US 20250263741A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2025/0263741 A1**

Krisky et al. (43) **Pub. Date: Aug. 21, 2025**

(54) **HSV VECTORS**

(71) Applicant: **PeriphaGen, Inc.**, Pittsburgh, PA (US)

(72) Inventors: **David M. Krisky**, Pittsburgh, PA (US); **James B. Wechuck**, Pittsburgh, PA (US)

(21) Appl. No.: **18/859,281**

(22) PCT Filed: **Apr. 26, 2022**

(86) PCT No.: **PCT/US2022/026323**

§ 371 (c)(1),
(2) Date: **Oct. 23, 2024**

Publication Classification

(51) **Int. Cl.**

C12N 15/86	(2006.01)
A61K 48/00	(2006.01)
C12Q 1/70	(2006.01)
G01N 33/569	(2006.01)

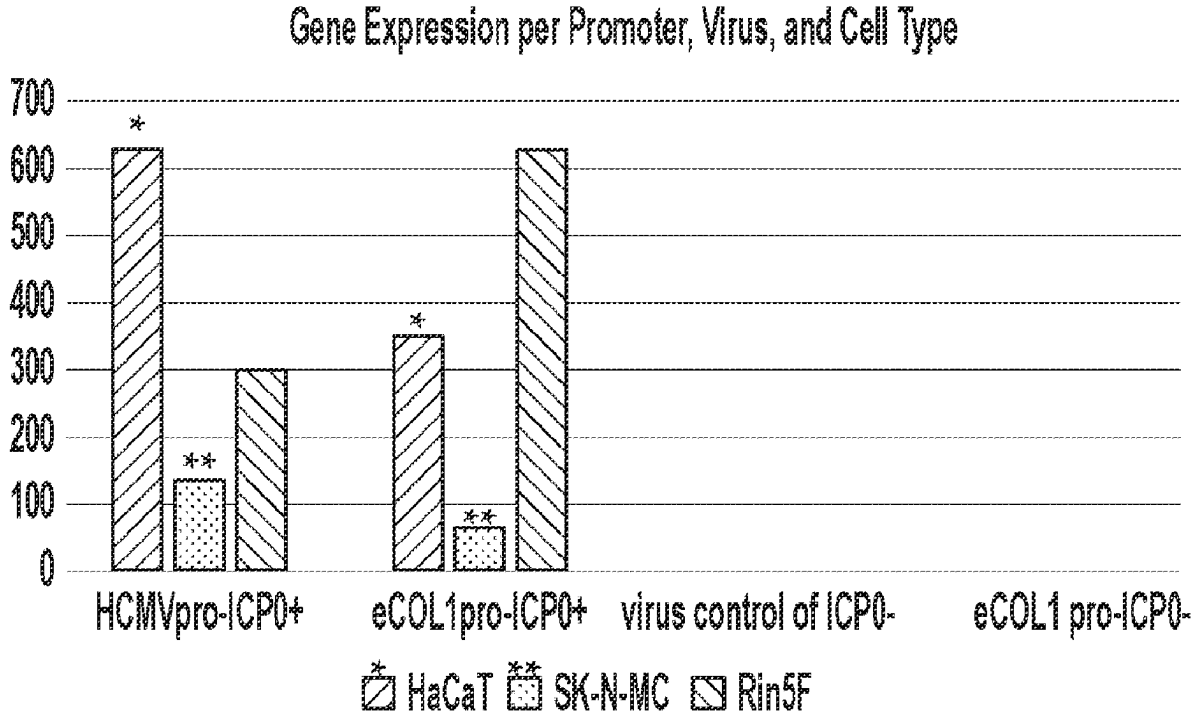
(52) **U.S. Cl.**

CPC ***C12N 15/86*** (2013.01); ***A61K 48/0058*** (2013.01); ***A61K 48/0075*** (2013.01); ***C12Q 1/705*** (2013.01); ***G01N 33/56994*** (2013.01); *C12N 2710/16643* (2013.01); *C12N 2710/16652* (2013.01); *C12N 2710/16671* (2013.01); *C12N 2830/008* (2013.01); *C12N 2830/50* (2013.01); *C12Q 2600/156* (2013.01); *G01N 2333/035* (2013.01)

(57) **ABSTRACT**

Disclosed herein are high transducing replication defective herpes simplex virus (HSV) vectors, and methods of using the same.

Specification includes a Sequence Listing.



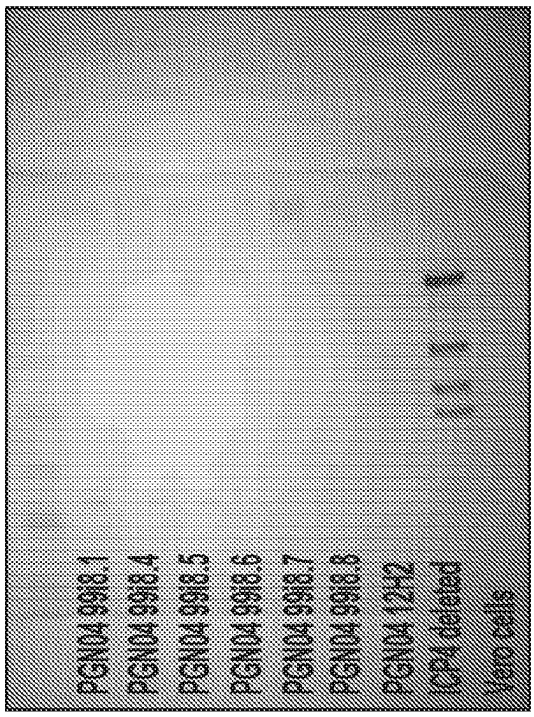

PGN04 9918.1
PGN04 9918.4
PGN04 9918.5
PGN04 9918.6
PGN04 9918.7
PGN04 9918.8
PGN04 1252
ICP4 deleted
Vero cells

FIG. 1

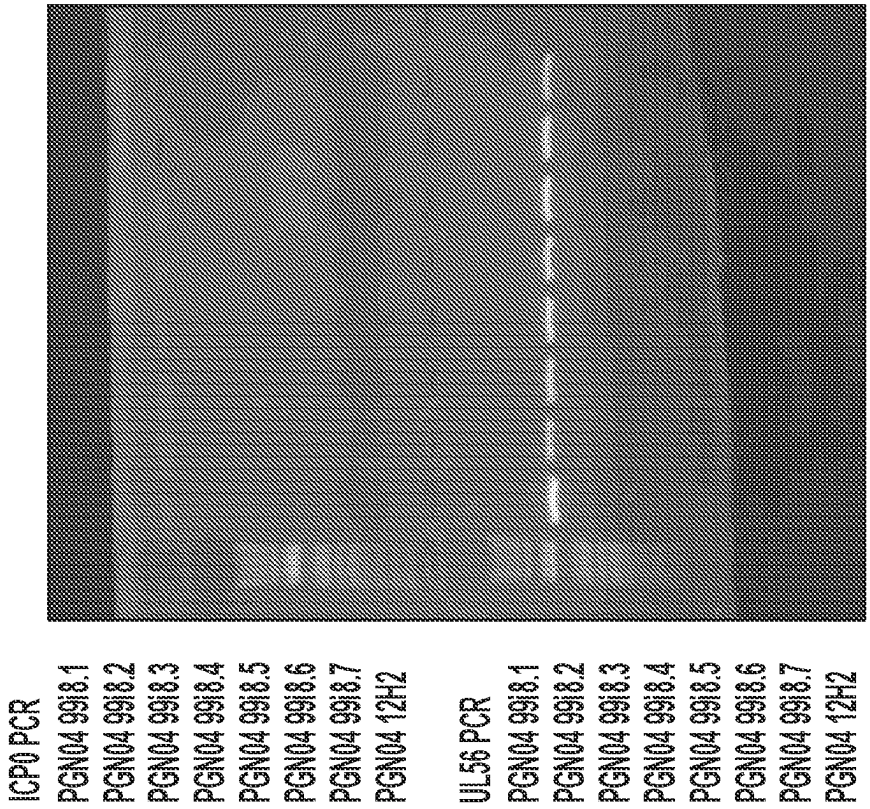

ICP0 PCR
PGN04 9918.1
PGN04 9918.2
PGN04 9918.3
PGN04 9918.4
PGN04 9918.5
PGN04 9918.6
PGN04 9918.7
PGN04 12H2
UL56 PCR
PGN04 9918.1
PGN04 9918.2
PGN04 9918.3
PGN04 9918.4
PGN04 9918.5
PGN04 9918.6
PGN04 9918.7
PGN04 12H2

FIG. 2A

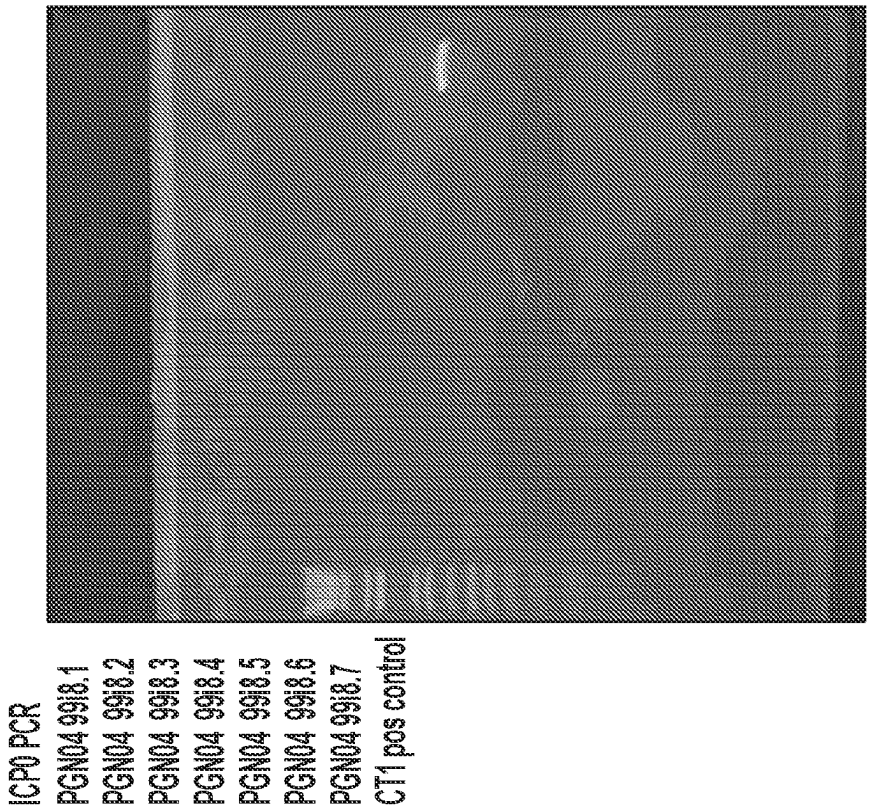

ICP0 PCR
PGN04 99I8.1
PGN04 99I8.2
PGN04 99I8.3
PGN04 99I8.4
PGN04 99I8.5
PGN04 99I8.6
PGN04 99I8.7
CT1 pos control

FIG. 2B

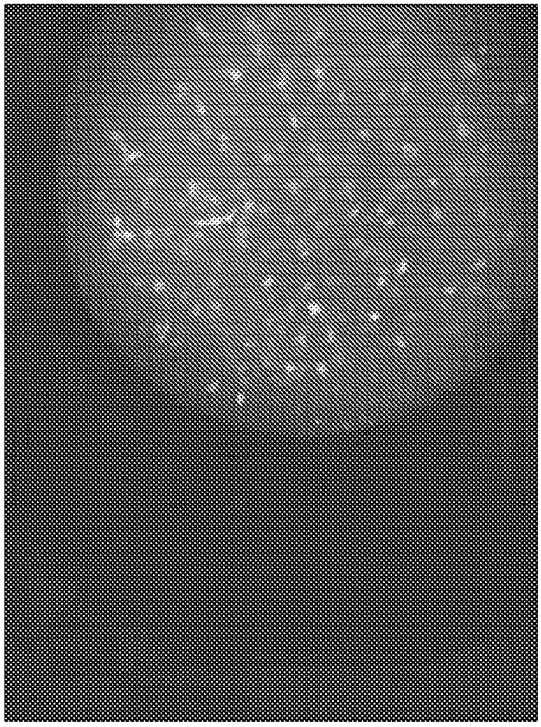


FIG. 3B

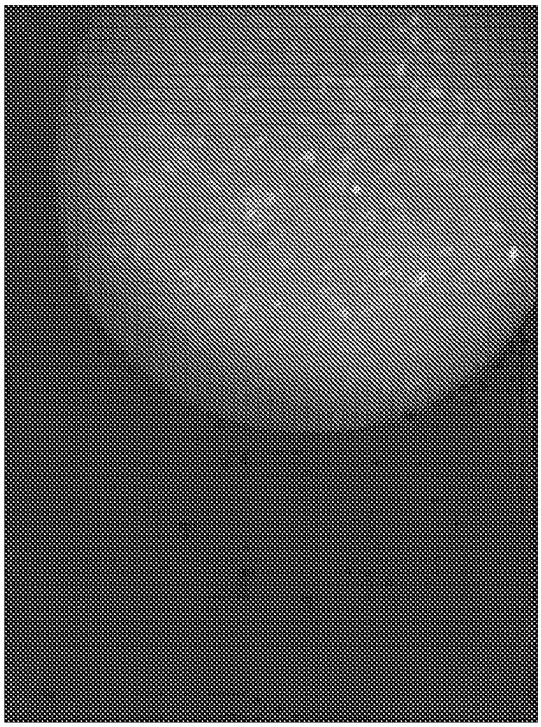


FIG. 3A

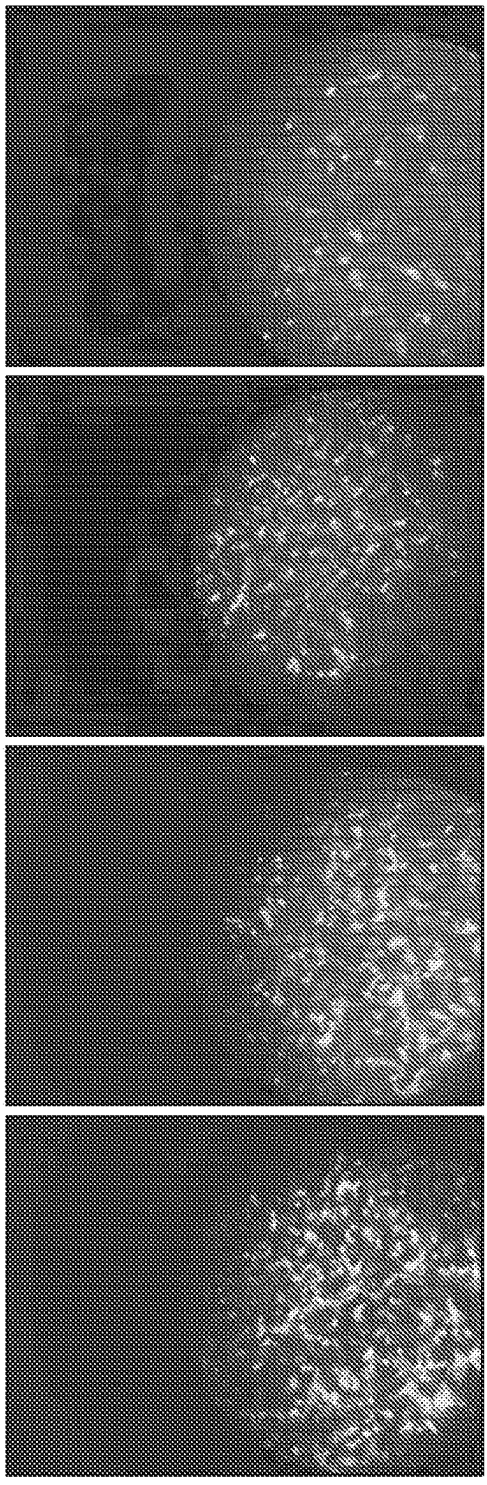


FIG. 4D

FIG. 4C

FIG. 4B

FIG. 4A

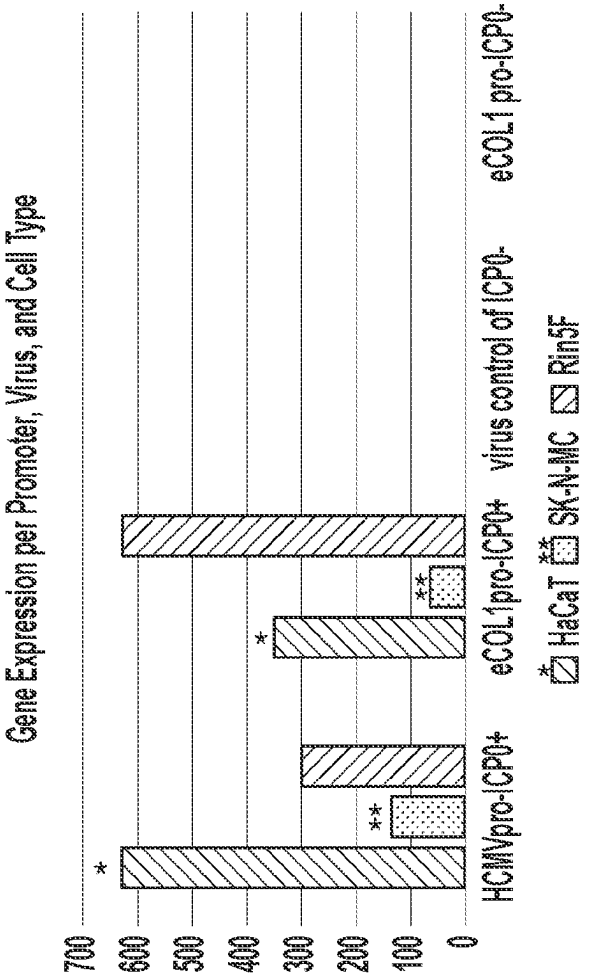

Gene Expression per Promoter, Virus, and Cell Type
700
600
500
400
300
200
100
0
*
**
*
**
HCMVpro-ICP0+
eCOL1pro-ICP0+
virus control of ICP0-
eCOL1 pro-ICP0-
* HaCaT ** SK-N-MC Rin5F

FIG. 5

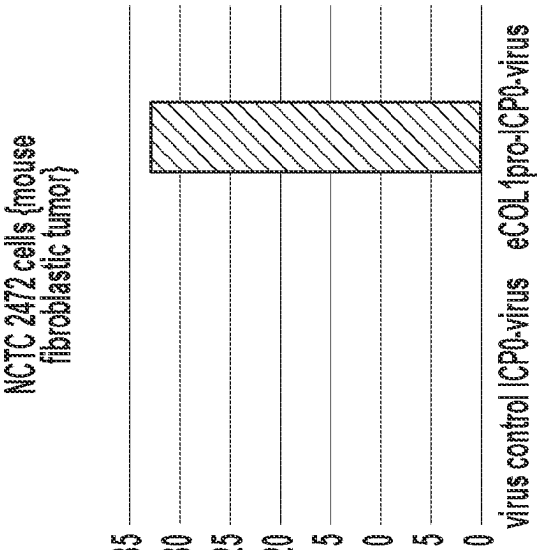

NCTC 2472 cells (mouse fibroblastic tumor)
35
30
25
20
15
10
5
0
virus control ICP0-virus
eCOL1pro-ICP0-virus

FIG. 6B

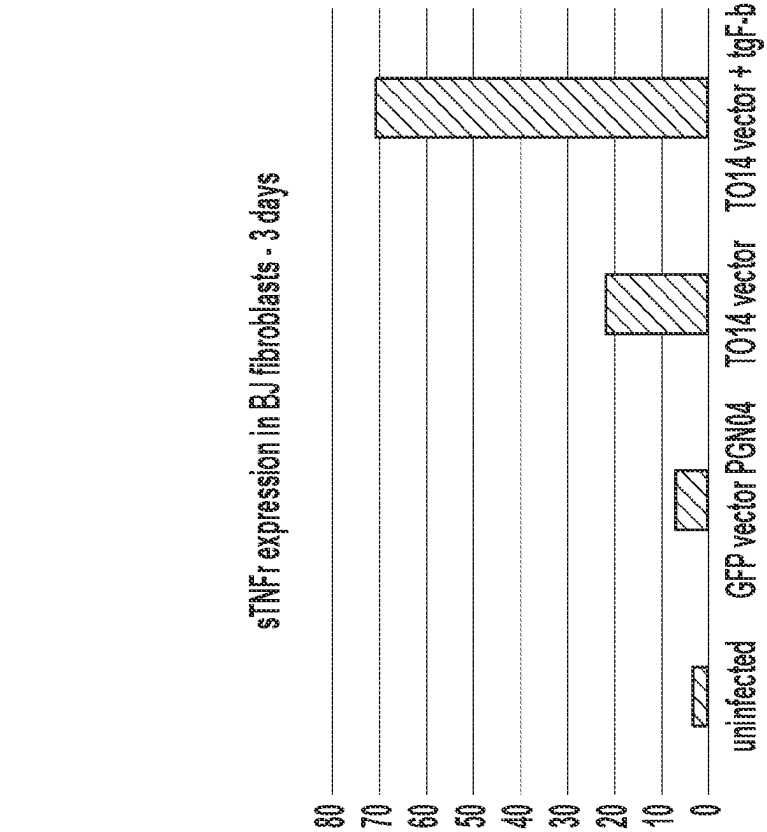

sTNFr expression in BJ fibroblasts - 3 days
80
70
60
50
40
30
20
10
0
uninfected
GFP vector PGN04
TO14 vector
TO14 vector + tgF-b

FIG. 6A

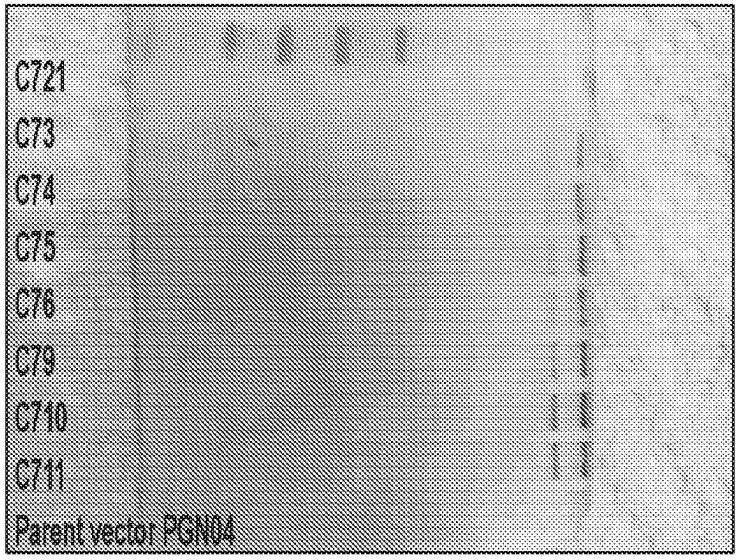


FIG. 7

HSV McKrae strain nucleotide sequence (SEQ ID NO:1)
Accession no. JQ730035.1

```
   1 gcagcccggg ccccccgcgc gcggggcggc gcgcaaaaaa ggcgggcggc ggtccgggcg
  61 gcgtgcgcgc gcgcggcggg cgtgggggggc ggggccgcgg gagcgggggga ggagccccac
 121 ccacagacgg ggaggagcgg gggaggagcg ggggaggagc gggggaggag ccccacccac
 181 agacggggag gagcgggggga ggagcggcca gaccccaaaa acgggccccc ccgaaacaca
 241 cccccgggg gtcgcgcgcg gccctttaaa gcgcggcggc gggcagcccg ggcccccccgc
 301 ggccgagact agcgagttag acaggcaagc actactcgcc tctgcacgca catgcttgcc
 361 tgtcaaactc taccacccccg gcacgctctc tgtctccatg gcccgccgcc gccgccatcg
 421 cggcccccgc cgcccccggc cgcccgggcc cacgggcgcc gtcccaaccg cacagtccca
 481 ggtaacctcc acgcccaact cggaacccgc ggtcaggagc gcgcccgcgg ccgcccccgcc
 541 gccgcccccc gccggtgggc ccccgccttc ttgttcgctg ctgctgcgcc agtggctcca
 601 cgttcccgag tccgcgtccg acgacgacga tgacgacgac tggccggaca gccccccgcc
 661 cgagccggcg ccagaggccc ggcccaccgc cgccgccccc cggccccggt ccccaccgcc
 721 cggcgtgggc ccgggggggcg gggctgaccc ctcccacccc ccctcgcgcc ccttccgcct
 781 tccgccgcgc ctcgccctcc gcctgcgcgt caccgcggag cacctggcgc gcctgcgcct
 841 gcgacgcgcg ggcggggagg gggcgccgga gccccccgcg accccgcgca ccccgcgac
 901 ccccgcgacc cccgcgaccc ccgcgacccc cgcgaccccc gcgacccccg cgcgggtgcg
 961 cttctcgccc cacgtccggg tgcgccacct ggtggtctgg gcctcggccg cccgcctggc
1021 gcgccgcggc tcgtgggccc gcgagcgggc cgaccgggct cggttccggc gccgggtggc
1081 ggaggccgag gcggtcatcg ggccgtgcct ggggcccgag gcccgtgccc gggccctggc
1141 ccgcggagcc ggcccggcga actcggtcta acgttacacc cgaggcgcct gggtcttccg
1201 cggagctccc gggagctccg caccaagccg ctctccggag agacgatggc aggagccgcg
1261 catatatacg ctgggagccg gtccgccccc aaggcgggcc cgcctcgggg gcgggactgg
1321 ccaatcggcg gccgccagcg cggcggggcc cggccaacca gcgtccgccg agtcttcggg
1381 gcccggccca ttgggcggga gttaccgccc aatgggccgg gccgcccact tcccggtatg
1441 gtaattaaaa acttgcaaga ggccttgttc cgcttcccgg tatggtaatt agaaactcat
1501 taatgggcgg ccccggccgc ccttcccgct tccggcaatt cccgcggccc ttaatgggca
```

FIG. 8A

1561 acccc ggtat tccccgcctc ccgcgccgcg cgtaaccact cccctggggt tccgggttat
1621 gctaattgct tttttggcgg aacacacggc ccctcgcgca ttggcccgcg ggtcgctcaa
1681 tgaacccgca ttggtccccc ggggttccgg gtatggtaat gagtttcttc gggaaggcgg
1741 gaagccccgg ggcaccgacg caggccaagc ccctgttgcg tcggcgggag gggcatgcta
1801 atggggttct ttgggggaca ccgggttggt ccccaaatc ggggggccggg ccgtgcatgc
1861 taatgatatt ctttgggggc gccgggttgg tccccgggga cggggccgcc ccgcggtggg
1921 cctgcctccc ctgggacgcg cggccattgg gggaatcgtc actgccgccc ctttggggag
1981 gggaaaggcg tggggtataa gttagccctg gcccgacggt ctggtcgcat ttgcacctcg
2041 gcactcggag cgagacgcag cagccaggca gactcgggcc gccccctctc cgcatcacca
2101 cagaagcccc gcctacgttg cgacccccag ggaccctccg tccgcgaccc tccagccgca
2161 tacgaccccc atggagcccc gccccggagc gagtacccgc cggcctgagg gccgcccccа
2221 gcgcgaggtg aggggccggg cgccatgtct ggggcgccat attggggggc gccatgttgg
2281 gggacccccg acccttaccc tggaaccggc ccccatgttg ggggaccccc actcatacac
2341 gggagccggg cgccatgttg gggcgccatg ttaggggggcg tggaaccccg tgacactata
2401 tatacaggga ccgggggcgc catgttaggg ggcgcggaac ccccctgaccc tatatataca
2461 gggaccgggg tcgccctgtt ggggggtcgcc atgtgacccc ctgactttat atatacagac
2521 ccccaacaca tacacatggc ccctttgact cagacgcagg gcccggggtc gccgtgggac
2581 cccctgactc atacacagag acacgccccc acaacaaaca cacagggacc ggggtcgccg
2641 tgttgggggc gtggtcccca ctgactcata cgcaggcccc ccttactcac acgcatctag
2701 gggggtgggg aggagccgcc cgccatattt gggggacgcc gtgggacccc cgactccggt
2761 gcgtctggag ggcgggagaa gagggaagaa gaggggtcgg gatccaaagg acggacccag
2821 accacctttg gttgcagacc cctttctccc ccctcttccg aggccagcag gggggcagga
2881 ctttgtgagg cgggggggga gaggggaac tcgtgggcgc tgattgacgc gggaaatccc
2941 ccccccattct tacccgcccc cctttttcc ccttagcccg ccccggatgt ctgggtgttt
3001 ccctgcgacc gagacctgcc ggacagcagc gactctgagg cggagaccga agtgggggggg
3061 cgggggacg ccgaccacca tgacgacgac tccgcctccg aggcggacag cacggacacg
3121 gaactgttcg agacggggct gctggggccg cagggcgtgg atgggggggc ggtctcgggg
3181 gggagccccc cccgcgagga agacccccggc agttgcgggg gcgccccccc tcgagaggac

FIG. 8B

3241 gggggagcg acgagggcga cgtgtgcgcc gtgtgcacgg atgagatcgc gccccacctg
3301 cgctgcgaca ccttcccgtg catgcaccgc ttctgcatcc cgtgcatgaa aacctggatg
3361 caattgcgca acacctgccc gctgtgcaac gccaagctgg tgtacctgat agtgggcgtg
3421 acgcccagcg ggtcgttcag caccatcccg atcgtgaacg acccccagac ccgcatggag
3481 gccgaggagg ccgtcagggc gggcacggcc gtggacttta tctggacggg caatcagcgg
3541 ttcgccccgc ggtacctgac cctgggggggg cacacggtga gggccctgtc gcccacccac
3601 ccggagccca ccacggacga ggatgacgac gacctggacg acggtgaggc ggggggcggc
3661 aaggaccctg ggggaggagg aggagggagg aatgggcggg cgggcgagga aagggcgggc
3721 cggggagggg gcgtaacctg atcgcgcccc ccgttgtctc ttgcagcaga ctacgtcccg
3781 cccgcccccc gccggacgcc ccgcgccccc ccacgcagag gcaccgccgc gccccccgtg
3841 acgggcgggg cgtctaacgc agcccccccag ccggccgcgg ctcggacagc gccccccctcg
3901 gcgcccatcg ggccacacgg cagcagtaac accaacacca ccaccaacag cagcggcggc
3961 ggcggctccc gccagtcgcg agccgcggcg ccgcgggggg cgtctggccc ctccgggggg
4021 gttggggttg gggttggggt tgttgaagcg gaggcggggc ggccgagggg ccggacgggc
4081 ccccttgtca acagacccgc ccccccttgca aacaacagag accccatagt gatcagcgac
4141 tccccccccgg cctctcccca caggcccccc gcggcgccca tgccaggctc cgcccccccgc
4201 cccgggcccc ccgcgtcctc ggccgcgtcg ggacccgcgc gccccgcgc ggccgtggcc
4261 ccgtgcgtgc gagcgccgcc tccggggccc ggcccccgcg ccccggcccc cggggcggag
4321 ccggccgccc gccccgcgga cgcgcgccgt gtgccccagt cgcactcgtc cctggctcag
4381 gccgcgaacc aagaacagag tctgtgccgg gcgcgtgcga cggtggcgcg cggctcgggg
4441 gggccgggcg tggagggtgg gcacgggccc tccgcggcc gcacccccctc cggcgccgcc
4501 ccgctccccct ccgccgtctc tgtcgagcag gaggcggcgg tgcgtccgag gaagaggcgc
4561 gggtcgggcc aggaaaaccc ctccccccag tccacgcgtc ccccctcgc gccggcaggg
4621 gccaagaggg cggcgacgca ccccccctcc gactcagggc cgggggggcg cggccagggt
4681 gggcccggga ccccccctgac gtcctcggcg gcctccgcct cttcctcctc tgcctcttcc
4741 tcctcggccc cgaccccccgc gggggccgcc tcttccgccg ccgggggccgc gtcctcctcc
4801 gcttccgcct cctcgggcgg ggccgtcggt gccctgggag ggagacaaga ggaaacctcc
4861 ctcggccccc gcgctgcttc tgggccgcgg gggccgagga agtgtgcccg gaagacgcgc

FIG. 8C

4921 cacgcggaga cttccggggc cgtccccgcg ggcggcctca cgcgctacct gcccatctcg
4981 ggggtctcta gcgtggtcgc cctgtcgcct tacgtgaaca agactatcac gggggactgc
5041 ctgcccatcc tggacatgga gacggggaac atcggggcgt acgtggtcct ggtggaccag
5101 acgggaaaca tggcgacccg gctgcgggcc gcggtccccg gctggagccg ccgcaccctg
5161 ctccccgaga ccgcgggtaa ccacgtgatg cccccgagt acccgacggc cccgcgtcg
5221 gagtggaaca gcctctggat gaccccgtg gggaacatgc tgttcgacca gggcacccta
5281 gtgggcgccc tggacttccg cagcctgcgg tctcggcacc cgtggtccgg ggagcagggg
5341 gcgtcgaccc gggacgaggg aaaacaataa gggacgcccc ccgtgtttgt ggggaggggg
5401 gtcgggtgct gggtggtctc tggccgcgcc cactacacca gccaatccgt gtcggggagg
5461 ggaaagtgaa agacacgggc accacacacc agcgggtctt tagtgttggc cctaataaaa
5521 aactcagggg atttttgctg tctattggga aataaaggtt tacttttgta tcttttccct
5581 gtctgtgttg gatggatctt gggggtgcgt gggagtgggg gtgcgtggga gtgggggtgc
5641 gtgggagtgg gggtgcgtgg gagtgggggt gcgtgggagt gggggtgcgt gggagtgggg
5701 gtgcgtggga gtgggggtgc gtgggagtgg gggtgcgtgg gagtgggggt gcgtgggagt
5761 gggggtgcgt gggagtgggg gtgcgtggga gtgggggtgc gtgggagtgg gggtgccatg
5821 ttgggcaggc tctggtgtta accacagagc cgcggcccgg gctgcctgac caccgatccc
5881 cgaaagcatc ctgccactgg catggagcca gaaccacagt gggctgggtg tgggtgttaa
5941 gtttccgcga gcgcctgccc gcccggactg acctggcctc tggccgccac aaagggcggg
6001 ggggggggtta actacactat agggcaacaa aggacgggag gggtggcggg acggggcgcc
6061 caaaaggggg tcggccacac cacagacgtg ggtgttgggg ggtggggcgg aggggtgggg
6121 gggagacaga aacaggaaca tagttagaaa acaagaatgc ggtgcagcca gagaatcaca
6181 ggagacgagg ggatgggcgt gttggttacc aacccacacc caggcatgct cggtggtatg
6241 aaggaggggg ggcggtgctt cttagagacc gccgggggac gtggggttgg tgtgcaaagg
6301 cacgcgcacc cgcgtcggcc aggtgggccg gtaccccatc ccccctccc ccgacccttc
6361 cccccccgcg tgccagagat cacccccgtc ccccggcacc cgccactcct ccatatcctc
6421 gctttaggaa caactttggg gggggggtac acacgcgccg tgcatttcct tccacacccc
6481 ccctcccccg catccccccc cccaggcagt aagacccaag catagagagc caggcacaaa
6541 aacacaggcg gggtgggaca catgccttct tggagtacgt gggtcattgg cgtgggggggt

FIG. 8D

6601 tacagcgaca ccggccgacc ccctggcggt cttccagccg gcccttagat aaggggggcag
6661 ttggtggtcg gacgggtaag taacagagtc tgactaaggg tgggaggggg ggaaaagaac
6721 gggctggtgt gctgtaacac gagcccaccc gcgagtggcg tggccgacct tagcctctgg
6781 ggcgcccccc gtcgtttggg tccccccccc tctattgggg agaagcaggt gtctaaccta
6841 cctggaaacg cggcgtcttt gttgaacgac accggggcgc cctcgacgag tgggataacg
6901 ggggaggaag ggagggagga gggtactggg ggtgaagaag gggggggggga agaagcgaga
6961 acaggaaagg cgacggagcc cggcagaaca ccgaggaaaa aaaaaacaca gcgcatgcgc
7021 cgggccgttg tggggccccg ggccggggcc ccttgggtcc gccggggccc cgggccgggc
7081 cgccacgggg gccggccgtt ggcggtaacc ccgattgttt atctcaggcc ccgggccggg
7141 aacccggaaa agcctccggg gggccttttt cgcgtcgcgt gccggcgagc gggcccggac
7201 ggggcccgga ccgccgcggt cgggggcccc tcgtcccggg ccgtacgcgg ccttcgcccc
7261 gtgaggggcc gaacgaacga aacatcccgg cgacggaacg aaaaacaccc cagacgggtt
7321 taaaaaacag aaaccgtaac ccccccccacc cccgaaacgg ggaaaacaaa aaacagacca
7381 gcggccggcc ggcgcttagg gggaggatgt cgccgacgcc ccttggccgc cccggctgca
7441 gggggcccg gagagccgcg gcacccggac gcgcccggaa agtctttcgc accacccgcg
7501 atcggcacgg ccgcgccccc gcttttataa aggctcagat gacgcagcaa aaacaggcca
7561 cagcaccacg tgggtaggtg atgtaatttt attttcctcg tctgcggcct aatggatttc
7621 cgggcgcggt gcccctgtct gcagagcact taacggattg atatctcgcg ggcacgcgcg
7681 cccttaatgg accggcgcgg ggcgggggggc cggataccca cacgggcggg ggggtgtcgc
7741 gggccgtctg ctggcccgcg gccacataaa caatgactcg gggcctttct gcctctgccg
7801 cttgtgtgtg cgcgcgccgg ctctgcggtg tcggcggcgg ctgcggcggc tgcggcggcc
7861 gccgtgttcg gtctcggtag ccggccggcg ggtggactcg cgggggccg gagggtggaa
7921 ggcaggggggg tgtaggatgg gtatcaggac ttccacttcc cgtccttcca tccccccgttc
7981 ccctcggttg ttcctcgcct cccccaacac cccgccgctt tccgttgggg ttgttattgt
8041 tgtcgggatc gtgcgggccg ggggtcgccg gggcaggggc gggggcgggg gtgctcgtcg
8101 atcgaccggg ctcagtgggg gcgtggggtg ggtgggaaaa ggcgaggaga ctggggtggg
8161 gggtgtcggg ggtggctgtt tttttgtggt tgtttttgt gtctgttccc gtccccccgtc
8221 acccccctcc ctccgtcccc ccgtcgcggg tgtttgtgtt tgtttattcc gacatcggtt

FIG. 8E

8281 tatttaaata aacacagccg ttctgcgtgt ctgttcttgc gtgtggctgg gggcttatat
8341 gtggggtccc gggggcggga tggggtttag cggcgggggg cggcgcgccg gacggggcgc
8401 tggagataac ggcccccggg gaacgggggа ccggggctgg gtctcccgcg gtgggtgggt
8461 gggcggcggt ggccgggccg ggccgggccg ggtgggcggg gtttggaaaa acgaggagga
8521 ggagaaggag gaggaggggg ggggagacgg ggggaaagca aggacacggc cccggggggg
8581 ggggagcgcg ggccgggccg cttggcaacc cccctgtttc ttccggaaac caggcttgtg
8641 gccccacccg acatcacaag ggacctcttg tcgggcctcc cgacgtacgc cgaggctatg
8701 tcggaccacc ccccaaccta agaggggaga ggggagaggg gagaggggag aggggagagg
8761 ggagagggga ggagaggggg tatataaacc aacgaaaagc gcgggaacgg ggatacgggg
8821 cttgtgtggc acgacgtcgt ggttgtgtta ctgggcaaac acttggggac tgtaggtttc
8881 tgtggtgccg accctaggcg ctatggggat tttgggttgg gttgggctta ttgccgttgg
8941 ggttttgtgt gtgcgggggg gcttgccttc aaccgaatat gttattcgga gtcgggtggc
9001 tcgagaggtg ggggatatat taaaggtgcc ttgtgtgccg ctcccgtctg acgatcttga
9061 ttggcgctac gagacccccc cggctataaa ctatgctttg atagacggta tatttttgcg
9121 ttatcactgt cccggattgg acacggtctt gtgggatagg cacgcccaga gggcgtattg
9181 ggttaacccc ttttgtttg gggcgggttt tttggaggac ttgagtcatc ccgcgtttcc
9241 tgccgacacc caggaaacag aaacgcgctt ggccctttat aaagagatac gccaggcgct
9301 ggacagtcgc aagcaggccg ccagccacac acctgtgaag gctgggtgtg tgaactttga
9361 ctattcgcgc acccgccgct gtgtagggcg ccaggatttg ggacttacca acagaacgtc
9421 tggacggacc ccggttctgc cgtcggacga tgaagcgggc ctgcagccga agcccctcac
9481 cacgccgtcg cccatcatcg ccacgtcgga ccccaccccg cgacgggacg ccgccacaaa
9541 aagcagacgc cgacgacccc attcccggcg catctaatga tgcctcgacg gaaaaccgtc
9601 cgggtttggg gggcgaaccg gccgcctgtc gctcgtcagg gccggcgggc gctcctcgcc
9661 gccctagagg ctgtcccgct ggtgtgacgt tttcctcgtc cgcgcccccc gaccctccca
9721 tggatttaac aaacgggggg gtgtcgcctg cggcgacctc ggcgcctctg gactggacca
9781 cgtttcggcg tgtgtttctg atcgacgacg cgtggcggcc cctgttggag cctgagctgg
9841 cgaacccctt aaccgcccac ctcctggccg aatataatcg tcggtgccag accgaagagg
9901 tgctgccgcc gcgggaggat gtgttttcgt ggactcgtta ttgcaccccc gacgaggtgc

FIG. 8F

9961 gcgtggttat catcggccag gacccatatc accaccccgg ccaggcgcac ggacttgcgt
10021 ttagcgtgcg cgcgaacgtg ccgcctcccc cgagtcttcg gaatgtcttg gcggccgtca
10081 agaactgtta tcccgaggca cggatgagcg gccacggttg cctggaaaag tgggcgcggg
10141 acggcgtcct gttactaaac acgaccctga ccgtcaagcg cggggcggcg gcgtcccact
10201 ctagaatcgg ttgggaccgc ttcgtgggcg gagttatccg ccggttggct gcgcgccgcc
10261 ccggcctggt gtttatgctc tggggcgcac atgcccagaa tgccatcagg ccggaccctc
10321 gggtccattg cgtcctcaag ttttcgcacc cgtcgccccc ctccaaggtt ccgttcggaa
10381 catgccagca tttcctcgtg gcgaatcgat atctcgagac ccggtcgatt tcacccatcg
10441 actggtcggt ttgaaaggca tcgacgtccg gggttttcgt ctgtgggggc ttttgggtat
10501 ttccgatgaa taaagacggt taatggttaa acctctggtc tcatacgggt cggtgatgtc
10561 gggcgtcggg ggagagggag ttcccctctgc gcttgcgatt ctagcctcgt ggggctggac
10621 gttcgacacg ccaaaccacg agtcagggat atcgccagat acgactcccg cagattccat
10681 tcgggggccc gctgtggcct cacctgacca acctttacac gggggcccgg aacgggaggc
10741 cacagcgccg tctttctccc caacgcgcgc ggatgacggc ccgccctgta ccgacgggcc
10801 ctacgtgacg tttgataccc tgtttatggt gtcgtcgatc gacgaattag ggcgtcgcca
10861 gctcaccgac accatccgca aggacctgcg gttgtcgctg gccaagttta gcattgcgtg
10921 caccaagacc tcctcgtttt cgggaaacgc cccgcgccac cacagacgcg gggcgttcca
10981 gcgcggcacg cgggcgccgc gcagcaacaa aagccttcag atgtttgtgt tgtgcaaacg
11041 cgcccacgcc gctcgagtgc gagagcagct tcgggtcgtt attcagtccc gcaagccgcg
11101 caagtattac acgcgatctt cggacgggcg gctctgcccc gccgtccccg tgttcgtcca
11161 cgagttcgtc tcgtccgagc caatgcgcct ccaccgagat aacgtcatgc tggcctcggg
11221 ggccgagtaa ccgccccccc cccgcgccac cctcactgcc cgtcgcgcgt gtttgatgtt
11281 aataaataac acataaattt ggctggttgt ttgttgtctt taatggaccg cccgcagggg
11341 gggtggcatt tcagtgtcgg gtgacgagcg cgatccggcc gggatcctag gaccccaaaa
11401 gtttgtctgc gtattccagg gcggggctca gttgaatctc ccgcagcacc tctaccagca
11461 ggtccgcggt gggctggaga aactcggccg tcccggggca ggcggtcgtc gggagtggag
11521 gcgcggcgcc caccccgtgt gccgcgcctg gcgtctcctc tgggggcgac ccgtaaatgg
11581 ttgcagtgat gtaaatggtg tccgcggtcc agaccacggt caaaatgccg gccgtggtgc

FIG. 8G

11641 tccgggcgct ttcgccgcgc gaggagctga cccaggagtc gaacggatac gcgtacatat
11701 gggcgtccca cccgcgttcg agcttctggt cgctgtcccg gcctataaag cggtaggcac
11761 aaaattcggc gcgacagtcg ataatcacca acagcccaat gggggtgtgc tggataacaa
11821 cgcctccgcg cggcaggcgg tcctggcgct cccggccccg taccataatc gcgcgggtgc
11881 cgtactcaaa aacatgcacc acctgcgcgg cgtcgggcag tgcgctggtc agcgaggccc
11941 tggcgtggca taggctatac gcgatggtcg tctgtggatt ggacatctcg cggtgggtag
12001 tgagtccccc gggccgggtt cggtagaact gtaaggggac ggcgggttaa tagacaatga
12061 ccacgttcgg atcgcgcaga gccgatagta tgtgctcact aatgacgtca tcgcgctcgt
12121 ggcgctcccg gagcggattt aagttcatgc gaaggaattc ggaggaggtg gtgcgggaca
12181 tggccacgta cgcgctgttg aggcgcaggt tgccgggcgt aaagcagatg gcgaccttgt
12241 ccaggctaag gccctgggag cgcgtgatgg tcatggcaag cttggagctg atgccgtagt
12301 cggcgtttat ggccatggcc agctccgtag agtcaatgga ctcgacaaac tcgctgatgt
12361 tggtgttgac gacggacatg aagccgtgtt ggtcccgcaa gaccacgtaa ggcaggggcg
12421 cctcttccag taactcggcc acgttggccg tcgcgtgccg cctccgcagc tcgtccgcaa
12481 aggcaaacac ccgtgcgtac gtgtatccca tgagcgtata attgtccgtc tgcagggcga
12541 cggacatcag ccccccgcgc ggcgagccgg tcagcatctc gcagcccccgg aagataacgt
12601 tgtccacgta cgtgctaaag ggggcgcctt caaatgcctc cccgaagagc tcttggagga
12661 ttcggaatct cccgaggaag gcccgcttca gcagcgcaaa ctgggtgtga acggcggcgg
12721 tggtctccgg ttccccgggg gtgtagtggc agtaaaacac gtcgagctgt tgttcgtcca
12781 gccccgcgaa aataacgtcg aggtcgtcgt cgggaaaatc gtccgggccc ccgtcccgcg
12841 gccccagttg cttaaaatca aacgcacgct cgccgggggc gcctgcgtcg gctattaccg
12901 acgcctgcgt cggcgccccc gaagatttgg ggcgcagaga cagaatctcc gccgttagtt
12961 ctcccatgcg ggcgtaggcg agggtcctct gggtcgcatc caggcccggg cgctgcagaa
13021 agttgtaaaa ggagataagc ccgctaaata tgagccgcga caggaacctg taggcaaact
13081 ccaccgaagt ctccccctga gtctttacaa agctgtcgtc acgcaacact gcctcgaagg
13141 cccggaacgt cccactaaac ccaaaaacca gttttcgcag gcgcgcggtc accgcgatct
13201 ggctgttgag gacgtaagtg acgtcgttgc gggccacgac cagctgctgt ttgctgtgca
13261 cctcgcagcg catgtgcccc gcgtcctggt cctggctctg cgagtagttg gtgatgcggc

FIG. 8H

13321 tggcgttggc cgtgagccac ttttcaatag tcaggccggg ctggtgtgtc agccgtcggt
13381 attcgtcaaa ctccttgacc gacacgaacg taagcacggg gagggtgaac acgacgaact
13441 ccccctcacg ggtcaccttc aggtaggcgt ggagcttggc catgtacgcg ctcacctctt
13501 tgtgggagga gaacagccgc gtccagccgg ggaggttggc ggggttggtg atgtagcttt
13561 ccgggacgac gaagcgatcc acgaactgca tgtgctcctc ggtgatgggc aggccgtact
13621 ccagcacctt catgaggtta ccgaactcgt gctcgacgca ccgtttgttg ttaataaaaa
13681 tggcccagct atacgagagg cgggcgtact cgcgcagcgt gcggttgcag atgaggtacg
13741 tgagcacgtt ctcgctctgg cggacggaac accgcagttt ctggtgctcg aaggtcgact
13801 ccagggacgc cgtctgcgtc ggcgagccca cacacaccaa cacgggccgc aggcgggccg
13861 catactgggg ggtgtggtac agggcgttaa tcatccacca gcaatacacc acggccgtga
13921 ggaggtgacg cccaaggagc ccggcctcgt ctatgacgat cacgttgctg cgggtaaagg
13981 ccggcagcgc cccgtgggtg gccggggcca accgcgtcag ggcgccctcg gccaacccca
14041 gggtccgttc cagggcggcc agggcgcgaa actcgttccg cgactcctcg ccccggagg
14101 cggccagggc gcgcttcgtg aggtccaaaa tcacctccca gtagtacgtc agatctcgtc
14161 gctgcaggtc ctccagcgag gcggggttgc tggtcagggt gtacgggtac tgtcccagtt
14221 gggcctggac gtgattcccg cgaaacccaa attcatgaaa gatggtgttg atgggtcggc
14281 tgagaaaggc gcccgagagt ttggcgtaca tgttttgggc cgcaatgcgc gtggcgcccg
14341 tcaccacaca gtccaagacc tcgttgattg tctgcacgca cgtgctcttt ccggagccag
14401 cgttgccggt gataagatac accgcgaacg gaaactccct gaggggcagg cctgcggggg
14461 actctaaggc cgccacgtcc cggaaccact gcagacgggg cacttgcgct ccgtcgagct
14521 gttgttgcga gagctctcgg atgcgcttaa ggattggctg caccccgtgc atagacgtaa
14581 aatttaaaaa ggcctcggcc ctccctggaa cggctggtcg gtccccgggt tgctgaaggt
14641 gcggcgggcc gggtctctgt ccgtctagct ggcgctcccc gccggccgcc gccatgaccg
14701 caccacgctc gcgggccccc actacgcgtg cgcgggggga cacggaagcg ctgtgctccc
14761 ccgaggacgg ctgggtaaag gttcacccca cccccggtac gatgctgttc cgtgagattc
14821 tccacgggca gctggggtat accgagggcc aggggggtgta caacgtcgtc cggtccagcg
14881 aggcgaccac ccggcagctg caggcggcga tctttcacgc gctcctcaac gccaccactt
14941 accgggacct cgaggcggac tggctcggcc acgtggcggc ccgcggtctg cagccccaac

FIG. 8I

15001 ggctggttcg ccggtacagg aacgcccggg aggcggatat cgccggggtg gccgagcggg
15061 tgttcgacac gtggcggaac acgcttagga cgacgctgct ggactttgcc cacgggttgg
15121 tcgcctgctt tgcgccgggc ggcccgagcg gcccgtcaag cttccccaaa tatatcgact
15181 ggctgacgtg cctggggctg gtccccatat tacgcaagcg acaagaaggg ggtgtgacgc
15241 agggtctgag ggcgtttctc aagcagcacc cgctgacccg ccagctggcc acggtcgcgg
15301 aggccgcgga gcgcgccggc cccgggtttt ttgagctggc gctggccttc gactccacgc
15361 gcgtggcgga ctacgaccgc gtgtatatct actacaacca ccgccggggc gactggctcg
15421 tgcgagaccc catcagcggg cagcgcggag aatgtctggt gctgtggcct cccttgtgga
15481 ccggggaccg tctggtcttc gattcgcccg tccagcggct gtttcccgag atcgtcgcgt
15541 gtcactccct ccgggaacac gcgcacgtct gccggctgcg caataccgcg tccgtcaagg
15601 tgctgctggg gcgcaagagc gacagcgagc gcggggtggc cggcgccgcg cgggtcgtta
15661 acaaggtgtt gggggaggac gacgagacca aggccgggtc ggccgcctca cgcctcgtgc
15721 ggcttatcat caacatgaag ggcatgcgcc acgtaggcga cattaacgac accgtgcgtg
15781 cctacctcga cgaggccggg gggcacctga tagacgcccc ggccgtcgac ggtaccctcc
15841 ctggattcgg caagggcgga aacagccgcg ggtctgcggg ccaggaccag gggggcgggg
15901 cgccgcagct tcgccaggcc ttccgcacgg ccgtggttaa caacatcaac ggcgtgttgg
15961 agggctatat aaataacctg tttggaacca tcgagcgcct gcgcgagacc aacgcgggcc
16021 tggcgaccca attgcaggag cgcgaccgcg agctccggcg cgcaacagcg ggggccctgg
16081 agcgccagca gcgcgcggcc gacctggcgg ccgagtccgt gaccggtgga tgcggcagcc
16141 gccctgcggg ggcggacctg ctccgggccg actatgacat tatcgacgtc agcaagtcca
16201 tggacgacga catgtacgtc gccaacagct ttcagcaccc gtacatccct tcgtacgccc
16261 aggacctgga gcgcctgtcg cgcctctggg agcacgagct ggtgcgctgt tttaaaattc
16321 tgtgtcaccg caacaaccag ggccaagaga cgtcgatctc gtactccagc ggggcgatcg
16381 ccgcattcgt cgccccctac tttgaggcag tgcttcgggc cccccgggta ggcgcgccca
16441 tcacgggctc cgatgtcatc ctgggggagg aggagttatg ggatgcggtg tttaagaaaa
16501 cctgcctgca aacgtacctg acagacatcg cggccctgtt cgtcgcggac gtccagcacg
16561 cagcgctgcc cccgcccccc tccccggtcg gcgccgattt ccggcccggc gcgtccccgc
16621 ggggccggtc cagatcgcgg tcgcccggaa gaactgcgcg aggcgcaccg gaccagggcg

FIG. 8J

16681 ggggcatcgg gcaccgggat ggccgccgcg acggccgacg atgaggggtc ggccgtcacc
16741 atcctcaagc aggccatcgc cggggaccgc agcctggtcg aggcggccga ggcgattagc
16801 cagcagacgc tgctccgcct ggcctgcgag gtgcgccagg tcggcgaccg ccagccgcgg
16861 tttaccgcca ccagcatcgc gcgcgtcgac gtcgcgcctg ggtgccggtt gcggttcgtt
16921 ctggacggga gtcccgagga cgcctatgtg acgtcggagg attactttaa gcgctgctgc
16981 ggccagtcca gttatcgcgg cttcgcggtg gcggtcctga cggccaacga ggaccacgtg
17041 cacagcctgg ccgtgccccc cctcgttctg ctgcaccggt tctccctgtt caaccccagg
17101 gacctcctgg actttgagct tgcctgtctg ctgatgtacc tggagaactg cccccgaagc
17161 cacgccaccc cgtcgacctt tgccaaggtt ctggcgtggc tcggggtcgc gggtcgccgc
17221 acgtccccat tcgaacgcgt tcgctgcctt ttcatccgca gttgccactg ggtcctaaac
17281 acactcatgt tcatggtgca cgtaaaaccg ttcgacgacg agttcgtcct gccccactgg
17341 tacatggccc ggtacctgct ggccaacaac ccgccccccg ttctctcggc cctgttctgt
17401 gccaccccga cgagctcctc attccggctg ccggggccgc ccccccgctc cgactgcgtg
17461 gcctataacc ccgccgggat catggggagc tgctgggcgt cggaggaggt gcgcgcgcct
17521 ctggtctatt ggtggctttc ggagacccca aaacgacaga cgtcgtcgct gttttatcag
17581 ttttgttgaa ttttagtaaa taaacccggt tttgtttcta tggcctcctg acggatgcgc
17641 gtgtccttac tccgttttgg tgggtgggtg gctgtgtatg gcgtcccatc tgtgcgggga
17701 gggggcaagt cggcacgtat tcggacagac tcaagcacac acgggggagc gctcttggct
17761 cagggcaatg tttttattgg tcaaactcag gcaaacagaa acaacatctt gtcgtcaaag
17821 ggatacacaa acttcccccc ctcgccccat actcccgcca gcaccccggt aaacaccaac
17881 tcaatctcgc gcaggatttc gcgcaggtga tgagcgcagt ccacgggggg gagcacaagg
17941 ggccgcgggt atagatcgac ggggacgccg accgactccc cgcctccggg acagacacgc
18001 acgacgcgcc gccagtagtg ctctgcgtcc aacaaggcgc cgccgcggaa ggcagtgggg
18061 ggcaaggggt cgctggcctc aaagggggac acccgaacgc tccagtactc cgcgtccaac
18121 cgtttattaa acgcgtccac gataaggcgg tcgcaggcgt cctccataag gccccgggcc
18181 gtgagcgcgt cctcctccgg cacgcctgcc gttgtcaggc ccaggacccg tcgcagcgtg
18241 tcgcgtacga ccccggccgc cgtggtgtac gcgggcccgc ggagaggaaa tccccaaga
18301 tggtcagtgt tgtcgcggga gttccagaac cacactcccg cttggctcca ggcgacggcg

FIG. 8K

18361 tgggtgtaga cgccctcgag cgccaggcac agtgggtgcc gcagccggag gccgttggcc
18421 ataagcacgg ctcccacggc cgtctcgatg gcccgccggg cgtcctcgat cacccggaa
18481 gccgcatccg cgtcttgggg gtccacgtta aagacacccc agaacgcacc cccatcgccc
18541 ccgcagaccg cgaacttaac cgagctggcc gtctcctcaa tctgcaggca gacggcggcc
18601 atcaccccgc ccaggagctg ccgcagcgca gggcaggcgt cgcacgtgtc cgggaccagg
18661 cgctccaaga cggccccggc ccagggctct gagggagcgg ccaccaccag cgcgtccagt
18721 cttgctaggc ccgtccggcc gtggggttcc gccagcccgc tccccccgag gtcggccagg
18781 gccgccagga gctgggcgcg aagtccgggg aagcaaaacc gcgccgtcca gacgggcccg
18841 acggccgcgg gcgggtctaa cagttggatg attttagtgg cgggatgcca ccgcgccacc
18901 gcctcccgca ccgcgggcag gaggcatccg gctgccgccg aggccacgcc gggccaggct
18961 cgcgggggga ggacgaccct ggcccccacc gcgggccagg cccccaggag cgcggcgtaa
19021 gcggccgcgg ccccgcgcac caggtcccgt gccgactcgg ccgtggccgg cacggtgaac
19081 gtgggccaac ccggaaaccc caggacggca aagtacggga cgggtccccc ccggacctca
19141 aactcgggcc ccagaaaggc aaagacgggg gccagggccc cgggggcggc gtggaccgtg
19201 gtatgccact gccggaaaag ggcgacgagc gccggcgcgg agaacttctc gccggcgctt
19261 acaaagtagt cgtaatcgcg gggcagcagc acccgtgccg tgactcgttg cgggtgcccg
19321 cgtggccgca ggcccacctc gcacacctcg accaggtccc cgaacgctcc ctccttcttg
19381 atcggcggaa acgcaagagt ctggtattcg cgcgcaaata gcgcggttcc ggtggtgatg
19441 ttaacggtca gcgaagcggc ggacgcgcac tgggggtgt cgcgaatggc cgccaggcgc
19501 gcccacgcca gccgcgcgtc gggatgctcg gcaacgcgcg ccgccagggc catagggtcg
19561 atgtcaatgt tggcctccgc gaccaggaga gcggcgcgag gggcggcggg cgggccccac
19621 gacgctctct caactttcac caccagtccc gtgcgtgggt ccgagccgat acgcagcggg
19681 gcgaacaggg ccaccggccc ggtctggcgc tccagggccg ccaggacgca cgcgtacagc
19741 gcccgccaca gagtcgggtt ctccaggggc tccagcgggg aggcggccgg cgtcgtcgcg
19801 gcgcgggcgg ccgccacgac ggcctggacg gagacgtccg cggagccgta gaaatcccgc
19861 agctccgtcg cggtgacgga gacctccgca aagcgcgcgc gaccctcccc tgcggcgttg
19921 cgacatacaa aatacaccag ggcgtggaag tactcgcgag cgcgggggcg cagccatacc
19981 gcgtaaaggg taatggcgct gacgctctcc tccacccaca cgatatctgc ggtgtccatc

FIG. 8L

20041 gcacggcccc taaggatcac gggcggtctg tgggtcccat gctgccgtgc ctggccgggc
20101 ccggtgggtc gcggaaaccg gtgacggggg ggggcggttt ttggggttgg ggtgggggtg
20161 ggaaacggcc cgggtccggg ggccaacttg gcccctcggt gcgttccggc aacagcgccg
20221 ccggtccgcg gacgaccacg taccgaacga gtgcggtccc gagacttata gggtgctaaa
20281 gttcaccgcc ccctgcatca tgggccaggc ctcggtgggg agctccgaca gcgccgcctc
20341 caggatgatg tcagcgttgg ggttggcgct ggatgagtgc gtgcgcaaac agcgccccca
20401 cgcgggcacg cgtagcttga agcgcgcgcc cgcaaactcc cgcttgtggg ccataagcag
20461 ggcgtacagc tgcctgtggg tccggcaggc gctgtggtcg atgtggtggg cgtccaacaa
20521 ccccacgatt gtctgtttgg tgaggttttt aacgcgcccc gccccgggaa acgtctgcgt
20581 gcttttggcc atctgcacgc caaacagttc gccccagatt atcttgaaca gcgccaccgc
20641 gtggtccgtc tcgctaacgg acccgcgcgg gggacagccg cttagggcgt cggcgacgcg
20701 cttgacggct tcctccgaga gcagaagtcc gtcggttacg ttacagtggc ccagttcgaa
20761 caccagctgc atgtagcggt cgtagtgggg ggtcagtagg tccagcacgt catcggggcc
20821 gaaggtcctc ccagatcccc cggccgccga gtcccaatgc aggcgcgcgg ccatggtgct
20881 gcacaggcac aacagctccc agacgggggt tacgttcagg gtgggggggca gggccacgag
20941 ctccagctct ccggtgacgt tgatcgtggg gatgacgccc gtggcgtagt ggtcatagat
21001 ccgccgaaat atggcgctgc tgcgggtggc catgggaacg cggagacagg cctccagcaa
21061 cgccaggtaa ataaaccgcg tgcgtcccat caggctgttg aggttgcgca tgagcgcgac
21121 aatttccgcc ggcgcgacat cggaccggag gtatttttcg acgaaaagac ccacctcctc
21181 cgtctcggcg gcctgggccg gcagcgacgc ctcgggatcc cggcaccgca gctcccgtag
21241 atcgcgctgg gccctgaggg cgtcgaaatg tacgccccgc aaaaacagac agaagtcctt
21301 tggggtcagg gtatcgtcgt gtccccagaa gcgcacgcgt atgcagttta gggtcagcag
21361 catgtgaagg atgttaaggc tgtccgagag acacgccagc gtgcatctct caaagtagtg
21421 tttgtaacgg aatttgttgt agatgcgcga cccccgcccc agcgacgtgt cgcatgccga
21481 cgcgtcacag cgccccttga accggcgaca cagcaggttt gtgacctggg agaactgcgc
21541 gggccactgg ccgcaggaac tgaccacgtg gttcaggagc atgggcgtaa agacgggctc
21601 cgagcgcgcc ccggagccgt ccatgtaaat cagtagctcc cccttgcgga gggtgcgcac
21661 ccgtcccagg gactggtaca cggacaccat gtccggtccg tagttcatgg gtttcacgta

FIG. 8M

21721 ggcgaacatg ccatcaaagt gcaggggatc gaagctgagg cccacggtta cgaccgtcgt
21781 gtatataacc acgcggtatt ggccccacgt ggtcacgtcc ccgaggggggg tgagcgagtg
21841 aagcaacagc acgcggtccg taaactgacg gcagaaccgg gccacgatct ccgcgaagga
21901 gaccgtcgat gaaaaaatgc agatgttatc gcccccgcca aggcgcgctt ccagctcccc
21961 aaagaacgtg gcccccggg cgtccggaga ggcgtccgga gacgggccgc ttggcggccc
22021 gggcgggcgc agggcagcct gcaggagctc ggtccccaga cgcgggagaa acaggcaccg
22081 gcgcgccgaa aacccgggca tggcgtactc gccgaccacc acatgcacgt ttttttcgcc
22141 ccggagaccg cacaggaagt ccaccaactg cgcgttggcg gttgcgtcca tggcgatgat
22201 ccgaggacag gtgcgcagca ggcgtagcat taacgcatcc acgcggccca gttgctgcat
22261 cgttggcgaa tagagctggc ccagcgtcga cataacctcg tccagaacga ggacgtcgta
22321 gttgttcaga aggttggggc ccacgcgatg aaggctttcc acctggacga taagtcggtg
22381 gaaggggcgg tcgttcataa tgtaattggt ggatgagaag taggtgacaa agtcgaccag
22441 gcctgactca gcgaaccgcg tcgccagggt ctgggtaaaa ctccgacgac aggagacgac
22501 gagcacactc gtgtccggag agtggatcgc ttcccgcagc cagcggatca gcgcggtagt
22561 ttttcccgac cccattggcg cgcggaccac agtcacgcac ctggccgtcg gggcgctcgc
22621 gttggggaag gtgacgggtc cgtgctgctg ccgctcgatc gttgttttcg ggtgaacccg
22681 gggcacccat tcggccaaat cccccccgta caacatccgc gctagcgata cgctcgacgt
22741 gtactgttcg cactcgtcgt ccccaatggg acgcccggcc cccagaggat cccccgactc
22801 cgcgcccccc acgaaaggca tgaccggggc gcggacggcg tggtgggtct ggtgtgtgca
22861 ggtggcgacg tttgtggtct ctgcggtctg cgtcacgggg cttctcgtcc tggcctctgt
22921 gttccgggca cggtttccct gcttttatgc cacggcgagc tcttatgccg gggtgaactc
22981 cacggccgag gtgcgcgggg gtgtagccgt gcccctcagg ttggacacgc agagccttgt
23041 gggcacttat gtaatcacgg ccgtgttgtt gttggccgcg gccgtgtatg ccgtggtcgg
23101 cgccgtgacc tcccgctacg accgcgccct ggacgcgggc cgccgtctgg ctgcggcccg
23161 catggccatg ccgcacgcca cgctgatcgc cggaaacgtc tgctcttggt tgctgcagat
23221 caccgtcctg ttgctggccc atcgcatcag ccagctggcc cacctggttt acgtcctgca
23281 ctttgcgtgt ctggtgtatt ttgcggccca tttttgcacc agggggggtcc tgagcgggac
23341 gtatctgcgt caggtgcacg gcctgatgga gccggccccg actcatcatc gcgtcgtcgg

FIG. 8N

23401 cccggctcga gccgtgctga caaacgcctt gctgttgggc gtcttcctgt gcacggccga
23461 cgccgcggta tccctgaata ccatcgccgc gttcaacttt aatttttcgg ccccgggcat
23521 gctcatatgc ctgaccgtgc tgttcgccct tctcgtcgta tcgctgttgt tggtggtcga
23581 gggggtgttg tgtcactacg tgcgcgtgtt ggtgggcccc cacctggggg ccgtggccgc
23641 cacgggcatc gtcggcctgg cctgcgagca ctattacacc aacggctact acgttgtgga
23701 gacgcagtgg ccgggggccc agacgggagt ccgcgtcgcc ctcgccctgg tcgccgcctt
23761 tgccctcggc atggccgtgc tccgctgcac ccgcgcctat ctgtatcaca ggcggcacca
23821 caccaaattt tttatgcgca tgcgcgacac gcgacaccgc gcacattccg ccctcaggcg
23881 cgtacgcagt tccatgcgcg gatcgcgaga cggccgccac aggcccgcac ccggcagccc
23941 gcccgggatt cccgaatatg cggaagaccc ctacgcgatc tcatacggcg gccagctcga
24001 ccggtacgga gattccgacg gggagccgat ttacgacgag gtggctgacg accaaaccga
24061 cgtattgtac gccaagatac aacacccgcg gcacctgccc gacgacgagc ccatctatga
24121 caccgttggg gggtacgacc ccgagcccgc cgaggacccc gtgtacagca ccgtccgccg
24181 ttggtagctg tttggttccg ttttaataaa ccgtttgtgt ttaacccgac cgtggtgtat
24241 gtctggtgtg tggcgtccga tcccgttact atcaccgttt cccccccccc cccctcaacc
24301 ccggcgattg tgggtttttt aaaaacgaca cgcgtgcgac cgtatacaga acattgtttt
24361 ggtttttatt cgctatcgga catgggggt ggaaactggg tggcggggca ggcgcctccg
24421 ggggtccgcc ggtgagtgtg gcgcgagggg gggtccgacg aacgcaggcg ctgtctcccc
24481 ggggcccgcg taaccacgcg catatccggg ggcacgtaga aattaccttc ctcttcggac
24541 tcgatatcca cgacatcaaa gtcgtgggcg gtcagcgaga cgacctcccc gtcgtcggtg
24601 atgaggacgt tgtttcggca gcagcagggc cgggccccgg agaacgagag gcccatagct
24661 cggcgagcgt gtcgtcgaac gccaggcggc tgcttcgctg gatggcctta tagatctccg
24721 gatcgatgcg gacgggggta atgatcaggg cgatcggaac ggcctggttc gggagaatgg
24781 acgccttgct gggtcctgcg gccccgagag ccccggcgcc gtcctccagg cggaacgtta
24841 cgccctcctc cgcgctggtg cggtgcctgc cgataaacgt caccagatgc gggtgggggg
24901 ggcagtcggg gaagtggctg tcgagcacgt agccctgcac caagatctgc ttaaagttcg
24961 ggtgacgggg gttcgcgaag acgggctcgc ggcggaccag atccccggag ctccaggaca
25021 cgggggagat ggtgtggcgt ccgaggtcgg gggcgccaaa cagaagcacc tccgagacaa

FIG. 8O

25081 cgccgctatt taactccacc aaggcccgat ccgcggcgga gcaccgcctt ttttcgcccg
25141 aggcgtgggc ctctgaccag gcctggtctt gcgtgacgag agcctcctcc gggccgggga
25201 cgcgcccggg cgcgaagtat cgcacgctgg gcttcgggat cgaccggata aatgcccgga
25261 acgcctccgg ggaccggtgt gtcatcaagt cctcgtacgc ggaggccgtg gggtcgctgg
25321 ggtccatggg gtcgaaagcg tacttggccc ggcatttgac ctcgtaaaag gccaggggcg
25381 tcttggggac tggggccagg tagccgtgaa tctcccgagg acagacgaga atatccaggg
25441 acgccccgac catccccgtg tgaccgtcca tgaggacccc acacgtatgc acgttctctt
25501 cggtgaggtc gctgggttcg tggaagataa agcgccgcgt gtcggcgccg gcctcgccgc
25561 cgtcgtccgc gcggcccacg cagtagcgaa acagcaggct tcgggccgtc ggctcgttca
25621 cccgcccgaa catcaccgcc gaagactgta catccggccg caggctggcg ttgtgcttca
25681 gccactgggg cgagaaacac ggaccctggg ggccccagcg gagggtggat gcggtcgtga
25741 ggccccgccg gagcagggcc catagctggc agtcggcctg gttttgcgtg gccgcctcgt
25801 aaaaccccat gaggggccgg ggcgccacgg cgtccgcggc ggccgggggc ccgcggcgcg
25861 tcaggcgcca taggtgccgg ccgagtccgc ggtccaccat acccgcctcc tcgaggacca
25921 cggccaggga acacagataa tccaggcggg cccagagggg accgatggcc agaggggcgc
25981 ggacgccgcg cagcaacccg cgcaggtggc gctcgaacgt ctcggctagt atatgggagg
26041 gcagcgcgtt ggggatcacc gacgccgacc acatagagtc aaggtccggg gagtcgggat
26101 cggcgtccgg gtcgcgggcg tgggtgcccc caggagatag cggaatgtct ggggtcggag
26161 gccctgaggc gtcagaaagt gccggcgacg cggcccgggg cttttcgtct gcggtgtcgg
26221 tggcgtgctg atcacgtggg gggttaacgg gcgaatggga gctcgggtcc acagctgacg
26281 tcgtctgggg tgggggggcc aggggacgga aggtggttgt cagcggaaga ctgttagggc
26341 gggggcgctt gggggggctg tcgggggccac gaggggtgtc ctcggccagg gcccagggac
26401 gcttagtcac ggtgcgtccc ggcggacatg ctgggcctac cgtggactcc atttccgaga
26461 cgacgtgggg gagcggtggt tgagcgcgcc gccgggtgaa cgctgattct cacgacagcg
26521 cgtgccgcgc gcacgggttg gtgtgacaca ggcgggacac cagcaccagg agaggcttaa
26581 gctcgggagg cagcgccacc gacgacagta tcgccttgtg tgtgtgctgg taatttatac
26641 accgatccgt aaacgcgcgc cgaatcttgg gattgcggag gtggcgccgg atgccctctg
26701 gtacgtcata cgccaggccg tgggtgttgg tctcggccga gttgacaaac agggctgggt

FIG. 8P

26761 gcagcacgca gcgataggcg agcagggcca gggcgaagtc cggcgacagc tggttgttga
26821 aatactggta accgggaaac cgggtcacgg gtacgcccag gctcggggcg acgtacacgc
26881 taaccaccaa ctccagcagc gtctggccaa gggcgtacag gtcaaccgct aacccgacgt
26941 cgtgcttcag gcggtggttg gtaaattcgg cccgttcgtt gttaaggtat ttcaccaaca
27001 gctccggggg ctggttatac ccgtgaccca ccagggtgtg aaagttggct gtggttaggg
27061 cggtgggcat gccaaacatc cgggggggact tgaggtccgg ctcctggagg caaaactgcc
27121 cccgggcgat cgtggagttg gagttgaggg tgacgaggct aaagtcggcg aggacggccc
27181 gccggagcga gacggcgtcc gaccgcagca tgacgaggat gttggcgcac ttgatatcca
27241 ggtggctgat cccgcaggtg gtgtttaaaa acacaacggc gcgggccagc tccgtgaagc
27301 actggtggag ggccgtcgag accgaggggt ttgttgtgcg cagggacgcc agttggccga
27361 tatacttacc gaggtccatg tcgtacgcgg ggaacactat ctgtcgttgt tgcagcgaga
27421 acccgagggg cgcgatgaag ccgcggatgt tgtgggtgcg gccggcgcgt agagcgcact
27481 ccccgaccaa cagggtcgcg atgagctcaa cggcaaacca ctcctttcc tttatggtct
27541 taacggcaag cttatgttcg cgaatcagtt ggacgtcgcc gtatccccca gacccccccga
27601 agcttcgggc cccggggatc tcgagggtcg tgtagtgtag ggcggggttg atggcgaaca
27661 cggggctgca tagcttgcgg atgcgcgtga gggtaaggat gtgcgagggg gacgaggggg
27721 gtgcggttaa cgccgcctgg gatctgcgca ggggcgggcg gttcagtttg gccgccgtac
27781 cgggcgtctc gggggacgcg cggcgatgag acgagcggct cattcgccat cgggatagtc
27841 ccgcgcgaag ccgctcgcgg aggccggatc ggtggcggga cccgtgggag gagcgggagc
27901 cggcggcgtc ctggagagag gggccgctgg ggcgcccgga ggccccgtgt gggttggagt
27961 gtatgtagga tgcgagccaa tccttgaagg accgttggcg tgcaccttgg gggctgaggt
28021 tagctgccac atgaccagca ggtcgctgtc tgcgggactc atccatcctt cggccaggtc
28081 gccgtctccc cacagagaag cgttggtcgc tgcttcctcg agttgctcct cctggtccgc
28141 aagacgatcg tccacggcgt ccaggcgctc accaagcgcc ggatcgaggt accgtcggtg
28201 tgcggttaga aagtcacgac gcgccgcttg ctcctccacg cgaattttaa cacaggtcgc
28261 gcgctgtcgc atcatctcta agcgcgcgcg ggactttagc cgcgcctcca attccaagtg
28321 ggccgccttt gcagccataa aggcgccaac aaaccgagga tcttgggtgc tgacgccctc
28381 ccggtgcagc tgcagggtct ggtccttgta aatctcggct cggaggtgcg tctcggccag

FIG. 8Q

28441 gcgtcggcgc agggccgcgt gggcggcatc tcggtccatt ccgccaccct gcgggcgacc
28501 cgggggggtgc tctgatagtc tcgcgtgccc aaggcccgtg atcggggtac ttcgccgccg
28561 cgacccgcca cccggtgtgc gcgatgtttg gtcagcagct ggcgtccgac gtccagcagt
28621 acctggagcg cctcgagaaa cagaggcaac ttaaggtggg cgcggacgag gcgtcggcgg
28681 gcctcacaat gggcggcgat gccctacgag tgccctttt agatttcgcg accgcgaccc
28741 ccaagcgcca ccagaccgtg gtcccgggcg tcgggacgct ccacgactgc tgcgagcact
28801 caccgctctt ctcggccgtg gcgcggcggc tgctgtttaa tagcctggtg ccggcgcaac
28861 taaaggggcg tgatttcggg ggcgaccaca cggccaagct ggaattcctg gccccgagt
28921 tggtacgggc ggtggcgcga ctgcggttta aggagtgcgc gccggcggac gtggtgcctc
28981 agcgtaacgc ctactatagc gttctgaaca cgtttcaggc cctccaccgc tccgaagcct
29041 ttcgccagct ggtgcacttt gtgcgggact ttgcccagct gcttaaaacc tccttccggg
29101 cctccagcct cacggagacc acgggccccc ccaaaaaacg ggccaaggtg gacgtggcca
29161 cccacggccg gacgtacggc acgctggagc tgttccaaaa aatgatcctt atgcacgcca
29221 cctactttct ggccgccgtg ctcctcgggg accacgcgga gcaggtcaac acgttcctgc
29281 gtctcgtgtt tgagatcccc ctgtttagcg acgcggccgt gcgccacttc cgccagcgcg
29341 ccaccgtgtt tctcgtcccc cggcgccacg gcaagacctg gtttctagtg ccctcatcg
29401 cgctgtcact ggcctccttt cgggggatca agatcggcta cacggcgcac atccgcaagg
29461 cgaccgagcc ggtgtttgag gagatcgacg cctgcctgcg gggctggttc ggttcggccc
29521 gagtggacca cgttaaaggg gaaaccatct ccttctcgtt tccggacggg tcgcgcagta
29581 ccatcgtgtt tgcctccagc cacaacacaa acgtaagtcc tcttttcttt cgcatggctc
29641 tcccaagggg ccccgggtcg acccgaccca cacccaccca cccacccaca tacacacaca
29701 accagacgcg ggaggaaagt ctgccccgtg ggcactgatt tttattcggg atcgcttgag
29761 gaggcccggg caacggcccg ggcaacggtg gggcaactcg tagcaaatag gcgactgatg
29821 tacgaagaga agacacacag gcgccacccg gcgctggtcg gggggatgtt gtccgcgccg
29881 caccgtcccc cgacgacctc ttgcagacgg tccgtgatgc aaggacggcg gggggcctgc
29941 agcagggtga ccgtatccac gggatggcca aagagaagcg gacacaggct agcatccccc
30001 tggaccgcca gggtacactg ggccatcttg gcccacagac acggggcgac gcagggacag
30061 gactccgtta cgacggagga gagccacagt gcgttggcgg aatcgatgtg gggcggcggg

