cccagtcac tgaagcaact agtgaggatg atagccaggaa ggtgcagct ctaactgtgt 2580  
 tggaaacagg agcatctgtt caagaaatcc ccagtgtat gttggaaaact aggcaacgtgg 2640  
 taaattacaa aaccagggtt gatgtgc ttgagttttt ctttggaga gctgtgtgt 2700  
 tcacaatccat atccctgacc aactcttca aagccggaga ggaaaaaaatg catttcaaca 2760  
 tatggaaat tatacataccca gacactgtcc agttacccggg aaattatgaa tttttcactgt 2820  
 attcagggtt tgatcttgc atgactttt tattcacaga gaactatctt agtacagcc 2880  
 gtggagaatggc gggaaacccat ggttacccaa tctatgtat tccacccagg gaccccccgg 2940  
 catatcttgc ggtatgtactt acatggcaat cctttcaaa ccctttccat ttctatcatgt 3000  
 atggaaatgc acctccaccc atgttcaatc cttacgttgg gattttccat gccttccat 3060  
 acttctacg tggcttgcg cgggtccac ttgagggtga gaacaccgtt gctggcgaca 3120  
 cgttgcgttcc tttagtgcg ataaatgtt ttggagggtt agcaggatggaa gcaatggaaacc 3180  
 gcaatgttcc acatatacaata cacaatctg tggatgttgc catggaaacca aaacccatc 3240  
 ggttgcgttcc ccccaacactt cctcgactgtt tattatcagg gggagggggg gtggacatgt 3300  
 tatccagtgc aattctactt ctggccaagg tagactcaat taccactttt gggtttggc 3360  
 atcagaacaa agcgtgtac gttccgggtt acaaggatggg caacttaccac ctggaaaccc 3420  
 caagtgtatcc cttgtatgcg atttagatgtt tatggggacag ggattttatgg tgggttggaa 3480  
 cttagccccc gggaaatgtt accatccggca gatgtgttgc cttgttggaa ttactatt 3540  
 gtgtatcttag gggggatgtt accctgtca cttttactgg cccaaacgttt cgattcatgg 3600  
 aagcaaaacgca ctactatcca gcaagatacc agtctccat gctgtatggg tgcggatgg 3660  
 cagaacccggg ggactgtcggtt gggatgtca ggtgcactca tggggtaat ggtatcatca 3720  
 ctggagggtt gtaaggggta tggcccttgc ctgacatggg agacccctgg gtgtatgaag 3780  
 aggaggccat ggaacaggaa ataaacaaatg acatccatc ttcggccaca gccttggcgg 3840  
 cagggttccat ccacacaatc agtggaaaatg tgactgaatt gacaacgttg gttaccggca 3900  
 ctatcacaga aaaactctgg tggaaatagt gtcggctca tggatgttgc 3960  
 tgagaaat tggggacactt accacgatcc tggcaacactt agcactactt ggggttgc 4020  
 tatacttgc tggatgttgc aagaaggaaatg catgtgttgc acttagagat cttatgtca 4080  
 tgcgccaagg tgatgggtt atggaaaat tcaacagggc gtgcgttgc gctaaagggt 4140  
 tagatgtgtt tagaccaactt atttcaactt ttatagttt gttggatgtt aaaatattcc 4200  
 cagacgttccat ggacaagggtt gaaatcttca ccaatgttgc agacgttgcg atgttggaa 4260  
 atcaatattgc aaccatccac caattttgc ccaatggcaaga acacaaaggatcc 4320  
 acaatgttgcg atgggttgc gtcggatccc gtcgggttgc accattatc gctgtggagg 4380  
 cacccggaaat taacaaaatg gagacccatca taaaacatca tatacgttcc aagaccaac 4440  
 acggatgttgc accatgttgc atgtccatc atgggttcc accgggggggaaatattatgg 4500  
 ctacttccat aataggatca gcaatggatcc agaaggaaatg cacaatgttgc ttttccatgc 4560  
 cacctgaccc atcttactt gatggctata aacaacaaagg ggttagtgcg atggacgacc 4620  
 taaaccaaaa ccccgatgtt atggacatgtt aactgttttgc ccaatgttgc tcaacatgtt 4680  
 agtttatttc tccaaatggcc tccatggggtgg agaagggtt tttgttccatc tctgttggat 4740  
 tcttgggttcc taccacactt ctttcaatgggg ttttccatgc agtggcttgc agtgtatgcct 4800  
 taaccacagc atttgcattt gatgtggggg ttttccatgc gttgttgc ttttccatgc 4860  
 gcaactgttcaat gtttccatgc ttttccatgc gttgttgc ttttccatgc 4920  
 ttttccatgc ttttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 4980  
 gacaaagggtt cactgtatgtt gatgttgc ttttccatgc gttgttgc ttttccatgc 5040  
 ccaatatccg ctttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5100  
 agatcgatgtt gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5160  
 atttcttcaaaa ggtttaggtt tactgttgc gtttccatgc gttgttgc ttttccatgc 5220  
 agatgttgcg atggaaaggaaatc atcaatggggtccatc gtttccatgc gttgttgc ttttccatgc 5280  
 tccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5340  
 ctttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5400  
 agggggccaggg attttgcattt gatgtggggg ttttccatgc gttgttgc ttttccatgc 5460  
 ccaccaacgggg ttttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5520  
 attttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5580  
 attttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5640  
 attttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5700  
 gtttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5760  
 ctttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5820  
 ttttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5880  
 ttttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 5940  
 ttttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 6000  
 ttttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 6060  
 ttttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 6120  
 ttttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 6180  
 ttttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 6240  
 ttttccatgc gtttccatgc gttgttgc ttttccatgc gttgttgc ttttccatgc 6300

