FIG. 1

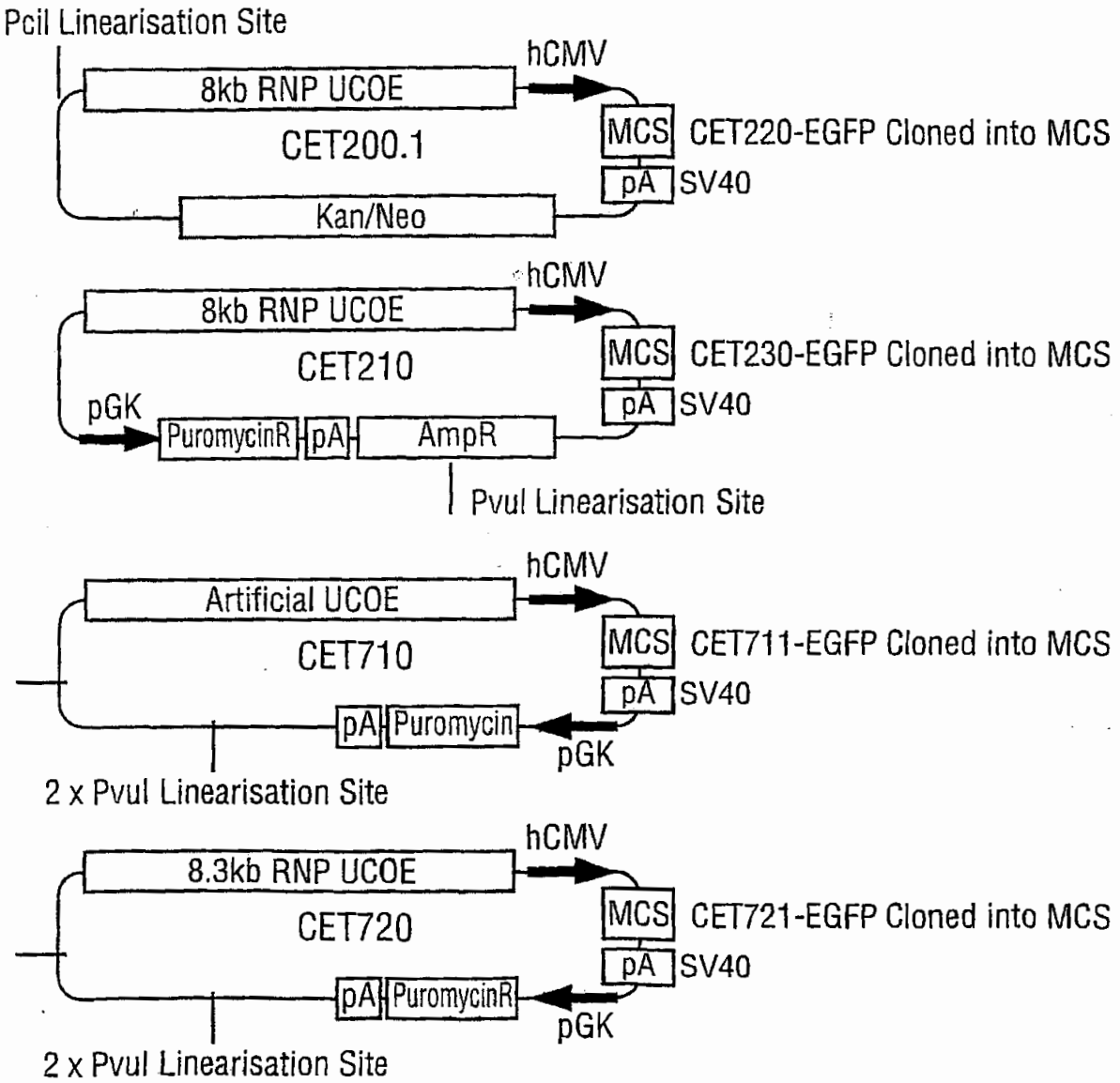

PciI Linearisation Site
hCMV
8kb RNP UCOE
CET200.1
MCS
CET220-EGFP Cloned into MCS
pA
SV40
Kan/Neo
hCMV
8kb RNP UCOE
CET210
MCS
CET230-EGFP Cloned into MCS
pA
SV40
pGK
PuromycinR
pA
AmpR
PvuI Linearisation Site
hCMV
Artificial UCOE
CET710
MCS
CET711-EGFP Cloned into MCS
pA
SV40
pA
Puromycin
pGK
2 x PvuI Linearisation Site
hCMV
8.3kb RNP UCOE
CET720
MCS
CET721-EGFP Cloned into MCS
pA
SV40
pA
PuromycinR
pGK
2 x PvuI Linearisation Site

FIG. 2

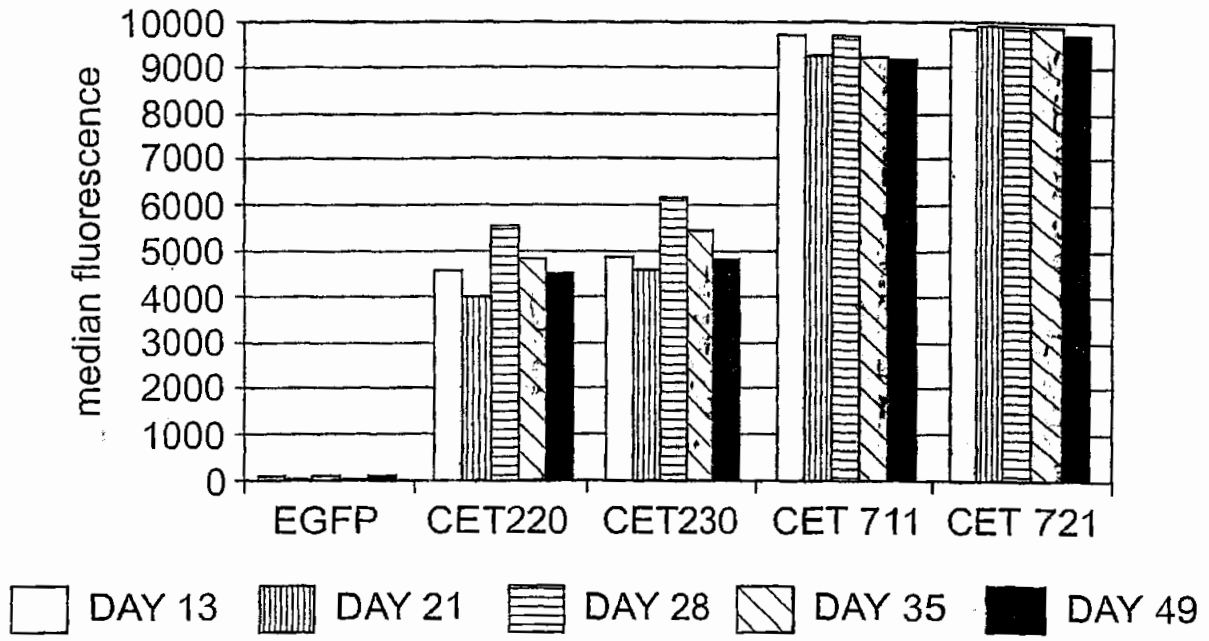

median fluorescence
10000
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
EGFP
CET220
CET230
CET 711
CET 721
DAY 13
DAY 21
DAY 28
DAY 35
DAY 49

FIG. 4

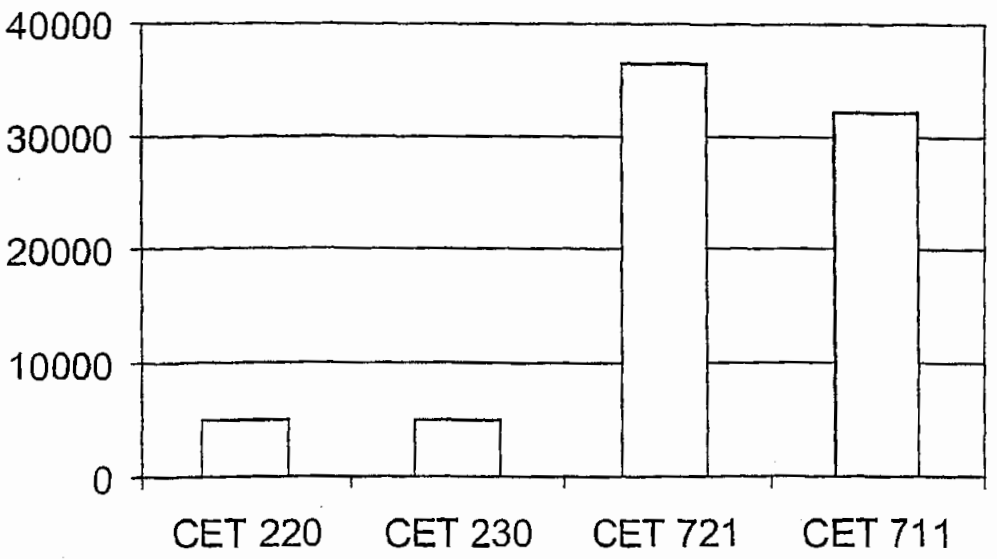

40000
30000
20000
10000
0
CET 220
CET 230
CET 721
CET 711

FIG. 3

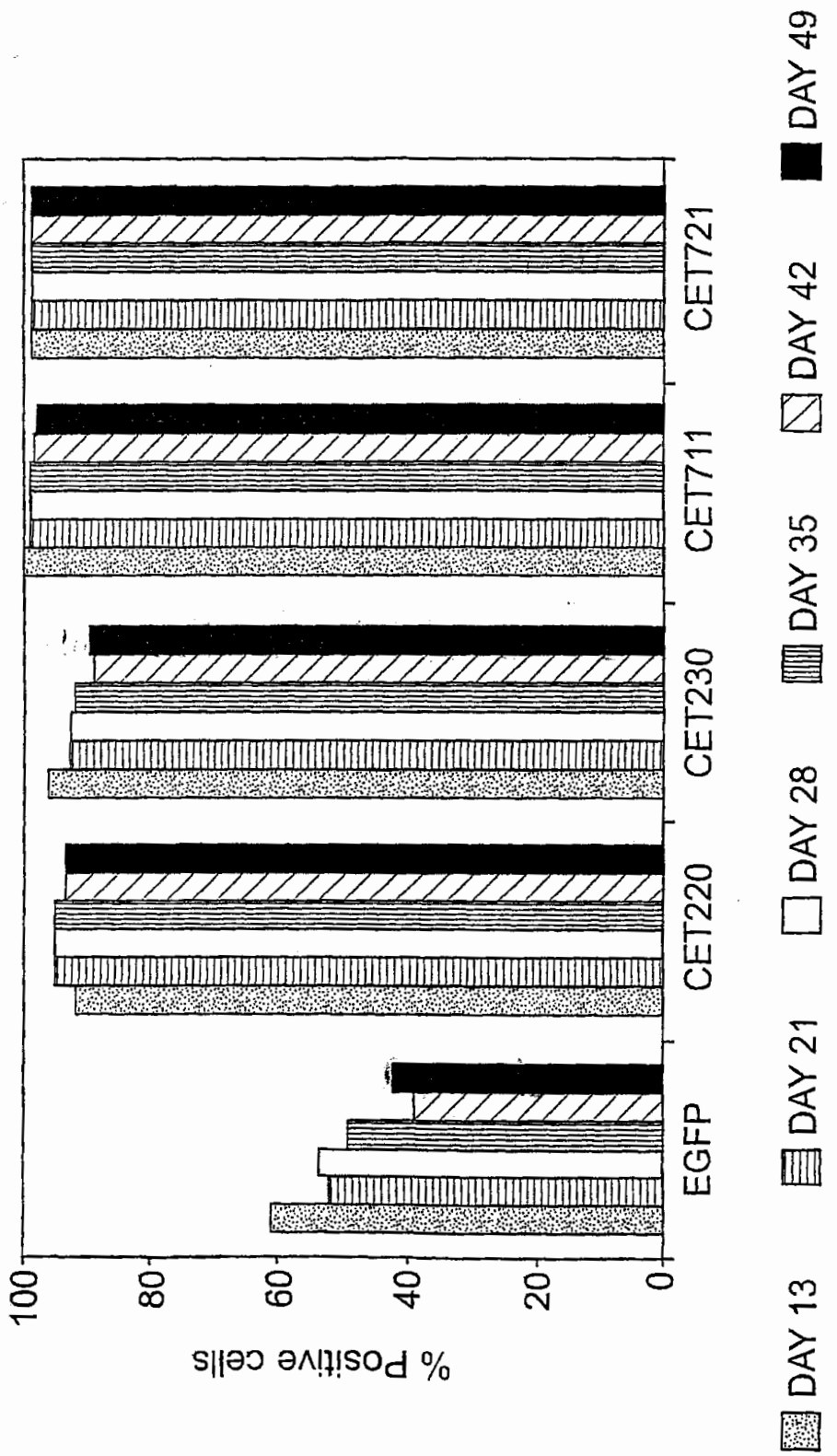

% Positive cells
100
80
60
40
20
0
EGFP
CET220
CET230
CET711
CET721
DAY 13
DAY 21
DAY 28
DAY 35
DAY 42
DAY 49

FIG. 5

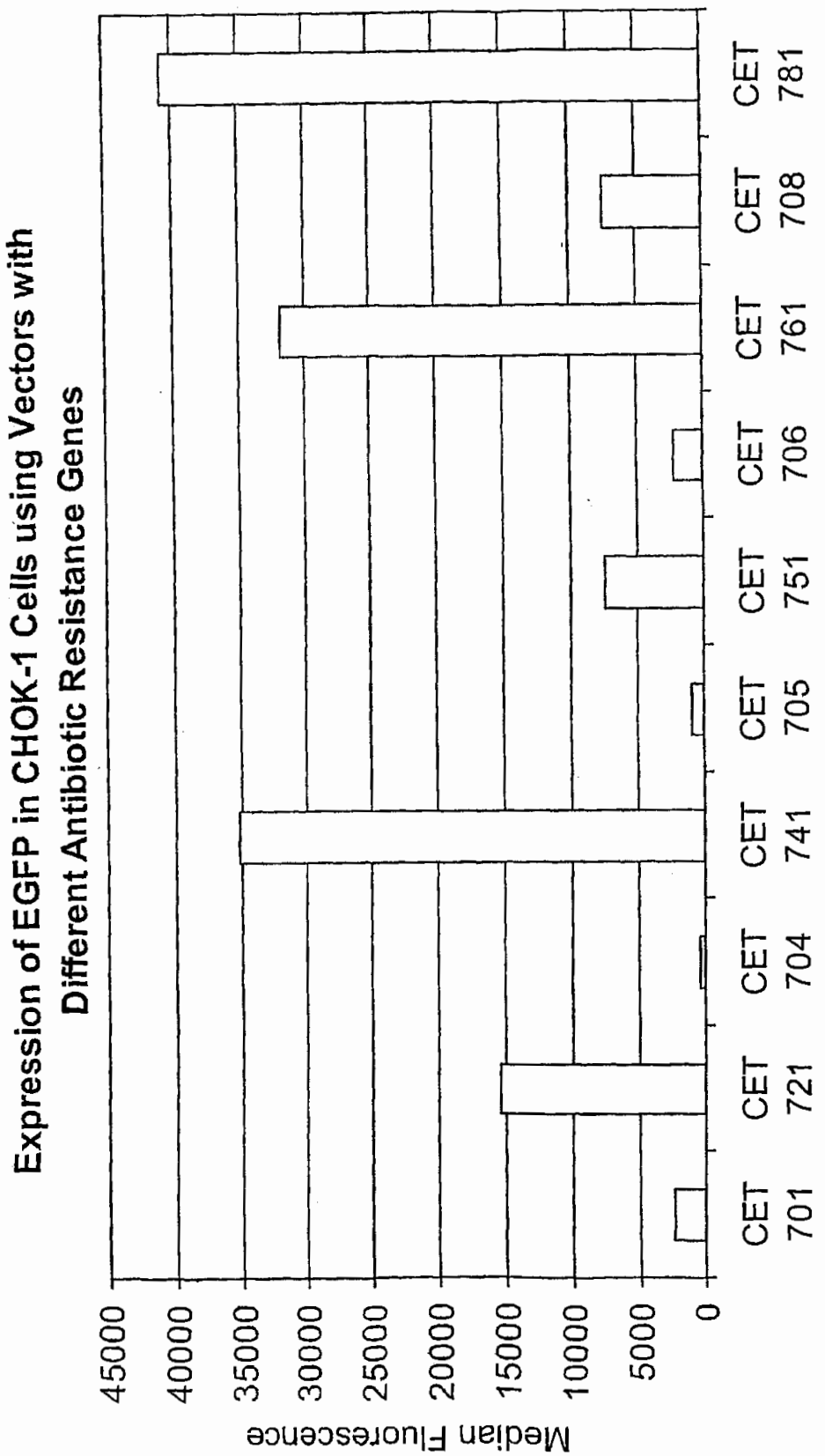

Expression of EGFP in CHOK-1 Cells using Vectors with Different Antibiotic Resistance Genes
Median Fluorescence
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
CET 701
CET 721
CET 704
CET 741
CET 705
CET 751
CET 706
CET 761
CET 708
CET 781

FIG. 6

Expression Levels of EGFP in CHOK-1 cells transfected with vectors encoding the Puromycin Gene

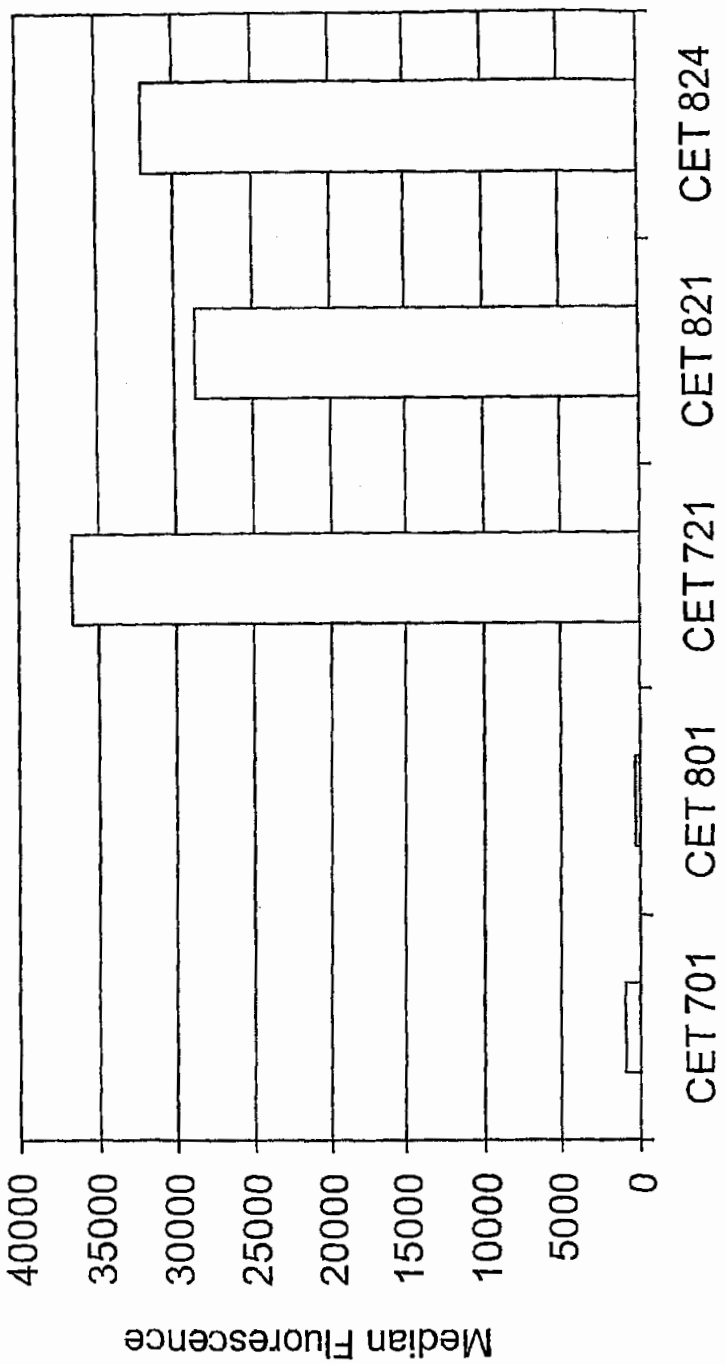
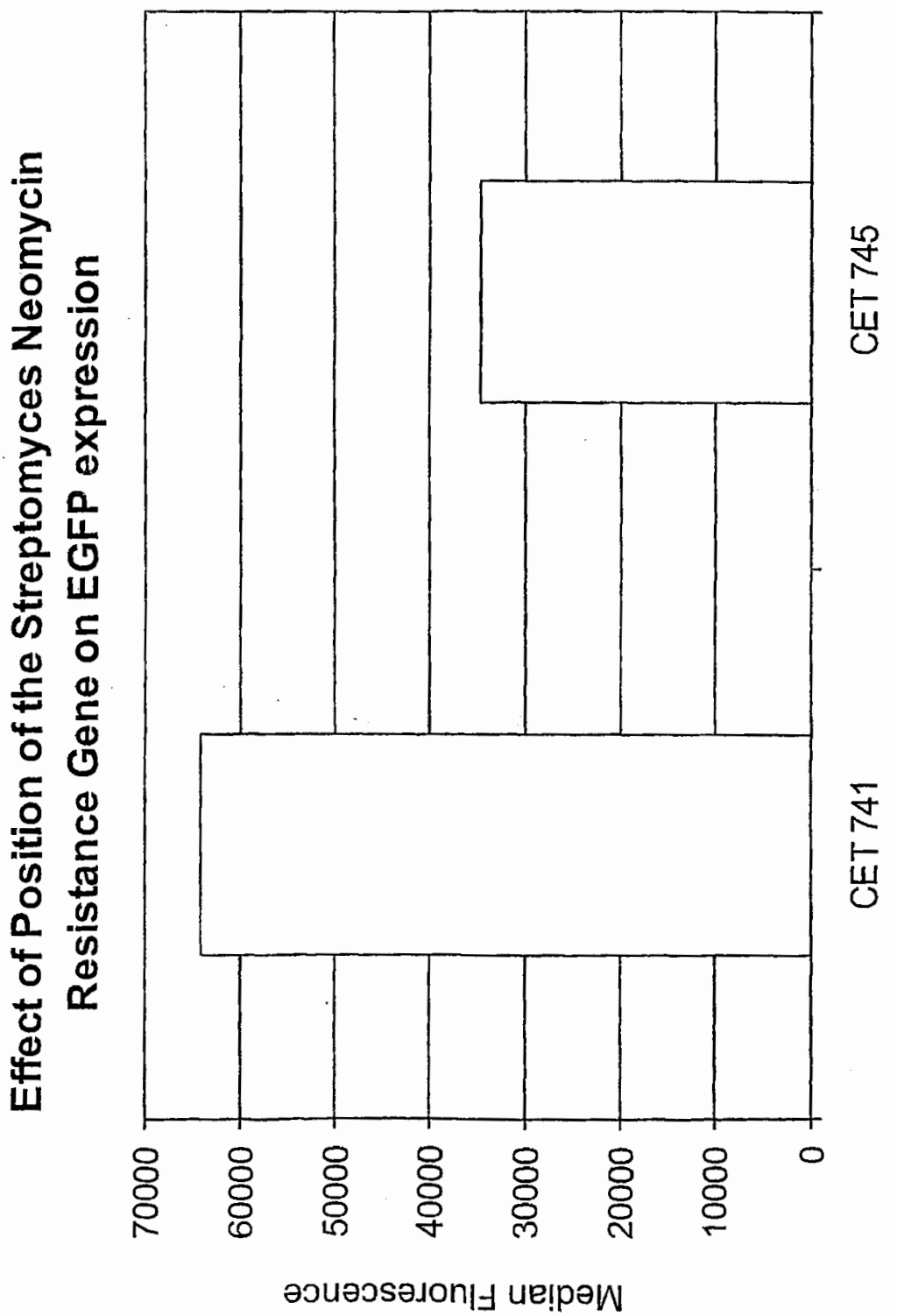


FIG. 7


Effect of Position of the Streptomyces Neomycin Resistance Gene on EGFP expression
Median Fluorescence
70000
60000
50000
40000
30000
20000
10000
0
CET 741
CET 745

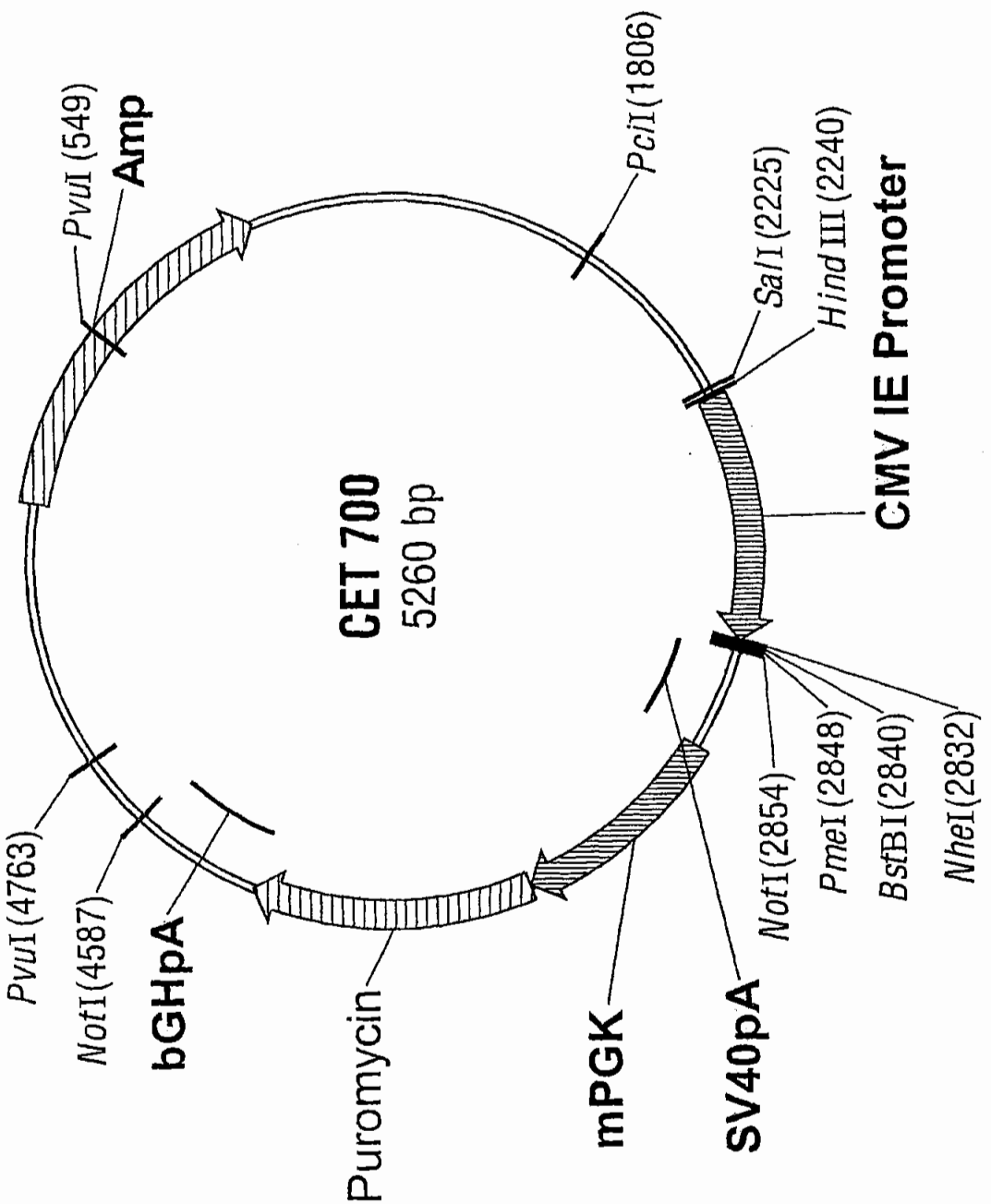

CET 700
5260 bp
PvuI (549)
Amp
PciI(1806)
SalI(2225)
HindIII(2240)
CMV IE Promoter
NheI(2832)
BstBI(2840)
PmeI(2848)
NotI(2854)
SV40pA
mPGK
Puromycin
bGHpA
NotI(4587)
PvuI(4763)

FIG. 8

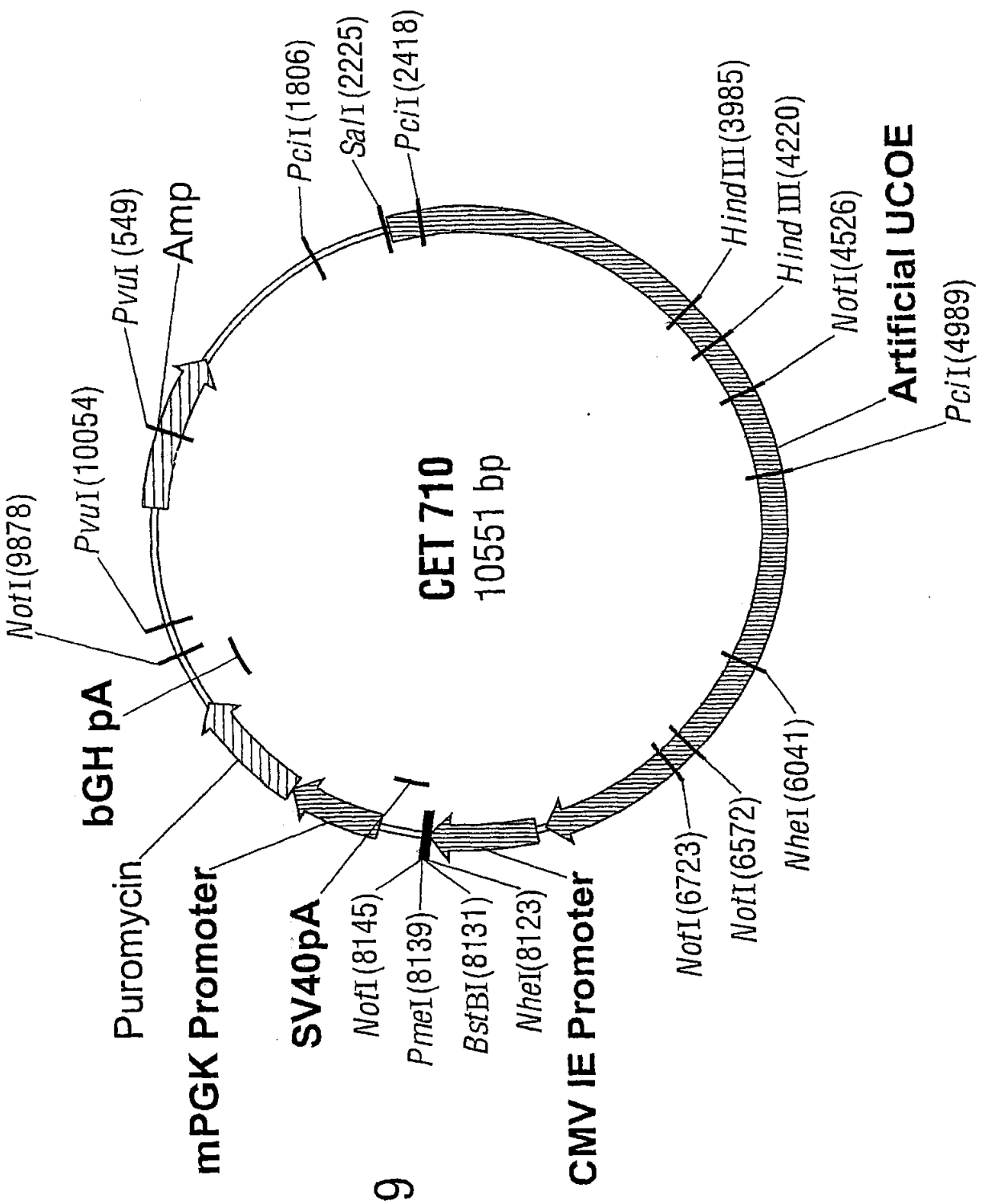

CET 710
10551 bp
Artificial UCOE
PciI(4989)
NotI(4526)
HindIII(4220)
HindIII(3985)
PciI(2418)
SalI(2225)
PciI(1806)
Amp
PvuI(549)
PvuI(10054)
NotI(9878)
bGH pA
Puromycin
mPGK Promoter
SV40pA
NotI(8145)
PmeI(8139)
BstBI(8131)
NheI(8123)
CMV IE Promoter
NotI(6723)
NotI(6572)
NheI(6041)