FIG. 8R

30121 gcgcaggact cgcagccccc cgggtggttg gtgatcctgg ccaggagcca tcccagatgg
30181 cgggccctgc ttcccggtgg acagagcgac cccaggtcgc tgtccatggc ccagcagtag
30241 atctggccgc tggggaggtg ccaccaggcc cccgggccca aggcgcagca cgcgcccggc
30301 tccgggggggg tcttcgcggg gaccagatac gcgccatcca gctcgccgac cactggctcc
30361 tccgcgagct gttcggtggt tgggtcgggg gtttcctccg ggggggtggc cgcccgtatg
30421 cgggcgaacg tgagggtgca caggagcggg gtcagggggt gcgtcacgct ccggaggtgg
30481 acgatcgcgc agtagcggcg ctcgcggtta aagaaaaaga gggcaaagaa ggtgttcggg
30541 ggcaaccgca gcgccttggg gcgcgtcaga tacagaaaaa tctcgcagaa gagggcgcgc
30601 ccggggtctg ggttaggaag ggccacctga cacagaggct cggtgaggac cgttagacac
30661 cgaaagatct tgagccgctc gtccgcccga acgacgcgcc acacaaagac ggagttgaca
30721 atgcgcgcga tagagtcgac gtccgtcccc aggtcgtcga ctctgtcgcg cgtgccgcga
30781 gctccggccc gggaatccgg ccggggcaag gtccccgggg gaccaggcgg cgccaggggc
30841 cgccggggtc ccagctgcgc catgccgggg gcggggggag ggcaaacccc agaggcgggg
30901 gccaacggcg cggggaggag tgggtgggcg aggtggccgg gggaaggcgc ccgctagcga
30961 gaacggccgt tcccggacga caccttgcga caaaacctaa ggacagcggc ccgcgcgacg
31021 gggtccgaga ggctaaggta ggccgcgatg ttaatggtga acgcaaagcc gccgggaaag
31081 acaactatgc cacagaggcg gcgattaaac cccaggcaga ggtaggcgta gctttccccg
31141 ggcaggtatt gctcgcagac cctgcgtggg gctgtggagg ggacggcctc catgaagcga
31201 catttactct gctcgcgttt actgacgtca ccatccatcg ccacggcgat tggacgattg
31261 ttaagccgca gcgtgtctcc gcttgtgctg tagtagtcaa aaacgtaatg gccgtcggag
31321 tcggcaaagc gggccgggag gtcgtcgccg agcgggacga cccgccgccc ccgaccgccc
31381 cgtcccccca ggtgtgccag gacggccagg gcatacgcgg tgtgaaaaaa ggagtcgggg
31441 gcggtcccct cgacggcgca catcaggttc tcgaggagaa tggggaagcg cctggtcacc
31501 tcccccaacc acgcgcgttg gtcggggcca aagtcatagc gcaggcgctg tgagattcgc
31561 gggccgccct gaagcgcggc ccggatggcc tggcccaggg cccggaggca cgccagatgt
31621 atgcgcgcgg taaaggcgac ctcggcggcg atgtcaaagg gcggcaggac ggggcgcggg
31681 tggcgcaggg gcacctcgag cgcgggaaag cgtagcagca gctccgcctg cccagcggga
31741 gacagctggt gggggcgcac gacgcgttct gcggcgcagg cctcggtcag ggccgtggcc

FIG. 8S

31801 agcgccgagg acagcagcgg agggcgggcg cgtcgcccgc cccacgccac ggagttctcg
31861 taggagacga cgacgaagcg ctgcttggtt ccgtagtggt ggcgcaggac cacggagata
31921 gaacgacggc tccacagcca gtccggccgg tcgccgccgg ccagggcttc ccatccgcga
31981 tccaaccact cgaccagcga ccgcggcttt gcggtaccag gggtcagggt tagaacgtcg
32041 ttcaggatgt cctcgccccc gggcccgtgg ggcacggggg ccacaaagcg gccccgcct
32101 gggggctcca gacccgccaa caccgcatct gcgtcagccg ccccatggc gccccgctg
32161 acggcctggt gaaccagggc gccctggcgg agccccgatg caacgccaca ggccgcacgc
32221 ccggtccgag cgcggaccgg gtggcggcgg gtgacgtcct gcactgcccg ctgaaccaac
32281 gcgaggatct cctcgttctc ctgcgcgatg gacacgtcct gggccgcggt cgtgtcgccg
32341 ccgggggccg tcagctgctc ctccggggag atgggggggt cggacgcccc gacgatgggc
32401 gggtctgcgg gcgccccgc gtggggccgg gccaagggct gcggacgcgg ggacgcgctt
32461 tccccagac ccatggacag gtgggccgca gcctccttcg cggccggcgg ggcggcggcg
32521 ccaagcagag cgacgtagcg gcacaaatgc cgacagacgc gcatgatgcg cgtgctgtcg
32581 gccgcgtagc gcgtgttggg ggggacgagc tcgtcgtaac taaacagaat cacgcgggca
32641 cagctcgccc ccgagcccca cgcaaggcgc agcgccgcca cggcgtacgg gtcatagacg
32701 ccctgcgcgt tacacaccac gggcagggag acgaacaacc ccccggcgct ggacgcacgc
32761 ggaaggaggc cagggtgtgc cggcacgacg ggggccagaa gctcccccac cgcatccgcg
32821 ggcacgtagg cggcaaacgc cgtgcaccac ggggtacagt cgccggtggc atgagcccga
32881 gtctggattt cgacctggaa gtttgcggcc gtcccgagtc cggggtggcc gcgcatcagg
32941 gcggccagag ggattcccgc ggccgccagg cactcgctgg atatgatgac gtgaaccaaa
33001 gacgagggcc gacccgggcc gtggccgaga tcgtactgga cctcgttggc caagtgcgcg
33061 ttcatggttc gggggtgggt gtgggtgtgt aggcgatgcg ggtcccccga gtccgcggga
33121 agggcgtggg tttggcgcgc gtatgcgtat tcgccaacgg aggcgtgcgt gcttatgcgc
33181 ggcgcgtttc ttctgtctct agggaatccg aggccaggac tttaacctgc tctttgtcga
33241 cgaggccaac tttattcgcc cggatgcggt ccagacgatt atgggctttc tcaaccaggc
33301 caactgcaag attatcttcg tgtcgtccac caacaccggg aaggccagta cgagcttttt
33361 gtacaacctc cgcgggccg cagacgagct tctcaacgtg gtgacctata tatgcgatga
33421 tcacatgccg agggtggtga cgcacacaaa cgccacggcc tgttcttgtt atatcctcaa

FIG. 8T

33481 caagcccgtt ttcatcacga tggacggggc ggttcgccgg accgccgatt tgtttctggc
33541 cgattccttc atgcaggaga tcatcgggggg ccaggccagg gagaccggcg acgaccggcc
33601 cgttctgacc aagtctgcgg gggagcggtt tctgttgtac cgcccctcga ccaccaccaa
33661 cagcggcctc atggcccccg atttgtacgt gtacgtggat cccgcgttca cggccaacac
33721 ccgagcctcc gggaccggcg tcgctgtcgt cgggcggtac cgcgacgatt atatcatctt
33781 cgccctggag cacttttttc tccgcgcgct cacgggctcg gcccccgccg acatcgcccg
33841 ctgcgtcgtc cacagtctga cgcaggtcct ggccctgcat cccggggcgt ttcgcggcgt
33901 ccgggtggcg gtcgagggaa atagcagcca ggactcggcc gtcgccatcg ccacgcacgt
33961 gcacacagag atgcaccgcc tactggcctc ggaggggggcc gacgcgggct cgggccccga
34021 gcttctcttc taccactgcg agcctcccgg gagcgcggtg ctgtaccccct ttttcctgct
34081 caacaaacag aagacgcccg cctttgaaca ctttattaaa aagtttaact ccgggggcgt
34141 catggcctcc caggagatcg tttccgcgac ggtgcgcctg cagaccgacc cggtcgagta
34201 tctgctcgag cagctaaata acctcaccga aaccgtctcc cccaacactg acgtccgtac
34261 gtattccgga aaacggaacg gcgcctcgga tgaccttatg gtcgccgtca ttatggccat
34321 ctacctcgcg gcccaggccg gacctccgca cacattcgct cctatcacac gcgtcttgtg
34381 agcgcccaat aaacacaccc aggtatgcta cgcacgacca cggtgtcgtc tgttaagggg
34441 gggggaagg gggtgttggc gggaagcgtg ggaacacggg ggattctctc acgaccggca
34501 ccagtaccac ccccctgtga acacagaaac cccaacccaa atcccataaa catacgacac
34561 acaggcatat tttggaattt cttaggtttt tatttattta ggtatgctgg ggtttctccc
34621 tggatgccca ccccacccc cccgtgggtc tagccgggcc ttagggatag cgtataacgg
34681 gggccatgtc tccggaccgc acaacggccg cgccgtcaaa ggtgcacacc cgaaccacgg
34741 gagccagggc caaggtgtct cctagttggc ccgcgtgggt cagccaggcg acgagcgcct
34801 cgtaaagcgg cagccttcgc tctccatcct gcatcagggc cggggcttcg gggtgaatga
34861 gctgggcggc ctcccgcgtg acactctgca tctgcaggag agcgttcacg tacccgtcct
34921 gggcacttag cgcaaagagc cgggggatta gcgtaaggat gatggtggtt ccctccgtga
34981 tcgagtaaac catgttaagg accagcgatc gcagctcggc gtttacggga ccgagttgtt
35041 ggacgtccgc cagcagcgag aggcgactcc cgttgtagta cagcacgttg aggtctggca
35101 gccctccggg gttttctgggg ctggggttca ggtcccggat gcccctggcc acgagccgcg

FIG. 8U

35161 ccacgatttc gcgcgccagg ggcgatggaa gcggaacggg aaaccgcaac gtgaggtcca
35221 gcgaatccag gcgcacgtcc gtcgcttggc cctcgaacac gggcgggacg aggctgatgg
35281 ggtccccgtt acagagatct acgggggagg tgttgcgaag gttaacggtg ccggcgtggg
35341 tgaggcccac gtccaggggg caggcgacga ttcgcgtggg aagcacccgg gtgatgaccg
35401 cggggaagcg ccttcggtac gccagcaaca accccaacgt gtcgggactg acgcctccgg
35461 agacgaagga ttcgtgcgcc acgtcggcca gcgtcagttg ccggcggatg gtcggcagga
35521 ataccacccg cccttcgcag cgctgcagcg ccgccgcatc ggggcgcgag atgcccgagg
35581 gtatcgcgat gtcagtttca aagccgtccg ccagcatggc gccgatccac gcggcaggga
35641 gtgcagtggt gggtcgggtg gcgggaggag cgcggtgggg gtcagcggcg tagcagagac
35701 gggcgaccaa cctcgcatag gacggggggt gggtcttagg gggttgggag gcgacaggga
35761 ccccagagca tgcgcgggga ggtctgtcgg gcccagacgc accgagagcg aatccgtccg
35821 cggagtcccg gcttgggttt tatggggccc ggccctcgga atcgcggctt gtcggcgggg
35881 acaaaggggg cggggctagg ggcttgcgga aacagaagac gcgtgggata aaagaatcgc
35941 actaccccaa ggaagggcgg ggcggtttat tacagagcca gtcccttgag cggggatgcg
36001 tcatagacga gatactgcgc gaagtgggtc tcccgcgcgt gggcttcccc gttgcgggcg
36061 ctgcggagga gggcggggtc gctggcgcag gtgagcgggt aggcctcctg aaacaggcca
36121 cacgggtcct ccacgagttc gcggcacccc gggggcgct taaactgtac gtcgctggcg
36181 gcggtggccg tggacaccgc cgaacccgtc tccacgatca ggcgctccag gcagcgatgt
36241 ttggcggcga tgtcggccga cgtaaagaac ttaaagcagg ggctgagcac cggcagggcc
36301 ccgttgaggt ggtaggcccc gttatagagc aggtccccgt acgaaaatcg ctgcgacgcc
36361 cacgggttgg ccgtggccgc gaaggcccgg gacgggtcgc tctggccgtg gtcgtacatg
36421 agggcggtga catccccctc cttgtccccc gcgtaaacgc ccccggcggc gcgtccccgg
36481 gggttgcagg gccggcggaa gtagttgacg tcggtcgaca cgggggtggc gataaactca
36541 cacacggcgt cctggccgtg gtccatcccc gcgcgccgcg gcacctgggc gcacccgaac
36601 acggggacgg gctgggccgg ccccaggcgg tttcccgcca cgaccgcgtt ccgcaggtac
36661 acggctgccg cgttgtccag gagaggggga gccccgcggc ccaggtaaaa gttttgggga
36721 aggttgccca tgtcggtgac ggggttgcgg acggttgccg tggccacgac ggcggtgtag
36781 cccacgccca ggtccacgtt cccgcgcggc tgggtgagcg tgaagtttac ccccccgcca

FIG. 8V

36841 gtttcgtgcc gggccacctg gagctggccc aggaagtacg cctccgacgc gcgctccgag
36901 aacagcacgt tctcagtcac aaagcggtcc tgtcggacga cggtgaaccc aaacccggga
36961 tggaggcccg tcttgagctg atgatgcaag gccacgggac tgatcttgaa gtaccccgcc
37021 atgagcgcgt aggtcagcgc gttctccccg gccgcgctct cgcggacgtg ctgcacgacg
37081 ggctgtcgga tcgacgaaaa gtagttggcc cccagagccg gggggaccag ggggacctgc
37141 cgcgacaggt cgcgcagggc cgggggggaaa ttgggcgcgt tcgccacgtg gtcggccccg
37201 gcgaacagcg cgtggacggg gaggggggtaa aaatagtcgc cattttggat ggtatggtcc
37261 agatgctggg gggccatcag caggattccg gcgtgcaacg ccccgtcgaa tatgcgcatg
37321 ttggtggtgg acgcggtgtt ggcgcccgcg tcgggcgccg ccgagcagag cagcgccgtt
37381 gtgcgttcgg ccatgttgtg ggccagcacc tgcagcgtga gcatggcggg cccgtccact
37441 accacgcgcc cgttgtgaaa catggcgttg accgtgttgg ccaccagatt ggccgggtgc
37501 aggggggtgcg cggggtccgt cacggggtcg ctggggcact cctcgccggg ggcgatctcc
37561 gggaccacca tgttctgcag ggtggcgtat acgcggtcga agcgaacccc cgcggtgcag
37621 cagcggcccc gcgagaaggc gggcaccatc acgtagtagt aaatcttgtg gtgcacggtc
37681 cagtccgccc cccggtgcgg ccggtcatcc gcggcgtccg cggctcgggc ctgggtgttg
37741 tgcagcagct ggccgtcgtt gcggttgaag tccgcggtcg ccacgttaca tgccgccgcg
37801 tacacggggt cgtggccccc cgcgctaacc cggcagtcgc gatggcggtc cagggccgcg
37861 cgccgcatca gggcgtcaca gtcccacacg aggggtggca gcagcgccgg gtctcgcatt
37921 aggtgattca gctcggcttg cgcctgcccg cccagctccg ggccggtcag ggtaaagtca
37981 tcaaccagct gggccagggc ctcgacgtgc gccaccaggt cccggtacac ggccatgcac
38041 tcctcgggaa ggtctccccc gaggtaggtc acgacgtacg agaccagcga gtagtcgttc
38101 acgaacgccg cgcaccgcgt gttgttccag tagctggtga tgcactggac cacgagccgg
38161 gccagggcgc agaagacgtg ctcgctgccg tgtatggcgg cctgcagcag gtaaaacacc
38221 gccgggtagt tgcggtcgtc gaacgccccg cgaacggcgg cgatggtggc gggggccatg
38281 gcgtggcgtc ccaccccag ctccaggccc cgggcgtccc ggaacgccgc cggacatagc
38341 gccaggggca agttgccgtt caccacgcgc caggtggcct ggatctcccc cgggccggcc
38401 gggggaacgt ccccccccgg cagctccacg tcggccaccc ccacaaagaa gtcgaacgcg
38461 gggtgcagct caagagccag gttggcgttg tcgggctgca taaactgctc cggggtcatc

FIG. 8W

38521 tggccttccg cgacccatcg gacccgcccg tgggccaggc gctgccccca ggcgttcaaa
38581 aacagctgct gcatgtctgc ggcggggccg gccgggccg ccacgtacgc cccgtacgga
38641 ttggcggctt cgacggggtc gcggttaagg cccccgaccg ccgcgtcaac gttcatcagc
38701 gaagggtggc acacggtccc gatcgcgtgt tccagagaca ggcgcagcac ctggcggtcc
38761 ttcccccaaa aaaacagctg gcgggggcggg aaggcgcggg gatccgggtg gccgggggcg
38821 gggactaggt ccccggcgtg cgcggcaaac cgttccatga ccggattgaa caggcccagg
38881 ggcaggacga acgtcaggtc catggcgccc accaggggggt agggaacgtt ggtggcggcg
38941 tagatgcgct tctccagggc ctccagaaag accagcttct cgccgatgga caccagatcc
39001 gcgcgcacgc gcgtcgtctg gggggcgctc tcgagctcgt ccagcgtctg ccggttcagg
39061 tcgagctgct cctcctgcat ctccagcagg tggcggccca cgtcgtccag acttcgcacg
39121 gccttgccca tcacgagcgc cgtgaccagg ttggccccgt tcaggaccat ctcgccgtac
39181 gtcaccggca cgtcggcttc ggtgtcctcc gctttcagga aggactgcag gaggcgctgt
39241 ttgatcgggg cggtggtgac gagcaccccg tcgaccggcc gcccgcgcgt gtcggcatgc
39301 gtcagacggg gcacggccac ggagggctgc gtggccgtgg tgaggtccac gagccaggcc
39361 tcgacggcct cccggcggtg gcccgccttg cccaggaaaa agctcgtctc gcagaagctt
39421 cgctttagct cggcgaccag ggtcgcccgg gccaccctgg tggccaggcg gccgttgtcc
39481 aggtatcgtt gcatcggcaa caacaaagcc aggggcggcg ccttttccag cagcacgtgc
39541 agcatctggt cggccgtgcc gcgctcaaac gccccgagga cggcctggac gttgcgagcg
39601 agctgttgga tggcgcgcaa ctggcgatgc gcgccgatac ccgtcccgtc cagggcctcc
39661 cccgtgagca gggcgatggc ctcggtggcc aggctgaagg cggcgttcag ggcccggcgg
39721 tcgataatct tggtcatgta attgtgtgtg ggttgctcga tggggtgcgg gccgtcgcgg
39781 gcaatcagcg gctggtggac ctcgaactgt acgcgcccct cgttcatgta ggccagctcc
39841 ggaaacttgg tacacacgca cgccaccgac aacccgagct ccagaaagcg cacgagcgac
39901 agggtgttgc aatacgaccc cagcagggcg tcgaactcga cgtcgtacag gctgtttgca
39961 tcggagcgca cgcgggaaaa aaaatcgaac aggcgtcgat gcgacgccac ctcgatcgtg
40021 ctaaggaggg acccggtcgg caccatggcc gcggcatacc ggtatcccgg agggtcgcgg
40081 ttgggagctg ccatggggtc gcgtggagat cggctggatc tagcgatatt tgcccgggga
40141 ggctaagatc caccccaacg cccggccacc cgtgtacgtg cccgacggcc caaggtccac

FIG. 8X

40201 cgaaagacac gacgggcccg gacccaaaaa ggcgggggat gctgtgtgag gggccgggtg
40261 tcggtcgggg gggaaaggca ccgggagaag gctgcggcct cgttccagga gaacccagtg
40321 tccccaacag acccggggac gtgggatccc aggccttata tacccccccc gccccacccc
40381 cgttagaacg cgacgggtgc attcaagatg gccctggtcc aaaagcgtgc caggaagaaa
40441 ttggcagagg cggcaaagct gtccgccgcc gccacccaca tcgaggcccc ggccgcgcag
40501 gctatcccca gggcccgtgt gcgcagggga tcggtgggcg gcagcatttg gttggtggcg
40561 ataaagtgga aaagcccgtc cggactgaag gtctcgtggg cggcggcgaa caaggcacac
40621 agggccgtgc ctcccaaaaa cacggacatc ccccaaaaca ctggcgccga caacggcaga
40681 cgatccctct tgatgttaac gtacaggagg agcgcccgca ccgcccacgt aacgtagtag
40741 ccgacgatgg cggccaggat acaggccggc gccaccaccc ttccggtcag cccgtaatac
40801 atgcccgctg ccaccatctc caacggcttc aggaccaaaa acgaccaaag gaacagaatc
40861 acgcgctttg aaaagaccgg ctgggtatgg ggcggaagac gcgagtatgc cgaactgaca
40921 aaaaaatcag aggtgccgta cgaggacaat gaaaactgtt cctccagcgg cagttctccc
40981 tcctcccccc cgaaggcggc ctcgtcgacc agatctcgat ccaccagagg aaggtcatcc
41041 cgcatggtca tggggtgtgc ggtggaggtg gggagaccga aaccgcaaag ggtcgcttac
41101 gtcagcagga tcccgagatc aaagacaccc gggttcttgc acaaacacca cccgggttgc
41161 atccgcggag gcgagtgttt tgataaggcc gttccgcgcc ttgatataac ctttgatgtt
41221 gaccacaaaa cccggaattt acgcctacgc cccaatgccc acgcaagatg aggtaggtaa
41281 cccccccgtg ggtgtgacgt tgcgtttagt tcattggagg ccaaggggaa aaatggggtg
41341 gggaggaaac ggaaaaccca gtaggccgtg tcgggaacac gcccggggtt gtcctcaaaa
41401 ggcagggtcc atactacgga agccgtcgtt gtattcgaga cctgcctgtg cgacgcacgt
41461 cggggttgcc tgtgtccggt tcggcccccа ccgcgtgcgg cacgcacgag gacgagtccg
41521 cgtgctttat tggcgttcca agcgttgccc tccagtttct gttgtcggtg ttcccccata
41581 cccacgccca catccaccgt agggggcctc tgggccgtgt tacgtcgccg cccgcgatgg
41641 agcttagcta cgccaccacc atgcactacc gggacgttgt gttttacgtc acaacggacc
41701 gaaaccgggc ctactttgtg tgcggggggt gtgtttattc cgtggggcgg ccgtgtgcct
41761 cgcagcccgg ggagattgcc aagtttggtc tggtcgttcg agggacaggc ccagacgacc
41821 gcgtggtcgc caactatgta cgaagcgagc tccgacaacg cggcctgcag gacgtgcgtc

FIG. 8Y

41881 ccattgggga ggacgaggtg tttctggaca gcgtgtgtct tctaaacccg aacgtgagct
41941 ccgagctgga tgtgattaac acgaacgacg tggaagtgct ggacgaatgt ctggccgagt
42001 actgcacctc gctgcgaacc agcccgggtg tgctaatatc cgggctgcgc gtgcgggcgc
42061 aggacagaat catcgagttg tttgaacacc caacgatagt caacgtttcc tcgcactttg
42121 tgtatacccc gtccccatac gtgttcgccc tggcccaggc gcacctcccc cggctcccga
42181 gctcgctgga ggccctggtg agcggcctgt ttgacggcat ccccgcccca cgccagccac
42241 ttgacgccca caacccgcgc acggatgtgg ttatcacggg ccgccgcgcc ccacgaccca
42301 tcgccgggtc gggggcgggg tcgggggcg cgggcgccaa gcgggccacc gtcagcgagt
42361 tcgtgcaagt caaacacatt gaccgcgtgg gccccgctgg cgtttcgccg gcgcctccgc
42421 caaacaacac cgactcgagt tccctggtgc ccggggccca ggattccgcc ccgcccggcc
42481 ccacgctaag ggagctgtgg tgggtgtttt atgccgcaga ccgggcgctg gaggagcccc
42541 gcgccgactc tggcctcacc cgcgaggagg tacgtgccgt acgtgggttc cgggagcagg
42601 cgtggaaact gtttggctcc gcgggggccc cgcgggcgtt tatcggggcc gcgttgggcc
42661 tgagcccccc ccaaaagctg gccgtttact actatatcat ccaccgagag aggcgcctgt
42721 ccccccttccc cgcgctagtc cggctcgtag gccggtacac acagcgccac ggcctgtacg
42781 tccctcggcc cgacgaccca gtcttggccg atgccatcaa cgggctggtt cgcgacgcgc
42841 tggcggccgg aaccacagcc gagcagctcc tcatgttcga ccttctcccc ccaaaggacg
42901 tgccggtggg aagcgacgtg caggccgaca gcaccgctct gctgcgcttt atagaatcgc
42961 aacgtctcgc cgtccccggg ggggtgatct cccccgagca cgtcgcgtac cttggtgcgt
43021 tcctgagcgt gctgtacgct ggccgcgggc gcatgtccgc agccacgcac accgcgcggc
43081 tgacaggggt gacctccctg gtgctagcgg tgggtgacgt ggaccgtctt tccgcgtttg
43141 accgcggagc ggcgggcgcg gccagccgca cgcgggccgc cgggtacctg gatgtgcttc
43201 ttaccgttcg tctcgctcgc tcccaacacg gacagtctgt gtaaaagacc ccaataaacg
43261 tatgtcgcta ctacacccct gtgtgtcaat ggacgcctct ccgggggggg gaagggaaag
43321 caaagagggg ctgggggagc ggcaccaccg gggcctgaac aaacaaacca cagacacggt
43381 tacagtttat tcggtcgggc ggagaaacgg ccgaagccac gcccacttta ttcgcgtctc
43441 caaaaaaacg ggacacttgt ccggagaacc tttaggatgc cagccagggc ggcggtaatc
43501 ataaccacgc ccagcgcaga ggcggccaga aacccgggcg caattgcggc cacgggctgc

FIG. 8Z

43561 gtgtcaaagg ctagcaaatg aatgacggtt ccgtttggaa atagcaacaa ggccgtggac
43621 ggcacgtcgc tcgaaaacac gcttggggcg ccctccgtcg gcccggcggc gatttgctgc
43681 tgtgtgttgt ccgtatccac cagcaacaca gacatgacct ccccggccgg ggtgtagcgc
43741 ataaacacgg cccccacgag ccccaggtcg cgctggtttt gggtgcgcac cagccgcttg
43801 gactcgatat cccgggtgga gccttcgcat gtcgcggtga ggtaggttag gaacagtggg
43861 cgtcggacgt cgacgccggt gagcttgtag ccgatccccc ggggcagagg ggagtgggtg
43921 acgacgtagc tggcgctgtg ggtgatgggt accaggatcc gtggctcgac gttggcagac
43981 tgcccccgc accgatgtga ggcctcaggg acgaaggcgc ggatcagggc gttgtagtgt
44041 gcccaacgcg tcagggtcga ggcgaggccg tgggtctgct gggccaggac ttcgaccggg
44101 gtctcggatc gggtggcttg agccagcgcg tccaggataa acacgctctc gtctagatca
44161 aagcgcaggg aggccgcgca tggcgaaaag tggtccggaa gccaaaagag ggttttctgg
44221 tggtcggccc gggccagcgc ggtccggagg tcggcgttgg tcgctgcggc gacgtcggac
44281 gtacacaggg ccgaggctat cagaaggctc cggcgggcgc gttcccgctg caccgccgag
44341 gggacgccag ccaagaacgg ctgccggagg acagccgagg cgtaaaatag cgcccggtgg
44401 acgaccgggg tggtcagcac gcggcccccc agaaactcgg catacagggc gtcgatgaga
44461 tgggctgcgc tgggcgccac tgcgtcgtac gccgaggggc tatccagcac gaaggccagc
44521 tgatagccca gcgcgtgtaa tgccaagctc tgttcgcgct ccagaatctc ggccaccagg
44581 tgctggagcc gagcctctag ctgcaggcgg gccgtgggat ccaagactga cacattaaaa
44641 aacacagaat ccgcggcaca gcccgcggcc ccgcgggcgg ccaacccggc aagcgcgcgc
44701 gagtgggcca aaaagcctag caggtcggag aggcagaccg cgccgtttgc gtgggcggcg
44761 ttcacgaaag caaaacccga cgtcgcgagc agccccgtta ggcgccagaa gagaggggcg
44821 cgcgggccct gctcggcgcc cgcgtccccc gagaaaaact ccgcgtatgc ccgcgacagg
44881 aactgggcgt agttcgtgcc ctcctccggg tagccgccca cgcggcggag ggcgtccagc
44941 gcggagccgt tgtcggcccg cgtcagggac cctaggacaa agacccgata ccgggggccg
45001 cccgggggcc cgggaagagc ccccgggggg ttttcgtccg cggggtcccc gacccgatct
45061 agcgtctggc ccgcggggac caccatcact tccaccggag ggctgtcgtg catggatatc
45121 acgagcccca tgaattcccg cccgtagcgc gcgcgcacca gcgcggcatc gcacccgagc
45181 accagctccc ccgtcgtcca gatgcccacg ggccacgtcg aggccgacgg ggagaaatac

FIG. 8AA

45241 acgtacctac ctggggatct caacaggccc cgggtggcca accaggtcgt ggacgcgttg
45301 tgcaggtgcg tgatgtccag ctccgtcgtc gggtgccgcc gggccccaac cggcggtcgg
45361 gggggcggtg tatcacgcgg cccgctcggg tggctcgccg tcgccacgtt gtctccccgc
45421 gggaacgtca gggcctcggg gtcagggacg gccgaaaacg ttacccaggc ccgggaacgc
45481 agcaacacgg aggcggctgg attgtgcaag agacccttaa ggggggcgac cgaggggga
45541 ggctgggcgg tcggctcgac cgtggtgggg gcgggcaggc tcgcgttcgg gggccggccg
45601 agcaggtagg tcttcgggat gtaaagcagc tggccggggt cccgcggaaa ctcggccgtg
45661 gtgaccaata caaaacaaaa gcgctcctcg taccagcgaa gaaggggcag agatgccgta
45721 gtcaggttta gttcgtccgg cggcgccaga aatccgcgcg gtggtttttg ggggtcgggg
45781 gtgtttggca gccacagacg cccggtgttc gtgtcgcgcc agtacatgcg gtccatgccc
45841 aggccatcca aaaaccatgg gtctgtctgc tcagtccagt cgtggacctg accccacgca
45901 acgcccaaaa taataacccc cacgaaccat aaaccattcc ccatggggga ccccgtccct
45961 aacccacggg gcccgtggct atggcagggc ttgccgcccc gacgttggct gcgagccctg
46021 ggccttcacc cgaacttggg gggtggggtg gggaaaagga agaaacgcgg gcgtattggc
46081 cccaatgggg tctcggtggg gtatcgacag agtgccagcc ctgggaccga accccgcgtt
46141 tatgaacaaa cgacccaaca cccgtgcgtt ttattctgtc tttttattgc cgtcatagcg
46201 cgggttactt ccggtattgt ctccttccgt gtttcagtta gcctcccccа tctcccgggc
46261 aaacgtgcgc gccaggtcgc agatcgtcgg tatggagccg ggggtggtga cgtgggtctg
46321 gaccatcccg gaggtaagtt gcagcagggc gtcccggcag ccggcgggcg attggtcgta
46381 atccaggata aagacgtgca tgggacggag gcgtttggcc aagacgtcca aggcccaggc
46441 aaacacgtta tacaggtcgc cgttgggggc cagcaactcg ggggcccgaa acagggtaaa
46501 taacgtgtcc ccgatatggg gttgtgggcc cgcgttgctc tggggctcgg cacccctgggg
46561 cggcacggcc gtccccgaaa gctgtcccca atcctcccgc cacgacccgc cgccctgcag
46621 ataccgcacc gtattggcaa gcagctcgta aacgcggcga atcgcggcca acatagccag
46681 gtcaagccgc tcgccggggc gctggcgttt ggccaggcgg tcgatgtgtc tgtcctccgg
46741 aagggccccc aacacgatgt ttgtgccggg caaggtcggc gggatgaggg ccacgaacgc
46801 cagcacggcc tgggggggtca tgctgcccat aaggtatcgc gcggccgggt aacacaggag
46861 ggcggcgatg ggatggcggt cgaagatgag ggtgagggcc gggggcgggg catgtgagct

FIG. 8AB

46921 cccagcctcc cccccgatat gaggagccag aacggcgtcg gtcacggcat aaggcatgcc
46981 cattgttatc tgggcgcttg tcattaccac cgccgcgtcc ccggccgata tctcaccctg
47041 gtcgaggcgg tgttgtgtgg tgtagatgtt cgcgattgtc tcggaagccc ccaacacccg
47101 ccagtaagtc atcggctcgg gtacgtagac gatatcgtcg cgcgaaccca gggccaccag
47461 cagttgcgtg gtggtggttt tccccatccc gtggggaccg tctatataaa cccgcagtag
47221 cgtgggcatt ttctgctcca ggcggacttc cgtggctttt tgctgccggc gagggcgcaa
47281 cgccgtacgt cggttgttat ggccgcgaga acgcgcagcc tggtcgaacg cagacgcgtg
47341 ttgatggcag gggtacgaag ccatacgcgc ttctacaagg cgctggccga agaggtgcgg
47401 gagtttcacg ccaccaagat ctgcggcacg ctgttgacgc tgttaagcgg gtcgctgcag
47161 ggtcgctcgg tattcgaggc cacacgcgtc accttaatat gcgaagtgga cctgggaccg
47521 cgccgccccg actgcatctg cgtgttcgaa ttcgccaatg acaagacgct gggcggggtt
47581 tgtgtcatca tagaactaaa gacatgcaaa tatatttctt ccggggacac cgccagcaaa
47641 cgcgagcaac gggccacggg gatgaagcag ctgcgccact ccctgaagct cctgcagtcc
47701 ctcgcgcctc cgggtgacaa gatagtgtac ctgtgccccg tcctggtgtt tgtcgcccaa
47761 cggacgctcc gcgtcagccg cgtgacccgg ctcgtcccgc agaaggtctc cggtaatatc
47821 accgcagtcg tgcggatgct ccagagcctg tccacgtata cggtccccat tgagcctagg
47881 acccagcgag cccgtcgccg ccgcggcggc gccgcccggg ggtctgcgag cagaccgaaa
47941 aggtcacact ctggggcgcg cgacccgccc gagtcagcgg cccgccagtt accacccgcc
48001 gaccaaaccc ccgcctccac ggagggcggg ggggtgctta agaggatcgc ggcgctcttc
48061 tgcgtgcccg tggccaccaa gaccaaaccc cgagccgcct ccgaatgaga gtgtttcgtt
48121 ccttcccccc cccccgcgt cagacaaacc ctaaccaccg cttaagcggc ccccgcgagg
48181 tccgaagact catttggatc cggcgggagc cacccgacaa cagcccccgg gttttcccac
48241 gccagacgcc ggtccgctgt gccatcgcgc cccctcatcc caccccccat cttgtcccca
48301 aataaaacaa ggtctggtag ttaggacaac gaccgcagtt ctcgtgtgtt attttcgctc
48361 tccgcctctc gcagatggac ccgtactgcc catttgacgc tctggacgtc tgggaacaca
48421 ggcgcttcat agtcgccgat tcccgaaact tcatcacccc cgagttcccc cgggactttt
48481 ggatgtcgcc cgtctttaac ctcccccggg agacggcggc ggagcaggtg gtcgtcctac
48541 aggcccagcg cacagcggct gccgctgccc tggagaacgc cgccatgcag gcggccgagc

FIG. 8AC

48601 tccccgtcga tatcgagcgc cggttacgcc cgatcgaacg gaacgtgcac gagatcgcag
48661 gcgccctgga ggcgctggag acggcggcgg ccgccgccga agaggcggat gccgcgcgcg
48721 gggatgagcc ggcgggtggg ggcgacgggg gggcgccccc gggtctggcc gtcgcggaga
48781 tggaggtcca gatcgtgcgc aacgacccgc cgctacgata cgacaccaac ctccccgtgg
48841 atctgctaca catggtgtac gcgggccgcg gggcgaccgg ctcgtcgggg gtggtgttcg
48901 ggacctggta ccgcactatc caggaccgca ccatcacgga ctttcccctg accacccgca
48961 gtgccgactt tcgggacggc cggatgtcca agaccttcat gacggcgctg gtcctgtccc
49021 tgcagtcgtg cggccggctg tatgtgggcc agcgccacta ttccgccttc gagtgcgccg
49081 tgttgtgtct ctacctgctg taccgaaaca cgcacggggc cgccgacgat agcgaccgcg
49141 ctccggtcac gttcggggat ctgctgggcc ggctgccccg ctacctggcg tgcctggccg
49201 cggtgatcgg gaccgagggc ggccggccac agtaccgcta ccgcgacgac aagctccccа
49261 agacgcagtt cgcggccggc gggggccgct acgaacacgg agcgctggcg tcgcacatcg
49321 tgatcgccac gctgatgcac cacggggtgc tcccggcggc cccgggggac gtcccccggg
49381 acgcgagtac ccacgttaac cccgacggcg tggcgcacca cgacgacata aaccgcgccg
49441 ccgccgcgtt cctcagccgg ggccacaacc tattcctgtg ggaggaccag actctgctgc
49501 gggcaaccgc gaacaccata acggccctgg gcgttatcca gcggctcctc gcgaacggca
49561 acgtgtacgc ggaccgcctc aacaaccgcc tgcagctggg catgctgatc cccggagccg
49621 tcccttcgga ggccatcgcc cgtggggcct ccgggtccga ctcgggggcc atcaagagcg
49681 gagacaacaa tctggaggcg ctatgtgcca attacgtgct tccgctgtac cgggccgacc
49741 cggcggtcga gctgacccag ctgtttcccg gcctggccgc cctgtgtctt gacgcccagg
49801 cggggcggcc ggtcgggtcg acgcggcggg tggtggatat gtcatcgggg gcccgccagg
49861 cggcgctggt gcgcctcacc gccctggaac tcatcaaccg cacccgcaca aaccccaccc
49921 ccgtgggggа ggttatccac gcccacgacg ccctggcgat ccaatacgaa cagggcttg
49981 gcctgctggc gcagcaggca cgcattggct tgggctccaa caccaagcgt ttctccgcgt
50041 tcaacgttag cagcgactac gacatgttgt acttttatg tctggggttc attccacagt
50101 acctgtcggc ggtttagtgg gtggtgggcg agggggggagg gggcattagg gagaaagaac
50161 aagagcctcc gttgggtttt ctttgtgcct gtactcaaaa ggtcataccc cgtaaacggc
50221 gggctccagt cccggcccgg tggttggcgt gaacgcaacg gcgggagctg ggttagcgtt

FIG. 8AD

50281 tagtttagca ttcgctctcg cctttccgcc cgcccccga ccgttgcgcc ttttttttcg
50341 tccaccaaag tctctgtggg tgcgcgcatg gcagccgatg ccccgggaga ccggatggag
50401 gagcccctgc ccgacagggc cgtgcccatt tacgtggctg ggtttttggc cctgtatgac
50461 agcggggact cgggcgagtt ggcattggat ccggatacgg tgcgggcggc cctgcctccg
50521 gataacccac tcccgattaa cgtggaccac cgcgctggct gcgaggtggg gcgggtgctg
50581 gccgtggtcg acgacccccg cgggccgttt tttgtggggc tgatcgcctg cgtgcagctg
50641 gagcgcgtcc tcgagacggc cgccagcgct gcgattttcg agcgccgcgg gccgccgctc
50701 tcccgggagg agcgcctgtt gtacctgatc accaactacc tgccctcggt ctccctggcc
50761 acaaaacgcc tgggggggcga ggcgcacccc gatcgcacgc tgttcgcgca cgtcgcgctg
50821 tgcgcgatcg ggcggcgcct cggcactatc gtcacctacg acaccggtct cgacgccgcc
50881 atcgcgccct ttcgccacct gtcgccggcg tctcgcgagg gggcgcggcg actggccgcc
50941 gaggccgaga tcgcgctgtc cgggcgcacc tgggcgcccg gcgtggaggc gctgacccac
51001 acgctgcttt ccaccgccgt taacaacatg atgctgcggg accgctggag cctggtggcc
51061 gagcggcggc ggcaggccgg gatcgccgga cacacctacc tccaggcgag cgaaaaattc
51121 aaaatgtggg gggcggagcc tgtttccgcg ccggcgcgcg ggtataagaa cggggccccg
51181 gagtccacgg acataccgcc cggctcgatc gctgccgcgc cgcagggtga ccggtgccca
51241 atcgtccgtc agcgcggggt cgccttgtcc ccggtactgc cccccatgaa ccccgttccg
51301 acatcgggca ccccggcccc cgcgccgccc ggcgacggga gctacctgtg gatcccggcc
51361 tcccattaca accagctcgt cgccggccat gccgcgcccc aaccccagcc gcattccgcg
51421 tttggtttcc cggctgcggc gggggccgtg gcctatgggc ctcacggcgc gggtctttcc
51481 cagcattacc ctccccacgt cgcccatcag tatcccgggg tgctgttctc gggacccagc
51541 ccactcgagg cgcagatagc cgcgttggtg ggggccatag ccgcggaccg ccaggcgggc
51601 ggtcagccgg ccgcgggaga ccctggggtc cgggggtcgg gaaagcgtcg ccggtacgag
51661 gcgggggccgt cggagtccta ctgcgaccag gacgaaccgg acgcggacta cccgtactac
51721 cccgggggagg ctcgaggcgg gccgcgcggg gtcgactctc ggcgcgcggc ccgccagtct
51781 cccgggacca acgagaccat cacggcgctg atgggggcgg tgacgtcttt gcagcaggaa
51841 ctggcgcaca tgcgggctag gaccagcgcc ccctatggga tgtacacgcc ggtggcgcac
51901 tatcgccctc aggtgggggga gccggaacca acaacgaccc acccggccct ttgtcccccg

FIG. 8AE

51961 gaggccgtgt atcgcccccc accacacagc gccccctacg gtcctcccca gggtccggcg
52021 tcccatgccc ccactccccc gtatgcccca gctgcctgcc cgccaggccc gccaccgccc
52081 ccatgtcctt ccacccagac gcgcgccccc ctaccgacgg agcccgcgtt cccccccgcc
52141 gccaccggat cccaaccgga ggcatccaac gcggaggccg gggcccttgt caacgccagc
52201 agcgcagcac acgtggacgt tgacacggcc cgcgccgccg atttgttcgt ctctcagatg
52261 atgggggccc gctgattcgc cccggtcttt ggtaccatgg gatgtcttac tgtatatctt
52321 tttaaataaa ccaggtaata ccaaagaaga cccattggtg tatgttcttt ttttattggg
52381 aggcgcgggt aggcgggtag ctttacaatg caaaagcctt cgacgtggag gaaggcgtgg
52441 gggggaatcg gcactgacca agggggtccg ttttgtcacg ggaaaggaaa gaggaaacag
52501 gccgcggaca cccgggggag tttatgtgtt ccctttttctt tcttcccaca cacacaaaag
52561 gcgtaccaaa caaacaaacc aaaagatgca catgcggttt aacacccgtg gtttttattt
52621 acaacaaacc ccccgtcaca ggtcgtcctc gtcggcgtca ccgtctttgt tgggaacttg
52681 ggtgtagttg gtgttgcggc gcttgcgcat gaccatgtcg gtgaccttgg cgctgagcag
52741 cgcgctcgtg ccctttcttct tggccttgtg ttccgtgcgc tccatggcag acaccagggc
52801 catgtaccgt atcatctccc gggcctcggc tagcttggcc tcgtcaaagt cgccgccctc
52861 ctcgccctcc ccggacgcgt ccgggttggt ggggttcttg agctccttgg tggttagcgg
52921 gtacagggcc ttcatggggt tgctctgcag ccgcatgacg tagcgaaagg cgaagaaggc
52981 cgccgccagg ccggccagga ccaacagacc cacggccagc gccccaaagg ggttggacat
53041 gaaggaggac acgcccgaca cggccgatac cacgccgccc acgatgccca tcaccacctt
53101 gccgaccgcg cgccccaggt cgcccatccc ctcgaagaac gcgcccaggc ccgcaaacat
53161 ggcggcgttg gcgtcggcgt ggatgaccgt gtcgatgtcg gcgaagcgca ggtcgtgcag
53221 ctggttgcgg cgctggacct ccgtgtagtc cagcaggccg ctgtccttga tctcgtggcg
53281 ggtgtacacc tccaggggga caaactcgtg atcctccagc atggtgatgt tgaggtcgat
53341 gaaggtgctg acggtggtga tgtcggcgcg gctcagctgg tgggagtacg cgtactcctc
53401 gaagtacacg tagccccccac cgaaggtgaa gtagcgccgg tgtcccacgg tacacggctc
53461 gatcgcatcg cgcgtcagcc gcagctcgtt gttctccccc agctgccccct cgaccaacgg
53521 gccctggtct tcgtaccgaa agctgaccag ggggcggctg tagcaggccc cgggccgcga
53581 gctgatgcgc atcgagtttt ggacgatcac gttgtccgcg gcgaccggca cgcacgtgga

FIG. 8AF

53641 gacggccatc acgtcgccga gcatccgcgc gctcacccgc cggcccacgg tggccgaggc
53701 gatggcgttg gggttcagct tgcgggcctc gttccacagg gtcagctcgt gattctgcag
53761 ctcgcaccac gcgatggcaa cgcggcccaa catatcgttg acatggcgct gtatgtggtt
53821 gtacgtaaac tgcagcctgg cgaactcgat ggaggaggtg gtcttgatgc gctccacgga
53881 cgcgttggcg ctggccccgg gcggcggggg cgtggggttt gggggcttgc ggctctgctc
53941 tcggaggtgt tcccgcacgt acagctccgc gagcgtgttg ctgagaaggg gctggtacgc
54001 gatcagaaag ccccccattgg ccaggtagta ctgcggctgg cccaccttga tgtgcgtcgc
54061 gttgtacctg cgggcgaaga tgcggtccat ggcgtcgcgg gcgtccttgc cgatgcagtc
54121 ccccaggtcc acgcgcgaga gcgggtactc ggtcaggttg gtggtgaagg tggtggatat
54181 ggcgtcggag gagaatcgga aggagccgcc gtactcggag cgcagcatct cgtccacctc
54241 ctgccacttg gtcatggtgc agaccgacgg gcgctttggc acccagtccc aggccacggt
54301 gaacttgggg gtcgtgagca ggttccgggt ggtcggcgcc gtggcccggg ccttggtggt
54361 gaggtcgcgc gcgtagaagc cgtcgacctg cttgaagcgg tcggcggcgt agctggtgtg
54421 ttcggtgtgc gacccctccc ggtagccgta aaacggggac atgtacacaa agtcgccagt
54481 cgccagcaca aactcgtcgt acgggtacac cgagcgcgcg tccacctcct cgacgatgca
54541 gtttaccgtc gtcccgtacc ggtggaacgc ctccacccgc gagggttgt acttgaggtc
54601 ggtggtgtgc cagccccggc tcgtgcgggt cgcggcgttg gccggtttca gctccatgtc
54661 ggtctcgtgg tcgtcccggt gaaacgcggt ggtctccagg ttgttgcgca cgtacttggc
54721 cgtggaccga cagacccct tggcgttgat cttgtcgatc acctcctcga aggggacggg
54781 ggcgcggtcc tcaaagatcc ccataaactg ggagtagcgg tggccgaacc acacctgcga
54841 aacggtgacg tctttgtagt acatggtggc cttgaacttg tacggggcga tgttctcctt
54901 gaagaccacc gcgatgccct ccgtgtagtt ctgaccctcg ggccgggtcg ggcagcggcg
54961 cggctgctcg aactgcacca ccgtggcgcc cgtgggggt gggcacacgt aaaagtttgc
55021 atcggtgttc tccgccttga tgtcccgcag gtgctcgcgc agggtggcgt ggcccgcggc
55081 gacggtcgcg ttgtcgccgg cggggcgcgg cggctttggg ggtttcggtt ttttgttctt
55141 cttcggtttc gggtcccccg ttgggggggc gccaggggcg ggcggcgccg gagtggcagg
55201 gcccccgttc gccgcctggg tcgcggccgc gaccccaggc gtgccgggggg aactcggagc
55261 cgccgacacc accaggaccc ccagcgtcaa ccccaagagc gcccatacga cgaaccaccg

FIG. 8AG

55321 gcgccccgc gcgggggcgc cctggcgcat ggcgggacta cgggggcccg tcgtgccccc
55381 cgtcaggtag cctgggggcg aggtgctgga ggaccgagta gaggatcgag aaaacgtctc
55441 ggtcgtagac cacgaccgac cgggggccga tacagccgtc gggggcgctc tcgacgatgg
55501 ccaccagcgg acagtcggag tcgtacgtga gatatacgcc gggcgggtaa cggtaacgac
55561 cttcggaggt cgggcggctg cagtccgggc ggcgcaactc gagctccccg caccggtaga
55621 ccgaggcaaa gagtgtggtg gcgataatca gctcgcgaat atatcgccag gcggcgcgct
55681 gagtgggcgt tattccggaa atgccgtcaa aacagtaaaa cctctgaaat tcgctgacgg
55741 cccaatcagc acccgagccc cccgccccca tgatgaaccg ggcgagctcc tccttcaggt
55801 gcggcaggag ccccacgttc tcgacgctgt aatacagcgc ggtgttgggg ggctgggcga
55861 agctgtgggt ggagtgatca aagaggggcc cgttgacgag ctcgaagaag cgatgggtga
55921 tgctggggag cagggccggg tccacctggt gtcgcaggag agacgctcgc atgaaccggt
55981 gcgcgtcgaa cacgcccggc gccgagcggt tgtcgatgac cgtgcccgcg cccgccgtca
56041 gggcgcagcc gcgcgcgcgc gccgcaaagc cgttggcgac cgcggcgaac gtcgcgggca
56101 gcacctcgaa gtggacgctg acccgcagca tcttctcgag ctccccgcgc tgctcgcgga
56161 cgcagcgccc caggctggcc aacgaccgct tcgtcaggcg gtccgcgtac agccgccgtc
56221 gctcccgcac gtccgcggcc gcttgcgtgg cgatgtcccc ccacgtctcg ggcccctgcc
56281 ccccgggccc gcggcgacgg tcttcgtcct cgccccgcc cccgggagct cccaacccca
56341 gtgcccctte ctctacggcg acacggtccc cgtcgtcgtc ggggcccgcg ccgcccttgg
56401 gcgcgtccgc cgcgcccccc gcccccatgc gcgccagcac gcgacgcagc gcctcctcgt
56461 cgcactgttc ggggctgacg aggcgccgca agagcggcgt cgtcaggtgg tggtcgtagc
56521 acgcgcggat gagcgcctcg atctgatcgt cgggtgacgt ggcctgaccg ccgattatta
56581 gggcgtccac catatccagc gccgccaggt ggctcccgaa cgcgcgatcg aaatgctccg
56641 cccgccgccc gaacagcgcc agttccacgg ccaccgcggc ggtctcctgc tgcaactcgc
56701 gccgcgccag cgcggtcagg ttgctggcaa acgcgtccat ggtggtctgg ccggcgcggt
56761 cgccggacgc gagccagaat cgcaattcgc tgatggcgta caggccgggc gtggtggcct
56821 gaaacacgtc gtgcgcctcc agcagggcgt cggcctcctt gcggaccgag tcattctcgg
56881 gcgacgggtg gggctgcccg tcgccccccg cggtccgggc cagcgcatgg tccaacacgg
56941 agagcgcccg cgcgcggtcg gcgtccgaca gcccggcggc gtggggcagg taccgccgca

FIG. 8AH

57001 gctcgttggc gtccagccgc acctgcgcct gctgggtgac gtggttacag atacggtccg
57061 ccaggcggcg ggcgatcgtc gccccctggt tcgccgtcac acacagttcc tcgaaacaga
57121 ccgcgcaggg gtgggacggg tcgctaagct ccgggggggac gataaggccc gaccccaccg
57181 ccccaccat aaactcccga acgcgctcca gcgcggcggt ggcgccgcgc gaggggggtga
57241 tgaggtggca gtagtttagc tgctttagaa agttctcgac gtcgtgcagg aaacacagct
57301 ccatatggac ggtcccgcca tacgtatcca gcctgacccg ttggtgatac ggacagggtc
57361 gggccaggcc catggtctcg gtgaaaaacg ccgcgacgtc tcccgcggtc gcgaacgtct
57421 ccaggctgcc caggagccgc tcgccctcgc gccacgcgta ctctagcagc aactccaggg
57481 tgaccgacag cggggtgaga aaggccccgg cctgggcctc caggcccggc ctcagacgac
57541 gccgcagcgc ccgcacctga agcgcgttca gcttcagttg ggggagcttc ccccgtccga
57601 tgtgggggtc gcaccgccgg agcagctcta tctgaaacac ataggtctgc acctgcccga
57661 gcagggctaa caacttttga cgggccacgg tgggctcgga caccggggcg gccatctcgc
57721 ggcgccgatc tgtaccgcgg ccggagtatg cggtggaccg aggcggtccg tacgctaccc
57781 ggcgtctggc tgagccccgg ggtccccctc ttcggggcgg cctcccgcgg gcccgccgac
57841 cggcaagccg ggagtcggcg gcgcgtgcgt ttctgctcta ttcccagaca ccgcggagag
57901 gaatcacggc ccgcccagag atatagacac ggaacacaaa caagcacgga tgtcgtagca
57961 ataatttatt ttacacacat tccccgcccc gccctaggtt ccccccaccc caaccccctca
58021 cagcatatcc aacgtcaggt ctccctttt gtcgggggggc ccctccccaa acgggtcatc
58081 cccgtggaac gcccgtttgc ggccggcaaa tgccggtccc ggggccccccg ggccgccgaa
58141 cggcgtcgcg ttgtcgtcct cgcagccaaa atccccaaag ttaaacacct ccccggcgtt
58201 gccgagttgg ctgactaggg cctcggcctc gtgcgccacc tccagggccg cgtccgtcga
58261 ccactcgccg ttgccgcgct ccagggcacg cgcggtcagc tccatcatct cctcgcttag
58321 gtactcgtcc tccaggagcg ccagccagtc ctcgatctgc agctgctggg tgcggggccc
58381 caggcttttc acggtcacca cgaacacgct actggcgacg gccgcccccgc cctcggagat
58441 aatgccccgg agctgctcgc acagcgagct ttcgtgcgct ccgccgccga ggctcgaggc
58501 cgcgcacaca aacccggccc ggggacaggc caggacgaac ttgcgggtgc ggtcaaaaat
58561 aaggagcggg cacgcgtttt tgccgcccat caggctggcc cagttcccgg cctgaaacac
58621 acggtcgttg ccggccatgc cgtagtactt gctgatgctc aaccccaaca cgaccatggg

FIG. 8AI

58681 gcgcgccgcc atgacgggcc gcagcaggtt gcagctggcg aacatggacg tccacgcgcc
58741 cggatgcgcg tccacggcgt ccatcagcgc gcgggccccg gcctccaggc ccgccccgcc
58801 ctgcgcggac cacgcggccg cagcctgcac gctgggggga cggcgggacc ccgcgatgat
58861 ggccgtaagg gtgttgatga agtacgtcga gtgatcgcag taccgcagaa tctggtttgc
58921 catgtagtac atcgccagct cgctcacgtt gttgggggcc aggttaataa agtttatcgc
58981 gccgtagtcc agggaaaact ttttaatgaa cgcgatggtc tcgatgtcct cgcgcgacag
59041 gagccgggcg ggaagctggt tgcgttggag ggccgtccag aaccactgcg ggttcggctg
59101 gttggacccc gggggcttgc cgttggggaa gatggccgcg tggaactgct tcagcagaaa
59161 gcccagcggt ccgaggagga tgtccacgcg cttgtcgggc ttctggtagg cgctctggag
59221 gctggcgacc cgcgccttgg cggcctcgga cgcgttggcg ctcgcgcccg cgaacaacac
59281 gcggctcttg acgcgcagct ccttgggaaa ccccagggtc acgcgggcaa cgtcgccctc
59341 gaagctgctc tcggcggggg ccgtctggcc ggccgttagg ctgggggcgc agatagccgc
59401 cccctccgag agcgcgaccg tcagcgtttt ggccgacaga aacccgttgt taaacatgtc
59461 catcacgcgc cgccgcagca ccggttggaa ttgattgcga aagttgcgcc cctcgaccga
59521 ctgccccggcg aacaccccgt ggcactggct cagggccagg tcctgataca cggcgaggtt
59581 ggatcgccgc ccgagaagct gaagcagggg gcatggcccg cacgctaccg ggtccagcgt
59641 cagggacatg gcgtggttgg cctcgcccag accgtcgcga aacttgaagt tcctcccctc
59701 caccaggttg cgcatcagct gctccacctc gcggtccacg acctgcctga cgttgttcac
59761 caccgtatgc agggcctcgc ggttggtgat gatggtctcc agccgcccca tggccgtggg
59821 gaccgcctgg tccacgtact gcagggtctc gagttcggcc atgacgcgct cggtcgccgc
59881 gcggtacgtc tcctgcatga tggtccgggc ggtctcggat ccgtccgcgc gcttcagggc
59941 cgagaaggcg gcgtagtttc ccagcacgtc gcagtcgctg tacatgctgt tcatggtccc
60001 gaagacgccg atggctccgc gggcggcgct ggcgaacttg ggatggcgcg cccggaggcg
60061 catgagcgtc gtgtgtacgc aggcgtggcg cgtgtcgaag gtgcacaggt tgcagggcac
60121 gtcggtctgg ttggagtccg cgacgtatcg aaacacgtcc atctcctggc gcccgacgat
60181 cacgccgccg tcgcagcgct ccaggtaaaa cagcatcttg gccagcagcg ccggggaaaa
60241 cccacacagc atggccaggt gctcgccggc aaattcctgg gttccgccga cgaggggcgc
60301 ggtgggccga ccctcgaacc cgggcaccac gtgtccctcg cggtccacct gtgggttggc

FIG. 8AJ

60361 cgccacgtgg gtcccgggca cgaggaagaa gcggtaaaag gagggtttgc tgtggtcctt
60421 tgggtccgcc ggaccggcgt cgtccacctc ggtgagatgg agggccgagt tggtgctaaa
60481 taccatggcc cccacgagtc ccgcggcgcg cgccaggtac gccccgacgg cgttggcgcg
60541 ggccgcggcc gtgtcctggc cctcgcacag cggccatgcg gagatgtcgg tgggcggctc
60601 gtcgaagacg gccatcgaca cgatagactc gagggccagg gcggcgtctc cggccatgac
60661 ggaggccagg cgctgttcga acccgcccgc cgggcccttg ccgccgccgt cgcgcccacc
60721 ccgcggggtc ttaccctggc tggcttcgaa ggccgtgaac gtaatgtcgg cggggagggc
60781 ggcgccctcg tggttttcgt caaacgccag gtgggcggcc gcgcgggcca cggcgtccac
60841 gtttcggcat cgcagtgcca cggcggcggg tcccacgacc gcctcgaaca ggaggcggtt
60901 gaggggcgcg ttaaaaaacg gaagcgggta ggtaaaattc tccccgatcg atcggtggtt
60961 ggcgttgaac ggctcggcga tgacccggct aaaatccggc atgaacagct gcaacggata
61021 cacgggtatg cggtgcacct ccgccccgcc tatggttacc ttgtccgagc ctcccaggtg
61081 cagaaaggtg ttgttgatgc acacggcctc cttgaagccc tcggtaacga ccagatacag
61141 gagggcgcgg tccgggtcca ggccgaggcg ctcacacagc gcctcccccg tcgtctcgtg
61201 tttgaggtcg ccgggccggg gggtgtagtc cgaaaagcca aaatggcggc gtgcccgctc
61261 gcagagtcgc gtcaggtttg gggcctgggt gttggggtcc aggtgccggc cgccgtgaaa
61321 gacgtacacg gacgagctgt agtgcgatgg cgtcagtttc agggacaccg cggtaccccc
61381 gagccccgtc gtgcgagaac ccacgaccac ggctacgttg gcctcaaagc cgctctccac
61441 ggtcaggccc acgaccaggg gcgccacggc gacgtcggca tcgccgctgc gcgccgacag
61501 taacgccaga agctcgatgc cttcggatgg acacgcgcga gcgtacacgt atcccagggg
61561 ccccgggggg gccttgatgg tggttgccgt cttgggcttt gtctccatgt cctcctggca
61621 atcggtccgc aaacggaggt aatcccggca cgacgacgga cgcccgacga ggtatgtctc
61681 ccgagcgtca aaatccgggg ggggcggcga cggtcaaggg gagggtggga gaccggggtt
61741 ggggaatgaa tccctaccct tcaccgacaa ccccccgggta accacggggt gccgatgaac
61801 cccggcggct ggcaacgcgg ggtccctgcg agaggcacag atgcttacgg tcaggtgctc
61861 cgggccgggt gcgtctgata tgcggttggt atatgtacac tttacctggg ggcgtgccgg
61921 accgcaccag cccctcccac acccccgcgcg tcatcagccg gtgggcgnnn nnnnnnnnnn
61981 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn

FIG. 8AK

62041 nnnnnnnnnn nnnnnnnnnn nnnnnnnnnt ttttataata gcggccacgc ccaccggcta
62101 cgtcacgctc ctgtcggccg ccggcggtcc ataagcccgg ccggccgggc cgacgcgaat
62161 aaaccgggcc gccggccggg gcgccgcgca gcagctcgcc gcccggatcc gccagacaaa
62221 caaggccctt gcacatgccg gcccgggcga gcctgggggt ccggtaattt tgccatccca
62281 cccaagcggc ttttggggtt tttcctcttc cccctcccc acctcccccc tctttagggg
62341 ttcgggtggg aacaaccgcg atgttttccg gtggcggcgg cccgctgtcc cccggaggaa
62401 agtcggcggc cagggcggcg tccgggtttt ttgcgcccgc cggccctcgc ggagccggcc
62461 ggggaccccc gccttgtttg aggcaaaact tttacaaccc ctacctcgcc ccagtcggga
62521 cgcaacagaa gccgaccggg ccaacccagc gccatacgta ctatagcgaa tgcgatgaat
62581 ttcgattcat cgccccgcgg gtgctggacg aggatgcccc cccggagaag cgcgccgggg
62641 tgcacgacgg tcacctcaag cgcgcccccа aggtgtactg cgggggggac gagcgcgacg
62701 tcctccgcgt cgggtcgggc ggcttctggc cgcggcgctc gcgcctgtgg ggcggcgtgg
62761 accacgcccc ggcggggttc gaccccaccg tcaccgtctt tcacgtgtat gacatcctgg
62821 agaacgtgga gcacgcgtac ggcatgcgcg cggcccagtt ccacgcgcgg tttatggacg
62881 ccatcacacc gacggggacc gtcatcacgc tcctgggcct gactccggaa ggccaccggg
62941 tggccgttca cgtttacggc acgcggcagt acttttacat gaacaaggag gaggttgaca
63001 ggcacctaca atgccgcgcc ccacgagatc tctgcgagcg catggccgcg gccctgcgcg
63061 agtccccggg cgcgtcgttc cgcggcatct ccgcggacca cttcgaggcg gaggtggtgg
63121 agcgcaccga cgtgtactac tacgagacgc gccccgctct gttttaccgc gtctacgtcc
63181 gaagcgggcg cgtgctgtcg tacctgtgcg acaacttctg cccggccatc aagaagtacg
63241 agggtggggt cgacgccacc acccggttca tcctggacaa ccccgggttc gtcaccttcg
63301 gctggtaccg tctcaaaccg ggccggaaca acacgctagc ccagccgcgg gccccgatgg
63361 ccttcgggac atccagcgat gtcgagttta actgtacggc ggacaacctg gccatcgagg
63421 ggggcatgag cgacctaccg gcatacaagc tcatgtgctt cgatatcgaa tgcaaggcgg
63481 gggggagga cgagctggcc tttccggtgg ccgggcaccc ggaggacctg gtcatccaga
63541 tatcctgtct gctctacgac ctgtccacca ccgccctgga gcacgtcctc ctgttttcgc
63601 tcggttcctg cgacctcccc gaatcccacc tgaacgagct ggcggccagg ggcctgccca
63661 cgcccgtggt tctggaattc gacagcgaat tcgagatgct gttggccttc atgaccctg

FIG. 8AL

63721 tgaaacagta cggccccgag ttcgtgaccg ggtacaacat catcaacttc gactggccct
63781 tcttgctggc caagctgacg gacatttaca aggtcccccc ggacgggtac ggccgcatga
63841 acggccgggg cgtgtttcgc gtgtgggaca taggccagag ccacttccag aagcgcagca
63901 agataaaggt gaacggcatg gtgaacatcg acatgtacgg gattataacc gacaagatca
63961 agctctcgag ctacaagctc aacgccgtgg ccgaagccgt cctgaaggac aagaagaagg
64021 acctgagcta tcgcgacatc cccgcctact acgccgccgg gcccgcgcaa cgcggggtga
64081 tcggcgagta ctgcatacag gattccctgc tggtgggcca gctgtttttt aagtttttgc
64141 cccatctgga gctctcggcc gtcgcgcgct tggcgggtat taacatcacc cgcaccatct
64201 acgacggcca gcagatccgc gtctttacgt gcctgctgcg cctggccgac cagaagggct
64261 ttattctgcc ggacacccag gggcgattta ggggcgccgg gggggaggcg cccaagcgtc
64321 cggccgcagc ccgggaggac gaggagcggc cagaggagga gggggaggac gagaacgaac
64381 gcgaggaggg cgggggcgag cgggagccgg agggcgcgcg ggagaccgcc ggccggcacg
64441 tggggtacca gggggccagg gtccttgacc ccacttccgg gtttcacgtg aaccccgtgg
64501 tggtgttcga ctttgccagc ctgtacccca gcatcatcca ggcccacaac ctgtgcttca
64561 gcacgctctc cctgagggcc gacgcagtgg cgcacctgga ggcgggcaag gactacctgg
64621 agatcgagat ggggggcga cggctgttct tcgtcaaggc tcacgtgcga gagagcctcc
64681 tcagcatcct cctgcgggac tggctcgcca tgcgaaagca gatccgctcg cggattcccc
64741 agagcagccc cgaggaggcc gtgctggtgg acaagcagca ggccgccatc aaggtcgtgt
64801 gtaactcggt gtacgggttc acgggagtgc agcacggact cctgccgtgc ctgcatgttg
64861 ccgcgacggt gacgaccatc ggccgcgaga tgctgctcgc gacccgcgag tacgtccacg
64921 cgcgctgggc ggccttcgaa cagctcctgg ccgatttccc ggaggcggcc gacatgcgcg
64981 ccccgggcc ctattccatg cgcatcatct acggggacac ggactccata tttgtgctgt
65041 gccgcggcct cacggccgcc gggctgacgg ccatgggcga caagatggcg agccacatct
65101 cgcgcgcgct gtttctgccc cccatcaaac tcgagtgcga aaagacgttc accaagctgc
65161 tgctgatcgc caagaaaaag tacatcggcg tcatctacgg gggtaagatg ctcatcaagg
65221 gcgtggatct ggtgcgcaaa aacaactgcg cgtttatcaa ccgcacctcc agggccctgg
65281 tcgacctgct gttttacgac gataccgtat ccggagcggc cgccgcgtta gccgagcgcc
65341 ccgcagagga gtggctggcg cgacccctgc ccgagggact gcaggcgttc ggggccgtcc

FIG. 8AM

65401 tcgtagacgc ccatcggcgc atcaccgacc cggagaggga catccaggac tttgtcctca
65461 ccgccgaact gagcagacac ccgcgcgcgt acaccaacaa gcgcctggcc cacctgacgg
65521 tgtattacaa gctcatggcc cgccgcgcgc aggtcccgtc catcaaggac cggatcccgt
65581 acgtgatcgt ggcccagacc cgcgaggtag aggagacggt cgcgcggctg gccgccctcc
65641 gcgagctaga cgccgccgcc ccaggggacg agcccgcccc ccccgcggcc ctgccctccc
65701 cggccaagcg cccccgggag acgccgtcgc atgccgaccc cccgggaggc gcgtccaagc
65761 cccgcaagct gctggtgtcc gagctggccg aggatcccgc atacgccatt gcccacggcg
65821 tcgccctgaa cacggactat tacttctccc acctgttggg ggcggcgtgc gtgacattca
65881 aggccctgtt tgggaataac gccaagatca ccgagagtct gttaaaaagg tttattcccg
65941 aagtgtggca ccccccggac gacgtggccg cgcggctccg ggccgcaggg ttcggggcgg
66001 tgggtgccgg cgctacggcg gaggaaactc gtcgaatgtt gcatagagcc tttgatactc
66061 tagcatgagc ccccgtcga agctgatgtc cctcatttta caataaatgt ctgcggccga
66121 cacggtcgga atctccgcgt ccgtgggttt ctctgcgttg cgccggacca cgagcacaaa
66181 cgtgctctgc cacacgtggg cgacgaacct gtaccccggg cacgcggtga gcatccggtc
66241 tatgagccgg tagtgcaggt gggcggacgt gccgggaaag atgacgtaca gcatgtggcc
66301 cccgtaagtg gggtccgggt aaaacaacag ccgcgggtcg cacgccccgc ctccgcgcag
66361 gatcgtgtgg acgaaaaaaa gctcgggttg gccaagaatc ccggccaaga ggtcctggag
66421 gggggcgttg tggcggtcgg ccaacacgac caaggaggcc aggaaggcgc gatgctcgaa
66481 tatcgtgttg atctgctgca cgaaggccag gattagggcc tcgcggctgg tggcggcgaa
66541 ccgcccgtct cccgcgttgc acgcgggaca gcaaccccgg atgcctaggt agtagcccat
66601 cccggagagg gtcaggcagt tgtcggccac ggtctggtcc agacagaagg gcagcgacac
66661 gggagtggtc ttcaccaggg gcaccgagag cgagcgcacg atggcgatct cctcggaggg
66721 cgtctgggcg agggcggcga aaaggccccg atagcgctgg cgctcgtgta aacacagctc
66781 ctgtttgcgg gcgtgaggcg gcaggctctt ccgggaggcc cgacgcacca cgcccagagt
66841 cccgccggcc gcagaggagc gcgaccgccg gcgctccttg ccgtgatagg gcccgggccg
66901 ggagccgcgg cgatgggggt cggtatcata cataggtaca cagggtgtgc tccagggaca
66961 ggagcgagat cgagtggcgt ctaagcagcg cgcccgcctc acggacaaat gtggcgagcg
67021 cggtgggctt tggtacaaat acctgatacg tcttgaaggt gtagatgagg gcacgcaacg

FIG. 8AN

67081 ctatgcagac acgcccctcg aactcgttcc cgcaggccag cttggccttg tggagcagca
67141 gctcgtcggg atgggtggcg gggggatggc cgaacagaac ccaggggtca acctccatct
67201 ccgtgatggc gcacatgggg tcacagaaca tgtgcttaaa gatggcctcg ggccccgcgg
67261 cccgcagcag gctcacaaac cggcccccgt ccccgggctg cgtctcgggg tccgcctcga
67321 gctggtcgac gacgggtacg atacagtcga agaggctcgt gttgttttcc gagtagcgga
67381 ccacggaggc ccggagtctg cgcagggcca gccagtaagc ccgcaccagt aacaggttac
67441 acagcaggca ttctccgccg gtgcgcccgc gcccccggcc gtgtttcagc acggtggcca
67501 tcagagggcc caggtcgagg tcgggctggg catcgggttc ggtaaactgc gcaaagcgcg
67561 gagccacgtc gcgcgtgcgt gccccgcgat gcgcttccca ggactggcgg accgtggcgc
67621 gacgggcctc cgcggcagcg cgcagctggg gccccgactc ccagacggcg ggggtgccgg
67681 cgaggagcag caggaccaga tccgcgtacg cccacgtatc cggcgactcc tccggctcgc
67741 ggtccccggc gaccgtctcg aattccccgt tgcgagcggc ggcgcgcgta cagcagctgt
67801 ccccgccccc gcgccgaccc tccgtgcagt ccaggagacg ggcgcaatcc ttccagttca
67861 tcagtgcggt ggtaagcgac ggctgcgtgc cggataccgc cgccgacccc gcccccctcct
67921 cgcccccgga ggccaaggtt ccgatgaggg cccgggtggc agactgcgcc aggaacgagt
67981 agttggagta ctgcaccttg gcggctcccg gggagggcga gggcttgggt tgcttctggg
68041 catgccgccc gggcaccccg ccgtcggtac ggaagcagca gtggagaaaa aagtgccggt
68101 ggatgtcgtt tatggtgagg gcaaagcgtg cgaaggagcc gaccagggtc gccttcttgg
68161 tgcgcagaaa gtggcggtcc atgacgtaca caaactcgaa cgcggccacg aagatgctag
68221 cggcgcagtg gggcgccccc aggcatttgg cacagagaaa cgcgtaatcg gccacccact
68281 gaggcgagag gcggtaggtt tgcttgtaca gctcgatggt gcggcagacc agacagggcc
68341 ggtccagcgc gaaggtgtcg atggccgccg cggaaaaggg cccggtgtcc aaaagccccct
68401 ccccacaggg atccgggggc gggttgcggg gtcctccgcg cccgcccgaa cccctccgt
68461 cgcccgcccc cccgcgggcc cttgaggggg cggtgaccac gtcggcggcg acgtcctcgt
68521 cgagcgtacc gacgggcggc acacctatca cgtgactggc cgtcaggagc tcggcgcaga
68581 gagcctcgtt aagagccagg aggctgggat cgaaggccac atacgcgcgc tcgaacgccc
68641 ccgccttcca gctgctgccg ggggactctt cgcacaccgc gacgctcgcc aggaccccgg
68701 ggggcgaagt tgccatggct gggcgggagg ggcgcacgcg ccagcgaact ttacgggaca

FIG. 8AO

68761 caatccccga ctgcgcgctg cggtcccaga ccctggagag tctagacgcg cgctacgtct
68821 cgcgagacgg cgcgcatgac gcggccgtct ggttcgagga tatgaccccc gccgagctgg
68881 aggttgtctt cccgactacg gacgccaagc taaactacct gtcgcggacg cagcggctgg
68941 cctccctcct gacgtacgcc gggcctataa aagcgcccga cgacgccgcc gccccgcaga
69001 ccccggacac cgcgtgtgtg cacggcgagc tgctcgcccg caagcgggaa agattcgcgg
69061 cggtcattaa ccggttcctg gacctgcacc agattctgcg gggctgacgc gcgcgctgtt
69121 gggcgggacg gttcgcgaac cctttggtgg gtttacgcgg gcacgcacgc tcccatcgcg
69181 ggcgccatgg cgggactggg caagccctac accggccacc caggtgacgc cttcgagggt
69241 ctcgttcagc gaattcggct tatcgtccca tctacgttgc ggggcgggga cggggaggcg
69301 ggcccctact ctcctccag cctcccctcc aggtgcgcct ttcagtttca tggccatgac
69361 gggtccgacg agtcgtttcc catcgagtat gtactgcggc ttatgaacga ctgggccgag
69421 gtcccgtgca acccttacct gcgcatacag aacaccggcg tgtcggtgct gtttcagggg
69481 ttttttcatc gcccacacaa cgccccgggg ggcgcgatta cgccagagcg gaccaatgtg
69541 atcctggggt ccaccgagac gacggggttg tccctcggcg acctggacac catcaagggg
69601 cggctcggcc tggatgcccg gccgatgatg gccagcatgt ggatcagctg ctttgtgcgc
69661 atgccccgcg tgcagctcgc gtttcggttc atgggccccg aagatgccgg acggacgaga
69721 cggatcctgt gccgcgccgc cgagcaggct attacccgtc gccgccgaac ccggcggtcc
69781 cgggaggcgt acggggccga ggccgggctg ggggtggccg gaacgggttt ccgggccagg
69841 ggggacggtt ttggcccgct ccccttgtta acccaagggc cctcccgccc gtggcaccag
69901 gccctgcggg gtcttaagca cctacggatt ggcccccccg cgctcgtttt ggcggcggga
69961 ctcgtcctgg gggccgctat ttggtgggtg gttggtgctg gcgcgcgcct ataaaaaagg
70021 acgcaccgcc gccctaatcg ccagtgcgtt ccggacgcct tcgccccaca cagccctccc
70081 gaccgacacc cccatatcgc tccccgacct ccggtcccga tggccgtccc gcaatttcac
70141 cgccccagca ccgttaccac cgatagcgtc cgggcgcttg gcatgcgcgg gctcgtcttg
70201 gccaccaata actctcagtt tatcatggat aacaaccacc cacacccccа gggcacccaa
70261 ggggccgtgc gggagtttct ccgcggtcag gcggcggcac tgacggacct tggtctggcc
70321 cacgcaaaca acacgtttac cccgcagcct atgttcgcgg gcgacgcacc ggccgcctgg
70381 ttgcggcccg cgtttggcct gcggcgcacc tattcacctt ttgtcgttcg agaaccttcg