-continued

---

gatatcctta	tgttgcata	ggaaaaaaga	aaagagacat	tctagataaa	cagaccagag	6360
atactaagg	gatcgacaga	cttttagata	cctatggaaat	caatctaca	ttatgtcacgt	6420
acgtgaaaga	tgaactcagg	tcaaagacta	aagtggaaac	aggaaagtca	agattgattg	6480
aaagcttcagg	cctaatagtat	tcagttgcac	tgagaatggc	ctttggcaat	ctttacgcag	6540
cttccacaa	gaatccagg	gtggtgacag	gatcagcagt	tggttgtgac	ccagattgt	6600
tttggagtaa	gataccagg	ctaattggaa	aaaaactctt	cgcttttgc	tacacagggt	6660
atgtatgc	actcagg	gettgggtt	aagctcttaa	aatgggtgtt	aaaaaaattt	6720
gatggcag	tagatgtac	tatataact	acctgaaacc	ctctcaccac	ctttacaaa	6780
acaagactt	ttgtgtcaaa	ggccgcatac	catccggctt	ctctggacc	tcaatttca	6840
actcaatgt	taacaacatg	atcattggaa	cgcttttact	gagaacccat	aaggccatag	6900
acttggacca	ttaaaaaatg	atgcctatg	gtgtgacgt	gatagtttcc	taccccatt	6960
aggttgacgc	tagtctctt	gcccatacg	gaaaagacta	tggactaacc	atgactccag	7020
cagataaaatc	agcaaccc	gaaacagtca	catgggagaa	tgtacattt	ctgaaaagat	7080
tttcagacgc	agatgagaac	tatccattt	ttgtgcatacc	agtgtatcca	atggaaagaa	7140
ttcacgaaatc	aatcagatgg	accaggacc	ctggaaacac	acaggatcac	gtacgcctgt	7200
tgtgcctatt	agcttgcac	aacgttga	aaqaatacaa	taaatttttca	gtttaaaaat	7260
gaagtgtgc	aatcggaa	gcttattgc	tcccagagta	ctctacatg	tacccgcgt	7320
ggctcgactt	attttagtaa	ggcgccgc	acagctgtt	gaaggggacc	aacgatacag	7380
ctgggtttag	ggggaccaat	ggaggccacat	acttccattt	atccatcac	ccacacactt	7440
ccttacattt	cacggcgcgc	cgacttatt	tagttaacc	acctcagtcg	gattggattt	7500
ggttatactg	ttgttaggg	aaattttct	ttaatttgcg	gaaaaaaaaa	aaaaaaaaaa	7560
gagctcccaa	tcaactgt	atcgccgc	gcttgcagg	cgaccatatg	ggagagctcc	7620
caacgcgtt	gatgcata	tttgcattt	tatagtgt	cttataatagc	tttgcgtt	7680
catggtcata	gtgtttctt	gttgcattt	gttgcattt	cacaatcca	cacaatata	7740
gagccggaa	cataaaatgt	aaagoctgg	gtgcctaa	agttagttaa	cttcacattaa	7800
ttgcgttgc	tcaactgt	gctttccat	cgggaaacct	gtcggtcc	ctgcatttaat	7860
gaatcgocca	acgcgggggg	agaggcggtt	tgcgttatttgc	gcttctctgc	7920	
tcactgact	gtgcgtct	gtcgttgcgc	tggtatcgat	cactcaaaagg	7980	
cggttaatc	gttatccata	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaag	8040
gccagcaaaa	ggccaggaa	cgtaaaaaagg	ccgcgttgc	ggcgcccc	cataggctcc	8100
gccccccatc	cgaggcat	aaaaatcgac	gtcaactgt	gagggtggcga	aaccggacag	8160
gactataaaat	ataccaggcg	tttccccctg	gaaagcttcc	cggtgcgtt	cctgttccga	8220
ccttgcgcgt	taccggatac	ctgttgcgtt	tttcccccttgc	gggaagctg	gcttcttctc	8280
atagctc	atcggttat	ctcagttcg	tgttaggtcg	tcgttccaa	ctgggtgtt	8340
tgacgcgaa	ccccgttca	cccgaccgtt	gccccttata	cggttaactat	ctgtttgtt	8400
ccaaacccgtt	aaagacacgac	ttatcgccat	ttgcgcgc	cactgttgc	aggatttagca	8460