FIG. 9

FIG.10 CET710 nucleotide sequence

```
   1  GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT GTTTATTTTT
      CCACCGTGAA AAGCCCCTTT ACACGCGCCT TGGGGATAAA CAAATAAAAA
  51  CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA CCCTGATAAA
      GATTTATGTA AGTTTATACA TAGGCGAGTA CTCTGTTATT GGGACTATTT
 101  TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT
      ACGAAGTTAT TATAACTTTT TCCTTCTCAT ACTCATAAGT TGTAAAGGCA
 151  GTCGCCCTTA TTCCCTTTTT TGCGGCATTT TGCCTTCCTG TTTTTGCTCA
      CAGCGGGAAT AAGGGAAAAA ACGCCGTAAA ACGGAAGGAC AAAAACGAGT
 201  CCCAGAAACG CTGGTGAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC
      GGGTCTTTGC GACCACTTTC ATTTTCTACG ACTTCTAGTC AACCCACGTG
 251  GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT CCTTGAGAGT
      CTCACCCAAT GTAGCTTGAC CTAGAGTTGT CGCCATTCTA GGAACTCTCA
 301  TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA AAGTTCTGCT
      AAAGCGGGGC TTCTTGCAAA AGGTTACTAC TCGTGAAAAT TTCAAGACGA
 351  ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC
      TACACCGCGC CATAATAGGG CATAACTGCG GCCCGTTCTC GTTGAGCCAG
 401  GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA
      CGGCGTATGT GATAAGAGTC TTACTGAACC AACTCATGAG TGGTCAGTGT
 451  GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC
      CTTTTCGTAG AATGCCTACC GTACTGTCAT TCTCTTAATA CGTCACGACG
 501  CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG
      GTATTGGTAC TCACTATTGT GACGCCGGTT GAATGAAGAC TGTTGCTAGC
 551  GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG GGATCATGTA
      CTCCTGGCTT CCTCGATTGG CGAAAAAACG TGTTGTACCC CCTAGTACAT
 601  ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA
      TGAGCGGAAC TAGCAACCCT TGGCCTCGAC TTACTTCGGT ATGGTTTGCT
 651  CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAAC
      GCTCGCACTG TGGTGCTACG GACATCGTTA CCGTTGTTGC AACGCGTTTG
 701  TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC
      ATAATTGACC GCTTGATGAA TGAGATCGAA GGGCCGTTGT TAATTATCTG
 751  TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC
      ACCTACCTCC GCCTATTTCA ACGTCCTGGT GAAGACGCGA GCCGGGAAGG
 801  GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC
      CCGACCGACC AAATAACGAC TATTTAGACC TCGGCCACTC GCACCCAGAG
 851  GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA
      CGCCATAGTA ACGTCGTGAC CCCGGTCTAC CATTCGGGAG GGCATAGCAT
 901  GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
      CAATAGATGT GCTGCCCCTC AGTCCGTTGA TACCTACTTG CTTTATCTGT
 951  GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC
      CTAGCGACTC TATCCACGGA GTGACTAATT CGTAACCATT GACAGTCTGG
1001  AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA TTTTTAATTT
      TTCAAATGAG TATATATGAA ATCTAACTAA ATTTTGAAGT AAAAATTAAA
1051  AAAAGGATCT AGGTGAAGAT CCTTTTTGAT AATCTCATGA CCAAAATCCC
      TTTTCCTAGA TCCACTTCTA GGAAAAACTA TTAGAGTACT GGTTTTAGGG
1101  TTAACGTGAG TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA
      AATTGCACTC AAAAGCAAGG TGACTCGCAG TCTGGGGCAT CTTTTCTAGT
1151  AAGGATCTTC TTGAGATCCT TTTTTTCTGC GCGTAATCTG CTGCTTGCAA
      TTCCTAGAAG AACTCTAGGA AAAAAAGACG CGCATTAGAC GACGAACGTT
1201  ACAAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG ATCAAGAGCT
      TGTTTTTTTG GTGGCGATGG TCGCCACCAA ACAAACGGCC TAGTTCTCGA
1251  ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA
      TGGTTGAGAA AAAGGCTTCC ATTGACCGAA GTCGTCTCGC GTCTATGGTT
1301  ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT
      TATGACAGGA AGATCACATC GGCATCAATC CGGTGGTGAA GTTCTTGAGA
```

FIG.10 (contd.)

```
1351  GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC
      CATCGTGGCG GATGTATGGA GCGAGACGAT TAGGACAATG GTCACCGACG
1401  TGCCAGTGGC GATAAGTCGT GTCTTACCGG GTTGGACTCA AGACGATAGT
      ACGGTCACCG CTATTCAGCA CAGAATGGCC CAACCTGAGT TCTGCTATCA
1451  TACCGGATAA GGCGCAGCGG TCGGGCTGAA CGGGGGGTTC GTGCACACAG
      ATGGCCTATT CCGCGTCGCC AGCCCGACTT GCCCCCCAAG CACGTGTGTC
1501  CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA
      GGGTCGAACC TCGCTTGCTG GATGTGGCTT GACTCTATGG ATGTCGCACT
1551  GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC
      CGATACTCTT TCGCGGTGCG AAGGGCTTCC CTCTTTCCGC CTGTCCATAG
1601  CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG
      GCCATTCGCC GTCCCAGCCT TGTCCTCTCG CGTGCTCCCT CGAAGGTCCC
1651  GGAAACGCCT GGTATCTTTA TAGTCCTGTC GGGTTTCGCC ACCTCTGACT
      CCTTTGCGGA CCATAGAAAT ATCAGGACAG CCCAAAGCGG TGGAGACTGA
1701  TGAGCGTCGA TTTTTGTGAT GCTCGTCAGG GGGGCGGAGC CTATGGAAAA
      ACTCGCAGCT AAAAACACTA CGAGCAGTCC CCCCGCCTCG GATACCTTTT
1751  ACGCCAGCAA CGCGGCCTTT TTACGGTTCC TGGCCTTTTG CTGGCCTTTT
      TGCGGTCGTT GCGCCGGAAA AATGCCAAGG ACCGGAAAAC GACCGGAAAA
1801  GCTCACATGT TCTTTCCTGC GTTATCCCCT GATTCTGTGG ATAACCGTAT
      CGAGTGTACA AGAAAGGACG CAATAGGGGA CTAAGACACC TATTGGCATA
1851  TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC
      ATGGCGGAAA CTCACTCGAC TATGGCGAGC GGCGTCGGCT TGCTGGCTCG
1901  GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT ACGCAAACCG
      CGTCGCTCAG TCACTCGCTC CTTCGCCTTC TCGCGGGTTA TGCGTTTGGC
1951  CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC ACGACAGGTT
      GGAGAGGGGC GCGCAACCGG CTAAGTAATT ACGTCGACCG TGCTGTCCAA
2001  TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC
      AGGGCTGACC TTTCGCCCGT CACTCGCGTT GCGTTAATTA CACTCAATCG
2051  TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG
      AGTGAGTAAT CCGTGGGGTC CGAAATGTGA AATACGAAGG CCGAGCATAC
2101  TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA
      AACACACCTT AACACTCGCC TATTGTTAAA GTGTGTCCTT TGTCGATACT
2151  CCATGATTAC GCCAAGCGCG CAATTAACCC TCACTAAAGG GAACAAAAGC
      GGTACTAATG CGGTTCGCGC GTTAATTGGG AGTGATTTCC CTTGTTTTCG
2201  TGGGTACCGG GCCCCCCCTC GAGGTCGACG GTATCGATAA GCTCGATAAG
      ACCCATGGCC CGGGGGGGAG CTCCAGCTGC CATAGCTATT CGAGCTATTC
2251  CTCATGGCAC CTGTATTGTA CTCTTATCAG TCATTATATG GACTTTAACT
      GAGTACCGTG GACATAACAT GAGAATAGTC AGTAATATAC CTGAAATTGA
2301  TCCCCAGATA TTATTTGGGC TCCTCCATAA GACTGTGAGC ATCTGACCAC
      AGGGGTCTAT AATAAACCCG AGGAGGTATT CTGACACTCG TAGACTGGTG
2351  TGGAGTGTTG CTTCCCATTA TATCCCTGTT ATCAAGCACA AGGTCAGGCA
      ACCTCACAAC GAAGGGTAAT ATAGGGACAA TAGTTCGTGT TCCAGTCCGT
2401  CAGAGTAAGA CTCAAAACAT GTTTTGGAAT GTATGACTGG TATGAACTAC
      GTCTCATTCT GAGTTTTGTA CAAAACCTTA CATACTGACC ATACTTGATG
2451  AAACCAGTAA GCTGATGTTT TCATTTTGAG TCTATAAATC TAATTTTGTG
      TTTGGTCATT CGACTACAAA AGTAAAACTC AGATATTTAG ATTAAAACAC
2501  GTGGTTTTGT GTATGGCTCA AGGCTCAAAT TGTAAAATTT AATATTATGT
      CACCAAAACA CATACCGAGT TCCGAGTTTA ACATTTTAAA TTATAATACA
2551  GACCAAAGAA AGTTATACCC AGAACCTCAA TTTCCTCACC TTCAAAATGG
      CTGGTTTCTT TCAATATGGG TCTTGGAGTT AAAGGAGTGG AAGTTTTACC
2601  GGCAGTTTCT CACTCATTGG TCTGCTGTCA CGATTTTAAT GAGCTCATGC
      CCGTCAAAGA GTGAGTAACC AGACGACAGT GCTAAAATTA CTCGAGTACG
2651  ACAAACAGCC CTTTATATAA GGTAAGTGCT GGATAAATGT TGGCTACTAT
      TGTTTGTCGG GAAATATATT CCATTCACGA CCTATTTACA ACCGATGATA
```

FIG.10 (contd.)

```
2701  AATAAAATAA GCCTCTAAGA TACTTGGTCA GCACAAGTAC TACCCAAGAG
      TTATTTTATT CGGAGATTCT ATGAACCAGT CGTGTTCATG ATGGGTTCTC
2751  TATGCACTGT AAGTAAACTG ACAAAATTGT GTATCTAAAA CTGGCCAGAT
      ATACGTGACA TTCATTTGAC TGTTTTAACA CATAGATTTT GACCGGTCTA
2801  GAAAGAGAAA CTTTTAAGGG GCCCTTCTGC GTGCCCGACA CTGTGCTAGG
      CTTTCTCTTT GAAAATTCCC CGGGAAGACG CACGGGCTGT GACACGATCC
2851  CACTCACACT ATCCCGACCC GAGAAACCGA TCTGCGACCC AGAGGAACTT
      GTGAGTGTGA TAGGGCTGGG CTCTTTGGCT AGACGCTGGG TCTCCTTGAA
2901  ACCAAGCCTC CAGCATCTTG TGCAGCCCTA CTCATGGGAC CATCTGGATA
      TGGTTCGGAG GTCGTAGAAC ACGTCGGGAT GAGTACCCTG GTAGACCTAT
2951  CCCACCCTTG TCTTTACAGG GAGCAGAACA CACCTCTTAT GTGTCAGAAA
      GGGTGGGAAC AGAAATGTCC CTCGTCTTGT GTGGAGAATA CACAGTCTTT
3001  ACAAAGTCCA GGAAGTATAT TTTTACCTGA GGCAATATCT GAAAATTGTA
      TGTTTCAGGT CCTTCATATA AAAATGGACT CCGTTATAGA CTTTTAACAT
3051  TGCTACAGCC TCCAAAGTGA GTCTTCCTCT CAGTACCTCT CTTCTAGGCA
      ACGATGTCGG AGGTTTCACT CAGAAGGAGA GTCATGGAGA GAAGATCCGT
3101  CATGGAGCCC TTTCTTCCAA GTATTATGTT TAACCACTTA ATGAATGAAG
      GTACCTCGGG AAAGAAGGTT CATAATACAA ATTGGTGAAT TACTTACTTC
3151  TCCTGAAACT GCTTACCCAT GCTCCCTATA ATCTCTGAGT AATCTTCCTT
      AGGACTTTGA CGAATGGGTA CGAGGGATAT TAGAGACTCA TTAGAAGGAA
3201  TTCCACAACC TCAGGCATAA TCTCATCTTC TGTTTCTATT ACAATTTCAA
      AAGGTGTTGG AGTCCGTATT AGAGTAGAAG ACAAAGATAA TGTTAAAGTT
3251  ATTCTGGAAA AAGGAAGTTG TGGTCTGGAA TTATATGGTC CAGATGATCT
      TAAGACCTTT TTCCTTCAAC ACCAGACCTT AATATACCAG GTCTACTAGA
3301  GAAACAAAAA GGACAGCACT ATTAGTAATC ATTTAGTTTT GAAGACAGTC
      CTTTGTTTTT CCTGTCGTGA TAATCATTAG TAAATCAAAA CTTCTGTCAG
3351  TAATAATTTG CTGTCTCTAA AGTACTATAT TCCCTATAGT TCTGGCATTT
      ATTATTAAAC GACAGAGATT TCATGATATA AGGGATATCA AGACCGTAAA
3401  TAGATAAAGG GTCATAAATT AAATGCCTAT ATGGTGACAT TATTCAGTGA
      ATCTATTTCC CAGTATTTAA TTTACGGATA TACCACTGTA ATAAGTCACT
3451  TTCAGACTTC ACAGCCTTTT TTTTTTTTTT ACAAAGGTGT TCCAGGCATG
      AAGTCTGAAG TGTCGGAAAA AAAAAAAAAA TGTTTCCACA AGGTCCGTAC
3501  AAAAATTTTA AAGTACTATA CCTTTCCTAA TTTTACCTTT AAAGTTGTCC
      TTTTTAAAAT TTCATGATAT GGAAAGGATT AAAATGGAAA TTTCAACAGG
3551  TGGAAATATC TGGGTTGACA AAGGCGATGA AACTGAACTG AGACTTAAAA
      ACCTTTATAG ACCCAACTGT TTCCGCTACT TTGACTTGAC TCTGAATTTT
3601  AAAAGATTAC CCACCTGGTT GTGCACAAGC CTGCTTATGT CCCAATCTCC
      TTTTCTAATG GGTGGACCAA CACGTGTTCG GACGAATACA GGGTTAGAGG
3651  AGTCTAGGGT CTGATGCTCC TTGCTGCAGT AATATGCTTT GTGGCATCTG
      TCAGATCCCA GACTACGAGG AACGACGTCA TTATACGAAA CACCGTAGAC
3701  GAGCACGTTT TGGGGCCTAA ACAGCCACAA ACCCTGCAGA GATGAGCACC
      CTCGTGCAAA ACCCCGGATT TGTCGGTGTT TGGGACGTCT CTACTCGTGG
3751  AGACTTAAGC TGGAGACACA CTGATTCTCC TGTTTCTGGG GGAGGATTCT
      TCTGAATTCG ACCTCTGTGT GACTAAGAGG ACAAAGACCC CCTCCTAAGA
3801  CAGAAGGTGG CTCATATGAG TAAAAATCGT TTTTCCTGGG TAGTTGATTC
      GTCTTCCACC GAGTATACTC ATTTTTAGCA AAAAGGACCC ATCAACTAAG
3851  CTAAAAACTA AAAAAGAATA CAGAGAAAAG TTTTATCTTC AAACAAAACA
      GATTTTTGAT TTTTTCTTAT GTCTCTTTTC AAAATAGAAG TTTGTTTTGT
3901  GCAATTCACA TATTTTATCC TCTGCACGTA AAACTGAAAA TAACAACAAC
      CGTTAAGTGT ATAAAATAGG AGACGTGCAT TTTGACTTTT ATTGTTGTTG
3951  AAAAAAGAAA TGAAAGTTTT TGCTTTCAGG AATAAGCTTT TAAAATCCAG
      TTTTTTCTTT ACTTTCAAAA ACGAAAGTCC TTATTCGAAA ATTTTAGGTC
4001  AAACTAGATT TCGTCCGGTA CACGCAACTG AGTTGCCTCC TAGAGGTGGT
      TTTGATCTAA AGCAGGCCAT GTGCGTTGAC TCAACGGAGG ATCTCCACCA
```

FIG.10 (contd.)

```
4051  TTGAGTTAAT CAAATTAATA AGACTGATCG TTAAGAACGA CTGCCAAAAA
      AACTCAATTA GTTTAATTAT TCTGACTAGC AATTCTTGCT GACGGTTTTT
4101  TACGAAAAAG CTACTGGGAT CCATCTTTCC AAGACAATTT CTATTATCTG
      ATGCTTTTTC GATGACCCTA GGTAGAAAGG TTCTGTTAAA GATAATAGAC
4151  AATTAACACC ATACCTGGTA CCCACTGATT AAAAGCTGGG GGTTACCAAT
      TTAATTGTGG TATGGACCAT GGGTGACTAA TTTTCGACCC CCAATGGTTA
4201  GCGCGTGGGC ACAGTTAGAA GCTTATGTAG CAAAAATGAG CACATCCTGG
      CGCGCACCCG TGTCAATCTT CGAATACATC GTTTTTACTC GTGTAGGACC
4251  AAGGGCCCGG GAGAAGGTGC TCCTGGGGCA GCGCGGAGAG GGAGCTCTGA
      TTCCCGGGCC CTCTTCCACG AGGACCCCGT CGCGCCTCTC CCTCGAGACT
4301  GGCTGGGGCG GCAGCGGTGC TTGCCGCCGT CCCCCTGGTC GCTCCCGGAA
      CCGACCCCGC CGTCGCCACG AACGGCGGCA GGGGGACCAG CGAGGGCCTT
4351  TTAACGCCGC GCACGCGTCG GAGGCATGGC CCCGTCCCGA CCCCGTTTGG
      AATTGCGGCG CGTGCGCAGC CTCCGTACCG GGGCAGGGCT GGGGCAAACC
4401  CGGCTCACCT CGCAGGCCGG CACAGCACGG CTGCTCGCGG CAGCAGAAGA
      GCCGAGTGGA GCGTCCGGCC GTGTCGTGCC GACGAGCGCC GTCGTCTTCT
4451  GGAAGATGCA GCGGTGGAAG GCGTCCGGGC GGCCAGGCAG CGGCGCATAC
      CCTTCTACGT CGCCACCTTC CGCAGGCCCG CCGGTCCGTC GCCGCGTATG
4501  ACCTGCAGCA GGAAGGAGAG CGGGCGGCCG CACAGCTCGC AGGCCAGGGC
      TGGACGTCGT CCTTCCTCTC GCCCGCCGGC GTGTCGAGCG TCCGGTCCCG
4551  CTGGGGCCCC GGCAGCCCGG CCGCGCCCAG CCATGCCGGC CGCCCGCCCA
      GACCCCGGGG CCGTCGGGCC GGCGCGGGTC GGTACGGCCG GCGGGCGGGT
4601  CCTTGCTGGG GAACTGCTCG CTGCGCAGTC GCCACGCCGG CGCCGACTCG
      GGAACGACCC CTTGACGAGC GACGCGTCAG CGGTGCGGCC GCGGCTGAGC
4651  GCGAAGCCCA GCTCCACAGG CCTGGCCCCG GCGGCAGCCA TGCGGGGCGC
      CGCTTCGGGT CGAGGTGTCC GGACCGGGGC CGCCGTCGGT ACGCCCCGCG
4701  GGGCTGGCGT GGGGCGCAGC CCACAGCTGG GTCGGAAGGC GGAAATCGGG
      CCCGACCGCA CCCCGCGTCG GGTGTCGACC CAGCCTTCCG CCTTTAGCCC
4751  CGCCGGGCCG GAAGGCAAGA GGCGGGCACC TTTCCGGAGG ACAGGAGGCG
      GCGGCCCGGC CTTCCGTTCT CCGCCCGTGG AAAGGCCTCC TGTCCTCCGC
4801  GAAACGCGTC TGACGGGAGC GGTTGCAGGA CCAATGCGAG GGAACGGGGC
      CTTTGCGCAG ACTGCCCTCG CCAACGTCCT GGTTACGCTC CCTTGCCCCG
4851  AGAGGAAACC TCTCGGCATC AGCCCCGCCC CTGGCGCCTC TGCCTCCGAG
      TCTCCTTTGG AGAGCCGTAG TCGGGGCGGG GACCGCGGAG ACGGAGGCTC
4901  CCGCTTTCCT GGTGCCTCCG GGTGCTCTGG GATGGTTCTG GTCTTTGGGA
      GGCGAAAGGA CCACGGAGGC CCACGAGACC CTACCAAGAC CAGAAACCCT
4951  GAGTGGCAGC TGGTGACGGC GCTCCGCTCA CCTCTGCACA TGTCTTGCTG
      CTCACCGTCG ACCACTGCCG CGAGGCGAGT GGAGACGTGT ACAGAACGAC
5001  TGGGCCTGCG GGTGGCCGCC AGGGAGGCAG AGCCCTCCCG CAAACCTTCC
      ACCCGGACGC CCACCGGCGG TCCCTCCGTC TCGGGAGGGC GTTTGGAAGG
5051  CTGCTGGTGT CCACCTCAGG GTGTGGGAAA CCTGTGCGCT GGCCGAGTGC
      GACGACCACA GGTGGAGTCC CACACCCTTT GGACACGCGA CCGGCTCACG
5101  TAACCAAGAG TAGGCAGTGA AAGACAAATG AAGGTTGAAC AGGTAAAGTG
      ATTGGTTCTC ATCCGTCACT TTCTGTTTAC TTCCAACTTG TCCATTTCAC
5151  AGGACCCTAC AGCGGAAACC AAGAATCCTG TGTGCCTGAG AGTAATGAAG
      TCCTGGGATG TCGCCTTTGG TTCTTAGGAC ACACGGACTC TCATTACTTC
5201  AAGCCTCTGC AGAAGAGTCT TTTCTGTCAG TCTTAAGGTC TCTGTTTTAA
      TTCGGAGACG TCTTCTCAGA AAAGACAGTC AGAATTCCAG AGACAAAATT
5251  TGTTAGTGCT GGCTTGCTGT ACCTGAATTC CAAGGGAGGA GTGTATAATG
      ACAATCACGA CCGAACGACA TGGACTTAAG GTTCCCTCCT CACATATTAC
5301  AGGCATGGCC AACCCCCACT TCCCATCATT GCCTGAACTA GTTTTTCAGG
      TCCGTACCGG TTGGGGGTGA AGGGTAGTAA CGGACTTGAT CAAAAAGTCC
5351  TTAACTTCAG AATGCCCTTG GTACCGCGGG CCCCCTCTGT GGTCCCACGC
      AATTGAAGTC TTACGGGAAC CATGGCGCCC GGGGGAGACA CCAGGGTGCG
```

FIG.10 (contd.)

```
5401  CACTGATCGC TGCATGCCCA CCACCTGGGT ACACACAGTC TGTGATTCCC
      GTGACTAGCG ACGTACGGGT GGTGGACCCA TGTGTGTCAG ACACTAAGGG
5451  GGAGCAGAAC GGACCCTGCC CACCCGGTCT TGTGTGCTAC TCAGTGGACA
      CCTCGTCTTG CCTGGGACGG GTGGGCCAGA ACACACGATG AGTCACCTGT
5501  GACCCAAGGC AAGAAAGGGT GACAAGGACA GGGTCTTCCC AGGCTGGCTT
      CTGGGTTCCG TTCTTTCCCA CTGTTCCTGT CCCAGAAGGG TCCGACCGAA
5551  TGAGTTCCTA GCACCGCCCC GCCCCCAATC CTCTGTGGCA CATGGAGTCT
      ACTCAAGGAT CGTGGCGGGG CGGGGGTTAG GAGACACCGT GTACCTCAGA
5601  TGGTCCCCAG AGTCCCCCAG CGGCCTCCAG ATGGTCTGGG AGGGCAGTTC
      ACCAGGGGTC TCAGGGGGTC GCCGGAGGTC TACCAGACCC TCCCGTCAAG
5651  AGCTGTGGCT GCGCATAGCA GACATACAAC GGACGGTGGG CCCAGACCCA
      TCGACACCGA CGCGTATCGT CTGTATGTTG CCTGCCACCC GGGTCTGGGT
5701  GGCTGTGTAG ACCCAGCCCC CCCGCCCCGC AGTGCCTAGG TCACCCACTA
      CCGACACATC TGGGTCGGGG GGGCGGGGCG TCACGGATCC AGTGGGTGAT
5751  ACGCCCCAGG CCTGGTCTTG GCTGGGCGTG ACTGTTACCC TCAAAAGCAG
      TGCGGGGTCC GGACCAGAAC CGACCCGCAC TGACAATGGG AGTTTTCGTC
5801  GCAGCTCCAG GGTAAAAGGT GCCCTGCCCT GTAGAGCCCA CTTCCTTCCC
      CGTCGAGGTC CCATTTTCCA CGGGACGGGA CATCTCGGGT GAAGGAAGGG
5851  AGGGCTGCGG CTGGGTAGGT TTGTAGCCTT CATCACGGGC CACCTCCAGC
      TCCCGACGCC GACCCATCCA AACATCGGAA GTAGTGCCCG GTGGAGGTCG
5901  CACTGGACCG CTGGCCCCTG CCCTGTCCTG GGGAGTGTGG TCCTGCGACT
      GTGACCTGGC GACCGGGGAC GGGACAGGAC CCCTCACACC AGGACGCTGA
5951  CTAATGGCCG CAAGCCACCT GACTCCCCCA ACACCACACT CTACCTCTCA
      GATTACCGGC GTTCGGTGGA CTGAGGGGGT TGTGGTGTGA GATGGAGAGT
6001  AGCCCAGGTC TCTCCCTAGT GACCCACCCA GCACATTTAG CTAGCTGAGC
      TCGGGTCCAG AGAGGGATCA CTGGGTGGGT CGTGTAAATC GATCGACTCG
6051  CCCACAGCCA GAGGTCCTCA GGCCCTGCTT TCAGGGCAGT TGCTCTGAAG
      GGGTGTCGGT CTCCAGGAGT CCGGGACGAA AGTCCCGTCA ACGAGACTTC
6101  TCGGCAAGGG GGAGTGACTG CCTGGCCACT CCATGCCCTC CAAGAGCTCC
      AGCCGTTCCC CCTCACTGAC GGACCGGTGA GGTACGGGAG GTTCTCGAGG
6151  TTCTGCAGGA GCGTACAGAA CCCAGGGCCC TGGCACCCGT GCAGACCCTG
      AAGACGTCCT CGCATGTCTT GGGTCCCGGG ACCGTGGGCA CGTCTGGGAC
6201  GCCCACCCCA CCTGGGCGCT CAGTGCCCAA GAGATGTCCA CACCTAGGAT
      CGGGTGGGGT GGACCCGCGA GTCACGGGTT CTCTACAGGT GTGGATCCTA
6251  GTCCCGCGGT GGGTGGGGGG CCCGAGAGAC GGGCAGGCCG GGGGCAGGCC
      CAGGGCGCCA CCCACCCCCC GGGCTCTCTG CCCGTCCGGC CCCCGTCCGG
6301  TGGCCATGCG GGGCCGAACC GGGCACTGCC CAGCGTGGGG CGCGGGGGCC
      ACCGGTACGC CCCGGCTTGG CCCGTGACGG GTCGCACCCC GCGCCCCCGG
6351  ACGGCGCGCG CCCCCAGCCC CCGGGCCCAG CACCCCAAGG CGGCCAACGC
      TGCCGCGCGC GGGGGTCGGG GGCCCGGGTC GTGGGGTTCC GCCGGTTGCG
6401  CAAAACTCTC CCTCCTCCTC TTCCTCAATC TCGCTCTCGC TCTTTTTTTT
      GTTTTGAGAG GGAGGAGGAG AAGGAGTTAG AGCGAGAGCG AGAAAAAAAA
6451  TTTCGCAAAA GGAGGGGAGA GGGGGTAAAA AAATGCTGCA CTGTGCGGCG
      AAAGCGTTTT CCTCCCCTCT CCCCCATTTT TTTACGACGT GACACGCCGC
6501  AAGCCGGTGA GTGAGCGGCG CGGGGCCAAT CAGCGTGCGC CGTTCCGAAA
      TTCGGCCACT CACTCGCCGC GCCCCGGTTA GTCGCACGCG GCAAGGCTTT
6551  GTTGCCTTTT ATGGCTCGAG CGGCCGCGGC GGCGCCCTAT AAAACCCAGC
      CAACGGAAAA TACCGAGCTC GCCGGCGCCG CCGCGGGATA TTTTGGGTCG
6601  GGCGCGACGC GCCACCACCG CCGAGACCGC GTCCGCCCGC GAGCACAGAG
      CCGCGCTGCG CGGTGGTGGC GGCTCTGGCG CAGGCGGGCG CTCGTGTCTC
6651  CCTCGCCTTT GCCGATCCGC CGCCCGTCCA CACCCGCCGC CAGGTAAGCC
      GGAGCGGAAA CGGCTAGGCG GCGGGCAGGT GTGGGCGGCG GTCCATTCGG
6701  CGGCCAGCCG ACCGGGGCAT GCGGCCGCGG CCCTTCGCCC GTGCAGAGCC
      GCCGGTCGGC TGGCCCCGTA CGCCGGCGCC GGGAAGCGGG CACGTCTCGG
```

FIG.10 (contd.)

```
6751  GCCGTCTGGG CCGCAGCGGG GGGCGCATGG GGCGGAACCG GACCGCCGTG
      CGGCAGACCC GGCGTCGCCC CCCGCGTACC CCGCCTTGGC CTGGCGGCAC
6801  GGGGGCGCGG GAGAAGCCCC TGGGCCTCCG GAGATGGGGG ACACCCCACG
      CCCCCGCGCC CTCTTCGGGG ACCCGGAGGC CTCTACCCCC TGTGGGGTGC
6851  CCAGTTCGCA GGCGCGAGGC CGCGCTCGGG CGGGCGCGCT CCGGGGGTGC
      GGTCAAGCGT CCGCGCTCCG GCGCGAGCCC GCCCGCGCGA GGCCCCCACG
6901  CGCTCTCGGG GCGGGGGCAA CCGGCGGGGT CTTTGTCTGA GCCGGGCTCT
      GCGAGAGCCC CGCCCCCGTT GGCCGCCCCA GAAACAGACT CGGCCCGAGA
6951  TGCCAATGGG GATCGCACGG TGGGCGCGGC GTAGCCCCCG TCAGGCCCGG
      ACGGTTACCC CTAGCGTGCC ACCCGCGCCG CATCGGGGGC AGTCCGGGCC
7001  TGGGGGCTGG GGCGCCATGC GCGTGCGCGC TGGTCCTTTG GGCGCTAACT
      ACCCCCGACC CCGCGGTACG CGCACGCGCG ACCAGGAAAC CCGCGATTGA
7051  GCGTGCGCGC TGGGAATTGG CGCTAATTGC GCGTGCGCGC TGGGACTCAA
      CGCACGCGCG ACCCTTAACC GCGATTAACG CGCACGCGCG ACCCTGAGTT
7101  TGGCGCTAAT CGCGCGTGCG TTCTGGGGCC CGGGCGCTTG CGCCACTTCC
      ACCGCGATTA GCGCGCACGC AAGACCCCGG GCCCGCGAAC GCGGTGAAGG
7151  TGCCCGAGCC GCTGGCGCCC GAGGGTGTGG CCGCTGCGTG CGCGCGCGCG
      ACGGGCTCGG CGACCGCGGG CTCCCACACC GGCGACGCAC GCGCGCGCGC
7201  ACCCGGTCGC TGTTTGAACC GGGCGGAGGC GGGGCTGGCG CCCGGTTGGG
      TGGGCCAGCG ACAAACTTGG CCCGCCTCCG CCCCGACCGC GGGCCAACCC
7251  AGGGGGTTGG GGCCTGGCTT CCTGCCGCGC GCCGCGGGGA CGCCTCCGAC
      TCCCCCAACC CCGGACCGAA GGACGGCGCG CGGCGCCCCT GCGGAGGCTG
7301  CAGTGTTTGC CTTTTATGGT AATAACGCGG CCGGCCCGGC TTCCTTTGTC
      GTCACAAACG GAAAATACCA TTATTGCGCC GGCCGGGCCG AAGGAAACAG
7351  CCCAATCTGG GCGCGCGCCG GCGCCCCCTG GCGGCCTAAG GACTCGGCGC
      GGGTTAGACC CGCGCGCGGC CGCGGGGGAC CGCCGGATTC CTGAGCCGCG
7401  GCCGGAAGTG GCCAGGGCGG GGGCGACTTC GGCTCACAGC GCGCCCGGCT
      CGGCCTTCAC CGGTCCCGCC CCCGCTGAAG CCGAGTGTCG CGCGGGCCGA
7451  ATTCTCGCAG CTCACCATGC CGGTCGCCAC CATGAGCTTG ATATCGAATT
      TAAGAGCGTC GAGTGGTACG GCCAGCGGTG GTACTCGAAC TATAGCTTAA
7501  CCTGCAGCCC GGGGGATCCA CTAGTTCTAG AGCTTGATTA ATAGTAATCA
      GGACGTCGGG CCCCCTAGGT GATCAAGATC TCGAACTAAT TATCATTAGT
7551  ATTACGGGGT CATTAGTTCA TAGCCCATAT ATGGAGTTCC GCGTTACATA
      TAATGCCCCA GTAATCAAGT ATCGGGTATA TACCTCAAGG CGCAATGTAT
7601  ACTTACGGTA AATGGCCCGC CTGGCTGACC GCCCAACGAC CCCCGCCCAT
      TGAATGCCAT TTACCGGGCG GACCGACTGG CGGGTTGCTG GGGGCGGGTA
7651  TGACGTCAAT AATGACGTAT GTTCCCATAG TAACGCCAAT AGGGACTTTC
      ACTGCAGTTA TTACTGCATA CAAGGGTATC ATTGCGGTTA TCCCTGAAAG
7701  CATTGACGTC AATGGGTGGA GTATTTACGG TAAACTGCCC ACTTGGCAGT
      GTAACTGCAG TTACCCACCT CATAAATGCC ATTTGACGGG TGAACCGTCA
7751  ACATCAAGTG TATCATATGC CAAGTACGCC CCCTATTGAC GTCAATGACG
      TGTAGTTCAC ATAGTATACG GTTCATGCGG GGGATAACTG CAGTTACTGC
7801  GTAAATGGCC CGCCTGGCAT TATGCCCAGT ACATGACCTT ATGGGACTTT
      CATTTACCGG GCGGACCGTA ATACGGGTCA TGTACTGGAA TACCCTGAAA
7851  CCTACTTGGC AGTACATCTA CGTATTAGTC ATCGCTATTA CCATGGTGAT
      GGATGAACCG TCATGTAGAT GCATAATCAG TAGCGATAAT GGTACCACTA
7901  GCGGTTTTGG CAGTACATCA ATGGGCGTGG ATAGCGGTTT GACTCACGGG
      CGCCAAAACC GTCATGTAGT TACCCGCACC TATCGCCAAA CTGAGTGCCC
7951  GATTTCCAAG TCTCCACCCC ATTGACGTCA ATGGGAGTTT GTTTTGGCAC
      CTAAAGGTTC AGAGGTGGGG TAACTGCAGT TACCCTCAAA CAAAACCGTG
8001  CAAAATCAAC GGGACTTTCC AAAATGTCGT AACAACTCCG CCCCATTGAC
      GTTTTAGTTG CCCTGAAAGG TTTTACAGCA TTGTTGAGGC GGGGTAACTG
8051  GCAAATGGGC GGTAGGCGTG TACGGTGGGA GGTCTATATA AGCAGAGCTG
      CGTTTACCCG CCATCCGCAC ATGCCACCCT CCAGATATAT TCGTCTCGAC
```