FIG. 8AP

70441 acgcccggga ccccgtgagg cccagggagt tccttctggg gtgttttaat caataaaaga
70501 ccacaccaac gcacgagcct tgcgtttaat gtcgtgttta ttcaagggag tgggataggg
70561 ttcgacggtt cgaaacttaa cacaccaaat aatcgagcgc gtctagccca gtaacatgcg
70621 cacgtgatgt aggctggtca gcacggcgtc gctgtgatga agcagcgccc ggcgggtccg
70681 ctgtaactgc tgttgtaggc ggtaacaggc gcggatcagc accgccaggg cgctacgacc
70741 ggtgcgttgc acgtagcgtc gcgacagaac tgcgtttgcc gatacgggcg gggggccgaa
70801 ttgtaagcgc gtcacctctt gggagtcatc ggcgtataac gcactgaatg gttcgttggt
70861 tatgggggag tgtggttccc cagggagtgg gtcgagcgcc tcggcctcgg aatccgagag
70921 gaacaacgag gtggcgtcgg agtcttcgtc gtcagagaca tacagggtct gaagcagcga
70981 cacgggcgtg ggggtagcgt cgatgtgtag cgcgagggag gatgcccacg aagacacccc
71041 agacaaggag ctgcccgtgc gtggatttgt ggaagacgcg gaagccggga cggatgggcg
71101 gttttgcggt gcccggaacc gaaccgccgg atactccccg ggtgctacat gcccgttttg
71161 gggctggggt tggggctggg gcgcggacag gcggctgacg gtcaaatgcc cccgggggcg
71221 cgcagatgtg gcgggcgtgg ccaccggctg ccgtgtagtg gggcggcggg aaaccgggcc
71281 tccgggcgta acaccgccct ccagcgtcaa gtatgtgggg ggcgggcctg acgtcggggg
71341 cggggtgacg ggttggaccg cgggaggcgg gggagaggga cctgcgggag aggatgaggt
71401 cggctcggcc gggttgcggc ctaaaacagg ggccgtgggg tcggcggggt cccagggtga
71461 agggagggat tcccgcgatt cggacagcga cgcgacagcg gggcgcgtaa ggcgccgctg
71521 cggcccgcct acgggaaccc tggggggggt tggcgcggga cccgaggtta gcgggggcg
71581 gcggttttcg ccccgggca aaaccgtgcc ggttgcgacc gggggcggaa cgggatcgat
71641 agggagagcg ggagaagcct ggccggcgga ctggggaccg agcgggaggg gcacaccaga
71701 caccaaagcg tggggcgctg gctctggggg tttgggaggg gccgggggcc gcgcgaaatc
71761 ggtaaccggg gcgaccgtgt cggggagggc aggcggccgc caaccctggg tggtcgcgga
71821 agcctgggtg gcgcgcgcca gggagcgtgc ccggcggtgt cggcgcgcgc gcgacccgga
71881 cgaagaagcg gcagaagcgc gggaggaggc gggggggcgg ggggcggtgg catcgggggg
71941 cgccggggaa ctttgggggg acggcaagcg ccggacgtcg tcgcgggggc ccacgggcgc
72001 cggccgcgtg ctttcggccg ggacgcccgg tcgtgcttcg cgagccgtga ctgccggccc
72061 aggggccgc ggtgcacact gggatgtggg gacggactga tcggcggtgg gcgaaagggg

FIG. 8AQ

72121 gtccggggca aggaggggcg cggggccgcc ggagtcgtca gacgcgagct cctccaggcc
72181 gtgaatccat gcccacatgc gaggggggac gggctcgccg ggggtggcgt cggtgaatag
72241 cgtgggggcc aggcttccgg gccccaacga gccctccgtc ccaacaaggt ccgccgggcc
72301 gggggtcggg ttcgggaccg aggggctctg gtcgtcgggg gcgcgctggt acaccggatg
72361 ccccgggaat agctcccccg acaggaggga ggcgtcgaac ggccgcccga ggatagctcg
72421 tgcgaggaag gggtcctcgt cggtggcgct ggcggcgagg acgtcctcgc cgcccgccac
72481 aaacgggagc tcctcggtgg cctcgctgcc aacaaaccgc acgtcggggg ggccgggggg
72541 tccgggtttt cccacaacac cgcgaccggg gtcatggaga tgtccacgag caccagacac
72601 ggcgggcccc gggcgagggg ccgctcggcg atgagcgcgg acaggcgcgg gagctgtgcc
72661 gccagacacg cgttttcgat cgggttcagg tcggcgtgca ggaggcggac ggcccacgtc
72721 tcgatgtcgg acgacacggc atcgcgcaag gcggcgtccg gcccgcgagc gcgtgagtca
72781 aacagcgtga ggcacagctc cagctccgac tcgcgggaaa aggccgtggt gttgcggagc
72841 gccacgacga cgggcgcgcc caggagcact gccgccagca ccaggtccat ggccgtaacg
72901 cgcgccgcgg gggtgcggtg ggtggcggcg gccggcacgg cgacgtgctg gcccgtgggc
72961 cggtagaggg cgttgggggg agcgggggct gacgcctcgc gccccccga ggggctcagc
73021 gtctgcccag attccagacg cgcggtcaga agggcgtcga aactgtcata ctctgtgtag
73081 tcgtccggaa acatgcaggt ccaaagagcg gccagcgcgg tgcttgggag acacatgcgc
73141 ccgaggacgc tcaccgccgc cagcgcctgg gcgggactca gctttcccag cgcggcgccg
73201 cgctcggttc ccagctcggg gaccgagcgc cagggcgcca gggggtcggt ttcggacaac
73261 ttgccgcggc gccagtctgc cagccgcgtg ccgaacatga ggccccgggt cggagggcct
73321 ccggccgaaa acgctggcag cacgcggatg cgggcgtctg gatgcggggt caggcgctgc
73381 acgaatagca tggaatctgc tgcgttctga aacgcacggg ggagggtgag atgcatgtac
73441 tcgtgttggc ggaccagatc caggcgccaa aaggtgtaaa tgtgttccgg ggagctggcc
73501 accagcgcca ccagcacgtc gttctcgtta aaggaaacgc ggtgcctagt ggagctctgg
73561 ggtccgagcg gcggccccgg ggccgccgcg tcaccccccc attccagctg ggcccagcga
73621 cacccaaact cgcgcgtgag agtggtcgcg acgagggcga cgtagagctc ggccgccgca
73681 tccatcgagg cccccccatct cgcctggcgg tggcgcacaa agcgtccgaa gagctgaaag
73741 ttggcggcct gggcgtcgct gagggccagc tgaagccggt tgatgacggt gaggacgtac

FIG. 8AR

73801 atggccgtga cggtcgaggc cgactccagg gtgtccgtcg gaagcggggg gcgaatgcat
73861 gccgcctcgg gacacatcag cagcgcgccg agcttgtcgg tcacggccgg gaagcagagc
73921 gcgtactgca gtggcgttcc atccgggacc aaaaagctgg gggcgaacgg ccgatccagc
73981 gtactggtgg cctcgcgcag caccaggggc cccgggcctc cgctcactcg caggtacgcc
74041 tcgccccggc ggcgcagcat ctgcgggtcg gcctcttggc cgggtggggc ggacgcccgg
74101 gcgcgggcgt ctagggcgcg aagatccacg agcaggggcg cgggcgcggc ggccgcgccc
74161 gcgcccgtct ggcctgtggc cttggcgtac gcgctatata agcccatgcg gcgttggatg
74221 agctcccgcg cgccccggaa ctcctccatc gcccatgggg ccaggtcccc ggccaccgcg
74281 tcgaattccg ccaacaggcc ccccagggta tcaaagttca tctcccaggc caccttggc
74341 accacctcgt cccgcagccg ggcgctcagg tcggcgtgtt gggccacgcg cccccgagc
74401 tcctccacgg ccccggcccg ctcggcgctc ttggcgccca ggacgccctg gtacttggcg
74461 ggaaggcgct cgtagtcccg ctgggctcgc agccccgaca cagtgttggt ggtgtcctgc
74521 agggcgcgaa gctgctcgca tgccgcgcga aatccctcgg gcgatttcca ggccccccg
74581 cgaacgcggc cgaagcgacc ccatacctcg tcccactccg cctcggcctc ctcgagagac
74641 ctccgcaggg cctcgacgcg gcgacgggtg tcgaagagcg cctgcaggcg cgcgccctgt
74701 cgcgtcagga ggcccgggcc gtcgctgctg gccgcgctta gcgggtgcgt ctcaaaggta
74761 cgctgggcat gttccaacca ggcgaccgcc tgcacgtcga gctcgcgcgc cttctccgtc
74821 tggtccaaca gaatttcgac ctgatccgcg atctcctccg ccgagcgcgc ctggtccagc
74881 gtcttggcca cggtcgccgg gacggcgacc accttcagca gggtcttcag attggccaga
74941 ccctcggcct cgagctgggc ccggcgctcg cgcgcggcca gcacatcccg cagccccgcc
75001 gtgacccgct cggtggcttc ggcgcgctgc tgtttggcgc gcaccacggc gtccttggta
75061 tcggccaggt cctgtcgggt cacgaatgcg acgtagtcgg agtacgccgt gtccttcacg
75121 gggctctggt ccacgcgctc cagcgccgcc acgcacgcca ccagcgcgtc ctcgctcggg
75181 cagggcaggg tgaccctgc ccggacaagc tcggcggccg ccgccgggtc gttgcgcacc
75241 gcggatatct cctccgcggc ggcggccagg tccagcgcca cgcttccgat cgcgcgccgc
75301 gcgtcggccc ggagggcgtc caggcgatcg cggatatcca cgtactcggc gtagcccttt
75361 tgaaaaaacg gcacgtactg gcgcagggcc ggcacgcccc ccaagtcttc cgacaggtgt
75421 agtacggcct cgtggtagtc gataaacccg tcgttcgcct gggcccgctc cagcagcccc

FIG. 8AS

75481 cccgcgagcc gcagaagccg cgccaggggc tcggtgtcca cccgaaacat gtcggcgtac
75541 gtgtcggccg cggccccaaa ggccgcgctc cagtcgatgc ggtgaatggc tgcgagcggg
75601 gggagcatgg ggtggcgctg gttctcgggg gtgtatgggt taaacgcaag ggccgtctcc
75661 agggcaaggg tcaccgcctt ggcgttggtt cccagcgcct gctcggcccg ctttcggaag
75721 tcccgggggt tgtagccgtg cgtgcccgcc agcgcctgca ggcgacggag ctcgaccacg
75781 tcaaactcgg cactgctttc cacgcggtcc agcacggcct ccacgtcggc ggcccagcgc
75841 tcgtggctac tgcgggcgcg ctgggccgcc atcttctctc tgaggtcggc ggtggcggcc
75901 tcaagttcgt cggcgcggcg tcgcgtggcg ccgatgacct ttcccagctc ctgcagggcg
75961 cgcccgctgg gggagtggtc cccggccgtc ccttcggcgt gcaacaggcc cccgaacctg
76021 ccctcgtggc ccgcgaggct ttcccgcgcg ccggtggtcg cgcgcgtcgc ggcttggatc
76081 agggaggcat gctctccctt cggttggttg gcggcccggc gcacctggac gacaaggtcg
76141 gcggcagccg accctaaggt cgtgagctgg gcgatggccc cccgcgcgtc cagggccaac
76201 cgagtcgcct tgacgtatcc cgcggcgctg tcggccatgg ccgctaggaa ggccaggggg
76261 gaggccgggt cgctggcggc cgcgcccagg gccgtcaccg cgtcgaccag gacgcggtgc
76321 gcccgcacgg ccgcatccac cgtcgacgcg gggtctgccg tcgcgacggc ggcgctgccg
76381 gcgttgatgg cgttcgagac ggcgtgggct atgatcgggg cgtgatcggc gaagaactgc
76441 aagagaaacg gagtctcggg ggcgtcggcg aacaggttct tcagcaccac cacgaagctg
76501 ggatgcaagc cggacagagc cgtcgccgtg tccggagtcg ggtgctccag ggcatctcgg
76561 tactgcccca gcagccccca catgtccgcc cgcagcgccg ccgtaacctc aggggcgcc
76621 ccccgaacgg cctcggggag gtccgaccag cccgccggca gggaggcccg cagggtcgcc
76681 aggacggccg gacaggcctt tagccccaca aagtcaggga gggggcgcag gacccccctgg
76741 agtttgtgca agaacttctc ccgggcgtcg cgggccacct tcgcccgctc ccgcgctccc
76801 tcgagcattg cctccaggga gcgcgcgcgc tcccgcaaac ggacacgcgc atcgggggcg
76861 agctctgccg tcagcttggc ggcatccatg gcccgcgcct gccgcagcgc ttcctcggcc
76921 atgcgcgtgg cctctggcga cagcccgccg tcgtcggggt agggcgacgc gccgggcgca
76981 ggaacaaagg ccgcgtcgct gtccagctgc tggcccaggg ccgcatctag ggcgtcgaag
77041 cgccgcagct cggccagacc cgagctgcgg cgcgcctgct ggtcgttaat gtcgcggatg
77101 ctgcgcgcca gctcgtccag cggcttgcgt tctatcagcc cttggttggc ggcgtccgtc

FIG. 8AT

77161 aggacggaga gccaggccgc caggtcctcg ggggcgtcca gcgtctggcc ccgctgtatc
77221 agatcccgca acaggatggc cgtggggctg gtcgcgatcg ggggcggggc gggaatggcg
77281 gcgcgctgcg cgatgtcccg cgtgtgctgg tcgaagacag gcagggactc gagcagctgg
77341 accacgggca cgacggcggc cgaagccacg tgaaaccggc ggtcgttgtt gtcgctggcc
77401 tgtagagcct tggcgctgta tacggccccc cggtaaaagt actccttaac cgccccctcg
77461 atcgcccgac gggcctgggt ccgcacctcc tccagccgaa cctgaacggc ctcgggggccc
77521 agggggggtg ggcgcggagc cccctgcggg gccgccccgg ccggggcggg cattacgccg
77581 aggggcccgg cgtgctgtga gaccgcgtcg accccgcgag cgagggcgtc gagggcctcg
77641 cgcatctggc gatcctccgc ctccacccta atctcttcgc cacgggcaaa tttggccaga
77701 gcctggactc tatacagaag cggttctggg tgcgtcgggg tggcgggggc aaaaagggtg
77761 tccgggtggg cctgcgagcg ctccagaagc cactcgccga ggcgtgtata cagattggcc
77821 ggcggggccg cgcgaagctg cagctccagg tccgcgagtt ccccgtaaaa ggcgtccgtc
77881 tcccgaatga catccctagc cacaaggatc agcttcgcca gcgccaggcg accgatcaga
77941 gagttttcgt ccagcacgtg ctggacgagg ggcagatggg cggccacgtc ggccaggctc
78001 aggcgcgtgg aggccagaaa gtccccccacg gccgttttcc ggggcagcat gctcagggta
78061 aactccaaca gggcggcggc cgggccggcc accccggcct gggtgtgcgt ccgggccccg
78121 ttctcgatga gaaaggcgag gacgcgttca aagaaaaaaa taacacagag ctccagcagc
78181 cccggagaag ccggatacgg cgaccgtaag gcgctgatgg tgagccgcga acacgcggcg
78241 acctcgcggg ccagggcggc ggagcacgcg gtgaacttaa ccgccgtggc ggccacgttt
78301 gggtgggcct cgaacagctg ggcaaggtct gcgcccgggg gctcgggtga gcggcgagtc
78361 ttcagcgcct cgagggcctg cgaggacgcc ggaaccgtgg gcccgtcgtc ctcgcccgcc
78421 tcggcgaccg gcggcccggc cgggtcgggg ggtgccgagg cgaggacagg ctccggaacg
78481 gaggcgggga ccgcggcccc gacgggggtt ttgcctttgg gggtggattt cttcttggtt
78541 ttggcagggg gggccgagcg tttcgttttc tccccccgaag tcaggtcttc gacgctggaa
78601 ggcggagtcc aggtgggtcg gcggcgcttg ggatggccgg ccgagtagcg tgcccggtgc
78661 cgaccaaccg ggacgacgcc catctccagg acccgcatgt cgtcgtcatc ttcttcggcc
78721 gcctctgcgg cgggggggctt gggggcggag ggaggcggtg gtgggatcgc ggagggtggg
78781 tcggcggagg ggggatccgt gggtggggta cccttcaggg ccaccgccca tacatcgtcg

FIG. 8AU

78841 ggcgcccgat tcgggcgctt ggcctctggt tttgccgacg gaccggccgt cccccgggat
78901 gtctcggagg ccctgtcgtc gcgacgggcc cgggtcggtg gcggcgactg ggcggctgtg
78961 ggcgggtgtg gccccgtgcc ccctaccccc tcccgggggc ccacgccgac gcagggctcc
79021 cccaggcccg cgatctcgcc ccgcaggggg tgcgtgatgg ccacgcgccg ttcgctgaac
79081 gcttcgtcct gcaggtaagt ctcgctggcc ccgtaaagat gcagagccgc ggccgtcaag
79141 tccgcaggag ccgcgggttc cgggcccgac ggcacgaaaa acaccatggc tcccgcccac
79201 cgtacgtccg ggcgatcgcg ggtgtaatac gtcaggtatg gatacatgtc ccccgcccgc
79261 actttggcga tgaacgcggg ggtgccctcc ggaaggccgt gcgggtcaaa aaggtatgcg
79321 gtgtcgccgt ccctgaacag ccccatcccc aggggccaa tggttaggag cgtgtacgac
79381 aggggcgca gggcccacgg gccggcgaag aacgtgtgtg cggggcattg tgtctccagc
79441 aggcccgccg cgggctcccc gaagaagccc acctcgccgt atacgcgcga gaagacacag
79501 cgcagtccgc cgcgcgcccc tgggtactcg aggaagttgg ggagctcgac gatcgaacac
79561 atgcgcggcg gcccagggct cgcggtcgcg cgcgtccact cgccccctc gaccaaacaa
79621 ccctcgatgg cctccgcgga cagaacgtcg cgagggccca catcaaatat gaggctgaga
79681 aaggacagcg acgagcgcat gcacgatacc gacccccccg gctccaggtc gggcgcgaac
79741 tggttccgag caccggtgac cacgatgtcg cgatcccccc cgcgttccat cgtggagtgc
79801 ggtggggtgc ccgcgatcat atgtgcccta ctggccagag acccggcctg tttatggacc
79861 ggaccccggg ggttagtgtt gtttccgcca cccatgcccc cgtaccatgg ccccggttcc
79921 cctgattagg ctacgagtcg cggtgatcgc ttcccaaaaa ccgagctgcg tttgtctgtc
79981 ttgatcttcc ccccccccc cgcccgcccg cccgcccgcc cgcccgcaca ccataacacc
80041 gagaacaaca cacgggggtg ggcgtaacat aataaagctt tattggtaac tagttaacgg
80101 caagtccgtg ggtggcgcga cggtgtcctc cgggctcatc tcgtcgtcct cgacgggggt
80161 gttggaatga ggcgcccccc cgcggtccgc ctggcgtggg ccgtgcccat cggcctccgg
80221 cttctgtgcg tccatgggca taggcgcggg gagactgttt ccggcgtcgc ggacctccag
80281 gtccctggga gactccggtc cggctaacgg acgaaacgcg gaagcgcgaa acacgccgtc
80341 ggtgacccgc aggagctcgt tcatcagtaa ccaatccata ctcagcgtaa cggccagccc
80401 ctggcgagac agatccacgg agtccggaac cgcggtcgtc tggcccaggg ggccgaggct
80461 gtagtccccc caggccccta ggtcgcgacg gctcgtaagc acgacgcggt cggccgcggg

FIG. 8AV

80521 gctttgcggg ggggcgtcct cgggcgcatg cgccattacc tctcggatgg ccgcggcgcg
80581 ctggtcggcc gagctgacca agggcgccac gaccacggcg cgctccgtct gcaggccctt
80641 ccacgtgtcg tggagttcct ggacaaactc ggccacgggc tcgggtcccg cggccgcgcg
80701 cgcggcttga tagcaggccg agagacgccg ccagcgcgct agaaactgac ccatgaagca
80761 aaacccgggg acctggtctc ccgacagcag cttcgacgcc cgggtgtgaa tgccggacac
80821 aacggacaga aacccgtgaa tttcgcgccg gaccacggcc agcacgttgt cctcgtgcga
80881 cacctgggcc gccagctcgt cgcacacctc caggtgcgcc gtggtttcgg tgatgacgga
80941 acgcaggctc gcgagggacg cgaccagcgc gcgcttggcg tcgtgataca tgctgcagta
81001 ctgactcacc gcgtccccca tggcctcggg gggccagggc cccaggcggt cgggagtgtc
81061 cccgaccacc gcatacaggc ggcgcccgtc gctctcgaac cgacactcga aaaaggcgga
81121 gagcgtgcgc atgtgcagcc gcagcagcac gatggcgtcc tccagttggc gaatcagggg
81181 gtctgcgcgc tcggcgaggt cctgcagcac cccccgggcg gccagggcgt acatgctaat
81241 caacaggagg ctggtgccca cctcgggggg cgggggggcc tgcagctgga ccaggggccg
81301 cagctgctcg acggcacccc tggagatcac gtacagctcc cggagcagct gctctatgtt
81361 gtcggccatc tgcatagtgg ggccgaggcc gccccgggcg gccggttcga ggagggtaat
81421 cagcgcgccc agtttggtgc gatggccctc gaccgtgggg agatagccca gcccaaagtc
81481 ccgggcccag gccaacacac gcagggcgaa ctcgaccggg cgtggaaggt aggccgcgct
81541 acacgtggcc cccaacgcgt ccccgaccac cagggccaga acgtagggga cgaagcccgg
81601 gtcggcgagg acgttggggt gaatgccctc gagggcgggg aagcggatct gggtcgccgc
81661 ggccaggtgg acagaggggg cgtggctggg ctgcccgacg gggagaagcg cggacagcgg
81721 cgtggccggg gtggtggggg tgatgtccca gtgggtctga ccatacacgt cgatccagat
81781 gagcgccgtc tcgcggagaa ggctgggttg accggaacta aagcggcgct cggccgtctc
81841 aaactccccc acgagcgccc gccgcaggct cgccagatgt tccgtcggca cggccggacc
81901 catgatacgc gccagtgtct ggctcagaac gccccccgac aggccgaccg cctcgcagag
81961 ccgcccgtgc gtgtgctcgc tggcgccctg gacccgcctg aaagttttta cgtagttggc
82021 atagtacccg tattcccgcg ccagaccaaa cacgttcgac cccgcgaggg caatgcaccc
82081 aaagagctgc tggacttcgc cgagtccgtg gccggcgggc gtccgcgcgg ggacgcccgc
82141 cgccagaaac ccctccaggg ccgaaaggta gtgcgtgcag tgcgagggcg tgaacccagc

FIG. 8AW

82201 gtcgatcagg gtgttgatca ccacggaggg cgaattggta ttctggatca acgtccacgt
82261 ctgctgcagc agagccagca gccgctgctg ggcgccggcg gagggctgct ccccgagctg
82321 cagcaggctg gagacggcag gctggaagac tgccagtgcc gacgaactca ggaacggcac
82381 gtcgggatca aacacggcca cgtccgtccg cacgcgcgcc attagcgtcc ccgggggcgc
82441 acaggccgag cgcgggctga cgcggctgag ggccgtcgac acgcgcacct cctcgcggct
82501 gcgaaccatc ttgttggcct ccagtggcgg aatcattatg gccgggtcga tctcccgcac
82561 ggtgtgctga aactgcgcca acaggggcgg cgggaccaca gccccccgct cgggggtcgt
82621 caggtactcg tccaccaggg ccaacgtaaa gagggcccgt gtgaggggag tgagggtcgc
82681 gtcgtctatg cgctggaggt gcgccgagaa cagcgtcacc cgattactca ccagggccaa
82741 gaaccggagg ccctcttgca cgaacggggc ggggaagagc aggctgtacg ccggggtggt
82801 aaggttcgcg ctgggctgcc ccaacgggac cggcgccatc ttgagtgacg tctccccaag
82861 ggcctcgatg gaggtccgcg ggctcatggc caagcagctc ttggtgacgg tttgccagcg
82921 gtctatccac tccacggcgc actggcggac gcggaccggc cccagggccg ccgcggtgcg
82981 caggccggcg gaatccagcg catgggacgt gtcggagccg gtgaccgcga ggatggtgtc
83041 cttgatgacc tccatctccc ggaaggcctg gtcgggggcc tcggggagag ccaccaccaa
83101 gcggtgtacg agcaacccgg ggaggttctc ggccaagagc gccgtctccg gaagcccgtg
83161 ggcccggtgg agcgcgcaca ggtgttccag cagcggccgc cagcatgccc gcgcgtctac
83221 cggggcaatg gccgttcccg acaacagaaa cgccgccatg gcggcgcgca gcttggccgt
83281 ggccagaaac gccgggtcgt ccgccccgtt tgccgtctcg gccgtggggg ttggcggttg
83341 gcgaaggccg gctaggctcg ccaataggcg ctgcataggt ccgtccgagg gcggaccggc
83401 gggtgaggtc gtgacgacgg gggcctcgga cgggagaccg cggtctgcca tgacgcccgg
83461 ctcgcgtggg tgggggacag cgtagaccaa cgacgagatc gggcgggaat gactgtcgtg
83521 cgctgtaggg agcggcgaat tatcgatccc ccgcggccct ccaggaaccc cgcaggcgtt
83581 gcgagtaccc cgcgtcttcg cggggtgtta tacggccact taagtccccgg catcccgttc
83641 gcggacccag gcccgggggа ttgtccggat gtgcgggcag cccggacggc gtgggttgcg
83701 gactttctgc ggggcggccc aaatggccct ttaaacgtgt gtatacggac gcgccgggcc
83761 agtcggccaa cacaacccac cggaggcggt agccgcgttt ggctgtgggg tgggtggttc
83821 cgccttgcgt gagtgtcctt tcgacccccc tccccgggt tttgttaggt cgcgatctgc

FIG. 8AX

```
83881 agtcgcaatg aagaccaatc cgctacccgc aacccccttcc gtgtggggcg ggagtaccgt
83941 ggaactcccc cccaccacac gcgataccgc gggacagggc ctgcttcggc gcgtcctgcg
84001 ccccccgatc tctcgccgcg acggcccagt gctccccagg gggtcgggac cccggagggc
84061 ggccagcacg ctgtggttgc ttggcctgga cggcacagac gcgccccctg gggcgctgac
84121 ccccaacgac gataccgaac aggccctgga caagatcctg cggggcacca tgcgcggggg
84181 ggcggccctg atctgctccc cgcgccatca tctaacccgc caagtgatcc tgacggatct
84241 gtgccaaccc aacgcggatc gtgccgggac gctgcttctg gcgctgcggc accccgccga
84301 cctgcctcac ctggcccacc agcgcgcccc gccaggccgg cagaccgagc ggctgggcga
84361 ggcctggggc cagctgatgg aggcgaccgc cctggggtcg gggcgagccg agagcgggtg
84421 cacgcgcgcg ggcctcgtgt cgtttaactt cctggtggcg gcgtgtgccg cctcgtacga
84481 cgcgcgcgac gccgccgatg cggtacgggc ccacgtcacg gccaactacc gcgggacgcg
84541 ggtgggggcg cgcctggatc gtttttccga gtgtctgcgc gccatggttc acacgcacgt
84601 cttccccccac gaggtcatgc ggtttttcgg ggggctggtg tcgtgggtca cccaggacga
84661 gctagcgagc gtcaccgccg tgtgcgccgg gccacaggag gcggcgcaca ccggccaccc
84721 gggccggccc cgctcggccg tgatcctccc ggcgtgtgcg ttcgtggacc tggacgccga
84781 gctggggctg gggggcccgg gtgcggcgtt tctgtacctg gtattcactt accgccagcg
84841 ccgggaccag gagctgtgtt gtgtgtacgt gatcaagagc cagctccccc cgcgcgggtt
84901 ggagccggcc ctggagcggc tgtttgggcg cctccggatc accaacacga ttcacggcac
84961 cgaggacatg acgccccccgg ccccaaaccg aaaccccgac ttcccccctcg cgggcctggc
85021 cgccaatccc caaaccccgc gttgctcggc tggccaggtc acgaaccccc agttcgccga
85081 caggctgtac cgctggcagc cggacctgcg ggggcgcccc accgcacgca cctgtacgta
85141 cgccgccttt gcagagctcg gcatgatgcc cgaggatagt ccccgctgcc tgcaccgcac
85201 cgagcgcttt ggggcggtca gcgtccccgt tgtcatcctg gaaggcgtgg tgtggcgccc
85261 cggcgagtgg cgggcatgcg cgtgagcgta gcaaacgccc cgcccacaca acgctccgcc
85321 cccaaccccct tccccgctgt cactcgtggt tcgttgaccc ggacgtccgc caaataaagc
85381 cactgaaacc cgaaacgcga gtgttgtaac gtcctttggg cgggaggaag ccacaaaatg
85441 caaatgggat acatggaagg aacacacccc cgtgactcag gacatcggcg tgtccttttg
85501 ggtttcactg aaactggccc gcgccccacc cctgcgcgat gtggataaaa agccagcgcg
```

FIG. 8AY

85561 ggtggtttag ggtaccacag gtgggtgctt tggaaacttg tcggtcgccg tgctcctgtg
85621 agcttgcgtc cctccccggt ttcctttgcg ctcccgcctt ccggacctgc tcttgcctat
85681 cttctttggc tctcggtgcg attcgtcagg cagcggcctt gtcgaatctc gaccccacca
85741 ctcgccggac ccgccgacgt cccctctcga gcccaccgaa acccgccgcg tctgttgaaa
85801 tggccagccg cccagccgca tcctctcccg tcgaagcgcg ggccccggtt gggggacagg
85861 aggccggcgg ccccagcgca gccacccagg gggaggccgc cggggccccc ctcgcccacg
85921 gccaccacgt gtactgccag cgagtcaatg gcgtgatggt gctttccgac aagacgcccg
85981 ggtccgcgtc ctaccgcatc agcgatagca actttgtcca atgtggttcc aactgcacca
86041 tgatcatcga cggagacgtg gtgcgcgggc gcccccagga cccgggggcc gcggcatccc
86101 ccgctccctt cgttgcggtg acaaacatcg gagccggcag cgacggcggg accgccgtcg
86161 tggcattcgg gggaacccca cgtcgctcgg cggggacgtc taccggtacc cagacggccg
86221 acgtccccac cgaggccctt gggggccccc ctcctcctcc ccgcttcacc ctgggtggcg
86281 gctgttgttc ctgtcgcgac acacggcgcc gctctgcggt attcgggggg gagggggatc
86341 cagtcggccc cgcggagttc gtctcggacg accggtcgtc cgattccgac tcggatgact
86401 cggaggacac ggactcggag acgctgtcac acgcctcctc ggacgtgtcc ggcggggcca
86461 cgtacgacga cgcccttgac tccgattcgt catcggatga ctccctgcag atagatggcc
86521 ccgtgtgtcg cccgtggagc aatgacaccg cgcccctgga tgtttgcccc gggacccccg
86581 gcccgggcgc cgacgccggt ggtccctcag cggtagaccc acacgcgccg acgccagagg
86641 ccggcgctgg tcttgcggcc gatcccgccg tggcccggga cgacgcggag gggctttcgg
86701 acccccggcc acgtctggga acgggcacgg cctaccccgt cccctggaa ctcacgcccg
86761 agaacgcgga ggccgtggcg cgctttctgg gagatgccgt gaaccgcgaa cccgcgctca
86821 tgctggagta cttttgccgg tgcgcccgcg aggaaaccaa gcgtgtcccc cccaggacat
86881 tcggcagccc ccctcgcctc acggaggacg actttgggct tctcaactac gcgctcgtgg
86941 agatgcagcg cctgtgtctg gacgttcctc cggtcccgcc gaacgcatac atgccctatt
87001 atctcaggga gtatgtgacg cggctggtca acgggttcaa gccgctggtg agccggtccg
87061 ctcgccttta ccgcatcctg ggggttctgg tgcacctgcg gatccggacc cgggaggcct
87121 cctttgagga gtggctgcga tccaaggaag tggccctgga ttttggcctg acggaaaggc
87181 ttcgcgagca cgaagcccag ctggtgatcc tggcccaggc tctggaccat tacgactgtc

FIG. 8AZ

87241 tgatccacag cacaccacac acgctggtcg agcgggggct gcaatcggcc ctgaagtatg
87301 aggagtttta cctaaagcgt tttggcgggc actacatgga gtccgtcttc cagatgtaca
87361 cccgcatcgc cggcttttg gcctgccggg ccacgcgcgg catgcgccac atcgccctgg
87421 ggcgagaggg gtcgtggtgg gaaatgttca agttcttttt ccaccgcctc tacgaccacc
87481 agatcgtacc gtcgaccccc gccatgctga acctggggac ccgcaactac tacacctcca
87541 gctgctacct ggtaaacccc caggccacca caaacaaggc gaccctgcgg gccatcacca
87601 gcaacgtcag tgccatcctc gcccgcaacg ggggcatcgg gctatgcgtg caggcgttta
87661 acgactccgg ccccgggacc gccagcgtca tgcccgccct caaggtcctt gactcgctgg
87721 tggcggcgca caacaaagag agcgcgcgtc cgaccggcgc gtgcgtgtac ctggagccgt
87781 ggcacaccga cgtgcgggcc gtgctccgga tgaaggggt cctcgccggc gaagaggccc
87841 agcgctgcga caatatcttc agcgccctct ggatgccaga cctgtttttc aagcgcctga
87901 ttcgccacct ggacggcgag aagaacgtca catggaccct gttcgaccgg gacaccagca
87961 tgtcgctcgc cgactttcac ggggaggagt tcgagaagct ctaccagcac ctcgaggtca
88021 tggggttcgg cgagcagata cccatccagg agctggccta tggcattgtg cgcagcgcgg
88081 ccacgaccgg gagccccttc gtcatgttca aagacgcggt gaaccgccac tacatctacg
88141 acacccaggg ggcggccatc gccggctcca acctctgcac cgagatcgtc catccggcct
88201 ccaagcgatc cagtggggtc tgcaacctgg gaagcgtgaa tctggcccga tgcgtctcca
88261 ggcagacgtt tgactttggg cggctccgcg acgccgtgca ggcgtgcgtg ctgatggtga
88321 acatcatgat cgacagcacg ctacaaccca cgccccagtg cacccgcggc aacgacaacc
88381 tgcggtccat gggaatcggc atgcagggcc tgcacacggc ctgcctgaag ctggggctgg
88441 atctggagtc tgccgaattt caggacctga acaaacacat cgccgaggtg atgctgctgt
88501 cggcgatgaa gaccagcaac gcgctgtgcg ttcgcggggc ccgtcccttc aaccacttta
88561 agcgcagcat gtatcgcgcc ggccgctttc actgggagcg ctttccggac gcccggccgc
88621 ggtacgaggg cgagtgggag atgctacgcc agagcatgat gaaacacggc ctgcgcaaca
88681 gccagtttgt cgcgctgatg cccaccgccg cctcggcgca gatctcggac gtcagcgagg
88741 gctttgcccc cctgttcacc aacctgttca gcaaggtgac ccgggacggc gagacgctgc
88801 gccccaacac gctcctgcta aaggaactgg aacgcacgtt tagcgggaag cgcctcctgg
88861 aggtgatgga cagtctcgac gccaagcagt ggtccgtggc gcaggcgctc ccgtgcctgg

FIG. 8BA

88921 agcccaccca ccccctccgg cgattcaaga ccgcgtttga ctacgaccag aagttgttga
88981 tcgacctgtg tgcggaccgc gccccctacg tcgaccatag ccaatccatg accctgtatg
89041 tcacggagaa ggcggacggg accctcccag cctccaccct ggtccgcctt ctggtccacg
89101 catataagcg cggactaaaa acagggatgt actactgcaa ggttcgcaag gcgaccaaca
89161 gcggggtctt tggcggcgac gacaacattg tctgcacgag ctgcgcgctg tgaccgacaa
89221 acccccctcg cgccaggccc gccgccactg tcgtcgccgt cccacgctct cccctgctgc
89281 catggattcc gcggccccag ccctctcccc cgctctgacg gcccttacgg gccagagcgc
89341 gacggcggac ctggcgatcc agattccaaa gtgccccgac cccgagaggt acttctacac
89401 ctcccagtgt cccgacatta accacctgcg ctccctcagc atccttaacc gctggctgga
89461 aaccgagctt gttttcgtgg gggacgagga ggacgtctcc aagctttccg agggcgagct
89521 cagcttttac cgcttcctct tcgctttcct gtcggccgcc gacgacctgg ttacggaaaa
89581 cctgggcggc ctctccggcc tgtttgagca gaaggacatt ctccactact acgtggagca
89641 ggaatgcatc gaagtcgtac actcgcgcgt gtacaacatc atccagctgg tgcttttcca
89701 caacaacgac caggcgcgcc gcgagtacgt ggccggcacc atcaaccacc cggccatccg
89761 cgccaaggtg gactggttgg aagcgcgggt gcgggaatgc gcctccgttc cggaaaagtt
89821 cattctcatg atcctcatcg agggcatctt ttttgccgcc tcgtttgccg ccatcgccta
89881 ccttcgcacc aacaaccttc tgcgggtcac ctgccagtca aacgacctca tcagccggga
89941 cgaggccgtg cacacgacgg cctcgtgtta catctacaac aactacctcg gcgggcacgc
90001 caagcccccg cccgaccgcg tgtacgggct gttccgccag gcggtcgaga tcgagatcgg
90061 atttatccga tcccaggcgc cgacggacag ccatatcctg agcccggcgg cgctggcggc
90121 catcgaaaac tacgtgcgat tcagcgcgga tcgcctgttg ggccttatcc acatgaagcc
90181 actgttttcc gccccacccc ccgacgccag ctttccgctg agcctcatgt ccaccgacaa
90241 acacaccaat tttttcgagt gtcgcagcac ctcctacgcc ggggcggtcg tcaacgatct
90301 gtgagtgtcg cggcgcgctt ctacccgtgt ttgcccataa taaacctctg aaccaaactt
90361 tgggtctcat tgtgattctt gtcagggacg cgggggtggg agaggataaa aggcggcgca
90421 aaaagcagta accaggtccg tccagattct gcgggcatag aataccataa ttttattggt
90481 gggtcgtttg ttcgggggaca agcgcgctcg tctgacgttt gggctactcg tcccagaatt
90541 tggccaggac gtccttgtag aacgcgggtg gggggcctg ggtccgcaac tgctccagaa

FIG. 8BB

90601 acctgtcggc gatatcaggg gccgtgatat gccgggtcac gatagatcgc gccaggtttt
90661 cgtcgcggat gtcctggtag ataggcaggc gtttcagaag agtccacggc ccccgctcct
90721 tggggccgat aagcgatatg acgtacttaa tgtagcggtg ttccaccagc tcggtgatgg
90781 tcatgggatc ggggagccag tccagggact ctggggcgtc gtggatgacg tggcgtcgcc
90841 ggttggccac ataactgcgg tgctcttcca gcagctgcgc gttcgggacc tggacgagct
90901 cgggcggggt gagtatctcc gaggaggacg acctggggcc ggggtggccc ccggtaacgt
90961 cccgggggatc caggggggagg tcctcgtcgt cttcgtatcc gccggcgatc tgttgggtta
91021 gaatttcggt ccacgagacg cgcgtctcgg tgccgccggc ggccggcggc agaggggccc
91081 tggtttccgt ggagcgcgag ctggtgtgtt cccggcggat ggcccgccgg gtctgagagc
91141 gactcggggg ggtccagtga cattcgcgca gcacatcctc cacggaggcg taggtgttat
91201 tgggatggag gtcggtgtgg cagcggacaa agagggccag gaactgggggg tagctcatct
91261 taaagtactt cagtatatcg cgacagttga tcgtgggaat gtagcaggcg ctaatatcca
91321 acacaatatc acagcccatc aacaggaggt cagtgtccgt ggtgtacacg tacgcgaccg
91381 tgttggtgtg atagaggttg gcgcaggcat cgtccgcctc caactgaccc gagttaatgt
91441 aggcgtaccc cagggcccgg agaacgcgaa tacagaacag atgcgccaga cgcagggccg
91501 gcttcgaggg cgcggcggac ggcagcgcgg ctccggaccc ggccgtcccc cgggtccccg
91561 aggccagaga ggtgccgcgt cggcgcatgt tggaaaaggc agagctgggt ctggagtcgg
91621 tgatgggggga aggcggtgga gaggcgtcca cgtcactggc ctcctcgtcc gtccggcact
91681 gggccgtcgt gcgggccagg atggccttgg ctccaaacac aaccggctcc atacaattga
91741 ccccgcgatc ggtaacgaag atggggaaaa gggacttttg ggtaaacacc tttaataagc
91801 gacagaggca gtgtagcgta atggcctcgc ggtcgtaact ggggtatcgg cgctgatatt
91861 tgaccaccaa cgtgtacatg acgttccaca ggtccacggc aatgggggtg aagtacccgg
91921 ccgggggcccc aaggccccgg cgcttgacca gatggtgtgt gtgggcaaac ttcatcatcc
91981 cgaacaaacc catgtcaggt cgattgtaac tgcggatcgg cctaactaag gcgtggttgg
92041 tgcgacggtc cgggacaccc gagcctgtct ctctgtgtat ggtgacccag acaacaacac
92101 cgacacaaga ggacaataat ccgttagggg acgctcttta taatttcgat ggcccaactc
92161 cacgcggatt ggtgcagcac cctgcatgcg ccggtgcggg ccaaccttcc ccccgctcat
92221 tgcctcttcc aaaagggtgt ggcctaacga gctgggggcg tatttaatca ggctagcgcg

FIG. 8BC

92281 gcgggcctgc cgtagtttct ggctcggtga gcgacggtcc ggttgcttgg gtcccctggc
92341 tgccatcaaa accccaccct cgcagcggca tacgcccccc ccgcgtcccg cacccgagac
92401 cccggcccgg ctgccctcac caccgaagcc cacctcgtca ctgtggggtg ttcccagccc
92461 gcgtcgggat gacggattcc cctggcggtg tggcccccgc ctcccacgtg gaggacgcgt
92521 cggacgcgtc cctcgggcag ccggaggagg gggcgccctg ccaggtggtc ctgcagggcg
92581 ccgaacttaa tggaatccta caggcgtttg ccccgctgcg cacgagcctt ctggactcgc
92641 ttctggttat gggcgaccgg ggcatcctta tccataacac gatctttggg gagcaggtgt
92701 tcctgccccc ggaacactcg caattcagtc ggtatcgctg gcgcggaccc acggcggcgt
92761 tcctgtctct cgtggaccag aagcgctccc tcctgagcgt gtttcgcgcc aaccagtacc
92821 cggacctacg tcgggtggag ttggcgatca cgggccaggc cccgtttcgc acgctggttc
92881 agcgcatatg gacgacgacg tccgacggcg aggccgttga gctagccagc gagacgctga
92941 tgaagcgcga actgacgagc tttgtggtgc tggttcccca gggaaccccc gacgttcagt
93001 tgcgcctgac gaggccgcag ctcaccaagg tccttaacgc gaccggggcc gatagtgcca
93061 cgcccaccac gttcgagctc ggggttaacg gcaaaatttc cgtgttcacc acgagtacct
93121 gcgtcacatt tgctgcccgc gaggagggcg tgtcgtccag caccagcacc caggtccaga
93181 tcctgtccaa cgcgctcacc aaggcgggcc aggcggccgc caacgccaag acggtgtacg
93241 gggaaaatac ccatcgcacc ttctctgtgg tcgtcgacga ttgcagcatg cgggcggtgc
93301 tccggcgact gcaggtcggc gggggcaccc tcaagttctt cctcacgacc cccgtcccca
93361 gtctgtgcgt caccgccacc ggtcccaacg ctgtatcggc ggtatttctc ctgaaacccc
93421 agaagatttg cctggactgg ctgggtcata gccaggggtc tccttcagcc gggagctcgg
93481 cctcccgggc ctctgggagc gagccaacag acagcaagga ctccgcgtcg gacgcggtca
93541 gccacggcga tccggaagac ctcgatggcg gtgcccgggc gggagaggcg ggggcctcgc
93601 acgcctgtcc gatgccgtcg tcgaccacgc gggtcactcc cacgaccaag cgggggcgct
93661 cgggggcga ggatgcgcgc gcggacacgg ccctaaagaa acctaagacg gggtcgccca
93721 ccgcaccccc gcccgcagat ccagtccccc tggacacgga ggacgactcc gatgcggcgg
93781 acgggacggc ggcccgtccc gccgctccag acgcccggag cggaagccgt tacgcgtgtt
93841 actttcgcga cctcccgacc ggagaagcaa gccccggcgc cttctccgcc ttccgggggg
93901 ggccccaaac cccgtatggt tttggattcc cctgacgggg cggggccttg gcggccgccc

FIG. 8BD

93961 aactctcgca ccatcccggg ttaatgtaaa taaacttggt attgcccaac actctcccgc
94021 gtgtcgcgtg tggttcatgt gtgtgcctgg cgtcccccac cctcgggttc gtgtatttcc
94081 tttccctgtc cttataaaag ccgtatgtgg ggcgctgacg gaaccacccc gcgtgccatc
94141 acggccaagg cgcgggatgc tccgcaacga cagccaccgg gccgcgtccc cggaggacgg
94201 ccagggacgg gtcgacgacg gacggccaca cctcgcgtgc gtgggggccc tggcgcgggg
94261 gttcatgcat atctggcttc aggccgccac gctgggtttt gcgggatcgg tcgttatgtc
94321 gcgcgggccg tacgcgaatg ccgcgtctgg ggcgttcgcc gtcgggtgcg ccgtgctggg
94381 ctttatgcgc gcaccccctc ccctcgcgcg gcccaccgcg cggatatacg cctggctcaa
94441 actggcggcc ggtggagcgg cccttgttct gtggagtctc ggggagcccg gcacgcagcc
94501 ggggggcccg gccccgggcc cggccaccca gtgcctggcg ctgggcgccg cctatgcggc
94561 gctcctggtg ctcgccgatg acgtctatcc gctctttctc ctcgccccgg ggcccctgtt
94621 cgtcggcacc ctggggatgg tcgtcggcgg gctgacgatc ggaggcagcg cgcgctactg
94681 gtggatcggt gggcccgccg cggccgcctt ggccgcggcg gtgttggcgg gcccgggggc
94741 gaccaccgcc agggactgct tctccagggc gtgccccgac caccgccgcg tctgcgtcat
94801 cgtcgcaggc gagtctgttt cccgccgccc cccggaggac ccagagcgac ccggggaccc
94861 cgggccaccg tccccccccga caccccaacg atcccagggg ccgccggccg atgaggtcgc
94921 accggccggg gtagcgcggc ccgaaaacgt ctgggtgccc gtggtcacct ttctgggggc
94981 gggcgcgctc gccgtcaaga cggtgcgaga acatgcccgg ggaacgccgg gcccgggcct
95041 gccgctgtgg ccccaggtgt ttctcggagg ccatgtggcg gtggccctga cggagctgtg
95101 tcaggcgctt gcgccctggg accttacgga cccgctgctg tttgttcacg ccggactgca
95161 ggtcatcaac ctcgggttgg tgtttcggtt ttccgaggtt gtcgtgtatg cggcgctagg
95221 gggtgccgtg tggatttcgt tggcgcaggt gctggggctc cggcgtcgcc tgcgcaggaa
95281 ggaccccggg gacggggccc ggttggcggc gacgcttcgg ggcctcttct tctccgtgta
95341 cgcgctgggg tttggggtgg gggcgctgct gtgccctccg gggtcaacgg gcgggcggtc
95401 gggcgattga tatattttc aataaaaggc attagtcccg aagaccgccg gtgtgtgatg
95461 atttcgccat aacacccaaa ccccggatgg ggcccgggta taaattccgg aaggggacac
95521 gggctaccct cactaccgag ggcgcttggt cgggaggccg catcgaacgc acacccccat
95581 ccggtggtcc gtgtggaggt cgttttcagt gcccggtctc gctttgccgg gaacgctagc

FIG. 8BE

95641 cgatccctcg cgagggggag gcgtagggca tggccccggg gcgggtgggc cttgccgtgg
95701 tcctgtggag cctgttgtgg ctcggggcgg gggtgtccgg gggctcggaa actgcctcca
95761 ccgggcccac gatcaccgcg ggagcggtga cgaacgcgag cgaggccccc acatcggggt
95821 ccccgggtc agccgccagc ccggaagtca cccccacatc gaccccaaac cccaacaatg
95881 tcacacaaaa caaaaccacc cccaccgagc cggccagccc cccaacaacc cccaagccca
95941 cctccacgcc caaaagcccc cccacgtcca cccccgaccc caaacccaag aacaacacca
96001 ccccccgccaa gtcgggccgc cccactaaac cccccgggcc cgtgtggtgc gaccgccgcg
96061 acccattggc ccggtacggc tcgcgggtgc agatccgatg ccggtttcgg aattccaccc
96121 gcatggagtt ccgcctccag atatggcgtt actccatggg tccgtccccc ccaatcgctc
96181 cggctcccga cctagaggag gtcctgacga acatcaccgc cccacccggg ggactcctgg
96241 tgtacgacag cgccccccaac ctgacggacc cccacgtgct ctgggcggag ggggccggcc
96301 cgggtgccga ccctccgttg tattctgtca ccgggccgct gccgacccag cggctgatta
96361 tcggcgaggt gacgcccgcg acccagggaa tgtattactt ggcctggggc cggatggaca
96421 gcccgcacga gtacgggacg tgggtgcgcg tccgcatgtt ccgccccccg tctctgaccc
96481 tccagcccca cgcggtgatg gagggtcagc cgttcaaggc gacgtgcacg gccgacgcct
96541 actacccgcg taacccccgtg gagttggtct ggttcgagga cgaccgccag gtgtttaacc
96601 cgggccagat cgacacgcag acgcacgagc accccgacgg gttcaccacc gtctctaccg
96661 tgacctccga ggctgtcggc ggccaggtcc ccccgcggac cttcacctgc cagatgacgt
96721 ggcaccgcga ctccgtgaca ttctcgcgac gcaatgccac cgggctggcc ctggtgctgc
96781 cgcggccaac catcaccatg gaatttgggg tccggcatgt ggtctgcacg gccggctgcg
96841 tccccgaggg cgtgacgttt gcctggttcc tgggggacga cccctcaccg gcggctaagt
96901 cggccgttac ggcccaggag tcatgcgacc accccgggct ggctacggtc cggtccaccc
96961 tgcccatttc gtacgactac agcgagtaca tctgtcggtt gaccggatat ccggccggga
97021 ttcccgttct agaacaccac ggcagtcacc agccccccacc cagggacccc accgagcggc
97081 aggtgatcga ggcgatcgag tgggtgggga ttggaatcgg ggttctcgcg gcgggggtcc
97141 tggtcgtaac ggcgatcgtg tacgtcgtcc gcacatcaca gtcgcggcag cgtcatcggc
97201 ggtaacgcga gacccccccg ttacctttttt aatatctata tagtttggtc cccctctat
97261 cccgcccacc gctgggcgct ataaagccgc cacccctctct tccctcaggt catccttggt

FIG. 8BF

97321 cgatcccgaa cgacacacgg cgtggagcaa aacgcctccc cctgagccgc tttcctacca
97381 acacaccggc atgcctctgc gggcatcgga acacgcctac cggcccctgg gccccgggac
97441 acccccccatg cgggctcggc tccccgccgc ggcctgggtt ggcgtcggga ccatcatcgg
97501 gggagttgtg atcattgccg cgttggtcct cgtgccctcg cgggcctcgt gggcactttc
97561 cccatgcgac agcggatggc acgagttcaa cctcgggtgc atatcctggg atccgacccc
97621 catggagcac gagcaggcgg tcggcggctg tagcgccccg gcgaccctga tccccccgcgc
97681 ggctgccaaa cagctggccg ccgtcgcacg cgtccagtcg gcaagatcct cgggctactg
97741 gtgggtgagc ggagacggca ttcgggcctg cctgcggctc gtcgacggcg tcggcggtat
97801 tgaccagttt tgcgaggagc ccgcccttcg catatgctac tatccccgca gtcccgggggg
97861 ctttgttcag tttgtaactt cgacccgcaa cgcgctgggg ctgccgtgag gcgcgtgtac
97921 tgcggtctgt ctcgtctcct cttctccccc tccctccccc tccgcatccc aggatcacac
97981 cggccaacga gggttggggg gtccggcacg gacccaaaat aataaacaca caatcacgtg
98041 cgataaaaag aacacgcggt cccctgtggt gtttttggtt atttttatta aatctcgtcg
98101 acaaacaggg ggaaaggggc gtggtctagc gacggcagca cgggcggagg cgttcaccgg
98161 ctccggcgtc cttcgcgttt aagcttggtc aggagggcgc tcagggcggc gacgttggtc
98221 gggccgtcgt tggtcagggc gttggctcga tggcgggcga ggacgggcga ggggctcaac
98281 ggcgggggcg ggggcccggt gcggcccggg ggggaaaata gggcggatcc ccccagtcg
98341 tacaggggat tttccgcctc aatgtacggg gaggccggcg ctgcattcgc cgtgttcacg
98401 cagacgtttt cgtagacccg catccatggt atttcctcgt agacacgccc cccgtcctcg
98461 cgcaccgtct cgtatattga ctcgtcgtcc tcgtaggggg cgtgccgttc gcgggccgag
98521 gcggcgtggg tggctttgcg gcgggcgtcg tcgtcgtcgt cgtcggccgt cagatacgtg
98581 gcttccatct ggtcgggttc tccctccggg gcgggtcccc acacccgtgg ccgatcgagg
98641 ctccccagag acgcgcgccg gacgaggagg gggcacgtcg ccgccggcgg tcgcctgtcg
98701 ggtcccgcga cgttacgggc cgggaggcgc gggggcacct cccccatgtg cgtgtaatac
98761 gtggccggct gtgcggccgc agcgggggggc tcggcgaccg ggtcgttcgc atccggaagc
98821 gggggccccg cgccgtccgc gcggcgcctc cggaacctcc gggtggacgc gggggtcgag
98881 tgtaggcgag gtcgggggag gggcgggggc tcgttgtcgc gccgcgcccg ctgaatcttt
98941 tcccgacagg tcccaccccc cgcgcgatgc cccccgggc cgctggccat gtcgtccggg

FIG. 8BG

99001 ggaggccccg cggaccacgt cgtccggcga gacgccacga gccgcaggat ggactcgtag
99061 tggagcgacg gcgccccgct gcggagcaga tccgcggcca gggcggcccc gaaccaagcc
99121 ttgatgctca actccatccg ggcccagctg ggggcggtca tcgtgggaa caggggggcg
99181 gtggtccgac agaaacgctc ctggctgtcc accgcggccc gcagatactc gttgttcagg
99241 ctgtcggtgg cccagacgcc gtacccggtg agggtcgcgt tgatgatata ctgggcgtgg
99301 tgatggacga tcgacagaac ctccaccgtg gatacgacgg tatccacggt cccgtacgta
99361 ccgccgctcc gcttgccggt ctgccacagg ttggctaggc gcgtcaggtg gcccaggacg
99421 tcgctgaccg ccgccctgag cgccatgcac tgcatggagc cggttgtgcc gctgggaccc
99481 cggtccagat ggcgcgcgaa cgtttccgcg ggcgcctccg ggctgccgcc gagcgggagg
99541 aaccggcgat tggagggact cagccggtga catacgtgct tgtctgtcgt ccacagcatc
99601 caggacgccc accggtacag cacggagacg taggccagga gctcgttgag ccgcagtgcg
99661 gtgtcggtgc tggggcggct tgggtccgcc gggcgcataa agaacatgta ctgctgaatc
99721 cgatggaggg cgtcgcgcag gccggccacg gtggcggcgt acttggccgc cacggccccg
99781 ctcttgaacg gggtgcgcgc cagcagcttt ggcgccaggg tgggccgcag cagcacgtga
99841 aggctggggt cgcagtcgcc cacggggtcc tcggggacgt ccaggccgct gggcaccacc
99901 gtctgcaggt acttccagta ctgcgtgagg atggcgcggc tcaactggcc gccgggcagc
99961 tccacctcgc ccagcgcctg ggtggcggcc gaagcgtagt gccggatgta ctcgtagtgc
100021 gggtcgctgg cgagcccgtc cacgatcaaa ctctcgggaa ccgtgttgtg ttgccgcgcg
100081 gccaaccgga cgctgcgatc ggtgcaggtc agaaacgccg gctgcgcgtc gtcggagcgc
100141 tgccgcaagg cgcccacggc cgcgctaagg agcccctccg gggtggggag cagacacccg
100201 ccgaagatgc gccgctcggg aacgcccgcg ttgtcgccgc ggatcaggtt ggcaggcgtc
100261 aggcaccgcg ccagccgcag ggagctcgcg ccgcgcgtcc ggcgctgcat ggtgacgccc
100321 gttcggtcgg gacccgccgg tcggagttat gccgcgtcca gggccatcgg ggcgcttttt
100381 atcgggagga gcttatgggc gtggcgggcc tcccagcccg gtcgcgcgcc tccccgacac
100441 gtgcgcccgc agggcggcgg ccccctcgtc tcccatcagc agtttcctaa actgggacat
100501 gatgtccacc acgcggaccc gcgggcccaa cacggacccg ccgcttacgg gggcgggggg
100561 gaagggctcc aggtccttga gaagaaaggc ggggtctgcc gtcccggaca cgggggcccg
100621 gggcgctgag gaggcggggc gcagatccac gtgctccgcg gccgcgcgga cgtccgccca

FIG. 8BH

100681 gaacttggcg ggggtggtgc gcgcgtacag gggctgggtc gctcggagga cgcacgcgta
100741 gcgcaggggg gtgtacgtgc ccacctcggg ggccgtgaat ccccgtcaa acgcggccag
100801 tgtcacgcac gccaccacgg tgtcggcaaa gcccagcagc cgctgcagga cgagcccggc
100861 ggccagaatg gcgcgcgtgg ccgccgcgtc gtcccggcgc cggtgcgcgt ccccgcacgc
100921 ccgggcgtac tttaaggtca cggtcgccag ggccgtgtgc agcgcgtaca ccgcagcgcc
100981 cagcacggcg ttgagcccgc tgttggcgag cagccggcgc gctgcggtgt cgcccagcgc
101041 ctcgtgctcg gcccccacga ccgcggggct tcccaggggc agggcgcgaa acagctcctc
101101 ccgcgccacg tccgcaaagg cggggtggtg cacgtgcggg tgcaggcgcg cccccacgac
101161 caccgagagc cactggaccg tctgctccgc catcaccgcc agcacatcca gcacgcgccc
101221 caggaaggcg gcctcccgcg tcaaaacgca ccggacggcg tcgggattga agcgggcgag
101281 cagggccccg gtggccaggt acgtcatgcg gccggcatag cgggcggcca cgcgacagtc
101341 gcggtccagc agcgcgcgca ccccgggcca gtacagcagg gaccccagcg agctgcggaa
101401 caccgcggcg tcggggccgg attgggggga cactaacccc cccgcgctca gtaacggcac
101461 ggccgcggcc ccgacgggac gcaacgccgt gaggctcgcg aactgccgcc tcagctcggc
101521 cgccctgtcg tccaggtcag acccgcgcgc ctccgcgtga aggcgcgtcc cgcacaccca
101581 cccgttgatg gccagccgca cgacggcatc cgccaaaaag ctcatcgcct gggcgggggct
101641 ggtttttgtt cgacgatccg tcaggtcaag aatcccatcg cccgtgatat accaggccaa
101701 cgcctcgccc tgctgcaggg tttggcggaa aaacaccgcg gggttgtcgg gggaggcgaa
101761 gtgcatgacc cccacgcgcg ataacccgaa cgcgctatcc ggacacgggt aaaacccggc
101821 cggatgcccc agggctaggg cggagcgcac ggactcgtcc cacacggcaa cctgaggggc
101881 cagtcgatcc aacgggaatg ccgcccggag ctccgggccc ggcacgcgtc cctccagaac
101941 ctccaccttg ggcggggaac gggccccgcc gccgtcctcc ggcccgacgg cttccgggta
102001 gtcgtcctcc tcgtactgca gctcctctag gaacagcggc gacggcgcca cccgcgaacc
102061 gccgacccgc cccaaaatag cccgcgcgtc gacgggaccc aggtatcccc cctgccgggc
102121 ctgcggagga ccgcggggaa cctcatcatc atcgtccagg cgaccgcgca ccgactggct
102181 acgggccgca tcgggcccgg ggcgctgccg ggacgctcgg cgatgggatg tgggcggggc
102241 ttccgacgcg cgccgtcgtc gggctcgcgg gccttcccgt cgacggcgca cgggcggctc
102301 gtcgcccgcc atctcctcca gagcctctag ctcgctgtcg tcatccccgc ggaacaccgc

FIG. 8BI

```
102361 acgcaggtac cccatgaacc ccaccccatc gcccgctggc tcgtccgcca cgggcgaggc
102421 gcgggggcgg gtggatgcgc gcctcctgcg ccccgcgggt tcgcgagccg acatggtggc
102481 gatagacgcg ggttatcgga tgtccgctac cccccaaaaa agaaaaagac cccacagcgc
102541 ggatggaggc cggggtaggt gccgccggac cccctcgcga tgggaatgga cgggagcgac
102601 ggggccggcg caaaaaaacg cagtatctcc cgcgaaggct accccgccgcc ccagcccccg
102661 gccaaatgcg gaaacggtcc cgcgctctcg cctttatacg cgggccgccc tgcgacacaa
102721 tcacccgtcc gtggtttcga atctacacga cagggccgca gacgcggcta acacacacgc
102781 cggcaaccca gaccccagtg ggttggttgc gcggtccgct ctcctggcta gttctttccc
102841 ccaccaccaa ataatcagac gacaaccgca ggttttgtaa tgtatgtgct cgtgtttatt
102901 gtggatacga accggtgacg ggaggggaaa acccagacgg gggatgcggg tccggtcgcg
102961 ccccctaccc accgtactcg tcaattccaa gggcatcggt aacaatctgc tcaaactcga
103021 agtcggccat atccagagcg ccgtaggggg cggagtcgtg gggggtaaat cccggccccg
103081 gggaatcccc gtcccccaac atgtccagat cgaaatcgtc tagcgcgtcg gcatgcgcca
103141 tcgccacgtc ctcgccgtct aagtggagct cgtcccccag gctgacatcg gtcgggggggg
103201 ccgtcgacag tctgcgcgtg tgtcccgcgg ggagaaagga caggcgcgga gccgccagcc
103261 ccgcctcttc ggggggcgtcg tcgtccggga gatcgagcag gccctcgatg gtagacccgt
103321 aattgttttt cgtacgcgcg cggctgtacg cgtgttcccg catgaccgcc tcggagggcg
103381 aggtcgtgaa gctggaatac gagtccaact tcgcccgaat caacaccata aagtacccag
103441 aggcgcgggc ctgggtgcca tgcagggtgg gaggggtcgt caacggcgcc cctggctcct
103501 ccgtagccgc gctgcgcacc agcgggaggt taaggtgctc gcgaatgtgg tttagctccc
103561 gcagccggcg ggcctcgatt ggcactcccc ggacggtgag cgctccgttg acgaacatga
103621 agggctggaa cagacccgcc aactgacgcc agctctccag gtcgcaacag aggcagtcaa
103681 acaggtcggg ccgcatcatc tgctcggcgt acgcggccca taggatctcg cgggtcaaaa
103741 atagatacaa atgcaaaaac agaacacgcg ccagacgagc ggtctctcgg tagtacctgt
103801 ccgcgatcgt ggcgcgcagc atttctccca ggtcgcgatc gcgtccgcgc atgtgcgcct
103861 ggcggtgcag ctgccggacg ctggcgcgca ggtaccggta cagggccgag cagaagttgg
103921 ccaacacggt tcgatagctc tcctcccgcg cccgtagctc ggcgtggaag aaacgagaga
103981 gcgcttcgta gtagagcccg aggccgtcgc gggtggccgg aagcgtcggg aaggccacgt
```

FIG. 8BJ

104041 cgccgtgggc gcgaatgtcg atttgggcgc gttcggggac gtacgcgtcc ccccattcca
104101 ccacatcgct gggcagcgtt gataggaatt tacactcccg gtacaggtcg gcgttggtcg
104161 gtagcgccga aaacagatcc tcgttccagg tatcgagcat ggtacatagc gcggggcccg
104221 cgctaaagcc caagtcgtcg aggagacggt taaagagggc ggcgggggcg acgggcatgg
104281 gtggggaggg catgagctgg gcctggctca ggcgcccccgt tgcgtacagc ggggggccg
104341 ccggggtgtt tttgggaccc ccggctgggc gggggggcgg tggcgaagcg ccgtccgcgt
104401 tcatgtcggc aaacagctcg tcgaccaaga ggtccattgg gtggggttga tacgggaaag
104461 acgatatcgg gcttttgatg cgatcgtccc cgcccgccca gagagtgtgg gacgcccgac
104521 ggcgcgggaa gagaaaaacc cccaaacgcg ttagaggacc ggacggacct tatgggggga
104581 agtgggcagc gggaaccccg tccgttcccg aggaatgaca gcccgtggtc gccaccacgc
104641 atttaagcaa cccgcacggg ccgccccgta cctcgtgact tccccccaca ttggctcctg
104701 tcacgtgaag gcgaaccgag ggcggctgtc caacccaccc cccgccaccc agtcccggtc
104761 cccgtcggat tgggaaacaa aggcacgcaa cgccaacacc gaatgaaccc ctgttggtgc
104821 tttattgtct gggtacggaa gttttcactc gacgggccgt ctggggcgag aagcggagcg
104881 ggctggggct cgaggtcgct cggtggggcg cgacgccgca gaacgccctc gagtcgccgt
104941 ggccgcgtcg acgtcctgca ccacgtctgg attcaccaac tcgttggcgc gctgaagcag
105001 gtttttgccc tcgcagaccg tcacgcggat ggtggtgatg ccaaggagtt cgttgaggtc
105061 ttcgtctgtg cgcggacgcg acatgtccca gagctggacc gccgccatcc gggcatgcat
105121 ggccgccagg cgcccgaccg cggcgcagaa gacgcgcttg ttaaagccgg ccacccgggg
105181 ggtccatggc gcgtcggggt ttgggggggc ggtgctaaag tgcagctttc tggccagccc
105241 ctgcgcgggt gtcttggatc gggttggcgc cgtcgacgcg ggggcgtctg ggagtgcggc
105301 ggattctggc tgggccgatt tcctgccgcg ggtggtctcc gccgccgggg ccgcgggggc
105361 cttagtcgcc acccgctggg ttcgggggggc ccgggggggcg gtggtgggtg tgcgtccggc
105421 ccctccggac ccagcgggtg gcggaggtgc ccgcgcaggc cccgggccgg acaaaaccgc
105481 cccggaaacg ggacgccgcg tccgggggac ctccgggtgt tcgtcgtctt cggatgacga
105541 gccccccgtag agggcataat ccgactcgtc gtactggacg aaacggacct cgcccctctg
105601 gcgcgagcgt gtctgtaggg cgccacggcg ggaggtgtca ggcggactat cgggactcgc
105661 catacctgaa gacggggtgt agtacagatc ctcgtactca tcgcgcggaa cctcccgcgg

FIG. 8BK

105721 acccgacttc acggagcggc gagaggtcat ggttccacga acacgctagg gtcggatgcg
105781 cggacaatta ggcctgggtt cggacggcgg gggtggtgca ggtgtggaga ggtcgagcga
105841 taggggcggc ccgggagaga agagagggtc cgcaaaaccc actggggatg cgtgagtggc
105901 cctctgtggg cggtggggga gagtcttata ggaagtgcat ataaccacaa cccatgggtc
105961 taaccaatcc ccaggggcca agaaacagac acgccccaaa cggtctcggt ttccgcgagg
106021 aaggggaagt cctgggacac cctccacccc cacccctcac cccacacagg gcgggttcag
106081 gcgtgcccgg cagccagtag cctctggcag atctgacaga cgtgtgcgat aatacacacg
106141 cccatcgagg ccatgcctac ataaaagggc accagggccc ccggggcaga catttggcca
106201 gtgttttggg tctcgcaccg cgcgcccccg atcccatcgc gcccgccctc ctcgccgggc
106261 ggctccccgt gcgggcccgc gtctcccgcc gctaaggcga cgagcaagac aaacaacagg
106321 cccgcccgac agacccttct ggggggggccc atcgtcccta acaggaagat gagtcagtgg
106381 ggatccgggg cgatccttgt ccagccggac agcttgggtc gggggtacga tggcgactgg
106441 cacacggccg tcgctactcg cgggggcgga gtcgtgcaac tgaacctggt caacaggcgc
106501 gcggtggctt ttatgccgaa ggttagcggg gactccggat gggccgtcgg gcgcgtctct
106561 ctggacctgc gaatggctat gccggctgac ttttgcgcga ttattcacgc ccccgcgcta
106621 gccagccccg ggcaccacgt aatactgggt cttatcgact cggggtaccg cggaaccgtt
106681 atggccgtgg tcgtagcgcc taaaaggacg cgggaatttg ccccgggac cctgtgggtc
106741 gacgtgacgt tcctggacat cctggcgacc ccccggcccc tcaccgagcc gatttccctg
106801 cggcagttcc cgcaactggc gccccccct ccaaccgggg ccgggatacg cgaagatcct
106861 tggttggagg gggcgctcgg ggccccaagc gtgactacgg ccctaccggc gcgacgccga
106921 gggcggtccc tcgtctatgc cggcgagctg acgccggttc agacggaaca cggggacggc
106981 gtacgagaag ccatcgcctt ccttccaaaa cgcgaggagg atgccggttt cgacattgtc
107041 gtccgtcgcc cggtcaccgt cccggcaaac ggcaccacgg tcgtgcagcc atccctccgc
107101 atgctccacg cggacgccgg gcccgcggcc tgctatgtgt tggggcggtc gtcgctcaac
107161 gcccgcggcc tcctggtcgt tcctacgcgc tggctccccg ggcacgtatg tgcgtttgtt
107221 gtttacaacc ttacgggggt tcctgtgacc ctcgaggccg gcgccaaggt cgcccagctc
107281 ctggttgcgg gggcggacgc tcttccttgg atccccccgg acaactttca cgggaccaaa
107341 gcgcttcgaa actaccccag gggtgttccg gactcaaccg ccgaacccag gaacccgccg

FIG. 8BL

```
107401 ctcctggtgt ttacgaacga gtttgacgcg gaggcccccc cgagcgagcg cgggaccggg
107461 ggttttggct ctaccggtat ttagcccaca gctttgggtt cgttccgggc aataaaaaac
107521 gtttgtatcg catctttcct gtgtgtagtt gtttatgttg gatgcctgtg ggtctatcac
107581 acccgccgct ccatcccaca aacacaaaac acacgggttg gatgaaaaca cgcatttatt
107641 gacccaaaac acacggagct gctcgagatg ggccagggcg aggtgcggtt ggggaggctg
107701 taggtctggg aacggacacg cggggacacg attccggttt ggggtccggg agggcgtcgc
107761 cgtttcgggc ggcaggcgcc agcgtaacct ccgggggcgg cgtgtggggg tgccccaagg
107821 agggcgcctc ggtcacccca atccccccgg accgggttcc cccggcaacc ccgaaggcgg
107881 agaggccaag ggcccgttcg gcgatggcca catcctccat gaccacgtca ctctcggcca
107941 tgctccgaat agcctgggag acgagcacat ccgcggactt gtcagccgcc cccacggaca
108001 tgtacatctg caggatggtg gccatacacg tgtccgccag gcgccgcatc ttgtcctgat
108061 gggccgccac ggccccgtcg atcgtggggg cctcgagccc ggggtggtgg cgcgccagtc
108121 gttctaggtt caccatgcag gcgtggtacg tgcgggccaa ggcgcgggcc ttcacgaggc
108181 gtcgggtgtc gtccagggac cccagggcgt catcgagcgt gatgggggcg ggaagtagcg
108241 cgttaacgac cgccagggcc tcctgcagcc gcggctccgc ctccgagggc ggaacggccg
108301 cgcggatcat ctcatattgt tcctcggggc gcgctcccca gccacatata gccccgagaa
108361 gagaagccat cgcgggcggg tactggccct tgggcgcgcg gacgcaatgg ggcaggaaga
108421 cgggaaccgc ggggagaggc gggcggccgg gactcccgtg gaggtgaccg cgctttatgc
108481 gaccgacggg tgcgttatta cctcttcgat cgccctcctc acaaactctc tactgggggc
108541 cgagccggtt tatatattca gctacgacgc atacacgcac gatggccgtg ctgacgggcc
108601 cacggagcaa gacaggttcg aagagagtcg ggcgctctac caagcgtcgg gcgggctaaa
108661 tggcgactcc ttccgagtaa ccttttgttt attggggacg gaagtgggtg ggacccacca
108721 ggcccgcggg cgaacccgac ccatgttcgt ctgtcgcttc gagcgagcgg acgacgtcgc
108781 cgcgctacag gacgccctgg cgcacgggac cccgctacaa ccggaccaca tcgccgccac
108841 cctggacgcg gaggccacgt tcgcgctgca tgcgaacatg atcctggctc tcaccgtggc
108901 catcaacaac gccagccccc gcaccggacg cgacgccgcc gcggcgcagt atgatcaggg
108961 cgcgtcccta cgctcgctcg tggggcgcac gtccctggga caacgcggcc ttaccacgct
109021 atacgtccac cacgaggcgc gcgtgcttgc cgcgtaccgc agggcgtatt atggaagcgc
```

FIG. 8BM

109081 gcagagtccc ttctggtttc ttagcaaatt cgggccggac gaaaaaagcc tggtgctcac
109141 cactcggtac tacctgcttc aggcccagcg tctgggggggc gcgggggcca cgtacgacct
109201 gcaggccatc aaggacatct gcgccaccta cgcgattccc cacgcccccc gccccgacac
109261 cgtcagcgct gcgtccctga cctcgtttgc cgccatcacg cggttctgtt gcacgagcca
109321 gtacgcccgc ggggccgcgg cggccgggtt tccgctttac gtggagcgcc gtattgcggc
109381 cgacgtccgc gagaccagtg cgctggagaa gttcataacc cacgatcgca gttgcctgcg
109441 cgtgtccgac cgtgaattca ttacgtacat ctacctggcc cattttgagt gtttcagccc
109501 cccgcgccta gccacgcatc ttcgggccgt gacgacccac gaccccaacc ccgcggccag
109561 cacggagcag ccctcgcccc tgggcaggga ggccgtggaa caatttttt gtcacgtgcg
109621 cgcccaactg aatatcgggg agtacgtcaa acacaacgtg acccccgggg agaccgtcct
109681 ggatggcgat acggccaagg cctacctgcg cgctcgcacg tacgcgcccg gggccctgac
109741 gcccgccccc gcgtattgcg gggccgtgga ctccgccacc aaaatgatgg ggcgtttggc
109801 ggacgccgaa aagctcctgg tccccgcgcg gtggcccgcg tttgcgcccg ccagtcccgg
109861 ggaggacacg gcgggcggca cgccgccccc acagacctgc ggaattgtca agcgcctcct
109921 gagactggcc gccacggaac agcagggcac cacacccccg gcgatcgcgg cgcttatccg
109981 taatgcggcg gtgcagactc ccctgcccgt ctaccggata tccatggtcc ccacgggaca
110041 ggcatttgcc gcgctggcct gggacgactg ggcccgcata acgcgggacg ctcgcctggc
110101 cgaagcggtc gtgtccgccg aagcggcggc gcaccccgac cacggcgcgc tgggcaggcg
110161 gctcacggat cgcatccgcg cccagggccc cgtgatgccc cctggcggcc tggatgccgg
110221 ggggcagatg tacgtgaatc gcaacgagat attcaacggc gcgctggcaa tcacaaacat
110281 catcctggat ctcgacatcg ccctgaagga gcccgtcccc tttcgccggc tccacgaggc
110341 cctgggccac tttaggcgcg gggctctggc tgcggttcag ctcctgtttc ccgcggcccg
110401 cgtggacccc gacgcatatc cctgttattt tttcaaaagc gcatgtcggc ccggcccggc
110461 atccgtgggt tccggcagcg gactcggcaa cgacgacgac ggggactggt ttccctgcta
110521 cgacgacgcc ggtgatgagg agtgggcgga ggacccgggc gccatggaca catcccacga
110581 tccccccggac gacgaggttg cctactttga cctgtgccac gaagtcggcc ccacggcgga
110641 acctcgcgaa acggattcgc ccgtgtgttc ctgcaccgac aagatcggac tgcgggtgtg
110701 catgcccgtc cccgccccgt acgtcgtcca cggttctcta acgatgcggg gggtggcacg

FIG. 8BN

110761 ggtcatccag caggcggtgc tgttggaccg agattttgtg gaggccatcg ggagctacgt
110821 aaaaaacttc ctgttgatcg atacgggagt gtacgcccac ggccacagcc tgcgcttgcc
110881 gtattttgcc aaaatcgccc ccgacgggcc tgcgtgcgga aggctgctgc cagtgtttgt
110941 gatccccccc gcctgcaaag acgttccggc gtttgtcgcc gcgcacgccg acccgcggcg
111001 cttccatttt cacgccccgc ccacctatct cgcttccccc cgggagatcc gtgtcctgca
111061 cagcctgggt ggggactatg tgagcttctt tgaaaggaag gcgtcccgca acgcgctgga
111121 acactttggg cgacgcgaga ccctgacgga ggtcctgggt cggtacaacg tacagccgga
111181 tgcgggaggg accgtcgagg ggttcgcatc ggaactgctg gggcggatag tcgcgtgcat
111241 cgaaacccac tttcccgaac acgccggcga atatcaggcc gtatccgtcc ggcgggccgt
111301 cagtaaggac gactgggtcc tcctacagct agtccccgtt cgcggtaccc tgcagcaaag
111361 cctgtcgtgt ctgcgcttta agcacggccg ggcgagtcgc gccacggcgc ggacattcgt
111421 cgcgctgagc gtcggggcca acaaccgcct gtgcgtgtcc ttgtgtcagc agtgctttgc
111481 cgccaaatgc gacagcaacc gcctgcacac gctgtttacc attgacgccg gtacgccatg
111541 ctcgccgtcc gttccctgca gcacctctca accgtcgtct tgataacggc gtacggcctc
111601 gtgctcgtgt ggtacaccgt cttcggtgcc agtccgctgc accgatgtat ttacgcggta
111661 cgccccaccg gcaccaacaa cgacaccgcc ctcgtgtgga tgaaaatgaa ccagacccta
111721 ttgtttctgg gggccccgac gcacccccc aacgggggct ggcgcaacca cgcccatatc
111781 tgctacgcca atcttatcgc gggtagggtc gtgcccttcc agtccccacc cgacgccatg
111841 aatcgtcgga tcatgaacgt ccacgaggca gttaactgtc tggagaccct atggtacaca
111901 cgggtgcgtc tggtggtcgt agggtggttc ctgtatctgg cgttcgtcgc cctccaccaa
111961 cgccgatgta tgtttggtgt cgtgagtccc gcccacaaga tggtggcccc ggccacctac
112021 ctcttgaact acgcaggccg catcgtatcg agcgtgttcc tgcagtaccc ctacacgaaa
112081 attacccgcc tgctctgcga gctgtcggtc cagcggcaaa acctggttca gttgtttgag
112141 acggacccgg tcaccttctt gtaccaccgc cccgccatcg gggtcatcgt aggctgcgag
112201 ttgatgctac gctttgtggc cgtgggtctc atcgtcggca ccgctttcat atcccgggcg
112261 gcatgtgcga tcacataccc cctgtttctg accatcacca cctggtgttt tgtctccacc
112321 atcggcctga cagagctgta ttgtattctg cggcggggcc cggccccaa gaacgcagac
112381 aaggccgccg ccccggggcg atccaagggg ctgtctggcg tctgcgggcg ctgttgttcc