gagcgcgggt	tgttaggcgt	gttacagat	tcttgcatttgc	gtggccaa	tacggatata	8520
ctagaagaac	agtatttgc	atctgcgtc	tgctgaacc	agtttaccc	ggaaaaaagag	8580
tttgcgttgc	tttgcgttgc	aaacaaaccc	ccgcgttgc	cggtgttttt	tttgcgttgc	8640
agcagcagat	tacgcgcaga	aaaaaaaggat	cttcaagaga	tccctttgttgc	ttttctacgg	8700
ggtctgacgc	tcaactgt	gaaaacttac	gtttaaggat	tttgcgttgc	agattatcaa	8760
aaaggatctt	cacccatgt	ttttttaaat	aaaaatggaa	ttttttaaatca	atctaaatgt	8820
tatatgtat	acttgcgtt	gacatgttcc	aatgttcaat	cgatggacca	cctatctca	8880
cgatctgtt	attcgttca	tccatgttcc	ctcgacttcc	cgatgttgc	ttactacga	8940
taaggggagg	tttccatct	ggccccatgt	ctgcaatgt	accgcgcag	ccacgcgttcc	9000
cggttccaga	tttacatgt	ataaaaccgc	cagccggaa	ggccgcgc	agaagtgtt	9060
ctgcaactt	atccgcgtt	atccgttca	tttattgttgc	ccggaaact	agatgttacta	9120
tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9180
gttcgtcggtt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9240
gtatccccat	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9300
gttaatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9360
tcatgcgtt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9420
aatatgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9480
cacatagcg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9540
caaggatctt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9600
cttcgcgtt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9660
ccgcgggggg	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9720
aatattatttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9780
tttagaaaa	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9840
cggtgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9900
taataatatttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	9960
aggccggaaat	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10020
tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10080
aaaaaaacgtt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10140
tgggtgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10200
cttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10260
gchgcttgcgtt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10320
ttaatgcgtt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10380
gcgatcggtt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10440
gcgatcggtt	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10500
tgaattgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10560
ccatggcgcc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	tttgcgttgc	10613