FIG.10 (contd.)

```
8101  GTTTAGTGAA CCGTCAGATC CGCTAGCGTT CGAAGTTTAA ACGCGGCCGC
      CAAATCACTT GGCAGTCTAG GCGATCGCAA GCTTCAAATT TGCGCCGGCG
8151  GACTCTAGAT CATAATCAGC CATACCACAT TTGTAGAGGT TTTACTTGCT
      CTGAGATCTA GTATTAGTCG GTATGGTGTA AACATCTCCA AAATGAACGA
8201  TTAAAAAACC TCCCACACCT CCCCCTGAAC CTGAAACATA AAATGAATGC
      AATTTTTTGG AGGGTGTGGA GGGGGACTTG GACTTTGTAT TTTACTTACG
8251  AATTGTTGTT GTTAACTTGT TTATTGCAGC TTATAATGGT TACAAATAAA
      TTAACAACAA CAATTGAACA AATAACGTCG AATATTACCA ATGTTTATTT
8301  GCAATAGCAT CACAAATTTC ACAAATAAAG CATTTTTTTC ACTGCATTCT
      CGTTATCGTA GTGTTTAAAG TGTTTATTTC GTAAAAAAAG TGACGTAAGA
8351  AGTTGTGGTT TGTCCAAACT CATCAATGTA TCTTAAATCG AATTCTACCG
      TCAACACCAA ACAGGTTTGA GTAGTTACAT AGAATTTAGC TTAAGATGGC
8401  GGTAGGGGAG GCGCTTTTCC CAAGGCAGTC TGGAGCATGC GCTTTAGCAG
      CCATCCCCTC CGCGAAAAGG GTTCCGTCAG ACCTCGTACG CGAAATCGTC
8451  CCCCGCTGGG CACTTGGCGC TACACAAGTG GCCTCTGGCC TCGCACACAT
      GGGGCGACCC GTGAACCGCG ATGTGTTCAC CGGAGACCGG AGCGTGTGTA
8501  TCCACATCCA CCGGTAGGCG CCAACCGGCT CCGTTCTTTG GTGGCCCCTT
      AGGTGTAGGT GGCCATCCGC GGTTGGCCGA GGCAAGAAAC CACCGGGGAA
8551  CGCGCCACCT TCTACTCCTC CCCTAGTCAG GAAGTTCCCC CCCGCCCCGC
      GCGCGGTGGA AGATGAGGAG GGGATCAGTC CTTCAAGGGG GGGCGGGGCG
8601  AGCTCGCGTC GTGCAGGACG TGACAAATGG AAGTAGCACG TCTCACTAGT
      TCGAGCGCAG CACGTCCTGC ACTGTTTACC TTCATCGTGC AGAGTGATCA
8651  CTCGTGCAGA TGGACAGCAC CGCTGAGCAA TGGAAGCGGG TAGGCCTTTG
      GAGCACGTCT ACCTGTCGTG GCGACTCGTT ACCTTCGCCC ATCCGGAAAC
8701  GGGCAGCGGC CAATAGCAGC TTTGCTCCTT CGCTTTCTGG GCTCAGAGGC
      CCCGTCGCCG GTTATCGTCG AAACGAGGAA GCGAAAGACC CGAGTCTCCG
8751  TGGGAAGGGG TGGGTCCGGG GGCGGGCTCA GGGGCGGGCT CAGGGGCGGG
      ACCCTTCCCC ACCCAGGCCC CCGCCCGAGT CCCCGCCCGA GTCCCCGCCC
8801  GCGGGCGCCC GAAGGTCCTC CGGAGGCCCG GCATTCTGCA CGCTTCAAAA
      CGCCCGCGGG CTTCCAGGAG GCCTCCGGGC CGTAAGACGT GCGAAGTTTT
8851  GCGCACGTCT GCCGCGCTGT TCTCCTCTTC CTCATCTCCG GGCCTTTCGA
      CGCGTGCAGA CGGCGCGACA AGAGGAGAAG GAGTAGAGGC CCGGAAAGCT
8901  CCAGCTTACC ATGACCGAGT ACAAGCCCAC GGTGCGCCTC GCCACCCGCG
      GGTCGAATGG TACTGGCTCA TGTTCGGGTG CCACGCGGAG CGGTGGGCGC
8951  ACGACGTCCC CAGGGCCGTA CGCACCCTCG CCGCCGCGTT CGCCGACTAC
      TGCTGCAGGG GTCCCGGCAT GCGTGGGAGC GGCGGCGCAA GCGGCTGATG
9001  CCCGCCACGC GCCACACCGT CGATCCGGAC CGCCACATCG AGCGGGTCAC
      GGGCGGTGCG CGGTGTGGCA GCTAGGCCTG GCGGTGTAGC TCGCCCAGTG
9051  CGAGCTGCAA GAACTCTTCC TCACGCGCGT CGGGCTCGAC ATCGGCAAGG
      GCTCGACGTT CTTGAGAAGG AGTGCGCGCA GCCCGAGCTG TAGCCGTTCC
9101  TGTGGGTCGC GGACGACGGC GCCGCGGTGG CGGTCTGGAC CACGCCGGAG
      ACACCCAGCG CCTGCTGCCG CGGCGCCACC GCCAGACCTG GTGCGGCCTC
9151  AGCGTCGAAG CGGGGGCGGT GTTCGCCGAG ATCGGCCCGC GCATGGCCGA
      TCGCAGCTTC GCCCCCGCCA CAAGCGGCTC TAGCCGGGCG CGTACCGGCT
9201  GTTGAGCGGT TCCCGGCTGG CCGCGCAGAA CAGATGGAAG GCCTCCTGGC
      CAACTCGCCA AGGGCCGACC GGCGCGTCTT GTCTACCTTC CGGAGGACCG
9251  GCCGCACCGG CCCAAGGAGC CCGCGTGGTT CCTGGCCACC GTCGCGTCTC
      CGGCGTGGCC GGGTTCCTCG GGCGCACCAA GGACCGGTGG CAGCGCAGAG
9301  GCCCGACCAC CAGGGCAAGG GTCTGGGCAG CGCCGTCGTG CTCCCCGGAG
      CGGGCTGGTG GTCCCGTTCC CAGACCCGTC GCGGCAGCAC GAGGGGCCTC
9351  TGGAGGCGGC CGAGCGCGCC GGGGTGCCCG CCTTCCTGGA GACCTCCGCG
      ACCTCCGCCG GCTCGCGCGG CCCCACGGGC GGAAGGACCT CTGGAGGCGC
9401  CCCCGCAACC TCCCCTTCTA CGAGCGGCTC GGCTTCACCG TCACCGCCGA
      GGGGCGTTGG AGGGGAAGAT GCTCGCCGAG CCGAAGTGGC AGTGGCGGCT
```

```
 9451  CGTCGAGGTG CCCGAAGGAC CGCGCACCTG GTGCATGACC CGCAAGCCCG
       GCAGCTCCAC GGGCTTCCTG GCGCGTGGAC CACGTACTGG GCGTTCGGGC
 9501  GTGCCTGACG CCCGCCCCAC GACCCGCAGC GCCCGACCGA AAGGAGCGCA
       CACGGACTGC GGGCGGGGTG CTGGGCGTCG CGGGCTGGCT TTCCTCGCGT
 9551  CGACCCCATG CATCGTAGAG CTCGCTGATC AGCCTCGACT GTGCCTTCTA
       GCTGGGGTAC GTAGCATCTC GAGCGACTAG TCGGAGCTGA CACGGAAGAT
 9601  GTTGCCAGCC ATCTGTTGTT TGCCCCTCCC CCGTGCCTTC CTTGACCCTG
       CAACGGTCGG TAGACAACAA ACGGGGAGGG GGCACGGAAG GAACTGGGAC
 9651  GAAGGTGCCA CTCCCACTGT CCTTTCCTAA TAAAATGAGG AAATTGCATC
       CTTCCACGGT GAGGGTGACA GGAAAGGATT ATTTTACTCC TTTAACGTAG
 9701  GCATTGTCTG AGTAGGTGTC ATTCTATTCT GGGGGGTGGG GTGGGGCAGG
       CGTAACAGAC TCATCCACAG TAAGATAAGA CCCCCCACCC CACCCCGTCC
 9751  ACAGCAAGGG GGGGGATTGG GRAGACAATA GCAGGCATGC TGGGGGGGCG
       TGTCGTTCCC CCCCCTAACC CYTCTGTTAT CGTCCGTACG ACCCCCCCGC
 9801  GTGGGGGCTA TGGCTTCTGA GGCGGAAAGA ACCAGCTGGG GCTCGAGATC
       CACCCCCGAT ACCGAAGACT CCGCCTTTCT TGGTCGACCC CGAGCTCTAG
 9851  CACTAGTTCT AGCCTCGAGG CTAGAGCGGC CGCCACCGCG GTGGAGCTCC
       GTGATCAAGA TCGGAGCTCC GATCTCGCCG GCGGTGGCGC CACCTCGAGG
 9901  AATTCGCCCT ATAGTGAGTC GTATTACGCG CGCTCACTGG CCGTCGTTTT
       TTAAGCGGGA TATCACTCAG CATAATGCGC GCGAGTGACC GGCAGCAAAA
 9951  ACAACGTCGT GACTGGGAAA ACCCTGGCGT TACCCAACTT AATCGCCTTG
       TGTTGCAGCA CTGACCCTTT TGGGACCGCA ATGGGTTGAA TTAGCGGAAC
10001  CAGCACATCC CCCTTTCGCC AGCTGGCGTA ATAGCGAAGA GGCCCGCACC
       GTCGTGTAGG GGGAAAGCGG TCGACCGCAT TATCGCTTCT CCGGGCGTGG
10051  GATCGCCCTT CCCAACAGTT GCGCAGCCTG AATGGCGAAT GGAAATTGTA
       CTAGCGGGAA GGGTTGTCAA CGCGTCGGAC TTACCGCTTA CCTTTAACAT
10101  AGCGTTAATA TTTTGTTAAA ATTCGCGTTA AATTTTTGTT AAATCAGCTC
       TCGCAATTAT AAAACAATTT TAAGCGCAAT TTAAAAACAA TTTAGTCGAG
10151  ATTTTTTAAC CAATAGGCCG AAATCGGCAA AATCCCTTAT AAATCAAAAG
       TAAAAAATTG GTTATCCGGC TTTAGCCGTT TTAGGGAATA TTTAGTTTTC
10201  AATAGACCGA GATAGGGTTG AGTGTTGTTC CAGTTTGGAA CAAGAGTCCA
       TTATCTGGCT CTATCCCAAC TCACAACAAG GTCAAACCTT GTTCTCAGGT
10251  CTATTAAAGA ACGTGGACTC CAACGTCAAA GGGCGAAAAA CCGTCTATCA
       GATAATTTCT TGCACCTGAG GTTGCAGTTT CCCGCTTTTT GGCAGATAGT
10301  GGGCGATGGC CCACTACGTG AACCATCACC CTAATCAAGT TTTTTGGGGT
       CCCGCTACCG GGTGATGCAC TTGGTAGTGG GATTAGTTCA AAAAACCCCA
10351  CGAGGTGCCG TAAAGCACTA AATCGGAACC CTAAAGGGAG CCCCCGATTT
       GCTCCACGGC ATTTCGTGAT TTAGCCTTGG GATTTCCCTC GGGGGCTAAA
10401  AGAGCTTGAC GGGGAAAGCC GGCGAACGTG GCGAGAAAGG AAGGGAAGAA
       TCTCGAACTG CCCCTTTCGG CCGCTTGCAC CGCTCTTTCC TTCCCTTCTT
10451  AGCGAAAGGA GCGGGCGCTA GGGCGCTGGC AAGTGTAGCG GTCACGCTGC
       TCGCTTTCCT CGCCCGCGAT CCCGCGACCG TTCACATCGC CAGTGCGACG
10501  GCGTAACCAC CACACCCGCC GCGCTTAATG CGCCGCTACA GGGCGCGTCA
       CGCATTGGTG GTGTGGGCGG CGCGAATTAC GCGGCGATGT CCCGCGCAGT
10551  G
       C
```

FIG.10 (contd.)

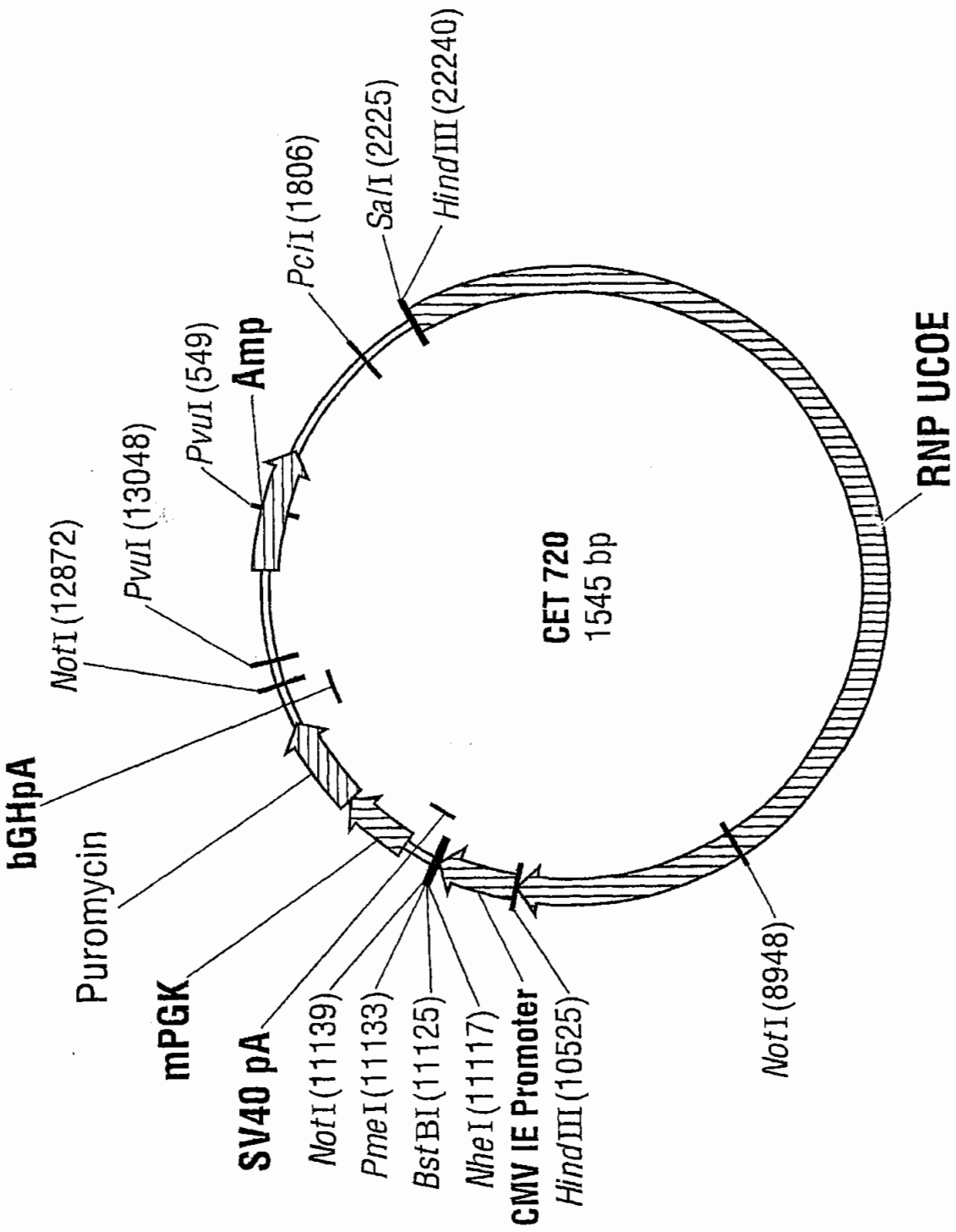

CET 720
1545 bp
RNP UCOE
Amp
bGHpA
Puromycin
mPGK
SV40 pA
CMV IE Promoter
NotI (12872)
PvuI (13048)
PvuI (549)
PciI (1806)
SalI (2225)
HindIII (22240)
NotI (11139)
PmeI (11133)
BstBI (11125)
NheI (11117)
HindIII (10525)
NotI (8948)

FIG. 11

FIG. 12 CET720 nucleotide sequence

```
   1  GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT GTTTATTTTT
      CCACCGTGAA AAGCCCCTTT ACACGCGCCT TGGGGATAAA CAAATAAAAA
  51  CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA CCCTGATAAA
      GATTTATGTA AGTTTATACA TAGGCGAGTA CTCTGTTATT GGGACTATTT
 101  TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT
      ACGAAGTTAT TATAACTTTT TCCTTCTCAT ACTCATAAGT TGTAAAGGCA
 151  GTCGCCCTTA TTCCCTTTTT TGCGGCATTT TGCCTTCCTG TTTTTGCTCA
      CAGCGGGAAT AAGGGAAAAA ACGCCGTAAA ACGGAAGGAC AAAAACGAGT
 201  CCCAGAAACG CTGGTGAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC
      GGGTCTTTGC GACCACTTTC ATTTTCTACG ACTTCTAGTC AACCCACGTG
 251  GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT CCTTGAGAGT
      CTCACCCAAT GTAGCTTGAC CTAGAGTTGT CGCCATTCTA GGAACTCTCA
 301  TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA AAGTTCTGCT
      AAAGCGGGGC TTCTTGCAAA AGGTTACTAC TCGTGAAAAT TTCAAGACGA
 351  ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC
      TACACCGCGC CATAATAGGG CATAACTGCG GCCCGTTCTC GTTGAGCCAG
 401  GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC ACCAGTCACA
      CGGCGTATGT GATAAGAGTC TTACTGAACC AACTCATGAG TGGTCAGTGT
 451  GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT GCAGTGCTGC
      CTTTTCGTAG AATGCCTACC GTACTGTCAT TCTCTTAATA CGTCACGACG
 501  CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG ACAACGATCG
      GTATTGGTAC TCACTATTGT GACGCCGGTT GAATGAAGAC TGTTGCTAGC
 551  GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG GGATCATGTA
      CTCCTGGCTT CCTCGATTGG CGAAAAAACG TGTTGTACCC CCTAGTACAT
 601  ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA
      TGAGCGGAAC TAGCAACCCT TGGCCTCGAC TTACTTCGGT ATGGTTTGCT
 651  CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAAC
      GCTCGCACTG TGGTGCTACG GACATCGTTA CCGTTGTTGC AACGCGTTTG
 701  TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC
      ATAATTGACC GCTTGATGAA TGAGATCGAA GGGCCGTTGT TAATTATCTG
 751  TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC
      ACCTACCTCC GCCTATTTCA ACGTCCTGGT GAAGACGCGA GCCGGGAAGG
 801  GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC
      CCGACCGACC AAATAACGAC TATTTAGACC TCGGCCACTC GCACCCAGAG
 851  GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA
      CGCCATAGTA ACGTCGTGAC CCCGGTCTAC CATTCGGGAG GGCATAGCAT
 901  GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
      CAATAGATGT GCTGCCCCTC AGTCCGTTGA TACCTACTTG CTTTATCTGT
 951  GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC
      CTAGCGACTC TATCCACGGA GTGACTAATT CGTAACCATT GACAGTCTGG
1001  AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA TTTTTAATTT
      TTCAAATGAG TATATATGAA ATCTAACTAA ATTTTGAAGT AAAAATTAAA
1051  AAAAGGATCT AGGTGAAGAT CCTTTTTGAT AATCTCATGA CCAAAATCCC
      TTTTCCTAGA TCCACTTCTA GGAAAAACTA TTAGAGTACT GGTTTTAGGG
1101  TTAACGTGAG TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA
      AATTGCACTC AAAAGCAAGG TGACTCGCAG TCTGGGGCAT CTTTTCTAGT
1151  AAGGATCTTC TTGAGATCCT TTTTTTCTGC GCGTAATCTG CTGCTTGCAA
      TTCCTAGAAG AACTCTAGGA AAAAAAGACG CGCATTAGAC GACGAACGTT
1201  ACAAAAAAAC CACCGCTACC AGCGGTGGTT TGTTTGCCGG ATCAAGAGCT
      TGTTTTTTTG GTGGCGATGG TCGCCACCAA ACAAACGGCC TAGTTCTCGA
1251  ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA
      TGGTTGAGAA AAAGGCTTCC ATTGACCGAA GTCGTCTCGC GTCTATGGTT
1301  ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT
      TATGACAGGA AGATCACATC GGCATCAATC CGGTGGTGAA GTTCTTGAGA
```

FIG. 12(contd.) 19/93

```
1351  GTAGCACCGC CTACATACCT CGCTCTGCTA ATCCTGTTAC CAGTGGCTGC
      CATCGTGGCG GATGTATGGA GCGAGACGAT TAGGACAATG GTCACCGACG
1401  TGCCAGTGGC GATAAGTCGT GTCTTACCGG GTTGGACTCA AGACGATAGT
      ACGGTCACCG CTATTCAGCA CAGAATGGCC CAACCTGAGT TCTGCTATCA
1451  TACCGGATAA GGCGCAGCGG TCGGGCTGAA CGGGGGGTTC GTGCACACAG
      ATGGCCTATT CCGCGTCGCC AGCCCGACTT GCCCCCCAAG CACGTGTGTC
1501  CCCAGCTTGG AGCGAACGAC CTACACCGAA CTGAGATACC TACAGCGTGA
      GGGTCGAACC TCGCTTGCTG GATGTGGCTT GACTCTATGG ATGTCGCACT
1551  GCTATGAGAA AGCGCCACGC TTCCCGAAGG GAGAAAGGCG GACAGGTATC
      CGATACTCTT TCGCGGTGCG AAGGGCTTCC CTCTTTCCGC CTGTCCATAG
1601  CGGTAAGCGG CAGGGTCGGA ACAGGAGAGC GCACGAGGGA GCTTCCAGGG
      GCCATTCGCC GTCCCAGCCT TGTCCTCTCG CGTGCTCCCT CGAAGGTCCC
1651  GGAAACGCCT GGTATCTTTA TAGTCCTGTC GGGTTTCGCC ACCTCTGACT
      CCTTTGCGGA CCATAGAAAT ATCAGGACAG CCCAAAGCGG TGGAGACTGA
1701  TGAGCGTCGA TTTTTGTGAT GCTCGTCAGG GGGGCGGAGC CTATGGAAAA
      ACTCGCAGCT AAAAACACTA CGAGCAGTCC CCCCGCCTCG GATACCTTTT
1751  ACGCCAGCAA CGCGGCCTTT TTACGGTTCC TGGCCTTTTG CTGGCCTTTT
      TGCGGTCGTT GCGCCGGAAA AATGCCAAGG ACCGGAAAAC GACCGGAAAA
1801  GCTCACATGT TCTTTCCTGC GTTATCCCCT GATTCTGTGG ATAACCGTAT
      CGAGTGTACA AGAAAGGACG CAATAGGGGA CTAAGACACC TATTGGCATA
1851  TACCGCCTTT GAGTGAGCTG ATACCGCTCG CCGCAGCCGA ACGACCGAGC
      ATGGCGGAAA CTCACTCGAC TATGGCGAGC GGCGTCGGCT TGCTGGCTCG
1901  GCAGCGAGTC AGTGAGCGAG GAAGCGGAAG AGCGCCCAAT ACGCAAACCG
      CGTCGCTCAG TCACTCGCTC CTTCGCCTTC TCGCGGGTTA TGCGTTTGGC
1951  CCTCTCCCCG CGCGTTGGCC GATTCATTAA TGCAGCTGGC ACGACAGGTT
      GGAGAGGGGC GCGCAACCGG CTAAGTAATT ACGTCGACCG TGCTGTCCAA
2001  TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT GTGAGTTAGC
      AGGGCTGACC TTTCGCCCGT CACTCGCGTT GCGTTAATTA CACTCAATCG
2051  TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG
      AGTGAGTAAT CCGTGGGGTC CGAAATGTGA AATACGAAGG CCGAGCATAC
2101  TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA ACAGCTATGA
      AACACACCTT AACACTCGCC TATTGTTAAA GTGTGTCCTT TGTCGATACT
2151  CCATGATTAC GCCAAGCGCG CAATTAACCC TCACTAAAGG GAACAAAAGC
      GGTACTAATG CGGTTCGCGC GTTAATTGGG AGTGATTTCC CTTGTTTTCG
2201  TGGGTACCGG GCCCCCCCTC GAGGTCGACG GTATCGATAA GCTTCAATGT
      ACCCATGGCC CGGGGGGGAG CTCCAGCTGC CATAGCTATT CGAAGTTACA
2251  TTTTAGCACC CTCTGTGTGG AGGAAAATAA TGCAGATTAT TCTAATTAGT
      AAAATCGTGG GAGACACACC TCCTTTTATT ACGTCTAATA AGATTAATCA
2301  GTAATATCTA ACCACATTAA AATATATTAC ATAGTAAACT ACACTCCATA
      CATTATAGAT TGGTGTAATT TTATATAATG TATCATTTGA TGTGAGGTAT
2351  ATTTTATAAA TTTGACTCCC CAGGGTAATA AACTAGTCTC TAGTCTGCTC
      TAAAATATTT AAACTGAGGG GTCCCATTAT TTGATCAGAG ATCAGACGAG
2401  ACCTTCAACT GTACAATAAA GTCTTGGTTC TTTTGAAATA GACCTCAAAT
      TGGAAGTTGA CATGTTATTT CAGAACCAAG AAAACTTTAT CTGGAGTTTA
2451  GAGACACCTA AAATTCAAAG TGTCTTTACA TTTAAAGACA CCTACAGGAA
      CTCTGTGGAT TTTAAGTTTC ACAGAAATGT AAATTTCTGT GGATGTCCTT
2501  AGCAGGTAAA AGAGCCAGGT TAAAAACAAA TTCTAAAACC ACTTAGCTGC
      TCGTCCATTT TCTCGGTCCA ATTTTTGTTT AAGATTTTGG TGAATCGACG
2551  AGTTAAACAT ATAGTAAAGA TGCACTAAAG TTTCTTACTC TGTAAATCCC
      TCAATTTGTA TATCATTTCT ACGTGATTTC AAAGAATGAG ACATTTAGGG
2601  TTCCACTTCA GGAAATATTC CACTTTCCCA TTCACTACAC GTCGATCTAG
      AAGGTGAAGT CCTTTATAAG GTGAAAGGGT AAGTGATGTG CAGCTAGATC
2651  TACTTTTTCC ACGACAAATT CTTCAGGCTC TGCCTCTTCA ACTTTTTTAC
      ATGAAAAAGG TGCTGTTTAA GAAGTCCGAG ACGGAGAAGT TGAAAAAATG
2701  TCTTTCCATT CTGTTTTTTT CCCATTTTTT GCTAAAATAA AACAAAAGAG
      AGAAAGGTAA GACAAAAAAA GGGTAAAAAA CGATTTTATT TTGTTTTCTC
2751  AAATTAAGAA ATATTCCTCT TGAATTTTGA GCACATTTTC AAGGCTCAAT
      TTTAATTCTT TATAAGGAGA ACTTAAAACT CGTGTAAAAG TTCCGAGTTA
```