FIG. 8BO

112441 atcatcctct cgggcatcgc agtgcgattg tgttatatcg ccgtggtggc cggggtggtg
112501 ctcgtggcgc ttcactacga gcaggagatc cagaggcgcc tgtttgatgt atgacgtcac
112561 atccaggccg gcggaaaccg gaacggcata tgcaaattgg aaactgtcct gtcttggggc
112621 ccacccaccc gacgcgtcat atgcaaatga aaatcggtcc cccgaggcca cgtgtagcct
112681 ggatcccaac gacccegccc atgggtccca attggccgtc ccgttaccaa gaccaaccca
112741 gccagcgtat ccaccccegc ccgggtcccc gcggaagcgg aacggtgtat gtgatatgct
112801 aattaaatac atgccacgta cttatggtgt ctgattggtc cttgtctgtg ccggaggtgg
112861 ggcgggggccc ccgcccgggg ggcggaacga ggaggggttt gggagagccg gccccggcac
112921 cacgggtata aggacatcca ccacccggcc ggtggtggtg tgcagccgtg ttccaaccac
112981 ggtcacgctt ctgtgcctct ccccgattcg ggcccggtcg ctcgctaccg gtgcaccacc
113041 accagaggcc atatccgaca ccccagcccc gacggcagcc gacagcccgg tcatggcgac
113101 tgacattgat atgctaattg acctcggcct ggacctctcc gacagcgatc tggacgagga
113161 ccccccccgag ccggcggaga gccgccgcga cgacctggca tcggacagca gcggggagtg
113221 ttcctcgtcg gacgaggaca tggaagaccc ccacggagag gacggaccgg agccgatact
113281 cgacgccgct cgcccggcgg tccgcccgtc tcgtccagaa gaccccggcg tacccagcac
113341 ccagacgcct cgtccgacgg agcggcaggg ccccaacgat cctcaaccag cgccccacag
113401 tgtgtggtcg cgcctcgggg cccggcgacc gtcttgctcc cccgagcagc acgggggcaa
113461 ggtggcccgc ctccaacccc caccgaccaa agcccagcct gcccgcggcg gacgccgtgg
113521 gcgtcgcagg ggtcggggtc gcggtggtcc cggggccgcc gatggtttgt cggacccccg
113581 ccggcgtgcc cccagaacca atcgcaaccc ggggggaccc cgccccgggg cggggtggac
113641 ggacggcccc ggcgcccccc atggcgaggc gtggcgcgga agtgagcagc ccgacccacc
113701 cggaggcccg cggacacggg gcgtgcgcca agcacccccc ccgctaatga cgctggcgat
113761 tgcccccccg cccgcggacc cccgcgcccc ggccccggag cgaaaggcgc ccgccgccga
113821 caccatcgac gccaccacgc ggttggtcct gcgctccatc tccgagcgcg cggcggtcga
113881 ccgcatcagc gagagctttg gccgcagcgc acaggtcatg cacgacccct ttgggggcga
113941 gccgtttccc gccgcgaata gcccctgggc cccggtgttg gcgggccaag gagggccctt
114001 tgacgccgag accagacggg tgtcctggga aaccttggtc gcccacggcc cgagcctcta
114061 tcgcactttt gccggcaatc ctcgggccgc atcgaccgcc aaggccatgc gcgactgcgt

FIG. 8BP

114121 gctgcgccaa gaaaatttca tcgaggcgct ggcctccgcc gacgagacgc tggcgtggtg
114181 caagatgtgc atccaccaca acctgccgct gcgcccccag gaccccatta tcgggacggc
114241 cgcggctgtg ctggataacc tcgccacgcg cctgcggccc tttctccagt gctacctgaa
114301 ggcgcgaggc ctgtgcggcc tggacgaact gtgttcgcgg cggcgtctgg cggacattaa
114361 ggacattgca tccttcgtgt ttgtcattct ggccaggctc gccaaccgcg tcgagcgtgg
114421 cgtcgcggag atcgactacg cgacccttgg tgtcggggtc ggagagaaga tgcatttcta
114481 cctccccggg gcctgcatgg cgggcctgat cgaaatccta gacacgcacc gccaggagtg
114541 ttcgagtcgt gtctgcgagt tgacggccag tcacatcgtc gcccccccgt acgtgcacgg
114601 caaatatttt tattgcaact ccctgtttta ggtacaataa aaacaaaaca tttcaaacaa
114661 atcgccccac gtgttgtcct tctttgctca tggccggcgg ggcgtgggtc acggcagatg
114721 gcgggggtgg gcccggcgta cggcctgggt gggcggaggg aactaaccca acgtataaat
114781 ccgtccccgc tccaaggccg gtgtcatagt gcccttagga gcttcccgcc cgggcgcatc
114841 ccccctttg cactatgaca gcgacccccc tcaccaacct gttcttacgg gccccggaca
114901 taacccacgt tgccccccct tactgcctca acgccacctg gcaggccgaa acggccatgc
114961 acaccagcaa aactgactcc gcttgcgtgg ccgtgtggag ttacctggtc cgcgcctcct
115021 gtgagaccag cggcacaatc cactgctttt tctttgtggt atacaaggac acccaccata
115081 cccctccgct gattaccgag ctccgcaact ttgcggacct ggttaaccac ccgccggtcc
115141 tacgcgaact ggaggataag cgcggggtgc ggctgcggtg tgcgcggccg tttagcgtcg
115201 ggacgattaa ggacgtctct gggtccggcg cgtcctcggc gggagagtac acgataaacg
115261 ggatcgtgta ccactgccac tgtcggtatc cgttctcaaa aacatgctgg atgggggcct
115321 ccgcggccct acagcacctg cgctccatca gctccagcgg catggccgcc cgcgcggcag
115381 agcatcgacg cgtcaagatt aaaattaagg cgtgatttcc aacccccccat gaatgtgtgt
115441 aacccccccc aaaaaaataa agagccgtaa cccaaccaaa ccaggcgtgg tgtgagtttg
115501 tggacccaaa gccctcagag acaacgcgac aggccagtat ggaccgtgag acttttattt
115561 attaactcac aggggcgctt accgccacag gaataccaga ataatgacca ccacaatcgc
115621 gaccacccca aatacagcat ggcgccccac cacgccacaa cagccctgtc gccggtatgg
115681 ggcatgatca gacgagccgc gagccgcgcg ttgggccctg tacagctcgc gcgaattgac
115741 cctaggaggc cgccacgcgc ccgagttttg cgttcgtcgc tggtcgtcgg gcaccaaagc

FIG. 8BQ

115801 cccggacggc tgttcggtcg aacgaacggc cacgacagtg gcataggttg gggggtggtc
115861 cgacatagcc tcggcgtacg tcgggaggcc cgacaagagg tcccttgtga tgtcgggtgg
115921 ggccacaagc ctggtttccg gaagaaacag gggggttgcc aataacccgc cagggccaaa
115981 actccggccc tggcgcacgt cgttcggcgc ggcgccgggc gcgccgagcg gctcgctggg
116041 cggcttggcg tgagcggccc cgctccgacg cctcgccctc tccggaggag gttggtggaa
116101 ttggcacgga cgacaggggc ccagcagagt acggtggagg tgggtccgtg ggggtgtcca
116161 gatcaataac gacaaacggc ccctcgttcc taccagacaa gctatcgtag gggggcgggg
116221 gatcaacaaa cgcgttcccc gcgctccata gacccgcgtc gggttgcgcc gcctccgaag
116281 ccatggatgc gccccaaagc cacgactccc gcgcgctagg tccttggggt aagggaaaag
116341 gccctactcc ccatccaagc cagccaagtt aacgggctac gccttcgggg atgggactgg
116401 caccccggcg gattttgttg ggctggcatg cgtcgcccaa ccgagggccg cgtccacggg
116461 acgcgccttt tataacccgg ggggtcattc ccaacgatca catgcaatct aactggctcc
116521 cctctcctcc cctctcccct ctcccctctc ccctctcccc tctcccctct cccctcttag
116581 gttgggggggt ggtccgacat agcctcggcg tacgtcggga ggcccgacaa gaggtccctt
116641 gtgatgtcgg gtggggccac aagcctggtt tccggaagaa acaggggggt tgccaagcgg
116701 cccggcccgc gctccccccc ccccggggcc gtgtccttgc tttccccccg tctccccccc
116761 cctcctcctc cttctcctcc tcctcgtttt tccaaacccc gcccacccgg cccggcccgg
116821 cccggccacc gccgcccacc cacccaccgc gggagaccca gccccggtcc cccgttcccc
116881 gggggccgtt atctccagcg ccccgtccgg cgcgccgccc cccgccgcta aaccccatcc
116941 cgcccccggg acccccacata taagccccca gccacacgca agaacagaca cgcagaacgg
117001 ctgtgtttat ttaaataaac cgatgtcgga ataaacaaac acaaacaccc gcgacggggg
117061 gacggaggga ggggggtgac ggggacgggg aacagacaca aaaaacaacc acaaaaaaac
117121 agccacccccc gacacccccc accccagtct cctcgccttt tcccacccac cccacgcccc
117181 cactgagccc ggtcgatcga cgagcacccc cgcccccgcc cctgccccgg cgacccccgg
117241 cccgcacgat cccgacaaca ataacaaccc caacggaaag cggcggggtg ttgggggagg
117301 cgaggaacaa ccgaggggaa cgggggatgg aaggacggga agtggaagtc ctgataccca
117361 tcctacaccc ccctgccttc caccctccgg ccccccgcga gtccacccgc cggccggcta
117421 ccgagaccga acacggcggc cgccgcagcc gccgcagccg ccgccgacac cgcagagccg

FIG. 8BR

117481 gcgcgcgcac acacaagcgg cagaggcaga aaggccccga gtcattgttt atgtggccgc
117541 gggccagcag acggcccgcg acaccccccc gcccgtgtgg gtatccggcc ccccgccccg
117601 cgccggtcca ttaagggcgc gcgtgcccgc gagatatcaa tccgttaagt gctctgcaga
117661 caggggcacc gcgcccggaa atccattagg ccgcagacga ggaaaataaa attacatcac
117721 ctacccacgt ggtgctgtgg cctgtttttg ctgcgtcatc tgagccttta taaaagcggg
117781 ggcgcggccg tgccgatcgc gggtggtgcg aaagactttc cgggcgcgtc cgggtgccgc
117841 ggctctccgg gcccccctgc agccgggccg gccaaggggc gtcggcgaca tcctcccccc
117901 aagcgccggc cggccgctgg tctgtttttt gttttccccg tttcgggggt ggggggggtt
117961 acggtttctg ttttttaaac ccgtctgggg tgtttttcgt tccgtcgccg ggatgtttcg
118021 ttcgttcggc ccctcacggg gcgaaggccg cgtacggccc gggacgaggg gccccccgacc
118081 gcggcggtcc gggccccgtc cgggcccgct cgccggcacg cgacgcgaaa aaggcccccc
118141 ggaggctttt ccgggttccc ggcccggggc ctgagataaa caatcggggt taccgccaac
118201 ggccggcccc cgtggcggcc cggcccgggg ccccggcgga cccaaggggc cccggcccgg
118261 ggccccacaa cggcccggcg catgcgctgt gttttttttt tcctcggtgt tctgccgggc
118321 tccgtcgcct ttcctgttct cgcttcttcc ccccccccct cttcaccccc agtaccctcc
118381 tccctccctt cctcccccgt tatcccactc gtcgagggcg ccccggtgtc gttcaacaaa
118441 gacgccgcgt ttccaggtag gttagacacc tgcttctccc caatagaggg gggggaccc
118501 aaacgacagg gggcgcccca gaggctaagg tcggccacgc cactcgcggg tgggctcgtg
118561 ttacagcaca ccagcccgtt cttttccccc cctcccaccc ttagtcagac tctgttactt
118621 acccgtccga ccaccaactg cccccttatc taagggccgg ctggaagacc gccaggggggt
118681 cggccggtgt cgctgtaacc ccccacgcca atgacccacg tactccaaga aggcatgtgt
118741 cccacccccgc ctgtgttttt gtgcctggct ctctatgctt gggtcttact gcctgggggg
118801 ggggatgcgg gggaggggggg gtgtggaagg aaatgcacgg cgcgtgtgta ccccccccccc
118861 aaagttgttc ctaaagcgag gatatggagg agtggcgggt gccgggggac gggggtgatc
118921 tctggcacgc gggggggggaa gggtcgggggg agggggggat ggggtaccgg cccacctggc
118981 cgacgcgggt acgcgtgcct ttgcacacca accccacgtc ccccggcggt ctctaagaag
119041 caccgccccc cctccttcat accaccgagc atgcctgggt gtgggttggt aaccaacacg
119101 cccatcccct cgtctcctgt gattctctgg ctgcaccgca ttcttgtttt ctaactatgt

FIG. 8BS

119161 tcctgtttct gtctcccccc cacccctccg ccccaccccc caacacccac gtctgtggtg
119221 tggccgaccc ccttttgggc gccccgtccc gccacccctc ccgtcctttg ttgccctata
119281 gtgtagttaa cccccccccc gcccttttgtg gcggccagag gccaggtcag tccgggcggg
119341 caggcgctcg cggaaactta acacccacac ccagcccact gtggttctgg ctccatgcca
119401 gtggcaggat gctttcgggg atcggtggtc aggcagcccg ggccgcggct ctgtggttaa
119461 caccagagcc tgcccaacat ggcacccccа ctcccacgca cccccactcc cacgcacccc
119521 cactcccacg cacccccact cccacgcacc cccactccca cgcacccccа ctcccacgca
119581 cccccactcc cacgcacccc cactcccacg cacccccact cccacgcacc cccactccca
119641 cgcacccccа ctcccacgca cccccactcc cacgcacccc cactcccacg cacccccaag
119701 atccatccaa cacagacagg gaaaagatac aaaagtaaac ctttatttcc caatagacag
119761 caaaaatccc ctgagttttt tattagggcc aacactaaag acccgctggt gtgtggtgcc
119821 cgtgtctttc actttccccc cccgacacg gattggctgg tgtagtgggc gcggccagag
119881 accacccagc acccgacccc cctccccaca aacacggggg gcgtcccctta ttgttttccc
119941 tcgtcccggg tcgacgcccc ctgctccccg gaccacgggt gccgagaccg caggctgcgg
120001 aagtccaggg cgcccactag ggtgccctgg tcgaacagca tgttccccac gggggtcatc
120061 cagaggctgt tccactccga cgcgggggcc gtcgggtact cgggggcat cacgtggtta
120121 cccgcggtct cggggagcag ggtgcggcgg ctccagccgg ggaccgcggc ccgcagccgg
120181 gtcgccatgt ttcccgtctg gtccaccagg accacgtacg ccccgatgtt ccccgtctcc
120241 atgtccagga tgggcaggca gtcccccgtg atagtcttgt tcacgtaagg cgacagggcg
120301 accacgctag agacccccga gatgggcagg tagcgcgtga ggccgcccgc ggggacggcc
120361 ccggaagtct ccgcgtggcg cgtcttccgg gcacacttcc tcggcccccg cggcccagaa
120421 gcagcgcggg ggccgaggga ggtttcctct tgtctccctc ccagggcacc gacggccccg
120481 cccgaggagg cggaagcgga ggaggacgcg gccccggcgg cggaagaggc ggcccccgcg
120541 ggggtcgggg ccgaggagga agaggcagag gaggaagagg cggaggccgc cgaggacgtc
120601 agggggggtcc cgggcccacc ctggccgcgc cccccggccc ctgagtcgga ggggggggtgc
120661 gtcgccgccc tcttggcccc tgccggcgcg agggggggac gcgtggactg gggggagggg
120721 ttttcctggc ccgaccgcg cctcttcctc ggacgcaccg ccgcctcctg ctcgacagag
120781 acggcggagg ggagcggggc ggcgccggag ggggtgcggc cgcgggaggg cccgtgccca

FIG. 8BT

120841 ccctccacgc ccggcccccc cgagccgcgc gccaccgtcg cacgcgcccg gcacagactc
120901 tgttcttggt tcgcggcctg agccagggac gagtgcgact ggggcacacg gcgcgcgtcc
120961 gcggggcggg cggccggctc cgccccgggg gccggggcgc gggggccggg ccccggaggc
121021 ggcgctcgca cgcacggggc cacggccgcg cgggggcgcg cgggtcccga cgcggccgag
121081 gacgcggggg gcccggggcg gggggcggag cctggcatgg gcgccgcggg gggcctgtgg
121141 ggagaggccg gggggagtc gctgatcact atggggtctc tgttgtttgc aaggggggcg
121201 ggtctgttga caaggggcc cgtccggccc ctcggccgcc ccgcctccgc ttcaacaacc
121261 ccaaccccaa ccccaacccc cccggagggg ccagacgccc cccgcggcgc cgcggctcgc
121321 gactggcggg agccgccgcc gccgctgctg ttggtggtgg tgttggtgtt actgctgccg
121381 tgtggcccga tgggcgccga gggggcgct gtccgagccg cggccggctg gggggctgcg
121441 ttagacgccc cgcccgtcac gggggcgcg gcggtgcctc tgcgtggggg ggcgcggggc
121501 gtccggcggg gggcgggcgg gacgtagtct gctgcaagag acaacggggg gcgcgatcag
121561 gttacgcccc ctccccggcc cgcccttcc tcgcccgccc gcccattcct ccctcctcct
121621 cctcccccag ggtccttgcc gcccccgcc tcaccgtcgt ccaggtcgtc gtcatcctcg
121681 tccgtggtgg gctccgggtg ggtgggcgac agggccctca ccgtgtgccc ccccagggtc
121741 aggtaccgcg gggcgaaccg ctgattgccc gtccagataa agtccacggc cgtgcccgcc
121801 ctgacggcct cctcggcctc catgcgggtc tgggggtcgt tcacgatcgg gatggtgctg
121861 aacgacccgc tgggcgtcac gcccactatc aggtacacca gcttggcgtt gcacagcggg
121921 caggtgttgc gcaattgcat ccaggttttc atgcacggga tgcagaagcg gtgcatgcac
121981 gggaaggtgt cgcagcgcag gtggggcgcg atctcatccg tgcacacggc gcacacgtcg
122041 ccctcgtcgc tccccccgtc ctctcgaggg ggggcgcccc cgcaactgcc ggggtcttcc
122101 tcgcgggggg ggctcccccc cgagaccgcc cccccatcca cgccctgcgg ccccagcagc
122161 cccgtctcga acagttccgt gtccgtgctg tccgcctcgg aggcggagtc gtcgtcatgg
122221 tggtcggcgt cccccccgccc ccccacttcg gtctccgcct cagagtcgct gctgtccggc
122281 aggtctcggt cgcagggaaa cacccagaca tccggggcgg gctaagggga aaaaaggggg
122341 gcgggtaaga atgggggggg atttcccgcg tcaatcagcg cccacgagtt cccccctctcc
122401 cccccccgcct cacaaagtcc tgcccccctg ctggcctcgg aagaggggggg agaaaggggt
122461 ctgcaaccaa aggtggtctg ggtccgtcct ttggatcccg acccctcttc ttcccctcttc

FIG. 8BU

122521 tcccgccctc cagacgcacc ggagtcgggg gtcccacggc gtcccccaaa tatggcgggc
122581 ggctcctccc cacccccta gatgcgtgtg agtaaggggg gcctgcgtat gagtcagtgg
122641 ggaccacgcc cccaacacgg cgacccggt ccctgtgtgt ttgttgtggg ggcgtgtctc
122701 tgtgtatgag tcaggggtc ccacggcgac cccgggccct gcgtctgagt caaaggggcc
122761 atgtgtatgt gttgggggtc tgtatatata aagtcagggg gtcacatggc gacccccaac
122821 agggcgaccc cggtccctgt atatataggg tcaggggggtt ccgcgccccc taacatggcg
122881 cccccggtcc ctgtatatat agtgtcacgg ggttccacgc cccctaacat ggcgccccaa
122941 catggcgccc ggctcccgtg tatgagtggg ggtcccccaa catgggggcc ggttccaggg
123001 taagggtcgg gggtccccca acatggcgcc ccccaatatg gcgccccaga catggcgccc
123061 ggcccctcac ctcgcgctgg gggcggccct caggccggcg ggtactcgct ccggggcggg
123121 gctccatggg ggtcgtatgc ggctggaggg tcgcggacgg agggtccctg ggggtcgcaa
123181 cgtaggcggg gcttctgtgg tgatgcggag agggggcggc ccgagtctgc ctggctgctg
123241 cgtctcgctc cgagtgccga ggtgcaaatg cgaccagacc gtcgggccag ggctaactta
123301 taccccacgc ctttcccctc cccaaagggg cggcagtgac gattcccccа atggccgcgc
123361 gtcccagggg aggcaggccc accgcggggc ggccccgtcc ccggggacca acccggcgcc
123421 cccaaagaat atcattagca tgcacggccc ggcccccgat ttgggggacc aaccgggtgt
123481 cccccaaaga acccccattag catgcccctc ccgccgacgc aacaggggct tggcctgcgt
123541 cggtgccccg gggcttcccg ccttcccgaa gaaactcatt accataccccg gaaccccagg
123601 ggaccaatgc gggttcattg agcgacccgc gggccaatgc gcgaggggcc gtgtgttccg
123661 ccaaaaaagc aattagcata acccggaacc ccaggggagt ggttacgcgc ggcgcgggag
123721 gcggggaata ccggggttgc ccattaaggg ccgcgggaat tgccggaagc gggaagggcg
123781 gccgggggccg cccattaatg agtttctaat taccataccg ggaagcggaa caaggcctct
123841 tgcaagtttt taattaccat accgggaagt gggcggcccg gcccattggg cggtaactcc
123901 cgcccaatgg gccgggcccc gaagactcgg cggacgctgg ttggccgggc cccgccgcgc
123961 tggcggccgc cgattggcca gtcccgcccc cgaggcgggc ccgccttggg ggcggaccgg
124021 ctcccagcgt atatatgcgc ggctcctgcc atcgtctctc cggagagcgg cttggtgcgg
124081 agctcccggg agctccgcgg aagacccagg cgcctcgggt gtaacgttag accgagttcg
124141 ccgggccggc tccgcgggcc agggcccggg cacgggcctc gggccccagg cacggcccga

FIG. 8BV

124201 tgaccgcctc ggcctccgcc acccggcgcc ggaaccgagc ccggtcggcc cgctcgcggg
124261 cccacgagcc gcggcgcgcc aggcgggcgg ccgaggccca gaccaccagg tggcgcaccc
124321 ggacgtgggg cgagaagcgc acccgcgcgg gggtcgcggg ggtcgcgggg gtcgcggggg
124381 tcgcgggggt cgcgggggtc gcgggggtcg cgggggtcgc gggggctcc ggcgcccccc
124441 ccccgcccgc gcgtcgcagg cgcaggcgcg ccaggtgctc cgcggtgacg cgcaggcgga
124501 gggcgaggcg cggcggaagg cggaaggggc gcgaggggggg gtgggagggg tcagccccgc
124561 cccccgggcc cacgccgggc ggtggggacc ggggccgggg ggcggcggcg gtgggccggg
124621 cctctggcgc cggctcgggc gggggctgt ccggccagtc gtcgtcatcg tcgtcgtcgg
124681 acgcggactc gggaacgtgg agccactggc gcagcagcag cgaacaagaa ggcgggggcc
124741 caccggcggg gggcggcggc ggggcggccg cgggcgcgct cctgaccgcg ggttccgagt
124801 tgggcgtgga ggttacctgg gactgtgcgg ttgggacggc gcccgtgggc ccgggcggcc
124861 gggggcggcg ggggccgcga tggcggcggc ggcgggccat ggagacagag agcgtgccgg
124921 ggtggtagag tttgacaggc aagcatgtgc gtgcagaggc gagtagtgct tgcctgtcta
124981 actcgctagt ctcggccgcg gggggcccgg gctgcccgcc gccgcgcttt aaagggccgc
125041 gcgcgacccc cgggggggtgt gtttcggggg ggcccgtttt tggggtctgg ccgctcctcc
125101 cccgctcctc cccgtctgtg ggtggggctc ctccccccgct cctcccccgc tcctcccccg
125161 ctcctcccccg tctgtgggtg gggctcctcc cccgctcccg cggccccgcc ccccacgccc
125221 gccgcgcgcg cgcacgccgc ccggaccgcc gcccgccttt tttgcgcgcc gccccgcgcg
125281 cgggggggccc gggctgccac aggtgtaaca acaccaacag aacaccaaca gcacggcgca
125341 ccggcgactc cggttcctca tccacacgtc acgtcatcca acacacctgc ccaacaacac
125401 aactcacagc gacaactcac cgcgcaacaa ctcctgttcc tcatccacac gtcaccgcgc
125461 acccccccgct cctccagacg tcccccagcg caacacgccg ctcctgtcac acaccacagc
125521 cccagccctc cccagcccca gccctcccca gccccagccc tccccagccc cagccctccc
125581 cagccccagc cctccccagc cccagccctc cccagcccca gccctcccca gccccagccc
125641 tccccagccc cagccctccc cagccccagc cctccccagc cccagccctc cccagcccca
125701 gccctcccca gccgcgtccc gcgctccctc gggggggttc gggcatctct acctcagtgc
125761 cgccaatctc aggtcagaga tccaaaccct ccgggggcgc ccgcgcacca ccaccgcccc
125821 tcgccccctc ccgccctcg ccccctcccg cccctcgccc cctcccgccc ctcgcccccc

FIG. 8BW

125881 cccgccccctc gccccctccc gcccctcgcc ccctcccgcc cctcgccccc tcccgccccct
125941 cgcccccctcc cgcccctcga ataaacaacg ctactgcaaa acttaatcag gtcgttgccg
126001 tttattgcgt cttcgggttt cacaagcgcc ccgccccgtc ccggcccgtt acagcacccc
126061 gtcccccctcg aacgcgccgc cgtcgtcttc gtcccaggcg ccttcccagt ccacaacgtc
126121 ccgtcgcggg ggcgtggcca agcccgcctc cgcccccagc acctccacgg ccccgccgc
126181 cgccagcacg gtgccgctgc ggcccgtggc cgaggcccag cgaatcccgg gcggcgccgg
126241 cggcagggcc cccgggccgt cgtcgtcgcc gcgcagcacc agcgggggggg cgtcgtcgtc
126301 gggctccagc agggcgcggg cgcaaaagtc cctccgcggc ccgcgccacc gggccgggcc
126361 ggcgcgcacc gcctcgcgcc ccagcgccac gtacacgggc cgcagcggcg cgcccaggcc
126421 ccagcgcgcg caggcgcggt gcgagtgggc ctcctcctcg cagaagtccg gcgcgccggg
126481 cgccatggcg tcggtggtcc ccgaggccgc cgcccggccg tccagcgccg gcagcacggc
126541 ccggcggtac tcgcgcgggg acatgggcac cggcgtgtcc gggccgaagc gcgtgcgcac
126601 gcggtagcgc acgttgccgc cgcggcacag gcgcagcggc ggcgcgtcgg ggtacaggcg
126661 cgcgtgcgcg gcctccacgc gcgcgaagac ccccgggccg aacacgcggc ccgaggccag
126721 caccgtgcgg cgcaggtcgc gcgccgccgg ccagcgcacg gcgcactgca cggcgggcag
126781 caggtcgcac gccaggtagg cgtgctgccg cgacaccgcg ggcccgtcgg cgggccagtc
126841 gcaggcgcgc acggtgttga ccacgatgag ccgccggtcg ccggcgctgg cgagcagccc
126901 cagaaactcc acggccccgg cgaaggccag gtcccgcgtg gacagcagca gcacgccctg
126961 cgcgcccagc gccgacacgt cgggggcgcc ggtccagttg cccgcccagg cggccgtgtc
127021 cggcccgcac agccggttgg ccagggccgc cagcaggcag gacagcccgc cgcgctcggc
127081 ggaccactcc ggcggccccc ccgaggcccc gccgccggcc aggtcctcgc ccggcagcgg
127141 cgagtacagc accaccacgc gcacgtcctc ggggtcgggg atctggcgca tccaggccgc
127201 catgcggcgc agcgggcccg aggcgcgcag ggggccaaag aggcggcccc cggcggcccc
127261 gtgggggtgg gggttctcgt cgtcgtcgcc gccgccgcac gcggcctggg cggcgggggc
127321 gggcccggcg caccgcgcgg cgatcgaggc cagggcccgc gggtcaaaca tgagggccgg
127381 tcgccagggg acggggaaca gcgggtggtc cgtgagctcg gccacggcgc gcggggagca
127441 gtaggcctcc agggcggcgg ccgcgggcgc cgccgtgtgg ctgggccccg ggggctgccg
127501 ccgccagccg cccagggggt cggggccctc ggcgggccgg cgcgacagcg ccacgggggcg

FIG. 8BX

127561 cgggcgggcc tgcgccgcgg cggcccgggg cgccgcgggc tgggcgggcg cgggctcggg
127621 ccccgggggc gtggaggggg gcgcgggcgc ggggaggggg gcgcgggcgt ccgagccggg
127681 ggcgtccgcg ccgctcttct tcgtcttcgg gggtcgcggg ccgccgcctc cgggcggccg
127741 ggccgggccg ggactcttgc gcttgcgccc ctcccgcggc gcggcggagg cggcggccgc
127801 gaccccgaa gacgaagaag agcggcgcgg acccgccgcc agcaggggcc gcaggctctg
127861 gttctcaaac agcaggtccg cggcggcggc ggccgcggag ctcggcaggc gcgggtcccg
127921 cggcagcgcg ggacccaggg ccccggcgac caggctcacg gcgcgcacgg cggccacggc
127981 ggcctcgctg ccgccggcca cgcgcaggtc cccgcgcagg cgcatgagca ccagcgcgtc
128041 gcgcacgaac cgcagctcgc gcagccacgc gcgcaggcgg ggcgcgtcgg cgtgcggcgg
128101 cggcggggaa gcggggcccg cgggtccctc cggccgcggg gggctggcgg gccgggcccc
128161 ggccagcccc gggacggccg ccaggtcgcc gtcgaagccc tcggccagcg cctccaggat
128221 cccgcggcag gcggccaggc actccacggc cacgcggccg gcctgggcgc ggcgcccggc
128281 gtcgtcgtcg gcgtcggcgt ggcgggcggc gtcggggtcg tcgccccccg cggggaggc
128341 gggcgcggcg gacagccgcc ccagggcggc gaggatcccc gcggcgccgt acccggcggg
128401 caccgcgcgc tcgcccggtg cggcggcggc ggcgacgacg gcggcggcga ccccctcgtc
128461 atctgcgccg gcgccggggc tccccgcggc cccgtcagc gccgcgttct cgcgcgccaa
128521 caggggcgcg taggcgcggc gcaggctggt cagcaggaag cccttctgcg cgcggtcgta
128581 tcggcggctc atggccacgg cggccgccgc gtgcgccagg ccccagccga agcggccggc
128641 cgccatggcg tagcccaggt ggggcacggc ccgcgccacg ctgccggtga tgaaggagct
128701 gctgttgcgc gcggcgcccg agatccggaa gcaggcctgg tccagcgcca cgtccccggg
128761 gaccacgcgc gggttctgga gccaccccat ggcctccgcg tccggggtgt acagcagccg
128821 cgtgatcagg gcgtactgct gcgcggcgtc gcccagctcg ggcgcccaca cggccgccgg
128881 ggcgcccgag gcctcgaacc ggcgtcgcgc ctcctccgcc tcgggcgccc cccagaggcc
128941 cgggcggctg tcgcccaggc cgccgtacag cacccgcccc gggggcgggg gcccggcgcc
129001 gggccacggc tccccgctga cgtacccgtc gcgatagcgc gcgtagaagg cgccggaggc
129061 cgcgtcggcg tccagctcga cccgccgggg ctgcccggcc gtgaagcggc ccgtggcgtc
129121 gcggccggcc accgccgcgc gggcccggcg gcgctcgatg cggcccgcgg aggccgcggg
129181 ggtcctcgcc gccgcccggg gcttgggcgc ggcctcggag agggggggtg gcccgggcgg

FIG. 8BY

```
129241 gggcggcgtc cgcccggggg cttccggcgc cgcgctcgac ggaccccgcc cgacggcccg
129301 cgcctcgcgt gcgtggtcgg ccgcgtcgtt gccgtcgtcg tcctcgtcct cgtcggacga
129361 cgaggacgaa gaggatgcgg acgacgagga cgaggacccg gagtccgacg aggtcgatga
129421 cgccgatggc cgccgccggc cgtgacgacg tctctgcggc ggctgggccg gcgggcgcgg
129481 cgacaggcgg tccgtggggt ccggatacgc gccgcgtagc ggggcctccc gtgcgcggcc
129541 ccgggccggg gcccggtcgc cggcggcgtc ggctgcgtcg tcgtactcgt ccccgtcatc
129601 gtcgtcggct cgaaaggcgg gggtccgggg cggcgaggcc gcggggtcgg gcgtcgggat
129661 cgtccggacg gcctcctcta ccatggaggc cagcagggcc agctgtcgcg gcgagacggc
129721 gtccccggcg tcctcgccgg cgtcggtgcc cgccgcgggg gccctcccgt cccgccgggc
129781 gtcgtcgagg tcgtgggggt ggtcggggtc gtggtcgggg tcgtccccgc cctcctccgt
129841 ctccgcgccc cacccgaggg cccccgctc gtcgcggtct gggctcgggg tgggcggcgg
129901 cccgtcggtg gggccgggg agccggggcg ctgcttgttc tccgacgcca tcgccgatgc
129961 ggggcgatcc tccggggata cgactgcgac ggcggacgta gcacggtagg tcacctacgg
130021 actctcgatg gggagggggc gagacccacg gaccccgacg acccccgccg tcgacgcgga
130081 actagcgcgg accggtcgat gcttgggtgg gaaaaaggac agggacggcc gatccccctc
130141 ccgcgcttcg tccgcgtatc ggcgtcccgg cgcggcgagc gtctgacggt ctgtctctgg
130201 cggtcccgcg tcgggtcgtg gatccgtgtc ggcagccgcg ctccgtgtgg acgatcgggg
130261 cgtcctcggg ctcatatagt cccaggggcc ggcgggaagg aggagcagcg gaggccgccg
130321 gccccccgcc cccacggcgg gcccgccccg aacggaattc cattatgcac gaccccgccc
130381 cgacgccggc acgccggggg cccgtggccg cggcccgttg gtcgaacccc cggccccgcc
130441 catccgcgcc atctgccatg ggcggggcgc tagggcgggt gggcccgcgc cccgccccgc
130501 atggcatctc attaccgccc gatccggcgg tttccgcttc cgttccgcat gctaacgagg
130561 aacgggcagg gggcgggggcc cgggccccga cttcccggtt cggcggtaat gagatacgag
130621 ccccgcgcgc ccgttggccg tccccgggcc cccggtcccg cccgccggac gccgggacca
130681 acgggacggc gggcggccca agggccgccc gccttgccgc ccccccattg gccggcgggc
130741 gggaccgccc caaggggcg gggccgccgg gtaaaagaag tgagaacgcg aagcgttcgc
130801 acttcgtccc aatatatata tattattagg gcgaagtgcg agcactggcg ccgtgcccga
130861 ctccgcgccg gccccggggg cgggcccggg cggcgggggg cgggtctctc cggcgcacat
```

FIG. 8BZ

130921 aaaggcccgg cgcgaccgac gcccgcagac ggcgccggcc acgaacgacg ggagctgctg
130981 cggagcacgc ggaccgggag cgggactcgc agagggccgt cggagcggac ggcgtcggca
131041 tcgcgacgcc ccggctcggg atcgggatcg catcggaaag ggacacgcgg aaagacccac
131101 ccaccccacc cacgaaacac aggggacgca ccccgggggc ctccgacgac agaaacccac
131161 cggtccgcct ttttgcacgg gtaagcacct tgggtgggcg gaggagggcg gaggaggggg
131221 gacgcggggg cggaggaggg gggacgcggg ggcggaggag gggggacgcg ggggcggagg
131281 agggggacg cgggggcgga ggaggggct cacccgcgtt cgtgccttcc cgcaggagga
131341 acgtcctcgt cgaggcgacc ggcggcgacc gttgcgtgga ccgcttcctg ctcgtcgggg
131401 cgaccggcgg cgaccgttgc gtggaccgct tcctgctcgt cggggcgacc ggcggcgacc
131461 gttgcgtgga ccgcttcctg ctcgtcgggg gggggggggg gaagccactg tggtcctccg
131521 ggacgttttc tggatggccg acatttcccc aggcgctttt gcgccttgtg taaaagcgcg
131581 gcgtcccgct ctccgatccc cgcccctggg cacgcgcaag cgcaagcgcc ctgcccgccc
131641 cctctcatcg gagtctgagg tcgaaaccga tacagccttg gagtctgagg tcgaatccga
131701 gacagcatcg gattcgaccg agtctgggga ccaggaggaa gccccccgca tcggtggccg
131761 tagggccccc cggaggcttg gggggcggtt ttttctggac atgtcggcgg aatccaccac
131821 ggggacggaa acggatacgg cggtgtcgga cgaccccgac gacacgtccg actggtctta
131881 tgacgacatt cccccacgac ccaagcgggc ccgggtaaac ctgcggctca cgagctctcc
131941 cgatcggcgg gatggggtta tttttcctaa gatggggcgg gtccggtcta cccgggaaac
132001 gcagccccgg gccccaaccc cgtcggcccc aagcccaaat gcaatgctac ggcgctcggt
132061 gcgccaggcc cagaggcgga gcagcgcacg atggacccca gacctgggct acatgcgcca
132121 gtgtatcaat cagctgtttc gggtcctgcg ggtcgcccgg gacccccacg gcagtgccaa
132181 ccgcctgcgc cacctgatac gcgactgtta cctgatggga tactgccgag cccgtctggc
132241 cccgcgcacg tggtgccgct tgctgcaggt gtccggcgga acctggggca tgcacctgcg
132301 caacaccata cgggaggtgg aggctcgatt cgacgccacc gcggaacccg tgtgcaaact
132361 tccttgtttg gaggccagac ggtacggccc ggagtgtgat cttagtaatc tcgagattca
132421 tctcagcgcg acaagcgatg atgaaatctc cgatgccacc gatctggagg ccgccggttc
132481 ggaccacacg ctcgcgtccc agtccgacac ggaggatgcc ccctcccccg ttacgctgga
132541 aaccccagaa ccccgcgggt ccctcgctgt gcgtctggag gatgagtttg gggagtttga

FIG. 8CA

132601 ctggaccccc caggagggct cccagccctg gctgtctgcg gtcgtggccg ataccagctc
132661 cgtggaacgc ccgggcccat ccgattctgg ggcgggtcgc gccgcagaag accgcaagtg
132721 tctggacggc tgccggaaaa tgcgcttctc caccgcctgc ccctatccgt gtagcgacac
132781 gtttctccgg ccgtgagtcc ggtcgccccg acccctttgt atgtcaccaa aataaaagac
132841 caaaatcaaa gcgtttgtcc cagcgtctta atggcgggaa gggcggagag aaacagacca
132901 cgcggacatg gggggtgttt gggggtttat tggcaccggg ggctaaaggg tggtaaccgg
132961 atagcagatg tgaggaagtc ggggccgttc gccgcgaacg gcgatcagag ggtcagtttc
133021 ttgcggacca cggcccggcg atgtgggttg ctcgtctggg acctcgggca tgcccataca
133081 cgcacaacac ggacgccgca ccggatggga cgtcgtaagg gggcctgggg tagctgggtg
133141 gggtttgtgc agagcaatca gggaccgcag ccagcgcata caatcgcgct cccgtccgtt
133201 tgtcccgggc agtaccacgc cgtactggta ttcgtaccgg ctgagcaggg tctccagggg
133261 gtggttgggg gccgcgggga acggggtcca cgccacggtc cactcgggca aaaaccgagt
133321 cggcacggcc cacggttctc ccacccacgc gtctggggtc ttgatggcga taaatcttac
133381 cccgagccgg atttttggg cgtattcgag aaacggcaca cacagatccg ccgcgcctac
133441 cacccacaag tggtagatgc gaggggggct gggttggtct cggtgcagca gtcggaagca
133501 cgccacggcg tccacgacct cggtgctctc caaggggctg tcctccgcaa acaggcccgt
133561 ggtggtgttt gggggcagc gacaggacct agtgcgcacg atcgggcggg tgggtttggg
133621 taagtccatc agcggctcgg ccaaccgtcg aaggttggcc ggacgaacga cgaccggggt
133681 acccaggggt tctgatgcca aaatgcggca ctgcctaagc aggaagctcc acagggccgg
133741 gcttgcgtcg acggaagtcc ggggcagggc gttgttctgg tcaaggaggg tcattacgtt
133801 gacgacaaca acgcccatgt tggtatatta caggcccgtg tccgatttgg ggcacttgca
133861 gatttgtaag gccacgcacg gcggggagac aggccgacgc gggggctgct ctaaaaattt
133921 aagggccta cggtccacag acccgccttc ccgggggggg ggcccttgga gcgaccggca
133981 gcgtaggcgt ccgggggagg ggagggtgat ttacgggggg gtaggtcagg gggtgggtcg
134041 tcaaactgcc gctccttaaa accccggggc ccgtcgttcg gggtgctcgt tggttggcac
134101 tcacggtgcg gcgaatggcc tgtcgtaagt tttgtcgcgt ttacgggggga cagggcagga
134161 ggaaggagga ggccgtcccg ccggagacaa agccgtcccg ggtgtttcct catggccccc
134221 tttataccc agccgaggac gcgtgcctgg actccccgcc cccggagacc cccaaacctt

FIG. 8CB

134281 cccacaccac accacccagc gaggccgagc gcctgtttca tctgcaggag atccttgccc
134341 agatgtacgg aaaccaggac tacccatag aggacgaccc cagcgcggat gccgcggacg
134401 atgtcgacga ggacgccccg gacgacgtgg cctatccgga ggaatacgca gaggagcttt
134461 ttctgcccgg ggacgcgacc ggtccccctta tcggggccaa cgaccacatc cctcccccgt
134521 gtggcgcatc tccccccggt atacgacgac gcagccggga tgagattggg gccacgggat
134581 ttaccgcgga agagctggac gccatggaca gggaggcggc tcgagccatc agccgcggcg
134641 gcaagccccc ctcgaccatg gccaagctgg tgactggcat gggctttacg atccacggag
134701 cgctcacccc aggatcggag gggtgtgtct ttgatagcag ccacccagat tacccccaac
134761 gggtaatcgt gaaggcgggg tggtacacga gcacgagcca cgaggcgcga ctgctgaggc
134821 gactggacca cccgcgatc ctgcccctcc tggacctgca tgtcgtctcc ggggtcacgt
134881 gtctggtcct ccccaagtac caggccgacc tgtataccta tctgagtagg cgcctgaacc
134941 cgctgggacg cccgcagatc gcagcggtct cccggcagct cctaagcgcc gttgactaca
135001 ttcaccgcca gggcattatc caccgcgaca ttaagaccga aaatattttt attaacaccc
135061 ccgaggacat ttgcctgggg gactttggtg ccgcgtgctt cgtgcagggt tcccgatcaa
135121 gccccttccc ctacggaatc gccggaacca tcgacaccaa cgcccccgag gtcctggccg
135181 gggatccgta taccacgacc gtcgacattt ggagcgccgg tctggtgatc ttcgagactg
135241 ccgtccacaa cgcgtccttg ttctcggccc cccgcggccc caaaaggggc ccgtgtgaca
135301 gtcagatcac ccgcatcatc cgacaggccc aggtccacgt tgacgagttt tccccgcatc
135361 cagaatcgcg cctcacctcg cgctaccgct cccgcgcggc cgggaacaat cgcccgcctt
135421 acacccgacc ggcctggacc cgctactaca agatggacat agacgtcgaa tatctggttt
135481 gcaaagccct caccttcgac ggcgcgcttc gccccagcgc cgcagagctg ctttgtttgc
135541 cgctgtttca acagaaatga ccgccccccgg ggggcggtgc tgtttgcggg ttggcacaaa
135601 aagaccccga cccgcgtctg tggtgttttt ggcatcatgt cgccgggcgc catgcgtgcc
135661 gttgttccca ttatcccatt ccttttggtt cttgtcggtg tatcgggggt tcccaccaac
135721 gtctcctcca ccacccaacc ccaactccag accaccggtc gtcctcgca tgaagccccc
135781 aacatgaccc agaccggcac caccgactct cccaccgcca tcagccttac cacgcccgac
135841 cacacacccc ccatgccaag tatcggactg gaggaggagg aggaagagga ggaggggcc
135901 ggggatggcg aacatcttaa ggggggagat gggacccgtg acaccctacc ccagtccccg

FIG. 8CC

135961 ggtccagccg tcccgttggc cggggatgac gagaaggaca aacccaaccg tcccgtagtc
136021 ccaccccccg gtcccaacaa ctccccgcg cgccccgaga ccagtcgacc gaagacaccc
136081 cccaccagta tcgggccgct ggcaactcga cccacgaccc aactccccctc aaaggggcga
136141 cccttggttc cgacgcctca acatacccag ctgttctcgt tcctcactgc ctcccccgcc
136201 ctggacaccc tcttcgtcgt cagcaccgtc atccacacct tatcgttttt gtgtattggt
136261 gcgatggcga cacacctgtg tggcggttgg tccagacgcg ggcgacgcac acaccctagc
136321 gtgcgttacg tgtgcctgcc gtccgaacgc gggtagggta tggggcgggg gatggggaga
136381 gcccacatgc ggaaagcaag aacaataaag gcggtggtat ctagttgata tgcatctctg
136441 ggtgtttttg gggtgtggcg gacgcggggc ggtcattgga cggggtgcag ttaaatacat
136501 gcccgggacc catgaagcat gcgcgacttc cgggcctcgg aacccacccg aaacggccaa
136561 cggacgtctg agccaggcct ggctatccgg agaaacagca cacgacttgg cgttctgtgt
136621 gtcgcgatgt ctctgcgcgc agtctggcat ctggggcttt tgggaagcct cgtgggggct
136681 gttcttgccg ccacccatcg gggacctgcg gccaacacaa cggacccctt aacgcacgcc
136741 ccagtgtccc ctcaccccag ccccctgggg ggctttgccg tccccctcgt agtcggtggg
136801 ctgtgcgccg tagtcctggg ggcggcgtgt ctgcttgagc tcctgcgtcg tacgtgccgc
136861 gggtgggggc gttaccatcc ctacatggac ccagttgtcg tataattccc cccccccctt
136921 ctccgcatgg gtgatgtcgg gtccaaactc ccgacaccac cagctggcat ggtataaatc
136981 accggtgcgc ccccaaacc atgtccggca gggggatggg ggggcgaatg cggagggcac
137041 ccaacaacac cgggctaacc aggaaatccg tggccccggc ccccaataaa gatcgcggta
137101 gcccggccgt gtgacactat cgtccatacc gaccacaccg acgaatcccc taaggggagg
137161 gggccatttt acgaggagga ggggtataac aaagtctgtc tttaaaaagc aggggttagg
137221 gagttgttcg gtcataagct tcagcgcgaa cgaccaacta ccccgatcat cagttatcct
137281 taaggtctct tttgtgtggt gcgttccggt atgggggggg ctgccgccag gttgggggcc
137341 gtgattttgt ttgtcgtcat agtgggcctc catggggtcc gcggcaaata tgccttggcg
137401 gatgcctctc tcaagatggc cgaccccaat cgctttcgcg gcaaagacct tccggtcctg
137461 gaccagctga ccgaccctcc gggggtccgg cgcgtgtacc acatccaggc gggcctaccg
137521 gacccgttcc agccccccag cctcccgatc acggtttact acgccgtgtt ggagcgcgcc
137581 tgccgcagcg tgctcctaaa cgcaccgtcg gaggcccccc agattgtccg cggggcctcc

FIG. 8CD

137641 gaagacgtcc ggaaacaacc ctacaacctg accatcgctt ggtttcggat gggaggcaac
137701 tgtgctatcc ccatcacggt catggagtac accgaatgct cctacaacaa gtctctgggg
137761 gcctgtccca tccgaacgca gccccgctgg aactactatg acagcttcag cgccgtcagc
137821 gaggataacc tggggttcct gatgcacgcc cccgcgtttg agaccgccgg cacgtacctg
137881 cggctcgtga agataaacga ctggacggag attacacagt ttatcctgga gcaccgagcc
137941 aagggctcct gtaagtacgc cctcccgctg cgcatccccc cgtcagcctg cctgtccccc
138001 caggcctacc agcaggggct gacggtggac agcatcggga tgctgccccg cttcatcccc
138061 gagaaccagc gcaccgtcgc cgtatacagc ttgaagatcg ccgggtggca cgggcccaag
138121 gccccataca cgagcaccct gctgcccccg gagctgtccg agacccccaa cgccacgcag
138181 ccagaactcg ccccggaaga ccccgaggat tcggccctct tggaggaccc cgtggggacg
138241 gtggcgccgc aaatcccacc aaactggcac ataccgtcga tccaggacgc cgcgacgcct
138301 taccatcccc cggccacccc gaacaacatg ggcctgatcg ccggcgcggt gggcggcagt
138361 ctcctggcag ccctggtcat ttgcggaatt gtgtactgga tgcgccgccg cactcaaaaa
138421 gccccaaagc gcatacgcct cccccacatc cgggaagacg accagccgtc ctcgcaccag
138481 cccttgtttt actagatacc cccccttaat gggtgcgggg gggtcaggtc tgcgggggttg
138541 ggatgggacc ttaactccat ataaagcgag tctggaaggg gggaaaggcg gacagtcgat
138601 aagtcggtag cgggggacgc gcacctgttc cgcctgtcgc acccacagct ttttttgcga
138661 accgtcccgt tccgggatgc cgtgccgccc gttgcagggc ctggtgctcg tgggcctctg
138721 ggtctgtgcc accagcctgg ttgtccgtgg ccccacggtc agtctggtat caaactcatt
138781 tgtggacgcc ggggccttgg ggcccgacgg cgtagtggag gaagacctgc ttattctcgg
138841 ggagcttcgc tttgtggggg accaggtccc ccacaccacc tactacgatg gggtcgtaga
138901 gctgtggcac taccccatgg gacacaaatg cccacgggtc gtgcatgtcg tcacggtgac
138961 cgcgtgccca cgtcgccccg ccgtggcttt cgccctgtgt cgcgcgaccg acaacactca
139021 cagccccgca tatcccaccc tggagctgaa tctggcccaa cagccgcttt tgcgggtccg
139081 gagggcgacg cgtgactatg ccggggtgta cgtgttacgc gtatgggtcg tggacgcacc
139141 aaacgccagc ctgtttgtcc tggggatggc catagccgcc gaagggactc tggcgtacaa
139201 cggctcggcc catggctcct gcgacccgaa actgcttccg tattcggccc cgcgtctggc
139261 cccggcgagc gtataccaac ccgccccctaa cccggcctcc acccccctcga ccaccacctc

FIG. 8CE

139321 cacccctcg accaccacct ccacccctc gaccaccacc tccacccct cgaccaccac
139381 ctccacccc tcgaccacca cctccaccc ctcgaccacc acctccaccc cctcgaccac
139441 catccccgct ccccaagcat cgaccacacc cttccccacg ggagacccaa aaccccaacc
139501 tcacggggtc aaccacgaac ccccatcgaa tgccacgcga gcgacccgcg actcgcgata
139561 cgcgctaacg gtgacccaga taatccagat agccatcccc gcgtccatta tagccctggt
139621 gtttctgggg agctgtattt gctttataca cagatgtcaa cgccgctacc gacgctcccg
139681 ccgcccgatt tacaaccccc agatacccac gggcatctca tgcgcggtga acgaagcggc
139741 catggcccgc ctcggagccg agctcaaatc gcatccgagc acccccccca aatcccggcg
139801 ccggtcgtca cgcacgccaa tgccctccct gacggccatc gccgaagagt cggagcccgc
139861 gggggcggct gggcttccga cgcccccgt ggaccccacg acatccaccc caacgcctcc
139921 cctgttggta taggtccacg gccactggcc gggggcacca cataaccgac cgcagtcact
139981 gagttgggaa taaaccggta ttatttacct atatccgtgt atgtccattt ctttcccccc
140041 ccccccccc cggaaaccca aagaaggaag caaagaatgg atgggaggag ttcaggaagc
140101 cggggagagg gcccgcggcg catttaaggc gttgttgtgt tgactttggc tcttctggcg
140161 ggttggtgcg gtgctgtttg ttgggctccc attttacccg aagatcggct gctatccccg
140221 ggacatggat cgcgggggcgg tggtggggtt tcttctcggt gtttgtgttg tatcgtgctt
140281 ggcgggaatg cccaaaacgt cctggagacg ggtgagtgtc ggcgaggacg tttcgttgct
140341 tccagctccg gggcctacgg ggcgcggccc gacccagaaa ctactatggg ccgtggaacc
140401 cctggatggg tgcggccccc tacacccgtc gtgggtctcg ctgatgcccc ccaagcaggt
140461 gcccgagacg gtcgtggatg cggcgtgcat gcgcgctccg gtcccgctgg cgatggcgta
140521 cgcccccccg gccccatctg cgaccgggggg tctacgaacg gacttcgtgt ggcaggagcg
140581 cgcggccgtg gttaaccgga gtctggttat tcacggggtc cgagagacgg acagcggcct
140641 gtataccctg tccgtgggcg acataaagga cccggctcgc caagtggcct cggtggtcct
140701 ggtggtgcaa ccggccccag ttccgacccc acccccgacc ccagccgatt acgacgagga
140761 tgacaatgac gagggcgagg gcgaggacga aagtctagcc ggcactcccg ccagcgggac
140821 cccccggctc ccgcctcccc ccgccccccc gaggtcttgg cccagcgccc ccgaagtctc
140881 acacgtgcgt ggggtgaccg tgcgtatgga gactccggaa gctatcctgt tttccccccgg
140941 ggaggcgttt agcacgaacg tctccatcca tgccatcgcc cacgacgacc agacctacac

FIG. 8CF

```
141001 catggacgtc gtctggttga ggttcgacgt gccgacctcg tgtgccgaga tgcgaatata
141061 cgaatcgtgt ctgtatcacc cgcagctccc agagtgtctg tccccggccg acgctccgtg
141121 cgccgcgagt acgtggacgt ctcgcctggc cgtccgcagc tacgcggggt gttccagaac
141181 aaaccccccg ccgcgctgtt cggccgaggc tcacatggag cccttcccgg ggctggcgtg
141241 gcaggcggcc tccgtcaatc tggagttccg ggacgcgtcc ccacaacact ccggcctgta
141301 tctgtgcgtg gtgtacgtca acgaccatat tcacgcatgg ggccacatta ccatcagcac
141361 cgcggcgcag taccggaacg cggtggtgga acagcccctc ccacagcgcg gcgcggattt
141421 ggccgagccc acccacccgc acgtcggggc ccctccccac gcgccccaa cccacggcgc
141481 cctgcggtta ggggcggtga tgggggccgc cctgctgctg tctgcgctgg ggttgtcggt
141541 gtgggcgtgt atgacctgtt ggcgcaggcg tgcctggcgg gcggttaaaa gcagggcctc
141601 gggtaagggg cccacgtaca ttcgcgtggc cgacagcgag ctgtacgcgg actggagctc
141661 ggacagcgag ggagaacgcg accaggtccc gtggctggcc ccccggaga gacccgactc
141721 tccctccacc aatggatccg gctttgagat cttatcacca acggctccgt ctgtataccc
141781 ccgtagcgat gggcatcaat ctcgccgcca gctcacaacc tttggatccg gaaggcccga
141841 tcgccgttac tcccaggcct ccgattcgtc cgtcttctgg taaggcgccc catcccgagg
141901 ccccacgtcg gtcgccgaac tgggcgaccg ccggcgaggt ggacgtcgga gacgagctaa
141961 tcgcgatttc cgacgaacgc ggaccccccc gacatgaccg cccgccctc gccacgtcga
142021 ccgcgccctc gccacacccg cgacccccgg gctacacggc cgttgtctcc ccgatggccc
142081 tccaggctgt cgacgccccc tccctgtttg tcgcctggct ggccgctcgg tggctccggg
142141 gggcttccgg cctgggggcc gtcttgtgtg ggattgcgtg gtatgtgacg tcaattgccc
142201 gaggcgcata aagggccggt ggtccgccta gccgcagcaa attaaaaatc gtgagtcact
142261 gcgaccgcaa cttcccaccc ggagctttct tccggcctcg atgacgtccc ggctctccga
142321 tcccaactcc tcagcgcgat ccgacatgtc cgtgccgctt tatcccacgg cctcgccagt
142381 ttcggtcgaa gcctactact cggaaagcga agacgaggcg gccaacgact tcctcgtacg
142441 catgggccgc caacagtcgg tattaaggcg tcgacgcaga cgcacccgct gcgtcggcat
142501 ggtgatcgcc tgtctcctcg tggccgttct gtcgggcgga tttggggcgc tcctgatgtg
142561 gctgctccgc taaaagaccg catcgacacg cgcgtccttc ttgtcgtctc tcttcccccc
142621 atcaccccgc aatttgcacc cagcctttaa ctacattaaa ttgggttcga ttggcaatgt
```

FIG. 8CG

142681 tgtctcccgg ttgatttttg ggtgggtggg gagtgggtgg gtggggagtg ggtgggtggg
142741 gagtgggtgg gtggggagtg ggtgggtggg gagtgggtgg gtggggagtg ggtgggtggg
142801 gagtgggtgg gtggggagtg ggtgggtggg gagtgggtgg gtggggagtg ggtgggtggg
142861 gagtgggtgg gtggggagtg ggtgggtggg gagtgggtgg gtggggagtg ggtgggtggg
142921 gagtgggtgg gtggggagtg gcaaggaaga aacaagcccg accaccagac agaaaatgta
142981 accataccca aaccgactct gggggctgtt tgtggggtcg gaaccatagg atgaacaaac
143041 caccccgtac ctcccgcacc cttgggtgcg ggtggctcat cggcatctgt ccggtatggg
143101 ttgttcccca cccactcgcg ttcggacgtc ttagaatcat ggcggtttct atgccgacat
143161 cggtttctcc cccgcaataa gacacgatgc gataaaatct gtttgtgaaa tttattaagg
143221 gtacaaattg ccctagcaca ggggtggggt tagggccggg tccccacacc caaacgcacc
143281 aaacagatgc aggcagtggg tcgagtacag ccccgcgtac gaacacgtcg atgcgtgtgt
143341 cagacagcac cagaaagcac aggccatcaa caggtcgtgc atgtgtcggt gggtttggac
143401 gcgggggccc atggtgggtg ataaagttaa tggccgccgt ccgccagggc cacaggggcg
143461 acgtctcttg gttggcccgg agccactggg tgtggaccag ccgcgcgtgg cggcccaaca
143521 tggcccctgt agccgggggc gggggatcgc gcacgtttgc agcgcacatg cgagacacct
143581 cgaccacggt tcggaagaag gcccggtggt ccgcgggcaa catcaccagg tgcgcaagcg
143641 cccgggcgtc cagagggtag agccctgagt catccgaggt tggctcatcg cccgggtcat
143701 gccgcaagtg cgtgtgggtt gggcttccgg tgggcgggac gcgaaccgcg gtgtggagcc
143761 cgacgcgggc ccgagcgtac gctccatctt gtggggagaa ggggtctggg ctcgccaggg
143821 gggcatactt gcccgggcta tacagacccg cgagccgtac gtggttcgcg gggggtgcgt
143881 ggggtccggg gctcccgggg aggccggggc tcccggggtt gtcgtggatc cctggggtca
143941 cgcggtaccc tggggtctct gggagctcgc ggtactctgg gttccctagg ttctcggggt
144001 ggtcgcggaa cccggggctc ccggggaaca cgcggtgtcc tggggattgt tggcggtcgg
144061 acggcttcag atggcttcga gatcgtagtg tccgcaccga ctcgtagtag acccgaatct
144121 ccacattgcc ctgccgcttg atcattatca ccccgttgcg ggggtccgga gatcatgcgc
144181 gggtgtcctc gaggtgcgtg aacacctctg gggtgcatgc cggcggacgg cacgcctttt
144241 aagtaaacat ctgggtcgcc cggcccaact ggggccgggg gttgggtctg gctcatctcg
144301 agagccacgg ggggaaccac cctccgccca gaaacttggg cgatggtcgt acccgggact

FIG. 8CH

144361 caacgggtta ccggattacg gggactgtcg gtcacggtcc cgccggttct tcgatgtgcc
144421 acacccaagg atgcgttggg ggcgattttg ggcagcagcc cgggagagcg cagcagagga
144481 cgctccgggt cgtgcacggc ggttctggcc gcctcccggt cctcacgccc ccttttattg
144541 atctcatcgc gtacgtcggc gtacgtcctg ggcccaaccc gcatgttgtc caggaaggtg
144601 tccgccattt ccagggccca cgacatgctc cccccccccc ccccgacgag caggaagcgg
144661 tccacgcaac ggtcgccgcc ggtcgccccg acgagcagga agcggtccac gcaacggtcg
144721 ccgccggtcg ccccgacgag caggaagcgg tccacgcaac ggtcgccgcc ggtcgcctcg
144781 acgaggacgt tcctcctgcg ggaaggcacg aacgcgggtg agccccctcc tccgcccccg
144841 cgtcccccct cctccgcccc cgcgtccccc ctcctccgcc cccgcgtccc ccctcctccg
144901 ccccccgcgtc ccccctcctc cgccctcctc cgcccaccca aggtgcttac ccgtgcaaaa
144961 aggcggaccg gtgggtttct gtcgtcggag gccccccgggg tgcgtcccct gtgtttcgtg
145021 ggtggggtgg gtgggtcttt ccgcgtgtcc ctttccgatg cgatcccgat cccgagccgg
145081 ggcgtcgcga tgccgacgcc gtccgctccg acggccctct gcgagtcccg ctcccggtcc
145141 gcgtgctccg cagcagctcc cgtcgttcgt ggccggcgcc gtctgcgggc gtcggtcgcg
145201 ccgggccttt atgtgcgccg gagagacccg ccccccgccg cccgggcccg cccccggggc
145261 cggcgcggag tcgggcacgg cgccagtgct cgcacttcgc cctaataata tatatatatt
145321 gggacgaagt gcgaacgctt cgcgttctca cttcttttac ccggcggccc cgccccccttg
145381 gggcggtccc gcccgccggc caatgggggg gcggcaaggc gggcggccct tgggccgccc
145441 gccgtcccgt tggtcccggc gtccggcggg cgggaccggg ggcccgggga cggccaacgg
145501 gcgcgcgggg ctcgtatctc attaccgccg aaccgggaag tcgggggccg ggccccgccc
145561 cctgcccgtt cctcgttagc atgcggaacg gaagcggaaa ccgccggatc gggcggtaat
145621 gagatgccat gcggggcggg gcgcgggccc acccgcccta gcgccccgcc catggcagat
145681 ggcgcggatg ggcggggccg ggggttcgac caacgggccg cggccacggg cccccggcgt
145741 gccggcgtcg gggcggggtc gtgcataatg gaattccgtt cggggcgggc ccgccgtggg
145801 ggcgggggggc cggcggcctc cgctgctcct ccttcccgcc ggcccctggg actatatgag
145861 cccgaggacg cccgatcgt ccacacggag cgcggctgcc gacacggatc cacgacccga
145921 cgcgggaccg ccagagacag acggtcagac gctcgccgcg ccgggacgcc gatacgcgga
145981 cgaagcgcgg gaggggatc ggccgtccct gtcctttttc ccacccaagc atcgaccggt

FIG. 8CI

146041 ccgcgctagt tccgcgtcga cggcgggggt cgtcggggtc cgtgggtctc gccccctccc
146101 catcgagagt ccgtaggtga cctaccgtgc tacgtccgcc gtcgcagtcg tatccccgga
146161 ggatcgcccc gcatcggcga tggcgtcgga gaacaagcag cgccccggct ccccgggccc
146221 caccgacggg ccgccgccca ccccgagccc agaccgcgac gagcgggggg ccctcgggtg
146281 gggcgcggag acggaggagg gcggggacga ccccgaccac gaccccgacc accccacga
146341 cctcgacgac gcccggcggg acgggagggc ccccgcggcg ggcaccgacg ccggcgagga
146401 cgccggggac gccgtctcgc cgcgacagct ggccctgctg gcctccatgg tagaggaggc
146461 cgtccggacg atcccgacgc ccgaccccgc ggcctcgccg ccccggaccc ccgcctttcg
146521 agccgacgac gatgacgggg acgagtacga cgacgcagcc gacgccgccg gcgaccgggc
146581 cccggcccgg ggccgcgcac gggaggcccc gctacgcggc gcgtatccgg accccacgga
146641 ccgcctgtcg ccgcgcccgc cggcccagcc gccgcagaga cgtcgtcacg gccggcggcg
146701 gccatcggcg tcatcgacct cgtcggactc cgggtcctcg tcctcgtcgt ccgcatcctc
146761 ttcgtcctcg tcgtccgacg aggacgagga cgacgacggc aacgacgcgg ccgaccacgc
146821 acgcgaggcg cgggccgtcg ggcggggtcc gtcgagcgcg gcgccggaag ccccgggcg
146881 gacgccgccc ccgcccgggc cacccccct ctccgaggcc gcgcccaagc cccgggcggc
146941 ggcgaggacc cccgcggcct ccgcgggccg catcgagcgc cgccgggccc gcgcggcggt
147001 ggccggccgc gacgccacgg gccgcttcac ggccgggcag ccccggcggg tcgagctgga
147061 cgccgacgcg gcctccggcg ccttctacgc gcgctatcgc gacgggtacg tcagcgggga
147121 gccgtggccc ggcgccgggc ccccgccccc ggggcgggtg ctgtacggcg gcctgggcga
147181 cagccgcccg ggcctctggg gggcgcccga ggcggaggag gcgcgacgcc ggttcgaggc
147241 ctcgggcgcc ccggcggccg tgtgggcgcc cgagctgggc gacgccgcgc agcagtacgc
147301 cctgatcacg cggctgctgt acaccccgga cgcggaggcc atggggtggc tccagaaccc
147361 gcgcgtggtc cccgggacg tggcgctgga ccaggcctgc ttccggatct cgggcgccgc
147421 gcgcaacagc agctccttca tcaccggcag cgtggcgcgg gccgtgcccc acctgggcta
147481 cgccatggcg gccggccgct tcggctgggg cctggcgcac gcggcggccg ccgtggccat
147541 gagccgccga tacgaccgcg cgcagaaggg cttcctgctg accagcctgc gccgcgccta
147601 cgcgcccctg ttggcgcgcg agaacgcggc gctgacgggg gccgcgggga gccccggcgc
147661 cggcgcagat gacgaggggg tcgccgccgc cgtcgtcgcc gccgccgccg caccgggcga

FIG. 8CJ

147721 gcgcgcggtg cccgccgggt acggcgccgc ggggatcctc gccgccctgg ggcggctgtc
147781 cgccgcgccc gcctccccgg cgggggggga cgaccccgac gccgcccgcc acgccgacgc
147841 cgacgacgac gccgggcgcc gcgcccaggc cggccgcgtg gccgtggagt gcctggccgc
147901 ctgccgcggg atcctggagg cgctggccga gggcttcgac ggcgacctgg cggccgtccc
147961 ggggctggcc ggggccggc ccgccagccc cccgcggccg gagggacccg cgggccccgc
148021 ttccccgccg ccgccgcacg ccgacgcgcc ccgcctgcgc gcgtggctgc gcgagctgcg
148081 gttcgtgcgc gacgcgctgg tgctcatgcg cctgcgcggg gacctgcgcg tggccggcgg
148141 cagcgaggcc gccgtggccg ccgtgcgcgc cgtgagcctg gtcgccgggg ccctgggtcc
148201 cgcgctgccg cgggacccgc gcctgccgag ctccgcggcc gccgccgccg cggacctgct
148261 gtttgagaac cagagcctgc gcccctgct ggcggcgggt ccgcgccgct cttcttcgtc
148321 ttcgggggtc gcggccgccg cctccgccgc gccgcgggag gggcgcaagc gcaagagtcc
148381 cggcccggcc cggccgcccg gaggcggcgg cccgcgaccc ccgaagacga agaagagcgg
148441 cgcggacgcc cccggctcgg acgcccgcgc ccccctcccc gcgcccgcgc cccctccac
148501 gccccgggg cccgagcccg ccccgcccca gcccgcggcg ccccgggccg ccgcggcgca
148561 ggcccgcccg cgccccgtgg cgctgtcgcg ccggcccgcc gagggccccg acccctgggg
148621 cggctggcgg cggcagcccc cggggcccag ccacacggcg gcgcccgcgg ccgccgccct
148681 ggaggcctac tgctccccgc gcgccgtggc cgagctcacg gaccacccgc tgttccccgt
148741 cccctggcga ccggccctca tgtttgaccc gcgggccctg gcctcgatcg ccgcgcggtg
148801 cgccgggccc gccccgccg cccaggccgc gtgcggcggc ggcgacgacg acgagaaccc
148861 ccaccccac ggggccgccg ggggccgcct ctttggcccc ctgcgcgcct cgggcccgct
148921 gcgccgcatg gcggcctgga tgcgccagat ccccgacccc gaggacgtgc gcgtggtggt
148981 gctgtactcg ccgctgccgg gcgaggacct ggccggcggc ggggcctcgg gggggccgcc
149041 ggagtggtcc gccgagcgcg gcgggctgtc ctgcctgctg gcggccctgg ccaaccggct
149101 gtgcgggccg gacacggccg cctgggcggg caactggacc ggcgcccccg acgtgtcggc
149161 gctgggcgcg cagggcgtgc tgctgctgtc cacgcgggac ctggccttcg ccggggccgt
149221 ggagtttctg gggctgctcg ccagcgccgg cgaccggcgg ctcatcgtgg tcaacaccgt
149281 gcgcgcctgc gactggcccg ccgacgggcc cgcggtgtcg cggcagcacg cctacctggc
149341 gtgcgacctg ctgcccgccg tgcagtgcgc cgtgcgctgg ccggcggcgc gcgacctgcg

FIG. 8CK

149401 ccgcacggtg ctggcctcgg gccgcgtgtt cggcccgggg gtcttcgcgc gcgtggaggc
149461 cgcgcacgcg cgcctgtacc ccgacgcgcc gccgctgcgc ctgtgccgcg gcggcaacgt
149521 gcgctaccgc gtgcgcacgc gcttcggccc ggacacgccg gtgcccatgt ccccgcgcga
149581 gtaccgccgg gccgtgctgc cggcgctgga cggccgggcg gcggcctcgg ggaccaccga
149641 cgccatggcg cccggcgcgc cggacttctg cgaggaggag gcccactcgc accgcgcctg
149701 cgcgcgctgg ggcctgggcg cgccgctgcg gcccgtgtac gtggcgctgg ggcgcgaggc
149761 ggtgcgcgcc ggcccggccc ggtggcgcgg gccgcggagg gacttttgcg cccgcgccct
149821 gctggagccc gacgacgacg cccccccgct ggtgctgcgc ggcgacgacg acggcccggg
149881 ggccctgccg ccggcgccgc ccgggattcg ctgggcctcg gccacgggcc gcagcggcac
149941 cgtgctggcg gcggcggggg ccgtggaggt gctgggggcg gaggcgggct tggccacgcc
150001 cccgcgacgg gacgttgtgg actgggaagg cgcctgggac gaagacgacg gcggcgcgtt
150061 cgaggggggac ggggtgctgt aacgggccgg gacggggcgg ggcgcttgtg aaacccgaag
150121 acgcaataaa cggcaacgac ctgattaagt tttgcagtag cgttgtttat tcgaggggcg
150181 ggaggggggcg aggggcggga gggggcgagg ggcgggaggg ggcgaggggc gggaggggggc
150241 gaggggcggg aggggggcgag gggcgggagg gggcgaggggg cgggaggggg cgaggggcgg
150301 gaggggggcga gggggcggtgg tggtgcgcgg gcgcccccgg agggtttgga tctctgacct
150361 gagattggcg gcactgaggt agagatgccc gaaccccccc gagggagcgc gggacgcggc
150421 tggggagggc tggggctggg gagggctggg gctggggagg gctggggctg gggagggctg
150481 gggctgggga gggctggggc tggggagggc tggggctggg gagggctggg gctggggagg
150541 gctggggctg gggagggctg gggctgggga gggctggggc tggggagggc tggggctggg
150601 gagggctggg gctgtggtgt gtgacaggag cggcgtgttg cgctgggggga cgtctggagg
150661 agcgggggggt gcgcggtgac gtgtggatga ggaacaggag ttgttgcgcg gtgagttgtc
150721 gctgtgagtt gtgttgttgg gcaggtgtgt tggatgacgt gacgtgtgga tgaggaaccg
150781 gagtcgccgg tgcgccgtgc tgttggtgtt ctgttggtgt tgttacacct gtggcagccc
150841 gggccccccg cgcgcggggc ggcgcgcaaa aaaggcgggc ggcggtccgg gcggcgtgcg
150901 cgcgcgcggc gggcgtgggg ggcggggccg cgggagcggg ggaggagccc cacccacaga
150961 cggggaggag cgggggagga gcgggggagg agcgggggag gagccccacc cacagacggg
151021 gaggagcggg ggaggagcgg ccagacccca aaaacgggcc ccccgaaac acacccccg

FIG. 8CL

151081 ggggtcgcgc gcggcccttt aaagcgcggc ggcgggcagc ccgggccccc cgcgg

FIG. 8CM

McKrae ICP4 amino acid sequence (SEQ ID NO: 2)

```
   1 masenkqrpg spgptdgppp tpspdrderg algwgaetee ggddpdhdpd hphdlddarr
  61 dgrapaagtd agedagdavs prqlallasm veeavrtipt pdpaaspprt pafraddddg
 121 deyddaadaa gdrapargra reaplrgayp dptdrlsprp paqppqrrrh grrrpsasst
 181 ssdsgsssss sasssssssd ededdd gnda adharearav grgpssaape apgrtppppg
 241 ppplseaapk praaartpaa sagrierrra raavagrdat grftagqprr veldadaasg
 301 afyaryrdgy vsgepwpgag ppppgrvlyg glgdsrpglw gapeaeearr rfeasgapaa
 361 vwapelgdaa qqyalitrll ytpdaeamgw lqnprvvpgd valdqacfri sgaarnsssf
 421 itgsvaravp hlgyamaagr fqwglahaaa avamsrrydr aqkgflltsl rrayapllar
 481 enaaltgaag spgagaddeg vaaavvaaaa apgeravpag ygaagilaal grlsaapasp
 541 aggddpdaar hadadddagr raqagrvave claacrgile alaegfdgdl aavpglagar
 601 paspprpegp agpaspppph adaprlrawl relrfvrdal vlmrlrgdlr vaggseaava
 661 avravslvag algpalprdp rlpssaaaaa adllfenqsl rpllaagprr sssssgvaaa
 721 asaapregrk rkspgparpp ggggprppkt kksgadapgs daraplpapa ppstppgpep
 781 apaqpaapra aaaqarprpv alsrrpacgp dplggwrrqp pgpshtaapa aaalcaycsp
 841 ravaeltdhp lfpvpwrpal mfdpralasi aarcagpapa aqaacgggdd denphphgaa
 901 ggrlfgplra sgplrrmaaw mrqipdpedv rvvvlysplp gedlagggas ggppewsaer
 961 gglscllaal anrlcgpdta awagnwtgap dvsalgaqgv lllstrdlaf agaveflgll
1021 asagdrrliv vntvracdwp adgpavsrqh aylacdllpa vqcavrwpaa rdlrrtvlas
1081 grvfgpgvfa rveaaharly pdapplrlcr ggnvryrvrt rfgpdtpvpm spreyrravl
1141 paldgraaas gttdamapga pdfceeeahs hracarwglg aplrpvyval greavragpa
1201 rwrgprrdfc arallepddd applvlrgdd dgpgalppap pgirwasatg rsgtvlaaag
1261 avevlgaeag latpprrdvv dwegawdedd ggafegdgvl
```

FIG. 9

HSV McKrae strain amino acid sequence of ICP22 (SEQ ID NO: 3)

```
   1 madispgafa pcvkarrpal rspplgtrkr krparplsse sevetdtale sevesetasd
  61 stesgdqeea priggrrapr rlggrffldm saesttgtet dtavsddpdd tsdwsyddip
 121 prpkrarvnl rltsspdrrd gvifpkmgrv rstretqpra ptpsapspna mlrrsvrqaq
 181 rrssarwtpd lgymrqcinq lfrvlrvard phgsanrlrn lirdcylmgy crarlaprtw
 241 crllqvsggt wqmhlrntir evearfdata epvcklpcle arrygpecdl snleihlsat
 301 sddeisdatd leaagsdhtl asqsdtedap spvtletpep rgslavrled efgefdwtpq
 361 egsqpwlsav vadtssverp gpsdsgagra aedrkcldgc rkmrfstacp ypcsdtflrp
```

FIG. 10

HSV McKrae strain amino acid sequence of ICP47 (SEQ ID NO: 4)

1 mswalemadt fldnmrvgpr tyadvrdein krgredreaa rtavhdperp llrspgllpk

61 iapnaslgva hrrtggtvtd sprnpvtr

FIG. 11

HSV McKrae strain nucleotide sequence of ICP4 (SEQ ID NO: 5)

126001 tttattgcgt cttcgggttt cacaagcgcc ccgccccgtc ccggcccgtt acagcacccc
126061 gtccccctcg aacgcgccgc cgtcgtcttc gtcccaggcg ccttcccagt ccacaacgtc
126121 ccgtcgcggg ggcgtggcca agcccgcctc cgcccccagc acctccacgg ccccgccgc
126181 cgccagcacg gtgccgctgc ggcccgtggc cgaggcccag cgaatcccgg gcggcgccgg
126241 cggcagggcc cccgggccgt cgtcgtcgcc gcgcagcacc agcgggggggg cgtcgtcgtc
126301 gggctccagc agggcgcggg cgcaaaagtc cctccgcggc ccgcgccacc gggccgggcc
126361 ggcgcgcacc gcctcgcgcc ccagcgccac gtacacgggc cgcagcggcg cgcccaggcc
126421 ccagcgcgcg caggcgcggt gcgagtgggc ctcctcctcg cagaagtccg gcgcgccggg
126481 cgccatggcg tcggtggtcc ccgaggccgc cgcccggccg tccagcgccg gcagcacggc
126541 ccggcggtac tcgcgcgggg acatgggcac cggcgtgtcc gggccgaagc gcgtgcgcac
126601 gcggtagcgc acgttgccgc cgcggcacag gcgcagcggc ggcgcgtcgg ggtacaggcg
126661 cgcgtgcgcg gcctccacgc gcgcgaagac ccccgggccg aacacgcggc ccgaggccag
126721 caccgtgcgg cgcaggtcgc gcgccgccgg ccagcgcacg gcgcactgca cggcgggcag
126781 caggtcgcac gccaggtagg cgtgctgccg cgacaccgcg ggcccgtcgg cgggccagtc
126841 gcaggcgcgc acggtgttga ccacgatgag ccgccggtcg ccggcgctgg cgagcagccc
126901 cagaaactcc acggcccccgg cgaaggccag gtcccgcgtg gacagcagca gcacgccctg
126961 cgcgcccagc gccgacacgt cgggggcgcc ggtccagttg cccgcccagg cggccgtgtc
127021 cggcccgcac agccggttgg ccagggccgc cagcaggcag gacagcccgc cgcgctcggc
127081 ggaccactcc ggcggccccc ccgaggcccc gccgccggcc aggtcctcgc ccggcagcgg
127141 cgagtacagc accaccacgc gcacgtcctc ggggtcgggg atctggcgca tccaggccgc
127201 catgcggcgc agcgggcccg aggcgcgcag ggggccaaag aggcggcccc cggcggcccc
127261 gtgggggtgg gggttctcgt cgtcgtcgcc gccgccgcac gcggcctggg cggcgggggc
127321 gggcccggcg caccgcgcgg cgatcgaggc cagggcccgc gggtcaaaca tgagggccgg
127381 tcgccagggg acggggaaca gcgggtggtc cgtgagctcg gccacggcgc gcggggagca
127441 gtaggcctcc agggcggcgg ccgcgggcgc cgccgtgtgg ctgggccccg ggggctgccg
127501 ccgccagccg cccaggggct cggggccctc ggcgggccgg cgcgacagcg ccacgggggcg
127561 cgggcgggcc tgcgccgcgg cggcccgggg cgccgcgggc tgggcgggggg cgggctcggg