SEQ ID NO: 11                  moltype = RNA    length = 21  
 FEATURE                        Location/Qualifiers  
 source                        1..21

---

-continued

---

```

mol_type = other RNA
organism = synthetic construct
SEQUENCE: 11
tggaatgtaa agaagtatgt a 21

SEQ ID NO: 12      moltype = RNA length = 22
FEATURE          Location/Qualifiers
source           1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 12
tagactg tgataactga ag 22

SEQ ID NO: 13      moltype = RNA length = 23
FEATURE          Location/Qualifiers
source           1..23
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 13
tggagtgtga caatggtgtt tgt 23

SEQ ID NO: 14      moltype = RNA length = 22
FEATURE          Location/Qualifiers
source           1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 14
ttaaggcacg cggtaatgc ca 22

SEQ ID NO: 15      moltype = RNA length = 23
FEATURE          Location/Qualifiers
source           1..23
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 15
tccctgagac ccttaacct gtg 23

SEQ ID NO: 16      moltype = RNA length = 21
FEATURE          Location/Qualifiers
source           1..21
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 16
tcgtaccgtg agtaataatg c 21

SEQ ID NO: 17      moltype = RNA length = 22
FEATURE          Location/Qualifiers
source           1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 17
tcggatccgt ctgagcttgg ct 22

SEQ ID NO: 18      moltype = RNA length = 22
FEATURE          Location/Qualifiers
source           1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 18
tcacagtcaa ccggtcttt tc 22

SEQ ID NO: 19      moltype = RNA length = 22
FEATURE          Location/Qualifiers
source           1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 19
cagtgcattt taaaaaggcg at 22

SEQ ID NO: 20      moltype = RNA length = 22
FEATURE          Location/Qualifiers
source           1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 20
taacagtcta cagccatggc cg 22

```

---

-continued

---

```

SEQ ID NO: 21      moltype = RNA  length = 22
FEATURE
source
1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 21
tttgtccctc tcaaccagct gt                                22

SEQ ID NO: 22      moltype = RNA  length = 21
FEATURE
source
1..21
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 22
tgtgacttgt tgaccaggagg g                                21

SEQ ID NO: 23      moltype = RNA  length = 23
FEATURE
source
1..23
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 23
tatggctttt tattcctatg tga                                23

SEQ ID NO: 24      moltype = RNA  length = 17
FEATURE
source
1..17
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 24
agtgggtttt gtgaatc                                     17

SEQ ID NO: 25      moltype = RNA  length = 20
FEATURE
source
1..20
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 25
cataaaagtag aaaggactac                                20

SEQ ID NO: 26      moltype = RNA  length = 23
FEATURE
source
1..23
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 26
tgttagtgttt cctactttat gga                             23

SEQ ID NO: 27      moltype = RNA  length = 22
FEATURE
source
1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 27
tgagatgaag cactgttagct ca                                22

SEQ ID NO: 28      moltype = RNA  length = 24
FEATURE
source
1..24
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 28
gtccagttt cccaggaatc cctt                                24

SEQ ID NO: 29      moltype = RNA  length = 22
FEATURE
source
1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 29
tcagtgcact acagaaccttt gt                                22

SEQ ID NO: 30      moltype = RNA  length = 22
FEATURE
source
1..22