FIG. 12(contd.) 20/93

```
2801  TGCTTATATT ATTATCACAT TCGACATAAA TTTTTACTTC TATATCCCAG
      ACGAATATAA TAATAGTGTA AGCTGTATTT AAAAATGAAG ATATAGGGTC
2851  GGCAGACACC TTCTGGAAAG ATTAAAAGTC AACAGACAAT AAAATAAAAG
      CCGTCTGTGG AAGACCTTTC TAATTTTCAG TTGTCTGTTA TTTTATTTTC
2901  AATGCTTTAT CTTGTTCATT TAGTTCAAAC TTACAACCCA CCACCAAAAT
      TTACGAAATA GAACAAGTAA ATCAAGTTTG AATGTTGGGT GGTGGTTTTA
2951  AATACAATAA AAAAACACTA TCTGGAAACA GTTATTTTTT TCCAGTCTTT
      TTATGTTATT TTTTTGTGAT AGACCTTTGT CAATAAAAAA AGGTCAGAAA
3001  TTTTTTGAGA CAGGGTCTCA CACTCTTGTC GCCCAGGCTG GAGTGCAGTG
      AAAAAACTCT GTCCCAGAGT GTGAGAACAG CGGGTCCGAC CTCACGTCAC
3051  GCGTGATCTC AGCTCACTGC AACCTCCGCC TCCCCAGGTT CAAGCAGTTC
      CGCACTAGAG TCGAGTGACG TTGGAGGCGG AGGGGTCCAA GTTCGTCAAG
3101  TCATGCCTCA GCCTCCAGAG TAGCTGGGAT TATAGGCGGA TGCCACCATG
      AGTACGGAGT CGGAGGTCTC ATCGACCCTA ATATCCGCCT ACGGTGGTAC
3151  CCGGGCTAAT TTTTTTTGTG TTTTTATTAG AAACAGGGTT TCACCATGTT
      GGCCCGATTA AAAAAAACAC AAAAATAATC TTTGTCCCAA AGTGGTACAA
3201  GACCAGGCTG GTCTCAAACT CCTGACCTGA AGTGATTCAC CAGCCTGGGC
      CTGGTCCGAC CAGAGTTTGA GGACTGGACT TCACTAAGTG GTCGGACCCG
3251  CTCCCAAAGT GCTGGCATTA CAGGCGTGAG CCACTGCGCC CGGCCCTGTA
      GAGGGTTTCA CGACCGTAAT GTCCGCACTC GGTGACGCGG GCCGGGACAT
3301  GTCTTAAAAG ACCAAGTTTA CTAATTTTCA CTCATTTTAA CAACACTGCA
      CAGAATTTTC TGGTTCAAAT GATTAAAAGT GAGTAAAATT GTTGTGACGT
3351  ACAAACAACT ATGCAGGAAG TACCTAAAGG GTGATCCAGA GAAGCAAGTA
      TGTTTGTTGA TACGTCCTTC ATGGATTTCC CACTAGGTCT CTTCGTTCAT
3401  GTAGTGACAG GTCTTAGGTG AACCTATGAC AGACCTTGTA TCCACCCCCA
      CATCACTGTC CAGAATCCAC TTGGATACTG TCTGGAACAT AGGTGGGGGT
3451  GATGGTAAAA GCCCCAGCCC CCTTCTCAAT TCAAATATTA ATGTCAAAAG
      CTACCATTTT CGGGGTCGGG GGAAGAGTTA AGTTTATAAT TACAGTTTTC
3501  CATCAATGAT ACAGAGAAAA GATAAATGCA GAATGAAAAC ATGGTTCAAA
      GTAGTTACTA TGTCTCTTTT CTATTTACGT CTTACTTTTG TACCAAGTTT
3551  ATCCTGATAC CAACTGCAGG GTCAACTATA GAGACCACTA GGAGGTTCAA
      TAGGACTATG GTTGACGTCC CAGTTGATAT CTCTGGTGAT CCTCCAAGTT
3601  TTAAAGGACA AGATTATTTT TCCATAATCT CTGTAGATAA TATTTCCTAC
      AATTTCCTGT TCTAATAAAA AGGTATTAGA GACATCTATT ATAAAGGATG
3651  CACTTAGAAC AAAACTATAA AGCTATCACT TCAAGAGACC AACATTACAA
      GTGAATCTTG TTTTGATATT TCGATAGTGA AGTTCTCTGG TTGTAATGTT
3701  ATTTATTTTA ATTCCCTAAG GTGAAAAAAA TCCTTCCTTC CTGGTTTCTC
      TAAATAAAAT TAAGGGATTC CACTTTTTTT AGGAAGGAAG GACCAAAGAG
3751  AAGAGAAAGT CTATACTGGT AACCAAATTC ACTTTAAACA GGCATTTTCT
      TTCTCTTTCA GATATGACCA TTGGTTTAAG TGAAATTTGT CCGTAAAAGA
3801  TTGGTATGAC ACTATTTAAG AGAAGCAGGA AACCAACGTG AACCAGCTCT
      AACCATACTG TGATAAATTC TCTTCGTCCT TTGGTTGCAC TTGGTCGAGA
3851  TTCCAATGGC TCAAGATTTC CTATGAGAGG ACTAAAAATG GGGAAAATTT
      AAGGTTACCG AGTTCTAAAG GATACTCTCC TGATTTTTAC CCCTTTTAAA
3901  TTATGAGAGG ATTAAAAATG GGGGAAAAAA AACCCTGAAA TGGTTAATCA
      AATACTCTCC TAATTTTTAC CCCCTTTTTT TTGGGACTTT ACCAATTAGT
3951  GAAGATCCTA TGGGCTGAGA AGGAATCCAT CTTAACATTT CATCTTAAAG
      CTTCTAGGAT ACCCGACTCT TCCTTAGGTA GAATTGTAAA GTAGAATTTC
4001  CAAATGCTAT TGCCGGGGGC AGTGGCTCAT GCCTGTAATC CCAGCACTTT
      GTTTACGATA ACGGCCCCCG TCACCGAGTA CGGACATTAG GGTCGTGAAA
4051  GGGAGGCCGA GGTGGGCAGA TCATCTGAGG TCAGGAGTTT GAGACCAGCC
      CCCTCCGGCT CCACCCGTCT AGTAGACTCC AGTCCTCAAA CTCTGGTCGG
4101  TGACCAACAT GGAGAAACCC CGTTTCTACT AAAAATACAA AATTAGCCAG
      ACTGGTTGTA CCTCTTTGGG GCAAAGATGA TTTTTATGTT TTAATCGGTC
4151  GCATAGTGGT GCATGCCTGT AATCCCAGCT ACTTGGGAGG CTGAGGCAGG
      CGTATCACCA CGTACGGACA TTAGGGTCGA TGAACCCTCC GACTCCGTCC
4201  AGAACTGCTT GAACCCAGGA GGCTTAAGTT GCGGTGAGCC AAGATCACGC
      TCTTGACGAA CTTGGGTCCT CCGAATTCAA CGCCACTCGG TTCTAGTGCG
```

FIG. 12(contd.) 21/93

```
4251  CATTGCACTC TAGCCTGGAC AACAAGAGAA AAACTCTGTC TCAAAAAAAC
      GTAACGTGAG ATCGGACCTG TTGTTCTCTT TTTGAGACAG AGTTTTTTTG
4301  ACAAAAACAA AAAACCCAAA TACTATTTAA AAAAGATAAA CCTTAATTGC
      TGTTTTTGTT TTTTGGGTTT ATGATAAATT TTTTCTATTT GGAATTAACG
4351  TCAATCATTA AAGCCATCCC ACAAGTAAAG CAGCAAGCAG AAAAAAGTTA
      AGTTAGTAAT TTCGGTAGGG TGTTCATTTC GTCGTTCGTC TTTTTTCAAT
4401  AGAACACCTC AAGGCTACAG AAGGACATTT CAAGCTATGC AGGCATATGA
      TCTTGTGGAG TTCCGATGTC TTCCTGTAAA GTTCGATACG TCCGTATACT
4451  AGTGTGCAGA CAGATATGTA AGAAAGGCCT CAAGACTGCA AAAGGGCATT
      TCACACGTCT GTCTATACAT TCTTTCCGGA GTTCTGACGT TTTCCCGTAA
4501  TCAAGCTATG CAAGCATATA GGTAACACAT ACACACACAC AAAATAAAAT
      AGTTCGATAC GTTCGTATAT CCATTGTGTA TGTGTGTGTG TTTTATTTTA
4551  CCCCTGAAAT ACAAAAACAT GCAGCAAACA CCTGACGTTT TTGGATACCA
      GGGGACTTTA TGTTTTTGTA CGTCGTTTGT GGACTGCAAA AACCTATGGT
4601  TTTCTAAGTC AGGTGTTATG ATTCTCATTA GTCAAGATAC TTGAGTACTG
      AAAGATTCAG TCCACAATAC TAAGAGTAAT CAGTTCTATG AACTCATGAC
4651  GGCCCAAACA GCTTTCTGCC ACTGTACAGT ACAAGAAGGT AGGAATAATG
      CCGGGTTTGT CGAAAGACGG TGACATGTCA TGTTCTTCCA TCCTTATTAC
4701  GTGGGAGGAG CAAAGACAAA CTGTAATAGA CAGAAGTGTA TCAGATACCT
      CACCCTCCTC GTTTCTGTTT GACATTATCT GTCTTCACAT AGTCTATGGA
4751  ATACTACATG AAAAACAAAA CAGCTACTGC CACAAAGGGA GAAGGCTAAC
      TATGATGTAC TTTTTGTTTT GTCGATGACG GTGTTTCCCT CTTCCGATTG
4801  AAAATAAAGT CAACAATAAA TACAGAAAAT GAAAAGGATA CACACTAAGG
      TTTTATTTCA GTTGTTATTT ATGTCTTTTA CTTTTCCTAT GTGTGATTCC
4851  TTTACAAAAA AAAAAAGGCA GACAAAATGC CATACAGTAT TCATTCACTA
      AAATGTTTTT TTTTTTCCGT CTGTTTTACG GTATGTCATA AGTAAGTGAT
4901  CTATGGCATT CATAAGCTAG TTTCAAATGC TCACTATTTT CTTTTATAGT
      GATACCGTAA GTATTCGATC AAAGTTTACG AGTGATAAAA GAAAATATCA
4951  ATATATTTGC CTTAACCCAG CACTTTTTTC CAAAAGTGGA TGAGTCAAAA
      TATATAAACG GAATTGGGTC GTGAAAAAAG GTTTTCACCT ACTCAGTTTT
5001  TAAATTTCCC ATTATTTAAG TGAAATTAAC AGCACACATA TCTCACAACA
      ATTTAAAGGG TAATAAATTC ACTTTAATTG TCGTGTGTAT AGAGTGTTGT
5051  CTAATGAATT TTTAAAATGG AAAGTTAAGA ACTTTTAAAG TGGCCAACCT
      GATTACTTAA AAATTTTACC TTTCAATTCT TGAAAATTTC ACCGGTTGGA
5101  GTGATCCTTC ACAAAATAAA CTAAATACAA TAACAGACCC CAAAGGCTAT
      CACTAGGAAG TGTTTTATTT GATTTATGTT ATTGTCTGGG GTTTCCGATA
5151  CAATTGCGTG CAAAAACAAC TTCTGTTTTC CAGGGTAAAC AGAATCTAAT
      GTTAACGCAC GTTTTTGTTG AAGACAAAAG GTCCCATTTG TCTTAGATTA
5201  GCAGAATCTA ATGCAGGGTA AACAGACTTA ATGCAGAATC TAATGATGGC
      CGTCTTAGAT TACGTCCCAT TTGTCTGAAT TACGTCTTAG ATTACTACCG
5251  ACAAATTAAA AATCACTAAC GTGCCCTTTT TAGTGTGAAA CCCAGAGAGA
      TGTTTAATTT TTAGTGATTG CACGGGAAAA ATCACACTTT GGGTCTCTCT
5301  GCACATACAA GCCAAAAACA AATGCTTTAT TTTACCTAGG AGACATTAAC
      CGTGTATGTT CGGTTTTTGT TTACGAAATA AAATGGATCC TCTGTAATTG
5351  ATTCACCTTT ACGTGTTTAA GATTAATGCA ATGTTAAATA TTGTGAAAAC
      TAAGTGGAAA TGCACAAATT CTAATTACGT TACAATTTAT AACACTTTTG
5401  TGTAACTTTG AATTTCATGA TTTTTATGTG AATATTCCAG GGTTTAAAAA
      ACATTGAAAC TTAAAGTACT AAAAATACAC TTATAAGGTC CCAAATTTTT
5451  AACTTGTAAC ATGACATGGC TGAATAAGAT AAAAAAAAAA TCTAGCCTTT
      TTGAACATTG TACTGTACCG ACTTATTCTA TTTTTTTTTT AGATCGGAAA
5501  TCTCCCTTCT GGCTCATATT TGCGATTTCG ATCATTTTGT TTAAAAAACA
      AGAGGGAAGA CCGAGTATAA ACGCTAAAGC TAGTAAAACA AATTTTTTGT
5551  AAACACTGCA ATGAATTAAA CTTAATATTC TTCTATGTTT TAGAGTAAGT
      TTTGTGACGT TACTTAATTT GAATTATAAG AAGATACAAA ATCTCATTCA
5601  TAAAACAAGA TAAAGTGACC AAAGTAATTT GAAAGATTCA ATGACTTTTG
      ATTTTGTTCT ATTTCACTGG TTTCATTAAA CTTTCTAAGT TACTGAAAAC
5651  CTCCAACCTA GGTGCACAAG GTACCTTGTT CTTTAAATTG GGCTTTAATG
      GAGGTTGGAT CCACGTGTTC CATGGAACAA GAAATTTAAC CCGAAATTAC
```

FIG. 12(contd.) 22/93

```
5701  AAAATACTTC TCCAGAATTC TGGGGATTTA AGAAAAATTA TGCCAACCAA
      TTTTATGAAG AGGTCTTAAG ACCCCTAAAT TCTTTTTAAT ACGGTTGGTT
5751  CAAGGGCTTT ACCATTTTAT GTAACATTTT TCAACGCTGC AAAAATGTGT
      GTTCCCGAAA TGGTAAAATA CATTGTAAAA AGTTGCGACG TTTTTACACA
5801  GTATTTCTAT TTGAAGATAA AAATCCTCAG CAAAATCCAC ATTGCACTGT
      CATAAAGATA AACTTCTATT TTTAGGAGTC GTTTTAGGTG TAACGTGACA
5851  CCTTCAAAGA TTAGCCTTCT TTGAACTAGT TAAGACACTA TTAAGCCAAG
      GGAAGTTTCT AATCGGAAGA AACTTGATCA ATTCTGTGAT AATTCGGTTC
5901  CCAGTATCTC CCTGTAATGA ATTCGTTTTT CTCTTAATTT TCCCCTGTAA
      GGTCATAGAG GGACATTACT TAAGCAAAAA GAGAATTAAA AGGGGACATT
5951  TTTACACTGG GAGAGCTGGG AAATATGTGG ATGTAAATTT CTCAGCCACA
      AAATGTGACC CTCTCGACCC TTTATACACC TACATTTAAA GAGTCGGTGT
6001  GAGATGCAAA GTTATACTGT GGGGAAAAAA AACTTGAGTT AAATCCTTAC
      CTCTACGTTT CAATATGACA CCCCTTTTTT TTGAACTCAA TTTAGGAATG
6051  ATATTTTAGG TTTTCATTAA CTTACCAATG TAGTTTTGTT GGAGGCCATT
      TATAAAATCC AAAAGTAATT GAATGGTTAC ATCAAAACAA CCTCCGGTAA
6101  TTTTTTATTG CAGACTTGAA GAGCTATTAC TAGAAAAATG CATGACAGTT
      AAAAAATAAC GTCTGAACTT CTCGATAATG ATCTTTTTAC GTACTGTCAA
6151  AAGGTAAGTT TGCATGACAC AAAAAAGGTA ACTAAATACA AATTCTGTTT
      TTCCATTCAA ACGTACTGTG TTTTTTCCAT TGATTTATGT TTAAGACAAA
6201  GGATTCCAAC CCCCAAGTAG AGAGCGCACA CTTTCAAACG TGAATACAAA
      CCTAAGGTTG GGGGTTCATC TCTCGCGTGT GAAAGTTTGC ACTTATGTTT
6251  TCCAGAGTAG ATCTGCGCTC CTACCTACAT TGCTTATGAT GTACTTAAGT
      AGGTCTCATC TAGACGCGAG GATGGATGTA ACGAATACTA CATGAATTCA
6301  ACGTGTCCTA ACCATGTGAG TCTAGAAAGA CTTTACTGGG GATCCTGGTA
      TGCACAGGAT TGGTACACTC AGATCTTTCT GAAATGACCC CTAGGACCAT
6351  CCTAAAACAG CTTCACATGG CTTAAAATAG GGGACCAATG TCTTTTCCAA
      GGATTTTGTC GAAGTGTACC GAATTTTATC CCCTGGTTAC AGAAAAGGTT
6401  TCTAAGTCCC ATTTATAATA AAGTCCATGT TCCATTTTTA AAGGACAATC
      AGATTCAGGG TAAATATTAT TTCAGGTACA AGGTAAAAAT TTCCTGTTAG
6451  CTTTCGGTTT AAAACCAGGC ACGATTACCC AAACAACTCA CAACGGTAAA
      GAAAGCCAAA TTTTGGTCCG TGCTAATGGG TTTGTTGAGT GTTGCCATTT
6501  GCACTGTGAA TCTTCTCTGT TCTGCAATCC CAACTTGGTT TCTGCTCAGA
      CGTGACACTT AGAAGAGACA AGACGTTAGG GTTGAACCAA AGACGAGTCT
6551  AACCCTCCCT CTTTCCAATC GGTAATTAAA TAACAAAAGG AAAAAACTTA
      TTGGGAGGGA GAAAGGTTAG CCATTAATTT ATTGTTTTCC TTTTTTGAAT
6601  AGATGCTTCA ACCCCGTTTC GTGACACTTT GAAAAAAGAA TCACCTCTTG
      TCTACGAAGT TGGGGCAAAG CACTGTGAAA CTTTTTTCTT AGTGGAGAAC
6651  CAAACACCCG CTCCCGACCC CCGCCGCTGA AGCCCGGCGT CCAGAGGCCT
      GTTTGTGGGC GAGGGCTGGG GGCGGCGACT TCGGGCCGCA GGTCTCCGGA
6701  AAGCGCGGGT GCCCGCCCCC ACCCGGGAGC GCGGGCCTCG TGGTCAGCGC
      TTCGCGCCCA CGGGCGGGGG TGGGCCCTCG CGCCCGGAGC ACCAGTCGCG
6751  ATCCGCGGGG AGAAACAAAG GCCGCGGCAC GGGGGCTCAA GGGCACTGCG
      TAGGCGCCCC TCTTTGTTTC CGGCGCCGTG CCCCCGAGTT CCCGTGACGC
6801  CCACACCGCA CGCGCCTACC CCCGCGCGGC CACGTTAACT GGCGGTCGCC
      GGTGTGGCGT GCGCGGATGG GGGCGCGCCG GTGCAATTGA CCGCCAGCGG
6851  GCAGCCTCGG GACAGCCGGC CGCGCGCCGC CAGGCTCGCG GACGCGGGAC
      CGTCGGAGCC CTGTCGGCCG GCGCGCGGCG GTCCGAGCGC CTGCGCCCTG
6901  CACGCGCCGC CCTCCGGGAG GCCCAAGTCT CGACCCAGCC CCGCGTGGCG
      GTGCGCGGCG GGAGGCCCTC CGGGTTCAGA GCTGGGTCGG GGCGCACCGC
6951  CTGGGGGAGG GGGCGCCTCC GCCGGAACGC GGGTGGGGGA GGGGAGGGGG
      GACCCCCTCC CCCGCGGAGG CGGCCTTGCG CCCACCCCCT CCCCTCCCCC
7001  AAATGCGCTT TGTCTCGAAA TGGGGCAACC GTCGCCACAG CTCCCTACCC
      TTTACGCGAA ACAGAGCTTT ACCCCGTTGG CAGCGGTGTC GAGGGATGGG
7051  CCTCGAGGGC AGAGCAGTCC CCCCACTAAC TACCGGGCTG GCCGCGCGCC
      GGAGCTCCCG TCTCGTCAGG GGGGTGATTG ATGGCCCGAC CGGCGCGCGG
7101  AGGCCAGCCG CGAGGCCACC GCCCGACCCT CCACTCCTTC CCGCAGCTCC
      TCCGGTCGGC GCTCCGGTGG CGGGCTGGGA GGTGAGGAAG GGCGTCGAGG
```

FIG. 12(contd.) 23/93

```
7151  CGGCGCGGGG TCCGGCGAGA AGGGGAGGGG AGGGGAGCGG AGAACCGGGC
      GCCGCGCCCC AGGCCGCTCT TCCCCTCCCC TCCCCTCGCC TCTTGGCCCG
7201  CCCCGGGACG CGTGTGGCAT CTGAAGCACC ACCAGCGAGC GAGAGCTAGA
      GGGGCCCTGC GCACACCGTA GACTTCGTGG TGGTCGCTCG CTCTCGATCT
7251  GAGAAGGAAA GCCACCGACT TCACCGCCTC CGAGCTGCTC CGGGTCGCGG
      CTCTTCCTTT CGGTGGCTGA AGTGGCGGAG GCTCGACGAG GCCCAGCGCC
7301  GTCTGCAGCG TCTCCGGCCC TCCGCGCCTA CAGCTCAAGC CACATCCGAA
      CAGACGTCGC AGAGGCCGGG AGGCGCGGAT GTCGAGTTCG GTGTAGGCTT
7351  GGGGGAGGGA GCCGGGAGCT GCGCGCGGGG CCGCCGGGGG GAGGGGTGGC
      CCCCCTCCCT CGGCCCTCGA CGCGCGCCCC GGCGGCCCCC CTCCCCACCG
7401  ACCGCCCACG CCGGGCGGCC ACGAAGGGCG GGGCAGCGGG CGCGCGCGCG
      TGGCGGGTGC GGCCCGCCGG TGCTTCCCGC CCCGTCGCCC GCGCGCGCGC
7451  GCGGGGGGAG GGGCCGGCGC CGCGCCCGCT GGGAATTGGG GCCCTAGGGG
      CGCCCCCCTC CCCGGCCGCG GCGCGGGCGA CCCTTAACCC CGGGATCCCC
7501  GAGGGCGGAG GCGCCGACGA CCGCGGCACT TACCGTTCGC GGCGTGGCGC
      CTCCCGCCTC CGCGGCTGCT GGCGCCGTGA ATGGCAAGCG CCGCACCGCG
7551  CCGGTGGTCC CCAAGGGGAG GGAAGGGGGA GGCGGGGCGA GGACAGTGAC
      GGCCACCAGG GGTTCCCCTC CCTTCCCCCT CCGCCCCGCT CCTGTCACTG
7601  CGGAGTCTCC TCAGCGGTGG CTTTTCTGCT TGGCAGCCTC AGCGGCTGGC
      GCCTCAGAGG AGTCGCCACC GAAAAGACGA ACCGTCGGAG TCGCCGACCG
7651  GCCAAAACCG GACTCCGCCC ACTTCCTCGC CCGCCGGTGC GAGGGTGTGG
      CGGTTTTGGC CTGAGGCGGG TGAAGGAGCG GGCGGCCACG CTCCCACACC
7701  AATCCTCCAG ACGCTGGGGG AGGGGGAGTT GGGAGCTTAA AAACTAGTAC
      TTAGGAGGTC TGCGACCCCC TCCCCCTCAA CCCTCGAATT TTTGATCATG
7751  CCCTTTGGGA CCACTTTCAG CAGCGAACTC TCCTGTACAC CAGGGGTCAG
      GGGAAACCCT GGTGAAAGTC GTCGCTTGAG AGGACATGTG GTCCCCAGTC
7801  TTCCACAGAC GCGGGCCAGG GGTGGGTCAT TGCGGCGTGA ACAATAATTT
      AAGGTGTCTG CGCCCGGTCC CCACCCAGTA ACGCCGCACT TGTTATTAAA
7851  GACTAGAAGT TGATTCGGGT GTTTCCGGAA GGGGCCGAGT CAATCCGCCG
      CTGATCTTCA ACTAAGCCCA CAAAGGCCTT CCCCGGCTCA GTTAGGCGGC
7901  AGTTGGGGCA CGGAAAACAA AAAGGGAAGG CTACTAAGAT TTTTCTGGCG
      TCAACCCCGT GCCTTTTGTT TTTCCCTTCC GATGATTCTA AAAAGACCGC
7951  GGGGTTATCA TTGGCGTAAC TGCAGGGACC ACCTCCCGGG TTGAGGGGGC
      CCCCAATAGT AACCGCATTG ACGTCCCTGG TGGAGGGCCC AACTCCCCCG
8001  TGGATCTCCA GGCTGCGGAT TAAGCCCCTC CCGTCGGCGT TAATTTCAAA
      ACCTAGAGGT CCGACGCCTA ATTCGGGGAG GGCAGCCGCA ATTAAAGTTT
8051  CTGCGCGACG TTTCTCACCT GCCTTCGCCA AGGCAGGGGC CGGGACCCTA
      GACGCGCTGC AAAGAGTGGA CGGAAGCGGT TCCGTCCCCG GCCCTGGGAT
8101  TTCCAAGAGG TAGTAACTAG CAGGACTCTA GCCTTCCGCA ATTCATTGAG
      AAGGTTCTCC ATCATTGATC GTCCTGAGAT CGGAAGGCGT TAAGTAACTC
8151  CGCATTTACG GAAGTAACGT CGGGTACTGT CTCTGGCCGC AAGGGTGGGA
      GCGTAAATGC CTTCATTGCA GCCCATGACA GAGACCGGCG TTCCCACCCT
8201  GGAGTACGCA TTTGGCGTAA GGTGGGGCGT AGAGCCTTCC CGCCATTGGC
      CCTCATGCGT AAACCGCATT CCACCCCGCA TCTCGGAAGG GCGGTAACCG
8251  GGCGGATAGG GCGTTTACGC GACGGCCTGA CGTAGCGGAA GACGCGTTAG
      CCGCCTATCC CGCAAATGCG CTGCCGGACT GCATCGCCTT CTGCGCAATC
8301  TGGGGGGGAA GGTTCTAGAA AAGCGGCGGC AGCGGCTCTA GCGGCAGTAG
      ACCCCCCCTT CCAAGATCTT TTCGCCGCCG TCGCCGAGAT CGCCGTCATC
8351  CAGCAGCGCC GGGTCCCGTG CGGAGGTGCT CCTCGCAGAG TTGTTTCTCG
      GTCGTCGCGG CCCAGGGCAC GCCTCCACGA GGAGCGTCTC AACAAAGAGC
8401  AGCAGCGGCA GTTCTCACTA CAGCGCCAGG ACGAGTCCGG TTCGTGTTCG
      TCGTCGCCGT CAAGAGTGAT GTCGCGGTCC TGCTCAGGCC AAGCACAAGC
8451  TCCGCGGAGA TCTCTCTCAT CTCGCTCGGC TGCGGGAAAT CGGGCTGAAG
      AGGCGCCTCT AGAGAGAGTA GAGCGAGCCG ACGCCCTTTA GCCCGACTTC
8501  CGACTGAGTC CGCGATGGAG GTAACGGGTT TGAAATCAAT GAGTTATTGA
      GCTGACTCAG GCGCTACCTC CATTGCCCAA ACTTTAGTTA CTCAATAACT
8551  AAAGGGCATG GCGAGGCCGT TGGCGCCTCA GTGGAAGTCG GCCAGCCGCC
      TTTCCCGTAC CGCTCCGGCA ACCGCGGAGT CACCTTCAGC CGGTCGGCGG
```

FIG. 12(contd.) 24/93

```
8601  TCCGTGGGAG AGAGGCAGGA AATCGGACCA ATTCAGTAGC AGTGGGGCTT
      AGGCACCCTC TCTCCGTCCT TTAGCCTGGT TAAGTCATCG TCACCCCGAA
8651  AAGGTTTATG AACGGGGTCT TGAGCGGAGG CCTGAGCGTA CAAACAGCTT
      TTCCAAATAC TTGCCCCAGA ACTCGCCTCC GGACTCGCAT GTTTGTCGAA
8701  CCCCACCCTC AGCCTCCCGG CGCCATTTCC CTTCACTGGG GGTGGGGGAT
      GGGGTGGGAG TCGGAGGGCC GCGGTAAAGG GAAGTGACCC CCACCCCCTA
8751  GGGGAGCTTT CACATGGCGG ACGCTGCCCC GCTGGGGTGA AAGTGGGGCG
      CCCCTCGAAA GTGTACCGCC TGCGACGGGG CGACCCCACT TTCACCCCGC
8801  CGGAGGCGGG AATTCTTATT CCCTTTCTAA AGCACGCTGC TTCGGGGGCC
      GCCTCCGCCC TTAAGAATAA GGGAAAGATT TCGTGCGACG AAGCCCCCGG
8851  ACGGCGTCTC CTCGGCGAGC GTTTCGGCGG GCAGCAGGTC CTCGTGAGCG
      TGCCGCAGAG GAGCCGCTCG CAAAGCCGCC CGTCGTCCAG GAGCACTCGC
8901  AGGCTGCGGA GCTTCCCCTC CCCCTCTCTC CCGGGAACCG ATTTGGCGGC
      TCCGACGCCT CGAAGGGGAG GGGGAGAGAG GGCCCTTGGC TAAACCGCCG
8951  CGCCATTTTC ATGGCTCGCC TTCCTCTCAG CGTTTTCCTT ATAACTCTTT
      GCGGTAAAAG TACCGAGCGG AAGGAGAGTC GCAAAAGGAA TATTGAGAAA
9001  TATTTTCTTA GTGTGCTTTC TCTATCAAGA AGTAGAAGTG GTTAACTATT
      ATAAAAGAAT CACACGAAAG AGATAGTTCT TCATCTTCAC CAATTGATAA
9051  TTTTTTTTCT TCTCGGGCTG TTTTCATATC GTTTCGAGGT GGATTTGGAG
      AAAAAAAAGA AGAGCCCGAC AAAAGTATAG CAAAGCTCCA CCTAAACCTC
9101  TGTTTTGTGA GCTTGGATCT TTAGAGTCCT GCGCACCTCA TTAAAGGCGC
      ACAAAACACT CGAACCTAGA AATCTCAGGA CGCGTGGAGT AATTTCCGCG
9151  TCAGCCTTCC CCTCGATGAA ATGGCGCCAT TGCGTTCGGA AGCCACACCG
      AGTCGGAAGG GGAGCTACTT TACCGCGGTA ACGCAAGCCT TCGGTGTGGC
9201  AAGAGCGGGG AGGGGGGGTG CTCCGGGTTT GCGGGCCCGG TTTCAGAGAA
      TTCTCGCCCC TCCCCCCCAC GAGGCCCAAA CGCCCGGGCC AAAGTCTCTT
9251  GATATCACCA CCCAGGGCGT CGGGCCGGGT TCAATGCGAG CCGTAGGACA
      CTATAGTGGT GGGTCCCGCA GCCCGGCCCA AGTTACGCTC GGCATCCTGT
9301  AAGAAACCAT TTTATGTTTT TCCTGTCTTT TTTTTCCTTT GAGTAACGGT
      TTCTTTGGTA AAATACAAAA AGGACAGAAA AAAAAGGAAA CTCATTGCCA
9351  TTTATCTGGG TCTGCAGTCA GTAAAACGAC AGATGAACCG CGGCAAAATA
      AAATAGACCC AGACGTCAGT CATTTTGCTG TCTACTTGGC GCCGTTTTAT
9401  AACATAAATT GGAAGCCATC GGCCACGAGG GGCAGGGACG AAGGTGGTTT
      TTGTATTTAA CCTTCGGTAG CCGGTGCTCC CCGTCCCTGC TTCCACCAAA
9451  TCTGGGCGGG GGAGGGATAT TCGCGTCAGA ATCCTTTACT GTTCTTAAGG
      AGACCCGCCC CCTCCCTATA AGCGCAGTCT TAGGAAATGA CAAGAATTCC
9501  ATTCCGTTTA AGTTGTAGAG CTGACTCATT TTAAGTAATG TTGTTACTGA
      TAAGGCAAAT TCAACATCTC GACTGAGTAA AATTCATTAC AACAATGACT
9551  GAAGTTTAAC CCTTACGGGA CAGATCCATG GACCTTTATA GATGATTACG
      CTTCAAATTG GGAATGCCCT GTCTAGGTAC CTGGAAATAT CTACTAATGC
9601  AGGAAAGTGA AATAACGATT TTGTCCTTAG TTATACTTCG ATTAAAACAT
      TCCTTTCACT TTATTGCTAA AACAGGAATC AATATGAAGC TAATTTTGTA
9651  GGCTTCAGAG GCTCCTTCCT GTAATGCGTA TGGATTGATG TGCAAAACTG
      CCGAAGTCTC CGAGGAAGGA CATTACGCAT ACCTAACTAC ACGTTTTGAC
9701  TTTTGGGCCT GGGCCGCTCT GTATTTGAAC TTTGTTACTT TTCTCATTTT
      AAAACCCGGA CCCGGCGAGA CATAAACTTG AAACAATGAA AAGAGTAAAA
9751  GTTTGCAATC TTGGTTGAAC ATTACATTGA TAAGCATAAG GTCTCAAGCG
      CAAACGTTAG AACCAACTTG TAATGTAACT ATTCGTATTC CAGAGTTCGC
9801  AAGGGGGTCT ACCTGGTTAT TTTTCTTTGA CCCTAAGCAC GTTTATAAAA
      TTCCCCCAGA TGGACCAATA AAAAGAAACT GGGATTCGTG CAAATATTTT
9851  TAACATTGTT TAAAATCGAT AGTGGACATC GGGTAAGTTT GGATAAATTG
      ATTGTAACAA ATTTTAGCTA TCACCTGTAG CCCATTCAAA CCTATTTAAC
9901  TGAGGTAAGT AATGAGTTTT TGCTTTTTGT TAGTGATTTG TAAAACTTGT
      ACTCCATTCA TTACTCAAAA ACGAAAAACA ATCACTAAAC ATTTTGAACA
9951  TATAAATGTA CATTATCCGT AATTTCAGTT TAGAGATAAC CTATGTGCTG
      ATATTTACAT GTAATAGGCA TTAAAGTCAA ATCTCTATTG GATACACGAC
10001 ACGACAATTA AGAATAAAAA CTAGCTGAAA AAATGAAAAT AACTATCGTG
      TGCTGTTAAT TCTTATTTTT GATCGACTTT TTTACTTTTA TTGATAGCAC
```