FIG. 12A

127621 ccccgggggc gtggaggggg gcgcgggcgc ggggaggggg gcgcgggcgt ccgagccggg
127681 ggcgtccgcg ccgctcttct tcgtcttcgg gggtcgcggg ccgccgcctc cgggcggccg
127741 ggccgggccg ggactcttgc gcttgcgccc ctcccgcggc gcggcggagg cggcggccgc
127801 gacccccgaa gacgaagaag agcggcgcgg acccgccgcc agcaggggcc gcaggctctg
127861 gttctcaaac agcaggtccg cggcggcggc ggccgcggag ctcggcaggc gcgggtcccg
127921 cggcagcgcg ggacccaggg ccccggcgac caggctcacg gcgcgcacgg cggccacggc
127981 ggcctcgctg ccgccggcca cgcgcaggtc cccgcgcagg cgcatgagca ccagcgcgtc
128041 gcgcacgaac cgcagctcgc gcagccacgc gcgcaggcgg ggcgcgtcgg cgtgcggcgg
128101 cggcggggaa gcggggcccg cgggtccctc cggccgcggg gggctggcgg gccgggcccc
128161 ggccagcccc gggacggccg ccaggtcgcc gtcgaagccc tcggccagcg cctccaggat
128221 cccgcggcag gcggccaggc actccacggc cacgcggccg gcctgggcgc ggcgcccggc
128281 gtcgtcgtcg gcgtcggcgt ggcgggcggc gtcggggtcg tcgccccccg cggggaggc
128341 gggcgcggcg gacagccgcc ccagggcggc gaggatcccc gcggcgccgt acccggcggg
128401 caccgcgcgc tcgcccggtg cggcggcggc ggcgacgacg gcggcggcga ccccctcgtc
128461 atctgcgccg gcgccggggc tccccgcggc ccccgtcagc gccgcgttct cgcgcgccaa
128521 cagggcgcgg taggcgcggc gcaggctggt cagcaggaag cccttctgcg cgcggtcgta
128581 tcggcggctc atggccacgg cggccgccgc gtgcgccagg ccccagccga agcggccggc
128641 cgccatggcg tagcccaggt ggggcacggc ccgcgccacg ctgccggtga tgaaggagct
128701 gctgttgcgc gcggcgcccg agatccggaa gcaggcctgg tccagcgcca cgtccccggg
128761 gaccacgcgc gggttctgga gccaccccat ggcctccgcg tccggggtgt acagcagccg
128821 cgtgatcagg gcgtactgct gcgcggcgtc gcccagctcg ggcgcccaca cggccgccgg
128881 ggcgcccgag gcctcgaacc ggcgtcgcgc ctcctccgcc tcgggcgccc cccagaggcc
128941 cgggcggctg tcgcccaggc cgccgtacag cacccgcccc gggggcgggg gcccggcgcc
129001 gggccacggc tccccgctga cgtacccgtc gcgatagcgc gcgtagaagg cgccggaggc
129061 cgcgtcggcg tccagctcga cccgccgggg ctgcccggcc gtgaagcggc ccgtggcgtc
129121 gcggccggcc accgccgcgc gggcccggcg gcgctcgatg cggcccgcgg aggccgcggg
129181 ggtcctcgcc gccgcccggg gctggggcgc ggcctcggag agggggggtg gcccgggcgg
129241 gggcggcgtc cgcccggggg cttccggcgc cgcgctcgac ggacccccgcc cgacggcccg

FIG. 12B

```
129301 cgcctcgcgt gcgtggtcgg ccgcgtcgtt gccgtcgtcg tcctcgtcct cgtcggacga
129361 cgaggacgaa gaggatgcgg acgacgagga cgaggacccg gagtccgacg aggtcgatga
129421 cgccgatggc cgccgccggc cgtgacgacg tctctgcggc ggctgggccg gcgggcgcgg
129481 cgacaggcgg tccgtggggt ccggatacgc gccgcgtagc ggggcctccc gtgcgcggcc
129541 ccgggccggg gcccggtcgc cggcggcgtc ggctgcgtcg tcgtactcgt ccccgtcatc
129601 gtcgtcggct cgaaaggcgg gggtccgggg cggcgaggcc gcggggtcgg gcgtcgggat
129661 cgtccggacg gcctcctcta ccatggaggc cagcagggcc agctgtcgcg gcgagacggc
129721 gtccccggcg tcctcgccgg cgtcggtgcc cgccgcgggg gccctcccgt cccgccgggc
129781 gtcgtcgagg tcgtgggggt ggtcggggtc gtggtcgggg tcgtccccgc cctcctccgt
129841 ctccgcgccc cacccgaggg cccccgctc gtcgcggtct gggctcgggg tgggcggcgg
129901 cccgtcggtg gggcccgggg agccggggcg ctgcttgttc tccgacgcca tcgccgatgc
129961 ggggcgatcc tccgggata cgactgcgac ggcggacgta gcacggtagg tcacctacgg_
```

FIG. 12C

HSV McKrae strain nucleotide sequence of ICP22 (SEQ ID NO: 6)

132481 ggtcctccgg gacgttttct ggatggccga catttcccca ggcgcttttg tgccttgtgt
132541 aaaagcgcgg cgtcccgctc tccgatcccc gccccctgggc acgcgcaagc gcaagcgccc
132601 tgcccgcccc ctctcatcgg agtctgaggt cgaatccgag acagccttgg agtctgaggt
132661 cgaatccgag acagcatcgg attcgaccga gtctggggac caggaggaag ccccccgcat
132721 cggtggccgt agggcccccc ggaggcttgg ggggcggttt tttctggaca tgtcggcgga
132781 atccaccacg gggacggaaa cggatgcgtc ggtgtcggac gaccccgacg acacgtccga
132841 ctggtcttgt gacgacattc ccccacgacc caagcgggcc cgggtaaacc tgcggctcac
132901 tagctctccc gatcggcggg atggggttat ttttcctaag atggggcggg tccggtctac
132961 ccgggaaacg cagccccggg cccccacccc gtcggccccа agcccaaatg caatgctccg
133021 gcgctcggtg cgccaggccc agaggcggag cagcgcacga tggacccccg acctgggcta
133081 catgcgccag tgtatcaatc agctgtttcg ggtcctgcgg gtcgcccggg acccccacgg
133141 cagtgccaac cgcctgcgcc acctgatacg cgactgttac ctgatgggat actgccgagc
133201 ccgtctggcc ccgcgcacgt ggtgccgctt gctgcaggtg tccggcggaa cctggggcat
133261 gcacctgcgc aacaccatac gggaggtgga ggctcgattc gacgccaccg cagaacccgt
133321 gtgcaagctt ccttgtttgg aggccagacg gtacggcccg gagtgtgatc ttagtaatct
133381 cgagattcat ctcagcgcga caagcgatga tgaaatctcc gatgccaccg atctggaggc
133441 cgccggttcg gaccacacgc tcgcgtccca gtccgacacg gaggatgccc cctccccgt
133501 tacgctggaa accccagaac cccgcgggtc cctcgctgtg cgtctggagg atgagtttgg
133561 ggaggttgac tggacccccc aggagggctc ccagccctgg ctgtctgcgg tcgtggccga
133621 taccagctcc gtggaacgcc cgggcccatc cgattctggg gcgggtcgcg cagcagaaga
133681 ccgcaagtgt ctggacggct gccggaaaat gcgcttctcc accgcctgcc cctatccgtg
133741 cagcgacacg tttctccggc cgtgagtccg gtcgccccga ccccttgta tgtccccaaa

FIG. 13

HSV McKrae strain nucleotide sequence of ICP47 (SEQ ID NO: 7)

145081 tccgcccaga gactcgggtg atggtcgtac ccgggactca acgggttacc ggattacggg
145141 gactgtcggt cacggtcccg ccggttcttc gatgtgccac acccaaggat gcgttggggg
145201 cgatttcggg cagcagcccg ggagagcgca gcaggggacg ctccgggtcg tgcacggcgg
145261 ttctggccgc ctcccggtcc tcacgccccc ttttattgat ctcatcgcgt acgtcggcgt
145321 acgtcctggg cccaacccgc atgttgtcca ggaaggtgtc cgccatttcc agggcccacg
145381 acatgctttt cccccccgacg agcaggaagc ggtccacgca acggtcgccg ccggtcgcct

FIG. 14

Human cytomegalovirus enhancer nucleotide sequence (SEQ ID NO: 8)

gaagatctttggttatatagcataaatcaatattggctattggccattgcatacgttgtatccatatcataatatgt
acatttatattggctcatgtccaacattaccgccatgttgacattgattattgactagttattaatagtaatcaatt
acggggtcattagttcatagcccatatatggagttccgcgttacataacttacggtaaatggcccgcctggctgacc
gcccaacgacccccgcccattgacgtcaataatgacgtatgttcccatagtaacgccaatagggactttccattgac
gtcaatgggtggagtatttacggtaaactgcccacttggcagtacatcaagtgtatcatatgccaagtacgccccct
attgacgtcaatgacggtaaatggcccgcctggcattatgcccagtacatgaccttatgggactttcctacttggca
gtacatctacgtattagtcatcgctattaccatggtgatgcggttttggcagtacatcaat

FIG. 15

Calcitonin gene-related peptide promoter (SEQ ID NO: 9)

aatgggtttg ggtgtgtgta aatgagtgtg
accggaagcg agtgtgagct tgatctaggc agggaccaca cagcactgtc acacctgcct
gctctttagt agaggactga agtgcggggg tgggggtacg gggccggaat agaatgtctc
tgggacatct tggcaaacag cagccggaag caaaggggca gctgtgcaaa cggctcaggc
aggtgatgga tggcagggta ggaggggga ggtccagagg tctggatgga ggcttccgca
tctgtacctt gcaactcacc cctcaggccc agcaggtcat cggccccctc ctcacacatg
taatgacgta gaagagtacc ccgggacagt ccgggggagat ggagattcgg aaagtatcca
tggagctctt acagaatccc ctgtgcggac caggaaactc ttgtagatcc ctgcctatct
gaggcccagg cgctgggctg tttctcacaa tattccttca agatgagatt gtggtcccca
tttcaaagat gagtacactg agcctctgtg aagttacttg cccatgatca cacaaccagg
aattgggcca actgtaattg aactcctgtc taacaaagtt cttgctccca gctccgtctc
ttgtttccca cgagccctgg ccctctgtgg gtaataccag ctactggagt cagatttctt
gggcccagaa cccacccttа ggggcattaa cctttaaaat ctcacttggg caggggtctg
ggatcagagt tggaagagtc cctacaatcc tggacccttt ccgccaaatc gtgaaaccag
gggtggagtg gggcgagggt tcaaaaccag gccggactga gaggtgaaat tcaccatgac
gtcaaactgc cctcaaattc ccgctcactt taagggcgtt acttgttggt gccccacca
tcccccacca tttccatcaa tgacctcaat gcaaatacaa gtgggacggt cctgctgacg
cctccaggtt ctggaagcat gagggtgacg caacccaggg gcaaaggacc cctccgccca
ttggttgctg tgcactggcg gaactttccc gacccacagc ggcgggaata agagcagtcg
ctggcgctgg gaggcatcag agacactgcc cagcccaagt gtcgccgccg cttccacagg
gctctggctg gacgccgccg ccgccgctgc

FIG. 16

Bovine growth hormone polyadenylation signal nucleotide sequence (SEQ ID NO: 10)

ggatcccgactgtgccttctagttgccagccatctgttgtttgcccctccccgtgccttccttgaccctggaaggt
gccactcccactgtcctttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctggg
gggtggggtggggcaggacagcaaggggaggattgggaagacaatagcaggcatgctggggaagatcttc

FIG. 17

STPSTTT sequence (SEQ ID NO: 11)

Ser Thr Pro Ser Thr Thr Thr
1 5

FIG. 18

Latency Associated Transcript intron variation 1 (SEQ ID NO: 12)

gcacccccac tcccac

FIG. 19

Latency Associated Transcript variation 2 (SEQ ID NO: 13)

ccccagcccc ccccag

FIG. 20

Latency Associated Transcript variation 3 (SEQ ID NO: 14)

ccccteqccc cctcccg

FIG. 21

Alanine rich region of KOS ICP4 amino acid sequence (SEQ ID NO: 15)

Ala Ala Ser Ala Pro Asp Ala Ala Asp Ala Leu Ala Ala Ala
1 5 10

FIG. 22

Serine rich region of McKrae ICP4 amino acid sequence (SEQ ID NO: 16)

Gly Pro Arg Arg Ser Ser Ser Ser Ser Gly Val Ala Ala
1 5 10

FIG. 23

McKrae strain amino acid sequence of ICP0 (SEQ ID NO: 17)

MEPRPGASTRRPEGRPQREPAPDVWVFPCDRDLPDSSDSEAETE

VGGRGDADHHDDDSASEADSTDTELFETGLLGPQGVDGGAVSGGSPPREEDPGSCGGA

PPREDGGSDEGDVCAVCTDEIAPHLRCDTFPCMHRFCIPCMKTWMQLRNTCPLCNAKL

VYLIVGVTPSGSFSTIPIVNDPQTRMEAEEAVRAGTAVDFIWTGNQRFAPRYLTLGGH

TVRALSPTHPEPTTDEDDDDLDDADYVPPAPRRTPRAPPRRGTAAPPVTGGASNAAPQ

PAAARTAPPSAPIGPHGSSNTNTTTNSSGGGGSRQSRAAAPRGASGPSGGVGVGVGVV

EAEAGRPRGRTGPLVNRPAPLANNRDPIVISDSPPASPHRPPAAPMPGSAPRPGPPAS

SAASGPARPRAAVAPCVRAPPPGPGPRAPAPGAEPAARPADARRVPQSHSSLAQAANQ

EQSLCRARATVARGSGGPGVEGGHGPSRGRTPSGAAPLPSAVSVEQEAAVRPRKRRGS

GQENPSPQSTRPPLAPAGAKRAATHPPSDSGPGGRGQGGPGTPLTSSAASASSSSASS

SSAPTPAGAASSAAGAASSSASASSGGAVGALGGRQEETSLGPRAASGPRGPRKCARK

TRHAETSGAVPAGGLTRYLPISGVSSVVALSPYVNKTITGDCLPILDMETGNIGAYVV

LVDQTGNMATRLRAAVPGWSRRTLLPETAGNHVMPPEYPTAPASEWNSLWMTPVGNML

FDQGTLVGALDFRSLRSRHPWSGEQGASTRDEGKQ

FIG. 24

McKrae strain DNA sequence of ICP0 (SEQ ID NO: 18)

genome sequence of ICP0
join (2171..2227,2977..3643,3767..5370) of SEQ ID NO:18 to
create expected cDNA

```
atggagcccc gccccggagc gagtacccgc cggcctgagg gccgccccca
     2221  gcgcgaggtg aggggccggg cgccatgtct ggggcgccat attgggggc
gccatgttgg
     2281  gggacccccg acccttaccc tggaaccggc ccccatgttg ggggacccc
actcatacac
     2341  gggagccggg cgccatgttg gggcgccatg ttaggggcg tggaaccccg
tgacactata
     2401  tatacaggga ccgggggcgc catgttaggg ggcgcggaac cccctgaccc
tatatataca
     2461  gggaccgggg tcgccctgtt gggggtcgcc atgtgacccc ctgactttat
atatacagac
     2521  ccccaacaca tacacatggc ccctttgact cagacgcagg gcccggggtc
gccgtgggac
     2581  ccctgactc atacacagag acacgccccc acaacaaaca cacagggacc
ggggtcgccg
     2641  tgttgggggc gtggtcccca ctgactcata cgcaggcccc ccttactcac
acgcatctag
     2701  gggggtgggg aggagccgcc cgccatattt gggggacgcc gtgggacccc
cgactccggt
     2761  gcgtctggag ggcgggagaa gagggaagaa gaggggtcgg gatccaaagg
acggacccag
     2821  accacctttg gttgcagacc cctttctccc ccctcttccg aggccagcag
gggggcagga
     2881  ctttgtgagg cgggggggga gaggggaac tcgtgggcgc tgattgacgc
gggaaatccc
     2941  ccccccattct tacccgcccc cctttttttcc ccttagcccg ccccggatgt
ctgggtgttt
     3001  ccctgcgacc gagacctgcc ggacagcagc gactctgagg cggagaccga
agtgggggggg
     3061  cgggggggacg ccgaccacca tgacgacgac tccgcctccg aggcggacag
cacggacacg
     3121  gaactgttcg agacggggct gctggggccg cagggcgtgg atgggggggc
ggtctcgggg
     3181  gggagccccc cccgcgagga agaccccggc agttgcgggg gcgccccccc
tcgagaggac
     3241  gggggggagcg acgagggcga cgtgtgcgcc gtgtgcacgg atgagatcgc
gccccacctg
     3301  cgctgcgaca ccttcccgtg catgcaccgc ttctgcatcc cgtgcatgaa
aacctggatg
     3361  caattgcgca acacctgccc gctgtgcaac gccaagctgg
```

FIG. 25A

tgtacctgat agtgggcgtg

3421 acgcccagcg ggtcgttcag caccatcccg atcgtgaacg acccccagac ccgcatggag

3481 gccgaggagg ccgtcagggc gggcacggcc gtggacttta tctggacggg caatcagcgg

3541 ttcgccccgc ggtacctgac cctgggggggg cacacggtga gggccctgtc gcccacccac

3601 ccggagccca ccacggacga ggatgacgac gacctggacg acggtgaggc ggggggcggc

3661 aaggaccctg ggggaggagg aggagggagg aatgggcggg cgggcgagga aagggcgggc

3721 cggggagggg gcgtaccctg atcgcgcccc ccgttgtctc ttgcagcaga ctacgtcccg

3781 cccgcccccc gccggacgcc ccgcgccccc ccacgcagag gcaccgccgc gccccccgtg

3841 acgggcgggg cgtctaacgc agcccccccag ccggccgcgg ctcggacagc gccccctcg

3901 gcgcccatcg ggccacacgg cagcagtaac accaacacca ccaccaacag cagcggcggc

3961 ggcggctccc gccagtcgcg agccgcggcg ccgcgggggg cgtctggccc ctccgggggg

4021 gttggggttg gggttggggt tgttgaagcg gaggcggggc ggccgagggg ccggacgggc

4081 cccttgtca acagacccgc ccccttgca aacaacagag accccatagt gatcagcgac

4141 tccccccggg cctctcccca caggcccccc gcggcgccca tgccaggctc cgccccccgc

4201 cccgggcccc ccgcgtcctc ggccgcgtcg ggacccgcgc gccccgcgc ggccgtggcc

4261 ccgtgcgtgc gagcgccgcc tccggggccc ggcccccgcg ccccggcccc cggggcggag

4321 ccggccgccc gccccgcgga cgcgcgccgt gtgccccagt cgcactcgtc cctggctcag

4381 gccgcgaacc aagaacagag tctgtgccgg gcgcgtgcga cggtggcgcg cggctcgggg

4441 gggccgggcg tggagggtgg gcacgggccc tcccgcggcc gcacccccctc cggcgccgcc

4501 ccgctccccct ccgccgtctc tgtcgagcag gaggcggcgg tgcgtccgag gaagaggcgc

4561 gggtcgggcc aggaaaaccc ctccccccag tccacgcgtc ccccctcgc gccggcaggg

4621 gccaagaggg cggcgacgca ccccccctcc gactcagggc cgggggggcg cggccagggt

4681 gggcccggga ccccctgac gtcctcggcg gcctccgcct cttcctcctc tgcctcttcc

4741 tcctcggccc cgacccccgc gggggccgcc tcttccgccg ccggggccgc gtcctcctcc

FIG. 25B

gcttccgcct cctcgggcgg ggccgtcggt gccctgggag ggagacaaga ggaaacctcc

4861 ctcggccccc gcgctgcttc tgggccgcgg gggccgagga agtgtgcccg gaagacgcgc

4921 cacgcggaga cttccggggc cgtccccgcg ggcggcctca cgcgctacct gcccatctcg

4981 ggggtctcta gcgtggtcgc cctgtcgcct tacgtgaaca agactatcac ggggggactgc

5041 ctgcccatcc tggacatgga gacggggaac atcggggcgt acgtggtcct ggtggaccag

5101 acgggaaaca tggcgacccg gctgcgggcc gcggtccccg gctggagccg ccgcaccctg

5161 ctccccgaga ccgcgggtaa ccacgtgatg ccccccgagt acccgacggc ccccgcgtcg

5221 gagtggaaca gcctctggat gaccccgtg gggaacatgc tgttcgacca gggcacccta

5281 gtgggcgccc tggacttccg cagcctgcgg tctcggcacc cgtggtccgg ggagcagggg

5341 gcgtcgaccc gggacgaggg aaaacaataa

FIG. 25C

McKrae strain Latency Associated Transcript nucleotide sequence
(SEQ ID NO: 19)

atcgc gggtggtgcg aaagactttc cgggcgcgtc cgggtgccgc
117841 ggctctccgg gccccctgc agccggggcg gccaaggggc gtcggcgaca
tcctcccccc
117901 aagcgccggc cggccgctgg tctgtttttt gttttccccg tttcgggggt
ggggggggtt
117961 acggtttctg ttttttaaac ccgtctgggg tgtttttcgt tccgtcgccg
ggatgtttcg
118021 ttcgttcggc ccctcacggg gcgaaggccg cgtacggccc gggacgaggg
gccccgacc
118081 gcggcggtcc gggccccgtc cgggcccgct cgccggcacg cgacgcgaaa
aaggcccccc
118141 ggaggctttt ccgggttccc ggcccggggc ctgagataaa caatcggggt
taccgccaac
118201 ggccggcccc cgtggcggcc cggcccgggg ccccggcgga cccaaggggc
cccggccccgg
118261 ggccccacaa cggcccggcg catgcgctgt gttttttttt tcctcggtgt
tctgccgggc
118321 tccgtcgcct ttcctgttct cgcttcttcc cccccccctt cttcaccccc
agtaccctcc
118381 tccctccctt cctcccccgt tatcccactc gtcgagggcg ccccggtgtc
gttcaacaaa
118441 gacgccgcgt ttccaggtag gttagacacc tgcttctccc caatagaggg
ggggggaccc
118501 aaacgacagg gggcgcccca gaggctaagg tcggccacgc cactcgcggg
tgggctcgtg
118561 ttacagcaca ccagcccgtt cttttccccc cctcccaccc ttagtcagac
tctgttactt
118621 acccgtccga ccaccaactg cccccttatc taagggccgg ctggaagacc
gccaggggggt
118681 cggccggtgt cgctgtaacc ccccacgcca atgacccacg tactccaaga
aggcatgtgt
118741 cccaccccgc ctgtgttttt gtgcctggct ctctatgctt gggtcttact
gcctgggggg
118801 ggggatgcgg gggaggggggg gtgtggaagg aaatgcacgg cgcgtgtgta
cccccccccc
118861 aaagttgttc ctaaagcgag gatatggagg agtggcgggt gccgggggac
ggggggtgatc
118921 tctggcacgc gggggggggaa gggtcggggg agggggggat ggggtaccgg
cccacctggc
118981 cgacgcgggt gcgcgtgcct

FIG. 26A

ttgcacacca accccacgtc ccccggcggt ctctaagaag

119041 caccgccccc cctccttcat accaccgagc atgcctgggt gtgggttggt aaccaacacg

119101 cccatcccct cgtctcctgt gattctctgg ctgcaccgca ttcttgtttt ctaactatgt

119161 tcctgtttct gtctcccccc caccccctccg ccccaccccc caacacccac gtctgtggtg

119221 tggccgaccc cctttgggc gccccgtccc gccacccctc ccgtcctttg ttgccctata

119281 gtgtagttaa cccccccccc gcccttgtg gcggccagag gccaggtcag tccgggcggg

119341 caggcgctcg cggaaactta acacccacac ccagcccact gtggttctgg ctccatgcca

119401 gtggcaggat gctttcgggg atcggtggtc aggcagcccg ggccgcggct ctgtggttaa

119461 caccagagcc tgcccaacat ggcaccccca ctcccacgca cccccactcc cacgcacccc

119521 cactcccacg cacccccact cccacgcacc cccactccca cgcaccccca ctcccacgca

119581 ccccactcc cacgcacccc cactcccacg cacccccact cccacgcacc cccactccca

119641 cgcaccccca ctcccacgca cccccactcc cacgcacccc cactcccacg cacccccaag

119701 atccatccaa cacagacagg gaaaagatac aaaagtaaac ctttatttcc caatagacag

119761 caaaaatccc ctgagttttt tattagggcc aacactaaag acccgctggt gtgtggtgcc

119821 cgtgtctttc actttccccc ccccgacacg gattggctgg tgtagtgggc gcggccagag

119881 accacccagc acccgacccc cctccccaca aacacggggg gcgtcccctta ttgttttccc

119941 tcgtcccggg tcgacgcccc ctgctccccg gaccacgggt gccgagaccg caggctgcgg

120001 aagtccaggg cgcccactag ggtgccctgg tcgaacagca tgttccccac ggggggtcatc

120061 cagaggctgt tccactccga cgcgggggcc gtcgggtact cgggggggcat cacgtggtta

120121 cccgcggtct cggggagcag ggtgcggcgg ctccagccgg ggaccgcggc ccgcagccgg

120181 gtcgccatgt ttcccgtctg

FIG. 26B

```
 gtccaccagg accacgtacg ccccgatgtt ccccgtctcc
   120241 atgtccagga tgggcaggca gtcccccgtg atagtcttgt tcacgtaagg
cgacagggcg
   120301 accacgctag agacccccga gatgggcagg tagcgcgtga ggccgcccgc
ggggacggcc
   120361 ccggaagtct ccgcgtggcg cgtcttccgg gcacacttcc tcggcccccg
cggcccagaa
   120421 gcagcgcggg ggccgaggga ggtttcctct tgtctccctc ccagggcacc
gacggccccg
   120481 cccgaggagg cggaagcgga ggaggacgcg gccccggcgg cggaagaggc
ggcccccgcg
   120541 ggggtcgggg ccgaggagga agaggcagag gaggaagagg cggaggccgc
cgaggacgtc
   120601 agggggtcc cgggcccacc ctggccgcgc ccccccggcc ctgagtcgga
ggggggtgc
   120661 gtcgccgccc tcttggcccc tgccggcgcg agggggggac gcgtggactg
ggggagggg
   120721 ttttcctggc ccgacccgcg cctcttcctc ggacgcaccg ccgcctcctg
ctcgacagag
   120781 acggcggagg ggagcggggc ggcgccggag ggggtgcggc cgcgggaggg
cccgtgccca
   120841 ccctccacgc ccggcccccc cgagccgcgc gccaccgtcg cacgcgcccg
gcacagactc
   120901 tgttcttggt tcgcggcctg agccagggac gagtgcgact ggggcacacg
gcgcgcgtcc
   120961 gcggggcggg cggccggctc cgccccgggg gccggggcgc ggggccgggg
ccccggaggc
   121021 ggcgctcgca cgcacggggc cacggccgcg cggggcgcgcg cgggtcccga
cgcggccgag
   121081 gacgcggggg gcccggggcg gggggcggag cctggcatgg gcgccgcggg
gggcctgtgg
   121141 ggagaggccg gggggggagtc gctgatcact atggggtctc tgttgtttgc
aaggggggcg
   121201 ggtctgttga caaggggggcc cgtccggccc ctcggccgcc ccgcctccgc
ttcaacaacc
   121261 ccaaccccaa ccccaacccc cccggagggg ccagacgccc cccgcggcgc
cgcggctcgc
   121321 gactggcggg agccgccgcc gccgctgctg ttggtggtgg tgttggtgtt
actgctgccg
   121381 tgtggcccga tgggcgccga
```

FIG. 26C

```
ggggggcgct gtccgagccg cggccggctg gggggctgcg
   121441 ttagacgccc cgcccgtcac gggggggcgcg gcggtgcctc tgcgtggggg
ggcgcggggc
   121501 gtccggcggg gggcgggcgg gacgtagtct gctgcaagag acaacggggg
gcgcgatcag
   121561 gttacgcccc ctccccggcc cgccctttcc tcgcccgccc gcccattcct
ccctcctcct
   121621 cctcccccag ggtccttgcc gccccccgcc tcaccgtcgt ccaggtcgtc
gtcatcctcg
   121681 tccgtggtgg gctccgggtg ggtgggcgac agggccctca ccgtgtgccc
ccccagggtc
   121741 aggtaccgcg gggcgaaccg ctgattgccc gtccagataa agtccacggc
cgtgcccgcc
   121801 ctgacggcct cctcggcctc catgcgggtc tggggtcgt tcacgatcgg
gatggtgctg
   121861 aacgacccgc tgggcgtcac gcccactatc aggtacacca gcttggcgtt
gcacagcggg
   121921 caggtgttgc gcaattgcat ccaggttttc atgcacggga tgcagaagcg
gtgcatgcac
   121981 gggaaggtgt cgcagcgcag gtggggcgcg atctcatccg tgcacacggc
gcacacgtcg
   122041 ccctcgtcgc tccccccgtc ctctcgaggg ggggcgcccc cgcaactgcc
ggggtcttcc
   122101 tcgcgggggg ggctcccccc cgagaccgcc cccccatcca cgccctgcgg
ccccagcagc
   122161 cccgtctcga acagttccgt gtccgtgctg tccgcctcgg aggcggagtc
gtcgtcatgg
   122221 tggtcggcgt cccccccgccc ccccacttcg gtctccgcct cagagtcgct
gctgtccggc
   122281 aggtctcggt cgcagggaaa cacccagaca tccggggcgg gctaagggga
aaaaaggggg
   122341 gcgggtaaga atgggggggg atttcccgcg tcaatcagcg cccacgagtt
ccccctctcc
   122401 cccccgcct cacaaagtcc tgcccccctg ctggcctcgg aagaggggggg
agaaaggggt
   122461 ctgcaaccaa aggtggtctg ggtccgtcct ttggatcccg acccctcttc
ttccctcttc
   122521 tcccgccctc cagacgcacc ggagtcgggg gtcccacggc gtcccccaaa
tatggcgggc
   122581 ggctcctccc cacccccccta
```

FIG. 26D

gatgcgtgtg agtaaggggg gcctgcgtat gagtcagtgg
122641 ggaccacgcc cccaacacgg cgaccccggt ccctgtgtgt ttgttgtggg ggcgtgtctc
122701 tgtgtatgag tcaggggggtc ccacggcgac cccgggcccct gcgtctgagt caaaggggcc
122761 atgtgtatgt gttgggggtc tgtatatata aagtcagggg gtcacatggc gacccccaac
122821 agggcgaccc cggtccctgt atatataggg tcaggggggtt ccgcgccccc taacatggcg
122881 ccccggtcc ctgtatatat agtgtcacgg ggttccacgc cccctaacat ggcgccccaa
122941 catggcgccc ggctcccgtg tatgagtggg ggtcccccaa catgggggcc ggttccaggg
123001 taagggtcgg gggtcccccca acatggcgcc ccccaatatg gcgccccaga catggcgccc
123061 ggcccctcac ctcgcgctgg gggcggccct caggccggcg ggtactcgct ccggggcggg
123121 gctccatggg ggtcgtatgc ggctggaggg tcgcggacgg agggtccctg ggggtcgcaa
123181 cgtaggcggg gcttctgtgg tgatgcggag agggggcggc ccgagtctgc ctggctgctg
123241 cgtctcgctc cgagtgccga ggtgcaaatg cgaccagacc gtcgggccag ggctaactta
123301 taccccacgc ctttcccctc cccaaagggg cggcagtgac gattccccca atggccgcgc
123361 gtcccagggg aggcaggccc accgcggggc ggccccgtcc ccggggacca acccggcgcc
123421 cccaaagaat atcattagca tgcacggccc ggcccccgat ttgggggacc aacccggtgt
123481 ccccaaaga accccattag catgcccctc ccgccgacgc aacaggggct tggcctgcgt
123541 cggtgccccg gggcttcccg ccttcccgaa gaaactcatt accatacccg gaaccccagg
123601 ggaccaatgc gggttcattg agcgacccgc gggccaatgc gcgaggggcc gtgtgttccg
123661 ccaaaaaagc aattagcata acccggaacc ccaggggagt ggttacgcgc ggcgcgggag
123721 gcggggaata ccggggttgc ccattaaggg ccgcgggaat tgccggaagc gggaagggcg
123781 gccggggccg cccattaatg

FIG. 26E

agtttctaat taccataccg ggaagcggaa caaggcctct

123841 tgcaagtttt taattaccat accgggaagt gggcggcccg gcccattggg cggtaactcc

123901 cgcccaatgg gccgggcccc gaagactcgg cggacgctgg ttggccgggc cccgccgcgc

123961 tggcggccgc cgattggcca gtcccgcccc cgaggcgggc ccgccttggg ggcggaccgg

124021 ctcccagcgt atatatgcgc ggctcctgcc atcgtctctc cggagagcgg cttggtgcgg

124081 agctcccggg agctccgcgg aagacccagg cgcctcgggt gtaacgttag accgagttcg

124141 ccgggccggc tccgcgggcc agggcccggg cacgggcctc gggccccagg cacggcccga

124201 tgaccgcctc ggcctccgcc acccggcgcc ggaaccgagc ccggtcggcc cgctcgcggg

124261 cccacgagcc gcggcgcgcc aggcgggcgg ccgaggccca gaccaccagg tggcgcaccc

124321 ggacgtgggg cgagaagcgc acccgcgcgg gggtcgcggg ggtcgcgggg gtcgcggggg

124381 tcgcgggggt cgcgggggtc gcgggggtcg cgggggtcgc gggggctcc ggcgcccccc

124441 ccccgcccgc gcgtcgcagg cgcaggcgcg ccaggtgctc cgcggtgacg cgcaggcgga

124501 gggcgaggcg cggcggaagg cggaagggcc gcgaggggcg gtgggagggg tcagccccgc

124561 ccccgggccc cacgccgggc ggtggggacc ggggccgggg ggcggcggcg gtgggccggg

124621 cctctggcgc cggctcgggc gggggctgt ccggccagtc gtcgtcatcg tcgtcgtcgg

124681 acgcggactc gggaacgtgg agccactggc gcagcagcag cgaacaagaa ggcgggggcc

124741 caccggcggg gggcggcggc ggggcggccg cgggcgcgct cctgaccgcg ggttccgagt

124801 tgggcgtgga ggttacctgg gactgtgcgg ttgggacggc gcccgtgggc ccgggcggcc

124861 gggggcggcg ggggccgcga tggcggcggc ggcgggccat ggagacagag agcgtgccgg

124921 ggtggtagag tttgacaggc aagcatgtgc gtgcagaggc gagtagtgct tgcctgtcta

124981 actcgctagt ctcggccgcg

FIG. 26F

```
 ggggggcccgg gctgcccgcc gccgcgcttt aaagggccgc
   125041 gcgcgacccc cgggggggtgt gtttcggggg ggcccgtttt tggggtgtgg
ccgctcctcc
   125101 cccgctcctc cccgtctgtg ggtggggctc ctcccccgct cctccccccgc
tcctcccccg
   125161 ctcctccccg tctgtgggtg gggctcctcc cccgctcccg cggccccgcc
ccccacgccc
   125221 gccgcgcgcg cgcacgccgc ccggaccgcc gcccgccttt tttgcgcgcc
gccccgcgcg
   125281 cgggggggccc gggctgccac aggtgtaaca acaccaacag aacaccaaca
gcacggcgca
   125341 ccggcgactc cggttcctca tccacacgtc acgtcatcca acacacctgc
ccaacaacac
   125401 aactcacagc gacaactcac cgcgcaacaa ctcctgttcc tcatccacac
gtcaccgcgc
   125461 acccccgct cctccagacg tcccccagcg caacacgccg ctcctgtcac
acaccacagc
   125521 cccagccctc cccagcccca gccctcccca gccccagccc tccccagccc
cagccctccc
   125581 cagccccagc cctccccagc cccagccctc cccagcccca gccctcccca
gccccagccc
   125641 tccccagccc cagccctccc cagccccagc cctccccagc cccagccctc
cccagcccca
   125701 gccctcccca gccgcgtccc gcgctccctc ggggggggttc gggcatctct
acctcagtgc
   125761 cgccaatctc aggtcagaga tccaaaccct ccgggggcgc ccgcgcacca
ccaccgcccc
   125821 tcgccccctc ccgccctcg ccccctcccg ccctcgccc cctcccgccc
ctcgcccccc
   125881 cccgccccctc gccccctccc gccccctcgcc ccctcccgcc cctcgccccc
tcccgccccc
   125941 cgccccctcc cgcccctcga ataaa
```

FIG. 26G

Human Collagen 1 promoter SEQ ID NO: 20

ga caggtcacta accctcatac taccaagccc tgcggcaccc
2941 tgccctagac caccactcta aatgtctgtt ccctccaaaa acaggacccc
tgtcgcctat
3001 tagggagggg ttctcttgga actgacccac agtagggggc aggactttgg
tgggttcaag
3061 aactgccatc tcagcacctc agccccctag tcctgccctg cagtcgctgg
cactaggcgg
3121 gggcagaccc tgggccacaa gttgctgcca catggtcggg ataattgatg
aaggtccatc
3181 cctccattgc tgtctccagc cctgcctctc tggaaactct atattttccc
tttaattata
3241 gcccctgcag tctccctctg ctgccccacc cgcaccgctc atcctggctg
cccacggcca
3301 gccggccagc cgacgtggct ccctcccctt ctgttccttt ttttcccct
ttgccttcgt
3361 tgcacaaaac cagctgggggg agggcgtgga gaggggcggg gggaggcaat
ggaatcttgg
3421 atggtttggg ggaggcggga ctccccgctt ccacgtttgc agctctggag
cacccggggt
3481 ggggagctgc acaggaggga gagaaatgaa cagggcactg caaggagacc
cccaggcctt
3541 ctctcagccc tacagagttt ctcaggacga ggtagattgg ggttgaggca
gagccttgtt
3601 ggggggaatgg gacatggagg aagaaaggac gtggagttct agagccatct
tccttagata
3661 tagcctgctg tccttcgggt ccccagaccc tttcagagtg tacagatgat
tctctctggt
3721 tcctaaggca tagagcaatg accgggattt tcaagaaaga gatgaggcag
tgggaagtag
3781 ccccctaaaac aaagtcaatc atcctctgca gcccatccca caccccccaaa
ggaaagtttc
3841 acccagacac ccaaaatatc ccatacatcc ccaacactga gtccaggtac
aactggagaa
3901 ggggctttat gcagctccca gaaagacacc cctttagcta agtgccctcc
ctccacccag
3961 gttctctctg gtttgactgt gctgggaagg agggtctcta agcagccccct
ggccacagcc
4021 atggcaaaca aaactcttct ctaagtcacc aatgatcaca ggcctcccac
taaaaatact
4081 tcccaactct ggggtggaag agtttgggggg atgaattttt aggggattgc
aagccccaat
4141 ccccacctct gtgtccctag aatcccccac ccctaccttg gctgctccat
cacccaacca

FIG. 27A

4201 ccaaagcttt cttctgcaga ggccacctag tcatgtttct caccctgcac ctcagcctcc

4261 ccactccatc tctcaatcat gcctagggtt tggaggaggg catttgattc tgttctggag

4321 cacagcagaa gaattgacat cctcaaaatt aaaactccct tgcctgcacc cctccctcag

4381 atatctgatt cttaatgtct agaaaggaat ctgtaaattg ttccccaaat attcctaagc

4441 tccatccccc agccacacca gaagacaccc ccaaacaggc acatcttttt aattcccagc

4501 ttcctctgtt ttggagaggt cctcagcatg cctctttatg cccctccctt agctcttgcc

4561 aggatatcag agggtgactg gggcacagcc aggaggaccc cctccccaac acccccaacc

4621 cttccacctt tggaagtctc cccacccagc tccccagttc cccagttcca cttcttctag

4681 attggaggtc ccaggaagag agcagagggg cacccctacc cactggttag cccacgccat

4741 tctgaggacc cagctgcacc cctaccacag cacctctggc ccaggctggg ctgggggggct

4801 ggggaggcag agctgcgaag aggggagatg tggggtggac tcccttccct cctcctcccc

4861 ctctccattc caactcccaa attgggggcc gggccaggca gctctgattg gctggggcac

4921 gggcggccgg ctccccctct ccgaggggca gggttcctcc ctgctctcca tcaggacagt

4981 ataaaagggg cccgggccag tcgtcggagc agacgggagt ttctc

FIG. 27B

Codon optimized human tumor necrosis factor alpha soluble receptor
SEQ ID NO: 21

CGCGGATCCGCGCAGACTTGCAAAAGAAGGCATGGCACCTGTCGCTGTGTGGGCAGCTCTGGCTGTCG
GACTGGAGCTGTGGGCAGCAGCTCACGCACTGCCAGCACAGGTGGCTTTCACTCCTTACGCACCAGAA
CCCGGCTCCACCTGTAGGCTGCGCGAGTACTATGATCAGACCGCTCAGATGTGCTGTTCCAAGTGCTC
TCCTGGGCAGCATGCAAAAGTCTTTTGTACAAAGACTTCTGACACAGTGTGCGATTCTTGTGAGGACA
GTACCTACACACAGCTGTGGAACTGGGTCCCAGAATGCCTGAGTTGTGGCTCAAGATGCTCCTCTGAC
CAGGTGGAGACTCAGGCCTGTACCCGGGAACAGAATAGAATCTGCACCTGTCGCCCCGGATGGTATTG
CGCTCTGTCCAAACAGGAGGGATGCCGACTGTGTGCACCTCTGCGCAAGTGTCGACCAGGGTTCGGAG
TGGCCAGACCTGGAACTGAAACCAGCGATGTGGTCTGCAAGCCATGTGCTCCCGGCACATTTTCCAAC
ACTACCAGTTCAACTGACATTTGCAGGCCACACCAGATCTGTAACGTGGTCGCCATTCCCGGGAATGC
CTCTATGGACGCTGTGTGCACATCCACTTCTCCCACCCGCAGTATGGCACCTGGAGCAGTCCACCTGC
CTCAGCCAGTGAGCACACGGTCCCAGCATACACAGCCCACTCCTGAGCCAAGCACAGCACCAAGTACT
TCATTTCTGCTGCCTATGGGACCATCACCACCTGCAGAGGGAAGCACCGGCGATGAACCCAAAAGCTG
CGACAAGACCCATACATGCCCACCATGTCCAGCACCTGAGCTGCTGGGAGGACCAAGCGTGTTCCTGT
TTCCTCCAAAGCCTAAAGATACACTGATGATCTCAAGAACCCCCGAGGTCACATGTGTGGTCGTGGAT
GTGAGCCACGAAGACCCTCAGGTCAAGTTCAACTGGTACGTGGACGGCGTCCAGGTGCATAATGCCAA
GACCAAACCCCGGGAGCAGCAGTACAACTCCACCTATAGAGTCGTGTCTGTCCTGACAGTGCTGCACC
AGAACTGGCTGGACGGGAAGGAATATAAGTGCAAAGTGTCCAACAAGGCTCTGCCAGCACCCATTGAG
AAGACAATCTCTAAGGCCAAAGGACAGCCCCGAGAACCTCAGGTGTACACTCTGCCCCCTTCCCGGGA
GGAAATGACCAAAAACCAGGTCTCTCTGACATGTCTGGTGAAGGGGTTCTATCCATCAGATATCGCCG
TGGAGTGGGAAAGCAATGGACAGCCCGAGAACAATTACAAGACAACTCCACCCGTCCTGGACTCTGAT
GGAAGTTTCTTTCTGTATTCCAAACTGACCGTGGACAAGTCTAGGTGGCAGCAGGGCAACGTCTTCAG
CTGCTCCGTGATGCACGAGGCCCTGCACAATCATTACACTCAGAAATCTCTGAGTCTGTCACCCGGCA
AGTGAGACATCAGGGCGCGGATCCGCG

FIG. 28

McKrae strain ICP0 promoter nucleotide sequence SEQ ID NO: 22

cggagctccc gggagctccg caccaagccg ctctccggag agacgatggc aggagccgcg
catatatacg ctgggagccg gtccgccccc aaggcgggcc cgcctcgggg
gcgggactgg
ccaatcggcg gccgccagcg cggcggggcc cggccaacca gcgtccgccg
agtcttcggg
gcccggccca ttgggcggga gttaccgccc aatgggccgg gccgcccact
tcccggtatg
gtaattaaaa acttgcaaga ggccttgttc cgcttcccgg tatggtaatt
agaaactcat
taatgggcgg ccccggccgc ccttcccgct tccggcaatt cccgcggccc
ttaatgggca
accccggtat tccccgcctc ccgcgccgcg cgtaaccact cccctggggt
tccgggttat
gctaattgct tttttggcgg aacacacggc ccctcgcgca ttggcccgcg
ggtcgctcaa
tgaacccgca ttggtccccc ggggttccgg gtatggtaat gagtttcttc
gggaaggcgg
gaagccccgg ggcaccgacg caggccaagc ccctgttgcg tcggcgggag
gggcatgcta
atggggttct ttgggggaca ccgggttggt ccccccaaatc gggggccggg
ccgtgcatgc
taatgatatt ctttgggggc gccgggttgg tccccgggga cggggccgcc
ccgcggtggg
cctgcctccc ctgggacgcg cggccattgg gggaatcgtc actgccgccc
ctttggggag
gggaaaggcg tggggtataa gttagccctg gcccgacggt ctggtcgcat
ttgcacctcg
gcactcggag cgagacgcag cagccaggca gactcgggcc gccccctctc
cgcatcacca
cagaagcccc gcctacgttg cgaccccccag ggaccctccg tccgcgaccc
tccagccgca
tacga

FIG. 29

Synthetic terminator nucleotide sequence SEQ ID NO: 23

AATAAAATATCTTTATTTTCATTACATCTGTGTGTTGGTTTTTTGTGTG

FIG. 30

Immediate early gene specific promoter sequence TAATGARAT
SEQ ID NO: 24

TAATGARAT where R is a purine

FIG. 31

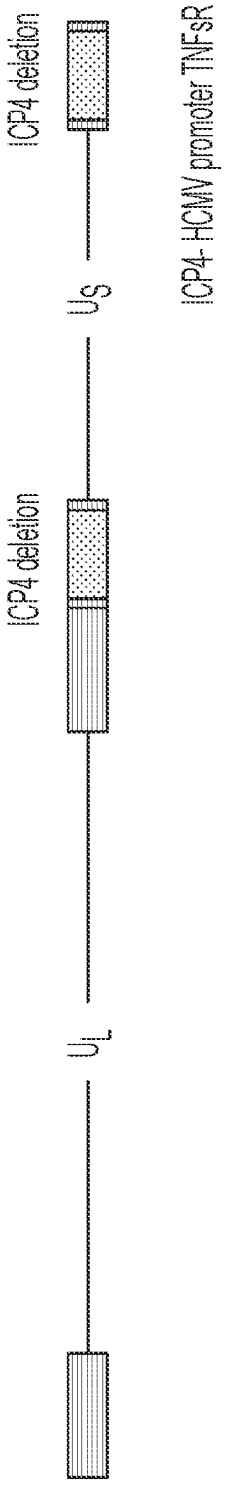


FIG. 32A

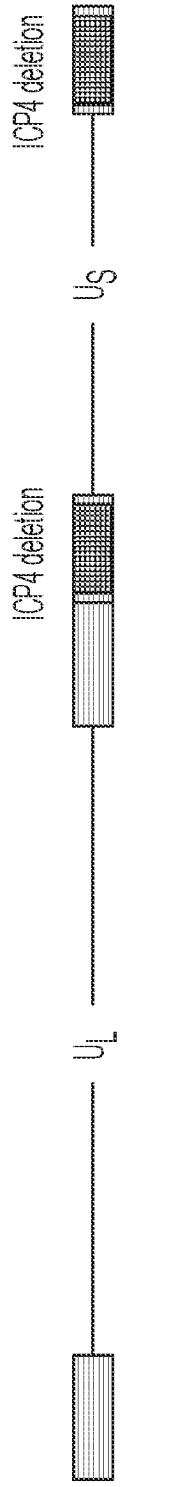


FIG. 32B

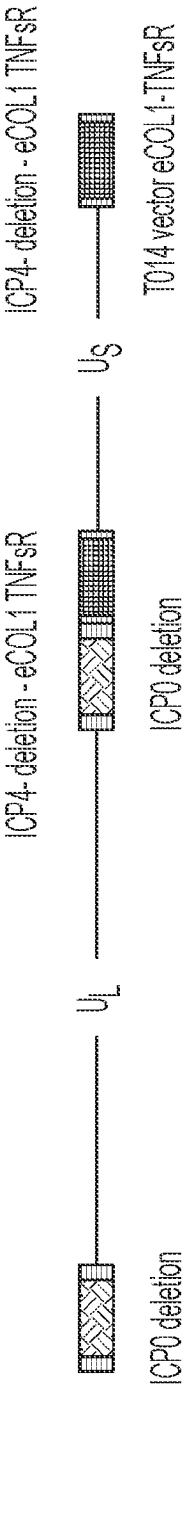


FIG. 32C

HSV VECTORS

SEQUENCE LISTING

[0001] In accordance with 37 C.F.R. 1.52 (e) (5), the present specification makes reference to a Sequence Listing (submitted electronically as a .txt file named "P113870003WO00-SEQ-ROS"). The .txt file was generated on Apr. 15, 2022, and is 96,682 bytes in size. The Sequence Listing is herein incorporated by reference in its entirety.

BACKGROUND

[0002] Systemic delivery of certain therapeutic agents can be problematic for agents with poor pharmacokinetics and/or a risk of off target adverse effects. Local injection at particular target sites may require highly invasive techniques or be infeasible. Delivery of agents by viral vectors allows the ability to specifically target cell populations to provide local production and/or delivery of agents.

SUMMARY OF THE INVENTION

[0003] The present disclosure provides compositions and methods for viral vector delivery of agents to target cells. In some embodiments, the disclosure provides vectors for use in administering nucleic acids that encode polypeptides for therapeutic or cosmetic indications to specific tissues (e.g., skin).

[0004] Aspects of the disclosure relate to a method of expressing a polypeptide in a tissue of a subject comprising administering to the subject a vector comprising a variant of a herpes simplex virus (HSV) strain whose genome contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins (e.g., "ICP0−/ICP4−").

[0005] In some embodiments, the HSV strain is an HSV-1 strain. In some embodiments, the HSV-1 strain is a McKrae strain.

[0006] In some embodiments, the variant fails to express functional ICP4 and ICP0 proteins characterized by the amino acid sequences of SEQ ID NO: 2 and SEQ ID NO: 17, respectively.

[0007] In some embodiments, the tissue comprises epidermis, dermis, or subcutaneous fat or muscle. In some embodiments, the tissue is skin tissue. In some embodiments, the skin tissue comprises fibroblasts, keratinocytes, adipocytes, or muscle cells.

[0008] In some embodiments, the vector comprises a tissue specific promoter. In some embodiments, the tissue is skin. In some embodiments, the vector comprises a collagen 1 promoter. In some embodiments, the vector comprises a human cytomegalovirus (HCMV) enhancer. In some embodiments, the vector comprises a bovine growth hormone (BGH) polyadenylation signal or an HSV viral polyadenylation signal.

[0009] In some embodiments, the vector comprises a nucleic acid that encodes a therapeutic polypeptide. In some embodiments, the vector comprises a nucleic acid that encodes a therapeutic polypeptide for use in treatment of skin.

[0010] In some embodiments, the vector is administered in vivo. In some embodiments, the vector is administered by contact with skin. In some embodiments, the vector is administered by intradermal injection.

[0011] In some embodiments, the alteration is a disruption or a deletion of the ICP0 and ICP4 genes. In some embodiments, the ICP0 and/or ICP4 gene is disrupted by deletion of its respective promoter.

[0012] In some embodiments, the disclosure provides variants of HSV McKrae strain whose genomes contain an alteration such that the variant fails to express functional ICP0 and ICP4 proteins. In some embodiments, the disclosure provides an HSV strain comprising a variant HSV strain genome which contains an alteration such that the variant fails to express functional ICP4 and ICP0 proteins characterized by the amino acid sequences of SEQ ID NO: 2 and SEQ ID NO: 17, respectively. In some embodiments, the variant is a replication-defective variant.

[0013] In some embodiments, vectors are provided comprising a variant HSV strain as described herein.

[0014] In some embodiments, the disclosure provides pharmaceutical compositions comprising a vector as described herein and a pharmaceutically acceptable carrier.

[0015] In some embodiments, the disclosure provides cells transduced with a vector as described herein.

[0016] In some embodiments, provided herein are methods of propagating a vector comprising a variant HSV genome which contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins, the methods comprising steps of (i) infecting cultured ICP0 and ICP4 complementing cells containing a nucleic acid (e.g., DNA) encoding HSV proteins ICP0 and ICP4 with the vector, and (ii) isolating supernatant from the culture of step (i).

[0017] In some embodiments, the methods comprise a step of purifying a vector in the supernatant by chromatography. In some embodiments, the methods comprise a step of concentrating the purified vector by tangential flow filtration.

[0018] In some embodiments, provided herein are methods of preparing a vector comprising a variant HSV genome which contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins, and wherein the vector expresses a marker element, the method comprising incubating cells transfected with: (a) a first nucleic acid molecule: (i) comprising a portion of HSV genome but does not encode functional ICP0 and ICP4 proteins; and (ii) comprising a first homology region (HR1) and a second homology region (HR2), and (b) a second nucleic acid molecule comprising a sequence that encodes a marker element, wherein the sequence is flanked by a first homology region (HR1') and a second homology region (HR2'), wherein HR1 is homologous to HR1' and HR2 is homologous to HR2' such that the sequence that encodes the marker element in the second nucleic acid molecule integrates into the first nucleic acid molecule via homologous recombination.

[0019] In some embodiments, the cells are ICP0 and/or ICP4 complementing cells. In some embodiments, the marker element is a polypeptide. In some embodiments, the polypeptide is a soluble tumor necrosis factor receptor.

[0020] In some embodiments, the polypeptide is quantified by enzyme linked immunosorbent assay (ELISA). In some embodiments, the polypeptide is detected by fluorescence.

[0021] In some embodiments, the methods comprise a step of purifying viral plaques that express the marker element.

[0022] In some embodiments, the disclosure provides methods of preparing a vector comprising a variant HSV genome which contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins, and wherein the vector expresses an agent of interest, the method comprising incubating cells transfected with: (a) a first nucleic acid molecule: (i) comprising a portion of HSV genome but does not encode functional ICP0 and ICP4 proteins; and (ii) comprising a sequence that encodes a marker element, wherein the sequence that encodes the marker element is flanked by a first homology region (HR1) and a second homology region (HR2); and (b) a second nucleic acid molecule comprising a sequence that encodes an agent of interest, wherein the sequence encoding the agent of interest is flanked by a first homology region (HR1') and a second homology region (HR2'), wherein HR1 is homologous to HR1' and HR2 is homologous to HR2' such the sequence encoding the agent of interest is integrated into the first nucleic acid molecule via homologous recombination.

[0023] In some embodiments, the cells are ICP0 and/or ICP4 complementing cells.

[0024] In some embodiments, the methods comprise a step of purifying viral plaques that do not express the marker element.

[0025] In some embodiments of the methods described herein, the HSV genome is an HSV-1 genome. In some embodiments, the HSV genome is a McKrae strain genome.

[0026] In some embodiments, the disclosure provides methods of measuring transduction efficiency of an HSV vector in a skin tissue, the methods comprising: (a) contacting the skin tissue of an animal with an HSV vector as described herein; (b) removing a portion of the skin tissue from the animal; and (c) assaying the number of HSV genomes transduced in the skin tissue.

[0027] In some embodiments, the skin tissue comprises fibroblasts, keratinocytes, adipocytes, muscle cells, epidermis, dermis, hypodermis, or underlying subcutaneous fat or muscle.

[0028] In some embodiments, the number of genomes is measured by an amplification technique. In some embodiments, the amplification technique is quantitative polymerase chain reaction (qPCR).

[0029] In some embodiments, the disclosure provides methods of measuring transduction efficiency of an HSV vector that contains an expression cassette comprising a polypeptide payload in a skin tissue, the methods comprising: (a) contacting the skin tissue of an animal with an HSV vector as described herein; (b) removing a portion of the skin tissue from the animal; and (c) assaying the amount of a polypeptide encoded by a nucleic acid of the expression cassette.

[0030] In some embodiments, the skin tissue comprises fibroblasts, keratinocytes, adipocytes, muscle cells, epidermis, dermis, hypodermis, or underlying subcutaneous fat or muscle.

[0031] In some embodiments, the amount of polypeptide is measured by an immunoassay. In some embodiments, the immunoassay is an ELISA or immunohistochemistry (IHC). In some embodiments, the ELISA or IHC is performed on tissue of the epidermis, dermis, subcutaneous tissue, subcutaneous fat, underlying muscle, or draining lymph node.

BRIEF DESCRIPTION OF THE DRAWINGS

[0032] The drawings are for illustration purposes only, not for limitation.

[0033] FIG. **1** depicts an exemplary western blot for ICP0 protein in McKrae HSV-1 mutants deleted for ICP0 and ICP4 (PGN04 99i8.1, PGN04 99i8.4, PGN04 99i8.5, PGN04 99i8.6, PGN04 99i8.7, and PGN04 99i8.8). The ICP0−/ICP4− mutants contain HCMV-mCherry in the ICP4 locus. No expression of ICP0 was detected in any of the ICP0−/ICP4− mutants. A control ICP4− mutant is also shown, which does express ICP0. The parent vector for the ICP0−/ICP4− mutants is an ICP4-McKrae strain vector, as previously described in International Publication No. WO2017165813A1, incorporated herein by reference in its entirety.

[0034] FIGS. **2**A and **2**B each depict an exemplary PCR analysis of ICP4-mutants sequentially deleted for ICP0. FIG. **2**A shows a polymerase chain reaction (PCR) experiment for ICP0 (top half) and for UL56 (bottom half) for mutants PGN04 99i8.1, PGN04 99i8.2, PGN04 99i8.3, PGN04 99i8.4, PGN04 99i8.5, PGN04 99i8.6, PGN04 99i8.7, and PGN04 12H2. FIG. **2**B shows a PCR experiment for ICP0 for mutants PGN04 99i8.1, PGN04 99i8.2, PGN04 99i8.3, PGN04 99i8.4, PGN04 99i8.5, PGN04 99i8.6, PGN04 99i8.7, and for a CT1 positive control. Vector PGN04 99.i8 and 12H2 viruses were negative for ICP0 DNA.

[0035] FIGS. **3**A and **3**B show that green fluorescent protein (GFP) expression is enhanced in an ICP0−/ICP4− mutant (PGN04). FIG. **3**A shows an ICP4− mutant (MOI=5, measured 18 hours after infection) with GFP expression driven by the human cytomegalovirus (HCMV) promoter. FIG. **3**B shows an ICP0−/ICP4− mutant (PNG04) (MOI=5, measured 18 hours after infection), with GFP expression driven by the ICP0 promoter. The deletion of multiple immediate early genes, including (1) ICP0−/ICP4−/ICP22−/ICP27−, (2) ICP0−/ICP4−/ICP22−, (3) ICP0−/ICP4−/ICP27−, and (4) ICP0−/ICP22−/ICP27−, results in little gene expression on non-complement cells. The ICP0−/ICP4− mutant (PGN04) (FIG. **3**B) shows good GFP expression that is similar to a virus that contains ICP0 (FIG. **3**A). Given that both the ICP0−/ICP4−/ICP27− and ICP0−/ICP4−/ICP22-mutants exhibited poor expression but the ICP0−/ICP4− mutant (PGN04) exhibited good expression, the data demonstrate that ICP22 and ICP27 are both required for useful gene expression in an ICP0− mutant. Thus, ICP22 and ICP27 do not need to be deleted to reduce vector toxicity.

[0036] FIGS. **4**A-**4**D each depict an exemplary photo that shows efficient gene expression of GFP driven by the ICP0 promoter, even in the absence of ICP0. Collectively, FIGS. **4**A-**4**D show the time course of gene expression for an ICP0−/ICP4− mutant (PGN04.6; MOI=10) in Vero cells (dividing) in culture. FIG. **4**A shows gene expression of the ICP0−/ICP4− mutant (PGN04.6; MOI=10) after 2 days. FIG. **4**B shows gene expression of the ICP0−/ICP4− mutant (PGN04.6; MOI=10) after 3 days. FIG. **4**C shows gene expression of the ICP0−/ICP4− mutant (PGN04.6; MOI=10) after 4 days. FIG. **4**D shows gene expression of the ICP0−/ICP4− mutant (PGN04.6; MOI=10) after 5 days. ICP0-viruses generally do not show good gene expression; however, the ICP0−/ICP4− mutant (PGN04.6) did (see FIGS. **4**A-**4**D). This retention of gene expression in the ICP0−/ICP4-mutant (PGN04.6) could be due to the retention of

ICP22 and ICP27, which play a role in the innate immune response evasion of the virus.

[0037] FIG. **5** demonstrates the effect of ICP0 on tissue-specific promoters, and depicts an exemplary graph that shows the relative expression of tumor necrosis factor alpha soluble receptor (sTNFR) from an ICP4− mutant (ICP0+) with a HCMV promoter, versus an ICP4− mutant (ICP0+) with a large collagen-1 (eCOL1) promoter, versus an ICP0−/ICP4-virus control (with ICP0-GFP and HCMV-Red) and an ICP0−/ICP4− mutant with a large collagen-1 promoter (eCOL1) in cell lines of epidermal and neuronal lineage. Gene expression results (in picograms per mL sTNFR (HCMV 100×)) are shown for each mutant for three different cell types: HaCaT (keratinocyte cell line), SK-N-MC (neuroblastoma cells), and Rin5F (insulinoma cells). Both SK-N-MC and Rin5F cells express neuronal markers. It is specifically shown that in the presence of ICP0, expression tissue specificity of promoter sequences is lost. Thus, ICP0 deletion is required for tissue specificity via a tissue-specific promoter. Additionally, the ICP0-mutant containing the eCOL1 promoter showed no activity in the keratinocyte (HaCaT) or neuronal cell types (SK-N-MC or Rin5F).

[0038] FIGS. **6**A and **6**B show that a PGN04 vector with an eCOL1 promoter only expresses in fibroblasts, and each depict an exemplary graph of tissue specific gene expression of sTNFR in a fibroblast tissue cell line and human fibroblasts from a collagen-1 promoter (eCOL1) in ICP0−/ICP4− mutants, as well as the further induction of that promoter by a transforming growth factor beta peptide (tgf-b). FIG. **6**A shows sTNFR expression in BJ fibroblasts (in picograms per mL sTNFR, measured after 3 days). FIG. **6**B shows STNFR expression in NCTC 2472 cells (mouse fibroblastic tumor) (in picograms per mL STNFR). All viruses shown in FIGS. **6**A and **6**B are ICP0−/ICP4− mutants. FIGS. **6**A and **6**B show that the eCOL1 promoter has expression which is limited to only fibroblasts in ICP0-viruses, and does not express in neurons or keratinocytes. This pattern of expression is ideal for skin indications and the results demonstrate a good safety profile as compared to ICP0+ viruses and the use of promoters like HCMV. The results also demonstrate that use of the eCOL1 promoter will protect neurons from any off-target expression. Lastly, the results show that expression from the eCOL1 promoter can be induced by TGF-β in media, and thus the expression is both tissue-specific and inducible.

[0039] FIG. **7** shows COL7 expression from a vector containing a tissue-specific promoter, and depicts an exemplary western blot that demonstrates an ICP0−/ICP4− mutant that expresses Collagen-7 (COL7) from a large Collagen-1 promoter. Viral isolates were grown on ICP0/ICP4 complementing cells. The results of FIG. **7** demonstrate that the tissue-specific promoter exhibited some expression while the virus was replicating.

[0040] FIGS. **8**A-**8**CM depict an exemplary HSV McKrae strain nucleotide sequence (SEQ ID NO: 1) which is identified as accession number JQ730035.1.

[0041] FIG. **9** depicts an exemplary HSV McKrae strain ICP4 amino acid sequence (SEQ ID NO: 2).

[0042] FIG. **10** depicts an exemplary HSV McKrae strain ICP22 amino acid sequence (SEQ ID NO: 3).

[0043] FIG. **11** depicts an exemplary HSV McKrae strain ICP47 amino acid sequence (SEQ ID NO: 4).

[0044] FIGS. **12**A-**12**C depict an exemplary HSV McKrae strain nucleotide sequence of ICP4 (SEQ ID NO: 5).

[0045] FIG. **13** depicts an exemplary HSV McKrae strain nucleotide sequence of ICP22 (SEQ ID NO: 6).

[0046] FIG. **14** depicts an exemplary HSV McKrae strain nucleotide sequence of ICP47 (SEQ ID NO: 7).

[0047] FIG. **15** depicts an exemplary human cytomegalovirus enhancer nucleotide sequence (SEQ ID NO: 8).

[0048] FIG. **16** depicts an exemplary calcitonin gene-related peptide promoter nucleotide sequence (SEQ ID NO: 9).

[0049] FIG. **17** depicts an exemplary bovine growth hormone polyadenylation signal nucleotide sequence (SEQ ID NO: 10).

[0050] FIG. **18** depicts an exemplary STPSTTT amino acid sequence (SEQ ID NO: 11).

[0051] FIG. **19** depicts an exemplary latency associated transcript intron variation 1 (SEQ ID NO: 12).

[0052] FIG. **20** depicts an exemplary latency associated transcript variation 2 (SEQ ID NO: 13).

[0053] FIG. **21** depicts an exemplary latency associated transcript variation 3 (SEQ ID NO: 14).

[0054] FIG. **22** depicts an exemplary Alanine rich region of KOS ICP4 amino acid sequence (SEQ ID NO: 15).

[0055] FIG. **23** depicts an exemplary Serine rich region of McKrae ICP4 amino acid sequence (SEQ ID NO: 16).

[0056] FIG. **24** depicts an exemplary McKrae strain amino acid sequence of ICP0 (SEQ ID NO: 17).

[0057] FIGS. **25**A-**25**C depict an exemplary McKrae strain DNA sequence of ICP0 (SEQ ID NO: 18).

[0058] FIGS. **26**A-**26**G depict an exemplary McKrae strain latency associated transcript nucleotide sequence (SEQ ID NO: 19).

[0059] FIGS. **27**A and **27**B depict an exemplary human Collagen 1 promoter nucleotide sequence (SEQ ID NO: 20).

[0060] FIG. **28** depicts an exemplary human tumor necrosis factor alpha soluble receptor amino acid sequence (SEQ ID NO: 21).

[0061] FIG. **29** depicts an exemplary McKrae strain ICP0 promoter sequence (SEQ ID NO: 22).

[0062] FIG. **30** depicts an exemplary synthetic terminator (SEQ ID NO: 23).

[0063] FIG. **31** depicts an exemplary TAATGARAT sequence (SEQ ID NO: 24).

[0064] FIGS. **32**A-**32**C depict a schematic of an exemplary replication defective McKrae strain viral vector. FIG. **32**A shows complete deletions of both copies of the viral ICP4 genes, and a human cytomegalovirus (HCMV) immediate early promoter driven expression cassette inserted within both copies of the deleted ICP4 loci. FIG. **32**B shows complete deletions of both copies of the viral ICP4 genes, and an extended Collagen-1 (COL1) promoter driven expression cassette inserted within both copies of the deleted ICP4 loci. FIG. **32**C shows complete deletions of both copies of the viral ICP0 and ICP4 genes, and an extended Collagen-1 (COL1) promoter driven expression cassette inserted within both copies of the deleted ICP4 loci. Each expression cassette in FIGS. **32**A-**32**C contains a payload of interest for expression in target cells.

DEFINITIONS

[0065] In this application, unless otherwise clear from context, (i) the term "a" may be understood to mean "at least one"; (ii) the term "or" may be understood to mean "and/or"; (iii) the terms "comprising" and "including" may be understood to encompass itemized components or steps whether

presented by themselves or together with one or more additional components or steps; and (iv) the terms "about" and "approximately" may be understood to permit standard variation as would be understood by those of ordinary skill in the art; and (v) where ranges are provided, endpoints are included.

[0066] Administration: As used herein, the term "administration" refers to the administration of a composition to a subject or system. Administration to an animal subject (e.g., to a human) may be by any appropriate route. For example, in some embodiments, administration may be bronchial (including by bronchial instillation), buccal, enteral, interdermal, intra-arterial, intradermal, intragastric, intramedullary, intramuscular, intranasal, intraperitoneal, intrathecal, intravenous, intraventricular, within a specific organ (e.g., intrahepatic), mucosal, nasal, oral, rectal, subcutaneous, sublingual, topical, tracheal (including by intratracheal instillation), transdermal, vaginal and vitreal. In some embodiments, administration may involve intermittent dosing. In some embodiments, administration may involve continuous dosing (e.g., perfusion) for at least a selected period of time.

[0067] Agent: As used herein, the term "agent" refers to a compound or entity of any chemical class including, for example, polypeptides, nucleic acids, saccharides, lipids, small molecules, or combinations thereof. In some embodiments, an agent is or comprises a natural product in that it is found in and/or is obtained from nature. In some embodiments, an agent is or comprises one or more entities that is man-made, in that it is designed, engineered, and/or produced through action of the hand of man and/or is not found in nature. Some particular embodiments of agents that may be utilized in accordance with the present invention include small molecules, antibodies, antibody fragments, aptamers, nucleic acids (e.g., siRNAs, shRNAs, DNA/RNA hybrids, antisense oligonucleotides, ribozymes), peptides, peptide mimetics, etc.

[0068] Amelioration: As used herein, the term "amelioration" refers to the prevention, reduction or palliation of a state, or improvement of the state of a subject. Amelioration includes, but does not require, complete recovery or complete prevention of a disease, disorder or condition.

[0069] Animal: As used herein, the term "animal" refers to any member of the animal kingdom. In some embodiments, "animal" refers to humans, of either sex and at any stage of development. In some embodiments, "animal" refers to non-human animals, at any stage of development. In certain embodiments, the non-human animal is a mammal (e.g., a rodent, a mouse, a rat, a rabbit, a monkey, a dog, a cat, a sheep, cattle, a primate, and/or a pig). In some embodiments, animals include, but are not limited to, mammals, birds, reptiles, amphibians, fish, insects, and/or worms. In some embodiments, an animal may be a transgenic animal, genetically engineered animal, and/or a clone.

[0070] Approximately: As used herein, the term "approximately" or "about," as applied to one or more values of interest, refers to a value that is similar to a stated reference value. In certain embodiments, the term "approximately" or "about" refers to a range of values that fall within 25%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, 1%, or less in either direction (greater than or less than) of the stated reference value unless otherwise stated or otherwise evident from the context (except where such number would exceed 100% of a possible value).

[0071] Characteristic sequence: As used herein, the term "characteristic sequence" or "conserved sequence" refers to a sequence that is found in all members of a family of polypeptides or nucleic acids, and therefore can be used by those of ordinary skill in the art to define members of the family.

[0072] Combination therapy: As used herein, the term "combination therapy" refers to those situations in which a subject is simultaneously exposed to two or more therapeutic regimens (e.g., two or more therapeutic agents). In some embodiments, two or more agents may be administered simultaneously; in some embodiments, such agents may be administered sequentially; in some embodiments, such agents are administered in overlapping dosing regimens.

[0073] Composition: As used herein, the term "composition" or a "pharmaceutical composition" refers to the combination of two or more agents as described herein for co-administration or administration as part of the same regimen. It is not required in all embodiments that the combination of agents result in physical admixture, that is, administration of each component of the composition as a separate co-agent is possible; however many patients or practitioners in the field may find it advantageous to prepare a composition that is an admixture of two or more of the ingredients in a pharmaceutically acceptable carrier, diluent, or excipient, making it possible to administer the component ingredients of the combination at the same time.

[0074] Engineered: As used herein, the term "engineered" refers to the aspect of having been manipulated by the hand of man. For example, a polynucleotide is considered to be "engineered" when two or more sequences, that are not linked together in that order in nature, are manipulated by the hand of man to be directly linked to one another in the engineered polynucleotide. For example, in some embodiments of the present disclosure, an engineered polynucleotide comprises a regulatory sequence which, in nature, is found in operative association with a first coding sequence but not in operative association with a second coding sequence, is linked by the hand of man so that it is operatively associated with the second coding sequence. Comparably, a cell or organism is considered to be "engineered" if it has been manipulated so that its genetic information is altered (e.g., new genetic material not previously present has been introduced, for example by transformation, mating, somatic hybridization, transfection, transduction, or other mechanism, or previously present genetic material is altered or removed, for example by substitution or deletion mutation, or by mating protocols). As is common practice and is understood by those in the art, progeny of an engineered polynucleotide or cell are typically still referred to as "engineered" even though the actual manipulation was performed on a prior entity.

[0075] Expression: As used herein, "expression" of a nucleic acid sequence refers to one or more of the following events: (1) production of an RNA template from a DNA sequence (e.g., by transcription); (2) processing of an RNA transcript (e.g., by splicing, editing, 5' cap formation, and/or 3' end formation); (3) translation of an RNA into a polypeptide or protein; and/or (4) post-translational modification of a polypeptide or protein.

[0076] Functional protein: As used herein, a "functional protein", e.g., a functional HSV immediate early (IE) protein, e.g., a functional ICP0, ICP4, ICP22, ICP27, and/or ICP47 protein, refers to a protein that exhibits at least 40%, 50%, 60%, 70%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of the activity of a corresponding full length, wild-type protein. For example, a functional ICP4 protein refers to a protein that exhibits at least 40%, 50%, 60%, 70%, 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% of the activity (e.g., DNA binding activity or transcriptional regulatory activity) of a corresponding full length, wild-type ICP4 protein.

[0077] Homology: As used herein, the term "homology" refers to the overall relatedness between polymeric molecules, e.g., between nucleic acid molecules (e.g., DNA molecules and/or RNA molecules) and/or between polypeptide molecules. In some embodiments, polymeric molecules are considered to be "homologous" to one another if their sequences are at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% identical. In some embodiments, polymeric molecules are considered to be "homologous" to one another if their sequences are at least 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 99% similar.

[0078] Isolated: As used herein, the term "isolated" refers to a substance and/or entity that has been (1) separated from at least some of the components with which it was associated when initially produced (whether in nature and/or in an experimental setting), and/or (2) designed, produced, prepared, and/or manufactured by the hand of man. Isolated substances and/or entities may be separated from about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more than about 99% of the other components with which they were initially associated. In some embodiments, isolated agents are about 80%, about 85%, about 90%, about 91%, about 92%, about 93%, about 94%, about 95%, about 96%, about 97%, about 98%, about 99%, or more than about 99% pure. As used herein, a substance is "pure" if it is substantially free of other components. In some embodiments, as will be understood by those skilled in the art, a substance may still be considered "isolated" or even "pure", after having been combined with certain other components such as, for example, one or more carriers or excipients (e.g., buffer, solvent, water, etc.); in such embodiments, percent isolation or purity of the substance is calculated without including such carriers or excipients. To give but one example, in some embodiments, a biological polymer such as a polypeptide or polynucleotide that occurs in nature is considered to be "isolated" when, a) by virtue of its origin or source of derivation it is not associated with some or all of the components that accompany it in its native state in nature; b) it is substantially free of other polypeptides or nucleic acids of the species that produces it in nature; c) is expressed by or is otherwise in association with components from a cell or other expression system that is not of the species that produces it in nature. Thus, for instance, in some embodiments, a polypeptide that is chemically synthesized or is synthesized in a cellular system different from that which produces it in nature is considered to be an "isolated" polypeptide. Alternatively or additionally, in some embodiments, a polypeptide that has been subjected to one or more purification techniques may be considered to be an "isolated" polypeptide to the extent that it has been separated from other components a) with which it is associated in nature; and/or b) with which it was associated when initially produced.

[0079] Marker element: As used herein, the term "marker element" refers to a detectable or selectable agent. In some embodiments, a "marker element" is a detectable or selectable nucleic acid sequence. In some embodiments a "marker element" is an expression product (e.g., RNA or protein) whose presence or absence is detectable and/or selectable in cells. In some embodiments, an expression product is or comprises an enzyme. In some embodiments, an expression product is a fluorophore.

[0080] Nucleic acid: As used herein, the term "nucleic acid" refers to any compound and/or substance that is or can be incorporated into an oligonucleotide chain. In some embodiments, a nucleic acid is a compound and/or substance that is or can be incorporated into an oligonucleotide chain via a phosphodiester linkage. As will be clear from context, in some embodiments, "nucleic acid" refers to individual nucleic acid residues (e.g., nucleotides and/or nucleosides); in some embodiments, "nucleic acid" refers to an oligonucleotide chain comprising individual nucleic acid residues. In some embodiments, a "nucleic acid" is or comprises RNA; in some embodiments, a "nucleic acid" is or comprises DNA. In some embodiments, a nucleic acid is, comprises, or consists of one or more natural nucleic acid residues. In some embodiments, a nucleic acid is, comprises, or consists of one or more nucleic acid analogs. In some embodiments, a nucleic acid analog differs from a nucleic acid in that it does not utilize a phosphodiester backbone. For example, in some embodiments, a nucleic acid is, comprises, or consists of one or more "peptide nucleic acids", which are known in the art and have peptide bonds instead of phosphodiester bonds in the backbone, and are considered within the scope of the present disclosure. Alternatively or additionally, in some embodiments, a nucleic acid has one or more phosphorothioate and/or 5'-N-phosphoramidite linkages rather than phosphodiester bonds. In some embodiments, a nucleic acid is, comprises, or consists of one or more natural nucleosides (e.g., adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxy guanosine, and deoxycytidine). In some embodiments, a nucleic acid is, comprises, or consists of one or more nucleoside analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3-methyl adenosine, 5-methylcytidine, C-5 propynyl-cytidine, C-5 propynyl-uridine, 2-aminoadenosine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5-propynyl-cytidine, C5-methylcytidine, 2-aminoadenosine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, 0 (6)-methylguanine, 2-thiocytidine, methylated bases, intercalated bases, and combinations thereof). In some embodiments, a nucleic acid comprises one or more modified sugars (e.g., 2'-fluororibose, ribose, 2'-deoxyribose, arabinose, and hexose) as compared with those in natural nucleic acids. In some embodiments, a nucleic acid has a nucleotide sequence that encodes a functional gene product such as an RNA or protein. In some embodiments, a nucleic acid includes one or more introns. In some embodiments, nucleic acids are prepared by one or more of: isolation from a natural source, enzymatic synthesis by

polymerization based on a complementary template (in vivo or in vitro), reproduction in a recombinant cell or system, and/or chemical synthesis. In some embodiments, a nucleic acid is at least 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000 or more residues long. In some embodiments, a nucleic acid is single stranded; in some embodiments, a nucleic acid is double stranded. In some embodiments a nucleic acid has a nucleotide sequence comprising at least one element that encodes, or is the complement of a sequence that encodes, a polypeptide. In some embodiments, a nucleic acid has enzymatic activity.