```

---

-continued

---

	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 30	
tagcagcaca taatggtttg tg	22
SEQ ID NO: 31	moltype = RNA length = 22
FEATURE	Location/Qualifiers
source	1..22
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 31	
tctcccaacc cttgttaccag tg	22
SEQ ID NO: 32	moltype = RNA length = 22
FEATURE	Location/Qualifiers
source	1..22
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 32	
actagactga agctccttga gg	22
SEQ ID NO: 33	moltype = RNA length = 22
FEATURE	Location/Qualifiers
source	1..22
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 33	
tcaagtgcatac acagaacctg gg	22
SEQ ID NO: 34	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 34	
ttgcatagtc aaaaaagtga	20
SEQ ID NO: 35	moltype = RNA length = 22
FEATURE	Location/Qualifiers
source	1..22
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 35	
ttaatgctaa tcgtatagg gg	22
SEQ ID NO: 36	moltype = RNA length = 22
FEATURE	Location/Qualifiers
source	1..22
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 36	
tagcagcacg taaatattgg cg	22
SEQ ID NO: 37	moltype = RNA length = 24
FEATURE	Location/Qualifiers
source	1..24
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 37	
caaagtgcattt acagtgcagg tagt	24
SEQ ID NO: 38	moltype = RNA length = 23
FEATURE	Location/Qualifiers
source	1..23
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 38	
aacattcaac gctgtcggtg agt	23
SEQ ID NO: 39	moltype = RNA length = 23
FEATURE	Location/Qualifiers
source	1..23
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 39	
tatggcactg gtagaattca ctg	23

---

-continued

---

```

SEQ ID NO: 40      moltype = RNA  length = 22
FEATURE
source
1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 40
taaggtgcat ctagtgcaga ta                                22

SEQ ID NO: 41      moltype = RNA  length = 22
FEATURE
source
1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 41
taaggtgcat ctagtgcagt ta                                22

SEQ ID NO: 42      moltype = RNA  length = 21
FEATURE
source
1..21
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 42
ctgacctatg aattgacagc c                                21

SEQ ID NO: 43      moltype = RNA  length = 22
FEATURE
source
1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 43
tgtAACAGCA ACTCCATGTG GA                                22

SEQ ID NO: 44      moltype = RNA  length = 21
FEATURE
source
1..21
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 44
tagcagcaca gaaatattgg c                                21

SEQ ID NO: 45      moltype = RNA  length = 23
FEATURE
source
1..23
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 45
cccaagtgttc agactaacctg ttc                                23

SEQ ID NO: 46      moltype = RNA  length = 23
FEATURE
source
1..23
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 46
tgtgcaaatac tatgaaaaac tga                                23

SEQ ID NO: 47      moltype = RNA  length = 23
FEATURE
source
1..23
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 47
tgtgcaaatac catgaaaaac tga                                23

SEQ ID NO: 48      moltype = RNA  length = 22
FEATURE
source
1..22
mol_type = other RNA
organism = synthetic construct
SEQUENCE: 48
ttccctttgt catcctatgc ct                                22

SEQ ID NO: 49      moltype = RNA  length = 22
FEATURE
source
1..22

```

---

-continued

---

	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 49	
ttcccttgt catcctatgc ct	22
SEQ ID NO: 50	moltype = RNA length = 22
FEATURE	Location/Qualifiers
source	1..22
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 50	
tggaatgtaa ggaagtgtgt gg	22
SEQ ID NO: 51	moltype = RNA length = 22
FEATURE	Location/Qualifiers
source	1..22
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 51	
ataagacgag caaaaagctt gt	22
SEQ ID NO: 52	moltype = RNA length = 21
FEATURE	Location/Qualifiers
source	1..21
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 52	
taaacgtctc cagtcacggc c	21
SEQ ID NO: 53	moltype = RNA length = 21
FEATURE	Location/Qualifiers
source	1..21
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 53	
atgacctatg aattgacaga c	21
SEQ ID NO: 54	moltype = RNA length = 21
FEATURE	Location/Qualifiers
source	1..21
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 54	
atgacctatg aattgacaga c	21
SEQ ID NO: 55	moltype = RNA length = 21
FEATURE	Location/Qualifiers
source	1..21
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 55	
taatctcagc tggcaactgt g	21
SEQ ID NO: 56	moltype = RNA length = 20
FEATURE	Location/Qualifiers
source	1..20
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 56	
tgattgtcca aacgcaattc	20
SEQ ID NO: 57	moltype = RNA length = 23
FEATURE	Location/Qualifiers
source	1..23
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 57	
agctacattg tctgctgggt ttc	23
SEQ ID NO: 58	moltype = RNA length = 24
FEATURE	Location/Qualifiers
source	1..24
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 58	
agctacatct ggctactggg tctc	24