FIG. 12(contd.) 25/93

```
10051   ACAAGTAACC ATTTCAAAAG ACTGCTTTGT GTCTCATAGG AGCTAGTTTG
        TGTTCATTGG TAAAGTTTTC TGACGAAACA CAGAGTATCC TCGATCAAAC
10101   ATCATTTCAG TTAATTTTTT CTTTAATTTT TACGAGTCAT GAAAACTACA
        TAGTAAAGTC AATTAAAAAA GAAATTAAAA ATGCTCAGTA CTTTTGATGT
10151   GGAAAAAAAA TCTGAACTGG GTTTTACCAC TACTTTTTAG GAGTTGGGAG
        CCTTTTTTTT AGACTTGACC CAAAATGGTG ATGAAAAATC CTCAACCCTC
10201   CATGCGAATG GAGGGAGAGC TCCGTAGAAC TGGGATGAGA GCAGCAATTA
        GTACGCTTAC CTCCCTCTCG AGGCATCTTG ACCCTACTCT CGTCGTTAAT
10251   ATGCTGCTTG CTAGGAACAA AAAATAATTG ATTGAAAATT ACGTGTGACT
        TACGACGAAC GATCCTTGTT TTTTATTAAC TAACTTTTAA TGCACACTGA
10301   TTTTAGTTTG CATTATGCGT TTGTAGCAGT TGGTCCTGGA TATCACTTTC
        AAAATCAAAC GTAATACGCA AACATCGTCA ACCAGGACCT ATAGTGAAAG
10351   TCTCGTTTGA GGTTTTTTAA CCTAGTTAAC TTTTAAGACA GGTTTCCTTA
        AGAGCAAACT CCAAAAAATT GGATCAATTG AAAATTCTGT CCAAAGGAAT
10401   ACATTCATAA GTGCCCAGAA TACAGCTGTG TAGTACAGCA TATAAAGATT
        TGTAAGTATT CACGGGTCTT ATGTCGACAC ATCATGTCGT ATATTTCTAA
10451   TCAGCTCTGA GGTTTTTCCT ATTGACTTGG AAAATTGTTT TGTGCCTGTC
        AGTCGAGACT CCAAAAAGGA TAACTGAACC TTTTAACAAA ACACGGACAG
10501   GCTTGCCACA TGGCCAATCA AGTAAGCTTG ATTAATAGTA ATCAATTACG
        CGAACGGTGT ACCGGTTAGT TCATTCGAAC TAATTATCAT TAGTTAATGC
10551   GGGTCATTAG TTCATAGCCC ATATATGGAG TTCCGCGTTA CATAACTTAC
        CCCAGTAATC AAGTATCGGG TATATACCTC AAGGCGCAAT GTATTGAATG
10601   GGTAAATGGC CCGCCTGGCT GACCGCCCAA CGACCCCCGC CCATTGACGT
        CCATTTACCG GGCGGACCGA CTGGCGGGTT GCTGGGGGCG GGTAACTGCA
10651   CAATAATGAC GTATGTTCCC ATAGTAACGC CAATAGGGAC TTTCCATTGA
        GTTATTACTG CATACAAGGG TATCATTGCG GTTATCCCTG AAAGGTAACT
10701   CGTCAATGGG TGGAGTATTT ACGGTAAACT GCCCACTTGG CAGTACATCA
        GCAGTTACCC ACCTCATAAA TGCCATTTGA CGGGTGAACC GTCATGTAGT
10751   AGTGTATCAT ATGCCAAGTA CGCCCCCTAT TGACGTCAAT GACGGTAAAT
        TCACATAGTA TACGGTTCAT GCGGGGGATA ACTGCAGTTA CTGCCATTTA
10801   GGCCCGCCTG GCATTATGCC CAGTACATGA CCTTATGGGA CTTTCCTACT
        CCGGGCGGAC CGTAATACGG GTCATGTACT GGAATACCCT GAAAGGATGA
10851   TGGCAGTACA TCTACGTATT AGTCATCGCT ATTACCATGG TGATGCGGTT
        ACCGTCATGT AGATGCATAA TCAGTAGCGA TAATGGTACC ACTACGCCAA
10901   TTGGCAGTAC ATCAATGGGC GTGGATAGCG GTTTGACTCA CGGGGATTTC
        AACCGTCATG TAGTTACCCG CACCTATCGC CAAACTGAGT GCCCCTAAAG
10951   CAAGTCTCCA CCCCATTGAC GTCAATGGGA GTTTGTTTTG GCACCAAAAT
        GTTCAGAGGT GGGGTAACTG CAGTTACCCT CAAACAAAAC CGTGGTTTTA
11001   CAACGGGACT TTCCAAAATG TCGTAACAAC TCCGCCCCAT TGACGCAAAT
        GTTGCCCTGA AAGGTTTTAC AGCATTGTTG AGGCGGGGTA ACTGCGTTTA
11051   GGGCGGTAGG CGTGTACGGT GGGAGGTCTA TATAAGCAGA GCTGGTTTAG
        CCCGCCATCC GCACATGCCA CCCTCCAGAT ATATTCGTCT CGACCAAATC
11101   TGAACCGTCA GATCCGCTAG CGTTCGAAGT TTAAACGCGG CCGCGACTCT
        ACTTGGCAGT CTAGGCGATC GCAAGCTTCA AATTTGCGCC GGCGCTGAGA
11151   AGATCATAAT CAGCCATACC ACATTTGTAG AGGTTTTACT TGCTTTAAAA
        TCTAGTATTA GTCGGTATGG TGTAAACATC TCCAAAATGA ACGAAATTTT
11201   AACCTCCCAC ACCTCCCCCT GAACCTGAAA CATAAAATGA ATGCAATTGT
        TTGGAGGGTG TGGAGGGGGA CTTGGACTTT GTATTTTACT TACGTTAACA
11251   TGTTGTTAAC TTGTTTATTG CAGCTTATAA TGGTTACAAA TAAAGCAATA
        ACAACAATTG AACAAATAAC GTCGAATATT ACCAATGTTT ATTTCGTTAT
11301   GCATCACAAA TTTCACAAAT AAAGCATTTT TTTCACTGCA TTCTAGTTGT
        CGTAGTGTTT AAAGTGTTTA TTTCGTAAAA AAAGTGACGT AAGATCAACA
11351   GGTTTGTCCA AACTCATCAA TGTATCTTAA ATCGAATTCT ACCGGGTAGG
        CCAAACAGGT TTGAGTAGTT ACATAGAATT TAGCTTAAGA TGGCCCATCC
11401   GGAGGCGCTT TTCCCAAGGC AGTCTGGAGC ATGCGCTTTA GCAGCCCCGC
        CCTCCGCGAA AAGGGTTCCG TCAGACCTCG TACGCGAAAT CGTCGGGGCG
11451   TGGGCACTTG GCGCTACACA AGTGGCCTCT GGCCTCGCAC ACATTCCACA
        ACCCGTGAAC CGCGATGTGT TCACCGGAGA CCGGAGCGTG TGTAAGGTGT
```

FIG. 12(contd.) 26/93

```
11501   TCCACCGGTA GGCGCCAACC GGCTCCGTTC TTTGGTGGCC CCTTCGCGCC
        AGGTGGCCAT CCGCGGTTGG CCGAGGCAAG AAACCACCGG GGAAGCGCGG
11551   ACCTTCTACT CCTCCCCTAG TCAGGAAGTT CCCCCCCGCC CCGCAGCTCG
        TGGAAGATGA GGAGGGGATC AGTCCTTCAA GGGGGGGCGG GGCGTCGAGC
11601   CGTCGTGCAG GACGTGACAA ATGGAAGTAG CACGTCTCAC TAGTCTCGTG
        GCAGCACGTC CTGCACTGTT TACCTTCATC GTGCAGAGTG ATCAGAGCAC
11651   CAGATGGACA GCACCGCTGA GCAATGGAAG CGGGTAGGCC TTTGGGGCAG
        GTCTACCTGT CGTGGCGACT CGTTACCTTC GCCCATCCGG AAACCCCGTC
11701   CGGCCAATAG CAGCTTTGCT CCTTCGCTTT CTGGGCTCAG AGGCTGGGAA
        GCCGGTTATC GTCGAAACGA GGAAGCGAAA GACCCGAGTC TCCGACCCTT
11751   GGGGTGGGTC CGGGGGCGGG CTCAGGGGCG GGCTCAGGGG CGGGGCGGGC
        CCCCACCCAG GCCCCCGCCC GAGTCCCCGC CCGAGTCCCC GCCCCGCCCG
11801   GCCCGAAGGT CCTCCGGAGG CCCGGCATTC TGCACGCTTC AAAAGCGCAC
        CGGGCTTCCA GGAGGCCTCC GGGCCGTAAG ACGTGCGAAG TTTTCGCGTG
11851   GTCTGCCGCG CTGTTCTCCT CTTCCTCATC TCCGGGCCTT TCGACCAGCT
        CAGACGGCGC GACAAGAGGA GAAGGAGTAG AGGCCCGGAA AGCTGGTCGA
11901   TACCATGACC GAGTACAAGC CCACGGTGCG CCTCGCCACC CGCGACGACG
        ATGGTACTGG CTCATGTTCG GGTGCCACGC GGAGCGGTGG GCGCTGCTGC
11951   TCCCCAGGGC CGTACGCACC CTCGCCGCCG CGTTCGCCGA CTACCCCGCC
        AGGGGTCCCG GCATGCGTGG GAGCGGCGGC GCAAGCGGCT GATGGGGCGG
12001   ACGCGCCACA CCGTCGATCC GGACCGCCAC ATCGAGCGGG TCACCGAGCT
        TGCGCGGTGT GGCAGCTAGG CCTGGCGGTG TAGCTCGCCC AGTGGCTCGA
12051   GCAAGAACTC TTCCTCACGC GCGTCGGGCT CGACATCGGC AAGGTGTGGG
        CGTTCTTGAG AAGGAGTGCG CGCAGCCCGA GCTGTAGCCG TTCCACACCC
12101   TCGCGGACGA CGGCGCCGCG GTGGCGGTCT GGACCACGCC GGAGAGCGTC
        AGCGCCTGCT GCCGCGGCGC CACCGCCAGA CCTGGTGCGG CCTCTCGCAG
12151   GAAGCGGGGG CGGTGTTCGC CGAGATCGGC CCGCGCATGG CCGAGTTGAG
        CTTCGCCCCC GCCACAAGCG GCTCTAGCCG GGCGCGTACC GGCTCAACTC
12201   CGGTTCCCGG CTGGCCGCGC AGCAACAGAT GGAAGGCCTC CTGGCGCCGC
        GCCAAGGGCC GACCGGCGCG TCGTTGTCTA CCTTCCGGAG GACCGCGGCG
12251   ACCGGCCCAA GGAGCCCGCG TGGTTCCTGG CCACCGTCGG CGTCTCGCCC
        TGGCCGGGTT CCTCGGGCGC ACCAAGGACC GGTGGCAGCC GCAGAGCGGG
12301   GACCACCAGG GCAAGGGTCT GGGCAGCGCC GTCGTGCTCC CCGGAGTGGA
        CTGGTGGTCC CGTTCCCAGA CCCGTCGCGG CAGCACGAGG GGCCTCACCT
12351   GGCGGCCGAG CGCGCCGGGG TGCCCGCCTT CCTGGAGACC TCCGCGCCCC
        CCGCCGGCTC GCGCGGCCCC ACGGGCGGAA GGACCTCTGG AGGCGCGGGG
12401   GCAACCTCCC CTTCTACGAG CGGCTCGGCT TCACCGTCAC CGCCGACGTC
        CGTTGGAGGG GAAGATGCTC GCCGAGCCGA AGTGGCAGTG GCGGCTGCAG
12451   GAGGTGCCCG AAGGACCGCG CACCTGGTGC ATGACCCGCA AGCCCGGTGC
        CTCCACGGGC TTCCTGGCGC GTGGACCACG TACTGGGCGT TCGGGCCACG
12501   CTGACGCCCG CCCCACGACC CGCAGCGCCC GACCGAAAGG AGCGCACGAC
        GACTGCGGGC GGGGTGCTGG GCGTCGCGGG CTGGCTTTCC TCGCGTGCTG
12551   CCCATGCATC GTAGAGCTCG CTGATCAGCC TCGACTGTGC CTTCTAGTTG
        GGGTACGTAG CATCTCGAGC GACTAGTCGG AGCTGACACG GAAGATCAAC
12601   CCAGCCATCT GTTGTTTGCC CCTCCCCCGT GCCTTCCTTG ACCCTGGAAG
        GGTCGGTAGA CAACAAACGG GGAGGGGGCA CGGAAGGAAC TGGGACCTTC
12651   GTGCCACTCC CACTGTCCTT TCCTAATAAA ATGAGGAAAT TGCATCGCAT
        CACGGTGAGG GTGACAGGAA AGGATTATTT TACTCCTTTA ACGTAGCGTA
12701   TGTCTGAGTA GGTGTCATTC TATTCTGGGG GGTGGGGTGG GGCAGGACAG
        ACAGACTCAT CCACAGTAAG ATAAGACCCC CCACCCCACC CCGTCCTGTC
12751   CAAGGGGGGG GATTGGGRAG ACAATAGCAG GCATGCTGGG GGGGCGGTGG
        GTTCCCCCCC CTAACCCYTC TGTTATCGTC CGTACGACCC CCCCGCCACC
12801   GGGCTATGGC TTCTGAGGCG GAAAGAACCA GCTGGGGCTC GAGATCCACT
        CCCGATACCG AAGACTCCGC CTTTCTTGGT CGACCCCGAG CTCTAGGTGA
12851   AGTTCTAGCC TCGAGGCTAG AGCGGCCGCC ACCGCGGTGG AGCTCCAATT
        TCAAGATCGG AGCTCCGATC TCGCCGGCGG TGGCGCCACC TCGAGGTTAA
12901   CGCCCTATAG TGAGTCGTAT TACGCGCGCT CACTGGCCGT CGTTTTACAA
        GCGGGATATC ACTCAGCATA ATGCGCGCGA GTGACCGGCA GCAAAATGTT
```

```
12951   CGTCGTGACT GGGAAAACCC TGGCGTTACC CAACTTAATC GCCTTGCAGC
        GCAGCACTGA CCCTTTTGGG ACCGCAATGG GTTGAATTAG CGGAACGTCG
13001   ACATCCCCCT TTCGCCAGCT GGCGTAATAG CGAAGAGGCC CGCACCGATC
        TGTAGGGGGA AAGCGGTCGA CCGCATTATC GCTTCTCCGG GCGTGGCTAG
13051   GCCCTTCCCA ACAGTTGCGC AGCCTGAATG GCGAATGGAA ATTGTAAGCG
        CGGGAAGGGT TGTCAACGCG TCGGACTTAC CGCTTACCTT TAACATTCGC
13101   TTAATATTTT GTTAAAATTC GCGTTAAATT TTTGTTAAAT CAGCTCATTT
        AATTATAAAA CAATTTTAAG CGCAATTTAA AAACAATTTA GTCGAGTAAA
13151   TTTAACCAAT AGGCCGAAAT CGGCAAAATC CCTTATAAAT CAAAAGAATA
        AAATTGGTTA TCCGGCTTTA GCCGTTTTAG GGAATATTTA GTTTTCTTAT
13201   GACCGAGATA GGGTTGAGTG TTGTTCCAGT TTGGAACAAG AGTCCACTAT
        CTGGCTCTAT CCCAACTCAC AACAAGGTCA AACCTTGTTC TCAGGTGATA
13251   TAAAGAACGT GGACTCCAAC GTCAAAGGGC GAAAAACCGT CTATCAGGGC
        ATTTCTTGCA CCTGAGGTTG CAGTTTCCCG CTTTTTGGCA GATAGTCCCG
13301   GATGGCCCAC TACGTGAACC ATCACCCTAA TCAAGTTTTT TGGGGTCGAG
        CTACCGGGTG ATGCACTTGG TAGTGGGATT AGTTCAAAAA ACCCCAGCTC
13351   GTGCCGTAAA GCACTAAATC GGAACCCTAA AGGGAGCCCC CGATTTAGAG
        CACGGCATTT CGTGATTTAG CCTTGGGATT TCCCTCGGGG GCTAAATCTC
13401   CTTGACGGGG AAAGCCGGCG AACGTGGCGA GAAAGGAAGG GAAGAAAGCG
        GAACTGCCCC TTTCGGCCGC TTGCACCGCT CTTTCCTTCC CTTCTTTCGC
13451   AAAGGAGCGG GCGCTAGGGC GCTGGCAAGT GTAGCGGTCA CGCTGCGCGT
        TTTCCTCGCC CGCGATCCCG CGACCGTTCA CATCGCCAGT GCGACGCGCA
13501   AACCACCACA CCCGCCGCGC TTAATGCGCC GCTACAGGGC GCGTCAG
        TTGGTGGTGT GGGCGGCGCG AATTACGCGG CGATGTCCCG CGCAGTC
```

FIG. 12(contd.)

FIG. 13

S alboniger puromycin N-acetyl transferase nucleotide sequence

```
   1  ATGACCGAGT ACAAGCCCAC GGTGCGCCTC GCCACCCGCG ACGACGTCCC
      TACTGGCTCA TGTTCGGGTG CCACGCGGAG CGGTGGGCGC TGCTGCAGGG
  51  CAGGGCCGTA CGCACCCTCG CCGCCGCGTT CGCCGACTAC CCCGCCACGC
      GTCCCGGCAT GCGTGGGAGC GGCGGCGCAA GCGGCTGATG GGGCGGTGCG
 101  GCCACACCGT CGATCCGGAC CGCCACATCG AGCGGGTCAC CGAGCTGCAA
      CGGTGTGGCA GCTAGGCCTG GCGGTGTAGC TCGCCCAGTG GCTCGACGTT
 151  GAACTCTTCC TCACGCGCGT CGGGCTCGAC ATCGGCAAGG TGTGGGTCGC
      CTTGAGAAGG AGTGCGCGCA GCCCGAGCTG TAGCCGTTCC ACACCCAGCG
 201  GGACGACGGC GCCGCGGTGG CGGTCTGGAC CACGCCGGAG AGCGTCGAAG
      CCTGCTGCCG CGGCGCCACC GCCAGACCTG GTGCGGCCTC TCGCAGCTTC
 251  CGGGGGCGGT GTTCGCCGAG ATCGGCCCGC GCATGGCCGA GTTGAGCGGT
      GCCCCCGCCA CAAGCGGCTC TAGCCGGGCG CGTACCGGCT CAACTCGCCA
 301  TCCCGGCTGG CCGCGCAGCA ACAGATGGAA GGCCTCCTGG CGCCGCACCG
      AGGGCCGACC GGCGCGTCGT TGTCTACCTT CCGGAGGACC GCGGCGTGGC
 351  GCCCAAGGAG CCCGCGTGGT TCCTGGCCAC CGTCGGCGTC TCGCCCGACC
      CGGGTTCCTC GGGCGCACCA AGGACCGGTG GCAGCCGCAG AGCGGGCTGG
 401  ACCAGGGCAA GGGTCTGGGC AGCGCCGTCG TGCTCCCCGG AGTGGAGGCG
      TGGTCCCGTT CCCAGACCCG TCGCGGCAGC ACGAGGGGCC TCACCTCCGC
 451  GCCGAGCGCG CCGGGGTGCC CGCCTTCCTG GAGACCTCCG CGCCCCGCAA
      CGGCTCGCGC GGCCCCACGG GCGGAAGGAC CTCTGGAGGC GCGGGGCGTT
 501  CCTCCCCTTC TACGAGCGGC TCGGCTTCAC CGTCACCGCC GACGTCGAGG
      GGAGGGGAAG ATGCTCGCCG AGCCGAAGTG GCAGTGGCGG CTGCAGCTCC
 551  TGCCCGAAGG ACCGCGCACC TGGTGCATGA CCCGCAAGCC CGGTGCCTGA
      ACGGGCTTCC TGGCGCGTGG ACCACGTACT GGGCGTTCGG GCCACGGACT
```

FIG. 14

Modified *S alboniger* puromycin N-acetyl transferase nucleotide sequence

```
   1  ATGACTGAAT ACAAACCAAC TGTTCGCCTG GCAACTCGTG ATGATGTTCC
      TACTGACTTA TGTTTGGTTG ACAAGCGGAC CGTTGAGCAC TACTACAAGG
  51  ACGTGCAGTT CGCACCCTGG CTGCTGCATT TGCTGACTAC CCTGCAACCC
      TGCACGTCAA GCGTGGGACC GACGACGTAA ACGACTGATG GGACGTTGGG
 101  GTCACACTGT GGACCCAGAC CGCCACATTG AACGTGTGAC TGAACTGCAG
      CAGTGTGACA CCTGGGTCTG GCGGTGTAAC TTGCACACTG ACTTGACGTC
 151  GAGCTGTTCC TGACCCGTGT GGGCCTGGAC ATTGGCAAAG TGTGGGTGGC
      CTCGACAAGG ACTGGGCACA CCCGGACCTG TAACCGTTTC ACACCCACCG
 201  AGATGATGGT GCTGCTGTGG CAGTGTGGAC CACCCCTGAA TCTGTTGAAG
      TCTACTACCA CGACGACACC GTCACACCTG GTGGGGACTT AGACAACTTC
 251  CTGGTGCAGT GTTTGCTGAG ATTGGCCCAC GCATGGCAGA ACTGTCTGGC
      GACCACGTCA CAAACGACTC TAACCGGGTG CGTACCGTCT TGACAGACCG
 301  AGCCGCCTGG CAGCACAACA GCAGATGGAA GGTCTGCTGG CACCACACCG
      TCGGCGGACC GTCGTGTTGT CGTCTACCTT CCAGACGACC GTGGTGTGGC
 351  CCCAAAAGAA CCTGCTTGGT TCCTGGCAAC TGTGGGTGTG AGCCCTGACC
      GGGTTTTCTT GGACGAACCA AGGACCGTTG ACACCCACAC TCGGGACTGG
 401  ACCAGGGTAA GGGCCTGGGC TCTGCAGTGG TGCTGCCTGG TGTGGAAGCA
      TGGTCCCATT CCCGGACCCG AGACGTCACC ACGACGGACC ACACCTTCGT
 451  GCTGAACGTG CAGGTGTGCC TGCTTTCCTG GAGACCTCAG CTCCACGCAA
      CGACTTGCAC GTCCACACGG ACGAAAGGAC CTCTGGAGTC GAGGTGCGTT
 501  CCTGCCTTTC TATGAACGCC TGGGCTTCAC TGTGACTGCT GATGTGGAAG
      GGACGGAAAG ATACTTGCGG ACCCGAAGTG ACACTGACGA CTACACCTTC
 551  TGCCAGAAGG CCCACGCACT TGGTGCATGA CTCGCAAACC AGGTGCTTAA
      ACGGTCTTCC GGGTGCGTGA ACCACGTACT GAGCGTTTGG TCCACGAATT
```

FIG. 15 *S fradiae* aminoglycoside phosphotransferase nucleotide sequence

```
   1  ATGGACGACA GCACGTTGCG CCGGAAGTAC CCGCACCACG AGTGGCACGC
      TACCTGCTGT CGTGCAACGC GGCCTTCATG GGCGTGGTGC TCACCGTGCG
  51  AGTGAACGAA GGAGACTCGG GCGCCTTCGT CTACCAGCTC ACCGGCGGCC
      TCACTTGCTT CCTCTGAGCC CGCGGAAGCA GATGGTCGAG TGGCCGCCGG
 101  CCGAGCCCCA GCCCGAGCTC TACGCGAAGA TCGCCCCCCG CGCCCCCGAG
      GGCTCGGGGT CGGGCTCGAG ATGCGCTTCT AGCGGGGGGC GCGGGGGCTC
 151  AACTCCGCCT TCGACCTGTC CGGCGAGGCC GACCGGCTGG AGTGGCTCCA
      TTGAGGCGGA AGCTGGACAG GCCGCTCCGG CTGGCCGACC TCACCGAGGT
 201  CCGCCACGGG ATCCCCGTCC CCCGCGTCGT CGAGCGCGGT GCCGACGACA
      GGCGGTGCCC TAGGGGCAGG GGGCGCAGCA GCTCGCGCCA CGGCTGCTGT
 251  CCGCCGCGTG GCTCGTCACG GAGGCCGTCC CCGGCGTCGC GGCGGCCGAG
      GGCGGCGCAC CGAGCAGTGC CTCCGGCAGG GGCCGCAGCG CCGCCGGCTC
 301  GAGTGGCCCG AGCACCAGCG GTTCGCCGTG GTCGAGGCGA TGGCGGAGCT
      CTCACCGGGC TCGTGGTCGC CAAGCGGCAC CAGCTCCGCT ACCGCCTCGA
 351  GGCCCGCGCC CTCCACGAGC TGCCCGTGGA GGACTGCCCC TCCGACCGGC
      CCGGGCGCGG GAGGTGCTCG ACGGGCACCT CCTGACGGGG AGGCTGGCCG
 401  GCCTCGACGC GGCGGTCGCC GAGGCCCGGC GGAACGTCGC CGAGGGCTTG
      CGGAGCTGCG CCGCCAGCGG CTCCGGGCCG CCTTGCAGCG GCTCCCGAAC
 451  GTGGACCTCG ACGACCTGCA GGAGGAGCGG GCCGGGTGGA CCGGCGACCA
      CACCTGGAGC TGCTGGACGT CCTCCTCGCC CGGCCCACCT GGCCGCTGGT
 501  GCTCCTGGCG GAGCTCGACC GCACCCGTCC CGAGAAGGAG GACCTGGTCG
      CGAGGACCGC CTCGAGCTGG CGTGGGCAGG GCTCTTCCTC CTGGACCAGC
 551  TCTGCCATGG CGACCTGTGC CCCAACAACG TCCTGCTCGA CCCCGGGACC
      AGACGGTACC GCTGGACACG GGGTTGTTGC AGGACGAGCT GGGGCCCTGG
 601  TGCCGGGTCA CCGGCGTGAT CGACGTCGGC CGCCTCGGGG TCGCCGACCG
      ACGGCCCAGT GGCCGCACTA GCTGCAGCCG GCGGAGCCCC AGCGGCTGGC
 651  CCACGCCGAC ATCGCCTTGG CCGCCCGCGA GCTGGAGATC GACGAGGACC
      GGTGCGGCTG TAGCGGAACC GGCGGGCGCT CGACCTCTAG CTGCTCCTGG
 701  CCTGGTTCGG CCCCGCCTAC GCCGAGCGGT TCCTGGAGCG GTACGGCGCC
      GGACCAAGCC GGGGCGGATG CGGCTCGCCA AGGACCTCGC CATGCCGCGG
 751  CACCGCGTCG ACAAGGAGAA GCTGGCCTTC TACCAGCTTC TCGACGAGTT
      GTGGCGCAGC TGTTCCTCTT CGACCGGAAG ATGGTCGAAG AGCTGCTCAA
 801  CTTCTAG
      GAAGATC
```

FIG. 16 *S hygroscopicus* hygromycin phosphotransferase nucleotide sequence

```
   1  ATGACACAAG AATCCCTGTT ACTTCTCGAC CGTATTGATT CGGATGATTC
      TACTGTGTTC TTAGGGACAA TGAAGAGCTG GCATAACTAA GCCTACTAAG
  51  CTACGCGAGC CTGCGGAACG ACCAGGAATT CTGGGAGCCG CTGGCCCGCC
      GATGCGCTCG GACGCCTTGC TGGTCCTTAA GACCCTCGGC GACCGGGCGG
 101  GAGCCCTGGA GGAGCTCGGG CTGCCGGTGC CGCCGGTGCT GCGGGTGCCC
      CTCGGGACCT CCTCGAGCCC GACGGCCACG GCGGCCACGA CGCCCACGGG
 151  GGCGAGAGCA CCAACCCCGT ACTGGTCGGC GAGCCCGACC CGGTGATCAA
      CCGCTCTCGT GGTTGGGGCA TGACCAGCCG CTCGGGCTGG GCCACTAGTT
 201  GCTGTTCGGC GAGCACTGGT GCGGTCCGGA GAGCCTCGCG TCGGAGTCGG
      CGACAAGCCG CTCGTGACCA CGCCAGGCCT CTCGGAGCGC AGCCTCAGCC
 251  AGGCGTACGC GGTCCTGGCG GACGCCCCGG TGCCGGTGCC CCGCCTCCTC
      TCCGCATGCG CCAGGACCGC CTGCGGGGCC ACGGCCACGG GGCGGAGGAG
 301  GGCCGCGGCG AGCTGCGGCC CGGCACCGGA GCCTGGCCGT GGCCCTACCT
      CCGGCGCCGC TCGACGCCGG GCCGTGGCCT CGGACCGGCA CCGGGATGGA
 351  GGTGATGAGC CGGATGACCG GCACCACCTG GCGGTCCGCG ATGGACGGCA
      CCACTACTCG GCCTACTGGC CGTGGTGGAC CGCCAGGCGC TACCTGCCGT
 401  CGACCGACCG GAACGCGCTG CTCGCCCTGG CCCGCGAACT CGGCCGGGTG
      GCTGGCTGGC CTTGCGCGAC GAGCGGGACC GGGCGCTTGA GCCGGCCCAC
 451  CTCGGCCGGC TGCACAGGGT GCCGCTGACC GGGAACACCG TGCTCACCCC
      GAGCCGGCCG ACGTGTCCCA CGGCGACTGG CCCTTGTGGC ACGAGTGGGG
 501  CCATTCCGAG GTCTTCCCGG AACTGCTGCG GGAACGCCGC GCGGCGACCG
      GGTAAGGCTC CAGAAGGGCC TTGACGACGC CCTTGCGGCG CGCCGCTGGC
 551  TCGAGGACCA CCGCGGGTGG GGCTACCTCT CGCCCCGGCT GCTGGACCGC
      AGCTCCTGGT GGCGCCCACC CCGATGGAGA GCGGGGCCGA CGACCTGGCG
 601  CTGGAGGACT GGCTGCCGGA CGTGGACACG CTGCTGGCCG GCCGCGAACC
      GACCTCCTGA CCGACGGCCT GCACCTGTGC GACGACCGGC CGGCGCTTGG
 651  CCGGTTCGTC CACGGCGACC TGCACGGGAC CAACATCTTC GTGGACCTGG
      GGCCAAGCAG GTGCCGCTGG ACGTGCCCTG GTTGTAGAAG CACCTGGACC
 701  CCGCGACCGA GGTCACCGGG ATCGTCGACT TCACCGACGT CTATGCGGGA
      GGCGCTGGCT CCAGTGGCCC TAGCAGCTGA AGTGGCTGCA GATACGCCCT
 751  GACTCCCGCT ACAGCCTGGT GCAACTGCAT CTCAACGCCT TCCGGGGCGA
      CTGAGGGCGA TGTCGGACCA CGTTGACGTA GAGTTGCGGA AGGCCCCGCT
 801  CCGCGAGATC CTGGCCGCGC TGCTCGACGG GGCGCAGTGG AAGCGGACCG
      GGCGCTCTAG GACCGGCGCG ACGAGCTGCC CCGCGTCACC TTCGCCTGGC
 851  AGGACTTCGC CCGCGAACTG CTCGCCTTCA CCTTCCTGCA CGACTTCGAG
      TCCTGAAGCG GGCGCTTGAC GAGCGGAAGT GGAAGGACGT GCTGAAGCTC
 901  GTGTTCGAGG AGACCCCGCT GGATCTCTCC GGCTTCACCG ATCCGGAGGA
      CACAAGCTCC TCTGGGGCGA CCTAGAGAGG CCGAAGTGGC TAGGCCTCCT
 951  ACTGGCGCAG TTCCTCTGGG GGCCGCCGGA CACCGCCCCC GGCGCCTGA
      TGACCGCGTC AAGGAGACCC CCGGCGGCCT GTGGCGGGGG CCGCGGACT
```