[0081] Operably linked: The term "operably linked" refers to a linkage of polynucleotide elements in a functional relationship. A nucleic acid sequence is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For instance, a promoter or enhancer is operably linked to a gene if it affects the transcription of the gene. Operably linked nucleotide sequences are typically contiguous. However, as enhancers generally function when separated from the promoter by several kilobases and intronic sequences may be of variable lengths, some polynucleotide elements may be operably linked but not directly flanked and may even function in trans from a different allele or chromosome.

[0082] Patient: As used herein, the term "patient" refers to any organism to which a provided composition is or may be administered, e.g., for experimental, diagnostic, prophylactic, cosmetic, and/or therapeutic purposes. Typical patients include animals (e.g., mammals such as mice, rats, rabbits, non-human primates, and/or humans). In some embodiments, a patient is a human. In some embodiments, a patient is suffering from or susceptible to one or more disorders or conditions. In some embodiments, a patient displays one or more signs or symptoms of a disorder or condition. In some embodiments, a patient has been diagnosed with one or more disorders or conditions. In some embodiments, the patient is receiving or has received certain therapy to diagnose and/or to treat a disease, disorder, or condition.

[0083] Pharmaceutical composition: As used herein, the term "pharmaceutical composition" refers to an active agent, formulated together with one or more pharmaceutically acceptable carriers. In some embodiments, active agent is present in a unit dose amount, appropriate for administration in a therapeutic regimen, that shows a statistically significant probability of achieving a predetermined therapeutic effect when administered to a relevant population. In some embodiments, pharmaceutical compositions may be specially formulated for administration in solid or liquid form, including those adapted for the following: oral administration, for example, drenches (aqueous or non-aqueous solutions or suspensions), tablets (e.g., those targeted for buccal, sublingual, and systemic absorption), boluses, powders, granules, pastes for application to the tongue, etc.; parenteral administration, for example, by subcutaneous, intramuscular, intravenous or epidural injection as, for example, a sterile solution or suspension, or sustained-release formulation, etc.; topical application, for example, as a cream, ointment, or a controlled-release patch or spray applied to the skin, lungs, or oral cavity, etc.; intravaginally or intrarectally, for example, as a pessary, cream, suppository, or foam, etc.; sublingual administration; ocular administration; transdermal administration; or nasal administration, pulmonary administration, and/or administration to other mucosal surfaces.

[0084] Pharmaceutically acceptable: As used herein, the term "pharmaceutically acceptable", as applied to the carrier, diluent, or excipient used to formulate a composition as disclosed herein, means that the carrier, diluent, or excipient must be compatible with the other ingredients of the composition and not deleterious to the recipient thereof.

[0085] Pharmaceutically acceptable carrier: As used herein, the term "pharmaceutically acceptable carrier" means a pharmaceutically-acceptable material, composition or vehicle, such as a liquid or solid filler, diluent, excipient, or solvent encapsulating material, that is involved in carrying or transporting the subject compound from one organ or portion of the body to another organ or portion of the body. Each carrier must be "acceptable" in the sense of being compatible with the other ingredients of the formulation and not injurious to the patient. Some examples of materials which can serve as pharmaceutically-acceptable carriers include: sugars, such as lactose, glucose and sucrose; starches, such as corn starch and potato starch; cellulose, and its derivatives, such as sodium carboxymethyl cellulose, ethyl cellulose and cellulose acetate; powdered tragacanth; malt; gelatin; talc; excipients, such as cocoa butter and suppository waxes; oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; glycols, such as propylene glycol; polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol; esters, such as ethyl oleate and ethyl laurate; agar; buffering agents, such as magnesium hydroxide and aluminum hydroxide; alginic acid; pyrogen-free water; isotonic saline; Ringer's solution; ethyl alcohol; pH buffered solutions; polyesters, polycarbonates and/or polyanhydrides; and other non-toxic compatible substances employed in pharmaceutical formulations.

[0086] Prevent or prevention: As used herein, the term "prevent" or "prevention", when used in connection with the occurrence of a disease, disorder, and/or condition, refers to reducing the risk of developing the disease, disorder and/or condition, and/or to delaying the onset of one or more characteristics or symptoms of the disease, disorder or condition. Prevention may be considered complete when onset of a disease, disorder or condition has been delayed for a predefined period of time.

[0087] Subject: As used herein, the term "subject" refers to a mammal (e.g., a human, in some embodiments including prenatal human forms). In some embodiments, a subject is suffering from a relevant disease, disorder or condition. In some embodiments, a subject is susceptible to a disease, disorder, or condition. In some embodiments, a subject displays one or more symptoms or characteristics of a disease, disorder or condition. In some embodiments, a subject does not display any symptom or characteristic of a disease, disorder, or condition. In some embodiments, a subject is someone with one or more features which are characteristic of a susceptibility to or a risk of a disease, disorder, or condition. In some embodiments, a subject is a patient. In some embodiments, a subject is an individual to whom diagnosis and/or therapy is and/or has been administered.

[0088] Therapeutic transgene: The term "transgene" refers to an exogenous gene or polynucleotide sequence. The term

"therapeutic transgene" refers to a transgene, which, when expressed in or by a virus, imparts a therapeutic effect in a target cell, body fluid, tissue, organ, physiological system, or subject.

[0089] Treatment: As used herein, the term "treatment" (also "treat" or "treating") refers to any administration of a substance that partially or completely alleviates, ameliorates, relieves, inhibits, delays onset of, reduces severity of, and/or reduces incidence of one or more symptoms, features, and/or causes of a particular disease, disorder, and/or condition (e.g., neuropathy). Such treatment may be administered to a subject who does not exhibit signs of the relevant disease, disorder and/or condition, and/or to a subject who exhibits only early signs of the disease, disorder, and/or condition. Alternatively or additionally, such treatment may be administered to a subject who exhibits one or more established signs of the relevant disease, disorder and/or condition. In some embodiments, treatment may be administered to a subject who has been diagnosed as suffering from the relevant disease, disorder, and/or condition. In some embodiments, treatment may be administered to a subject known to have one or more susceptibility factors that are statistically correlated with increased risk of development of the relevant disease, disorder, and/or condition.

[0090] Vector: As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments may be ligated. Another type of vector is a viral vector, wherein additional DNA segments may be ligated to a viral genome or portion thereof. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication, episomal mammalian vectors, herpes simplex virus (HSV) vectors, etc.). Other vectors (e.g., non-episomal mammalian vectors) can be integrated into the genome of a host cell upon introduction of the vectors into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as "expression vectors."

[0091] Standard techniques may be used for recombinant DNA, oligonucleotide synthesis, and tissue culture and transformation (e.g., electroporation, lipofection). Enzymatic reactions and purification techniques may be performed according to manufacturer's specifications or as commonly accomplished in the art or as described herein. The foregoing techniques and procedures may be generally performed according to conventional methods well known in the art and as described in various general and more specific references that are cited and discussed throughout the present specification. See, e.g., Sambrook, et al., Molecular Cloning: A Laboratory Manual (2d ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)), which is incorporated herein by reference for any purpose.

DETAILED DESCRIPTION

[0092] Various aspects of the invention are described in detail in the following sections. The use of sections is not meant to limit the invention. Each section can apply to any aspect of the invention. In this application, the use of "or" means "and/or" unless stated otherwise or clear from context to be disjunctive.

[0093] The present disclosure provides, among other things, compositions comprising HSV vectors and methods for use and production of the same. In some embodiments, HSV vectors are provided for the delivery of payloads to fibroblasts, keratinocytes, adipocytes, muscle cells, or any other tissues that comprise the epidermis, dermis, or underlying subcutaneous fat or muscle, and that may be accessible by administration of the vector into or onto the skin. In some embodiments, McKrae strain HSV vectors are provided. Certain HSV (e.g., McKrae strain) vectors are also useful for expression of one or more polypeptides in neurons, genetic modification of cells generically through the use of homologous recombination alone or in combination with DNA modifying enzymes, and oncologic indications where the ability to specifically express several gene products under one or many promoter sequences may be advantageous. The genetic modifications to the vector result in a non-toxic, high capacity (>18 kilobase capacity) vector with the ability for true tissue specific and inducible capabilities.

Viral Vectors and HSV

[0094] Viral vectors can be used to facilitate the transfer of nucleic acids into cells. Known viral vectors include those derived from retroviruses, adenoviruses, adeno-associated virus (AAV), vaccinia virus, and baculovirus. Vectors derived from herpes simplex viruses (HSV), such as herpes simplex virus 1 (HSV-1) and herpes simplex virus-2 (HSV-2) are particularly useful for delivery of agents to specifically targeted tissues. Considerations for choosing a particular vector and delivery system include, for example, characteristics of target cells, desired longevity of expression, virulence and invasiveness of the vector, and size of the genetic material to be transferred.

[0095] HSV-1 vectors can typically accommodate up to 25 kb of foreign DNA sequences. HSV-1 has an approximately 152-kb double-stranded linear DNA genome that can be maintained episomally in the nucleus of a cell. The HSV-1 virion is enveloped and is approximately 110 nm in diameter. Viral infection is initiated in epithelial cells of the skin or mucosal membranes by binding of the viral envelope glycoproteins to heparin sulfate moieties on the plasma membrane. HSV is particularly well suited for the delivery of genes to the nervous system and possesses a natural tropism for sensory neurons. The virus can establish a latent state in which viral genomes persist for the life of the host as an intranuclear episomal element. The life-long persistence of latent genomes in human trigeminal ganglia without the development of sensory loss or histologic damage to the ganglia exemplifies the effectiveness of the latency mechanisms. Wild-type HSV virus may be reactivated from latency under the influence of a variety of stresses. However, recombinant viral vectors that are rendered replication defective retain the ability to establish a persistent quiescent state, for example, in neurons, yet are unable to replicate (or reactivate) in the nervous system.

[0096] Vectors based upon HSV-1 may have one or more HSV genes necessary for replication rendered nonfunctional (e.g., by deletion or disruption). HSV genes necessary for replication can include, for example, immediate early genes such as ICP0, ICP4 and/or ICP27. In some embodiments, the disclosure provides replication defective HSV vectors with

one or more of ICP0, ICP4, ICP22, ICP27, and/or ICP47 deleted or disrupted. In some embodiments, the disclosure provides HSV vectors with nonfunctional ICP0 and ICP4 genes (e.g., ICP0−/ICP4−). In some embodiments, the disclosure provides replication defective HSV vectors with ICP0 and ICP4 deleted or disrupted (e.g., ICP0−/ICP4−). In some embodiments, the ICP0 and/or ICP4 gene is disrupted by deletion of its respective promoter.

[0097] In some embodiments, the disclosure provides HSV vectors with nonfunctional ICP0, ICP4, ICP22, and ICP47 genes. In some embodiments, the disclosure provides an HSV vector with ICP0 and ICP4 deleted, and ICP22 and ICP47 disrupted. In some embodiments, the disclosure provides an HSV vector with ICP0 and ICP4 deleted, and expression of ICP22 and ICP47 disrupted or delayed. In some embodiments, the disclosure provides an HSV vector with ICP0 and ICP4 deleted, and ICP22, ICP27, and/or ICP47 not expressed as immediate early proteins. In some embodiments, the disclosure provides an HSV vector with ICP0 and ICP4 deleted, and with ICP22, ICP27, and ICP47 expressed using undisrupted promoter sequences, thereby providing a unique advantage with respect to proper vector functioning in light of innate immunity as well as evasion of cellular and humoral immunity.

[0098] HSV-1 vectors that have deleted HSV genes can be produced in cell lines that express the deficient protein in trans. In some embodiments, HSV-1 vectors are produced in a mammalian cell line. In some embodiments, the cell line expresses ICP0 and/or ICP4. In some embodiments, the cell line expresses one or more of ICP0, ICP4, ICP22, ICP27, and/or ICP47. In some embodiments, the cells complement ICP0 and/or ICP4, and at least one other viral gene. In some embodiments, the cell line expresses ICP0 and/or ICP4, and at least one additional immediate early gene. In some embodiments, the cell line expresses ICP0 and/or ICP4, and also expresses other viral or non-viral genes that modify innate immune responses. In some embodiments, the cell line expresses ICP0 and/or ICP4, and has been modified to remove Type I or Type III interferon response pathways. In some embodiments, the cell line expresses ICP0 and/or ICP4, and has been modified to remove genes associated with the stimulator of interferon genes (STING) pathway.

[0099] In some embodiments, the cell line expresses ICP0, ICP4, ICP22, and ICP47. In some embodiments, the cell line expresses ICP0, ICP4, ICP22, and UL55. In some embodiments, the cell line expresses ICP0, ICP4, ICP27, and UL55. In some embodiments, the cell line expresses ICP0, ICP4, ICP27, and UL55. In some embodiments the cell line expresses ICP0 and ICP4, as well as ICP22, ICP27, or ICP27 and ICP27. In some embodiments, the cell line comprises a nucleic acid molecule having a simian virus 40 polyadenylation signal (SV40 pA).

[0100] In some embodiments, HSV-1 vectors are produced in a mammalian cell line of Vero lineage. In some embodiments, viral vectors are produced in Vero D cells. In some embodiments, viral vectors are produced in Vero 6-5C cells. In some embodiments viral vectors are produced in Vero 7240 cells. Cells (e.g., cells of a Vero lineage) may be complementing cells, for example ICP0 and/or ICP4 complementing cells. In some embodiments, the cells are ICP0 and/or ICP4 complementing cells. In some embodiments, the cells complement ICP0 and/or ICP4, and at least one other viral gene. In some embodiments, the cells complement ICP0 and/or ICP4, and at least one immediate early gene. In some embodiments, the cells are ICP0, ICP4, ICP27, and UL55 complementing cells. In some embodiments, the cells are ICP0, ICP4, ICP22, and ICP47 complementing cells.

[0101] In some embodiments, HSV-1 vectors are produced in 1-D4 cells. In some embodiments, HSV-1 vectors are produced in 5B4 cells.

HSV-1 Strains

[0102] At least 17 strains of HSV-1 have been isolated, including, but not limited to, McKrae, strain 17, strain F, H129, HF10, MacIntyre, Strain HF, ATCC 2011 and KOS (for review, see Watson et al., *Virology* (2012)). In some embodiments, the disclosure provides HSV viral vectors with deletion of genes that render HSV replication defective, but do not reduce invasiveness. In some embodiments, an HSV vector is an HSV-1 vector, and may comprise any one of: McKrae, strain 17, strain F, H129, HF10, MacIntyre, Strain HF, ATCC 2011, and KOS, or any other HSV-1 strain known in the art.

[0103] The various strains of HSV-1 exhibit certain differences. For example, inter-strain differences in HSV-1 peripheral replication and virulence are observed after injection into animals. HSV genes also influence viral characteristics and phenotype, and each strain may exhibit unique genes and/or non-coding sequences. Therefore, it will be appreciated that the HSV-1 strain which is utilized according to the methods of the present invention may influence the expression characteristics of the mutant vector, for example by localizing the expression of a payload of interest to specific cell or tissue types. Selection of a particular strain may depend on the payload of interest, targeted cell or tissue types, subject identity, etc.

McKrae Strain

[0104] In some embodiments, an HSV vector is an HSV-1 vector comprising a variant of a McKrae strain. A McKrae strain was isolated from a patient with herpes simplex keratitis and subsequently passaged in tissue culture. McKrae undergoes spontaneous or induced reactivation at a higher frequency than other known strains, and is among the most virulent HSV-1 strains. McKrae is also more neuroinvasive than other known strains, such as strain 17, KOS, F, and H129. A partial genome sequence of McKrae is shown in FIG. **9** (SEQ ID NO: 1) (accession number JQ730035).

[0105] There are at least 9 genes and several non-coding sequences unique to McKrae strain. For example, in McKrae, RL1 (ICP34.5) has an extended P-A-T repeat between residues 159 and 160 that results in 8 iterations, while other strains contain only 3-5 iterations. In addition to those associated with pathogenesis and latency reactivations, such as RL1, RS1, and RL2, three UL genes (UL36, UL49A, UL56) and three US genes (US7, US10, and US11) are unique for McKrae strain. In addition to gene variations, non-coding sequences such as LAT, 'a' sequence, and miRNAs contain variations unique to McKrae.

[0106] McKrae strain also contains an extended repeat element of six iterations of the internal tandem repeat STPSTTT (SEQ ID NO: 11) located within the coding sequence of US07 (gI). Additionally, in McKrae, UL36 contains a premature stop codon introduced due to a G nucleotide deletion in a mononucleotide string encoding amino acid residue 2453 (nt 72,535), and UL 56 (180 aa)

contains a single base pair insertion at nucleotide 115,992 (amino acid 97). McKrae strain also contains an extended ORF in US10 resulting from a single bp insertion at nucleotide 143,416, and the frameshift causes a stop codon loss in McKrae and a unique C-terminal protein sequence. McKrae has amino acid differences at UL49A at residues 28 and 51 compared to other strains. McKrae has histidine and threonine at residues 28 and 51, respectively, whereas strain 17 has arginine and threonine and other strains (e.g., KOS) have histidine and alanine. Also, McKrae strain contains reduced tandem repeats found at the UL-RL junction (49 bp in McKrae, as opposed to 181 bp in strain 17 and KOS) and approximately 330 nucleotides missing immediately following the UL-RL junction repeat. McKrae also contains unique variations within the 'a' sequence direct repeat 2 (DR2) array. Instead of a series of unbroken tandem repeats, the McKrae DR2 repeats are interrupted twice by identical guanine-rich sequences.

[0107] Major variation within the LAT intron between strains is due to differences in a repeat element (GCACCCC-CACTCCCAC) (SEQ ID NO: 12) that varies in iteration number beginning at nucleotide 119,482 in McKrae strain. McKrae contains 13 repeats, while strains F, H129 and 17 contain 9 repeats, and KOS contains 15 repeats. Also, tandem repeat variation between strains is found beginning in McKrae at base 125,520. McKrae repeat elements include twelve iterations of CCCCAGCCCTCCCCAG (SEQ ID NO: 13) and eight iterations of CCCCTCGCCCCCTCCCG (SEQ ID NO: 14). The first repeat unit is unique from other strains in that it contains a G-A transition, and McKrae contains three iterations more than any other strain. The McKrae strain second repeat element is collapsed, missing 188 nucleotides relative to all other strains, and separated from the upstream repeat by a 100% conserved sequence of 105 bp containing miR-H5.

[0108] McKrae further contains a unique coding sequence for ICP4 that is not found in other known strains (see Watson et al., Virology (2012)). ICP4 is an immediate early transcriptional regulator and has been implicated in reactivation. Whereas other strains contain an alanine rich region (AASAPDAADALAAA) (SEQ ID NO: 15) between residues 707 and 720, in McKrae the alanine rich region is replaced by a serine rich sequence (GPRRSSSSSGVAA) (SEQ ID NO: 16). The serine rich block of substitutions present in McKrae is adjacent to the nuclear localization signal (NLS) (amino acid 728-734). A change in conformation of this region may alter the NLS and in turn affect localization of not only ICP4, but also other viral proteins (e.g., ICP0, ICP8) that are affected by ICP4 localization (Knipe and Smith, 1986). Thus, this region may influence viral phenotype in part by altering the localization of proteins to the nucleus.

[0109] In some aspects, the HSV vectors of the present invention comprise a variant herpes simplex virus (HSV) McKrae strain genome which contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins. In some embodiments, the alteration is a deletion or disruption of the ICP0 and ICP4 genes. In some embodiments, the ICP0 and/or ICP4 gene is disrupted by deletion of its respective promoter.

[0110] In some embodiments, the ICP0 protein expressed from the wild type McKrae strain, and which is not expressed from the variants produced according to the methods described herein, is encoded by a cDNA sequence comprising nucleotides 2171-2227, 2977-3642, and 3767-5370 of a polynucleotide sequence comprising:

(SEQ ID NO: 18)

ATGGAGCCCCGCCCCGGAGCGAGTACCCGCCGGCCTGAGGGCCGC

CCCCAGCGCGAGGIGAGGGGCCGGGCGCCATGTCTGGGGCGCCAT

ATTGGGGGGCGCCATGTTGGGGGACCCCCGACCCTTACCCTGGAA

CCGGCCCCCATGTTGGGGGACCCCCACTCATACACGGGAGCCGGG

CGCCATGTTGGGGCGCCATGTTAGGGGGCGTGGAACCCCGTGACA

CTATATATACAGGGACCGGGGGCGCCATGTTAGGGGGCGCGGAAC

CCCCTGACCCTATATATACAGGGACCGGGGTCGCCCTGTTGGGGG

TCGCCATGTGACCCCCTGACTTTATATATACAGACCCCCAACACA

TACACATGGCCCCTTTGACTCAGACGCAGGGCCCGGGGTCGCCGT

GGGACCCCCTGACTCATACACAGAGACACGCCCCCACAACAAACA

CACAGGGACCGGGGTCGCCGTGTTGGGGGCGTGGTCCCCACTGAC

TCATACGCAGGCCCCCCTTACTCACACGCATCTAGGGGGGTGGGG

AGGAGCCGCCCGCCATATTTGGGGGACGCCGTGGGACCCCCGACT

CCGGTGCGTCTGGAGGGCGGGAGAAGAGGGAAGAAGAGGGGTCGG

GATCCAAAGGACGGACCCAGACCACCTTTGGTTGCAGACCCCTTT

CTCCCCCCTCTTCCGAGGCCAGCAGGGGGGCAGGACTTTGTGAGG

CGGGGGGGGAGAGGGGGAACTCGTGGGCGCTGATTGACGCGGGAA

ATCCCCCCCCCATTCTTACCCGCCCCCCTTTTTTCCCCTTAGCCCG

CCCCGGATGTCTGGGTGTTTCCCTGCGACCGAGACCTGCCGGACA

GCAGCGACTCTGAGGCGGAGACCGAAGTGGGGGGGCGGGGGGACG

CCGACCACCATGACGACGACTCCGCCTCCGAGGCGGACAGCACGG

ACACGGAACTGTTCGAGACGGGGCTGCTGGGGCCGCAGGGCGTGG

ATGGGGGGGCGGTCTCGGGGGGGAGCCCCCCCCGCGAGGAAGACC

CCGGCAGTTGCGGGGGCGCCCCCCCTCGAGAGGACGGGGGGAGCG

ACGAGGGCGACGTGTGCGCCGTGTGCACGGATGAGATCGCGCCCC

ACCTGCGCTGCGACACCTTCCCGTGCATGCACCGCTTCTGCATCC

CGTGCATGAAAACCTGGATGCAATTGCGCAACACCTGCCCGCTGT

GCAACGCCAAGCTGGTGTACCTGATAGTGGGCGTGACGCCCAGCG

GGTCGTTCAGCACCATCCCGATCGTGAACGACCCCCAGACCCGCA

TGGAGGCCGAGGAGGCCGTCAGGGCGGGCACGGCCGTGGACTTTA

TCTGGACGGGCAATCAGCGGTTCGCCCCGCGGTACCTGACCCTGG

GGGGGCACACGGTGAGGGCCCTGTCGCCCACCCACCCGGAGCCCA

CCACGGACGAGGATGACGACGACCTGGACGACGGTGAGGCGGGGG

GCGGCAAGGACCCTGGGGGAGGAGGAGGAGGGAGGAATGGGCGGG

CGGGCGAGGAAAGGGCGGGCCGGGGAGGGGGCGTAACCTGATCGC

GCCCCCCGTTGTCTCTTGCAGCAGACTACGTCCCGCCCGCCCCCC

GCCGGACGCCCCGCGCCCCCCCACGCAGAGGCACCGCCGCGCCCC

-continued

CCGTGACGGGCGGGGCGTCTAACGCAGCCCCCCAGCCGGCCGCGG

CTCGGACAGCGCCCCCCTCGGCGCCCATCGGGCCACACGGCAGCA

GTAACACCAACACCACCACCAACAGCAGCGGCGGCGGCGGCTCCC

GCCAGTCGCGAGCCGCGGCGCCGCGGGGGGCGTCTGGCCCCTCCG

GGGGGGTTGGGGTTGGGGTTGGGGTTGTTGAAGCGGAGGCGGGGC

GGCCGAGGGGCCGGACGGGCCCCCTTGTCAACAGACCCGCCCCCC

TTGCAAACAACAGAGACCCCATAGTGATCAGCGACTCCCCCCCGG

CCTCTCCCCACAGGCCCCCCGCGGCGCCCATGCCAGGCTCCGCCC

CCCGCCCCGGGCCCCCCGCGTCCTCGGCCGCGTCGGGACCCGCGC

GCCCCCGCGCGGCCGTGGCCCCGTGCGTGCGAGCGCCGCCTCCGG

GGCCCGGCCCCCGCGCCCCGGCCCCCGGGGCGGAGCCGGCCGCCC

GCCCCGCGGACGCGCGCCGTGTGCCCCAGTCGCACTCGTCCCTGG

CTCAGGCCGCGAACCAAGAACAGAGTCTGTGCCGGGCGCGTGCGA

CGGTGGCGCGCGGCTCGGGGGGGCCGGGCGTGGAGGGTGGGCACG

GGCCCTCCCGCGGCCGCACCCCCTCCGGCGCCGCCCCGCTCCCCT

CCGCCGTCTCTGTCGAGCAGGAGGCGGCGGTGCGTCCGAGGAAGA

GGCGCGGGTCGGGCCAGGAAAACCCCTCCCCCCAGTCCACGCGTC

CCCCCCTCGCGCCGGCAGGGGCCAAGAGGGCGGCGACGCACCCCC

CCTCCGACTCAGGGCCGGGGGGGCGCGGCCAGGGTGGGCCCGGGA

CCCCCCTGACGTCCTCGGCGGCCTCCGCCTCTTCCTCCTCTGCCT

CTTCCTCCTCGGCCCCGACCCCCGCGGGGGCCGCCTCTTCCGCCG

CCGGGGCCGCGTCCTCCTCCGCTTCCGCCTCCTCGGGGGGGGCCG

TCGGTGCCCTGGGAGGGAGACAAGAGGAAACCTCCCTCGGCCCCC

GCGCTGCTTCTGGGCCGCGGGGGCCGAGGAAGTGTGCCCGGAAGA

CGCGCCACGCGGAGACTTCCGGGGCCGTCCCCGCGGGCGGCCTCA

CGCGCTACCTGCCCATCTCGGGGGTCTCTAGCGTGGTCGCCCTGT

CGCCTTACGTGAACAAGACTATCACGGGGGACTGCCTGCCCATCC

TGGACATGGAGACGGGGAACATCGGGGCGTACGTGGTCCTGGTGG

ACCAGACGGGAAACATGGCGACCCGGCTGCGGGCCGCGGTCCCCG

GCTGGAGCCGCCGCACCCTGCTCCCCGAGACCGCGGGTAACCACG

TGATGCCCCCCGAGTACCCGACGGCCCCCGCGTCGGAGTGGAACA

GCCTCTGGATGACCCCCGTGGGGAACATGCTGTTCGACCAGGGCA

CCCTAGTGGGCGCCCTGGACTTCCGCAGCCTGCGGTCTCGGCACC

CGTGGTCCGGGGAGCAGGGGGCGTCGACCCGGGACGAGGGAAAAC

AATAA.

[0111] In some embodiments, the ICP0 protein expressed from the wild type McKrae strain, and which is not expressed from the variants produced according to the methods described herein, comprises:

(SEQ ID NO: 17)

MEPRPGASTRRPEGRPQREPAPDVWVFPCDRDLPDSSDSEAETEV

GGRGDADHHDDDSASEADSTDTELFETGLLGPQGVDGGAVSGGSP

PREEDPGSCGGAPPREDGGSDEGDVCAVCTDEIAPHLRCDTFPCM

HRFCIPCMKTWMQLRNTCPLCNAKLVYLIVGVTPSGSFSTIPIVN

DPQTRMEAEEAVRAGTAVDFIWTGNQRFAPRYLTLGGHTVRALSP

THPEPTIDEDDDDLDDADYVPPAPRRTPRAPPRRGTAAPPVIGGA

SNAAPQPAAARTAPPSAPIGPHGSSNINTTINSSGGGGSRQSRAA

APRGASGPSGGVGVGVGVVEAEAGRPRGRTGPLVNRPAPLANNRD

PIVISDSPPASPHRPPAAPMPGSAPRPGPPASSAASGPARPRAAV

APCVRAPPPGPGPRAPAPGAEPAARPADARRVPQSHSSLAQAANQ

EQSLCRARATVARGSGGPGVEGGHGPSRGRTPSGAAPLPSAVSVE

QEAAVRPRKRRGSGQENPSPQSTRPPLAPAGAKRAATHPPSDSGP

GGRGQGGPGTPLTSSAASASSSSASSSSAPTPAGAASSAAGAASS

SASASSGGAVGALGGRQEETSLGPRAASGPRGPRKCARKTRHAET

SGAVPAGGLTRYLPISGVSSVVALSPYVNKTITGDCLPILDMETG

NIGAYVVLVDQTGNMATRLRAAVPGWSRRILLPETAGNHVMPPEY

PTAPASEWNSLWMTPVGNMLFDQGTLVGALDFRSLRSRHPWSGEQ

GASTRDEGKQ.

[0112] In some embodiments, the ICP4 protein expressed from the wild type McKrae strain, and which is not expressed from the variants produced according to the methods described herein, is encoded by a polynucleotide sequence comprising:

(SEQ ID NO: 25)

ATGGCGTCGGAGAACAAGCAGCGCCCCGGCTCCCCGGGCCCCACC

GACGGGCCGCCGCCCACCCCGAGCCCAGACCGCGACGAGCGGGGG

GCCCTCGGGTGGGGCGCGGAGACGGAGGAGGGCGGGGACGACCCC

GACCACGACCCCGACCACCCCCACGACCTCGACGACGCCCGGCGG

GACGGGAGGGCCCCCGCGGCGGGCACCGACGCCGGCGAGGACGCC

GGGGACGCCGTCTCGCCGCGACAGCTGGCCCTGCTGGCCTCCATG

GTAGAGGAGGCCGTCCGGACGATCCCGACGCCCGACCCCGCGGCC

TCGCCGCCCCGGACCCCCGCCTTTCGAGCCGACGACGATGACGGG

GACGAGTACGACGACGCAGCCGACGCCGCCGGCGACCGGGCCCCG

GCCCGGGGCCGCGCACGGGAGGCCCCGCTACGCGGCGCGTATCCG

GACCCCACGGACCGCCTGTCGCCGCGCCCGCCGGCCCAGCCGCCG

CAGAGACGTCGTCACGGCCGGCGGCGGCCATCGGCGTCATCGACC

TCGTCGGACTCCGGGTCCTCGTCCTCGTCGTCCGCATCCTCTTCG

TCCTCGTCGTCCGACGAGGACGAGGACGACGACGGCAACGACGCG

GCCGACCACGCACGCGAGGCGCGGGCCGTCGGGGGGGGTCCGTCG

AGCGCGGCGCCGGAAGCCCCCGGGCGGACGCCGCCCCCGCCCGGG

-continued

CCACCCCCCCTCTCCGAGGCCGCGCCCAAGCCCCGGGCGGCGGCG

AGGACCCCCGCGGCCTCCGCGGGCCGCATCGAGCGCCGCCGGGCC

CGCGCGGCGGTGGCCGGCCGCGACGCCACGGGCCGCTTCACGGCC

GGGCAGCCCCGGCGGGTCGAGCTGGACGCCGACGCGGCCTCCGGC

GCCTTCTACGCGCGCTATCGCGACGGGTACGTCAGCGGGGAGCCG

TGGCCCGGCGCCGGGCCCCCGCCCCCGGGGGGGGTGCTGTACGGC

GGCCTGGGCGACAGCCGCCCGGGCCTCTGGGGGGCGCCCGAGGCG

GAGGAGGCGCGACGCCGGTTCGAGGCCTCGGGCGCCCCGGCGGCC

GTGTGGGCGCCCGAGCTGGGCGACGCCGCGCAGCAGTACGCCCTG

ATCACGCGGCTGCTGTACACCCCGGACGCGGAGGCCATGGGGTGG

CTCCAGAACCCGCGCGTGGTCCCCGGGGACGTGGCGCTGGACCAG

GCCTGCTTCCGGATCTCGGGCGCCGCGCGCAACAGCAGCTCCTTC

ATCACCGGCAGCGTGGCGCGGGCCGTGCCCCACCTGGGCTACGCC

ATGGCGGCCGGCCGCTTCGGCTGGGGCCTGGCGCACGCGGCGGCC

GCCGTGGCCATGAGCCGCCGATACGACCGCGCGCAGAAGGGCTTC

CTGCTGACCAGCCTGCGCCGCGCCTACGCGCCCCTGTTGGCGCGC

GAGAACGCGGCGCTGACGGGGGCCGCGGGGAGCCCCGGCGCCGGC

GCAGATGACGAGGGGGTCGCCGCCGCCGTCGTCGCCGCCGCCGCC

GCACCGGGCGAGCGCGCGGTGCCCGCCGGGTACGGCGCCGCGGGG

ATCCTCGCCGCCCTGGGGCGGCTGTCCGCCGCGCCCGCCTCCCCC

GCGGGGGGCGACGACCCCGACGCCGCCCGCCACGCCGACGCCGAC

GACGACGCCGGGCGCCGCGCCCAGGCCGGCCGCGTGGCCGTGGAG

TGCCTGGCCGCCTGCCGCGGGATCCTGGAGGCGCTGGCCGAGGGC

TTCGACGGCGACCTGGCGGCCGTCCCGGGGCTGGCCGGGGCCCGG

CCCGCCAGCCCCCCGCGGCCGGAGGGACCCGCGGGCCCCGCTTCC

CCGCCGCCGCCGCACGCCGACGCGCCCCGCCTGCGCGCGTGGCTG

CGCGAGCTGCGGTTCGTGCGCGACGCGCTGGTGCTCATGCGCCTG

CGCGGGGACCTGCGCGTGGCCGGCGGCAGCGAGGCCGCCGTGGCC

GCCGTGCGCGCCGTGAGCCTGGTCGCCGGGGCCCTGGGTCCCGCG

CTGCCGCGGGACCCGCGCCTGCCGAGCTCCGCGGCCGCCGCCGCC

GCGGACCTGCTGTTTGAGAACCAGAGCCTGCGCCCCCTGCTGGCG

GCGGGTCCGCGCCGCTCTTCTTCGTCTTCGGGGGTCGCGGCCGCC

GCCTCCGCCGCGCCGCGGGAGGGGCGCAAGCGCAAGAGTCCCGGC

CCGGCCCGGCCGCCCGGAGGCGGCGGCCCGCGACCCCCGAAGACG

AAGAAGAGCGGCGCGGACGCCCCCGGCTCGGACGCCCGCGCCCCC

CTCCCCGCGCCCGCGCCCCCCTCCACGCCCCCGGGGCCCGAGCCC

GCCCCCGCCCAGCCCGCGGCGCCCCGGGCCGCCGCGGCGCAGGCC

CGCCCGCGCCCCGTGGCGCTGTCGCGCCGGCCCGCCGAGGGCCCC

GACCCCCTGGGCGGCTGGCGGCGGCAGCCCCCGGGGCCCAGCCAC

-continued

ACGGCGGCGCCCGCGGCCGCCGCCCTGGAGGCCTACTGCTCCCCG

CGCGCCGTGGCCGAGCTCACGGACCACCCGCTGTTCCCCGTCCCC

TGGCGACCGGCCCTCATGTTTGACCCGCGGGCCCTGGCCTCGATC

GCCGCGCGGTGCGCCGGGCCCGCCCCCGCCGCCCAGGCCGCGTGC

GGCGGCGGCGACGACGACGAGAACCCCCACCCCCACGGGGCCGCC

GGGGGCCGCCTCTTTGGCCCCCTGCGCGCCTCGGGCCCGCTGCGC

CGCATGGCGGCCTGGATGCGCCAGATCCCCGACCCCGAGGACGTG

CGCGTGGTGGTGCTGTACTCGCCGCTGCCGGGCGAGGACCTGGCC

GGCGGCGGGGCCTCGGGGGGGCCGCCGGAGTGGTCCGCCGAGCGC

GGCGGGCTGTCCTGCCTGCTGGCGGCCCTGGCCAACCGGCTGTGC

GGGCCGGACACGGCCGCCTGGGCGGGCAACTGGACCGGCGCCCCC

GACGTGTCGGCGCTGGGCGCGCAGGGCGTGCTGCTGCTGTCCACG

CGGGACCTGGCCTTCGCCGGGGCCGTGGAGTTTCTGGGGCTGCTC

GCCAGCGCCGGCGACCGGCGGCTCATCGTGGTCAACACCGTGCGC

GCCTGCGACTGGCCCGCCGACGGGCCCGCGGTGTCGCGGCAGCAC

GCCTACCTGGCGTGCGACCTGCTGCCCGCCGTGCAGTGCGCCGTG

CGCTGGCCGGCGGCGCGCGACCTGCGCCGCACGGTGCTGGCCTCG

GGCCGCGTGTTCGGCCCGGGGGTCTTCGCGCGCGTGGAGGCCGCG

CACGCGCGCCTGTACCCCGACGCGCCGCCGCTGCGCCTGTGCCGC

GGCGGCAACGTGCGCTACCGCGTGCGCACGCGCTTCGGCCCGGAC

ACGCCGGTGCCCATGTCCCCGCGCGAGTACCGCCGGGCCGTGCTG

CCGGCGCTGGACGGCCGGGCGGCGGCCTCGGGGACCACCGACGCC

ATGGCGCCCGGCGCGCCGGACTTCTGCGAGGAGGAGGCCCACTCG

CACCGCGCCTGCGCGCGCTGGGGCCTGGGCGCGCCGCTGCGGCCC

GTGTACGTGGCGCTGGGGCGCGAGGCGGTGCGCGCCGGCCCGGCC

CGGTGGCGCGGGCCGCGGAGGGACTTTTGCGCCCGCGCCCTGCTG

GAGCCCGACGACGACGCCCCCCCGCTGGTGCTGCGCGGCGACGAC

GACGGCCCGGGGGCCCTGCCGCCGGCGCCGCCCGGGATTCGCTGG

GCCTCGGCCACGGGCCGCAGCGGCACCGTGCTGGCGGCGGCGGGG

GCCGTGGAGGTGCTGGGGGCGGAGGCGGGCTTGGCCACGCCCCCG

CGACGGGACGTTGTGGACTGGGAAGGCGCCTGGGACGAAGACGAC

GGCGGCGCGTTCGAGGGGGACGGGGTGCTGTAA.

[0113] In some embodiments, the ICP4 protein expressed from the wild type McKrae strain, and which is not expressed from the variants produced according to the methods described herein, is encoded by a polynucleotide sequence comprising:

(SEQ ID NO: 5)

TTTATTGCGTCTTCGGGTTTCACAAGCGCCCCGCCCCGTCCCGGC

CCGTTACAGCACCCCGTCCCCCTCGAACGCGCCGCCGTCGTCTTC

-continued

GTCCCAGGCGCCTTCCCAGTCCACAACGTCCCGTCGCGGGGGCGT

GGCCAAGCCCGCCTCCGCCCCCAGCACCTCCACGGCCCCCGCCGC

CGCCAGCACGGTGCCGCTGCGGCCCGTGGCCGAGGCCCAGCGAAT

CCCGGGCGGCGCCGGCGGCAGGGCCCCCGGGCCGTCGTCGTCGCC

GCGCAGCACCAGCGGGGGGGCGTCGTCGTCGGGCTCCAGCAGGGC

GCGGGCGCAAAAGTCCCTCCGCGGCCCGCGCCACCGGGCCGGGCC

GGCGCGCACCGCCTCGCGCCCCAGCGCCACGTACACGGGCCGCAG

CGGCGCGCCCAGGCCCCAGCGCGCGCAGGCGCGGTGCGAGTGGGC

CTCCTCCTCGCAGAAGTCCGGCGCGCCGGGCGCCATGGCGTCGGT

GGTCCCCGAGGCCGCCGCCCGGCCGTCCAGCGCCGGCAGCACGGC

CCGGCGGTACTCGCGCGGGGACATGGGCACCGGCGTGTCCGGGCC

GAAGCGCGTGCGCACGCGGTAGCGCACGTTGCCGCCGCGGCACAG

GCGCAGCGGCGGCGCGTCGGGGTACAGGCGCGCGTGCGCGGCCTC

CACGCGCGCGAAGACCCCCGGGCCGAACACGCGGCCCGAGGCCAG

CACCGTGCGGCGCAGGTCGCGCGCCGCCGGCCAGCGCACGGCGCA

CTGCACGGCGGGCAGCAGGTCGCACGCCAGGTAGGCGTGCTGCCG

CGACACCGCGGGCCCGTCGGCGGGCCAGTCGCAGGCGCGCACGGT

GTTGACCACGATGAGCCGCCGGTCGCCGGCGCTGGCGAGCAGCCC

CAGAAACTCCACGGCCCCGGCGAAGGCCAGGTCCCGCGTGGACAG

CAGCAGCACGCCCTGCGCGCCCAGCGCCGACACGTCGGGGGCGCC

GGTCCAGTTGCCCGCCCAGGCGGCCGTGTCCGGCCCGCACAGCCG

GTTGGCCAGGGCCGCCAGCAGGCAGGACAGCCCGCCGCGCTCGGC

GGACCACTCCGGCGGCCCCCCCGAGGCCCCGCCGCCGGCCAGGTC

CTCGCCCGGCAGCGGCGAGTACAGCACCACCACGCGCACGTCCTC

GGGGTCGGGGATCTGGCGCATCCAGGCCGCCATGCGGCGCAGCGG

GCCCGAGGCGCGCAGGGGGCCAAAGAGGCGGCCCCCGGCGGCCCC

GTGGGGGTGGGGGTTCTCGTCGTCGTCGCCGCCGCCGCACGCGGC

CTGGGCGGCGGGGGCGGGCCCGGCGCACCGCGCGGCGATCGAGGC

CAGGGCCCGCGGGTCAAACATGAGGGCCGGTCGCCAGGGGACGGG

GAACAGCGGGTGGTCCGTGAGCTCGGCCACGGCGCGCGGGGAGCA

GTAGGCCTCCAGGGCGGCGGCCGCGGGCGCCGCCGTGTGGCTGGG

CCCCGGGGGCTGCCGCCGCCAGCCGCCCAGGGGGTCGGGGCCCTC

GGCGGGCCGGCGCGACAGCGCCACGGGGCGCGGGCGGGCCTGCGC

CGCGGCGGCCCGGGGCGCCGCGGGCTGGGCGGGGGCGGGCTCGGG

CCCCGGGGGCGTGGAGGGGGGCGCGGGCGCGGGGAGGGGGGCGCG

GGCGTCCGAGCCGGGGGCGTCCGCGCCGCTCTTCTTCGTCTTCGG

GGGTCGCGGGCCGCCGCCTCCGGGCGGCCGGGCCGGGCCGGGACT

CTTGCGCTTGCGCCCCTCCCGCGGCGCGGCGGAGGCGGCGGCCGC

GACCCCCGAAGACGAAGAAGAGCGGCGCGGACCCGCCGCCAGCAG

-continued

GGGGCGCAGGCTCTGGTTCTCAAACAGCAGGTCCGCGGCGGCGGC

GGCCGCGGAGCTCGGCAGGCGCGGGTCCCGCGGCAGCGCGGGACC

CAGGGCCCCGGCGACCAGGCTCACGGCGCGCACGGCGGCCACGGC

GGCCTCGCTGCCGCCGGCCACGCGCAGGTCCCCGCGCAGGCGCAT

GAGCACCAGCGCGTCGCGCACGAACCGCAGCTCGCGCAGCCACGC

GCGCAGGCGGGGCGCGTCGGCGTGCGGCGGCGGCGGGGAAGCGGG

GCCCGCGGGTCCCTCCGGCCGCGGGGGGCTGGCGGGCCGGGCCCC

GGCCAGCCCCGGGACGGCCGCCAGGTCGCCGTCGAAGCCCTCGGC

CAGCGCCTCCAGGATCCCGCGGCAGGCGGCCAGGCACTCCACGGC

CACGCGGCCGGCCTGGGCGCGGCGCCCGGCGTCGTCGTCGGCGTC

GGCGTGGCGGGCGGCGTCGGGGTCGTCGCCCCCCGCGGGGGAGGC

GGGCGCGGCGGACAGCCGCCCCAGGGCGGCGAGGATCCCCGCGGC

GCCGTACCCGGCGGGCACCGCGCGCTCGCCCGGTGCGGCGGCGGC

GGCGACGACGGCGGCGGCGACCCCCTCGTCATCTGCGCCGGCGCC

GGGGCTCCCCGCGGCCCCCGTCAGCGCCGCGTTCTCGCGCGCCAA

CAGGGGCGCGTAGGCGCGGCGCAGGCTGGTCAGCAGGAAGCCCTT

CTGCGCGCGGTCGTATCGGCGGCTCATGGCCACGGCGGCCGCCGC

GTGCGCCAGGCCCCAGCCGAAGCGGCCGGCCGCCATGGCGTAGCC

CAGGTGGGGCACGGCCCGCGCCACGCTGCCGGTGATGAAGGAGCT

GCTGTTGCGCGCGGCGCCCGAGATCCGGAAGCAGGCCTGGTCCAG

CGCCACGTCCCCGGGGACCACGCGCGGGTTCTGGAGCCACCCCAT

GGCCTCCGCGTCCGGGGTGTACAGCAGCCGCGTGATCAGGGCGTA

CTGCTGCGCGGCGTCGCCCAGCTCGGGCGCCCACACGGCCGCCGG

GGCGCCCGAGGCCTCGAACCGGCGTCGCGCCTCCTCCGCCTCGGG

CGCCCCCCAGAGGCCCGGGCGGCTGTCGCCCAGGCCGCCGTACAG

CACCCGCCCCGGGGGCGGGGGCCCGGCGCCGGGCCACGGCTCCCC

GCTGACGTACCCGTCGCGATAGCGCGCGTAGAAGGCGCCGGAGGC

CGCGTCGGCGTCCAGCTCGACCCGCCGGGGCTGCCCGGCCGTGAA

GCGGCCCGTGGCGTCGCGGCCGGCCACCGCCGCGCGGGCCCGGCG

GCGCTCGATGCGGCCCGCGGAGGCCGCGGGGGTCCTCGCCGCCGC

CCGGGGCTTGGGCGCGGCCTCGGAGAGGGGGGGTGGCCCGGGCGG

GGGCGGCGTCCGCCCGGGGGCTTCCGGCGCCGCGCTCGACGGACC

CCGCCCGACGGCCCGCGCCTCGCGTGCGTGGTCGGCCGCGTCGTT

GCCGTCGTCGTCCTCGTCCTCGTCGGACGACGAGGACGAAGAGGA

TGCGGACGACGAGGACGAGGACCCGGAGTCCGACGAGGTCGATGA

CGCCGATGGCCGCCGCCGGCCGTGACGACGTCTCTGCGGCGGCTG

GGCCGGCGGGCGCGGCGACAGGCGGTCCGTGGGGTCCGGATACGC

GCCGCGTAGCGGGGCCTCCCGTGCGCGGCCCCGGGCCGGGGCCCG

GTCGCCGGCGGCGTCGGCTGCGTCGTCGTACTCGTCCCCGTCATC

-continued

GTCGTCGGCTCGAAAGGCGGGGGTCCGGGGCGGCGAGGCCGCGGG

GTCGGGCGTCGGGATCGTCCGGACGGCCTCCTCTACCATGGAGGC

CAGCAGGGCCAGCTGTCGCGGCGAGACGGCGTCCCCGGCGTCCTC

GCCGGCGTCGGTGCCCGCCGCGGGGGCCCTCCCGTCCCGCCGGGC

GTCGTCGAGGICGTGGGGGTGGTCGGGGTCGTGGTCGGGGTCGTC

CCCGCCCTCCTCCGTCTCCGCGCCCCACCCGAGGGCCCCCCGCTC

GTCGCGGTCTGGGCTCGGGGTGGGCGGCGGCCCGTCGGTGGGGCC

CGGGGAGCCGGGGCGCTGCTTGTTCTCCGACGCCATCGCCGATGC

GGGGCGATCCTCCGGGGATACGACTGCGACGGCGGACGTAGCACG

GTAGGTCACCTACGG.

[0114] In some embodiments, the ICP4 protein expressed from the wild type McKrae strain, and which is not expressed from the variants produced according to the methods described herein, comprises:

(SEQ ID NO: 2)

MASENKQRPGSPGPTDGPPPTPSPDRDERGALGWGAETEEGGDDP

DHDPDHPHDLDDARRDGRAPAAGTDAGEDAGDAVSPRQLALLASM

VEEAVRTIPTPDPAASPPRTPAFRADDDDGDEYDDAADAAGDRAP

ARGRAREAPLRGAYPDPTDRLSPRPPAQPPQRRRHGRRRPSASST

SSDSGSSSSSSASSSSSSSDEDEDDDGNDAADHAREARAVGRGPS

SAAPEAPGRIPPPPGPPPLSEAAPKPRAAARTPAASAGRIERRRA

RAAVAGRDATGRFTAGQPRRVELDADAASGAFYARYRDGYVSGEP

WPGAGPPPPGRVLYGGLGDSRPGLWGAPEAEEARRRFEASGAPAA

VWAPELGDAAQQYALITRLLYTPDAEAMGWLQNPRVVPGDVALDQ

ACFRISGAARNSSSFITGSVARAVPHLGYAMAAGREGWGLAHAAA

AVAMSRRYDRAQKGFLLTSLRRAYAPLLARENAALTGAAGSPGAG

ADDEGVAAAVVAAAAAPGERAVPAGYGAAGILAALGRLSAAPASP

AGGDDPDAARHADADDDAGRRAQAGRVAVECLAACRGILEALAEG

FDGDLAAVPGLAGARPASPPRPEGPAGPASPPPPHADAPRLRAWL

RELRFVRDALVLMRLRGDLRVAGGSEAAVAAVRAVSLVAGALGPA

LPRDPRLPSSAAAAAADLLFENQSLRPLLAAGPRRSSSSSGVAAA

ASAAPREGRKRKSPGPARPPGGGGPRPPKTKKSGADAPGSDARAP

LPAPAPPSTPPGPEPAPAQPAAPRAAAAQARPRPVALSRRPAEGP

DPLGGWRRQPPGPSHTAAPAAAALEAYCSPRAVAELTDHPLFPVP

WRPALMFDPRALASIAARCAGPAPAAQAACGGGDDDENPHPHGAA

GGRLFGPLRASGPLRRMAAWMRQIPDPEDVRVVVLYSPLPGEDLA

GGGASGGPPEWSAERGGLSCLLAALANRLCGPDTAAWAGNWTGAP

DVSALGAQGVLLLSTRDLAFAGAVEFLGLLASAGDRRLIVVNTVR

ACDWPADGPAVSRQHAYLACDLLPAVQCAVRWPAARDLRRTVLAS

-continued

GRVFGPGVFARVEAAHARLYPDAPPLRLCRGGNVRYRVRTRFGPD

TPVPMSPREYRRAVLPALDGRAAASGTTDAMAPGAPDFCEEEAHS

HRACARWGLGAPLRPVYVALGREAVRAGPARWRGPRRDFCARALL

EPDDDAPPLVLRGDDDGPGALPPAPPGIRWASATGRSGTVLAAAG

AVEVLGAEAGLATPPRRDVVDWEGAWDEDDGGAFEGDGVL.

Replication Defective McKrae Vector

McKrae Backbone

[0115] Viral genes are expressed in a tightly regulated, ordered cascade, which begins with the production of the immediate-early (IE) genes. The resulting IE proteins, which include infected cell proteins ICP0, ICP4, ICP22, ICP27, and ICP47, are responsible for regulating viral gene expression during subsequent phases of the replication cycle. Replication-defective variant viruses are defective for one or more functions that are essential for viral genome replication or synthesis and assembly of viral particles. Such viruses can be propagated in complementing cell lines expressing the missing gene product(s); however, in normal (e.g., non-complementing) cells, the viruses express viral gene products, but do not replicate to form progeny virions.

[0116] Replication-defective viruses can be created through various methods known in the art for modifying genes. In some embodiments, one or more nucleotides are rendered different relative to the wild-type sequence. In some embodiments, one or more nucleotides are deleted. In some embodiments, the deletion of one or more nucleotides creates a premature stop codon. In some embodiments, the deletion of one or more nucleotides creates a gene encoding a truncated polypeptide. In some embodiments, the deletion of one or more nucleotides creates a gene encoding a nonfunctional polypeptide. In some embodiments, the deletion of one or more nucleotides renders a gene nonfunctional by disruption. In some embodiments, a gene is disrupted by deletion of its promoter.

[0117] In some embodiments, one or more genes are deleted to render a virus replication defective. In some embodiments, the genes encoding ICP0 and ICP4 are each independently fully or partially deleted. In some embodiments, the genes encoding ICP0 and ICP4 are each partially deleted. In some embodiments, the genes encoding ICP0 and ICP4 are each fully deleted. In some embodiments, the gene encoding ICP0 is partially deleted and the gene encoding ICP4 is fully deleted. In some embodiments, the gene encoding ICP0 is fully deleted and the gene encoding ICP4 is partially deleted. In some embodiments, the genes encoding ICP0 and ICP4 are deleted, and the upstream promoter sequences of ICP22 and ICP47 are modified. In some embodiments, the upstream promoter sequences of ICP22 and ICP47 are modified to change the timing of gene expression and/or amount of genes expressed. In some embodiments, the genes encoding ICP0 and ICP4 are deleted, and gene expression of ICP22, ICP27, and/or ICP47 is modified. In some embodiments, the deletion of the genes encoding ICP0 and ICP4 leads to the modification of ICP22, ICP27, and/or ICP47 expression. In some embodiments, the modification of ICP22, ICP27, and/or ICP47 expression

improves gene expression, reduces toxicity, and/or reduces immune responses (e.g., innate, cell mediated, or humoral immune responses).

[0118] HSV-1 IE promoters contain one or more copies of an IE-specific regulatory sequence of consensus TAATGARAT (SEQ ID NO: 24) (where R is a purine). These motifs are normally located within a few hundred base pairs of the proximal IE promoter sequences, but in conjunction with their flanking sequences they are discrete functional entities which can confer IE-specific regulation to other proximal promoter elements of different temporal classes. In some embodiments, replication-defective viruses are created by deleting nucleotides in an IE-specific regulatory sequence. In some embodiments, an IE-specific regulatory sequence contains an internal deletion. In some embodiments, an IE-specific regulatory sequence contains a terminal deletion. In some embodiments, an IE-specific regulatory sequence is completely deleted.

[0119] A schematic of an exemplary replication defective McKrae strain viral vector is depicted in FIGS. **32**A-**32**C. FIG. **32**A shows complete deletions of both copies of the viral ICP4 genes, and a human cytomegalovirus (HCMV) immediate early promoter driven expression cassette inserted within both copies of the deleted ICP4 loci. FIG. **32**B shows complete deletions of both copies of the viral ICP4 genes, and an extended Collagen-1 (COL1) promoter driven expression cassette inserted within both copies of the deleted ICP4 loci. FIG. **32**C shows complete deletions of both copies of the viral ICP0 and ICP4 genes, and an extended Collagen-1 (COL1) promoter driven expression cassette inserted within both copies of the deleted ICP4 loci. Each expression cassette in FIGS. **32**A-**32**C contains a payload of interest for expression in target cells.

[0120] In some embodiments, the extent of the ICP0 and/or ICP4 deletion results in the modification (e.g., removal) of the upstream promoter sequences of two additional immediate early viral genes: ICP22 and ICP47. In some embodiments, such a modification may modify the timing or level of expression of the gene products, but not necessarily the gene function.

Payload

[0121] Viral vectors in accordance with the present disclosure may contain a nucleic acid molecule comprising the payload of the vector. In some embodiments, a payload comprises a nucleic acid molecule that encodes a protein. In some embodiments, a payload comprises a nucleic acid molecule that comprises a sequence complementary to a nucleic acid sequence that encodes a protein. In some embodiments, a payload encodes a nucleic acid molecule that is regulatory. In some embodiments, a payload encodes a small interfering RNA (siRNA) polynucleotide. In some embodiments, a payload encodes a micro RNA (miRNA) polynucleotide.

[0122] In some embodiments, the payload is a nucleic acid molecule that encodes a protein that is exogenous to the target tissue or subject to which the vector is administered. In some embodiments, the payload is a nucleic acid molecule that encodes a protein that is endogenous to the target tissue or subject to which the vector is administered. In some embodiments, a nucleic acid molecule is codon optimized.

Regulatory Elements

[0123] The inclusion of non-native regulatory sequences, gene control sequences, promoters, non-coding sequences, introns, or coding sequences in a nucleic acid of the present disclosure is contemplated herein. The inclusion of nucleic acid tags or signaling sequences, or nucleic acids encoding protein tags or protein signaling sequences, is further contemplated herein. Typically, the coding region is operably linked with one or more regulatory nucleic acid components.

[0124] A promoter included in a nucleic acid of the present disclosure can be a tissue- or cell type-specific promoter, a promoter specific to multiple tissues or cell types, an organ-specific promoter, a promoter specific to multiple organs, a systemic or ubiquitous promoter, or a nearly systemic or ubiquitous promoter. Promoters having stochastic expression, inducible expression, conditional expression, or otherwise discontinuous, inconstant, or unpredictable expression are also included within the scope of the present disclosure. A promoter of the present disclosure may include any of the above characteristics or other promoter characteristics known in the art.

[0125] Examples of known promoters include, but are not limited to, the cytomegalovirus (CMV) promoter CMV/human beta 3 globin promoter, glial fibrillary acidic protein (GFAP) promoter, chicken beta actin (CBA) promoter, β-glucuronidase (GUSB) promoter, collagen 1 promoter (eCOL1), and ubiquitin promoters such as those isolated from human ubiquitin A, human ubiquitin B, and human ubiquitin C.

[0126] In some embodiments, a promoter is a cell type-specific promoter. In some embodiments, a promoter is a neuron-specific promoter. In some embodiments, a promoter is a neuron-specific promoter in that it is a promoter having specific expression in neurons, preferential expression in neurons, or that typically drives expression of an associated coding sequence in neurons or a subset of neurons but not in one or more other tissues or cell types. Examples of such promoters include calcitonin gene-related peptide (CGRP), synapsin I (SYN), calcium/calmodulin-dependent protein kinase II, tubulin alpha I, neuron-specific enolase, microtubule-associated protein 1B (MAP1B), and platelet-derived growth factor beta chain promoters, as well as derivatives thereof. In some embodiments, the promoter is a calcitonin gene-related peptide (CGRP) promoter or derivative thereof.

[0127] In some embodiments, a promoter is a chimeric of one or more promoters or regulatory elements found in nature. In some embodiments, the viral vectors comprise a payload whose expression is driven by a calcitonin gene-related peptide (CGRP) promoter with an HCMV enhancer sequence.

[0128] In some embodiments, a promoter includes genomic sequences. In some embodiments, a promoter sequences contains more than two kilobases of genomic sequences. In some embodiments, a promoter contains more than five kilobases of genomic sequences. In some embodiments, a promoter includes genomic sequences comprising the human collagen-1 promoter. In some embodiments, a promoter expresses genes in a fashion similar to certain tissues in the body, instead of uncontrolled gene expression that is seen with small segments of human genomic promoters or viral promoters.

[0129] In some embodiments, a promoter is a tissue specific promoter (e.g., is a promoter that has activity in only certain cell types). In some embodiments, a promoter is tissue specific for fibroblasts. In some embodiments, a tissue specific promoter for fibroblasts is a collagen-1 (eCOL1) promoter. In some embodiments, a promoter is tissue spe-

cific for fibroblasts, and is inducible by a small molecule and/or peptides. In some embodiments, a promoter is tissue specific for cell lines of different lineages that are found in skin or in the underlying subcutaneous tissue, including muscle.

[0130] Other regulatory elements may additionally be operatively linked to the payload, such as an enhancer and a polyadenylation site. In some embodiments, an enhancer comprises a human cytomegalovirus (HCMV) sequence. In some embodiments, a polyadenylation site comprises a bovine growth hormone (BGH) polyadenylation signal.

Preparation of Vectors

[0131] The present disclosure relates particularly to HSV vectors (e.g., McKrae strain vectors) that are replication defective. In some embodiments, viral vectors are generated by deletion or disruption of one or more immediate early genes. Viral genes may be deleted or disrupted using methods of recombinant technology known in the art. In some embodiments, a viral vector of the present disclosure may be rendered replication defective as a result of a homologous recombination event. In some embodiments, replication defective viral vectors are generated by deletion or disruption of an ICP4 gene. In some embodiments, replication defective vectors are generated by deletion or disruption of the ICP0 and ICP4 genes. In some embodiments, the ICP0 and/or ICP4 gene is disrupted by deletion of its respective promoter. In some embodiments, replication defective viral vectors are generated by deletion of ICP0 and ICP4 genes and deletion of a promoter for one or more other immediate early genes (e.g., ICP22 and/or ICP47).

[0132] In some embodiments, viral vectors of the present disclosure are generated by deletion of loci encoding one or more ICPs (e.g., ICP0 and/or ICP4) through homologous recombination. In some embodiments, generation of a viral vector of the present disclosure includes a step of homologous recombination. In some embodiments, homologous recombination comprises recombination of a first plasmid with a second plasmid. In some embodiments, the first plasmid contains nucleic acid sequences homologous to regions of an HSV genome that are adjacent to a nucleic acid region of an HSV genome that is intended to be replaced. In some embodiments, the first plasmid contains a nucleic acid sequence encoding a gene of interest between the homologous nucleic acid sequences. In some embodiments, the gene of interest comprises a marker protein that is detectable by fluorescence, chemiluminescence, or any other detectable property, which can in some embodiments identify vectors resulting from successful homologous recombination. In some embodiments, the second plasmid contains an HSV genome, or fragment thereof.

[0133] In some embodiments, homologous recombination comprises recombination between a plasmid and a viral genome (e.g., an HSV genome). In some embodiments, a viral vector of the present disclosure is generated by homologous recombination of a first plasmid with a circular episome or linear configuration of an HSV McKrae strain genome. In some embodiments, the first plasmid comprises a nucleic acid sequence homologous to regions upstream of the ICP4 promoter, including the viral origin contained within the short inverted repeat regions of HSV.

[0134] In some embodiments, a vector is made by first replacing both copies of the ICP4 loci by homologous recombination using plasmid SASB3 and screening for green fluorescent protein (GFP)-expressing plaques. In some embodiments, a plasmid is constructed by cloning the Sph I to Afl III (Sal I linkered) fragment (1928 bp) of the HSV-1 KOS strain genome (nucleotides 124485-126413) into Sph I/Sal I digested pSP72 followed by insertion of the 695 bp Bgl II to BamH I fragment (nucleotides 131931 to 132626) containing regions upstream of the ICP4 promoter, including the viral origin contained within the short inverted repeat regions into the Bgl II to BamH I sites of the vector plasmid. In some embodiments, a plasmid is constructed by cloning a HCMV-eGFP fragment in the BamHI site of a plasmid as described above. In some embodiments, a plasmid as described above is then recombined into a specific locus of a wild-type McKrae virus. In some embodiments, the resulting vector is isolated using a stable cell line that expresses one or more genes deleted or disrupted in the HSV genome that are required for replication.

[0135] In some embodiments, a vector is made by first replacing both copies of the ICP4 loci by homologous recombination using plasmid SDAXB and screening for green fluorescent protein (GFP)-expressing plaques. In some embodiments, a plasmid is constructed by cloning the Sph I to Afl III fragment (1928 bp) of the HSV-1 KOS strain genome (nucleotides 124346 to 126273 of accession KT899744) into Sph I/Afl III digested pSP72 to make SDA followed by changing the Afl III site to a BamHI site (SDAB). A BamHI to Bgl II DNA PCR fragment containing regions upstream of the ICP4 promoter including the viral origin (nucleotides 144933 to 145534 of accession JQ730035) contained within the short inverted repeat regions was cloned into the BamHI site of SDAB to make SDAXB. In some embodiments, a plasmid is constructed by cloning a HCMV-eGFP fragment in the BamHI site of a plasmid as described above. In some embodiments, a plasmid as described above is then recombined into a specific locus of a wild-type McKrae virus. In some embodiments, the resulting vector is isolated using a stable cell line that expresses one or more genes deleted or disrupted in the HSV genome that are required for replication.

[0136] In some embodiments, a vector is made by first replacing both copies of the ICP0 loci by homologous recombination using plasmid OE2 and screening for green fluorescent protein (GFP)-expressing plaques. In some embodiments, a plasmid containing the sequences flanking the ICP0 gene base pairs 1201 to 2165 and 6021 to 6971 of JQ730035.1 presentation of a McKrae strain genome. In some embodiments, a plasmid as described above is then recombined into a specific locus of a wild-type McKrae virus. In some embodiments, the resulting vector is isolated using a stable cell line that expresses one or more genes deleted or disrupted in the HSV genome that are required for replication.

[0137] In some embodiments, a viral vector of the present disclosure is generated by homologous recombination of a first plasmid containing a nucleic acid sequence homologous to regions upstream of the ICP0 gene sequence, which may in some embodiments contain a promoter element, with a circular episome or linear configuration of an HSV McKrae strain genome that is deleted for ICP4 (e.g., as described above) in order to create a viral vector that is deleted for both ICP0 and ICP4. In some embodiments, a viral vector of the present disclosure is generated by homologous recombination of a first plasmid containing a nucleic acid sequence homologous to regions upstream of the ICP4 promoter

including the viral origin contained within the long inverted repeat regions of HSV, with a circular episome or linear configuration of an HSV McKrae strain genome that is deleted for ICP0 (e.g., as described above) in order to create a viral vector that is deleted for both ICP0 and ICP4.

Characterization of Vectors

[0138] Viral vectors in accordance with the present disclosure can be characterized by genomic sequencing in order to determine if the expected vector was successfully created. Any method of sequencing known in the art is acceptable for this purpose. Methods of sequencing include, for example, nanopore sequencing, single molecule real time sequencing (SMRT), DNA nanoball (DNB) sequencing, pyrosequencing and using DNA arrays.

[0139] The expression of a payload from a viral vector can be detected by any method known in the art for detecting proteins or nucleic acids. Methods of detecting protein expression include immunohistochemistry, flow cytometry, Western blotting, enzyme-linked immunosorbent assay (ELISA), immune-electron microscopy, individual protein immunoprecipitation (IP), protein complex immunoprecipitation (Co-IP), chromatin immunoprecipitation (ChIP), RNA immunoprecipitation (RIP), immunoelectrophoresis, spectrophotometry, and bicinchoninic acid assay (BCA). Methods of detecting nucleic acid expression include Southern blotting, Northern blotting, polymerase chain reaction (PCR), quantitative PCR, and RT-PCR.

[0140] In some embodiments, the present disclosure provides methods for testing the ability of viral vectors to transduce cells of a specific type (e.g., neurons, keratinocytes, muscle cells, etc.). In some embodiments, the present disclosure provides methods for testing the ability of viral vectors to transduce neurons. In some embodiments, the neurons are peripheral neurons. In some embodiments, the neurons are sensory neurons. In some embodiments, the neurons comprise dorsal root ganglia (DRG). In some embodiments, a viral vector preparation may be injected into the one or more dermatomes corresponding to a section of DRG (e.g., the left and right L4, L5, and L6 DRG). In some embodiments, DRG are removed, and DNA is isolated from the DRG and analyzed for vector genome copies using a qPCR assay that targets a sequence within HSV-1. In some embodiments, a qPCR assay targets a sequence within the HSV-1 glycoprotein (UL-22) gene.

[0141] In some embodiments, the present disclosure provides for testing the ability of viral vectors to transduce keratinocytes, fibroblasts, adipocytes, muscle cells, and all other cells that are capable of being transduced via administration to skin (e.g., by topical administration or injection). In some embodiments, a vector preparation may be administered to skin via topical application. In some embodiments, a vector preparation may be administered to skin via an injection. In some embodiments, injection may comprise an intradermal or subcutaneous method of injection. In some embodiments, vector expression in skin may be measured by ELISA, histology, or analysis for vector genomic DNA or RNA.

[0142] In some embodiments, the present disclosure provides for testing the ability of viral vectors to transduce tumor cells.

[0143] In some embodiments, the present disclosure provides for the testing the ability of viral vectors to localize to other organs, including the liver, lung, heart, kidney, bladder, and prostate.

Applications/Uses

[0144] Viral vectors in accordance with the present disclosure are useful for a wide variety of therapeutic applications. In some embodiments, vectors as described herein are useful to deliver one or more payloads to one or more target cells. In some embodiments, target cells reside in tissues that are poorly vascularized and difficult to reach by systemic circulation. In some embodiments, target cells are cells susceptible to infection by HSV. In some embodiments, target cells are particularly susceptible to infection by a McKrae strain of HSV. In some embodiments, target cells are neuronal cells. In some embodiments, target cells are dorsal root ganglion (DRG) cells.

[0145] In some embodiments, target cells are skin cells. In some embodiments, target cells are keratinocytes, fibroblasts, adipocytes, and/or muscle cells. In some embodiments, target cells are all other cells that present in the epidermis, dermis, and subcutaneous tissue that may be reached by topical (e.g., applied to the skin) or intradermal administration of an HSV vector of the present disclosure.

[0146] In some embodiments, target cells are tumor cells. In some embodiments, target cells reside in the liver, lung, heart, kidney, bladder, or prostate.

Gene Therapy

[0147] Viral vectors in accordance with the present disclosure are useful in any context in which gene therapy is contemplated. For example, viral vectors comprising a heterologous nucleic acid segment operably linked to a promoter are useful for any disease or clinical condition associated with reduction or absence of the protein encoded by the heterologous nucleic acid segment, or any disease or clinical condition that can be effectively treated by expression of the encoded protein within the subject.

[0148] Viral vectors that contain an expression cassette for synthesis of an RNAi agent (e.g., one or more siRNAs or shRNAs) are useful in treating any disease or clinical condition associated with overexpression of a transcript or its encoded protein in a subject, or any disease or clinical condition that may be treated by causing reduction of a transcript or its encoded protein in a subject.

[0149] Viral vectors that comprise an expression cassette for synthesis of one or more RNAs that self-hybridize or hybridize with each other to form an RNAi agent targeted to a transcript encoding a cytokine may be used to regulate immune system responses (e.g., responses responsible for organ transplant rejection, allergy, autoimmune diseases, inflammation, etc.).

[0150] Viral vectors that provide a template for synthesis of one or more RNAs that self-hybridize or hybridize with each other to form an RNAi agent targeted to a transcript of an infectious agent or targeted to a cellular transcript whose encoded product is necessary for or contributes to any aspect of the infectious process may be used in the treatment of infectious diseases.

[0151] Viral vectors that contain DNA modifying enzyme genomic DNA sequences may be used in the treatment of genetic disease. Viral vectors that contain genomic DNA

sequences may be used in the treatment of genetic diseases. Viral vectors that contain genetic material from other viruses may be used as a treatment or vaccine.

Administration

[0152] Compositions comprising viral vectors as described herein may be formulated for delivery by any available route including, but not limited to parenteral (e.g., intravenous), intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, rectal, and vaginal. Preferred routes of delivery include intradermal and transdermal. In some embodiments, pharmaceutical compositions include a viral vector in combination with a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" includes solvents, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. Supplementary active compounds can also be incorporated into the compositions. In some embodiments, viral vectors are formulated in glycerol. In some embodiments, viral vectors are formulated in approximately 10% glycerol in phosphate buffered saline (PBS).

[0153] It is advantageous to formulate compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of a viral vector calculated to produce the desired therapeutic effect in association with a pharmaceutical carrier.

[0154] The pharmaceutical composition can be administered at various intervals and over different periods of time as required, e.g., one time per week for between about 1 to 10 weeks, between about 2 to 8 weeks, between about 3 to 7 weeks, between about 4, 5, or 6 weeks, etc. The skilled artisan will appreciate that certain factors can influence the dosage and timing required to effectively treat a subject, including but not limited to the severity of the disease or disorder, previous treatments, the general health and/or age of the subject, and/or other diseases present in the subject. Treatment of a subject with a viral vector can include a single treatment or, in many cases, can include a series of treatments.

Compositions

[0155] In some embodiments, the active agents (e.g., a viral vector of the disclosure and/or other agents to be administered together with a viral vector of the disclosure) are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such compositions will be apparent to those skilled in the art. In some embodiments the composition is targeted to particular cell types or to cells that are infected by a virus.

Combination Therapy

[0156] According to the present disclosure, provided compositions may be administered in combination with one or more other active agents and/or therapeutic modalities, such as known therapeutic agents and/or independently active biologically active agents. In some embodiments, provided compositions include one or more such other active agents; in some embodiments, such other active agents are provided as part of distinct compositions. In some embodiments, combination therapy involves simultaneous administration of one or more doses or units of two or more different active agents and/or therapeutic modalities; in some embodiments, combination therapy involves simultaneous exposure to two or more different active agents and/or therapeutic modalities, for example through overlapping dosing regimens.

[0157] In some embodiments, provided compositions include or are administered in combination with one or more other active agents useful for the treatment of the relevant disease, disorder and/or condition.

Enumerated Embodiments

[0158] Certain embodiments are set forth in the enumerated clauses below.

[0159] Clause 1. A variant of a herpes simplex virus (HSV) strain whose genome contains an alteration such that the variant fails to express one or more functional ICP0 and ICP4 proteins.

[0160] Clause 2. The variant HSV strain of clause 1, wherein the HSV strain is an HSV-1 strain.

[0161] Clause 3. The variant HSV strain of clause 1, wherein the HSV strain is a McKrae strain.

[0162] Clause 4. A variant HSV strain comprising a variant herpes simplex virus (HSV) strain genome which contains an alteration such that the variant fails to express functional ICP4 and ICP0 proteins characterized by the amino acid sequences of SEQ ID NO: 2 and SEQ ID NO: 17, respectively.

[0163] Clause 5. The variant HSV strain of any one of clauses 1-4, wherein the HSV strain comprises a deletion of substantially all or all of the nucleic acids encoding the ICP4 and/or ICP0 proteins.

[0164] Clause 6. The variant HSV strain of any one of clauses 1-4, wherein the HSV strain contains an alteration such that the variant fails to express a functional ICP47 protein, e.g., where the HSV strain comprises a deletion of substantially all or all of the nucleic acids encoding the ICP47 protein.

[0165] Clause 7. The variant HSV strain of any one of clauses 1-5, wherein the HSV strain expresses one or more functional ICP22, ICP27, and/or ICP47 proteins, for example, wherein the ICP22, ICP27, and/or ICP47 proteins are substantially wild-type or wild-type proteins.

[0166] Clause 8. A vector comprising the variant HSV strain of any of clauses 1-7.

[0167] Clause 9. The vector of clause 8, wherein the vector comprises a cell or tissue specific promoter.

[0168] Clause 10. The vector of clause 9, wherein the tissue is skin.

[0169] Clause 11. The vector of clause 9, wherein the tissue specific promoter is a collagen 1 promoter.

[0170] Clause 12. The vector of any of clauses 8-11, wherein the vector comprises a human cytomegalovirus (HCMV) enhancer.

[0171] Clause 13. The vector of any of clauses 8-12, wherein the vector comprises a bovine growth hormone (BGH) polyadenylation signal or an HSV viral polyadenylation signal.

[0172] Clause 14. The vector of any of clauses 8-13, further comprising a nucleic acid that encodes one or more therapeutic transgenes.

[0173] Clause 15. The vector of any of clauses 8-14, where the nucleic acid that encodes one or more therapeutic transgenes is at least 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, 10 kb, 15 kb, 20 kb, 25 kb, 30 kb, 35 kb, from about 1 kb to about 35 kb, from about 1 kb to about 35 kb, from about 1 kb to about 30 kb, from about 1 kb to about 25 kb, from about 1 kb to about 20 kb, from about 1 kb to about 15 kb, from about 1 kb to about 10 kb, from about 1 kb to about 5 kb, from about 5 kb to about 35 kb, from about 5 kb to about 30 kb, from about 5 kb to about 25 kb, from about 5 kb to about 20 kb, from about 5 kb to about 15 kb, from about 5 kb to about 10 kb, from about 10 kb to about 35 kb, from about 10 kb to about 30 kb, from about 10 kb to about 25 kb, from about 10 kb to about 20 kb, from about 10 kb to about 15 kb, from about 15 kb to about 35 kb, from about 15 kb to about 30 kb, from about 15 kb to about 25 kb, from about 15 kb to about 20 kb, from about 20 kb to about 35 kb, from about 20 kb to about 30 kb, from about 20 kb to about 25 kb, from about 25 kb to about 35 kb, from about 25 kb to about 30 kb, or from about 30 kb to about 25 kb.

[0174] Clause 16. The vector of any of clauses 8-15, wherein the one or more therapeutic transgenes encode one or more therapeutic polypeptides.

[0175] Clause 17. The vector of any of clauses 8-16, wherein the tissue specific promoter is operably linked to the one more therapeutic transgenes.

[0176] Clause 18. The vector of any of clauses 8-17, wherein the vector is capable of expressing the one more therapeutic transgenes in a target cell, body fluid, tissue, organ, or physiological system of a subject.

[0177] Clause 19. The vector of any of clauses 8-18, wherein the vector is capable of expressing the one more therapeutic transgenes in the target cell, body fluid, tissue, organ, or physiological system of the subject for at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 20, 22, 24, 26, 28, or 30 days.

[0178] Clause 20. The vector of any of clauses 8-19, wherein the vector is non-toxic to a cell, body fluid, tissue, organ, or physiological system of a subject.

[0179] Clause 21. The vector of any of clauses 8-20, wherein the vector is less toxic to a cell, body fluid, tissue, organ, or physiological system of a subject relative to an HSV vector that does express one or more functional ICP0 and ICP4 proteins.

[0180] Clause 22. A cell transduced with a vector according to any one of clauses 8-21.

[0181] Clause 23. A pharmaceutical composition comprising a variant of a herpes simplex virus (HSV) strain according to any one of clauses 1-7, or a vector according to any one of clauses 8-22 and a pharmaceutically acceptable carrier.

[0182] Clause 24. A method of propagating a vector comprising a variant herpes simplex virus (HSV) genome which contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins, the method comprising steps of:

[0183] (a) infecting cultured ICP0 and ICP4 complementing cells containing a nucleic acid (e.g., DNA) encoding HSV proteins ICP0 and ICP4 with the vector, and

[0184] (b) isolating supernatant from the culture of step (a).