---

-continued

---

SEQ ID NO: 59 moltype = RNA length = 21  
FEATURE Location/Qualifiers  
source 1..21  
mol\_type = other RNA  
organism = synthetic construct  
SEQUENCE: 59  
tgtcagttt tcaaatacc c 21

SEQ ID NO: 60 moltype = RNA length = 22  
FEATURE Location/Qualifiers  
source 1..22  
mol\_type = other RNA  
organism = synthetic construct  
SEQUENCE: 60  
tggctcagg cagcaggaac ag 22

SEQ ID NO: 61 moltype = RNA length = 22  
FEATURE Location/Qualifiers  
source 1..22  
mol\_type = other RNA  
organism = synthetic construct  
SEQUENCE: 61  
cattgcactt gtctcggtct ga 22

SEQ ID NO: 62 moltype = RNA length = 22  
FEATURE Location/Qualifiers  
source 1..22  
mol\_type = other RNA  
organism = synthetic construct  
SEQUENCE: 62  
tgtaaacatc ctacactcag ct 22

SEQ ID NO: 63 moltype = RNA length = 23  
FEATURE Location/Qualifiers  
source 1..23  
mol\_type = other RNA  
organism = synthetic construct  
SEQUENCE: 63  
tgtaaacatc ctacactctc agc 23

SEQ ID NO: 64 moltype = RNA length = 21  
FEATURE Location/Qualifiers  
source 1..21  
mol\_type = other RNA  
organism = synthetic construct  
SEQUENCE: 64  
tattgcacat tactaagttg c 21

SEQ ID NO: 65 moltype = RNA length = 22  
FEATURE Location/Qualifiers  
source 1..22  
mol\_type = other RNA  
organism = synthetic construct  
SEQUENCE: 65  
tttgttcgtt cggctcgctg ga 22

SEQ ID NO: 66 moltype = RNA length = 22  
FEATURE Location/Qualifiers  
source 1..22  
mol\_type = other RNA  
organism = synthetic construct  
SEQUENCE: 66  
tggaaagacta gtgattttgt tg 22

SEQ ID NO: 67 moltype = RNA length = 23  
FEATURE Location/Qualifiers  
source 1..23  
mol\_type = other RNA  
organism = synthetic construct

-continued

---

SEQUENCE: 67	
tctttggta tctagctgta tga	23
SEQ ID NO: 68	moltype = RNA length = 22
FEATURE	Location/Qualifiers
source	1..22
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 68	
ttcAACGGGT atttattttag ca	22
SEQ ID NO: 69	moltype = RNA length = 22
FEATURE	Location/Qualifiers
source	1..22
	mol_type = other RNA
	organism = synthetic construct
SEQUENCE: 69	
caccCGtaga accgaccttg cg	22

---

What is claimed is:

1. A nucleic acid construct comprising an infectious nucleic acid comprising a picornavirus genome comprising one or more heterologous sequence elements of 20 or more bases, wherein the specific infectivity of said construct is sufficient to initiate a spreading picornavirus infection when administered to a living mammal, wherein said specific infectivity of said construct is of similar magnitude to the specific infectivity of a comparable construct lacking said one or more heterologous sequence elements.
2. The construct of claim 1, wherein said mammal is a human.
3. The construct of claim 1, wherein said construct is formulated as plasmid DNA.
4. The construct of claim 1, wherein said construct is formulated as an RNA molecule.
5. The construct of claim 1, wherein at least one of said one or more heterologous sequence elements is a microRNA response element.
6. The construct of claim 5, wherein a microRNA target element of said microRNA response element comprises at least a region of complementarity to a microRNA present in non-cancer cells.
7. The construct of claim 1, wherein at least one of said one or more heterologous sequence elements is inserted into the 5' non-coding region of said picornavirus genome as a substitution for nucleotides within a scanning region.
8. The construct of claim 1, wherein said picornavirus genome comprises a microRNA target element for miR-133.
9. The construct of claim 1, wherein said picornavirus genome comprises more than one microRNA target element for miR-133.
10. The construct of claim 1, wherein said picornavirus genome comprises a microRNA target element for miR-206.
11. The construct of claim 1, wherein said picornavirus genome comprises more than one microRNA target element for miR-206.
12. A method of reducing the number of cancer cells within a living mammal, wherein said method comprises administering a construct to the mammal, wherein said construct comprises an infectious nucleic acid comprising a picornavirus genome comprising one or more heterologous sequence elements of 20 or more bases, wherein the specific infectivity of said construct is sufficient to initiate a spreading picornavirus infection when administered to a living

mammal, wherein said specific infectivity of said construct is of similar magnitude to the specific infectivity of a comparable construct lacking said one or more heterologous sequence elements.

13. The method of claim 12, wherein said mammal is a human.

14. The method of claim 12, wherein said cancer cells are melanoma, pancreatic, prostate, bladder, non-small cell lung, myeloma, or breast cancer cells.

15. The method of claim 12, wherein said administering step results in a reduced number of non-cancerous cells present within said mammal undergoing cell lysis following said administering step as compared to the number of non-cancerous cells that undergo lysis when said comparable construct is administered to a comparable mammal.

16. The method of claim 12, wherein said administering step results in a similar number or an increased number of cancerous cells present within said living mammal undergoing cell lysis following said administering step as compared to the number of cancerous cells that undergo lysis when said comparable construct is administered to a comparable mammal.

17. A method for making infectious RNA comprising a picornavirus genome comprising one or more heterologous sequence elements of 20 or more bases, wherein the specific infectivity of said infectious RNA is sufficient to initiate a spreading picornavirus infection when administered to a living mammal, wherein said specific infectivity of said infectious RNA is of similar magnitude to the specific infectivity of a comparable infectious RNA lacking said one or more heterologous sequence elements, wherein said method comprises:

(a) providing an DNA construct encoding said infectious RNA, wherein said DNA construct encodes a ribozyme, and wherein said nucleic acid construct comprises a restriction endonuclease cut site, and

(b) contacting said DNA construct with a restriction endonuclease under conditions wherein at least a portion of said DNA construct is removed, thereby producing a restriction endonuclease-cleaved DNA construct,

wherein said restriction endonuclease-cleaved DNA construct encodes an infectious RNA having non-picorna-virus RNA located at a 5' end, and wherein said

ribozyme cleaves said non-picornavirus RNA located at the 5' end to generate said infectious RNA.

**18.** The method of claim **17**, wherein said infectious RNA encoded by said restriction endonuclease-cleaved DNA construct comprises a 3' end with less than 10 ribonucleotides that are not present in a picornavirus genome.

**19.** The method of claim **17**, wherein said infectious RNA encoded by said restriction endonuclease-cleaved DNA construct comprises a 3' end with no ribonucleotides that are not present in a picornavirus genome.

**20.** The method of claim **17**, wherein said infectious RNA comprises a 5' end with no ribonucleotides that are not present in a picornavirus genome.

\* \* \* \* \*