FIG. 17 *E coli* aminocyclitol phosphotransferase(hygro^r)nucleotide sequence

```
   1  ATGAAAAAGC CTGAACTCAC CGCGACGTCT GTCGCGAAGT TTCTGATCGA
      TACTTTTTCG GACTTGAGTG GCGCTGCAGA CAGCGCTTCA AAGACTAGCT
  51  AAAGTTCGAC AGCGTCTCCG ACCTGATGCA GCTCTCGGAG GGCGAAGAAT
      TTTCAAGCTG TCGCAGAGGC TGGACTACGT CGAGAGCCTC CCGCTTCTTA
 101  CTCGTGCTTT CAGCTTCGAT GTAGGAGGGC GTGGATATGT CCTGCGGGTA
      GAGCACGAAA GTCGAAGCTA CATCCTCCCG CACCTATACA GGACGCCCAT
 151  AATAGCTGCG CCGATGGTTT CTACAAAGAT CGTTATGTTT ATCGGCACTT
      TTATCGACGC GGCTACCAAA GATGTTTCTA GCAATACAAA TAGCCGTGAA
 201  TGCATCGGCC GCGCTCCCGA TTCCGGAAGT GCTTGACATT GGGGAATTCA
      ACGTAGCCGG CGCGAGGGCT AAGGCCTTCA CGAACTGTAA CCCCTTAAGT
 251  GCGAGAGCCT GACCTATTGC ATCTCCCGCC GTGCACAGGG TGTCACGTTG
      CGCTCTCGGA CTGGATAACG TAGAGGGCGG CACGTGTCCC ACAGTGCAAC
 301  CAAGACCTGC CTGAAACCGA ACTGCCCGCT GTTCTGCAAC CCGTCGCGGA
      GTTCTGGACG GACTTTGGCT TGACGGGCGA CAAGACGTTG GGCAGCGCCT
 351  GCTCATGGAT GCGATCGCTG CGGCCGATCT TAGCCAGACG AGCGGGTTCG
      CGAGTACCTA CGCTAGCGAC GCCGGCTAGA ATCGGTCTGC TCGCCCAAGC
 401  GCCCATTCGG ACCGCAAGGA ATCGGTCAAT ACACTACATG GCGTGATTTC
      CGGGTAAGCC TGGCGTTCCT TAGCCAGTTA TGTGATGTAC CGCACTAAAG
 451  ATATGCGCGA TTGCTGATCC CCATGTGTAT CACTGGCAAA CTGTGATGGA
      TATACGCGCT AACGACTAGG GGTACACATA GTGACCGTTT GACACTACCT
 501  CGACACCGTC AGTGCGTCCG TCGCGCAGGC TCTCGATGAG CTGATGCTTT
      GCTGTGGCAG TCACGCAGGC AGCGCGTCCG AGAGCTACTC GACTACGAAA
 551  GGGCCGAGGA CTGCCCCGAA GTCCGGCACC TCGTGCACGC GGATTTCGGC
      CCCGGCTCCT GACGGGGCTT CAGGCCGTGG AGCACGTGCG CCTAAAGCCG
 601  TCCAACAATG TCCTGACGGA CAATGGCCGC ATAACAGCGG TCATTGACTG
      AGGTTGTTAC AGGACTGCCT GTTACCGGCG TATTGTCGCC AGTAACTGAC
 651  GAGCGAGGCG ATGTTCGGGG ATTCCCAATA CGAGGTCGCC AACATCTTCT
      CTCGCTCCGC TACAAGCCCC TAAGGGTTAT GCTCCAGCGG TTGTAGAAGA
 701  TCTGGAGGCC GTGGTTGGCT TGTATGGAGC AGCAGACGCG CTACTTCGAG
      AGACCTCCGG CACCAACCGA ACATACCTCG TCGTCTGCGC GATGAAGCTC
 751  CGGAGGCATC CGGAGCTTGC AGGATCGCCG CGGCTCCGGG CGTATATGCT
      GCCTCCGTAG GCCTCGAACG TCCTAGCGGC GCCGAGGCCC GCATATACGA
 801  CCGCATTGGT CTTGACCAAC TCTATCAGAG CTTGGTTGAC GGCAATTTCG
      GGCGTAACCA GAACTGGTTG AGATAGTCTC GAACCAACTG CCGTTAAAGC
 851  ATGATGCAGC TTGGGCGCAG GGTCGATGCG ACGCAATCGT CCGATCCGGA
      TACTACGTCG AACCCGCGTC CCAGCTACGC TGCGTTAGCA GGCTAGGCCT
 901  GCCGGGACTG TCGGGCGTAC ACAAATCGCC CGCAGAAGCG CGGCCGTCTG
      CGGCCCTGAC AGCCCGCATG TGTTTAGCGG GCGTCTTCGC GCCGGCAGAC
 951  GACCGATGGC TGTGTAGAAG TACTCGCCGA TAGTGGAAAC CGACGCCCCA
      CTGGCTACCG ACACATCTTC ATGAGCGGCT ATCACCTTTG GCTGCGGGGT
1001  GCACTCGTCC GAGGGCAAAG GAATGA
      CGTGAGCAGG CTCCCGTTTC CTTACT
```

FIG. 18 Tn5 (*Klebsiella pneumoniae*) neomycin phosphotransferease nucleotide sequence

```
   1  ATGATTGAAC AAGATGGATT GCACGCAGGT TCTCCGGCCG CTTGGGTGGA
      TACTAACTTG TTCTACCTAA CGTGCGTCCA AGAGGCCGGC GAACCCACCT
  51  GAGGCTATTC GGCTATGACT GGGCACAACA GACAATCGGC TGCTCTGATG
      CTCCGATAAG CCGATACTGA CCCGTGTTGT CTGTTAGCCG ACGAGACTAC
 101  CCGCCGTGTT CCGGCTGTCA GCGCAGGGGC GCCCGGTTCT TTTTGTCAAG
      GGCGGCACAA GGCCGACAGT CGCGTCCCCG CGGGCCAAGA AAAACAGTTC
 151  ACCGACCTGT CCGGTGCCCT GAATGAACTG CAAGACGAGG CAGCGCGGCT
      TGGCTGGACA GGCCACGGGA CTTACTTGAC GTTCTGCTCC GTCGCGCCGA
 201  ATCGTGGCTG GCCACGACGG GCGTTCCTTG CGCAGCTGTG CTCGACGTTG
      TAGCACCGAC CGGTGCTGCC CGCAAGGAAC GCGTCGACAC GAGCTGCAAC
 251  TCACTGAAGC GGGAAGGGAC TGGCTGCTAT TGGGCGAAGT GCCGGGGCAG
      AGTGACTTCG CCCTTCCCTG ACCGACGATA ACCCGCTTCA CGGCCCCGTC
 301  GATCTCCTGT CATCTCACCT TGCTCCTGCC GAGAAAGTAT CCATCATGGC
      CTAGAGGACA GTAGAGTGGA ACGAGGACGG CTCTTTCATA GGTAGTACCG
 351  TGATGCAATG CGGCGGCTGC ATACGCTTGA TCCGGCTACC TGCCCATTCG
      ACTACGTTAC GCCGCCGACG TATGCGAACT AGGCCGATGG ACGGGTAAGC
 401  ACCACCAAGC GAAACATCGC ATCGAGCGAG CACGTACTCG GATGGAAGCC
      TGGTGGTTCG CTTTGTAGCG TAGCTCGCTC GTGCATGAGC CTACCTTCGG
 451  GGTCTTGTCG ATCAGGATGA TCTGGACGAA GAGCATCAGG GGCTCGCGCC
      CCAGAACAGC TAGTCCTACT AGACCTGCTT CTCGTAGTCC CCGAGCGCGG
 501  AGCCGAACTG TTCGCCAGGC TCAAGGCGAG CATGCCCGAC GGCGAGGATC
      TCGGCTTGAC AAGCGGTCCG AGTTCCGCTC GTACGGGCTG CCGCTCCTAG
 551  TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT GGTGGAAAAT
      AGCAGCACTG GGTACCGCTA CGGACGAACG GCTTATAGTA CCACCTTTTA
 601  GGCCGCTTTT CTGGATTCAT CGACTGTGGC CGGCTGGGTG TGGCGGACCG
      CCGGCGAAAA GACCTAAGTA GCTGACACCG GCCGACCCAC ACCGCCTGGC
 651  CTATCAGGAC ATAGCGTTGG CTACCCGTGA TATTGCTGAA GAGCTTGGCG
      GATAGTCCTG TATCGCAACC GATGGGCACT ATAACGACTT CTCGAACCGC
 701  GCGAATGGGC TGACCGCTTC CTCGTGCTTT ACGGTATCGC CGCTCCCGAT
      CGCTTACCCG ACTGGCGAAG GAGCACGAAA TGCCATAGCG GCGAGGGCTA
 751  TCGCAGCGCA TCGCCTTCTA TCGCCTTCTT GACGAGTTCT TCTGA
      AGCGTCGCGT AGCGGAAGAT AGCGGAAGAA CTGCTCAAGA AGACT
```

FIG. 19 Mouse hnRNP *Hind*III fragment nucleotide sequence

```
   1  AAGCTTTTAA CCCTCTATCC CTTTAAACTT CCTTGATCCA GTGTAAGCAC
      TTCGAAAATT GGGAGATAGG GAAATTTGAA GGAACTAGGT CACATTCGTG
  51  CTCCTAGAAA GTCAGTAGAC AATAAAACAA AAGTTCTGCT TCACCGATTT
      GAGGATCTTT CAGTCATCTG TTATTTTGTT TTCAAGACGA AGTGGCTAAA
 101  ACATTTATAA CCAAATACCC TTCACCAATA CAATAAAAAA ACAAAACAAC
      TGTAAATATT GGTTTATGGG AAGTGGTTAT GTTATTTTTT TGTTTTGTTG
 151  AAAAAACCCC AACCATCTGA GAAATAATCT TCTCCTTTCC CAGCTTTATT
      TTTTTTGGGG TTGGTAGACT CTTTATTAGA AGAGGAAAGG GTCGAAATAA
 201  CCCAGGATTC TACATGACCA AATTACCAGA GTCACCACTC ATTTTAATCA
      GGGTCCTAAG ATGTACTGGT TTAATGGTCT CAGTGGTGAG TAAAATTAGT
 251  CAACATAGTG TCAAATAACT AGAAAACATG AGACAACAAT GGAGAGCTGA
      GTTGTATCAC AGTTTATTGA TCTTTTGTAC TCTGTTGTTA CCTCTCGACT
 301  GTAACTATTA GTAGTAGTAC TTTACCAGAG AATGGCCTCT ATAGGCTCAC
      CATTGATAAT CATCATCATG AAATGGTCTC TTACCGGAGA TATCCGAGTG
 351  ATGTAGGAAT GGTTGGTCCC CAGGTGGTAG GTAGAGCTGT TTGAGGATTA
      TACATCCTTA CCAACCAGGG GTCCACCATC CATCTCGACA AACTCCTAAT
 401  CGTGGCCTTC TTGGATGGGG GGTGGGGGTG GGGTGGGAGG GTTGGGTGGT
      GCACCGGAAG AACCTACCCC CCACCCCCAC CCCACCCTCC CAACCCACCA
 451  GGGTACTTAA GAGGTTTCAA AAGTCAATAT TGTTTGCATT TAGCTCTTCC
      CCCATGAATT CTCCAAAGTT TTCAGTTATA ACAAACGTAA ATCGAGAAGG
 501  TTGTACTTGT GGATCAAACA CAACCTGTCA GCTACTGCTT CAAATGTCAT
      AACATGAACA CCTAGTTTGT GTTGGACAGT CGATGACGAA GTTTACAGTA
 551  GCCTGCTGCC ATCTTCTCAG CAGGATGGTC ATGGCCTCAC CCTCTTCAAC
      CGGACGACGG TAGAAGAGTC GTCCTACCAG TACCGGAGTG GGAGAAGTTG
 601  TGTAAATCTT TCTTTCTTTT CTTCTTTTTC TTTTGGTTTC GAGACAGGGT
      ACATTTAGAA AGAAAGAAAA GAAGAAAAAG AAAACCAAAG CTCTGTCCCA
 651  TTCTCTGTAT AGTCCTGGCT GTCCTGGAAC TCACTTTGTA GACCAGGCTG
      AAGAGACATA TCAGGACCGA CAGGACCTTG AGTGAAACAT CTGGTCCGAC
 701  GCCTTGAACT CAGAAATCCG CCTGCCTCTG CCTCCCTAGC ACTGGGATTA
      CGGAACTTGA GTCTTTAGGC GGACGGAGAC GGAGGGATCG TGACCCTAAT
 751  AAGGCGTGCG CCACCACGCC CAGCTTTCAA CTGGAAATCT TAATAAACTT
      TTCCGCACGC GGTGGTGCGG GTCGAAAGTT GACCTTTAGA ATTATTTGAA
 801  TCCTAGAAGT GGCCTTGGTT ATGGGAGCTT ATCACAGCAA TAGAACAGCA
      AGGATCTTCA CCGGAACCAA TACCCTCGAA TAGTGTCGTT ATCTTGTCGT
 851  ATTATGACTG GAGTATGATA GTTAAAAACA AGCAAGCAAG CAAGCAAACA
      TAATACTGAC CTCATACTAT CAATTTTTGT TCGTTCGTTC GTTCGTTTGT
 901  CACACACCAA AACAACAAAA CCCCAAGACA GAGTCACATG TAGCCCAGGC
      GTGTGTGGTT TTGTTGTTTT GGGGTTCTGT CTCAGTGTAC ATCGGGTCCG
 951  TAGCCTCCAA ATTCACTATA TAACTGAAGA AGACCCCTAA TTCCCATTCC
      ATCGGAGGTT TAAGTGATAT ATTGACTTCT TCTGGGGATT AAGGGTAAGG
1001  TCTAGAATCT ATACCTCAAG TACTGAATGG CTTGGTTCAC AATACCCCAC
      AGATCTTAGA TATGGAGTTC ATGACTTACC GAACCAAGTG TTATGGGGTG
1051  TAAATGATTG GTCTTACTAA GTGCAACAAG GTAAACCTAA AACTTCAGCC
      ATTTACTAAC CAGAATGATT CACGTTGTTC CATTTGGATT TTGAAGTCGG
1101  CTCAGACATC CCTTTTCCAG TATCAATTTA TAAAATTAGA TCCCAAGGAT
      GAGTCTGTAG GGAAAAGGTC ATAGTTAAAT ATTTTAATCT AGGGTTCCTA
1151  AAAAATTAAT TGTAAAGTAA AATCAGAGTT CTAGCATCAA CTACAGGCTC
      TTTTTAATTA ACATTTCATT TTAGTCTCAA GATCGTAGTT GATGTCCGAG
1201  AACCATGGGG ACCACAAATA AACTAAAAGG GATAAGACTG GCTTCCCCAT
      TTGGTACCCC TGGTGTTTAT TTGATTTTCC CTATTCTGAC CGAAGGGGTA
1251  AATTATTACA TTTAGATAAT TTTCCTGACT ACTCAACAAA GCTAAAATAT
      TTAATAATGT AAATCTATTA AAAGGACTGA TGAGTTGTTT CGATTTTATA
1301  CACCACTGGT TTATTTTCTC CTTCTAGGGT TTAAGCTCAC TCTGAGGAGG
      GTGGTGACCA AATAAAAGAG GAAGATCCCA AATTCGAGTG AGACTCCTCC
```

FIG. 19(contd.)

```
1351  GGCATGCGGC ACACACTCAT AGCATCCAGG AAATAGAAAT ATGGTGACTA
      CCGTACGCCG TGTGTGAGTA TCGTAGGTCC TTTATCTTTA TACCACTGAT
1401  TCATGGGTTC AGGGCCAACC TAGGCTTTAG AGAAAAACCT TGTCCCACAA
      AGTACCCAAG TCCCGGTTGG ATCCGAAATC TCTTTTTGGA ACAGGGTGTT
1451  ACCAAAAATG TCTCTTTTTT ATTCTATCAG GGGTGGATGG ATTTGTTAAA
      TGGTTTTTAC AGAGAAAAAA TAAGATAGTC CCCACCTACC TAAACAATTT
1501  GAAGTGCTTT TAAAAACCTT GAGATGGTTA TTTAGAAGTC CCCATGGGAT
      CTTCACGAAA ATTTTTGGAA CTCTACCAAT AAATCTTCAG GGGTACCCTA
1551  ACCAAAATAA CCCACTATTT ATATGCCCAA GCATTTCACC TCCACAACAG
      TGGTTTTATT GGGTGATAAA TATACGGGTT CGTAAAGTGG AGGTGTTGTC
1601  TGCTATGCAC CCTTTAACAT TTTTGAGACA GTAGCCCAGT CTAGTCTTTA
      ACGATACGTG GGAAATTGTA AAAACTCTGT CATCGGGTCA GATCAGAAAT
1651  ACTTGCAGTG ATTTTTCCTG ATTCAGCTTC TCCCAGTGCT GGAATTATAG
      TGAACGTCAC TAAAAAGGAC TAAGTCGAAG AGGGTCACGA CCTTAATATC
1701  GTATGCACCA CCATGTGTAA CTACAGATGC TACTTAAAAA TTTTTTAAAG
      CATACGTGGT GGTACACATT GATGTCTACG ATGAATTTTT AAAAAATTTC
1751  GAATCACAAA AATAACCCCC TATCAAATGC CTAGTCCCTC TAACCATCAC
      CTTAGTGTTT TTATTGGGGG ATAGTTTACG GATCAGGGAG ATTGGTAGTG
1801  CAAGTGAAGG ATCACGCAGG AAAAAAAAAA TCACCAGCAG CACCTCAGAA
      GTTCACTTCC TAGTGCGTCC TTTTTTTTTT AGTGGTCGTC GTGGAGTCTT
1851  CCAGGATACT CAGTCCATCA GCATCCAGGG CCATACCCAC ACTCACAGCA
      GGTCCTATGA GTCAGGTAGT CGTAGGTCCC GGTATGGGTG TGAGTGTCGT
1901  TCTCCACAGT TTACCAGATG ATTCATGCTT ATCACTGTAT TGGGTCATCT
      AGAGGTGTCA AATGGTCTAC TAAGTACGAA TAGTGACATA ACCCAGTAGA
1951  AAGAGTGACC ATCAGGGCTT CTGATCACAG AATCTAGTCC ACTTTGCAGA
      TTCTCACTGG TAGTCCCGAA GACTAGTGTC TTAGATCAGG TGAAACGTCT
2001  CCAGTTGAAG TCATGCACTA TATGAGATAG AAATACCCTC TTGCTCATTT
      GGTCAACTTC AGTACGTGAT ATACTCTATC TTTATGGGAG AACGAGTAAA
2051  TGGTCAGAAA TTCAAGGATA AAAACCCATG TTTTGTTAAT GCACACCTCC
      ACCAGTCTTT AAGTTCCTAT TTTTGGGTAC AAAACAATTA CGTGTGGAGG
2101  ATATGATTGA GATCAATGTG TCCTAATTAA TGTAGAAACC ACAACTGTAA
      TATACTAACT CTAGTTACAC AGGATTAATT ACATCTTTGG TGTTGACATT
2151  ATTTCACTCT TTTGACATGA ATCTTTTTCT AGACAGGGTC TTGGATGCAG
      TAAAGTGAGA AAACTGTACT TAGAAAAAGA TCTGTCCCAG AACCTACGTC
2201  CCCCGACTAC CCAGAATTTT GGAATCCAGG CTAGCCTCAA ACTCAAGGCA
      GGGGCTGATG GGTCTTAAAA CCTTAGGTCC GATCGGAGTT TGAGTTCCGT
2251  ATCTGCTTGC TTCAGCTTCT CACAGGCTGG ATCACAAACA TACACCTTCA
      TAGACGAACG AAGTCGAAGA GTGTCCGACC TAGTGTTTGT ATGTGGAAGT
2301  GACCCATTTT TTTTTCCTCC CTCCGTTTTT GGTTTCTCTG TGTAGCCCTG
      CTGGGTAAAA AAAAAGGAGG GAGGCAAAAA CCAAAGAGAC ACATCGGGAC
2351  GGTGTCCGTG GACTCGCTGT GTAGATCTAT CTACCAGCCT CTGTCTTGGA
      CCACAGGCAC CTGAGCGACA CATCTAGATA GATGGTCGGA GACAGAACCT
2401  GTACTGGGAT TAAAGTTGTG GGCTACCACT GCCTGGCTGA CCCAGTTTTA
      CATGACCCTA ATTTCAACAC CCGATGGTGA CGGACCGACT GGGTCAAAAT
2451  TTTATTTTAA ATATAACTTG ACAAAAATAA ATTTGTCTAA CTTACTAGAA
      AAATAAAATT TATATTGAAC TGTTTTTATT TAAACAGATT GAATGATCTT
2501  ATCCCAAGAA AACTAACACT GGATTTAGCA ACAGTCAGAA ATCGCTGAAA
      TAGGGTTCTT TTGATTGTGA CCTAAATCGT TGTCAGTCTT TAGCGACTTT
2551  AGAAACAGAA TTGATCTAAC AGTCTTAGAT CACTCCTAGA CAGTTTGTAA
      TCTTTGTCTT AACTAGATTG TCAGAATCTA GTGAGGATCT GTCAAACATT
2601  TTCTTGCTCA TGGCAACGTG AGCTCTATCT AACTCACTCT CTGTGCACTA
      AAGAACGAGT ACCGTTGCAC TCGAGATAGA TTGAGTGAGA GACACGTGAT
2651  ATGAATGCTC AGTGTCTCCA GAACAGCACA GCTTCCAGGG TAATCATGCC
      TACTTACGAG TCACAGAGGT CTTGTCGTGT CGAAGGTCCC ATTAGTACGG
2701  AACCCACAAG ACTTTTATAG AGCTGTCCAC GACTCTTCCC CCATTCAGCT
      TTGGGTGTTC TGAAAATATC TCGACAGGTG CTGAGAAGGG GGTAAGTCGA
2751  CATTAACAAT ATGATGGAGC TCCTGTGTGG AAATCAAGGC ACACTCTGGT
      GTAATTGTTA TACTACCTCG AGGACACACC TTTAGTTCCG TGTGAGACCA
```

FIG. 19(contd.)

```
2801  AGAAACTTGT TTTTTCTTTC CACTTTTCCT TGGGCTCTGA AGATTGAGCT
      TCTTTGAACA AAAAAGAAAG GTGAAAAGGA ACCCGAGACT TCTAACTCGA
2851  GTTTTATAAC CCACAAACAT GCATTTTTTA CCTCAAAAGC ATCCAGCAAA
      CAAAATATTG GGTGTTTGTA CGTAAAAAAT GGAGTTTTCG TAGGTCGTTT
2901  AACTGTACAA CGCTTTTTCA AAAAAATGTA TTGTGATCCT CCTTAAGAAA
      TTGACATGTT GCGAAAAAGT TTTTTTACAT AACACTAGGA GGAATTCTTT
2951  AGCCTTACTT AGTGTTAATT CCTTTTTCTT TAGAATGCTG GTAAATACAA
      TCGGAATGAA TCACAATTAA GGAAAAAGAA ATCTTACGAC CATTTATGTT
3001  GGACTTAGGT AGGCTGGCTT CTAACAGCAA TTCACCCACT TATGATGGGA
      CCTGAATCCA TCCGACCGAA GATTGTCGTT AAGTGGGTGA ATACTACCCT
3051  TTAAAGGAAG GCACAACCAT GTCCACCACA GGTTCTAGCT CCCCCACCCA
      AATTTCCTTC CGTGTTGGTA CAGGTGGTGT CCAAGATCGA GGGGGTGGGT
3101  CACGCCCAGA GAGGGTTTTT CTGTGTAGCT CTGACTATTC TGGAATTCAC
      GTGCGGGTCT CTCCCAAAAA GACACATCGA GACTGATAAG ACCTTAAGTG
3151  ACTGCAGACC AGGCTGGTCT CGAACTCAGA GATCCACCAC CACATGGTTT
      TGACGTCTGG TCCGACCAGA GCTTGAGTCT CTAGGTGGTG GTGTACCAAA
3201  CTTAATTGTA ATTTTAAAGA AAAAAAAAAA TCCTTCAGTT AAGATTCTTA
      GAATTAACAT TAAAATTTCT TTTTTTTTTT AGGAAGTCAA TTCTAAGAAT
3251  TGTTCTAGGT TTTCACAAAC TTACCAATGT AGTTTTATTG GAGGCCATTT
      ACAAGATCCA AAAGTGTTTG AATGGTTACA TCAAAATAAC CTCCGGTAAA
3301  TTTAAATTTA ATCGGAGACT TGAAGAGCTA TTGCAAGAAA AAAAATGTAG
      AAATTTAAAT TAGCCTCTGA ACTTCTCGAT AACGTTCTTT TTTTTACATC
3351  GACAGTTAAA ATTTCATGAC ACACAAAAGG CAGCTACAAG TTTTGTGTGG
      CTGTCAATTT TAAAGTACTG TGTGTTTTCC GTCGATGTTC AAAACACACC
3401  ATTTCAACAT GTAAATTTCG GGTAAAAATG CAGGAAAACA GTTGAGTTCC
      TAAAGTTGTA CATTTAAAGC CCATTTTTAC GTCCTTTTGT CAACTCAAGG
3451  CGTGTTATTA GTATGTTACT AATAATTTCA GTATGTTAGT GAAAATAATC
      GCACAATAAT CATACAATGA TTATTAAAGT CATACAATCA CTTTTATTAG
3501  TTACTAAAAC ACTGGTACCT CAGACAACTT TACATGGTGA CGATTGTTAC
      AATGATTTTG TGACCATGGA GTCTGTTGAA ATGTACCACT CCTAACAATG
3551  TTTCCCAATC CATATAGAAT TTTAACAATT TTAGTGTTTA TTTTGGATGA
      AAAGGGTTAG GTATATCTTA AAATTGTTAA AATCACAAAT AAAACCTACT
3601  AAGGAAATGA CTATCTTTTG TTAGCAAATT ACCATAAGAT CTTTTTCTTT
      TTCCTTTACT GATAGAAAAC AATCGTTTAA TGGTATTCTA GAAAAAGAAA
3651  AGATTTCTGA ATACTCCAAG GAGCTCATAT AATTCCATCC TTATTTTTTC
      TCTAAAGACT TATGAGGTTC CTCGAGTATA TTAAGGTAGG AATAAAAAAG
3701  AGAGGCCCTC CCTGTTCAAT CACGGTATAA AAAAAGGAAC ACATTAAGAT
      TCTCCGGGAG GGACAAGTTA GTGCCATATT TTTTTCCTTG TGTAATTCTA
3751  GTCCCAGTCC TATTTTCTGG CTTTTTTTTT CCGGGGGTGG TGGTGCGGTA
      CAGGGTCAGG ATAAAAGACC GAAAAAAAAA GGCCCCCACC ACCACGCCAT
3801  ATCACTCTCT ATAGTCCAGT CTGGGCTTCA ACGCCTGGCA ATCCCCAGCC
      TAGTGAGAGA TATCAGGTCA GACCCGAAGT TGCGGACCGT TAGGGGTCGG
3851  TCAAGCTCCC AAGTACTGTC CTGATAAGGA TAGAAGGAGT CGACCTCCTT
      AGTTCGAGGG TTCATGACAG GACTATTCCT ATCTTCCTCA GCTGGAGGAA
3901  CACGCTCCCC TCCGAGGAGG GCTCCTTCCC AGCTCCATTC CCCGGTCGGG
      GTGCGAGGGG AGGCTCCTCC CGAGGAAGGG TCGAGGTAAG GGGCCAGCCC
3951  AGCCCGTCCC CCACCCGAGA GCGCGGGCCT CGTGGTCAGC GCCTCCGCGG
      TCGGGCAGGG GGTGGGCTCT CGCGCCCGGA GCACCAGTCG CGGAGGCGCC
4001  GGAGAAACAA AGGCGGCGGC GGGGGCTCAA GGGCACTGCG CCACGGGCCC
      CCTCTTTGTT TCCGCCGCCG CCCCCGAGTT CCCGTGACGC GGTGCCCGGG
4051  GCGCCTCCCC CATCCGGCGG CGGCCACGTA GCCGGGAGCG CGCCGCAGCC
      CGCGGAGGGG GTAGGCCGCC GCCGGTGCAT CGGCCCTCGC GCGGCGTCGG
4101  CGGAGCCTCG GGCCTCGCAG CTGCAGAGCC TGAACCGCTC TCTCCCTGCG
      GCCTCGGAGC CCGGAGCGTC GACGTCTCGG ACTTGGCGAG AGAGGGACGC
4151  GGCCTGCGAC GAGGCTGGGG GAGGGGAGGC CCGCGCTTTG TCTGGAGTCT
      CCGGACGCTG CTCCGACCCC CTCCCCTCCG GGCGCGAAAC AGACCTCAGA
4201  CGGTAGCTGT CATCCGGCTC CCACCCTCAT GCACAATTGT CCCATCTCCC
      GCCATCGACA GTAGGCCGAG GGTGGGAGTA CGTGTTAACA GGGTAGAGGG
```

FIG. 19(contd.)

```
4251  CCACGCACCG GCGCGGCGCC CGCCTCAGCG AGGCCCCAGC CGGTTTCCCG
      GGTGCGTGGC CGCGCCGCGG GCGGAGTCGC TCCGGGGTCG GCCAAAGGGC
4301  CAGCCCGCGG CCCACGGGGC TCGCAGCCTC CCCGCAAGCT CGGACGCACG
      GTCGGGCGCC GGGTGCCCCG AGCGTCGGAG GGGCGTTCGA GCCTGCGTGC
4351  GAGCATCCTA AACCCCACCA CACGCAAGAT CGAAAAAAAG CAAAGGCACG
      CTCGTAGGAT TTGGGGTGGT GTGCGTTCTA GCTTTTTTTC GTTTCCGTGC
4401  AACTTCACCG CTCCGATGCT CAGGGCCGCG GATCCTGCAG AGTCTCCCGC
      TTGAAGTGGC GAGGCTACGA GTCCCGGCGC CTAGGACGTC TCAGAGGGCG
4451  CTGCGCGCTT CGGTTCAGCC ACATCCGAGG GGAGGGGGCG CGGGCAGCTC
      GACGCGCGAA GCCAAGTCGG TGTAGGCTCC CCTCCCCCGC GCCCGTCGAG
4501  CGCCGGGGGG GAGGGGGAGC ACCGCCCACG CCCTGGCCGC GCGGGGCCCG
      GCGGCCCCCC CTCCCCCTCG TGGCGGGTGC GGGACCGGCG CGCCCCGGGC
4551  CCGGGAACGC GTCCTGCGGG GGGCGGCGCG CGCAATGCTC ACCGTCCGCG
      GGCCCTTGCG CAGGACGCCC CCCGCCGCGC GCGTTACGAG TGGCAGGCGC
4601  GCGTGGCGCC CAGGGGGTCT CCTGGCTGGG GGGAGGGGGG GGAAGGCGGG
      CGCACCGCGG GTCCCCCAGA GGACCGACCC CCCTCCCCCC CCTTCCGCCC
4651  CAGGAAGGAC CGCGGAGGCC TCTCTGCGTC TCGGAGCGCG CCAAAGCGGG
      GTCCTTCCTG GCGCCTCCGG AGAGACGCAG AGCCTCGCGC GGTTTCGCCC
4701  GCTCCACCCA CCTCCTTGCC CGGATCTTGA AGGCCGGGGA GATAAACAGC
      CGAGGTGGGT GGAGGAACGG GCCTAGAACT TCCGGCCCCT CTATTTGTCG
4751  GGGGTTCTTT AAGCACCACC TCTCACTAGG CGCGGGATCC CAAGGCTTGT
      CCCCAAGAAA TTCGTGGTGG AGAGTGATCC GCGCCCTAGG GTTCCGAACA
4801  GGCATCCGGG GTGGTACTTG GACTAAAAGT CCTTCTGGGA GGGACCGAGT
      CCGTAGGCCC CACCATGAAC CTGATTTTCA GGAAGACCCT CCCTGGCTCA
4851  GAGAACCCCT TTGGGACGTG TAGAAATATT TGTGTGGTTC GAGAATATTT
      CTCTTGGGGA AACCCTGCAC ATCTTTATAA ACACACCAAG CTCTTATAAA
4901  GTGCGGACGG GCTTGGCAAA GGCGTAGCTG CAGAGAGCAC GCTTGGGTGG
      CACGCCTGCC CGAACCGTTT CCGCATCGAC GTCTCTCGTG CGAACCCACC
4951  AGAGGGCCGC ACGCCCCAGC GCCGGCCTAA GCCCCTCCCG ACGGCGTTAT
      TCTCCCGGCG TGCGGGGTCG CGGCCGGATT CGGGGAGGGC TGCCGCAATA
5001  TTCAAACTGC GCGACCGTTT CTCCGCTCCC TACGCGGAGG TGGGGGCCGG
      AAGTTTGACG CGCTGGCAAA GAGGCGAGGG ATGCGCCTCC ACCCCCGGCC
5051  ACCTAGTTCC GGACGTAGTA ACACGCCGAG CGCGAGCCTT CCGCAATTCA
      TGGATCAAGG CCTGCATCAT TGTGCGGCTC GCGCTCGGAA GGCGTTAAGT
5101  CGGAACACAG TTGCGCAAGT GATGTAAAGC AGTCCCGCTG TACCTAAAGG
      GCCTTGTGTC AACGCGTTCA CTACATTTCG TCAGGGCGAC ATGGATTTCC
5151  GGGAGTGTCA CGTACTTGGC GTAAGGAGAG TGTAGGCCCT TCCCGCCATT
      CCCTCACAGT GCATGAACCG CATTCCTCTC ACATCCGGGA AGGGCGGTAA
5201  GGCGGCGGTT AGGGCGTTTA CGTAACGGCG TGACGTAAGC GGAGACGCGT
      CCGCCGCCAA TCCCGCAAAT GCATTGCCGC ACTGCATTCG CCTCTGCGCA
5251  TAGTGGGGGG AAGGTTCTAG AAAAGCGGCG GTCTCGGCTC CAGCGGCAGT
      ATCACCCCCC TTCCAAGATC TTTTCGCCGC CAGAGCCGAG GTCGCCGTCA
5301  AGCAGCGGCG CCGGTCCCGT GTGCAGGAGC TCCTTTGCGG CCCAGTTTCT
      TCGTCGCCGC GGCCAGGGCA CACGTCCTCG AGGAAACGCC GGGTCAAAGA
5351  TGGCCATCGC CTGCTCTCCC CACAGCGCCA GGACGAGTCC CGTGCGCGTC
      ACCGGTAGCG GACGAGAGGG GTGTCGCGGT CCTGCTCAGG GCACGCGCAG
5401  CGTCCGCGGA GGTCTTTCTC ATCTCGCTCG GCTGCGGGAA ATCGGGCTGA
      GCAGGCGCCT CCAGAAAGAG TAGAGCGAGC CGACGCCCTT TAGCCCGACT
5451  AGCGACTGAG TCCGCGATGG AGGTAACGGG TTTGAAATCA ATGAGTTATT
      TCGCTGACTC AGGCGCTACC TCCATTGCCC AAACTTTAGT TACTCAATAA
5501  AAAAATGGCA TGGCGAGGCC GTAGGCACCG CAATGGAAAC CGGCCACCCG
      TTTTTACCGT ACCGCTCCGG CATCCGTGGC GTTACCTTTG GCCGGTGGGC
5551  CCTCCGTGGT CCGGCGGAGG GGATGCGGCC ACTCGAGTGG CGGTTGGCCT
      GGAGGCACCA GGCCGCCTCC CCTACGCCGG TGAGCTCACC GCCAACCGGA
5601  TGGCGAGTTT CTGAGGGGTC GTTGGAGGAG GCCTCTGATT GTCCGACCGC
      ACCGCTCAAA GACTCCCCAG CAACCTCCTC CGGAGACTAA CAGGCTGGCG
5651  CTTCCCCGCC CTCAGCCGCC CGGCGCCATT TCCCTCAGTT GGGGTGGGGG
      GAAGGGGCGG GAGTCGGCGG GCCGCGGTAA AGGGAGTCAA CCCCACCCCC
```