[0185] Clause 25. The method of clause 24, further comprising a step of purifying vector in the supernatant by chromatography.

[0186] Clause 26. The method of clause 24 or clause 25, further comprising a step of concentrating the purified vector by tangential flow filtration.

[0187] Clause 27. A method of preparing a vector comprising a variant herpes simplex virus (HSV) genome which contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins, and wherein the vector expresses a marker element, the method comprising incubating cells transfected with:

[0188] (a) a first nucleic acid molecule:

[0189] (i) comprising a portion of HSV genome but does not encode functional ICP0 and ICP4 proteins; and

[0190] (ii) comprising a first homology region (HR1) and a second homology region (HR2), and

[0191] (b) a second nucleic acid molecule comprising a sequence that encodes a marker element, wherein the sequence is flanked by a first homology region (HR1') and a second homology region (HR2'),

[0192] wherein HR1 is homologous to HR1' and HR2 is homologous to HR2' such that the sequence that encodes the marker element in the second nucleic acid molecule integrates into the first nucleic acid molecule via homologous recombination.

[0193] Clause 28. The method of clause 27, wherein the cells are ICP0 and/or ICP4 complementing cells.

[0194] Clause 29. The method of clause 27 or 28, wherein the marker element is a polypeptide.

[0195] Clause 30. The method of clause 29, wherein the polypeptide is a soluble tumor necrosis factor receptor.

[0196] Clause 31. The method of clause 29, wherein the polypeptide is quantified by enzyme linked immunosorbent assay (ELISA).

[0197] Clause 32. The method of any one of clauses 29-31, wherein the polypeptide is detected by fluorescence.

[0198] Clause 33. The method of any one of clauses 27-32, further comprising a step of purifying viral plaques that express the marker element.

[0199] Clause 34. A method of preparing a vector comprising a variant herpes simplex virus (HSV) genome which contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins, and wherein the vector expresses an agent of interest, the method comprising incubating cells transfected with:

[0200] (a) a first nucleic acid molecule:

[0201] (i) comprising a portion of HSV genome but does not encode functional ICP0 and ICP4 proteins; and

[0202] (ii) comprising a sequence that encodes a marker element, wherein the sequence that encodes the marker element is flanked by a first homology region (HR1) and a second homology region (HR2); and

[0203] (b) a second nucleic acid molecule comprising a sequence that encodes an agent of interest, wherein the sequence encoding the agent of interest is flanked by a first homology region (HR1') and a second homology region (HR2'),

[0204] wherein HR1 is homologous to HR1' and HR2 is homologous to HR2' such the sequence encoding the agent of interest is integrated into the first nucleic acid molecule via homologous recombination.

[0205] Clause 35. The method of clause 34, wherein the cells are ICP0 and/or ICP4 complementing cells.

[0206] Clause 36. The method of clause 34 or clause 35, further comprising a step of purifying viral plaques that do not express the marker element.

[0207] Clause 37. The method of any of clauses 24-36, wherein the HSV genome is an HSV-1 genome.

[0208] Clause 38. The method of any of clauses 24-36, wherein the HSV genome is a McKrae strain genome.

[0209] Clause 39. A method of expressing a transgene (e.g., a polypeptide) in a target cell, body fluid, tissue, organ, or physiological system of a subject comprising administering to the subject a variant of a herpes simplex virus (HSV) strain according to any one of clauses 1-7 or a vector according to any one of clauses 8-21.

[0210] Clause 40. The method of clause 39, wherein the tissue is selected from skin tissue, eye tissue, or tumor tissue.

[0211] Clause 41. The method of clause 39, wherein the tissue is skin tissue.

[0212] Clause 42. The method of clause 39, wherein the tissue comprises epidermis, dermis, or subcutaneous fat or muscle.

[0213] Clause 43. The method of clause 41, wherein the skin tissue comprises fibroblasts, keratinocytes, adipocytes, or muscle cells.

[0214] Clause 44. The method of any of clauses 39-43, wherein the vector is administered in vivo.

[0215] Clause 45. The method of any of clauses 39-43, wherein the vector is administered by contact with skin.

[0216] Clause 46. The method of any of clauses 39-43, wherein the vector is administered by intradermal injection.

[0217] Clause 47. A method of measuring transduction efficiency of an HSV vector in a skin tissue, the method comprising:

[0218] (a) contacting the skin tissue of an animal with a vector according to any one of clauses 8-21;

[0219] (b) removing a portion of the skin tissue from the animal; and

[0220] (c) assaying the number of HSV genomes transduced in the skin tissue.

[0221] Clause 48. The method of clause 47, wherein the skin tissue comprises fibroblasts, keratinocytes, adipocytes, muscle cells, epidermis, dermis, hypodermis, or underlying subcutaneous fat or muscle.

[0222] Clause 49. The method of clause 47 or clause 48, wherein the number of genomes is measured by an amplification technique.

[0223] Clause 50. The method of clause 49, wherein the amplification technique is quantitative polymerase chain reaction (qPCR).

[0224] Clause 51. A method of measuring transduction efficiency of an HSV vector that contains an expression cassette comprising a polypeptide payload in a skin tissue, the method comprising:

[0225] (a) contacting the skin tissue of an animal with a vector according to any one of clauses 8-21;

[0226] (b) removing a portion of the skin tissue from the animal; and

[0227] (c) assaying the amount of a polypeptide encoded by a nucleic acid of the expression cassette.

[0228] Clause 52. The method of clause 51, wherein the skin tissue comprises fibroblasts, keratinocytes, adipocytes, muscle cells, epidermis, dermis, hypodermis, or underlying subcutaneous fat or muscle.

[0229] Clause 53. The method of clause 51 or clause 52, wherein the amount of polypeptide is measured by an immunoassay.

[0230] Clause 54. The method of clause 53, wherein the immunoassay is an enzyme linked immunosorbent assay (ELISA) or immunohistochemistry (IHC).

[0231] Clause 55. The method of clause 53, wherein the ELISA or IHC is performed on tissue of the epidermis, dermis, subcutaneous tissue, subcutaneous fat, underlying muscle, or draining lymph node.

EXAMPLES

Example 1: Western Blot Demonstrating Deletion of ICP0 from an HSV ICP4− Mutant

[0232] This Example shows an exemplary method for testing the expression of ICP0 protein from infected cell cultures using Western blot.

[0233] Viral vector preparations were added to non-complementing Vero cells for one hour, and then whole media was added to cells. After 24 hours, cells were washed with phosphate buffered saline (PBS) and then lysed and collected in RIPA buffer. Thirty (30) μL of RIPA buffer lysate was added to 10 μL 4×LDS sample buffer (Invitrogen) and heated at 95° C. for 5 minutes, before PAGE electrophoresis on a 7% Tris Acetate PAGE gel. The PAGE gel was run at 140V for 90 minutes.

[0234] Next, the protein was transferred to immobilon membrane, and was then incubated in 2% non-fat dry milk in tris buffered saline (TBS). The primary antibody (mouse anti-ICP0) was added at 1:3000 dilution overnight at room temperature. The following day, the membrane was washed three times in TBS, and then incubated in 2% non-fat dry milk in TBS for 15 minutes. A secondary anti-mouse antibody conjugated to alkaline phosphatase was added to the incubation at 1:10000 dilution, and then was incubated at room temperature for 2 hours. The membrane was then washed 3 times in 20 mL TBS over the course of 30 minutes.

[0235] The membrane was developed by the addition NBT/BCIP solution in 0.5M Tris Buffer pH 9.5 for 30 minutes to visualize bands. Upon visualization, none of the McKrae HSV-1 mutants deleted for ICP0 and ICP4 (PGN04 99i8.1, PGN04 99i8.4, PGN04 99i8.5, PGN04 99i8.6, PGN04 99i8.7, and PGN04 99i8.8) expressed ICP0 (FIG. **1**). A control ICP4-mutant is also shown as a control, which does express ICP0. These results demonstrate that the ICP0−/ICP4− mutants developed according to the present disclosure do not express ICP0.

Example 2: PCR Confirmation of Deletion of ICP0 DNA Sequence

[0236] This example demonstrates PCR analysis of viral vector DNA preparations that confirm deletion of ICP0 sequences in the ICP0−/ICP4− mutants developed according to the present disclosure.

[0237] Viral vector preparations were added to non-complementing Vero cells for one hour, and then whole media was added to cells. After 24 hours, vector and genomic DNA were collected using a New England Biolabs

Monarch Genomic DNA Kit as described in the user manual. PCR analysis was performed on each DNA sample using PCR primers specific for ICP0 sequences (as the test sequence) and UL56 primers (to confirm the presence of viral DNA). An ICP4 mutant vector CT1 was included as a positive control for an ICP0 containing virus.

[0238] The PCR results demonstrate that the ICP0−/ICP4− mutants developed according to the present disclosure do not contain the ICP0 DNA sequence (FIG. **2**).

Example 3: Preparation of Vectors

[0239] This example describes methods of preparing and formulating exemplary vectors for gene therapy.

Genetic Structure of Vector

[0240] A vector is made by first replacing both copies of the ICP0 loci by homologous recombination in a McKrae strain ICP4− mutant using a plasmid, and screening for marker element expressing plaques. The plasmid is constructed by cloning a fragment of an HSV-1 genome comprising regions of the ICP0 promoter, as well as a sequence downstream of the 3' end of the ICP0 gene. The plasmid is further modified by cloning a marker element, for example eGFP, fragment into the plasmid. This plasmid is then recombined into the ICP0 locus of a McKrae strain ICP4 mutant HSV virus. The resulting vector is isolated using a stable ICP0 and ICP4 expressing Vero cell line, such as 7240. Vero 7240 cells are complementing cells that express ICP0, ICP4, ICP27, and UL55.

[0241] In order to make a vector capable of expressing a gene of interest, a plasmid is constructed by cloning a promoter and gene sequence into a plasmid that contains ICP0 or ICP4 flanking sequences. Plaques which do not express the marker element are isolated and tested by ELISA for GOI expression. The same method is used to recombine human tissue specific promoters in the ICP4 locus of the ICP0−/ICP4− mutant virus.

[0242] The plasmids used to generate the human genomic promoter sequences are generated by PCR of human genomic DNA that is cloned into the plasmid containing the ICP4 flanking sequences.

Production of Crude Vector

[0243] ICP0/ICP4 or ICP0/ICP4/ICP27/UL55 complementing Vero cells are cultured in tissue culture flasks using complete media (DMEM supplemented with FBS, HEPES, and Pen Strep), and expanded into 6-12× T175 flasks at a seeding density of 3-4× 10^4 cells/cm^2. The culture flasks are incubated at 37° C./7.5% CO2 for 3-4 days.

[0244] When cells are confluent, they are infected at a multiplicity of infection (MOI) of ~0.1 with a virus stock of known concentration. The infection is initiated by removing the culture supernatant from each flask and infecting with a total of 2.5 mL of complete media containing the appropriate amount of a virus stock. The virus is adsorbed on the cell monolayers by incubating the cultures for 1.5-2 hours, and shaking and rotating the flasks every 15-20 minutes. After the adsorption step, an additional 10 mL of complete medium is added to each flask and the cultures are incubated again at 37° C./7.5% CO2.

[0245] Approximately 72 hours after initiating the infection, the flasks are viewed by microscope to confirm that the cells show signs of cytopathogenic effect and detachment from the flask surface. At that point, the cells and supernatant are harvested, pooled together, and centrifuged at ~1500×g for ~10 min. The supernatant is removed from the cell pellet and held separately for later processing.

[0246] The cell pellet is resuspended in 4-5 mL of complete media, homogenized, and then frozen at −80° C. After the cell suspension has been frozen for >20 minutes, it is thawed and centrifuged at ~1500×g for ~10 min. This second cell pellet supernatant is removed and combined with the first collected supernatant.

[0247] The pooled supernatant is aliquoted into centrifuge tubes. The virus is then centrifuged at ~40,000×g for ~30 minutes at 2-8° C. in order to pellet the virus. After the centrifugation step is completed, the supernatant from the tubes is removed and discarded. The following day, the virus pellets are homogenized by pipetting, and pooled together. The resuspended virus stock is then aliquoted into cryovials typically at volumes of ~120 μL per vial. Complete medium (200-300 μL) is added to the virus pellets in order to cover them with liquid, and the cells are stored at 2-8° C. overnight to loosen the virus particles. The vials are then labeled and frozen at −80° C. Later, a frozen vial is thawed in order to perform a virus plaque titration assay to determine the concentration of the prepared virus stock prior to using in any in vivo or in vitro studies.

Manufacture of Clarified Vector

Cell Thaw and Expansion

[0248] Vero cells (e.g., ICP0/ICP4 or ICP0/ICP4/ICP27/UL55 complementing) from a working cell bank are thawed at 37° C., transferred to a conical tube, and pooled. The cells are vialed at approximately 1.0×10^7 viable cells/mL/tube. The cells are gradually diluted with complete medium, and a sample is removed to obtain viable cell counts. The cells are plated in tissue culture flasks at a density of 3.0-5.0× 10^4 cells/cm^2.

[0249] The cells are incubated at 37° C., 7.5% CO2 and examined periodically by phase microscopy. The cells are passaged while subconfluent. The complete medium is removed, rinsed with PBS, and the cells are dissociated. The flasks are incubated until the cells detach, then they are re-suspended in complete medium, pooled, counted, and seeded into new flasks at a density between 1.0-4.0×10^4 cells/cm^2. The cells are expanded and allowed to extend to 1-2 days post-confluence prior to infection.

Infection with Vector

[0250] When the cells reach the desired confluence, a model flask is subcultured and the cells are counted to estimate the number of cells per cell factory. A virus bank vector inoculum is prepared by thawing the appropriate volume required to obtain a multiplicity of infection (MOI) of 0.1 and diluting the stock with complete medium up to the target volume desired for the infection. The cell factories are infected by an initial adsorption period followed by incubation for the first day of infection in complete medium. After approximately 24 hours, the culture medium is removed and replaced with an equal volume of serum-free medium. The cultures are monitored daily and the percent cytopathic effect estimated by visual inspection.

Crude Viral Harvest and Clarification

[0251] The infection is stopped by placing the cell factories in a biosafety cabinet, and pooling the supernatant and

cell debris into a sterile bag. This bulk unclarified harvest is sampled for adventitious agents. After sampling, the sodium chloride level of the harvest is increased, and then it is mixed. The harvest is then aliquoted into centrifuge tubes and the cell debris is removed by centrifugation. The supernatant is pooled into a sterile bag. After pre-treatment of a clarification filter capsule with sterile water, the virus-containing supernatant is then pumped through the filter capsule into another sterile bag, followed by sterile water to recover remaining virus in the capsule. The bag is mixed and the filtrate is stored overnight at 4° C.

[0252] Afterwards, the filtrate is warmed and adjusted to ~2 mM MgCl2 by addition of 2 volumes of 3 mM MgCl2 in sterile water. The diluted filtrate is mixed and treated with an endonuclease.

Cation Exchange Column Chromatography

[0253] A BPG 400 column is packed with SP high performance resin, sanitized with 0.5N NaOH, and equilibrated with wash buffer (PBS, pH 7.0) and strip buffer (1M NaCl-PBS, pH 7.0) before loading endonuclease treated virus.

[0254] The process bag containing the endonuclease-treated filtrate is connected to the inlet using a tubing welder, and the virus is loaded onto the column. The flow through is collected in a sterile bag. The virus capture step is followed by washing with PBS until the UV absorbance returns to baseline. The pump is stopped and a process bag containing 0.45 M NaCl-PBS (pH 7.0) is connected to the inlet. The outlet tubing is transferred to a sterile container in a biosafety cabinet. The buffer is pumped into the column and when the UV absorbance begins to increase sharply, the column outlet is transferred to a new sterile container to collect the eluted virus. The collection is stopped after the UV absorbance returns to near baseline. This is the purified viral elute fraction. A process bag containing strip buffer is connected to the inlet and the end of the outlet tubing is transferred into a sterile bottle to collect the strip fraction. The buffer is pumped through the column until UV absorbance reaches a peak and returns to near baseline. The collected elute is stored at 4° C. overnight.

Tangential Flow Filtration

[0255] The tangential-flow filtration system, using a 0.1 micrometer pore size hollow fiber filter cartridge, is prepared by assembling the tubing and cartridge and sterilizing the system by autoclaving. The system is flushed with sterile PBS (pH 7.0) and the virus eluate fraction is added to the system reservoir and equilibrated by recirculation. After equilibration, the permeate collection pump is turned on and filtrate is collected. The system is run until the loaded volume is reduced to approximately 500 mL. The retentate in the reservoir is diluted with DPBS (pH 7.0) with continuous constant volume diafiltration, and the product in the retentate is recovered when the permeate conductivity is within 10% of the diafiltering buffer (DPBS pH 7.0).

Formulation, Final Filtration and Packaging

[0256] The recovered retentate is adjusted to 10% final volume with sterile glycerol and mixed well prior to filtering through a 0.45 μm disc filter unit. The product is dispensed into labeled cryovials for storage at ≤−65° C.

Example 4: Analysis of ICP0−/ICP4-Replication Defective HSV Viral Vectors

[0257] This example demonstrates that HSV strains with ICP0 and ICP4 deletions (ICP0−/ICP4−) are superior for skin applications. ICP0−/ICP4− mutants were found to be less toxic to skin cells and more suitable for delivery to skin than other mutants (e.g., ICP4-(ICP0+)). Additionally, ICP0−/ICP4− mutants, but not ICP4− (ICP0+) mutants, maintained the tissue specificity of a tissue-specific promoter (e.g., a collagen 1 promoter).

ICP0−/ICP4− Mutants Deliver Stronger Payload Expression

[0258] To assess the role of ICP0 in payload expression, GFP expression was measured in ICP4− (ICP0+) and ICP0−/ICP4− mutants (MOI=5). GFP expression was quantified 18 hours after infection. In the ICP4− (ICP0+) mutant, GFP expression was driven by an HCMV promoter. In the ICP0−/ICP4− mutant, GFP expression was driven by the ICP0 promoter. Deletion of ICP0 did not reduce GFP expression (FIG. **3**B), relative to the GFP expression observed in the ICP4− (ICP0+) mutant (FIG. **3**A). These results demonstrate that strong gene expression is retained in the ICP0−/ICP4− mutants, despite the deletion of ICP0.

[0259] Additionally, the role of ICP22 and ICP27 in gene expression was assessed. The deletion of multiple immediate early genes, including (1) ICP0−/ICP4−/ICP22−/ICP27−, (2) ICP0−/ICP4−/ICP22−, (3) ICP0−/ICP4−/ICP27−, and (4) ICP0−/ICP22−/ICP27−, resulted in little gene expression on non-complement cells. Given that both the ICP0−/ICP4−/ICP27− and ICP0−/ICP4−/ICP22-mutants exhibited poor expression but the ICP0−/ICP4− mutant (PGN04) exhibited good expression (FIG. **3**B), the data indicate that ICP22 and ICP27 are both useful for strong gene expression in an ICP0− mutant, and need not be deleted to reduce vector toxicity.

ICP0−/ICP4− Mutants Provide Persistent Expression

[0260] The persistence of gene expression was assessed by measuring GFP expression for an ICP0−/ICP4− mutant (PGN04.6; MOI=10) in Vero cells (dividing) in culture over five days (FIGS. **4**A-**4**D). Expression of GFP was driven by the ICP0 promoter. ICP0-viruses generally do not show good gene expression; however, the ICP0−/ICP4− mutant (PGN04.6) did (see FIGS. **4**A-**4**D). This retention of gene expression in the ICP0−/ICP4− mutant (PGN04.6) could be due to the retention of ICP22 and ICP27, which play a role in the innate immune response evasion of the virus.

ICP0−/ICP4− Mutants Facilitate Tissue Specific Delivery

[0261] ICP0−/ICP4− mutants designed to express a payload driven by a collagen-1 promoter were found to express the payload only in skin cells (fibroblasts), with no expression observed in neurons or keratinocytes (FIGS. **6**A and **6**B). This selective expression pattern makes ICP0−/ICP4− mutants ideal for use in skin indications.

[0262] Data was collected for an ICP4− (ICP0+) mutant with a HCMV promoter, an ICP4− (ICP0+) mutant with a large collagen-1 (eCOL1) promoter, an ICP0−/ICP4− virus control (with ICP0-GFP and HCMV-Red) and an ICP0−/ICP4− mutant with a large collagen-1 promoter (eCOL1) in cell lines of epidermal and neuronal lineage. Gene expression results (in picograms per mL sTNFR (HCMV 100×))

were measured for each mutant in three different cell types: HaCaT (keratinocyte cell line), SK-N-MC (neuroblastoma cells), and Rin5F (insulinoma cells). Both SK-N-MC and Rin5F cells express neuronal markers. ICP0− mutants containing the eCOL1 promoter showed no activity in the keratinocyte (HaCaT) or neuronal cell types (SK-N-MC or Rin5F). In the presence of ICP0, however, tissue specificity of the promoters was lost (FIG. **5**). Thus, ICP0 deletion facilitates tissue-specific expression by tissue-specific promoters.

ICP0−/ICP4− Provides Inducible and Tissue Specific Expression

[0263] Tissue specific expression and inducibility of an ICP0−/ICP4− mutant was assessed by expressing sTNFR using a fibroblast-specific promoter (collagen 1 promoter). Expression was measured in BJ fibroblasts and NCTC 2472 cells (mouse fibroblastic tumor). STNFR expression was measured in picograms per mL sTNFR. The collagen 1 promoter (eCOL1) promoter had expression that was limited to only fibroblasts in ICP0−/ICP4-viruses, and did not express in neurons or keratinocytes (FIGS. **6**A and **6**B). Thus, the use of an ICP0−/ICP4-mutant with an eCOL1 promoter can protect neurons from any off-target expression. These data also show that expression from the eCOL1 promoter can be induced by TGF-β in media, providing expression that is both tissue-specific and inducible.

EQUIVALENTS

[0264] Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. The scope of the present invention is not intended to be limited to the above Description, but rather is as set forth in the following claims:

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 25

<210> SEQ ID NO 1
<211> LENGTH: 25869
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 1

gcagcccggg ccccccgcgc gcggggcggc gcgcaaaaaa ggcgggcggc ggtccgggcg 60

gcgtgcgcgc gcgcggcggg cgtgggggggc ggggccgcgg gagcgggggga ggagccccac 120

ccacagacgg ggaggagcgg gggaggagcg ggggaggagc gggggaggag ccccacccac 180

agacggggag gagcgggggа ggagcggcca gaccccaaaa acgggccccc ccgaaacaca 240

ccccccgggg gtcgcgcgcg gccctttaaa gcgcggcggc gggcagcccg ggcccccgc 300

ggccgagact agcgagttag acaggcaagc actactcgcc tctgcacgca catgcttgcc 360

tgtcaaactc taccaccccg gcacgctctc tgtctccatg gcccgccgcc gccgccatcg 420

cggcccccgc cgcccccggc cgcccgggcc cacgggcgcc gtcccaaccg cacagtccca 480

ggtaacctcc acgcccaact cggaacccgc ggtcaggagc gcgcccgcgg ccgcccgcc 540

gccgcccccc gccggtgggc ccccgccttc ttgttcgctg ctgctgcgcc agtggctcca 600

cgttcccgag tccgcgtccg acgacgacga tgacgacgac tggccggaca gccccccgcc 660

cgagccggcg ccagaggccc ggcccaccgc cgccgccccc cggccccggt ccccaccgcc 720

cggcgtgggc ccgggggggcg gggctgaccc ctcccacccc ccctcgcgcc ccttccgcct 780

tccgccgcgc ctcgccctcc gcctgcgcgt caccgcggag cacctggcgc gcctgcgcct 840

gcgacgcgcg ggcggggagg gggcgccgga gccccccgcg acccccgcga ccccgcgac 900

ccccgcgacc cccgcgaccc ccgcgacccc cgcgaccccc gcgacccccg cgcgggtgcg 960

cttctcgccc cacgtccggg tgcgccacct ggtggtctgg gcctcggccg cccgcctggc 1020

gcgccgcggc tcgtgggccc gcgagcgggc cgaccgggct cggttccggc gccgggtggc 1080

ggaggccgag gcggtcatcg ggccgtgcct ggggcccgag gcccgtgccc gggccctggc 1140

ccgcggagcc ggcccggcga actcggtcta acgttacacc cgaggcgcct gggtcttccg 1200

cggagctccc gggagctccg caccaagccg ctctccggag agacgatggc aggagccgcg 1260

-continued

catatatacg ctgggagccg gtccgccccc aaggcgggcc cgcctcgggg gcgggactgg 1320

ccaatcggcg gccgccagcg cggcggggcc cggccaacca gcgtccgccg agtcttcggg 1380

gcccggccca ttgggcggga gttaccgccc aatgggccgg gccgcccact tcccggtatg 1440

gtaattaaaa acttgcaaga ggccttgttc cgcttcccgg tatggtaatt agaaactcat 1500

taatgggcgg ccccggccgc ccttcccgct tccggcaatt cccgcggccc ttaatgggca 1560

accccggtat tccccgcctc ccgcgccgcg cgtaaccact cccctggggt tccgggttat 1620

gctaattgct tttttggcgg aacacacggc ccctcgcgca ttggcccgcg ggtcgctcaa 1680

tgaacccgca ttggtcccct ggggttccgg gtatggtaat gagtttcttc gggaaggcgg 1740

gaagccccgg ggcaccgacg caggccaagc ccctgttgcg tcggcgggag gggcatgcta 1800

atggggttct ttgggggaca ccgggttggt cccccaaatc gggggccggg ccgtgcatgc 1860

taatgatatt ctttgggggc gccgggttgg tccccgggga cggggccgcc ccgcggtggg 1920

cctgcctccc ctgggacgcg cggccattgg gggaatcgtc actgccgccc ctttggggag 1980

gggaaaggcg tggggtataa gttagccctg gcccgacggt ctggtcgcat ttgcacctcg 2040

gcactcggag cgagacgcag cagccaggca gactcgggcc gccccctctc cgcatcacca 2100

cagaagcccc gcctacgttg cgacccccag ggaccctccg tccgcgaccc tccagccgca 2160

tacgaccccc atggagcccc gccccggagc gagtacccgc cggcctgagg gccgccccca 2220

gcgcgaggtg aggggccggg cgccatgtct ggggcgccat attggggggc gccatgttgg 2280

gggacccccg acccttaccc tggaaccggc ccccatgttg ggggaccccc actcatacac 2340

gggagccggg cgccatgttg gggcgccatg ttaggggggcg tggaaccccg tgacactata 2400

tatacaggga ccgggggcgc catgttaggg ggcgcggaac cccctgaccc tatatataca 2460

gggaccgggg tcgccctgtt gggggtcgcc atgtgacccc ctgactttat atatacagac 2520

ccccaacaca tacacatggc ccctttgact cagacgcagg gcccggggtc gccgtgggac 2580

cccctgactc atacacagag acacgccccc acaacaaaca cacagggacc ggggtcgccg 2640

tgttgggggc gtggtcccca ctgactcata cgcaggcccc ccttactcac acgcatctag 2700

gggggtgggg aggagccgcc cgccatattt gggggacgcc gtgggacccc cgactccggt 2760

gcgtctggag ggcgggagaa gagggaagaa gaggggtcgg gatccaaagg acggacccag 2820

accacctttg gttgcagacc cctttctccc ccctcttccg aggccagcag gggggcagga 2880

ctttgtgagg cgggggggga gaggggaac tcgtgggcgc tgattgacgc gggaaatccc 2940

cccccattct tacccgcccc cctttttttcc ccttagcccg ccccggatgt ctgggtgttt 3000

ccctgcgacc gagacctgcc ggacagcagc gactctgagg cggagaccga agtgggggggg 3060

cgggggggacg ccgaccacca tgacgacgac tccgcctccg aggcggacag cacggacacg 3120

gaactgttcg agacggggct gctggggccg cagggcgtgg atgggggggc ggtctcgggg 3180

gggagccccc cccgcgagga agaccccggc agttgcgggg gcgcccccccc tcgagaggac 3240

gggggggagcg acgagggcga cgtgtgcgcc gtgtgcacgg atgagatcgc gccccacctg 3300

cgctgcgaca ccttcccgtg catgcaccgc ttctgcatcc cgtgcatgaa aacctggatg 3360

caattgcgca acacctgccc gctgtgcaac gccaagctgg tgtacctgat agtgggcgtg 3420

acgcccagcg ggtcgttcag caccatcccg atcgtgaacg acccccagac ccgcatggag 3480

gccgaggagg ccgtcagggc gggcacggcc gtggacttta tctggacggg caatcagcgg 3540

-continued

ttcgccccgc ggtacctgac cctgggggggg cacacggtga gggccctgtc gcccacccac 3600

ccggagccca ccacggacga ggatgacgac gacctggacg acggtgaggc ggggggcggc 3660

aaggaccctg ggggaggagg aggagggagg aatgggcggg cgggcgagga aagggcgggc 3720

cggggagggg gcgtaacctg atcgcgcccc ccgttgtctc ttgcagcaga ctacgtcccg 3780

cccgcccccc gccggacgcc ccgcgccccc ccacgcagag gcaccgccgc gccccccgtg 3840

acgggcgggg cgtctaacgc agcccccag ccggccgcgg ctcggacagc gccccctcg 3900

gcgcccatcg ggccacacgg cagcagtaac accaacacca ccaccaacag cagcggcggc 3960

ggcggctccc gccagtcgcg agccgcggcg ccgcgggggg cgtctggccc ctccgggggg 4020

gttggggttg gggttggggt tgttgaagcg gaggcggggc ggccgagggg ccggacgggc 4080

ccccttgtca acagacccgc ccccttgca aacaacagag accccatagt gatcagcgac 4140

tccccccgg cctctcccca caggcccccc gcggcgccca tgccaggctc cgcccccgc 4200

cccgggcccc ccgcgtcctc ggccgcgtcg ggacccgcgc gcccccgcgc ggccgtggcc 4260

ccgtgcgtgc gagcgccgcc tccggggccc ggcccccgcg ccccggcccc cggggcggag 4320

ccggccgccc gccccgcgga cgcgcgccgt gtgccccagt cgcactcgtc cctggctcag 4380

gccgcgaacc aagaacagag tctgtgccgg gcgcgtgcga cggtggcgcg cggctcgggg 4440

gggccgggcg tggagggtgg gcacgggccc tcccgcggcc gcaccccctc cggcgccgcc 4500

ccgctcccct ccgccgtctc tgtcgagcag gaggcggcgg tgcgtccgag gaagaggcgc 4560

gggtcgggcc aggaaaaccc ctccccccag tccacgcgtc ccccctcgc gccggcaggg 4620

gccaagaggg cggcgacgca ccccccctcc gactcagggc cgggggggcg cggccagggt 4680

gggcccggga ccccctgac gtcctcggcg gcctccgcct cttcctcctc tgcctcttcc 4740

tcctcggccc cgacccccgc gggggccgcc tcttccgccg ccggggccgc gtcctcctcc 4800

gcttccgcct cctcgggcgg ggccgtcggt gccctgggag ggagacaaga ggaaacctcc 4860

ctcggccccc gcgctgcttc tgggccgcgg gggccgagga agtgtgcccg gaagacgcgc 4920

cacgcggaga cttccggggc cgtccccgcg ggcggcctca cgcgctacct gcccatctcg 4980

ggggtctcta gcgtggtcgc cctgtcgcct tacgtgaaca agactatcac gggggactgc 5040

ctgcccatcc tggacatgga gacggggaac atcggggcgt acgtggtcct ggtggaccag 5100

acgggaaaca tggcgacccg gctgcgggcc gcggtccccg gctggagccg ccgcaccctg 5160

ctccccgaga ccgcgggtaa ccacgtgatg ccccccgagt acccgacggc ccccgcgtcg 5220

gagtggaaca gcctctggat gacccccgtg gggaacatgc tgttcgacca gggcacccta 5280

gtgggcgccc tggacttccg cagcctgcgg tctcggcacc cgtggtccgg ggagcagggg 5340

gcgtcgaccc gggacgaggg aaaacaataa gggacgcccc ccgtgtttgt ggggaggggg 5400

gtcgggtgct gggtggtctc tggccgcgcc cactacacca gccaatccgt gtcggggagg 5460

ggaaagtgaa agacacgggc accacacacc agcgggtctt tagtgttggc cctaataaaa 5520

aactcagggg atttttgctg tctattggga aataaaggtt tacttttgta tcttttccct 5580

gtctgtgttg gatggatctt gggggtgcgt gggagtgggg gtgcgtggga gtgggggtgc 5640

gtgggagtgg gggtgcgtgg gagtgggggt gcgtgggagt gggggtgcgt gggagtgggg 5700

gtgcgtggga gtgggggtgc gtgggagtgg gggtgcgtgg gagtgggggt gcgtgggagt 5760

gggggtgcgt gggagtgggg gtgcgtggga gtgggggtgc gtgggagtgg gggtgccatg 5820

-continued

ttgggcaggc tctggtgtta accacagagc cgcggcccgg gctgcctgac caccgatccc 5880

cgaaagcatc ctgccactgg catggagcca gaaccacagt gggctgggtg tgggtgttaa 5940

gtttccgcga gcgcctgccc gcccggactg acctggcctc tggccgccac aaagggcggg 6000

ggggggggtta actacactat agggcaacaa aggacgggag gggtggcggg acggggcgcc 6060

caaaaggggg tcggccacac cacagacgtg ggtgttgggg ggtggggcgg aggggtgggg 6120

gggagacaga aacaggaaca tagttagaaa acaagaatgc ggtgcagcca gagaatcaca 6180

ggagacgagg ggatgggcgt gttggttacc aacccacacc caggcatgct cggtggtatg 6240

aaggaggggg ggcggtgctt cttagagacc gccgggggac gtggggttgg tgtgcaaagg 6300

cacgcgcacc cgcgtcggcc aggtgggccg gtaccccatc ccccccctccc ccgacccttc 6360

ccccccccgcg tgccagagat cacccccgtc ccccggcacc cgccactcct ccatatcctc 6420

gctttaggaa caactttggg gggggggtac acacgcgccg tgcatttcct tccacacccc 6480

ccctcccccg catccccccc cccaggcagt aagacccaag catagagagc caggcacaaa 6540

aacacaggcg gggtgggaca catgccttct tggagtacgt gggtcattgg cgtgggggggt 6600

tacagcgaca ccggccgacc ccctggcggt cttccagccg gcccttagat aaggggggcag 6660

ttggtggtcg gacgggtaag taacagagtc tgactaaggg tgggaggggg ggaaaagaac 6720

gggctggtgt gctgtaacac gagcccaccc gcgagtggcg tggccgacct tagcctctgg 6780

ggcgcccccct gtcgtttggg tccccccccc tctattgggg agaagcaggt gtctaaccta 6840

cctggaaacg cggcgtcttt gttgaacgac accggggcgc cctcgacgag tgggataacg 6900

ggggaggaag ggagggagga gggtactggg ggtgaagaag gggggggggga agaagcgaga 6960

acaggaaagg cgacggagcc cggcagaaca ccgaggaaaa aaaaaacaca gcgcatgcgc 7020

cgggccgttg tggggccccg ggccgggggcc ccttgggtcc gccggggccc cgggccgggc 7080

cgccacgggg gccggccgtt ggcggtaacc ccgattgttt atctcaggcc ccgggccggg 7140

aacccggaaa agcctccggg gggccttttt cgcgtcgcgt gccggcgagc gggcccggac 7200

ggggcccgga ccgccgcggt cgggggcccc tcgtcccggg ccgtacgcgg ccttcgcccc 7260

gtgaggggcc gaacgaacga aacatcccgg cgacggaacg aaaaacaccc cagacgggtt 7320

taaaaaacag aaaccgtaac ccccccacc cccgaaacgg ggaaaacaaa aaacagacca 7380

gcggccggcc ggcgcttagg gggaggatgt cgccgacgcc ccttggccgc cccggctgca 7440

gggggcccg gagagccgcg gcacccggac gcgcccggaa agtctttcgc accacccgcg 7500

atcggcacgg ccgcgccccc gcttttataa aggctcagat gacgcagcaa aaacaggcca 7560

cagcaccacg tgggtaggtg atgtaatttt attttcctcg tctgcggcct aatggatttc 7620

cgggcgcggt gcccctgtct gcagagcact taacggattg atatctcgcg ggcacgcgcg 7680

cccttaatgg accggcgcgg ggcgggggggc cggataccca cacgggcggg ggggtgtcgc 7740

gggccgtctg ctggcccgcg gccacataaa caatgactcg gggcctttct gcctctgccg 7800

cttgtgtgtg cgcgcgccgg ctctgcggtg tcggcggcgg ctgcggcggc tgcggcggcc 7860

gccgtgttcg gtctcggtag ccggccggcg ggtggactcg cgggggggccg gagggtggaa 7920

ggcaggggggg tgtaggatgg gtatcaggac ttccacttcc cgtccttcca tcccccgttc 7980

ccctcggttg ttcctcgcct cccccaacac cccgccgctt tccgttgggg ttgttattgt 8040

tgtcgggatc gtgcgggccg ggggtcgccg gggcaggggc gggggcgggg gtgctcgtcg 8100

-continued

atcgaccggg ctcagtgggg gcgtggggtg ggtgggaaaa ggcgaggaga ctggggtggg 8160

gggtgtcggg ggtggctgtt tttttgtggt tgtttttttgt gtctgttccc gtcccccgtc 8220

acccccctcc ctccgtcccc ccgtcgcggg tgtttgtgtt tgtttattcc gacatcggtt 8280

tatttaaata aacacagccg ttctgcgtgt ctgttcttgc gtgtggctgg gggcttatat 8340

gtggggtccc gggggcggga tggggtttag cggcgggggg cggcgcgccg gacggggcgc 8400

tggagataac ggccccccggg gaacgggggga ccggggctgg gtctcccgcg gtgggtgggt 8460

gggcggcggt ggccgggccg ggccgggccg ggtgggcggg gtttggaaaa acgaggagga 8520

ggagaaggag gaggaggggg ggggagacgg ggggaaagca aggacacggc cccgggggggg 8580

ggggagcgcg ggccgggccg cttggcaacc cccctgtttc ttccggaaac caggcttgtg 8640

gccccacccg acatcacaag ggacctcttg tcgggcctcc cgacgtacgc cgaggctatg 8700

tcggaccacc ccccaaccta agaggggaga ggggagaggg gagaggggag aggggagagg 8760

ggagagggga ggagaggggg tatataaacc aacgaaaagc gcgggaacgg ggatacgggg 8820

cttgtgtggc acgacgtcgt ggttgtgtta ctgggcaaac acttggggac tgtaggtttc 8880

tgtggtgccg accctaggcg ctatggggat tttgggttgg gttgggctta ttgccgttgg 8940

ggttttgtgt gtgcgggggg gcttgccttc aaccgaatat gttattcgga gtcgggtggc 9000

tcgagaggtg ggggatatat taaaggtgcc ttgtgtgccg ctcccgtctg acgatcttga 9060

ttggcgctac gagacccccct cggctataaa ctatgctttg atagacggta tatttttgcg 9120

ttatcactgt cccggattgg acacggtctt gtgggatagg cacgcccaga gggcgtattg 9180

ggttaacccc tttttgtttg gggcgggttt tttggaggac ttgagtcatc ccgcgtttcc 9240

tgccgacacc caggaaacag aaacgcgctt ggccctttat aaagagatac gccaggcgct 9300

ggacagtcgc aagcaggccg ccagccacac acctgtgaag gctgggtgtg tgaactttga 9360

ctattcgcgc acccgccgct gtgtagggcg ccaggatttg ggacttacca acagaacgtc 9420

tggacggacc ccggttctgc cgtcggacga tgaagcgggc ctgcagccga agcccctcac 9480

cacgccgtcg cccatcatcg ccacgtcgga ccccaccccg cgacgggacg ccgccacaaa 9540

aagcagacgc cgacgacccc attcccggcg catctaatga tgcctcgacg gaaaaccgtc 9600

cgggtttggg gggcgaaccg gccgcctgtc gctcgtcagg gccggcgggc gctcctcgcc 9660

gccctagagg ctgtcccgct ggtgtgacgt tttcctcgtc cgcgcccccc gaccctccca 9720

tggatttaac aaacgggggg gtgtcgcctg cggcgacctc ggcgcctctg gactggacca 9780

cgtttcggcg tgtgtttctg atcgacgacg cgtggcggcc cctgttggag cctgagctgg 9840

cgaacccctt aaccgcccac ctcctggccg aatataatcg tcggtgccag accgaagagg 9900

tgctgccgcc gcgggaggat gtgttttcgt ggactcgtta ttgcaccccc gacgaggtgc 9960

gcgtggttat catcggccag gacccatatc accaccccgg ccaggcgcac ggacttgcgt 10020

ttagcgtgcg cgcgaacgtg ccgcctcccc cgagtcttcg gaatgtcttg gcggccgtca 10080

agaactgtta tcccgaggca cggatgagcg gccacggttg cctggaaaag tgggcgcggg 10140

acggcgtcct gttactaaac acgaccctga ccgtcaagcg cggggcggcg gcgtcccact 10200

ctagaatcgg ttgggaccgc ttcgtgggcg gagttatccg ccggttggct gcgcgccgcc 10260

ccggcctggt gtttatgctc tggggcgcac atgcccagaa tgccatcagg ccggaccctc 10320

gggtccattg cgtcctcaag ttttcgcacc cgtcgccccct ctccaaggtt ccgttcggaa 10380

-continued

catgccagca tttcctcgtg gcgaatcgat atctcgagac ccggtcgatt tcacccatcg 10440

actggtcggt ttgaaaggca tcgacgtccg gggttttcgt ctgtgggggc ttttgggtat 10500

ttccgatgaa taaagacggt taatggttaa acctctggtc tcatacgggt cggtgatgtc 10560

gggcgtcggg ggagagggag ttccctctgc gcttgcgatt ctagcctcgt ggggctggac 10620

gttcgacacg ccaaaccacg agtcagggat atcgccagat acgactcccg cagattccat 10680

tcgggggccc gctgtggcct cacctgacca acctttacac gggggcccgg aacgggaggc 10740

cacagcgccg tctttctccc caacgcgcgc ggatgacggc ccgccctgta ccgacgggcc 10800

ctacgtgacg tttgataccc tgtttatggt gtcgtcgatc gacgaattag ggcgtcgcca 10860

gctcacggac accatccgca aggacctgcg gttgtcgctg gccaagttta gcattgcgtg 10920

caccaagacc tcctcgtttt cgggaaacgc cccgcgccac cacagacgcg gggcgttcca 10980

gcgcggcacg cgggcgccgc gcagcaacaa aagccttcag atgtttgtgt tgtgcaaacg 11040

cgcccacgcc gctcgagtgc gagagcagct tcgggtcgtt attcagtccc gcaagccgcg 11100

caagtattac acgcgatctt cggacgggcg gctctgcccc gccgtccccg tgttcgtcca 11160

cgagttcgtc tcgtccgagc caatgcgcct ccaccgagat aacgtcatgc tggcctcggg 11220

ggccgagtaa ccgccccccc cccgcgccac cctcactgcc cgtcgcgcgt gtttgatgtt 11280

aataaataac acataaattt ggctggttgt ttgttgtctt taatggaccg cccgcagggg 11340

gggtggcatt tcagtgtcgg gtgacgagcg cgatccggcc gggatcctag gaccccaaaa 11400

gtttgtctgc gtattccagg gcggggctca gttgaatctc ccgcagcacc tctaccagca 11460

ggtccgcggt gggctggaga aactcggccg tcccggggca ggcggtcgtc gggagtggag 11520

gcgcggcgcc caccccgtgt gccgcgcctg gcgtctcctc tgggggcgac ccgtaaatgg 11580

ttgcagtgat gtaaatggtg tccgcggtcc agaccacggt caaaatgccg gccgtggtgc 11640

tccgggcgct ttcgccgcgc gaggagctga cccaggagtc gaacggatac gcgtacatat 11700

gggcgtccca cccgcgttcg agcttctggt cgctgtcccg gcctataaag cggtaggcac 11760

aaaattcggc gcgacagtcg ataatcacca acagcccaat gggggtgtgc tggataacaa 11820

cgcctccgcg cggcaggcgg tcctggcgct cccggccccg taccataatc gcgcgggtgc 11880

cgtactcaaa aacatgcacc acctgcgcgg cgtcgggcag tgcgctggtc agcgaggccc 11940

tggcgtggca taggctatac gcgatggtcg tctgtggatt ggacatctcg cggtgggtag 12000

tgagtccccc gggccgggtt cggtagaact gtaaggggac ggcgggttaa tagacaatga 12060

ccacgttcgg atcgcgcaga gccgatagta tgtgctcact aatgacgtca tcgcgctcgt 12120

ggcgctcccg gagcggattt aagttcatgc gaaggaattc ggaggaggtg gtgcgggaca 12180

tggccacgta cgcgctgttg aggcgcaggt tgccgggcgt aaagcagatg gcgaccttgt 12240

ccaggctaag gccctgggag cgcgtgatgg tcatggcaag cttggagctg atgccgtagt 12300

cggcgtttat ggccatggcc agctccgtag agtcaatgga ctcgacaaac tcgctgatgt 12360

tggtgttgac gacggacatg aagccgtgtt ggtcccgcaa gaccacgtaa ggcaggggggg 12420

cctcttccag taactcggcc acgttggccg tcgcgtgccg cctccgcagc tcgtccgcaa 12480

aggcaaacac ccgtgcgtac gtgtatccca tgagcgtata attgtccgtc tgcagggcga 12540

cggacatcag ccccccgcgc ggcgagccgg tcagcatctc gcagccccgg aagataacgt 12600

tgtccacgta cgtgctaaag ggggcgcctt caaatgcctc cccgaagagc tcttggagga 12660

-continued

ttcggaatct cccgaggaag gcccgcttca gcagcgcaaa ctgggtgtga acggcggcgg 12720

tggtctccgg ttccccgggg gtgtagtggc agtaaaacac gtcgagctgt tgttcgtcca 12780

gccccgcgaa aataacgtcg aggtcgtcgt cgggaaaatc gtccgggccc ccgtcccgcg 12840

gccccagttg cttaaaatca aacgcacgct cgccgggggc gcctgcgtcg gctattaccg 12900

acgcctgcgt cggcgccccc gaagatttgg ggcgcagaga cagaatctcc gccgttagtt 12960

ctcccatgcg ggcgtaggcg agggtcctct gggtcgcatc caggcccggg cgctgcagaa 13020

agttgtaaaa ggagataagc ccgctaaata tgagccgcga caggaacctg taggcaaact 13080

ccaccgaagt ctccccctga gtctttacaa agctgtcgtc acgcaacact gcctcgaagg 13140

cccggaacgt cccactaaac ccaaaaacca gttttcgcag gcgcgcggtc accgcgatct 13200

ggctgttgag gacgtaagtg acgtcgttgc gggccacgac cagctgctgt ttgctgtgca 13260

cctcgcagcg catgtgcccc gcgtcctggt cctggctctg cgagtagttg gtgatgcggc 13320

tggcgttggc cgtgagccac ttttcaatag tcaggccggg ctggtgtgtc agccgtcggt 13380

attcgtcaaa ctccttgacc gacacgaacg taagcacggg gagggtgaac acgacgaact 13440

ccccctcacg ggtcaccttc aggtaggcgt ggagcttggc catgtacgcg ctcacctctt 13500

tgtgggagga gaacagccgc gtccagccgg ggaggttggc ggggttggtg atgtagcttt 13560

ccgggacgac gaagcgatcc acgaactgca tgtgctcctc ggtgatgggc aggccgtact 13620

ccagcacctt catgaggtta ccgaactcgt gctcgacgca ccgtttgttg ttaataaaaa 13680

tggcccagct atacgagagg cgggcgtact cgcgcagcgt gcggttgcag atgaggtacg 13740

tgagcacgtt ctcgctctgg cggacggaac accgcagttt ctggtgctcg aaggtcgact 13800

ccagggacgc cgtctgcgtc ggcgagccca cacacaccaa cacgggccgc aggcgggccg 13860

catactgggg ggtgtggtac agggcgttaa tcatccacca gcaatacacc acggccgtga 13920

ggaggtgacg cccaaggagc ccggcctcgt ctatgacgat cacgttgctg cgggtaaagg 13980

ccggcagcgc cccgtgggtg gccggggcca accgcgtcag ggcgccctcg gccaacccca 14040

gggtccgttc cagggcggcc agggcgcgaa actcgttccg cgactcctcg cccccggagg 14100

cggccagggc gcgcttcgtg aggtccaaaa tcacctccca gtagtacgtc agatctcgtc 14160

gctgcaggtc ctccagcgag gcggggttgc tggtcagggt gtacgggtac tgtcccagtt 14220

gggcctggac gtgattcccg cgaaacccaa attcatgaaa gatggtgttg atgggtcggc 14280

tgagaaaggc gcccgagagt ttggcgtaca tgttttgggc cgcaatgcgc gtggcgcccg 14340

tcaccacaca gtccaagacc tcgttgattg tctgcacgca cgtgctcttt ccggagccag 14400

cgttgccggt gataagatac accgcgaacg gaaactccct gaggggcagg cctgcggggg 14460

actctaaggc cgccacgtcc cggaaccact gcagacgggg cacttgcgct ccgtcgagct 14520

gttgttgcga gagctctcgg atgcgcttaa ggattggctg caccccgtgc atagacgtaa 14580

aatttaaaaa ggcctcggcc ctccctggaa cggctggtcg gtccccgggt tgctgaaggt 14640

gcggcgggcc gggtctctgt ccgtctagct ggcgctcccc gccggccgcc gccatgaccg 14700

caccacgctc gcgggccccc actacgcgtg cgcgggggga cacggaagcg ctgtgctccc 14760

ccgaggacgg ctgggtaaag gttcacccca cccccggtac gatgctgttc cgtgagattc 14820

tccacgggca gctggggtat accgagggcc agggggtgta caacgtcgtc cggtccagcg 14880

aggcgaccac ccggcagctg caggcggcga tctttcacgc gctcctcaac gccaccactt 14940

-continued

```
accgggacct cgaggcggac tggctcggcc acgtggcggc ccgcggtctg cagccccaac  15000

ggctggttcg ccggtacagg aacgcccggg aggcggatat cgccggggtg gccgagcggg  15060

tgttcgacac gtggcggaac acgcttagga cgacgctgct ggactttgcc cacgggttgg  15120

tcgcctgctt tgcgccgggc ggcccgagcg gcccgtcaag cttccccaaa tatatcgact  15180

ggctgacgtg cctggggctg gtccccatat tacgcaagcg acaagaaggg ggtgtgacgc  15240

agggtctgag ggcgtttctc aagcagcacc cgctgacccg ccagctggcc acggtcgcgg  15300

aggccgcgga gcgcgccggc cccgggtttt ttgagctggc gctggccttc gactccacgc  15360

gcgtggcgga ctacgaccgc gtgtatatct actacaacca ccgccggggc gactggctcg  15420

tgcgagaccc catcagcggg cagcgcggag aatgtctggt gctgtggcct cccttgtgga  15480

ccggggaccg tctggtcttc gattcgcccg tccagcggct gtttcccgag atcgtcgcgt  15540

gtcactccct ccgggaacac gcgcacgtct gccggctgcg caataccgcg tccgtcaagg  15600

tgctgctggg gcgcaagagc gacagcgagc gcggggtggc cggcgccgcg cgggtcgtta  15660

acaaggtgtt gggggaggac gacgagacca aggccgggtc ggccgcctca cgcctcgtgc  15720

ggcttatcat caacatgaag ggcatgcgcc acgtaggcga cattaacgac accgtgcgtg  15780

cctacctcga cgaggccggg gggcacctga tagacgcccc ggccgtcgac ggtaccctcc  15840

ctggattcgg caagggcgga aacagccgcg ggtctgcggg ccaggaccag gggggcgggg  15900

cgccgcagct tcgccaggcc ttccgcacgg ccgtggttaa caacatcaac ggcgtgttgg  15960

agggctatat aaataacctg tttggaacca tcgagcgcct gcgcgagacc aacgcgggcc  16020

tggcgaccca attgcaggag cgcgaccgcg agctccggcg cgcaacagcg ggggccctgg  16080

agcgccagca gcgcgcggcc gacctggcgg ccgagtccgt gaccggtgga tgcggcagcc  16140

gccctgcggg ggcggacctg ctccgggccg actatgacat tatcgacgtc agcaagtcca  16200

tggacgacga catgtacgtc gccaacagct ttcagcaccc gtacatccct tcgtacgccc  16260

aggacctgga gcgcctgtcg cgcctctggg agcacgagct ggtgcgctgt tttaaaattc  16320

tgtgtcaccg caacaaccag ggccaagaga cgtcgatctc gtactccagc ggggcgatcg  16380

ccgcattcgt cgccccctac tttgaggcag tgcttcgggc cccccgggta ggcgcgccca  16440

tcacgggctc cgatgtcatc ctgggggagg aggagttatg ggatgcggtg tttaagaaaa  16500

cctgcctgca aacgtacctg acagacatcg cggccctgtt cgtcgcggac gtccagcacg  16560

cagcgctgcc cccgcccccc tccccggtcg gcgccgattt ccggcccggc gcgtccccgc  16620

ggggccggtc cagatcgcgg tcgcccggaa gaactgcgcg aggcgcaccg gaccagggcg  16680

ggggcatcgg gcaccgggat ggccgccgcg acggccgacg atgaggggtc ggccgtcacc  16740

atcctcaagc aggccatcgc cggggaccgc agcctggtcg aggcggccga ggcgattagc  16800

cagcagacgc tgctccgcct ggcctgcgag gtgcgccagg tcggcgaccg ccagccgcgg  16860

tttaccgcca ccagcatcgc gcgcgtcgac gtcgcgcctg ggtgccggtt gcggttcgtt  16920

ctggacggga gtcccgagga cgcctatgtg acgtcggagg attactttaa gcgctgctgc  16980

ggccagtcca gttatcgcgg cttcgcggtg gcggtcctga cggccaacga ggaccacgtg  17040

cacagcctgg ccgtgccccc cctcgttctg ctgcaccggt tctccctgtt caaccccagg  17100

gacctcctgg actttgagct tgcctgtctg ctgatgtacc tggagaactg cccccgaagc  17160

cacgccaccc cgtcgacctt tgccaaggtt ctggcgtggc tcggggtcgc gggtcgccgc  17220
```

-continued

acgtccccat tcgaacgcgt tcgctgcctt ttcatccgca gttgccactg ggtcctaaac 17280

acactcatgt tcatggtgca cgtaaaaccg ttcgacgacg agttcgtcct gccccactgg 17340

tacatggccc ggtacctgct ggccaacaac ccgccccccg ttctctcggc cctgttctgt 17400

gccaccccga cgagctcctc attccggctg ccggggccgc ccccccgctc cgactgcgtg 17460

gcctataacc ccgccgggat catggggagc tgctgggcgt cggaggaggt gcgcgcgcct 17520

ctggtctatt ggtggctttc ggagacccca aaacgacaga cgtcgtcgct gttttatcag 17580

ttttgttgaa ttttagtaaa taaacccggt tttgtttcta tggcctcctg acggatgcgc 17640

gtgtccttac tccgttttgg tgggtgggtg gctgtgtatg gcgtcccatc tgtgcgggga 17700

gggggcaagt cggcacgtat tcggacagac tcaagcacac acgggggagc gctcttggct 17760

cagggcaatg tttttattgg tcaaactcag gcaaacagaa acaacatctt gtcgtcaaag 17820

ggatacacaa acttcccccc ctcgccccat actcccgcca gcaccccggt aaacaccaac 17880

tcaatctcgc gcaggatttc gcgcaggtga tgagcgcagt ccacgggggg gagcacaagg 17940

ggccgcgggt atagatcgac ggggacgccg accgactccc cgcctccggg acagacacgc 18000

acgacgcgcc gccagtagtg ctctgcgtcc aacaaggcgc cgccgcggaa ggcagtgggg 18060

ggcaaggggt cgctggcctc aaaggggggac acccgaacgc tccagtactc cgcgtccaac 18120

cgtttattaa acgcgtccac gataaggcgg tcgcaggcgt cctccataag gccccgggcc 18180

gtgagcgcgt cctcctccgg cacgcctgcc gttgtcaggc ccaggacccg tcgcagcgtg 18240

tcgcgtacga ccccggccgc cgtggtgtac gcgggcccgc ggagaggaaa tcccccaaga 18300

tggtcagtgt tgtcgcggga gttccagaac cacactcccg cttggctcca ggcgacggcg 18360

tgggtgtaga cgccctcgag cgccaggcac agtgggtgcc gcagccggag gccgttggcc 18420

ataagcacgg ctcccacggc cgtctcgatg gcccgccggg cgtcctcgat caccccggaa 18480

gccgcatccg cgtcttgggg gtccacgtta aagacacccc agaacgcacc cccatcgccc 18540

ccgcagaccg cgaacttaac cgagctggcc gtctcctcaa tctgcaggca gacggcggcc 18600

atcaccccgc ccaggagctg ccgcagcgca gggcaggcgt cgcacgtgtc cgggaccagg 18660

cgctccaaga cggccccggc ccagggctct gagggagcgg ccaccaccag cgcgtccagt 18720

cttgctaggc ccgtccggcc gtggggttcc gccagcccgc tccccccgag gtcggccagg 18780

gccgccagga gctgggcgcg aagtccgggg aagcaaaacc gcgccgtcca gacgggcccg 18840

acggccgcgg gcgggtctaa cagttggatg attttagtgg cgggatgcca ccgcgccacc 18900

gcctcccgca ccgcgggcag gaggcatccg gctgccgccg aggccacgcc gggccaggct 18960

cgcgggggga ggacgaccct ggcccccacc gcgggccagg cccccaggag cgcggcgtaa 19020

gcggccgcgg ccccgcgcac caggtcccgt gccgactcgg ccgtggccgg cacggtgaac 19080

gtgggccaac ccggaaaccc caggacggca aagtacggga cgggtccccc ccggacctca 19140

aactcgggcc ccagaaaggc aaagacgggg gccagggccc cgggggcggc gtggaccgtg 19200

gtatgccact gccggaaaag ggcgacgagc gccggcgcgg agaacttctc gccggcgctt 19260

acaaagtagt cgtaatcgcg gggcagcagc acccgtgccg tgactcgttg cgggtgcccg 19320

cgtggccgca ggcccacctc gcacacctcg accaggtccc cgaacgctcc ctccttcttg 19380

atcggcggaa acgcaagagt ctggtattcg cgcgcaaata gcgcggttcc ggtggtgatg 19440

ttaacggtca gcgaagcggc ggacgcgcac tgggggggtgt cgcgaatggc cgccaggcgc 19500

-continued

gcccacgcca gccgcgcgtc gggatgctcg gcaacgcgcg ccgccagggc catagggtcg 19560

atgtcaatgt tggcctccgc gaccaggaga gcggcgcgag gggcggcggg cgggccccac 19620

gacgctctct caactttcac caccagtccc gtgcgtgggt ccgagccgat acgcagcggg 19680

gcgaacaggg ccaccggccc ggtctggcgc tccagggccg ccaggacgca cgcgtacagc 19740

gcccgccaca gagtcgggtt ctccaggggc tccagcgggg aggcggccgg cgtcgtcgcg 19800

gcgcgggcgg ccgccacgac ggcctggacg gagacgtccg cggagccgta gaaatcccgc 19860

agctccgtcg cggtgacgga gacctccgca aagcgcgcgc gaccctcccc tgcggcgttg 19920

cgacatacaa aatacaccag ggcgtggaag tactcgcgag cgcgggggggg cagccatacc 19980

gcgtaaaggg taatggcgct gacgctctcc tccacccaca cgatatctgc ggtgtccatc 20040

gcacggcccc taaggatcac gggcggtctg tgggtcccat gctgccgtgc ctggccgggc 20100

ccggtgggtc gcggaaaccg gtgacggggg ggggcggttt ttggggttgg ggtgggggtg 20160

ggaaacggcc cgggtccggg ggccaacttg gcccctcggt gcgttccggc aacagcgccg 20220

ccggtccgcg gacgaccacg taccgaacga gtgcggtccc gagacttata gggtgctaaa 20280

gttcaccgcc ccctgcatca tgggccaggc ctcggtgggg agctccgaca gcgccgcctc 20340

caggatgatg tcagcgttgg ggttggcgct ggatgagtgc gtgcgcaaac agcgccccca 20400

cgcgggcacg cgtagcttga agcgcgcgcc cgcaaactcc cgcttgtggg ccataagcag 20460

ggcgtacagc tgcctgtggg tccggcaggc gctgtggtcg atgtggtggg cgtccaacaa 20520

ccccacgatt gtctgtttgg tgaggttttt aacgcgcccc gccccgggaa acgtctgcgt 20580

gcttttggcc atctgcacgc caaacagttc gccccagatt atcttgaaca gcgccaccgc 20640

gtggtccgtc tcgctaacgg acccgcgcgg gggacagccg cttagggcgt cggcgacgcg 20700

cttgacggct tcctccgaga gcagaagtcc gtcggttacg ttacagtggc ccagttcgaa 20760

caccagctgc atgtagcggt cgtagtgggg ggtcagtagg tccagcacgt catcggggcc 20820

gaaggtcctc ccagatcccc cggccgccga gtcccaatgc aggcgcgcgg ccatggtgct 20880

gcacaggcac aacagctccc agacgggggt tacgttcagg gtgggggggca gggccacgag 20940

ctccagctct ccggtgacgt tgatcgtggg gatgacgccc gtggcgtagt ggtcatagat 21000

ccgccgaaat atggcgctgc tgcgggtggc catgggaacg cggagacagg cctccagcaa 21060

cgccaggtaa ataaaccgcg tgcgtcccat caggctgttg aggttgcgca tgagcgcgac 21120

aatttccgcc ggcgcgacat cggaccggag gtatttttcg acgaaaagac ccacctcctc 21180

cgtctcggcg gcctgggccg gcagcgacgc ctcgggatcc cggcaccgca gctcccgtag 21240

atcgcgctgg gccctgaggg cgtcgaaatg tacgccccgc aaaaacagac agaagtcctt 21300

tggggtcagg gtatcgtcgt gtccccagaa gcgcacgcgt atgcagttta gggtcagcag 21360

catgtgaagg atgttaaggc tgtccgagag acacgccagc gtgcatctct caaagtagtg 21420

tttgtaacgg aatttgttgt agatgcgcga cccccgcccc agcgacgtgt cgcatgccga 21480

cgcgtcacag cgccccttga accggcgaca cagcaggttt gtgacctggg agaactgcgc 21540

gggccactgg ccgcaggaac tgaccacgtg gttcaggagc atgggcgtaa agacgggctc 21600

cgagcgcgcc ccggagccgt ccatgtaaat cagtagctcc cccttgcgga gggtgcgcac 21660

ccgtcccagg gactggtaca cggacaccat gtccggtccg tagttcatgg gtttcacgta 21720

ggcgaacatg ccatcaaagt gcaggggatc gaagctgagg cccacggtta cgaccgtcgt 21780

-continued

gtatataacc acgcggtatt ggccccacgt ggtcacgtcc ccgaggggggg tgagcgagtg 21840

aagcaacagc acgcggtccg taaactgacg gcagaaccgg gccacgatct ccgcgaagga 21900

gaccgtcgat gaaaaaatgc agatgttatc gcccccgcca aggcgcgctt ccagctcccc 21960

aaagaacgtg gccccccggg cgtccggaga ggcgtccgga gacgggccgc ttggcggccc 22020

gggcgggcgc agggcagcct gcaggagctc ggtccccaga cgcgggagaa acaggcaccg 22080

gcgcgccgaa aacccgggca tggcgtactc gccgaccacc acatgcacgt ttttttcgcc 22140

ccggagaccg cacaggaagt ccaccaactg cgcgttggcg gttgcgtcca tggcgatgat 22200

ccgaggacag gtgcgcagca ggcgtagcat taacgcatcc acgcggccca gttgctgcat 22260

cgttggcgaa tagagctggc ccagcgtcga cataacctcg tccagaacga ggacgtcgta 22320

gttgttcaga aggttggggc ccacgcgatg aaggctttcc acctggacga taagtcggtg 22380

gaaggggcgg tcgttcataa tgtaattggt ggatgagaag taggtgacaa agtcgaccag 22440

gcctgactca gcgaaccgcg tcgccagggt ctgggtaaaa ctccgacgac aggagacgac 22500

gagcacactc gtgtccggag agtggatcgc ttcccgcagc cagcggatca gcgcggtagt 22560

ttttcccgac cccattggcg cgcggaccac agtcacgcac ctggccgtcg gggcgctcgc 22620

gttggggaag gtgacgggtc cgtgctgctg ccgctcgatc gttgttttcg ggtgaacccg 22680

gggcacccat tcggccaaat ccccccccgta caacatccgc gctagcgata cgctcgacgt 22740

gtactgttcg cactcgtcgt ccccaatggg acgcccggcc cccagaggat cccccgactc 22800

cgcgcccccc acgaaaggca tgaccggggc gcggacggcg tggtgggtct ggtgtgtgca 22860

ggtggcgacg tttgtggtct ctgcggtctg cgtcacgggg cttctcgtcc tggcctctgt 22920

gttccgggca cggtttccct gcttttatgc cacggcgagc tcttatgccg gggtgaactc 22980

cacggccgag gtgcgcgggg gtgtagccgt gcccctcagg ttggacacgc agagccttgt 23040

gggcacttat gtaatcacgg ccgtgttgtt gttggccgcg gccgtgtatg ccgtggtcgg 23100

cgccgtgacc tcccgctacg accgcgccct ggacgcgggc cgccgtctgg ctgcggcccg 23160

catggccatg ccgcacgcca cgctgatcgc cggaaacgtc tgctcttggt tgctgcagat 23220

caccgtcctg ttgctggccc atcgcatcag ccagctggcc cacctggttt acgtcctgca 23280

ctttgcgtgt ctggtgtatt ttgcggccca tttttgcacc agggggggtcc tgagcgggac 23340

gtatctgcgt caggtgcacg gcctgatgga gccggccccg actcatcatc gcgtcgtcgg 23400

cccggctcga gccgtgctga caaacgcctt gctgttgggc gtcttcctgt gcacggccga 23460

cgccgcggta tccctgaata ccatcgccgc gttcaacttt aatttttcgg ccccgggcat 23520

gctcatatgc ctgaccgtgc tgttcgccct tctcgtcgta tcgctgttgt tggtggtcga 23580

gggggtgttg tgtcactacg tgcgcgtgtt ggtgggcccc cacctggggg ccgtggccgc 23640

cacgggcatc gtcggcctgg cctgcgagca ctattacacc aacggctact acgttgtgga 23700

gacgcagtgg ccgggggccc agacgggagt ccgcgtcgcc ctcgccctgg tcgccgcctt 23760

tgccctcggc atggccgtgc tccgctgcac ccgcgcctat ctgtatcaca ggcggcacca 23820

caccaaattt tttatgcgca tgcgcgacac gcgacaccgc gcacattccg ccctcaggcg 23880

cgtacgcagt tccatgcgcg gatcgcgaga cggccgccac aggcccgcac ccggcagccc 23940

gcccgggatt cccgaatatg cggaagaccc ctacgcgatc tcatacggcg gccagctcga 24000

ccggtacgga gattccgacg gggagccgat ttacgacgag gtggctgacg accaaaccga 24060

-continued

cgtattgtac gccaagatac aacacccgcg gcacctgccc gacgacgagc ccatctatga 24120

caccgttggg gggtacgacc ccgagcccgc cgaggacccc gtgtacagca ccgtccgccg 24180

ttggtagctg tttggttccg ttttaataaa ccgtttgtgt ttaacccgac cgtggtgtat 24240

gtctggtgtg tggcgtccga tcccgttact atcaccgttt ccccccccc cccctcaacc 24300

ccggcgattg tgggtttttt aaaaacgaca cgcgtgcgac cgtatacaga acattgtttt 24360

ggtttttatt cgctatcgga catgggggt ggaaactggg tggcggggca ggcgcctccg 24420

ggggtccgcc ggtgagtgtg gcgcgagggg gggtccgacg aacgcaggcg ctgtctcccc 24480

ggggcccgcg taaccacgcg catatccggg ggcacgtaga aattaccttc ctcttcggac 24540

tcgatatcca cgacatcaaa gtcgtgggcg gtcagcgaga cgacctcccc gtcgtcggtg 24600

atgaggacgt tgtttcggca gcagcagggc cgggccccgg agaacgagag gcccatagct 24660

cggcgagcgt gtcgtcgaac gccaggcggc tgcttcgctg gatggcctta tagatctccg 24720

gatcgatgcg gacgggggta atgatcaggg cgatcggaac ggcctggttc gggagaatgg 24780

acgccttgct gggtcctgcg gccccgagag ccccggcgcc gtcctccagg cggaacgtta 24840

cgccctcctc cgcgctggtg cggtgcctgc cgataaacgt caccagatgc gggtgggggg 24900

ggcagtcggg gaagtggctg tcgagcacgt agccctgcac caagatctgc ttaaagttcg 24960

ggtgacgggg gttcgcgaag acgggctcgc ggcggaccag atccccggag ctccaggaca 25020

cgggggagat ggtgtggcgt ccgaggtcgg gggcgccaaa cagaagcacc tccgagacaa 25080

cgccgctatt taactccacc aaggcccgat ccgcggcgga gcaccgcctt ttttcgcccg 25140

aggcgtgggc ctctgaccag gcctggtctt gcgtgacgag agcctcctcc gggccgggga 25200

cgcgcccggg cgcgaagtat cgcacgctgg gcttcgggat cgaccggata aatgcccgga 25260

acgcctccgg ggaccggtgt gtcatcaagt cctcgtacgc ggaggccgtg gggtcgctgg 25320

ggtccatggg gtcgaaagcg tacttggccc ggcatttgac ctcgtaaaag gccaggggcg 25380

tcttggggac tggggccagg tagccgtgaa tgtcccgagg acagacgaga atatccaggg 25440

acgccccgac catccccgtg tgaccgtcca tgaggacccc acacgtatgc acgttctctt 25500

cggtgaggtc gctgggttcg tggaagataa agcgccgcgt gtcggcgccg gcctcgccgc 25560

cgtcgtccgc gcggcccacg cagtagcgaa acagcaggct tcgggccgtc ggctcgttca 25620

cccgcccgaa catcaccgcc gaagactgta catccggccg caggctggcg ttgtgcttca 25680

gccactgggg cgagaaacac ggaccctggg ggccccagcg gagggtggat gcggtcgtga 25740

ggccccgccg gagcagggcc catagctggc agtcggcctg gttttgcgtg gccgcctcgt 25800

aaaaccccat gaggggccgg ggcgccacgg cgtccgcggc ggccgggggc ccgcggcgcg 25860

tcaggcgcc 25869

<210> SEQ ID NO 2
<211> LENGTH: 1300
<212> TYPE: PRT
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 2

Met Ala Ser Glu Asn Lys Gln Arg Pro Gly Ser Pro Gly Pro Thr Asp
1 5 10 15

Gly Pro Pro Pro Thr Pro Ser Pro Asp Arg Asp Glu Arg Gly Ala Leu
20 25 30

-continued

```
Gly Trp Gly Ala Glu Thr Glu Glu Gly Gly Asp Asp Pro Asp His Asp
        35                  40                  45

Pro Asp His Pro His Asp Leu Asp Asp Ala Arg Arg Asp Gly Arg Ala
    50                  55                  60

Pro Ala Ala Gly Thr Asp Ala Gly Glu Asp Ala Gly Asp Ala Val Ser
65                  70                  75                  80

Pro Arg Gln Leu Ala Leu Leu Ala Ser Met Val Glu Glu Ala Val Arg
                85                  90                  95

Thr Ile Pro Thr Pro Asp Pro Ala Ala Ser Pro Pro Arg Thr Pro Ala
            100                 105                 110

Phe Arg Ala Asp Asp Asp Asp Gly Asp Glu Tyr Asp Asp Ala Ala Asp
        115                 120                 125

Ala Ala Gly Asp Arg Ala Pro Ala Arg Gly Arg Ala Arg Glu Ala Pro
    130                 135                 140

Leu Arg Gly Ala Tyr Pro Asp Pro Thr Asp Arg Leu Ser Pro Arg Pro
145                 150                 155                 160

Pro Ala Gln Pro Pro Gln Arg Arg Arg His Gly Arg Arg Arg Pro Ser
                165                 170                 175

Ala Ser Ser Thr Ser Ser Asp Ser Gly Ser Ser Ser Ser Ser Ser Ala
            180                 185                 190

Ser Ser Ser Ser Ser Ser Ser Asp Glu Asp Glu Asp Asp Asp Gly Asn
        195                 200                 205

Asp Ala Ala Asp His Ala Arg Glu Ala Arg Ala Val Gly Arg Gly Pro
    210                 215                 220

Ser Ser Ala Ala Pro Glu Ala Pro Gly Arg Thr Pro Pro Pro Pro Gly
225                 230                 235                 240

Pro Pro Pro Leu Ser Glu Ala Ala Pro Lys Pro Arg Ala Ala Ala Arg
                245                 250                 255

Thr Pro Ala Ala Ser Ala Gly Arg Ile Glu Arg Arg Arg Ala Arg Ala
            260                 265                 270

Ala Val Ala Gly Arg Asp Ala Thr Gly Arg Phe Thr Ala Gly Gln Pro
        275                 280                 285

Arg Arg Val Glu Leu Asp Ala Asp Ala Ala Ser Gly Ala Phe Tyr Ala
    290                 295                 300

Arg Tyr Arg Asp Gly Tyr Val Ser Gly Glu Pro Trp Pro Gly Ala Gly
305                 310                 315                 320

Pro Pro Pro Pro Gly Arg Val Leu Tyr Gly Gly Leu Gly Asp Ser Arg
                325                 330                 335

Pro Gly Leu Trp Gly Ala Pro Glu Ala Glu Glu Ala Arg Arg Arg Phe
            340                 345                 350

Glu Ala Ser Gly Ala Pro Ala Ala Val Trp Ala Pro Glu Leu Gly Asp
        355                 360                 365

Ala Ala Gln Gln Tyr Ala Leu Ile Thr Arg Leu Leu Tyr Thr Pro Asp
    370                 375                 380

Ala Glu Ala Met Gly Trp Leu Gln Asn Pro Arg Val Val Pro Gly Asp
385                 390                 395                 400

Val Ala Leu Asp Gln Ala Cys Phe Arg Ile Ser Gly Ala Ala Arg Asn
                405                 410                 415

Ser Ser Ser Phe Ile Thr Gly Ser Val Ala Arg Ala Val Pro His Leu
            420                 425                 430

Gly Tyr Ala Met Ala Ala Gly Arg Phe Gly Trp Gly Leu Ala His Ala
```

-continued

```
        435                 440                 445

Ala Ala Ala Val Ala Met Ser Arg Arg Tyr Asp Arg Ala Gln Lys Gly
    450                 455                 460

Phe Leu Leu Thr Ser Leu Arg Arg Ala Tyr Ala Pro Leu Leu Ala Arg
465                 470                 475                 480

Glu Asn Ala Ala Leu Thr Gly Ala Ala Gly Ser Pro Gly Ala Gly Ala
                485                 490                 495

Asp Asp Glu Gly Val Ala Ala Ala Val Val Ala Ala Ala Ala Ala Pro
            500                 505                 510

Gly Glu Arg Ala Val Pro Ala Gly Tyr Gly Ala Ala Gly Ile Leu Ala
        515                 520                 525

Ala Leu Gly Arg Leu Ser Ala Ala Pro Ala Ser Pro Ala Gly Gly Asp
    530                 535                 540

Asp Pro Asp Ala Ala Arg His Ala Asp Ala Asp Asp Asp Ala Gly Arg
545                 550                 555                 560

Arg Ala Gln Ala Gly Arg Val Ala Val Glu Cys Leu Ala Ala Cys Arg
                565                 570                 575

Gly Ile Leu Glu Ala Leu Ala Glu Gly Phe Asp Gly Asp Leu Ala Ala
            580                 585                 590

Val Pro Gly Leu Ala Gly Ala Arg Pro Ala Ser Pro Pro Arg Pro Glu
        595                 600                 605

Gly Pro Ala Gly Pro Ala Ser Pro Pro Pro Pro His Ala Asp Ala Pro
    610                 615                 620

Arg Leu Arg Ala Trp Leu Arg Glu Leu Arg Phe Val Arg Asp Ala Leu
625                 630                 635                 640

Val Leu Met Arg Leu Arg Gly Asp Leu Arg Val Ala Gly Gly Ser Glu
                645                 650                 655

Ala Ala Val Ala Ala Val Arg Ala Val Ser Leu Val Ala Gly Ala Leu
            660                 665                 670

Gly Pro Ala Leu Pro Arg Asp Pro Arg Leu Pro Ser Ser Ala Ala Ala
        675                 680                 685

Ala Ala Ala Asp Leu Leu Phe Glu Asn Gln Ser Leu Arg Pro Leu Leu
    690                 695                 700

Ala Ala Gly Pro Arg Arg Ser Ser Ser Ser Ser Gly Val Ala Ala Ala
705                 710                 715                 720

Ala Ser Ala Ala Pro Arg Glu Gly Arg Lys Arg Lys Ser Pro Gly Pro
                725                 730                 735

Ala Arg Pro Pro Gly Gly Gly Gly Pro Arg Pro Pro Lys Thr Lys Lys
            740                 745                 750

Ser Gly Ala Asp Ala Pro Gly Ser Asp Ala Arg Ala Pro Leu Pro Ala
        755                 760                 765

Pro Ala Pro Pro Ser Thr Pro Pro Gly Pro Glu Pro Ala Pro Ala Gln
    770                 775                 780

Pro Ala Ala Pro Arg Ala Ala Ala Ala Gln Ala Arg Pro Arg Pro Val
785                 790                 795                 800

Ala Leu Ser Arg Arg Pro Ala Glu Gly Pro Asp Pro Leu Gly Gly Trp
                805                 810                 815

Arg Arg Gln Pro Pro Gly Pro Ser His Thr Ala Ala Pro Ala Ala Ala
            820                 825                 830

Ala Leu Glu Ala Tyr Cys Ser Pro Arg Ala Val Ala Glu Leu Thr Asp
        835                 840                 845
```