FIG. 19(contd.)

```
5701  ATGGGAAGTG CCCGCCGCGA CCGGGCTGGA CCGCTAAAGT AGCGCGTGAG
      TACCCTTCAC GGGCGGCGCT GGCCCGACCT GGCGATTTCA TCGCGCACTC
5751  CGGGCCATCG CTGGCCTTTC GATGTGCGCG GGCCTAGGGG CTCGGTTGTG
      GCCCGGTAGC GACCGGAAAG CTACACGCGC CCGGATCCCC GAGCCAACAC
5801  TTCGCGGCGG AACGTTTCTG GGGCCCCCCC GGCTTCCCGG AGCGAGTCTG
      AAGCGCCGCC TTGCAAAGAC CCCGGGGGGG CCGAAGGGCC TCGCTCAGAC
5851  CGAAGCTAGC TTCCCCTCCC CCCTCTCCCG GGAACCGGAT TTGGCGGCCG
      GCTTCGATCG AAGGGGAGGG GGGAGAGGGC CCTTGGCCTA AACCGCCGGC
5901  CCATTTTCCC GTCTCCTTCC TCGCCACGAT TTTGCTTTCA ACGCTTTAGG
      GGTAAAAGGG CAGAGGAAGG AGCGGTGCTA AAACGAAAGT TGCGAAATCC
5951  TTTACTAGTT TGGTTTTCTT TTTTCACCAC TGCGTAGACG TGTTTAGCGA
      AAATGATCAA ACCAAAAGAA AAAAGTGGTG ACGCATCTGC ACAAATCGCT
6001  TTTTCCTTTC TTTTGGAAGT CTTCATACCG TTTCGAGGTG GATTTAGCGT
      AAAAGGAAAG AAAACCTTCA GAAGTATGGC AAAGCTCCAC CTAAATCGCA
6051  TTTGAGCTTG GGTCTTCAGC GTCCTGCGCA CCTCGCTAAA GGCTCTCTGC
      AAACTCGAAC CCAGAAGTCG CAGGACGCGT GGAGCGATTT CCGAGAGACG
6101  CTTCCCCTCG ACGAAATGGC GCCATTGCTT TCTGAAGCCA CCGAGGCGCG
      GAAGGGGAGC TGCTTTACCG CGGTAACGAA AGACTTCGGT GGCTCCGCGC
6151  GGGTGGGGGC GGGGTGGCGG CGCTCCACGA GCTTTACTGG AACAGGCAGA
      CCCACCCCCG CCCCACCGCC GCGAGGTGCT CGAAATGACC TTGTCCGTCT
6201  GAGAACGTAG TACAACCGAG GCCTGGGCGG GTGGCTGAAG GCAGCGTCGC
      CTCTTGCATC ATGTTGGCTC CGGACCCGCC CACCGACTTC CGTCGCAGCG
6251  TGCAAAGAGA CCGTTTTATT TTTCATAATA CGTAAGATTA CGGGTGCTGT
      ACGTTTCTCT GGCAAAATAA AAAGTATTAT GCATTCTAAT GCCCACGACA
6301  AGTAAAGCAC TTGAGCATTA GTATAGTAGG AGGAAGTCAA AGTGGAAAAA
      TCATTTCGTG AACTCGTAAT CATATCATCC TCCTTCAGTT TCACCTTTTT
6351  ATGGGAGCGC TCATCAGGAA GCTAGGGAGG CTATGTTGAG TGCAGGGTTA
      TACCCTCGCG AGTAGTCCTT CGATCCCTCC GATACAACTC ACGTCCCAAT
6401  CTTTCCTTTT ATTGCAGAAC TTTTATCTGC TTAAAGGATC CTCGGATCGA
      GAAAGGAAAA TAACGTCTTG AAAATAGACG AATTTCCTAG GAGCCTAGCT
6451  AATAATTCAA ATTATAAGCA TTTTTAAGGG AATCTTCGAA TTTGTTGGTA
      TTATTAAGTT TAATATTCGT AAAAATTCCC TTAGAAGCTT AAACAACCAT
6501  AAGTCAACGG ATCCTTAGCA CGTGGTGTTC ACTTTAAGGA AGTGAAATAG
      TTCAGTTGCC TAGGAATCGT GCACCACAAG TGAAATTCCT TCACTTTATC
6551  CTGACTTTTC ATAGTTAGCC TTCGCTTAAA GCCTGGTTCA GTGGACGAAA
      GACTGAAAAG TATCAATCGG AAGCGAATTT CGGACCAAGT CACCTGCTTT
6601  ATCCACGTCC TGGCTATATA AAAACTTAGT TTGGGGTCAC AGTGTTTGAG
      TAGGTGCAGG ACCGATATAT TTTTGAATCA AACCCCAGTG TCACAAACTC
6651  CGTGGTCATT CGGTTTTTTT ATTTTTTATT TGTTTGAAAT TATGATGCAT
      GCACCAGTAA GCCAAAAAAA TAAAAAATAA ACAAACTTTA ATACTACGTA
6701  CATTACACTG ATAAGCATTA GCTTTCGAAT TGAAAGGGGT CTCCTTGGTT
      GTAATGTGAC TATTCGTAAT CGAAAGCTTA ACTTTCCCCA GAGGAACCAA
6751  ATTTTCTTTG ACTCTAAGCA CACTTATAAA TAAAATAACC TTGTTTATAA
      TAAAAGAAAC TGAGATTCGT GTGAATATTT ATTTTATTGG AACAAATATT
6801  TCGATAGTGG ACGTCTGGTA AGTTTGGAAA AAACCCGAGG TAAGTAAAGA
      AGCTATCACC TGCAGACCAT TCAAACCTTT TTTGGGCTCC ATTCATTTCT
6851  GCTTTTGCTT TCGTTAGTGA TATGAAAAAA CAAGGTGTAT TTAATACTTG
      CGAAAACGAA AGCAATCACT ATACTTTTTT GTTCCACATA AATTATGAAC
6901  CAACTTAGTT TAAGGAAAGC CAATTTACTG ACATTTTAGT AGAGCTACCA
      GTTGAATCAA ATTCCTTTCG GTTAAATGAC TGTAAAATCA TCTCGATGGT
6951  GAAACACTAT TTGGAGTCCT GATTAAGGCT TTTGTAACTA TTTTGACTAT
      CTTTGTGATA AACCTCAGGA CTAATTCCGA AAACATTGAT AAAACTGATA
7001  TTAAAACAAT TTTGGTCGTT TTTATTAAAC ATTTCAAAAC CTAAAAATTG
      AATTTTGTTA AAACCAGCAA AAATAATTTG TAAAGTTTTG GATTTTTAAC
7051  TAAACATTGG CTTTTTGAGC ACATTTTGGA GAAACTTACA AATTTAGGCT
      ATTTGTAACC GAAAAACTCG TGTAAAACCT CTTTGAATGT TTAAATCCGA
7101  ATACAGTAAA ATAACGGATT TGTTTTATAA TTTTGCTTTT TCATTTCGTT
      TATGTCATTT TATTGCCTAA ACAAAATATT AAAACGAAAA AGTAAAGCAA
```

```
7151    GTGCAGTCAT AGGTCCTGGA TAGTATGACC TAATTTATGA ACATCTTGAT
        CACGTCAGTA TCCAGGACCT ATCATACTGG ATTAAATACT TGTAGAACTA
7201    AAGTTTTTGT ACTTAGCTAT TGGAAAGCCA GTATTAAGTG CCTGACAAAA
        TTCAAAAACA TGAATCGATA ACCTTTCGGT CATAATTCAC GGACTGTTTT
7251    CCAGATTTAA GGTGATATCT GGAGTTTCAG CATTCTTCAT GGAGCTTGTT
        GGTCTAAATT CCACTATAGA CCTCAAAGTC GTAAGAAGTA CCTCGAACAA
7301    TCAGAGTTGC AGGATTTTTT TTTTTCATCT TGAGATACTT ACAATTAACA
        AGTCTCAACG TCCTAAAAAA AAAAAGTAGA ACTCTATGAA TGTTAATTGT
7351    CCAGAGGGGG CAGCTCAGGG AAAAGCAAAT ATGCCACTTT TCAGAAACTG
        GGTCTCCCCC GTCGAGTCCC TTTTCGTTTA TACGGTGAAA AGTCTTTGAC
7401    AATCTTGGAA GTGGTGAATT TGGAAACAGG TTTTTTAAAT TTTTTTTAAA
        TTAGAACCTT CACCACTTAA ACCTTTGTCC AAAAAATTTA AAAAAAATTT
7451    TCTAAAAAGT AGTAAATTTT GGACTTGGGT TGTAGAATTT AATGAATTAC
        AGATTTTTCA TCATTTAAAA CCTGAACCCA ACATCTTAAA TTACTTAATG
7501    AAAAGAATTC TTTAATACCC TTTAAATGAC CTAAGAGCTG GGTATGGTTT
        TTTTCTTAAG AAATTATGGG AAATTTACTG GATTCTCGAC CCATACCAAA
7551    TTCTGAATTT TTTTGAAGAA AATCTAAGAA AGTTTACGTG AATTAGAAGT
        AAGACTTAAA AAAACTTCTT TTAGATTCTT TCAAATGCAC TTAATCTTCA
7601    TAGATCGAAT ATTAGTGACT TTGAAACTTG TATAGCTCAG GCAATTTTTG
        ATCTAGCTTA TAATCACTGA AACTTTGAAC ATATCGAGTC CGTTAAAAAC
7651    GTGTAACACA ACTAATATGC AGTTTAACAT ATGGTTTAAA TTTGATGTAA
        CACATTGTGT TGATTATACG TCAAATTGTA TACCAAATTT AAACTACATT
7701    GTTTTTTTTC TCCCCCCCAG AAAACTTTAG AAACTGTTCC TTTGGAGAGG
        CAAAAAAAAG AGGGGGGGTC TTTTGAAATC TTTGACAAGG AAACCTCTCC
7751    AAAAAGGTAC TCTACCAGCA GGTCACCTCA TATTTAAGAA TTTAATTTCC
        TTTTTCCATG AGATGGTCGT CCAGTGGAGT ATAAATTCTT AAATTAAAGG
7801    TGCATACAAA GAAAGTGTAA ATAAAAATTG AAATGGTATT TCCCTTTGCA
        ACGTATGTTT CTTTCACATT TATTTTTAAC TTTACCATAA AGGGAAACGT
7851    GAGAGAAAAG GAACAGTTCC GAAAGCTCTT TATTGGTGGC TTAAGCTT
        CTCTCTTTTC CTTGTCAAGG CTTTCGAGAA ATAACCACCG AATTCGAA
```

FIG. 19(contd.)

FIG. 20

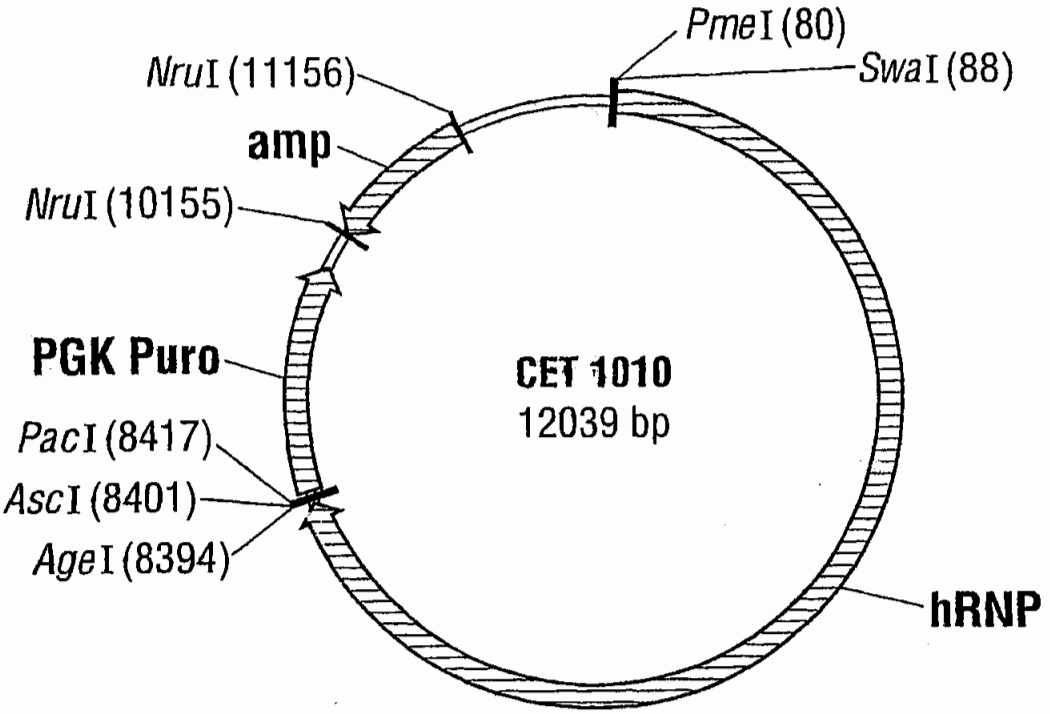

PmeI (80)
SwaI (88)
NruI (11156)
amp
NruI (10155)
PGK Puro
PacI (8417)
AscI (8401)
AgeI (8394)
CET 1010
12039 bp
hRNP

FIG. 22

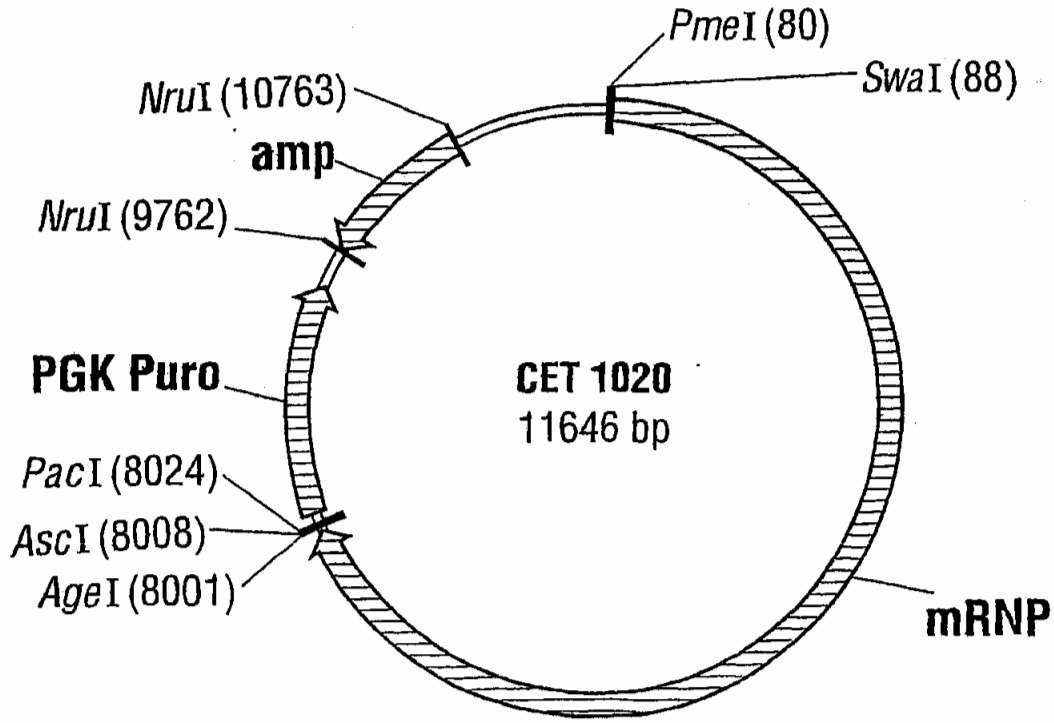

PmeI (80)
SwaI (88)
NruI (10763)
amp
NruI (9762)
PGK Puro
PacI (8024)
AscI (8008)
AgeI (8001)
CET 1020
11646 bp
mRNP

FIG. 21 CET1010 nucleotide sequence

```
   1  CGTTGTAAAA CGACGGCCAG TGAATTGTAA TACGACTCAC TATAGGGCGA
      GCAACATTTT GCTGCCGGTC ACTTAACATT ATGCTGAGTG ATATCCCGCT
  51  ATTGGGTACC GGGCCCCCCC TCGAAGTTTA AACATTTAAA TCTAGAAGCT
      TAACCCATGG CCCGGGGGGG AGCTTCAAAT TTGTAAATTT AGATCTTCGA
 101  TCAATGTTTT TAGCACCCTC TGTGTGGAGG AAAATAATGC AGATTATTCT
      AGTTACAAAA ATCGTGGGAG ACACACCTCC TTTTATTACG TCTAATAAGA
 151  AATTAGTGTA ATATCTAACC ACATTAAAAT ATATTACATA GTAAACTACA
      TTAATCACAT TATAGATTGG TGTAATTTTA TATAATGTAT CATTTGATGT
 201  CTCCATAATT TTATAAATTT GACTCCCCAG GGTAATAAAC TAGTCTCTAG
      GAGGTATTAA AATATTTAAA CTGAGGGGTC CCATTATTTG ATCAGAGATC
 251  TCTGCTCACC TTCAACTGTA CAATAAAGTC TTGGTTCTTT TGAAATAGAC
      AGACGAGTGG AAGTTGACAT GTTATTTCAG AACCAAGAAA ACTTTATCTG
 301  CTCAAATGAG ACACCTAAAA TTCAAAGTGT CTTTACATTT AAAGACACCT
      GAGTTTACTC TGTGGATTTT AAGTTTCACA GAAATGTAAA TTTCTGTGGA
 351  ACAGGAAAGC AGGTAAAAGA GCCAGGTTAA AAACAAATTC TAAAACCACT
      TGTCCTTTCG TCCATTTTCT CGGTCCAATT TTTGTTTAAG ATTTTGGTGA
 401  TAGCTGCAGT TAAACATATA GTAAAGATGC ACTAAAGTTT CTTACTCTGT
      ATCGACGTCA ATTTGTATAT CATTTCTACG TGATTTCAAA GAATGAGACA
 451  AAATCCCTTC CACTTCAGGA AATATTCCAC TTTCCCATTC ACTACACGTC
      TTTAGGGAAG GTGAAGTCCT TTATAAGGTG AAAGGGTAAG TGATGTGCAG
 501  GATCTAGTAC TTTTTCCACG ACAAATTCTT CAGGCTCTGC CTCTTCAACT
      CTAGATCATG AAAAAGGTGC TGTTTAAGAA GTCCGAGACG GAGAAGTTGA
 551  TTTTTACTCT TTCCATTCTG TTTTTTTCCC ATTTTTTGCT AAAATAAAAC
      AAAAATGAGA AAGGTAAGAC AAAAAAAGGG TAAAAAACGA TTTTATTTTG
 601  AAAAGAGAAA TTAAGAAATA TTCCTCTTGA ATTTTGAGCA CATTTTCAAG
      TTTTCTCTTT AATTCTTTAT AAGGAGAACT TAAAACTCGT GTAAAAGTTC
 651  GCTCAATTGC TTATATTATT ATCACATTCG ACATAAATTT TTACTTCTAT
      CGAGTTAACG AATATAATAA TAGTGTAAGC TGTATTTAAA AATGAAGATA
 701  ATCCCAGGGC AGACACCTTC TGGAAAGATT AAAAGTCAAC AGACAATAAA
      TAGGGTCCCG TCTGTGGAAG ACCTTTCTAA TTTTCAGTTG TCTGTTATTT
 751  ATAAAAGAAT GCTTTATCTT GTTCATTTAG TTCAAACTTA CAACCCACCA
      TATTTTCTTA CGAAATAGAA CAAGTAAATC AAGTTTGAAT GTTGGGTGGT
 801  CCAAAATAAT ACAATAAAAA AACACTATCT GGAAACAGTT ATTTTTTTCC
      GGTTTTATTA TGTTATTTTT TTGTGATAGA CCTTTGTCAA TAAAAAAAGG
 851  AGTCTTTTTT TTTGAGACAG GGTCTCACAC TCTTGTCGCC CAGGCTGGAG
      TCAGAAAAAA AAACTCTGTC CCAGAGTGTG AGAACAGCGG GTCCGACCTC
 901  TGCAGTGGCG TGATCTCAGC TCACTGCAAC CTCCGCCTCC CCAGGTTCAA
      ACGTCACCGC ACTAGAGTCG AGTGACGTTG GAGGCGGAGG GGTCCAAGTT
 951  GCAGTTCTCA TGCCTCAGCC TCCAGAGTAG CTGGGATTAT AGGCGGATGC
      CGTCAAGAGT ACGGAGTCGG AGGTCTCATC GACCCTAATA TCCGCCTACG
1001  CACCATGCCG GGCTAATTTT TTTTGTGTTT TTATTAGAAA CAGGGTTTCA
      GTGGTACGGC CCGATTAAAA AAAACACAAA AATAATCTTT GTCCCAAAGT
1051  CCATGTTGAC CAGGCTGGTC TCAAACTCCT GACCTGAAGT GATTCACCAG
      GGTACAACTG GTCCGACCAG AGTTTGAGGA CTGGACTTCA CTAAGTGGTC
1101  CCTGGGCCTC CCAAAGTGCT GGCATTACAG GCGTGAGCCA CTGCGCCCGG
      GGACCCGGAG GGTTTCACGA CCGTAATGTC CGCACTCGGT GACGCGGGCC
1151  CCCTGTAGTC TTAAAAGACC AAGTTTACTA ATTTTCACTC ATTTTAACAA
      GGGACATCAG AATTTTCTGG TTCAAATGAT TAAAAGTGAG TAAAATTGTT
1201  CACTGCAACA AACAACTATG CAGGAAGTAC CTAAAGGGTG ATCCAGAGAA
      GTGACGTTGT TTGTTGATAC GTCCTTCATG GATTTCCCAC TAGGTCTCTT
1251  GCAAGTAGTA GTGACAGGTC TTAGGTGAAC CTATGACAGA CCTTGTATCC
      CGTTCATCAT CACTGTCCAG AATCCACTTG GATACTGTCT GGAACATAGG
1301  ACCCCCAGAT GGTAAAAGCC CCAGCCCCCT TCTCAATTCA AATATTAATG
      TGGGGGTCTA CCATTTTCGG GGTCGGGGGA AGAGTTAAGT TTATAATTAC
```

FIG. 21(contd.) 42/93

```
1351  TCAAAAGCAT CAATGATACA GAGAAAAGAT AAATGCAGAA TGAAAACATG
      AGTTTTCGTA GTTACTATGT CTCTTTTCTA TTTACGTCTT ACTTTTGTAC
1401  GTTCAAAATC CTGATACCAA CTGCAGGGTC AACTATAGAG ACCACTAGGA
      CAAGTTTTAG GACTATGGTT GACGTCCCAG TTGATATCTC TGGTGATCCT
1451  GGTTCAATTA AAGGACAAGA TTATTTTTCC ATAATCTCTG TAGATAATAT
      CCAAGTTAAT TTCCTGTTCT AATAAAAAGG TATTAGAGAC ATCTATTATA
1501  TTCCTACCAC TTAGAACAAA ACTATAAAGC TATCACTTCA AGAGACCAAC
      AAGGATGGTG AATCTTGTTT TGATATTTCG ATAGTGAAGT TCTCTGGTTG
1551  ATTACAAATT TATTTTAATT CCCTAAGGTG AAAAAAATCC TTCCTTCCTG
      TAATGTTTAA ATAAAATTAA GGGATTCCAC TTTTTTTAGG AAGGAAGGAC
1601  GTTTCTCAAG AGAAAGTCTA TACTGGTAAC CAAATTCACT TTAAACAGGC
      CAAAGAGTTC TCTTTCAGAT ATGACCATTG GTTTAAGTGA AATTTGTCCG
1651  ATTTTCTTTG GTATGACACT ATTTAAGAGA AGCAGGAAAC CAACGTGAAC
      TAAAAGAAAC CATACTGTGA TAAATTCTCT TCGTCCTTTG GTTGCACTTG
1701  CAGCTCTTTC CAATGGCTCA AGATTTCCTA TGAGAGGACT AAAAATGGGG
      GTCGAGAAAG GTTACCGAGT TCTAAAGGAT ACTCTCCTGA TTTTTACCCC
1751  AAAATTTTTA TGAGAGGATT AAAAATGGGG GAAAAAAAAC CCTGAAATGG
      TTTTAAAAAT ACTCTCCTAA TTTTTACCCC CTTTTTTTTG GGACTTTACC
1801  TTAATCAGAA GATCCTATGG GCTGAGAAGG AATCCATCTT AACATTTCAT
      AATTAGTCTT CTAGGATACC CGACTCTTCC TTAGGTAGAA TTGTAAAGTA
1851  CTTAAAGCAA ATGCTATTGC CGGGGGCAGT GGCTCATGCC TGTAATCCCA
      GAATTTCGTT TACGATAACG GCCCCCGTCA CCGAGTACGG ACATTAGGGT
1901  GCACTTTGGG AGGCCGAGGT GGGCAGATCA TCTGAGGTCA GGAGTTTGAG
      CGTGAAACCC TCCGGCTCCA CCCGTCTAGT AGACTCCAGT CCTCAAACTC
1951  ACCAGCCTGA CCAACATGGA GAAACCCCGT TTCTACTAAA AATACAAAAT
      TGGTCGGACT GGTTGTACCT CTTTGGGGCA AAGATGATTT TTATGTTTTA
2001  TAGCCAGGCA TAGTGGTGCA TGCCTGTAAT CCCAGCTACT TGGGAGGCTG
      ATCGGTCCGT ATCACCACGT ACGGACATTA GGGTCGATGA ACCCTCCGAC
2051  AGGCAGGAGA ACTGCTTGAA CCCAGGAGGC TTAAGTTGCG GTGAGCCAAG
      TCCGTCCTCT TGACGAACTT GGGTCCTCCG AATTCAACGC CACTCGGTTC
2101  ATCACGCCAT TGCACTCTAG CCTGGACAAC AAGAGAAAAA CTCTGTCTCA
      TAGTGCGGTA ACGTGAGATC GGACCTGTTG TTCTCTTTTT GAGACAGAGT
2151  AAAAAACACA AAAACAAAAA ACCCAAATAC TATTTAAAAA AGATAAACCT
      TTTTTTGTGT TTTTGTTTTT TGGGTTTATG ATAAATTTTT TCTATTTGGA
2201  TAATTGCTCA ATCATTAAAG CCATCCCACA AGTAAAGCAG CAAGCAGAAA
      ATTAACGAGT TAGTAATTTC GGTAGGGTGT TCATTTCGTC GTTCGTCTTT
2251  AAAGTTAAGA ACACCTCAAG GCTACAGAAG GACATTTCAA GCTATGCAGG
      TTTCAATTCT TGTGGAGTTC CGATGTCTTC CTGTAAAGTT CGATACGTCC
2301  CATATGAAGT GTGCAGACAG ATATGTAAGA AAGGCCTCAA GACTGCAAAA
      GTATACTTCA CACGTCTGTC TATACATTCT TTCCGGAGTT CTGACGTTTT
2351  GGGCATTTCA AGCTATGCAA GCATATAGGT AACACATACA CACACACAAA
      CCCGTAAAGT TCGATACGTT CGTATATCCA TTGTGTATGT GTGTGTGTTT
2401  ATAAAATCCC CTGAAATACA AAAACATGCA GCAAACACCT GACGTTTTTG
      TATTTTAGGG GACTTTATGT TTTTGTACGT CGTTTGTGGA CTGCAAAAAC
2451  GATACCATTT CTAAGTCAGG TGTTATGATT CTCATTAGTC AAGATACTTG
      CTATGGTAAA GATTCAGTCC ACAATACTAA GAGTAATCAG TTCTATGAAC
2501  AGTACTGGGC CCAAACAGCT TTCTGCCACT GTACAGTACA AGAAGGTAGG
      TCATGACCCG GGTTTGTCGA AAGACGGTGA CATGTCATGT TCTTCCATCC
2551  AATAATGGTG GGAGGAGCAA AGACAAACTG TAATAGACAG AAGTGTATCA
      TTATTACCAC CCTCCTCGTT TCTGTTTGAC ATTATCTGTC TTCACATAGT
2601  GATACCTATA CTACATGAAA AACAAAACAG CTACTGCCAC AAAGGGAGAA
      CTATGGATAT GATGTACTTT TTGTTTTGTC GATGACGGTG TTTCCCTCTT
2651  GGCTAACAAA ATAAAGTCAA CAATAAATAC AGAAAATGAA AAGGATACAC
      CCGATTGTTT TATTTCAGTT GTTATTTATG TCTTTTACTT TTCCTATGTG
2701  ACTAAGGTTT ACAAAAAAAA AAAGGCAGAC AAAATGCCAT ACAGTATTCA
      TGATTCCAAA TGTTTTTTTT TTTCCGTCTG TTTTACGGTA TGTCATAAGT
2751  TTCACTACTA TGGCATTCAT AAGCTAGTTT CAAATGCTCA CTATTTTCTT
      AAGTGATGAT ACCGTAAGTA TTCGATCAAA GTTTACGAGT GATAAAAGAA
```