-continued

```
His Pro Leu Phe Pro Val Pro Trp Arg Pro Ala Leu Met Phe Asp Pro
    850                 855                 860

Arg Ala Leu Ala Ser Ile Ala Ala Arg Cys Ala Gly Pro Ala Pro Ala
865                 870                 875                 880

Ala Gln Ala Ala Cys Gly Gly Gly Asp Asp Asp Glu Asn Pro His Pro
                885                 890                 895

His Gly Ala Ala Gly Gly Arg Leu Phe Gly Pro Leu Arg Ala Ser Gly
            900                 905                 910

Pro Leu Arg Arg Met Ala Ala Trp Met Arg Gln Ile Pro Asp Pro Glu
        915                 920                 925

Asp Val Arg Val Val Val Leu Tyr Ser Pro Leu Pro Gly Glu Asp Leu
    930                 935                 940

Ala Gly Gly Gly Ala Ser Gly Gly Pro Pro Glu Trp Ser Ala Glu Arg
945                 950                 955                 960

Gly Gly Leu Ser Cys Leu Leu Ala Ala Leu Ala Asn Arg Leu Cys Gly
                965                 970                 975

Pro Asp Thr Ala Ala Trp Ala Gly Asn Trp Thr Gly Ala Pro Asp Val
            980                 985                 990

Ser Ala Leu Gly Ala Gln Gly Val  Leu Leu Leu Ser Thr  Arg Asp Leu
        995                 1000                1005

Ala Phe  Ala Gly Ala Val Glu  Phe Leu Gly Leu Leu  Ala Ser Ala
    1010                1015                1020

Gly Asp  Arg Arg Leu Ile Val  Val Asn Thr Val Arg  Ala Cys Asp
    1025                1030                1035

Trp Pro  Ala Asp Gly Pro Ala  Val Ser Arg Gln His  Ala Tyr Leu
    1040                1045                1050

Ala Cys  Asp Leu Leu Pro Ala  Val Gln Cys Ala Val  Arg Trp Pro
    1055                1060                1065

Ala Ala  Arg Asp Leu Arg Arg  Thr Val Leu Ala Ser  Gly Arg Val
    1070                1075                1080

Phe Gly  Pro Gly Val Phe Ala  Arg Val Glu Ala Ala  His Ala Arg
    1085                1090                1095

Leu Tyr  Pro Asp Ala Pro Pro  Leu Arg Leu Cys Arg  Gly Gly Asn
    1100                1105                1110

Val Arg  Tyr Arg Val Arg Thr  Arg Phe Gly Pro Asp  Thr Pro Val
    1115                1120                1125

Pro Met  Ser Pro Arg Glu Tyr  Arg Arg Ala Val Leu  Pro Ala Leu
    1130                1135                1140

Asp Gly  Arg Ala Ala Ala Ser  Gly Thr Thr Asp Ala  Met Ala Pro
    1145                1150                1155

Gly Ala  Pro Asp Phe Cys Glu  Glu Glu Ala His Ser  His Arg Ala
    1160                1165                1170

Cys Ala  Arg Trp Gly Leu Gly  Ala Pro Leu Arg Pro  Val Tyr Val
    1175                1180                1185

Ala Leu  Gly Arg Glu Ala Val  Arg Ala Gly Pro Ala  Arg Trp Arg
    1190                1195                1200

Gly Pro  Arg Arg Asp Phe Cys  Ala Arg Ala Leu Leu  Glu Pro Asp
    1205                1210                1215

Asp Asp  Ala Pro Pro Leu Val  Leu Arg Gly Asp Asp  Asp Gly Pro
    1220                1225                1230
```

-continued

```
Gly Ala  Leu Pro Pro Ala Pro  Pro Gly Ile Arg Trp  Ala Ser Ala
    1235                1240                1245

Thr Gly  Arg Ser Gly Thr Val  Leu Ala Ala Ala Gly  Ala Val Glu
    1250                1255                1260

Val Leu  Gly Ala Glu Ala Gly  Leu Ala Thr Pro Pro  Arg Arg Asp
    1265                1270                1275

Val Val  Asp Trp Glu Gly Ala  Trp Asp Glu Asp Asp  Gly Gly Ala
    1280                1285                1290

Phe Glu  Gly Asp Gly Val Leu
    1295                1300


<210> SEQ ID NO 3
<211> LENGTH: 420
<212> TYPE: PRT
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 3

Met Ala Asp Ile Ser Pro Gly Ala Phe Ala Pro Cys Val Lys Ala Arg
1               5                   10                  15

Arg Pro Ala Leu Arg Ser Pro Pro Leu Gly Thr Arg Lys Arg Lys Arg
            20                  25                  30

Pro Ala Arg Pro Leu Ser Ser Glu Ser Glu Val Glu Thr Asp Thr Ala
        35                  40                  45

Leu Glu Ser Glu Val Glu Ser Glu Thr Ala Ser Asp Ser Thr Glu Ser
    50                  55                  60

Gly Asp Gln Glu Glu Ala Pro Arg Ile Gly Gly Arg Arg Ala Pro Arg
65                  70                  75                  80

Arg Leu Gly Gly Arg Phe Phe Leu Asp Met Ser Ala Glu Ser Thr Thr
                85                  90                  95

Gly Thr Glu Thr Asp Thr Ala Val Ser Asp Asp Pro Asp Asp Thr Ser
            100                 105                 110

Asp Trp Ser Tyr Asp Asp Ile Pro Pro Arg Pro Lys Arg Ala Arg Val
        115                 120                 125

Asn Leu Arg Leu Thr Ser Ser Pro Asp Arg Arg Asp Gly Val Ile Phe
    130                 135                 140

Pro Lys Met Gly Arg Val Arg Ser Thr Arg Glu Thr Gln Pro Arg Ala
145                 150                 155                 160

Pro Thr Pro Ser Ala Pro Ser Pro Asn Ala Met Leu Arg Arg Ser Val
                165                 170                 175

Arg Gln Ala Gln Arg Arg Ser Ser Ala Arg Trp Thr Pro Asp Leu Gly
            180                 185                 190

Tyr Met Arg Gln Cys Ile Asn Gln Leu Phe Arg Val Leu Arg Val Ala
        195                 200                 205

Arg Asp Pro His Gly Ser Ala Asn Arg Leu Arg His Leu Ile Arg Asp
    210                 215                 220

Cys Tyr Leu Met Gly Tyr Cys Arg Ala Arg Leu Ala Pro Arg Thr Trp
225                 230                 235                 240

Cys Arg Leu Leu Gln Val Ser Gly Gly Thr Trp Gly Met His Leu Arg
                245                 250                 255

Asn Thr Ile Arg Glu Val Glu Ala Arg Phe Asp Ala Thr Ala Glu Pro
            260                 265                 270

Val Cys Lys Leu Pro Cys Leu Glu Ala Arg Arg Tyr Gly Pro Glu Cys
        275                 280                 285
```

-continued

```
Asp Leu Ser Asn Leu Glu Ile His Leu Ser Ala Thr Ser Asp Asp Glu
    290                 295                 300

Ile Ser Asp Ala Thr Asp Leu Glu Ala Ala Gly Ser Asp His Thr Leu
305                 310                 315                 320

Ala Ser Gln Ser Asp Thr Glu Asp Ala Pro Ser Pro Val Thr Leu Glu
                325                 330                 335

Thr Pro Glu Pro Arg Gly Ser Leu Ala Val Arg Leu Glu Asp Glu Phe
            340                 345                 350

Gly Glu Phe Asp Trp Thr Pro Gln Glu Gly Ser Gln Pro Trp Leu Ser
        355                 360                 365

Ala Val Val Ala Asp Thr Ser Ser Val Glu Arg Pro Gly Pro Ser Asp
    370                 375                 380

Ser Gly Ala Gly Arg Ala Ala Glu Asp Arg Lys Cys Leu Asp Gly Cys
385                 390                 395                 400

Arg Lys Met Arg Phe Ser Thr Ala Cys Pro Tyr Pro Cys Ser Asp Thr
                405                 410                 415

Phe Leu Arg Pro
            420


<210> SEQ ID NO 4
<211> LENGTH: 88
<212> TYPE: PRT
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 4

Met Ser Trp Ala Leu Glu Met Ala Asp Thr Phe Leu Asp Asn Met Arg
1               5                   10                  15

Val Gly Pro Arg Thr Tyr Ala Asp Val Arg Asp Glu Ile Asn Lys Arg
            20                  25                  30

Gly Arg Glu Asp Arg Glu Ala Ala Arg Thr Ala Val His Asp Pro Glu
        35                  40                  45

Arg Pro Leu Leu Arg Ser Pro Gly Leu Leu Pro Lys Ile Ala Pro Asn
    50                  55                  60

Ala Ser Leu Gly Val Ala His Arg Arg Thr Gly Gly Thr Val Thr Asp
65                  70                  75                  80

Ser Pro Arg Asn Pro Val Thr Arg
                85


<210> SEQ ID NO 5
<211> LENGTH: 4020
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 5

tttattgcgt cttcgggttt cacaagcgcc ccgccccgtc ccggcccgtt acagcacccc     60

gtccccctcg aacgcgccgc cgtcgtcttc gtcccaggcg ccttcccagt ccacaacgtc    120

ccgtcgcggg ggcgtggcca agcccgcctc cgcccccagc acctccacgg cccccgccgc    180

cgccagcacg gtgccgctgc ggcccgtggc cgaggcccag cgaatcccgg gcggcgccgg    240

cggcagggcc cccgggccgt cgtcgtcgcc gcgcagcacc agcgggggggg cgtcgtcgtc    300

gggctccagc agggcgcggg cgcaaaagtc cctccgcggc ccgcgccacc gggccgggcc    360

ggcgcgcacc gcctcgcgcc ccagcgccac gtacacgggc cgcagcggcg cgcccaggcc    420
```

-continued

ccagcgcgcg caggcgcggt gcgagtgggc ctcctcctcg cagaagtccg gcgcgccggg 480

cgccatggcg tcggtggtcc ccgaggccgc cgcccggccg tccagcgccg gcagcacggc 540

ccggcggtac tcgcgcgggg acatgggcac cggcgtgtcc gggccgaagc gcgtgcgcac 600

gcggtagcgc acgttgccgc cgcggcacag gcgcagcggc ggcgcgtcgg ggtacaggcg 660

cgcgtgcgcg gcctccacgc gcgcgaagac ccccgggccg aacacgcggc ccgaggccag 720

caccgtgcgg cgcaggtcgc gcgccgccgg ccagcgcacg gcgcactgca cggcgggcag 780

caggtcgcac gccaggtagg cgtgctgccg cgacaccgcg ggcccgtcgg cgggccagtc 840

gcaggcgcgc acggtgttga ccacgatgag ccgccggtcg ccggcgctgg cgagcagccc 900

cagaaactcc acggccccgg cgaaggccag gtcccgcgtg gacagcagca gcacgccctg 960

cgcgcccagc gccgacacgt cgggggcgcc ggtccagttg cccgcccagg cggccgtgtc 1020

cggcccgcac agccggttgg ccagggccgc cagcaggcag gacagcccgc cgcgctcggc 1080

ggaccactcc ggcggccccc ccgaggcccc gccgccggcc aggtcctcgc ccggcagcgg 1140

cgagtacagc accaccacgc gcacgtcctc ggggtcgggg atctggcgca tccaggccgc 1200

catgcggcgc agcgggcccg aggcgcgcag ggggccaaag aggcggcccc cggcggcccc 1260

gtgggggtgg gggttctcgt cgtcgtcgcc gccgccgcac gcggcctggg cggcgggggc 1320

gggcccggcg caccgcgcgg cgatcgaggc cagggcccgc gggtcaaaca tgagggccgg 1380

tcgccagggg acggggaaca gcgggtggtc cgtgagctcg gccacggcgc gcggggagca 1440

gtaggcctcc agggcggcgg ccgcgggcgc cgccgtgtgg ctgggccccg ggggctgccg 1500

ccgccagccg cccagggggt cggggccctc ggcgggccgg cgcgacagcg ccacggggcg 1560

cgggcgggcc tgcgccgcgg cggcccgggg cgccgcgggc tgggcggggg cgggctcggg 1620

ccccgggggc gtggaggggg gcgcgggcgc ggggaggggg gcgcgggcgt ccgagccggg 1680

ggcgtccgcg ccgctcttct tcgtcttcgg gggtcgcggg ccgccgcctc cgggcggccg 1740

ggccgggccg ggactcttgc gcttgcgccc ctcccgcggc gcggcggagg cggcggccgc 1800

gacccccgaa gacgaagaag agcggcgcgg acccgccgcc agcagggggc gcaggctctg 1860

gttctcaaac agcaggtccg cggcggcggc ggccgcggag ctcggcaggc gcgggtcccg 1920

cggcagcgcg ggacccaggg ccccggcgac caggctcacg gcgcgcacgg cggccacggc 1980

ggcctcgctg ccgccggcca cgcgcaggtc cccgcgcagg cgcatgagca ccagcgcgtc 2040

gcgcacgaac cgcagctcgc gcagccacgc gcgcaggcgg ggcgcgtcgg cgtgcggcgg 2100

cggcggggaa gcggggcccg cgggtccctc cggccgcggg gggctggcgg gccgggcccc 2160

ggccagcccc gggacggccg ccaggtcgcc gtcgaagccc tcggccagcg cctccaggat 2220

cccgcggcag gcggccaggc actccacggc cacgcggccg gcctgggcgc ggcgcccggc 2280

gtcgtcgtcg gcgtcggcgt ggcgggcggc gtcggggtcg tcgccccccg cgggggaggc 2340

gggcgcggcg gacagccgcc ccagggcggc gaggatcccc gcggcgccgt acccggcggg 2400

caccgcgcgc tcgcccggtg cggcggcggc ggcgacgacg gcggcggcga ccccctcgtc 2460

atctgcgccg gcgccggggc tccccgcggc ccccgtcagc gccgcgttct cgcgcgccaa 2520

caggggcgcg taggcgcggc gcaggctggt cagcaggaag cccttctgcg cgcggtcgta 2580

tcggcggctc atggccacgg cggccgccgc gtgcgccagg ccccagccga agcggccggc 2640

cgccatggcg tagcccaggt ggggcacggc ccgcgccacg ctgccggtga tgaaggagct 2700

-continued

gctgttgcgc gcggcgcccg agatccggaa gcaggcctgg tccagcgcca cgtccccggg 2760

gaccacgcgc gggttctgga gccaccccat ggcctccgcg tccggggtgt acagcagccg 2820

cgtgatcagg gcgtactgct gcgcggcgtc gcccagctcg ggcgcccaca cggccgccgg 2880

ggcgcccgag gcctcgaacc ggcgtcgcgc ctcctccgcc tcgggcgccc cccagaggcc 2940

cgggcggctg tcgcccaggc cgccgtacag cacccgcccc gggggcgggg gcccggcgcc 3000

gggccacggc tccccgctga cgtacccgtc gcgatagcgc gcgtagaagg cgccggaggc 3060

cgcgtcggcg tccagctcga cccgccgggg ctgcccggcc gtgaagcggc ccgtggcgtc 3120

gcggccggcc accgccgcgc gggcccggcg gcgctcgatg cggcccgcgg aggccgcggg 3180

ggtcctcgcc gccgcccggg gcttgggcgc ggcctcggag agggggggtg gcccgggcgg 3240

gggcggcgtc cgcccggggg cttccggcgc cgcgctcgac ggaccccgcc cgacggcccg 3300

cgcctcgcgt gcgtggtcgg ccgcgtcgtt gccgtcgtcg tcctcgtcct cgtcggacga 3360

cgaggacgaa gaggatgcgg acgacgagga cgaggacccg gagtccgacg aggtcgatga 3420

cgccgatggc cgccgccggc cgtgacgacg tctctgcggc ggctgggccg gcgggcgcgg 3480

cgacaggcgg tccgtggggt ccggatacgc gccgcgtagc ggggcctccc gtgcgcggcc 3540

ccgggccggg gcccggtcgc cggcggcgtc ggctgcgtcg tcgtactcgt ccccgtcatc 3600

gtcgtcggct cgaaaggcgg gggtccgggg cggcgaggcc gcggggtcgg gcgtcgggat 3660

cgtccggacg gcctcctcta ccatggaggc cagcagggcc agctgtcgcg gcgagacggc 3720

gtccccggcg tcctcgccgg cgtcggtgcc cgccgcgggg gccctcccgt cccgccgggc 3780

gtcgtcgagg tcgtgggggt ggtcggggtc gtggtcgggg tcgtccccgc cctcctccgt 3840

ctccgcgccc cacccgaggg ccccccgctc gtcgcggtct gggctcgggg tgggcggcgg 3900

cccgtcggtg gggcccgggg agccggggcg ctgcttgttc tccgacgcca tcgccgatgc 3960

ggggcgatcc tccggggata cgactgcgac ggcggacgta gcacggtagg tcacctacgg 4020

<210> SEQ ID NO 6
<211> LENGTH: 1320
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 6

ggtcctccgg gacgttttct ggatggccga catttcccca ggcgcttttg tgccttgtgt 60

aaaagcgcgg cgtcccgctc tccgatcccc gcccctgggc acgcgcaagc gcaagcgccc 120

tgcccgcccc ctctcatcgg agtctgaggt cgaatccgag acagccttgg agtctgaggt 180

cgaatccgag acagcatcgg attcgaccga gtctggggac caggaggaag ccccccgcat 240

cggtggccgt agggcccccc ggaggcttgg ggggcggttt tttctggaca tgtcggcgga 300

atccaccacg gggacggaaa cggatgcgtc ggtgtcggac gaccccgacg acacgtccga 360

ctggtcttgt gacgacattc ccccacgacc caagcgggcc cgggtaaacc tgcggctcac 420

tagctctccc gatcggcggg atggggttat ttttcctaag atggggcggg tccggtctac 480

ccgggaaacg cagccccggg cccccacccc gtcggcccca agcccaaatg caatgctccg 540

gcgctcggtg cgccaggccc agaggcggag cagcgcacga tggacccccg acctgggcta 600

catgcgccag tgtatcaatc agctgtttcg ggtcctgcgg gtcgcccggg acccccacgg 660

cagtgccaac cgcctgcgcc acctgatacg cgactgttac ctgatgggat actgccgagc 720

-continued

ccgtctggcc ccgcgcacgt ggtgccgctt gctgcaggtg tccggcggaa cctggggcat 780

gcacctgcgc aacaccatac gggaggtgga ggctcgattc gacgccaccg cagaacccgt 840

gtgcaagctt ccttgtttgg aggccagacg gtacggcccg gagtgtgatc ttagtaatct 900

cgagattcat ctcagcgcga caagcgatga tgaaatctcc gatgccaccg atctggaggc 960

cgccggttcg gaccacacgc tcgcgtccca gtccgacacg gaggatgccc cctcccccgt 1020

tacgctggaa accccagaac cccgcgggtc cctcgctgtg cgtctggagg atgagtttgg 1080

ggagtttgac tggacccccc aggagggctc ccagccctgg ctgtctgcgg tcgtggccga 1140

taccagctcc gtggaacgcc cgggcccatc cgattctggg gcgggtcgcg cagcagaaga 1200

ccgcaagtgt ctggacggct gccggaaaat gcgcttctcc accgcctgcc cctatccgtg 1260

cagcgacacg tttctccggc cgtgagtccg gtcgccccga ccccttgta tgtccccaaa 1320

<210> SEQ ID NO 7
<211> LENGTH: 360
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 7

tccgcccaga gactcgggtg atggtcgtac ccgggactca acgggttacc ggattacggg 60

gactgtcggt cacggtcccg ccggttcttc gatgtgccac acccaaggat gcgttggggg 120

cgatttcggg cagcagcccg ggagagcgca gcaggggacg ctccgggtcg tgcacggcgg 180

ttctggccgc ctcccggtcc tcacgccccc ttttattgat ctcatcgcgt acgtcggcgt 240

acgtcctggg cccaacccgc atgttgtcca ggaaggtgtc cgccatttcc agggcccacg 300

acatgctttt ccccccgacg agcaggaagc ggtccacgca acggtcgccg ccggtcgcct 360

<210> SEQ ID NO 8
<211> LENGTH: 523
<212> TYPE: DNA
<213> ORGANISM: Human cytomegalovirus

<400> SEQUENCE: 8

gaagatcttt ggttatatag cataaatcaa tattggctat tggccattgc atacgttgta 60

tccatatcat aatatgtaca tttatattgg ctcatgtcca acattaccgc catgttgaca 120

ttgattattg actagttatt aatagtaatc aattacgggg tcattagttc atagcccata 180

tatggagttc cgcgttacat aacttacggt aaatggcccg cctggctgac cgcccaacga 240

cccccgccca ttgacgtcaa taatgacgta tgttcccata gtaacgccaa tagggacttt 300

ccattgacgt caatgggtgg agtatttacg gtaaactgcc cacttggcag tacatcaagt 360

gtatcatatg ccaagtacgc cccctattga cgtcaatgac ggtaaatggc ccgcctggca 420

ttatgcccag tacatgacct tatgggactt tcctacttgg cagtacatct acgtattagt 480

catcgctatt accatggtga tgcggttttg gcagtacatc aat 523

<210> SEQ ID NO 9
<211> LENGTH: 1199
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

aatgggtttg ggtgtgtgta aatgagtgtg accggaagcg agtgtgagct tgatctaggc 60

-continued

agggaccaca cagcactgtc acacctgcct gctctttagt agaggactga agtgcggggg 120

tgggggtacg gggccggaat agaatgtctc tgggacatct tggcaaacag cagccggaag 180

caaaggggca gctgtgcaaa cggctcaggc aggtgatgga tggcagggta ggaaggggga 240

ggtccagagg tctggatgga ggcttccgca tctgtacctt gcaactcacc cctcaggccc 300

agcaggtcat cggccccctc ctcacacatg taatgacgta gaagagtacc ccgggacagt 360

ccggggagat ggagattcgg aaagtatcca tggagctctt acagaatccc ctgtgcggac 420

caggaaactc ttgtagatcc ctgcctatct gaggcccagg cgctgggctg tttctcacaa 480

tattccttca agatgagatt gtggtcccca tttcaaagat gagtacactg agcctctgtg 540

aagttacttg cccatgatca cacaaccagg aattgggcca actgtaattg aactcctgtc 600

taacaaagtt cttgctccca gctccgtctc ttgtttccca cgagccctgg ccctctgtgg 660

gtaataccag ctactggagt cagatttctt gggcccagaa cccaccctta ggggcattaa 720

cctttaaaat ctcacttggg caggggtctg ggatcagagt tggaagagtc cctacaatcc 780

tggacccttt ccgccaaatc gtgaaaccag gggtggagtg gggcgagggt tcaaaaccag 840

gccggactga gaggtgaaat tcaccatgac gtcaaactgc cctcaaattc ccgctcactt 900

taagggcgtt acttgttggt gcccccacca tcccccacca tttccatcaa tgacctcaat 960

gcaaatacaa gtgggacggt cctgctgacg cctccaggtt ctggaagcat gagggtgacg 1020

cacccagggg caaaggaccc ctccgcccat tggttgctgt gcactggcgg aactttcccg 1080

acccacagcg gcgggaataa gagcagtcgc tggcgctggg aggcatcaga gacactgccc 1140

agcccaagtg tcgccgccgc ttccacaggg ctctggctgg acgccgccgc cgccgctgc 1199

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

<400> SEQUENCE: 10

ggatcccgac tgtgccttct agttgccagc catctgttgt ttgcccctcc cccgtgcctt 60

ccttgaccct ggaaggtgcc actcccactg tcctttcctа ataaaatgag gaaattgcat 120

cgcattgtct gagtaggtgt cattctattc tgggggggtgg ggtggggcag gacagcaagg 180

gggaggattg ggaagacaat agcaggcatg ctggggaaga tcttc 225

<210> SEQ ID NO 11
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 11

Ser Thr Pro Ser Thr Thr Thr
1 5

<210> SEQ ID NO 12
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 12

gcacccccac tcccac 16

-continued

<210> SEQ ID NO 13
<211> LENGTH: 16
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 13

ccccagcccc ccccag 16

<210> SEQ ID NO 14
<211> LENGTH: 17
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 14

cccctcgccc cctcccg 17

<210> SEQ ID NO 15
<211> LENGTH: 14
<212> TYPE: PRT
<213> ORGANISM: human herpesvirus 1 strain KOS

<400> SEQUENCE: 15

Ala Ala Ser Ala Pro Asp Ala Ala Asp Ala Leu Ala Ala Ala
1 5 10

<210> SEQ ID NO 16
<211> LENGTH: 13
<212> TYPE: PRT
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 16

Gly Pro Arg Arg Ser Ser Ser Ser Ser Gly Val Ala Ala
1 5 10

<210> SEQ ID NO 17
<211> LENGTH: 775
<212> TYPE: PRT
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 17

Met Glu Pro Arg Pro Gly Ala Ser Thr Arg Arg Pro Glu Gly Arg Pro
1 5 10 15

Gln Arg Glu Pro Ala Pro Asp Val Trp Val Phe Pro Cys Asp Arg Asp
20 25 30

Leu Pro Asp Ser Ser Asp Ser Glu Ala Glu Thr Glu Val Gly Gly Arg
35 40 45

Gly Asp Ala Asp His His Asp Asp Asp Ser Ala Ser Glu Ala Asp Ser
50 55 60

Thr Asp Thr Glu Leu Phe Glu Thr Gly Leu Leu Gly Pro Gln Gly Val
65 70 75 80

Asp Gly Gly Ala Val Ser Gly Gly Ser Pro Pro Arg Glu Glu Asp Pro
85 90 95

Gly Ser Cys Gly Gly Ala Pro Pro Arg Glu Asp Gly Gly Ser Asp Glu
100 105 110

Gly Asp Val Cys Ala Val Cys Thr Asp Glu Ile Ala Pro His Leu Arg
115 120 125

Cys Asp Thr Phe Pro Cys Met His Arg Phe Cys Ile Pro Cys Met Lys

-continued

```
    130                 135                 140

Thr Trp Met Gln Leu Arg Asn Thr Cys Pro Leu Cys Asn Ala Lys Leu
145                 150                 155                 160

Val Tyr Leu Ile Val Gly Val Thr Pro Ser Gly Ser Phe Ser Thr Ile
                165                 170                 175

Pro Ile Val Asn Asp Pro Gln Thr Arg Met Glu Ala Glu Glu Ala Val
            180                 185                 190

Arg Ala Gly Thr Ala Val Asp Phe Ile Trp Thr Gly Asn Gln Arg Phe
        195                 200                 205

Ala Pro Arg Tyr Leu Thr Leu Gly Gly His Thr Val Arg Ala Leu Ser
    210                 215                 220

Pro Thr His Pro Glu Pro Thr Thr Asp Glu Asp Asp Asp Asp Leu Asp
225                 230                 235                 240

Asp Ala Asp Tyr Val Pro Pro Ala Pro Arg Arg Thr Pro Arg Ala Pro
                245                 250                 255

Pro Arg Arg Gly Thr Ala Ala Pro Pro Val Thr Gly Gly Ala Ser Asn
            260                 265                 270

Ala Ala Pro Gln Pro Ala Ala Ala Arg Thr Ala Pro Pro Ser Ala Pro
        275                 280                 285

Ile Gly Pro His Gly Ser Ser Asn Thr Asn Thr Thr Thr Asn Ser Ser
    290                 295                 300

Gly Gly Gly Gly Ser Arg Gln Ser Arg Ala Ala Ala Pro Arg Gly Ala
305                 310                 315                 320

Ser Gly Pro Ser Gly Gly Val Gly Val Gly Val Gly Val Val Glu Ala
                325                 330                 335

Glu Ala Gly Arg Pro Arg Gly Arg Thr Gly Pro Leu Val Asn Arg Pro
            340                 345                 350

Ala Pro Leu Ala Asn Asn Arg Asp Pro Ile Val Ile Ser Asp Ser Pro
        355                 360                 365

Pro Ala Ser Pro His Arg Pro Pro Ala Ala Pro Met Pro Gly Ser Ala
    370                 375                 380

Pro Arg Pro Gly Pro Pro Ala Ser Ser Ala Ala Ser Gly Pro Ala Arg
385                 390                 395                 400

Pro Arg Ala Ala Val Ala Pro Cys Val Arg Ala Pro Pro Pro Gly Pro
                405                 410                 415

Gly Pro Arg Ala Pro Ala Pro Gly Ala Glu Pro Ala Ala Arg Pro Ala
            420                 425                 430

Asp Ala Arg Arg Val Pro Gln Ser His Ser Ser Leu Ala Gln Ala Ala
        435                 440                 445

Asn Gln Glu Gln Ser Leu Cys Arg Ala Arg Ala Thr Val Ala Arg Gly
    450                 455                 460

Ser Gly Gly Pro Gly Val Glu Gly Gly His Gly Pro Ser Arg Gly Arg
465                 470                 475                 480

Thr Pro Ser Gly Ala Ala Pro Leu Pro Ser Ala Val Ser Val Glu Gln
                485                 490                 495

Glu Ala Ala Val Arg Pro Arg Lys Arg Arg Gly Ser Gly Gln Glu Asn
            500                 505                 510

Pro Ser Pro Gln Ser Thr Arg Pro Pro Leu Ala Pro Ala Gly Ala Lys
        515                 520                 525

Arg Ala Ala Thr His Pro Pro Ser Asp Ser Gly Pro Gly Gly Arg Gly
    530                 535                 540
```

-continued

```
Gln Gly Gly Pro Gly Thr Pro Leu Thr Ser Ser Ala Ala Ser Ala Ser
545                 550                 555                 560

Ser Ser Ser Ala Ser Ser Ser Ser Ala Pro Thr Pro Ala Gly Ala Ala
                565                 570                 575

Ser Ser Ala Ala Gly Ala Ala Ser Ser Ser Ala Ser Ala Ser Ser Gly
            580                 585                 590

Gly Ala Val Gly Ala Leu Gly Gly Arg Gln Glu Glu Thr Ser Leu Gly
        595                 600                 605

Pro Arg Ala Ala Ser Gly Pro Arg Gly Pro Arg Lys Cys Ala Arg Lys
    610                 615                 620

Thr Arg His Ala Glu Thr Ser Gly Ala Val Pro Ala Gly Gly Leu Thr
625                 630                 635                 640

Arg Tyr Leu Pro Ile Ser Gly Val Ser Ser Val Val Ala Leu Ser Pro
                645                 650                 655

Tyr Val Asn Lys Thr Ile Thr Gly Asp Cys Leu Pro Ile Leu Asp Met
            660                 665                 670

Glu Thr Gly Asn Ile Gly Ala Tyr Val Val Leu Val Asp Gln Thr Gly
        675                 680                 685

Asn Met Ala Thr Arg Leu Arg Ala Ala Val Pro Gly Trp Ser Arg Arg
    690                 695                 700

Thr Leu Leu Pro Glu Thr Ala Gly Asn His Val Met Pro Pro Glu Tyr
705                 710                 715                 720

Pro Thr Ala Pro Ala Ser Glu Trp Asn Ser Leu Trp Met Thr Pro Val
                725                 730                 735

Gly Asn Met Leu Phe Asp Gln Gly Thr Leu Val Gly Ala Leu Asp Phe
            740                 745                 750

Arg Ser Leu Arg Ser Arg His Pro Trp Ser Gly Glu Gln Gly Ala Ser
        755                 760                 765

Thr Arg Asp Glu Gly Lys Gln
    770                 775


<210> SEQ ID NO 18
<211> LENGTH: 4430
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 18

atggagcccc gccccggagc gagtacccgc cggcctgagg gccgccccca gcgcgaggtg     60

aggggccggg cgccatgtct ggggcgccat attgggggggc gccatgttgg gggacccccg    120

acccttaccc tggaaccggc ccccatgttg ggggaccccc actcatacac gggagccggg    180

cgccatgttg gggcgccatg ttaggggggcg tggaaccccg tgacactata tatacaggga    240

ccgggggcgc catgttaggg ggcgcggaac cccctgaccc tatatataca gggaccgggg    300

tcgccctgtt gggggtcgcc atgtgacccc ctgactttat atatacagac ccccaacaca    360

tacacatggc ccctttgact cagacgcagg gcccggggtc gccgtgggac cccctgactc    420

atacacagag acacgccccc acaacaaaca cacagggacc ggggtcgccg tgttgggggc    480

gtggtcccca ctgactcata cgcaggcccc ccttactcac acgcatctag gggggtgggg    540

aggagccgcc cgccatattt gggggacgcc gtgggacccc cgactccggt gcgtctggag    600

ggcgggagaa gagggaagaa gaggggtcgg gatccaaagg acggacccag accacctttg    660
```

-continued

gttgcagacc cctttctccc ccctcttccg aggccagcag gggggcagga ctttgtgagg 720

cgggggggga gaggggaac tcgtgggcgc tgattgacgc gggaaatccc cccccattct 780

tacccgcccc cctttttttcc ccttagcccg ccccggatgt ctgggtgttt ccctgcgacc 840

gagacctgcc ggacagcagc gactctgagg cggagaccga agtgggggggg cgggggacg 900

ccgaccacca tgacgacgac tccgcctccg aggcggacag cacggacacg gaactgttcg 960

agacggggct gctggggccg cagggcgtgg atgggggggc ggtctcgggg gggagccccc 1020

cccgcgagga agaccccggc agttgcgggg gcgccccccc tcgagaggac gggggggagcg 1080

acgagggcga cgtgtgcgcc gtgtgcacgg atgagatcgc gccccacctg cgctgcgaca 1140

ccttcccgtg catgcaccgc ttctgcatcc cgtgcatgaa aacctggatg caattgcgca 1200

acacctgccc gctgtgcaac gccaagctgg atggagcccc gccccggagc gagtacccgc 1260

cggcctgagg gccgccccca gcgcgaggtg aggggccggg cgccatgtct ggggcgccat 1320

attgggggggc gccatgttgg gggacccccg acccttaccc tggaaccggc ccccatgttg 1380

ggggaccccc actcatacac gggagccggg cgccatgttg gggcgccatg ttaggggggcg 1440

tggaaccccg tgacactata tatacaggga ccgggggcgc catgttaggg ggcgcggaac 1500

cccctgaccc tatatatata gggaccgggg tcgccctgtt gggggtcgcc atgtgacccc 1560

ctgactttat atatacagac ccccaacaca tacacatggc ccctttgact cagacgcagg 1620

gcccggggtc gccgtgggac cccctgactc atacacagag acacgccccc acaacaaaca 1680

cacagggacc ggggtcgccg tgttgggggc gtggtcccca ctgactcata cgcaggcccc 1740

ccttactcac acgcatctag gggggtgggg aggagccgcc cgccatattt gggggacgcc 1800

gtgggacccc cgactccggt gcgtctggag ggcgggagaa gagggaagaa gaggggtcgg 1860

gatccaaagg acggacccag accacctttg gttgcagacc cctttctccc ccctcttccg 1920

aggccagcag gggggcagga ctttgtgagg cgggggggga gaggggaac tcgtgggcgc 1980

tgattgacgc gggaaatccc cccccattct tacccgcccc cctttttttcc ccttagcccg 2040

ccccggatgt ctgggtgttt ccctgcgacc gagacctgcc ggacagcagc gactctgagg 2100

cggagaccga agtgggggggg cgggggggacg ccgaccacca tgacgacgac tccgcctccg 2160

aggcggacag cacggacacg gaactgttcg agacggggct gctggggccg cagggcgtgg 2220

atgggggggc ggtctcgggg gggagccccc cccgcgagga agaccccggc agttgcgggg 2280

gcgcccccccc tcgagaggac gggggggagcg acgagggcga cgtgtgcgcc gtgtgcacgg 2340

atgagatcgc gccccacctg cgctgcgaca ccttcccgtg catgcaccgc ttctgcatcc 2400

cgtgcatgaa aacctggatg caattgcgca acacctgccc gctgtgcaac gccaagctgg 2460

tgtacctgat agtgggcgtg acgcccagcg ggtcgttcag caccatcccg atcgtgaacg 2520

acccccagac ccgcatggag gccgaggagg ccgtcagggc gggcacggcc gtggacttta 2580

tctggacggg caatcagcgg ttcgccccgc ggtacctgac cctgggggggg cacacggtga 2640

gggccctgtc gcccacccac ccggagccca ccacggacga ggatgacgac gacctggacg 2700

acggtgaggc gggggggcggc aaggaccctg ggggaggagg aggagggagg aatgggcggg 2760

cgggcgagga aagggcgggc cggggagggg gcgtaacctg atcgcgcccc ccgttgtctc 2820

ttgcagcaga ctacgtcccg cccgcccccc gccggacgcc ccgcgccccc ccacgcagag 2880

gcaccgccgc gccccccgtg acgggcgggg cgtctaacgc agcccccccag ccggccgcgg 2940

-continued

ctcggacagc gcccccctcg gcgcccatcg ggccacacgg cagcagtaac accaacacca 3000

ccaccaacag cagcggcggc ggcggctccc gccagtcgcg agccgcggcg ccgcgggggg 3060

cgtctggccc ctccgggggg gttggggttg gggttggggt tgttgaagcg gaggcggggc 3120

ggccgagggg ccggacgggc ccccttgtca acagacccgc ccccccttgca aacaacagag 3180

accccatagt gatcagcgac tccccccggg cctctcccca caggcccccc gcggcgccca 3240

tgccaggctc cgcccccccgc cccgggcccc ccgcgtcctc ggccgcgtcg ggacccgcgc 3300

gcccccgcgc ggccgtggcc ccgtgcgtgc gagcgccgcc tccggggccc ggcccccgcg 3360

ccccggcccc cggggcggag ccggccgccc gccccgcgga cgcgcgccgt gtgccccagt 3420

cgcactcgtc cctggctcag gccgcgaacc aagaacagag tctgtgccgg gcgcgtgcga 3480

cggtggcgcg cggctcgggg gggccgggcg tggagggtgg gcacgggccc tcccgcggcc 3540

gcacccccctc cggcgccgcc ccgctcccct ccgccgtctc tgtcgagcag gaggcggcgg 3600

tgcgtccgag gaagaggcgc gggtcgggcc aggaaaaccc ctccccccag tccacgcgtc 3660

ccccctcgc gccggcaggg gccaagaggg cggcgacgca ccccccctcc gactcagggc 3720

cgggggggcg cggccagggt gggcccggga ccccccctgac gtcctcggcg gcctccgcct 3780

cttcctcctc tgcctcttcc tcctcggccc cgacccccgc gggggccgcc tcttccgccg 3840

ccggggccgc gtcctcctcc gcttccgcct cctcgggcgg ggccgtcggt gccctgggag 3900

ggagacaaga ggaaacctcc ctcggccccc gcgctgcttc tgggccgcgg gggccgagga 3960

agtgtgcccg gaagacgcgc cacgcggaga cttccggggc cgtccccgcg ggcggcctca 4020

cgcgctacct gcccatctcg ggggtctcta gcgtggtcgc cctgtcgcct tacgtgaaca 4080

agactatcac gggggactgc ctgcccatcc tggacatgga gacggggaac atcggggcgt 4140

acgtggtcct ggtggaccag acgggaaaca tggcgacccg gctgcgggcc gcggtccccg 4200

gctggagccg ccgcaccctg ctccccgaga ccgcgggtaa ccacgtgatg ccccccgagt 4260

acccgacggc ccccgcgtcg gagtggaaca gcctctggat gacccccgtg gggaacatgc 4320

tgttcgacca gggcacccta gtgggcgccc tggacttccg cagcctgcgg tctcggcacc 4380

cgtggtccgg ggagcagggg gcgtcgaccc gggacgaggg aaaacaataa 4430

<210> SEQ ID NO 19
<211> LENGTH: 8170
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 19

atcgcgggtg gtgcgaaaga ctttccgggc gcgtccgggt gccgcggctc tccgggcccc 60

cctgcagccg gggcggccaa ggggcgtcgg cgacatcctc cccctaagcg ccggccggcc 120

gctggtctgt tttttgtttt ccccgtttcg ggggtggggg gggttacggt ttctgttttt 180

taaacccgtc tggggtgttt ttcgttccgt cgccgggatg tttcgttcgt tcggcccctc 240

acggggcgaa ggccgcgtac ggcccgggac gaggggcccc cgaccgcggc ggtccgggcc 300

ccgtccgggc ccgctcgccg gcacgcgacg cgaaaaaggc cccccggagg cttttccggg 360

ttcccggccc ggggcctgag ataaacaatc ggggttaccg ccaacggccg gcccccgtgg 420

cggcccggcc cggggccccg gcggacccaa ggggccccgg cccgggggccc cacaacggcc 480

cggcgcatgc gctgtgtttt ttttttcctc ggtgttctgc cgggctccgt cgcctttcct 540

-continued

gttctcgctt cttccccccc cccttcttca cccccagtac cctcctccct cccttcctcc 600

cccgttatcc cactcgtcga gggcgccccg gtgtcgttca acaaagacgc cgcgtttcca 660

ggtaggttag acacctgctt ctccccaata gaggggggggg gacccaaacg acaggggggcg 720

ccccagaggc taaggtcggc cacgccactc gcgggtgggc tcgtgttaca gcacaccagc 780

ccgttctttt ccccccctcc cacccttagt cagactctgt tacttacccg tccgaccacc 840

aactgccccc ttatctaagg gccggctgga agaccgccag ggggtcggcc ggtgtcgctg 900

taacccccca cgccaatgac ccacgtactc caagaaggca tgtgtcccac cccgcctgtg 960

tttttgtgcc tggctctcta tgcttgggtc ttactgcctg gggggggggga tgcgggggag 1020

ggggggtgtg gaaggaaatg cacggcgcgt gtgtaccccc cccccaaagt tgttcctaaa 1080

gcgaggatat ggaggagtgg cgggtgccgg gggacggggg tgatctctgg cacgcggggg 1140

gggaagggtc gggggagggg gggatggggt accggcccac ctggccgacg cgggtgcgcg 1200

tgcctttgca caccaacccc acgtcccccg gcggtctcta agaagcaccg ccccccctcc 1260

ttcataccac cgagcatgcc tgggtgtggg ttggtaacca acacgcccat cccctcgtct 1320

cctgtgattc tctggctgca ccgcattctt gttttctaac tatgttcctg tttctgtctc 1380

ccccccaccc ctccgcccca ccccccaaca cccacgtctg tggtgtggcc gacccccttt 1440

tgggcgcccc gtcccgccac ccctcccgtc ctttgttgcc ctatagtgta gttaaccccc 1500

cccccgccct ttgtggcggc cagaggccag gtcagtccgg gcgggcaggc gctcgcggaa 1560

acttaacacc cacacccagc ccactgtggt tctggctcca tgccagtggc aggatgcttt 1620

cggggatcgg tggtcaggca gcccgggccg cggctctgtg gttaacacca gagcctgccc 1680

aacatggcac ccccactccc acgcaccccc actcccacgc acccccactc ccacgcaccc 1740

ccactcccac gcaccccccac tcccacgcac ccccactccc acgcaccccc actcccacgc 1800

acccccactc ccacgcaccc ccactcccac gcaccccccac tcccacgcac ccccactccc 1860

acgcaccccc actcccacgc acccccactc ccacgcaccc ccaagatcca tccaacacag 1920

acagggaaaa gatacaaaag taaaccttta tttcccaata gacagcaaaa atcccctgag 1980

ttttttatta gggccaacac taaagacccg ctggtgtgtg gtgcccgtgt ctttcacttt 2040

cccctccccg acacggattg gctggtgtag tgggcgcggc cagagaccac ccagcacccg 2100

acccccctcc ccacaaacac ggggggcgtc ccttattgtt ttccctcgtc ccgggtcgac 2160

gccccctgct ccccggacca cgggtgccga gaccgcaggc tgcggaagtc cagggcgccc 2220

actagggtgc cctggtcgaa cagcatgttc cccacggggg tcatccagag gctgttccac 2280

tccgacgcgg gggccgtcgg gtactcgggg ggcatcacgt ggttacccgc ggtctcgggg 2340

agcagggtgc ggcggctcca gccggggacc gcggcccgca gccgggtcgc catgtttccc 2400

gtctggtcca ccaggaccac gtacgccccg atgttccccg tctccatgtc caggatgggc 2460

aggcagtccc ccgtgatagt cttgttcacg taaggcgaca gggcgaccac gctagagacc 2520

cccgagatgg gcaggtagcg cgtgaggccg cccgcgggga cggccccgga agtctccgcg 2580

tggcgcgtct tccgggcaca cttcctcggc ccccgcggcc cagaagcagc gcgggggccg 2640

agggaggttt cctcttgtct ccctcccagg gcaccgacgg ccccgcccga ggaggcggaa 2700

gcggaggagg acgcggcccc ggcggcggaa gaggcggccc ccgcgggggt cggggccgag 2760

gaggaagagg cagaggagga agaggcggag gccgccgagg acgtcagggg ggtcccgggc 2820

-continued

ccaccctggc cgcgcccccc cggccctgag tcggaggggg ggtgcgtcgc cgccctcttg 2880

gcccctgccg gcgcgagggg gggacgcgtg gactgggggg aggggttttc ctggcccgac 2940

ccgcgcctct tcctcggacg caccgccgcc tcctgctcga cagagacggc ggaggggagc 3000

ggggcggcgc cggagggggt gcggccgcgg gagggcccgt gcccaccctc cacgcccggc 3060

ccccccgagc cgcgcgccac cgtcgcacgc gcccggcaca gactctgttc ttggttcgcg 3120

gcctgagcca gggacgagtg cgactggggc acacggcgcg cgtccgcggg gcgggcggcc 3180

ggctccgccc cgggggccgg ggcgcggggg ccgggccccg gaggcggcgc tcgcacgcac 3240

ggggccacgg ccgcgcgggg gcgcgcgggt cccgacgcgg ccgaggacgc ggggggcccg 3300

gggcgggggg cggagcctgg catgggcgcc gcgggggccc tgtggggaga ggccgggggg 3360

gagtcgctga tcactatggg gtctctgttg tttgcaaggg gggcgggtct gttgacaagg 3420

gggcccgtcc ggcccctcgg ccgccccgcc tccgcttcaa caaccccaac cccaacccca 3480

acccccccgg aggggccaga cgcccccgc ggcgccgcgg ctcgcgactg gcgggagccg 3540

ccgccgccgc tgctgttggt ggtggtgttg gtgttactgc tgccgtgtgg cccgatgggc 3600

gccgaggggg gcgctgtccg agccgcggcc ggctgggggg ctgcgttaga cgccccgccc 3660

gtcacggggg gcgcggcggt gcctctgcgt ggggggggcgc ggggcgtccg gcgggggggcg 3720

ggcgggacgt agtctgctgc aagagacaac ggggggcgcg atcaggttac gccccctccc 3780

cggcccgccc tttcctcgcc cgcccgccca ttcctccctc ctcctcctcc cccagggtcc 3840

ttgccgcccc ccgcctcacc gtcgtccagg tcgtcgtcat cctcgtccgt ggtgggctcc 3900

gggtgggtgg gcgacagggc cctcaccgtg tgccccccca gggtcaggta ccgcggggcg 3960

aaccgctgat tgcccgtcca gataaagtcc acggccgtgc ccgccctgac ggcctcctcg 4020

gcctccatgc gggtctgggg gtcgttcacg atcgggatgg tgctgaacga cccgctgggc 4080

gtcacgccca ctatcaggta caccagcttg gcgttgcaca gcgggcaggt gttgcgcaat 4140

tgcatccagg ttttcatgca cgggatgcag aagcggtgca tgcacgggaa ggtgtcgcag 4200

cgcaggtggg gcgcgatctc atccgtgcac acggcgcaca cgtcgccctc gtcgctcccc 4260

ccgtcctctc gaggggggggc gcccccgcaa ctgccggggt cttcctcgcg gggggggctc 4320

ccccccgaga ccgcccccccc atccacgccc tgcggccccca gcagccccgt ctcgaacagt 4380

tccgtgtccg tgctgtccgc ctcggaggcg gagtcgtcgt catggtggtc ggcgtccccc 4440

cgcccccccca cttcggtctc cgcctcagag tcgctgctgt ccggcaggtc tcggtcgcag 4500

ggaaacaccc agacatccgg ggcgggctaa ggggaaaaaa gggggggcggg taagaatggg 4560

gggggatttc ccgcgtcaat cagcgcccac gagttccccc tctccccccc cgcctcacaa 4620

agtcctgccc ccctgctggc ctcggaagag gggggagaaa ggggtctgca accaaaggtg 4680

gtctgggtcc gtcctttgga tcccgacccc tcttcttccc tcttctcccg ccctccagac 4740

gcaccggagt cgggggtccc acggcgtccc ccaaatatgg cgggcggctc ctccccaccc 4800

ccctagatgc gtgtgagtaa ggggggcctg cgtatgagtc agtggggacc acgcccccaa 4860

cacggcgacc ccggtccctg tgtgtttgtt gtgggggcgt gtctctgtgt atgagtcagg 4920

gggtcccacg gcgaccccgg gccctgcgtc tgagtcaaag gggccatgtg tatgtgttgg 4980

gggtctgtat atataaagtc agggggtcac atggcgaccc ccaacagggc gaccccggtc 5040

cctgtatata tagggtcagg gggttccgcg ccccctaaca tggcgccccc ggtccctgta 5100

-continued

tatatagtgt cacggggttc cacgccccct aacatggcgc cccaacatgg cgcccggctc 5160

ccgtgtatga gtgggggtcc cccaacatgg gggccggttc cagggtaagg gtcgggggtc 5220

ccccaacatg gcgcccccca atatggcgcc ccagacatgg cgcccggccc ctcacctcgc 5280

gctgggggcg gccctcaggc cggcgggtac tcgctccggg gcggggctcc atgggggtcg 5340

tatgcggctg gagggtcgcg gacggagggt ccctgggggt cgcaacgtag gcggggcttc 5400

tgtggtgatg cggagagggg gcggcccgag tctgcctggc tgctgcgtct cgctccgagt 5460

gccgaggtgc aaatgcgacc agaccgtcgg gccagggcta acttataccc cacgcctttc 5520

ccctccccaa aggggcggca gtgacgattc ccccaatggc cgcgcgtccc aggggaggca 5580

ggcccaccgc ggggcggccc cgtccccggg gaccaacccg gcgcccccaa agaatatcat 5640

tagcatgcac ggcccggccc ccgatttggg ggaccaaccc ggtgtccccc aaagaacccc 5700

attagcatgc ccctcccgcc gacgcaacag gggcttggcc tgcgtcggtg ccccggggct 5760

tcccgccttc ccgaagaaac tcattaccat acccggaacc ccaggggacc aatgcgggtt 5820

cattgagcga cccgcgggcc aatgcgcgag gggccgtgtg ttccgccaaa aaagcaatta 5880

gcataacccg gaaccccagg ggagtggtta cgcgcggcgc gggaggcggg gaataccggg 5940

gttgcccatt aagggccgcg ggaattgccg gaagcgggaa gggcggccgg ggccgcccat 6000

taatgagttt ctaattacca taccgggaag cggaacaagg cctcttgcaa gtttttaatt 6060

accataccgg gaagtgggcg gcccggccca ttgggcggta actcccgccc aatgggccgg 6120

gccccgaaga ctcggcggac gctggttggc cgggccccgc cgcgctggcg gccgccgatt 6180

ggccagtccc gcccccgagg cgggcccgcc ttgggggcgg accggctccc agcgtatata 6240

tgcgcggctc ctgccatcgt ctctccggag agcggcttgg tgcggagctc ccgggagctc 6300

cgcggaagac ccaggcgcct cgggtgtaac gttagaccga gttcgccggg ccggctccgc 6360

gggccagggc ccgggcacgg gcctcgggcc ccaggcacgg cccgatgacc gcctcggcct 6420

ccgccacccg gcgccggaac cgagcccggt cggcccgctc gcgggcccac gagccgcggc 6480

gcgccaggcg ggcggccgag gcccagacca ccaggtggcg cacccggacg tggggcgaga 6540

agcgcacccg cgcgggggtc gcgggggtcg cgggggtcgc gggggtcgcg ggggtcgcgg 6600

gggtcgcggg ggtcgcgggg gtcgcggggg gctccggcgc cccctccccg cccgcgcgtc 6660

gcaggcgcag gcgcgccagg tgctccgcgg tgacgcgcag gcggagggcg aggcgcggcg 6720

gaaggcggaa ggggcgcgag ggggggtggg aggggtcagc cccgcccccc gggcccacgc 6780

cgggcggtgg ggaccggggc cgggggggcgg cggcggtggg ccgggcctct ggcgccggct 6840

cgggcggggg gctgtccggc cagtcgtcgt catcgtcgtc gtcggacgcg gactcgggaa 6900

cgtggagcca ctggcgcagc agcagcgaac aagaaggcgg gggcccaccg gcgggggggcg 6960

gcggcggggc ggccgcgggc gcgctcctga ccgcgggttc cgagttgggc gtggaggtta 7020

cctgggactg tgcggttggg acggcgcccg tgggcccggg cggccggggg cggcgggggc 7080

cgcgatggcg gcggcggcgg gccatggaga cagagagcgt gccggggtgg tagagtttga 7140

caggcaagca tgtgcgtgca gaggcgagta gtgcttgcct gtctaactcg ctagtctcgg 7200

ccgcgggggg cccgggctgc ccgccgccgc gctttaaagg gccgcgcgcg acccccgggg 7260

ggtgtgtttc gggggggccc gtttttgggg tctggccgct cctccccgc tcctccccgt 7320

ctgtgggtgg ggctcctccc ccgctcctcc cccgctcctc ccccgctcct ccccgtctgt 7380

-continued

gggtggggct cctcccccgc tcccgcggcc ccgcccccca cgcccgccgc gcgcgcgcac 7440

gccgcccgga ccgccgcccg ccttttttgc gcgccgcccc gcgcgcgggg ggcccgggct 7500

gccacaggtg taacaacacc aacagaacac caacagcacg gcgcaccggc gactccggtt 7560

cctcatccac acgtcacgtc atccaacaca cctgcccaac aacacaactc acagcgacaa 7620

ctcaccgcgc aacaactcct gttcctcatc cacacgtcac cgcgcacccc ccgctcctcc 7680

agacgtcccc cagcgcaaca cgccgctcct gtcacacacc acagccccag ccctccccag 7740

ccccagccct ccccagcccc agccctcccc agccccagcc ctccccagcc ccagccctcc 7800

ccagccccag ccctccccag ccccagccct ccccagcccc agccctcccc agccccagcc 7860

ctccccagcc ccagccctcc ccagccccag ccctccccag ccccagccct ccccagccgc 7920

gtcccgcgct ccctcggggg ggttcgggca tctctacctc agtgccgcca atctcaggtc 7980

agagatccaa accctccggg ggcgcccgcg caccaccacc gcccctcgcc ccctcccgcc 8040

cctcgccccc tcccgccccc cgcccccctcc cgcccctcgc cccctcccgc ccctcgcccc 8100

ctcccgcccc tcgcccccctc ccgcccctcg ccccctcccg cccctcgccc cctcccgccc 8160

ctcgaataaa 8170

<210> SEQ ID NO 20
<211> LENGTH: 2127
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 20

gacaggtcac taaccctcat actaccaagc cctgcggcac cctgccctag accaccactc 60

taaatgtctg ttccctccaa aaacaggacc cctgtcgcct attagggagg ggttctcttg 120

gaactgaccc acagtagggg gcaggacttt ggtgggttca agaactgcca tctcagcacc 180

tcagcccccc agtcctgccc tgcagtcgct ggcactaggc gggggcagac cctgggccac 240

aagttgctgc cacatggtcg ggataattga tgaaggtcca tccctccatt gctgtctcca 300

gccctgcctc tctggaaact ctatattttc cctttaatta tagcccctgc agtctccctc 360

tgctgcccca cccgcaccgc tcatcctggc tgcccacggc cagccggcca gccgacgtgg 420

ctccctcccc ttctgttcct ttttttccc ctttgccttc gttgcacaaa accagctggg 480

ggagggcgtg gagaggggcg gggggaggca atggaatctt ggatggtttg ggggaggcgg 540

gactccccgc ttccacgttt gcagctctgg agcacccggg gtggggagct gcacaggagg 600

gagagaaatg aacagggcac tgcaaggaga cccccaggcc ttctctcagc cctacagagt 660

ttctcaggac gaggtagatt ggggttgagg cagagccttg ttgggggaat gggacatgga 720

ggaagaaagg acgtggagtt ctagagccat cttccttaga tatagcctgc tgtccttcgg 780

gtccccagac cctttcagag tgtacagatg attctctctg gttcctaagg catagagcaa 840

tgaccgggat tttcaagaaa gagatgaggc agtgggaagt agcccctaaa acaaagtcaa 900

tcatcctctg cagcccatcc cacaccccca aaggaaagtt tcacccagac acccaaaata 960

tcccatacat ccccaacact gagtccaggt acaactggag aaggggcttt atgcagctcc 1020

cagaaagaca ccccctttagc taagtgccct ccctccaccc aggttctctc tggtttgact 1080

gtgctgggaa ggagggtctc taagcagccc ctggccacag ccatggcaaa caaaactctt 1140

ctctaagtca ccaatgatca caggcctccc actaaaaata cttcccaact ctggggtgga 1200

-continued

agagtttggg ggatgaattt ttaggggatt gcaagcccca atccccacct ctgtgtccct 1260

agaatccccc acccctacct tggctgctcc atcacccaac caccaaagct ttcttctgca 1320

gaggccacct agtcatgttt ctcaccctgc acctcagcct ccccactcca tctctcaatc 1380

atgcctaggg tttggaggaa ggcatttgat tctgttctgg agcacagcag aagaattgac 1440

atcctcaaaa ttaaaactcc cttgcctgca cccctccctc agatatctga ttcttaatgt 1500

ctagaaagga atctgtaaat tgttccccaa atattcctaa gctccatccc ctagccacac 1560

cagaagacac ccccaaacag gcacatcttt ttaattccca gcttcctctg ttttggagag 1620

gtcctcagca tgcctcttta tgcccctccc ttagctcttg ccaggatatc agagggtgac 1680

tggggcacag ccaggaggac cccctcccca acacccccaa cccttccacc tttggaagtc 1740

tccccaccca gctccccagt tccccagttc cacttcttct agattggagg tcccaggaag 1800

agagcagagg ggcaccccta cccactggtt agcccacgcc attctgagga cccagctgca 1860

cccctaccac agcacctctg gcccaggctg ggctgggggg ctggggaggc agagctgcga 1920

agaggggaga tgtggggtgg actcccttcc ctcctcctcc ccctctccat tccaactccc 1980

aaattggggg ccgggccagg cagctctgat tggctggggc acgggcggcc ggctcccccc 2040

ctccgagggg cagggttcct ccctgctctc catcaggaca gtataaaagg ggcccgggcc 2100

agtcgtcgga gcagacggga gtttctc 2127

<210> SEQ ID NO 21
<211> LENGTH: 1523
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 21

cgcggatccg cgcagacttg caaaagaagg catggcacct gtcgctgtgt gggcagctct 60

ggctgtcgga ctggagctgt gggcagcagc tcacgcactg ccagcacagg tggctttcac 120

tccttacgca ccagaacccg gctccacctg taggctgcgc gagtactatg atcagaccgc 180

tcagatgtgc tgttccaagt gctctcctgg gcagcatgca aaagtctttt gtacaaagac 240

ttctgacaca gtgtgcgatt cttgtgagga cagtacctac acacagctgt ggaactgggt 300

cccagaatgc ctgagttgtg gctcaagatg ctcctctgac caggtggaga ctcaggcctg 360

tacccgggaa cagaatagaa tctgcacctg tcgccccgga tggtattgcg ctctgtccaa 420

acaggaggga tgccgactgt gtgcacctct gcgcaagtgt cgaccagggt tcggagtggc 480

cagacctgga actgaaacca gcgatgtggt ctgcaagcca tgtgctcccg gcacattttc 540

caacactacc agttcaactg acatttgcag gccacaccag atctgtaacg tggtcgccat 600

tcccgggaat gcctctatgg acgctgtgtg cacatccact tctcccaccc gcagtatggc 660

acctggagca gtccacctgc ctcagccagt gagcacacgg tcccagcata cacagcccac 720

tcctgagcca agcacagcac caagtacttc atttctgctg cctatgggac catcaccacc 780

tgcagaggga agcaccggcg atgaacccaa aagctgcgac aagacccata catgcccacc 840

atgtccagca cctgagctgc tgggaggacc aagcgtgttc ctgtttcctc caaagcctaa 900

agatacactg atgatctcaa gaacccccga ggtcacatgt gtggtcgtgg atgtgagcca 960

cgaagaccct caggtcaagt tcaactggta cgtggacggc gtccaggtgc ataatgccaa 1020

gaccaaaccc cgggagcagc agtacaactc cacctataga gtcgtgtctg tcctgacagt 1080

-continued

gctgcaccag aactggctgg acgggaagga atataagtgc aaagtgtcca acaaggctct 1140

gccagcaccc attgagaaga caatctctaa ggccaaagga cagccccgag aacctcaggt 1200

gtacactctg ccccttccc gggaggaaat gaccaaaaac caggtctctc tgacatgtct 1260

ggtgaagggg ttctatccat cagatatcgc cgtggagtgg gaaagcaatg gacagcccga 1320

gaacaattac aagacaactc cacccgtcct ggactctgat ggaagtttct ttctgtattc 1380

caaactgacc gtggacaagt ctaggtggca gcagggcaac gtcttcagct gctccgtgat 1440

gcacgaggcc ctgcacaatc attacactca gaaatctctg agtctgtcac ccggcaagtg 1500

agacatcagg gcgcggatcc gcg 1523

<210> SEQ ID NO 22
<211> LENGTH: 965
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 22

cggagctccc gggagctccg caccaagccg ctctccggag agacgatggc aggagccgcg 60

catatatacg ctgggagccg gtccgccccc aaggcgggcc cgcctcgggg gcgggactgg 120

ccaatcggcg gccgccagcg cggcggggcc cggccaacca gcgtccgccg agtcttcggg 180

gcccggccca ttgggcggga gttaccgccc aatgggccgg gccgcccact tcccggtatg 240

gtaattaaaa acttgcaaga ggccttgttc cgcttcccgg tatggtaatt agaaactcat 300

taatgggcgg ccccggccgc ccttcccgct tccggcaatt cccgcggccc ttaatgggca 360

accccggtat tccccgcctc ccgcgccgcg cgtaaccact cccctggggt tccgggttat 420

gctaattgct tttttggcgg aacacacggc ccctcgcgca ttggcccgcg ggtcgctcaa 480

tgaacccgca ttggtcccct ggggttccgg gtatggtaat gagtttcttc gggaaggcgg 540

gaagccccgg ggcaccgacg caggccaagc ccctgttgcg tcggcgggag gggcatgcta 600

atggggttct ttgggggaca ccgggttggt cccccaaatc gggggccggg ccgtgcatgc 660

taatgatatt ctttgggggc gccgggttgg tccccgggga cggggccgcc ccgcggtggg 720

cctgcctccc ctgggacgcg cggccattgg gggaatcgtc actgccgccc ctttggggag 780

gggaaaggcg tggggtataa gttagccctg gcccgacggt ctggtcgcat ttgcacctcg 840

gcactcggag cgagacgcag cagccaggca gactcgggcc gccccctctc cgcatcacca 900

cagaagcccc gcctacgttg cgacccccag ggaccctccg tccgcgaccc tccagccgca 960

tacga 965

<210> SEQ ID NO 23
<211> LENGTH: 49
<212> TYPE: DNA
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Synthetic

<400> SEQUENCE: 23

aataaaatat ctttattttc attacatctg tgtgttggtt ttttgtgtg 49

<210> SEQ ID NO 24
<211> LENGTH: 9
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1

-continued

<400> SEQUENCE: 24

taatgarat 9

<210> SEQ ID NO 25
<211> LENGTH: 3903
<212> TYPE: DNA
<213> ORGANISM: human herpesvirus 1 strain McKrae

<400> SEQUENCE: 25

atggcgtcgg agaacaagca gcgccccggc tccccgggcc ccaccgacgg gccgccgccc 60

accccgagcc cagaccgcga cgagcggggg gccctcgggt ggggcgcgga gacggaggag 120

ggcggggacg accccgacca cgaccccgac caccccacg acctcgacga cgcccggcgg 180

gacgggaggg cccccgcggc gggcaccgac gccggcgagg acgccgggga cgccgtctcg 240

ccgcgacagc tggccctgct ggcctccatg gtagaggagg ccgtccggac gatcccgacg 300

cccgaccccg cggcctcgcc gccccggacc cccgcctttc gagccgacga cgatgacggg 360

gacgagtacg acgacgcagc cgacgccgcc ggcgaccggg ccccggcccg gggccgcgca 420

cgggaggccc cgctacgcgg cgcgtatccg gaccccacgg accgcctgtc gccgcgcccg 480

ccggcccagc cgccgcagag acgtcgtcac ggccggcggc ggccatcggc gtcatcgacc 540

tcgtcggact ccgggtcctc gtcctcgtcg tccgcatcct cttcgtcctc gtcgtccgac 600

gaggacgagg acgacgacgg caacgacgcg gccgaccacg cacgcgaggc gcgggccgtc 660

gggcggggtc cgtcgagcgc ggcgccggaa gcccccgggc ggacgccgcc cccgcccggg 720

ccacccccc tctccgaggc cgcgcccaag ccccgggcgg cggcgaggac ccccgcggcc 780

tccgcgggcc gcatcgagcg ccgccgggcc cgcgcggcgg tggccggccg cgacgccacg 840

ggccgcttca cggccgggca gccccggcgg gtcgagctgg acgccgacgc ggcctccggc 900

gccttctacg cgcgctatcg cgacgggtac gtcagcgggg agccgtggcc cggcgccggg 960

cccccgcccc cggggcgggt gctgtacggc ggcctgggcg acagccgccc gggcctctgg 1020

ggggcgcccg aggcggagga ggcgcgacgc cggttcgagg cctcgggcgc cccggcggcc 1080

gtgtgggcgc ccgagctggg cgacgccgcg cagcagtacg ccctgatcac gcggctgctg 1140

tacaccccgg acgcggaggc catggggtgg ctccagaacc cgcgcgtggt ccccggggac 1200

gtggcgctgg accaggcctg cttccggatc tcgggcgccg cgcgcaacag cagctccttc 1260

atcaccggca gcgtggcgcg ggccgtgccc cacctgggct acgccatggc ggccggccgc 1320

ttcggctggg gcctggcgca cgcggcggcc gccgtggcca tgagccgccg atacgaccgc 1380

gcgcagaagg gcttcctgct gaccagcctg cgccgcgcct acgcgcccct gttggcgcgc 1440

gagaacgcgg cgctgacggg ggccgcgggg agccccggcg ccggcgcaga tgacgagggg 1500

gtcgccgccg ccgtcgtcgc cgccgccgcc gcaccgggcg agcgcgcggt gcccgccggg 1560

tacggcgccg cggggatcct cgccgccctg gggcggctgt ccgccgcgcc cgcctccccc 1620

gcgggggggcg acgaccccga cgccgcccgc cacgccgacg ccgacgacga cgccgggcgc 1680

cgcgcccagg ccggccgcgt ggccgtggag tgcctggccg cctgccgcgg gatcctggag 1740

gcgctggccg agggcttcga cggcgacctg gcggccgtcc cggggctggc cggggcccgg 1800

-continued

cccgccagcc ccccgcggcc ggagggaccc gcgggccccg cttccccgcc gccgccgcac 1860

gccgacgcgc cccgcctgcg cgcgtggctg cgcgagctgc ggttcgtgcg cgacgcgctg 1920

gtgctcatgc gcctgcgcgg ggacctgcgc gtggccggcg gcagcgaggc cgccgtggcc 1980

gccgtgcgcg ccgtgagcct ggtcgccggg gccctgggtc ccgcgctgcc gcgggacccg 2040

cgcctgccga gctccgcggc cgccgccgcc gcggacctgc tgtttgagaa ccagagcctg 2100

cgcccctgc tggcggcggg tccgcgccgc tcttcttcgt cttcgggggt cgcggccgcc 2160

gcctccgccg cgccgcggga ggggcgcaag cgcaagagtc ccggcccggc ccggccgccc 2220

ggaggcggcg gcccgcgacc cccgaagacg aagaagagcg gcgcggacgc ccccggctcg 2280

gacgcccgcg ccccctccc cgcgcccgcg ccccctcca cgcccccggg gcccgagccc 2340

gcccccgccc agcccgcggc gccccgggcc gccgcggcgc aggcccgccc gcgccccgtg 2400

gcgctgtcgc gccggcccgc cgagggcccc gacccctgg gcggctggcg gcggcagccc 2460

ccggggccca gccacacggc ggcgcccgcg gccgccgccc tggaggccta ctgctccccg 2520

cgcgccgtgg ccgagctcac ggaccacccg ctgttccccg tcccctggcg accggccctc 2580

atgtttgacc cgcgggccct ggcctcgatc gccgcgcggt gcgccgggcc cgcccccgcc 2640

gcccaggccg cgtgcggcgg cggcgacgac gacgagaacc cccaccccca cggggccgcc 2700

gggggccgcc tctttggccc cctgcgcgcc tcgggcccgc tgcgccgcat ggcggcctgg 2760

atgcgccaga tccccgaccc cgaggacgtg cgcgtggtgg tgctgtactc gccgctgccg 2820

ggcgaggacc tggccggcgg cggggcctcg ggggggccgc cggagtggtc cgccgagcgc 2880

ggcgggctgt cctgcctgct ggcggccctg gccaaccggc tgtgcgggcc ggacacggcc 2940

gcctgggcgg gcaactggac cggcgccccc gacgtgtcgg cgctgggcgc gcagggcgtg 3000

ctgctgctgt ccacgcggga cctggccttc gccggggccg tggagtttct ggggctgctc 3060

gccagcgccg gcgaccggcg gctcatcgtg gtcaacaccg tgcgcgcctg cgactggccc 3120

gccgacgggc ccgcggtgtc gcggcagcac gcctacctgg cgtgcgacct gctgcccgcc 3180

gtgcagtgcg ccgtgcgctg gccggcggcg cgcgacctgc gccgcacggt gctggcctcg 3240

ggccgcgtgt tcggcccggg ggtcttcgcg cgcgtggagg ccgcgcacgc gcgcctgtac 3300

cccgacgcgc cgccgctgcg cctgtgccgc ggcggcaacg tgcgctaccg cgtgcgcacg 3360

cgcttcggcc cggacacgcc ggtgcccatg tccccgcgcg agtaccgccg ggccgtgctg 3420

ccggcgctgg acggccgggc ggcggcctcg gggaccaccg acgccatggc gcccggcgcg 3480

ccggacttct gcgaggagga ggcccactcg caccgcgcct gcgcgcgctg gggcctgggc 3540

gcgccgctgc ggcccgtgta cgtggcgctg gggcgcgagg cggtgcgcgc cggcccggcc 3600

cggtggcgcg ggccgcggag ggacttttgc gcccgcgccc tgctggagcc cgacgacgac 3660

gccccccgc tggtgctgcg cggcgacgac gacggcccgg gggccctgcc gccggcgccg 3720

cccgggattc gctgggcctc ggccacgggc cgcagcggca ccgtgctggc ggcggcgggg 3780

gccgtggagg tgctgggggc ggaggcgggc ttggccacgc ccccgcgacg ggacgttgtg 3840

gactgggaag gcgcctggga cgaagacgac ggcggcgcgt tcgaggggga cggggtgctg 3900

taa 3903

What is claimed is:

1. A method of expressing a polypeptide in a tissue of a subject comprising administering to the subject a vector comprising a variant of a herpes simplex virus (HSV) strain whose genome contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins.

2. The method of claim **1**, wherein the HSV strain is an HSV-1 strain.

3. The method of claim **1**, wherein the HSV strain is a McKrae strain.

4. The method of claim **1**, wherein the variant fails to express functional ICP4 and ICP0 proteins characterized by the amino acid sequences of SEQ ID NO: 2 and SEQ ID NO: 17, respectively.

5. The method of claim **1**, wherein the tissue comprises epidermis, dermis, or subcutaneous fat or muscle.

6. The method of claim **1**, wherein the vector comprises a tissue specific promoter.

7. The method of claim **6**, wherein the tissue is skin.

8. The method of claim **6**, wherein the tissue specific promoter is a collagen 1 promoter.

9. The method of claim **1**, wherein the vector comprises a human cytomegalovirus (HCMV) enhancer.

10. The method of claim **1**, wherein the vector comprises a bovine growth hormone (BGH) polyadenylation signal or an HSV viral polyadenylation signal.

11. The method of claim **1**, further comprising a nucleic acid that encodes a therapeutic polypeptide.

12. The method of claim **1**, wherein the tissue is skin tissue.

13. The method of claim **12**, wherein the skin tissue comprises fibroblasts, keratinocytes, adipocytes, or muscle cells.

14. The method of claim **1**, wherein the vector is administered in vivo.

15. The method of claim **1**, wherein the vector is administered by contact with skin.

16. The method of claim **1**, wherein the vector is administered by intradermal injection.

17. A method of propagating a vector comprising a variant herpes simplex virus (HSV) genome which contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins, the method comprising steps of:

(a) infecting cultured ICP0 and ICP4 complementing cells containing DNA encoding HSV proteins ICP0 and ICP4 with the vector, and

(b) isolating supernatant from the culture of step (a).

18. The method of claim **17**, further comprising a step of purifying vector in the supernatant by chromatography.

19. The method of claim **18**, further comprising a step of concentrating the purified vector by tangential flow filtration.

20. A method of preparing a vector comprising a variant herpes simplex virus (HSV) genome which contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins, and wherein the vector expresses a marker element, the method comprising incubating cells transfected with:

(a) a first nucleic acid molecule:

(i) comprising a portion of HSV genome but does not encode functional ICP0 and ICP4 proteins; and

(ii) comprising a first homology region (HR1) and a second homology region (HR2), and

(b) a second nucleic acid molecule comprising a sequence that encodes a marker element, wherein the sequence is flanked by a first homology region (HR1') and a second homology region (HR2'),

wherein HR1 is homologous to HR1' and HR2 is homologous to HR2' such that the sequence that encodes the marker element in the second nucleic acid molecule integrates into the first nucleic acid molecule via homologous recombination.

21. The method of claim **20**, wherein the cells are ICP0 and/or ICP4 complementing cells.

22. The method of claim **20**, wherein the marker element is a polypeptide.

23. The method of claim **22**, wherein the polypeptide is a soluble tumor necrosis factor receptor.

24. The method of claim **22**, wherein the polypeptide is quantified by enzyme linked immunosorbent assay (ELISA).

25. The method of claim **22**, wherein the polypeptide is detected by fluorescence.

26. The method of claim **20**, further comprising a step of purifying viral plaques that express the marker element.

27. A method of preparing a vector comprising a variant herpes simplex virus (HSV) genome which contains an alteration such that the variant fails to express functional ICP0 and ICP4 proteins, and wherein the vector expresses an agent of interest, the method comprising incubating cells transfected with:

(a) a first nucleic acid molecule:

(i) comprising a portion of HSV genome but does not encode functional ICP0 and ICP4 proteins; and

(ii) comprising a sequence that encodes a marker element, wherein the sequence that encodes the marker element is flanked by a first homology region (HR1) and a second homology region (HR2); and

(b) a second nucleic acid molecule comprising a sequence that encodes an agent of interest, wherein the sequence encoding the agent of interest is flanked by a first homology region (HR1') and a second homology region (HR2'),

wherein HR1 is homologous to HR1' and HR2 is homologous to HR2' such the sequence encoding the agent of interest is integrated into the first nucleic acid molecule via homologous recombination.

28. The method of claim **27**, wherein the cells are ICP0 and/or ICP4 complementing cells.

29. The method of claim **27**, further comprising a step of purifying viral plaques that do not express the marker element.

30. The method of any of claims **17-29**, wherein the HSV genome is an HSV-1 genome.

31. The method of any of claims **17-29**, wherein the HSV genome is a McKrae strain genome.

32. A variant of a herpes simplex virus (HSV) strain whose genome contains an alteration such that the variant fails to express one or more functional ICP0 and ICP4 proteins.

33. The variant HSV strain of claim **32**, wherein the HSV strain is an HSV-1 strain.

34. The variant HSV strain of claim **32**, wherein the HSV strain is a McKrae strain.

35. An HSV strain comprising a variant herpes simplex virus (HSV) strain genome which contains an alteration such that the variant fails to express functional ICP4 and ICP0

proteins characterized by the amino acid sequences of SEQ ID NO: 2 and SEQ ID NO: 17, respectively.

36. A vector comprising the variant HSV strain of any of claims **32-35**.

37. The vector of claim **36**, wherein the vector comprises a tissue specific promoter.

38. The vector of claim **37**, wherein the tissue is skin.

39. The vector of claim **37**, wherein the tissue specific promoter is a collagen 1 promoter.

40. The vector of any of claims **36-39**, wherein the vector comprises a human cytomegalovirus (HCMV) enhancer.

41. The vector of any of claims **36-40**, wherein the vector comprises a bovine growth hormone (BGH) polyadenylation signal or an HSV viral polyadenylation signal.

42. The vector of any of claims **36-41**, further comprising a nucleic acid that encodes a therapeutic polypeptide.

43. A method of measuring transduction efficiency of an HSV vector in a skin tissue, the method comprising:

(a) contacting the skin tissue of an animal with a vector according to any one of claims **36-42**;

(b) removing a portion of the skin tissue from the animal; and

(c) assaying the number of HSV genomes transduced in the skin tissue.

44. The method of claim **43**, wherein the skin tissue comprises fibroblasts, keratinocytes, adipocytes, muscle cells, epidermis, dermis, hypodermis, or underlying subcutaneous fat or muscle.

45. The method of claim **43** or claim **44**, wherein the number of genomes is measured by an amplification technique.

46. The method of claim **45**, wherein the amplification technique is quantitative polymerase chain reaction (qPCR).

47. A method of measuring transduction efficiency of an HSV vector that contains an expression cassette comprising a polypeptide payload in a skin tissue, the method comprising:

(a) contacting the skin tissue of an animal with a vector according to any one of claims **36-42**;

(b) removing a portion of the skin tissue from the animal; and

(c) assaying the amount of a polypeptide encoded by a nucleic acid of the expression cassette.

48. The method of claim **47**, wherein the skin tissue comprises fibroblasts, keratinocytes, adipocytes, muscle cells, epidermis, dermis, hypodermis, or underlying subcutaneous fat or muscle.

49. The method of claim **47** or claim **48**, wherein the amount of polypeptide is measured by an immunoassay.

50. The method of claim **49**, wherein the immunoassay is an enzyme linked immunosorbent assay (ELISA) or immunohistochemistry (IHC).

51. The method of claim **50**, wherein the ELISA or IHC is performed on tissue of the epidermis, dermis, subcutaneous tissue, subcutaneous fat, underlying muscle, or draining lymph node.

52. A cell transduced with a vector according to any one of claims **36-42**.

53. A pharmaceutical composition comprising a vector according to any one of claims **36-41** and a pharmaceutically acceptable carrier.

54. The method of any one of claim **1-31** or **43-51**, the variant of any one of claims **32-34**, the HSV strain of claim **35**, the vector of any one of claims **36-42**, the cell of claim **52**, or the pharmaceutical composition of claim **53**, wherein the alteration is a disruption or a deletion of the ICP0 and ICP4 genes.

55. The method, variant, HSV strain, vector, cell, or pharmaceutical composition of claim **54**, wherein the ICP0 and/or ICP4 gene is disrupted by deletion of its respective promoter.

56. The method of any one of claim **1-31**, **43-51**, **54**, or **55**, the variant of any one of claim **32-34**, **54**, or **55**, the HSV strain of any one of claim **35**, **54**, or **55**, the vector of any one of claim **36-42**, **54**, or **55**, the cell of any one of claim **52**, **54**, or **55**, or the pharmaceutical composition of any one of claim **53**, **54**, or **55**, wherein the variant is a replication-defective variant.

* * * * *