FIG. 21 (contd.)


```
2801  TTATAGTATA TATTTGCCTT AACCCAGCAC TTTTTTCCAA AAGTGGATGA
      AATATCATAT ATAAACGGAA TTGGGTCGTG AAAAAAGGTT TTCACCTACT
2851  GTCAAAATAA ATTTCCCATT ATTTAAGTGA AATTAACAGC ACACATATCT
      CAGTTTTATT TAAAGGGTAA TAAATTCACT TTAATTGTCG TGTGTATAGA
2901  CACAACACTA ATGAATTTTT AAAATGGAAA GTTAAGAACT TTTAAAGTGG
      GTGTTGTGAT TACTTAAAAA TTTTACCTTT CAATTCTTGA AAATTTCACC
2951  CCAACCTGTG ATCCTTCACA AAATAAACTA AATACAATAA CAGACCCCAA
      GGTTGGACAC TAGGAAGTGT TTTATTTGAT TTATGTTATT GTCTGGGGTT
3001  AGGCTATCAA TTGCGTGCAA AAACAACTTC TGTTTTCCAG GGTAAACAGA
      TCCGATAGTT AACGCACGTT TTTGTTGAAG ACAAAAGGTC CCATTTGTCT
3051  ATCTAATGCA GAATCTAATG CAGGGTAAAC AGACTTAATG CAGAATCTAA
      TAGATTACGT CTTAGATTAC GTCCCATTTG TCTGAATTAC GTCTTAGATT
3101  TGATGGCACA AATTAAAAAT CACTAACGTG CCCTTTTTAG TGTGAAACCC
      ACTACCGTGT TTAATTTTTA GTGATTGCAC GGGAAAAATC ACACTTTGGG
3151  AGAGAGAGCA CATACAAGCC AAAAACAAAT GCTTTATTTT ACCTAGGAGA
      TCTCTCTCGT GTATGTTCGG TTTTTGTTTA CGAAATAAAA TGGATCCTCT
3201  CATTAACATT CACCTTTACG TGTTTAAGAT TAATGCAATG TTAAATATTG
      GTAATTGTAA GTGGAAATGC ACAAATTCTA ATTACGTTAC AATTTATAAC
3251  TGAAAACTGT AACTTTGAAT TTCATGATTT TTATGTGAAT ATTCCAGGGT
      ACTTTTGACA TTGAAACTTA AAGTACTAAA AATACACTTA TAAGGTCCCA
3301  TTAAAAAAAC TTGTAACATG ACATGGCTGA ATAAGATAAA AAAAAAATCT
      AATTTTTTTG AACATTGTAC TGTACCGACT TATTCTATTT TTTTTTAGA
3351  AGCCTTTTCT CCCTTCTGGC TCATATTTGC GATTTCGATC ATTTTGTTTA
      TCGGAAAAGA GGGAAGACCG AGTATAAACG CTAAAGCTAG TAAAACAAAT
3401  AAAAACAAAA CACTGCAATG AATTAAACTT AATATTCTTC TATGTTTTAG
      TTTTTGTTTT GTGACGTTAC TTAATTTGAA TTATAAGAAG ATACAAAATC
3451  AGTAAGTTAA AACAAGATAA AGTGACCAAA GTAATTTGAA AGATTCAATG
      TCATTCAATT TTGTTCTATT TCACTGGTTT CATTAAACTT TCTAAGTTAC
3501  ACTTTTGCTC CAACCTAGGT GCACAAGGTA CCTTGTTCTT TAAATTGGGC
      TGAAAACGAG GTTGGATCCA CGTGTTCCAT GGAACAAGAA ATTTAACCCG
3551  TTTAATGAAA ATACTTCTCC AGAATTCTGG GGATTTAAGA AAAATTATGC
      AAATTACTTT TATGAAGAGG TCTTAAGACC CCTAAATTCT TTTTAATACG
3601  CAACCAACAA GGGCTTTACC ATTTTATGTA ACATTTTTCA ACGCTGCAAA
      GTTGGTTGTT CCCGAAATGG TAAAATACAT TGTAAAAAGT TGCGACGTTT
3651  AATGTGTGTA TTTCTATTTG AAGATAAAAA TCCTCAGCAA AATCCACATT
      TTACACACAT AAAGATAAAC TTCTATTTTT AGGAGTCGTT TTAGGTGTAA
3701  GCACTGTCCT TCAAAGATTA GCCTTCTTTG AACTAGTTAA GACACTATTA
      CGTGACAGGA AGTTTCTAAT CGGAAGAAAC TTGATCAATT CTGTGATAAT
3751  AGCCAAGCCA GTATCTCCCT GTAATGAATT CGTTTTTCTC TTAATTTTCC
      TCGGTTCGGT CATAGAGGGA CATTACTTAA GCAAAAAGAG AATTAAAAGG
3801  CCTGTAATTT ACACTGGGAG AGCTGGGAAA TATGTGGATG TAAATTTCTC
      GGACATTAAA TGTGACCCTC TCGACCCTTT ATACACCTAC ATTTAAAGAG
3851  AGCCACAGAG ATGCAAAGTT ATACTGTGGG GAAAAAAAAC TTGAGTTAAA
      TCGGTGTCTC TACGTTTCAA TATGACACCC CTTTTTTTTG AACTCAATTT
3901  TCCTTACATA TTTTAGGTTT TCATTAACTT ACCAATGTAG TTTTGTTGGA
      AGGAATGTAT AAAATCCAAA AGTAATTGAA TGGTTACATC AAAACAACCT
3951  GGCCATTTTT TTTATTGCAG ACTTGAAGAG CTATTACTAG AAAAATGCAT
      CCGGTAAAAA AAATAACGTC TGAACTTCTC GATAATGATC TTTTTACGTA
4001  GACAGTTAAG GTAAGTTTGC ATGACACAAA AAAGGTAACT AAATACAAAT
      CTGTCAATTC CATTCAAACG TACTGTGTTT TTTCCATTGA TTTATGTTTA
4051  TCTGTTTGGA TTCCAACCCC CAAGTAGAGA GCGCACACTT TCAAACGTGA
      AGACAAACCT AAGGTTGGGG GTTCATCTCT CGCGTGTGAA AGTTTGCACT
4101  ATACAAATCC AGAGTAGATC TGCGCTCCTA CCTACATTGC TTATGATGTA
      TATGTTTAGG TCTCATCTAG ACGCGAGGAT GGATGTAACG AATACTACAT
4151  CTTAAGTACG TGTCCTAACC ATGTGAGTCT AGAAAGACTT TACTGGGGAT
      GAATTCATGC ACAGGATTGG TACACTCAGA TCTTTCTGAA ATGACCCCTA
4201  CCTGGTACCT AAAACAGCTT CACATGGCTT AAAATAGGGG ACCAATGTCT
      GGACCATGGA TTTTGTCGAA GTGTACCGAA TTTTATCCCC TGGTTACAGA
```

FIG. 21 (contd.)


```
4251  TTTCCAATCT AAGTCCCATT TATAATAAAG TCCATGTTCC ATTTTTAAAG
      AAAGGTTAGA TTCAGGGTAA ATATTATTTC AGGTACAAGG TAAAAATTTC
4301  GACAATCCTT TCGGTTTAAA ACCAGGCACG ATTACCCAAA CAACTCACAA
      CTGTTAGGAA AGCCAAATTT TGGTCCGTGC TAATGGGTTT GTTGAGTGTT
4351  CGGTAAAGCA CTGTGAATCT TCTCTGTTCT GCAATCCCAA CTTGGTTTCT
      GCCATTTCGT GACACTTAGA AGAGACAAGA CGTTAGGGTT GAACCAAAGA
4401  GCTCAGAAAC CCTCCCTCTT TCCAATCGGT AATTAAATAA CAAAAGGAAA
      CGAGTCTTTG GGAGGGAGAA AGGTTAGCCA TTAATTTATT GTTTTCCTTT
4451  AAACTTAAGA TGCTTCAACC CCGTTTCGTG ACACTTTGAA AAAAGAATCA
      TTTGAATTCT ACGAAGTTGG GGCAAAGCAC TGTGAAACTT TTTTCTTAGT
4501  CCTCTTGCAA ACACCCGCTC CCGACCCCCG CCGCTGAAGC CCGGCGTCCA
      GGAGAACGTT TGTGGGCGAG GGCTGGGGGC GGCGACTTCG GGCCGCAGGT
4551  GAGGCCTAAG CGCGGGTGCC CGCCCCCACC CGGGAGCGCG GGCCTCGTGG
      CTCCGGATTC GCGCCCACGG GCGGGGGTGG GCCCTCGCGC CCGGAGCACC
4601  TCAGCGCATC CGCGGGGAGA AACAAAGGCC GCGGCACGGG GGCTCAAGGG
      AGTCGCGTAG GCGCCCCTCT TTGTTTCCGG CGCCGTGCCC CCGAGTTCCC
4651  CACTGCGCCA CACCGCACGC GCCTACCCCC GCGCGGCCAC GTTAACTGGC
      GTGACGCGGT GTGGCGTGCG CGGATGGGGG CGCGCCGGTG CAATTGACCG
4701  GGTCGCCGCA GCCTCGGGAC AGCCGGCCGC GCGCCGCCAG GCTCGCGGAC
      CCAGCGGCGT CGGAGCCCTG TCGGCCGGCG CGCGGCGGTC CGAGCGCCTG
4751  GCGGGACCAC GCGCCGCCCT CCGGGAGGCC CAAGTCTCGA CCCAGCCCCG
      CGCCCTGGTG CGCGGCGGGA GGCCCTCCGG GTTCAGAGCT GGGTCGGGGC
4801  CGTGGCGCTG GGGGAGGGGG CGCCTCCGCC GGAACGCGGG TGGGGGAGGG
      GCACCGCGAC CCCCTCCCCC GCGGAGGCGG CCTTGCGCCC ACCCCCTCCC
4851  GAGGGGGAAA TGCGCTTTGT CTCGAAATGG GGCAACCGTC GCCACAGCTC
      CTCCCCCTTT ACGCGAAACA GAGCTTTACC CCGTTGGCAG CGGTGTCGAG
4901  CCTACCCCCT CGAGGGCAGA GCAGTCCCCC CACTAACTAC CGGGCTGGCC
      GGATGGGGGA GCTCCCGTCT CGTCAGGGGG GTGATTGATG GCCCGACCGG
4951  GCGCGCCAGG CCAGCCGCGA GGCCACCGCC CGACCCTCCA CTCCTTCCCG
      CGCGCGGTCC GGTCGGCGCT CCGGTGGCGG GCTGGGAGGT GAGGAAGGGC
5001  CAGCTCCCGG CGCGGGGTCC GGCGAGAAGG GGAGGGGAGG GGAGCGGAGA
      GTCGAGGGCC GCGCCCCAGG CCGCTCTTCC CCTCCCCTCC CCTCGCCTCT
5051  ACCGGGCCCC CGGGACGCGT GTGGCATCTG AAGCACCACC AGCGAGCGAG
      TGGCCCGGGG GCCCTGCGCA CACCGTAGAC TTCGTGGTGG TCGCTCGCTC
5101  AGCTAGAGAG AAGGAAAGCC ACCGACTTCA CCGCCTCCGA GCTGCTCCGG
      TCGATCTCTC TTCCTTTCGG TGGCTGAAGT GGCGGAGGCT CGACGAGGCC
5151  GTCGCGGGTC TGCAGCGTCT CCGGCCCTCC GCGCCTACAG CTCAAGCCAC
      CAGCGCCCAG ACGTCGCAGA GGCCGGGAGG CGCGGATGTC GAGTTCGGTG
5201  ATCCGAAGGG GGAGGGAGCC GGGAGCTGCG CGCGGGGCCG CCGGGGGGAG
      TAGGCTTCCC CCTCCCTCGG CCCTCGACGC GCGCCCCGGC GGCCCCCCTC
5251  GGGTGGCACC GCCCACGCCG GGCGGCCACG AAGGGCGGGG CAGCGGGCGC
      CCCACCGTGG CGGGTGCGGC CCGCCGGTGC TTCCCGCCCC GTCGCCCGCG
5301  GCGCGCGGCG GGGGGAGGGG CCGGCGCCGC GCCCGCTGGG AATTGGGGCC
      CGCGCGCCGC CCCCCTCCCC GGCCGCGGCG CGGGCGACCC TTAACCCCGG
5351  CTAGGGGGAG GGCGGAGGCG CCGACGACCG CGGCACTTAC CGTTCGCGGC
      GATCCCCCTC CCGCCTCCGC GGCTGCTGGC GCCGTGAATG GCAAGCGCCG
5401  GTGGCGCCCG GTGGTCCCCA AGGGGAGGGA AGGGGGAGGC GGGGCGAGGA
      CACCGCGGGC CACCAGGGGT TCCCCTCCCT TCCCCCTCCG CCCCGCTCCT
5451  CAGTGACCGG AGTCTCCTCA GCGGTGGCTT TTCTGCTTGG CAGCCTCAGC
      GTCACTGGCC TCAGAGGAGT CGCCACCGAA AAGACGAACC GTCGGAGTCG
5501  GGCTGGCGCC AAAACCGGAC TCCGCCCACT TCCTCGCCCG CCGGTGCGAG
      CCGACCGCGG TTTTGGCCTG AGGCGGGTGA AGGAGCGGGC GGCCACGCTC
5551  GGTGTGGAAT CCTCCAGACG CTGGGGGAGG GGGAGTTGGG AGCTTAAAAA
      CCACACCTTA GGAGGTCTGC GACCCCCTCC CCCTCAACCC TCGAATTTTT
5601  CTAGTACCCC TTTGGGACCA CTTTCAGCAG CGAACTCTCC TGTACACCAG
      GATCATGGGG AAACCCTGGT GAAAGTCGTC GCTTGAGAGG ACATGTGGTC
5651  GGGTCAGTTC CACAGACGCG GGCCAGGGGT GGGTCATTGC GGCGTGAACA
      CCCAGTCAAG GTGTCTGCGC CCGGTCCCCA CCCAGTAACG CCGCACTTGT
```

FIG. 21(contd.) 45/93

```
5701  ATAATTTGAC TAGAAGTTGA TTCGGGTGTT TCCGGAAGGG GCCGAGTCAA
      TATTAAACTG ATCTTCAACT AAGCCCACAA AGGCCTTCCC CGGCTCAGTT
5751  TCCGCCGAGT TGGGGCACGG AAAACAAAAA GGGAAGGCTA CTAAGATTTT
      AGGCGGCTCA ACCCCGTGCC TTTTGTTTTT CCCTTCCGAT GATTCTAAAA
5801  TCTGGCGGGG GTTATCATTG GCGTAACTGC AGGGACCACC TCCCGGGTTG
      AGACCGCCCC CAATAGTAAC CGCATTGACG TCCCTGGTGG AGGGCCCAAC
5851  AGGGGGCTGG ATCTCCAGGC TGCGGATTAA GCCCCTCCCG TCGGCGTTAA
      TCCCCCGACC TAGAGGTCCG ACGCCTAATT CGGGGAGGGC AGCCGCAATT
5901  TTTCAAACTG CGCGACGTTT CTCACCTGCC TTCGCCAAGG CAGGGGCCGG
      AAAGTTTGAC GCGCTGCAAA GAGTGGACGG AAGCGGTTCC GTCCCCGGCC
5951  GACCCTATTC CAAGAGGTAG TAACTAGCAG GACTCTAGCC TTCCGCAATT
      CTGGGATAAG GTTCTCCATC ATTGATCGTC CTGAGATCGG AAGGCGTTAA
6001  CATTGAGCGC ATTTACGGAA GTAACGTCGG GTACTGTCTC TGGCCGCAAG
      GTAACTCGCG TAAATGCCTT CATTGCAGCC CATGACAGAG ACCGGCGTTC
6051  GGTGGGAGGA GTACGCATTT GGCGTAAGGT GGGGCGTAGA GCCTTCCCGC
      CCACCCTCCT CATGCGTAAA CCGCATTCCA CCCCGCATCT CGGAAGGGCG
6101  CATTGGCGGC GGATAGGGCG TTTACGCGAC GGCCTGACGT AGCGGAAGAC
      GTAACCGCCG CCTATCCCGC AAATGCGCTG CCGGACTGCA TCGCCTTCTG
6151  GCGTTAGTGG GGGGGAAGGT TCTAGAAAAG CGGCGGCAGC GGCTCTAGCC
      CGCAATCACC CCCCCTTCCA AGATCTTTTC GCCGCCGTCG CCGAGATCGC
6201  GCAGTAGCAG CAGCGCCGGG TCCCGTGCGG AGGTGCTCCT CGCAGAGTTG
      CGTCATCGTC GTCGCGGCCC AGGGCACGCC TCCACGAGGA GCGTCTCAAC
6251  TTTCTCGAGC AGCGGCAGTT CTCACTACAG CGCCAGGACG AGTCCGGTTC
      AAAGAGCTCG TCGCCGTCAA GAGTGATGTC GCGGTCCTGC TCAGGCCAAG
6301  GTGTTCGTCC GCGGAGATCT CTCTCATCTC GCTCGGCTGC GGGAAATCGG
      CACAAGCAGG CGCCTCTAGA GAGAGTAGAG CGAGCCGACG CCCTTTAGCC
6351  GCTGAAGCGA CTGAGTCCGC GATGGAGGTA ACGGGTTTGA AATCAATGAG
      CGACTTCGCT GACTCAGGCG CTACCTCCAT TGCCCAAACT TTAGTTACTC
6401  TTATTGAAAA GGGCATGGCG AGGCCGTTGG CGCCTCAGTG GAAGTCGGCC
      AATAACTTTT CCCGTACCGC TCCGGCAACC GCGGAGTCAC CTTCAGCCGG
6451  AGCCGCCTCC GTGGGAGAGA GGCAGGAAAT CGGACCAATT CAGTAGCAGT
      TCGGCGGAGG CACCCTCTCT CCGTCCTTTA GCCTGGTTAA GTCATCGTCA
6501  GGGGCTTAAG GTTTATGAAC GGGGTCTTGA GCGGAGGCCT GAGCGTACAA
      CCCCGAATTC CAAATACTTG CCCCAGAACT CGCCTCCGGA CTCGCATGTT
6551  ACAGCTTCCC CACCCTCAGC CTCCCGGCGC CATTTCCCTT CACTGGGGGT
      TGTCGAAGGG GTGGGAGTCG GAGGGCCGCG GTAAAGGGAA GTGACCCCCA
6601  GGGGGATGGG GAGCTTTCAC ATGGCGGACG CTGCCCCGCT GGGGTGAAAG
      CCCCCTACCC CTCGAAAGTG TACCGCCTGC GACGGGGCGA CCCCACTTTC
6651  TGGGGCGCGG AGGCGGGAAT TCTTATTCCC TTTCTAAAGC ACGCTGCTTC
      ACCCCGCGCC TCCGCCCTTA AGAATAAGGG AAAGATTTCG TGCGACGAAG
6701  GGGGGCCACG GCGTCTCCTC GGCGAGCGTT TCGGCGGGCA GCAGGTCCTC
      CCCCCGGTGC CGCAGAGGAG CCGCTCGCAA AGCCGCCCGT CGTCCAGGAG
6751  GTGAGCGAGG CTGCGGAGCT TCCCCTCCCC CTCTCTCCCG GGAACCGATT
      CACTCGCTCC GACGCCTCGA AGGGGAGGGG GAGAGAGGGC CCTTGGCTAA
6801  TGGCGGCCGC CATTTTCATG GCTCGCCTTC CTCTCAGCGT TTTCCTTATA
      ACCGCCGGCG GTAAAAGTAC CGAGCGGAAG GAGAGTCGCA AAAGGAATAT
6851  ACTCTTTTAT TTTCTTAGTG TGCTTTCTCT ATCAAGAAGT AGAAGTGGTT
      TGAGAAAATA AAAGAATCAC ACGAAAGAGA TAGTTCTTCA TCTTCACCAA
6901  AACTATTTTT TTTTTCTTCT CGGGCTGTTT TCATATCGTT TCGAGGTGGA
      TTGATAAAAA AAAAAGAAGA GCCCGACAAA AGTATAGCAA AGCTCCACCT
6951  TTTGGAGTGT TTTGTGAGCT TGGATCTTTA GAGTCCTGCG CACCTCATTA
      AAACCTCACA AAACACTCGA ACCTAGAAAT CTCAGGACGC GTGGAGTAAT
7001  AAGGCGCTCA GCCTTCCCCT CGATGAAATG GCGCCATTGC GTTCGGAAGC
      TTCCGCGAGT CGGAAGGGGA GCTACTTTAC CGCGGTAACG CAAGCCTTCG
7051  CACACCGAAG AGCGGGGAGG GGGGGTGCTC CGGGTTTGCG GGCCCGGTTT
      GTGTGGCTTC TCGCCCCTCC CCCCCACGAG GCCCAAACGC CCGGGCCAAA
7101  CAGAGAAGAT ATCACCACCC AGGGCGTCGG GCCGGGTTCA ATGCGAGCCG
      GTCTCTTCTA TAGTGGTGGG TCCCGCAGCC CGGCCCAAGT TACGCTCGGC
```

FIG. 21 (contd.)


```
7151  TAGGACAAAG AAACCATTTT ATGTTTTTCC TGTCTTTTTT TTCCTTTGAG
      ATCCTGTTTC TTTGGTAAAA TACAAAAAGG ACAGAAAAAA AAGGAAACTC
7201  TAACGGTTTT ATCTGGGTCT GCAGTCAGTA AAACGACAGA TGAACCGCGG
      ATTGCCAAAA TAGACCCAGA CGTCAGTCAT TTTGCTGTCT ACTTGGCGCC
7251  CAAAATAAAC ATAAATTGGA AGCCATCGGC CACGAGGGGC AGGGACGAAG
      GTTTTATTTG TATTTAACCT TCGGTAGCCG GTGCTCCCCG TCCCTGCTTC
7301  GTGGTTTTCT GGGCGGGGGA GGGATATTCG CGTCAGAATC CTTTACTGTT
      CACCAAAAGA CCCGCCCCCT CCCTATAAGC GCAGTCTTAG GAAATGACAA
7351  CTTAAGGATT CCGTTTAAGT TGTAGAGCTG ACTCATTTTA AGTAATGTTG
      GAATTCCTAA GGCAAATTCA ACATCTCGAC TGAGTAAAAT TCATTACAAC
7401  TTACTGAGAA GTTTAACCCT TACGGGACAG ATCCATGGAC CTTTATAGAT
      AATGACTCTT CAAATTGGGA ATGCCCTGTC TAGGTACCTG GAAATATCTA
7451  GATTACGAGG AAAGTGAAAT AACGATTTTG TCCTTAGTTA TACTTCGATT
      CTAATGCTCC TTTCACTTTA TTGCTAAAAC AGGAATCAAT ATGAAGCTAA
7501  AAAACATGGC TTCAGAGGCT CCTTCCTGTA ATGCGTATGG ATTGATGTGC
      TTTTGTACCG AAGTCTCCGA GGAAGGACAT TACGCATACC TAACTACACG
7551  AAAACTGTTT TGGGCCTGGG CCGCTCTGTA TTTGAACTTT GTTACTTTTC
      TTTTGACAAA ACCCGGACCC GGCGAGACAT AAACTTGAAA CAATGAAAAG
7601  TCATTTTGTT TGCAATCTTG GTTGAACATT ACATTGATAA GCATAAGGTC
      AGTAAAACAA ACGTTAGAAC CAACTTGTAA TGTAACTATT CGTATTCCAG
7651  TCAAGCGAAG GGGGTCTACC TGGTTATTTT TCTTTGACCC TAAGCACGTT
      AGTTCGCTTC CCCCAGATGG ACCAATAAAA AGAAACTGGG ATTCGTGCAA
7701  TATAAAATAA CATTGTTTAA AATCGATAGT GGACATCGGG TAAGTTTGGA
      ATATTTTATT GTAACAAATT TTAGCTATCA CCTGTAGCCC ATTCAAACCT
7751  TAAATTGTGA GGTAAGTAAT GAGTTTTTGC TTTTTGTTAG TGATTTGTAA
      ATTTAACACT CCATTCATTA CTCAAAAACG AAAAACAATC ACTAAACATT
7801  AACTTGTTAT AAATGTACAT TATCCGTAAT TTCAGTTTAG AGATAACCTA
      TTGAACAATA TTTACATGTA ATAGGCATTA AAGTCAAATC TCTATTGGAT
7851  TGTGCTGACG ACAATTAAGA ATAAAAACTA GCTGAAAAAA TGAAAATAAC
      ACACGACTGC TGTTAATTCT TATTTTTGAT CGACTTTTTT ACTTTTATTG
7901  TATCGTGACA AGTAACCATT TCAAAAGACT GCTTTGTGTC TCATAGGAGC
      ATAGCACTGT TCATTGGTAA AGTTTTCTGA CGAAACACAG AGTATCCTCG
7951  TAGTTTGATC ATTTCAGTTA ATTTTTTCTT TAATTTTTAC GAGTCATGAA
      ATCAAACTAG TAAAGTCAAT TAAAAAAGAA ATTAAAAATG CTCAGTACTT
8001  AACTACAGGA AAAAAAATCT GAACTGGGTT TTACCACTAC TTTTTAGGAG
      TTGATGTCCT TTTTTTTAGA CTTGACCCAA AATGGTGATG AAAAATCCTC
8051  TTGGGAGCAT GCGAATGGAG GGAGAGCTCC GTAGAACTGG GATGAGAGCA
      AACCCTCGTA CGCTTACCTC CCTCTCGAGG CATCTTGACC CTACTCTCGT
8101  GCAATTAATG CTGCTTGCTA GGAACAAAAA ATAATTGATT GAAAATTACG
      CGTTAATTAC GACGAACGAT CCTTGTTTTT TATTAACTAA CTTTTAATGC
8151  TGTGACTTTT TAGTTTGCAT TATGCGTTTG TAGCAGTTGG TCCTGGATAT
      ACACTGAAAA ATCAAACGTA ATACGCAAAC ATCGTCAACC AGGACCTATA
8201  CACTTTCTCT CGTTTGAGGT TTTTTAACCT AGTTAACTTT TAAGACAGGT
      GTGAAAGAGA GCAAACTCCA AAAAATTGGA TCAATTGAAA ATTCTGTCCA
8251  TTCCTTAACA TTCATAAGTG CCCAGAATAC AGCTGTGTAG TACAGCATAT
      AAGGAATTGT AAGTATTCAC GGGTCTTATG TCGACACATC ATGTCGTATA
8301  AAAGATTTCA GCTCTGAGGT TTTTCCTATT GACTTGGAAA ATTGTTTTGT
      TTTCTAAAGT CGAGACTCCA AAAAGGATAA CTGAACCTTT TAACAAAACA
8351  GCCTGTCGCT TGCCACATGG CCAATCAAGT AAGCTTATCG ATACCGGTGG
      CGGACAGCGA ACGGTGTACC GGTTAGTTCA TTCGAATAGC TATGGCCACC
8401  CGCGCCAATT GTTAATTAAG ATCTGGCCCA ATGGGCCGTA CGAATTCCTT
      GCGCGGTTAA CAATTAATTC TAGACCGGGT TACCCGGCAT GCTTAAGGAA
8451  AGGCTACCGG GTAGGGGAGG CGCTTTTCCC AAGGCAGTCT GGAGCATGCG
      TCCGATGGCC CATCCCCTCC GCGAAAAGGG TTCCGTCAGA CCTCGTACGC
8501  CTTTAGCAGC CCCGCTGGGC ACTTGGCGCT ACACAAGTGG CCTCTGGCCT
      GAAATCGTCG GGGCGACCCG TGAACCGCGA TGTGTTCACC GGAGACCGGA
8551  CGCACAGATT CCACATCCAC CGGCCGGTAG GCGCCAACCG GCTCCGTTCT
      GCGTGTGTAA GGTGTAGGTG GCCGGCCATC CGCGGTTGGC CGAGGCAAGA
```

FIG. 21 (contd.) 47/93

```
8601   TTGGTGGCCC CTTCGCGCCA CCTTCTACTC CTCCCCTAGT CAGGAAGTTC
       AACCACCGGG GAAGCGCGGT GGAAGATGAG GAGGGGATCA GTCCTTCAAG
8651   CCCCCCGCCC CGCAGCTCGC GTCGTGCAGG ACGTGACAAA TGGAAGTAGC
       GGGGGGCGGG GCGTCGAGCG CAGCACGTCC TGCACTGTTT ACCTTCATCG
8701   ACGTCTCACT AGTCTCGTGC AGATGGACAG CACCGCTGAG CAATGGAAGC
       TGCAGAGTGA TCAGAGCACG TCTACCTGTC GTGGCGACTC GTTACCTTCG
8751   GGGTAGGCCT TTGGGGCAGC GGCCAATAGC AGCTTTGCTC CTTCGCTTTC
       CCCATCCGGA AACCCCGTCG CCGGTTATCG TCGAAACGAG GAAGCGAAAG
8801   TGGGCTCAGA GGCTGGGAAG GGGTGGGTCC GGGGGCGGGC TCAGGGGCGG
       ACCCGAGTCT CCGACCCTTC CCCACCCAGG CCCCCGCCCG AGTCCCCGCC
8851   GCTCAGGGGC GGGGCGGGCG CCCGAAGGTC CTCCGGAGGC CCGGCATTCT
       CGAGTCCCCG CCCCGCCCGC GGGCTTCCAG GAGGCCTCCG GGCCGTAAGA
8901   GCACGCTTCA AAAGCGCACG TCTGCCGCGC TGTTCTCCTC TTCCTCATCT
       CGTGCGAAGT TTTCGCGTGC AGACGGCGCG ACAAGAGGAG AAGGAGTAGA
8951   CCGGGCCTTT CGACCAGCTT ACCATGACCG AGTACAAGCC CACGGTGCGC
       GGCCCGGAAA GCTGGTCGAA TGGTACTGGC TCATGTTCGG GTGCCACGCG
9001   CTCGCCACCC GCGACGACGT CCCCAGGGCC GTACGCACCC TCGCCGCCGC
       GAGCGGTGGG CGCTGCTGCA GGGGTCCCGG CATGCGTGGG AGCGGCGGCG
9051   GTTCGCCGAC TACCCCGCCA CGCGCCACAC CGTCGATCCG GACCGCCACA
       CAAGCGGCTG ATGGGGCGGT GCGCGGTGTG GCAGCTAGGC CTGGCGGTGT
9101   TCGAGCGGGT CACCGAGCTG CAAGAACTCT TCCTCACGCG CGTCGGGCTC
       AGCTCGCCCA GTGGCTCGAC GTTCTTGAGA AGGAGTGCGC GCAGCCCGAG
9151   GACATCGGCA AGGTGTGGGT CGCGGACGAC GGCGCCGCGG TGGCGGTCTG
       CTGTAGCCGT TCCACACCCA GCGCCTGCTG CCGCGGCGCC ACCGCCAGAC
9201   GACCACGCCG GAGAGCGTCG AAGCGGGGGC GGTGTTCGCC GAGATCGGCC
       CTGGTGCGGC CTCTCGCAGC TTCGCCCCCG CCACAAGCGG CTCTAGCCGG
9251   CGCGCATGGC CGAGTTGAGC GGTTCCCGGC TGGCCGCGCA GCAACAGATG
       GCGCGTACCG GCTCAACTCG CCAAGGGCCG ACCGGCGCGT CGTTGTCTAC
9301   GAAGGCCTCC TGGCGCCGCA CCGGCCCAAG GAGCCCGCGT GGTTCCTGGC
       CTTCCGGAGG ACCGCGGCGT GGCCGGGTTC CTCGGGCGCA CCAAGGACCG
9351   CACCGTCGGC GTCTCGCCCG ACCACCAGGG CAAGGGTCTG GGCAGCGCCG
       GTGGCAGCCG CAGAGCGGGC TGGTGGTCCC GTTCCCAGAC CCGTCGCGGC
9401   TCGTGCTCCC CGGAGTGGAG GCGGCCGAGC GCGCCGGGGT GCCCGCCTTC
       AGCACGAGGG GCCTCACCTC CGCCGGCTCG CGCGGCCCCA CGGGCGGAAG
9451   CTGGAGACCT CCGCGCCCCG CAACCTCCCC TTCTACGAGC GGCTCGGCTT
       GACCTCTGGA GGCGCGGGGC GTTGGAGGGG AAGATGCTCG CCGAGCCGAA
9501   CACCGTCACC GCCGACGTCG AGGTGCCCGA AGGACCGCGC ACCTGGTGCA
       GTGGCAGTGG CGGCTGCAGC TCCACGGGCT TCCTGGCGCG TGGACCACGT
9551   TGACCCGCAA GCCCGGTGCC TGACGCCCGC CCCACGACCC GCAGCGCCCG
       ACTGGGCGTT CGGGCCACGG ACTGCGGGCG GGGTGCTGGG CGTCGCGGGC
9601   ACCGAAAGGA GCGCACGACC CCATGCATCG TAGAGCTCGC TGATCAGCCT
       TGGCTTTCCT CGCGTGCTGG GGTACGTAGC ATCTCGAGCG ACTAGTCGGA
9651   CGACTGTGCC TTCTAGTTGC CAGCCATCTG TTGTTTGCCC CTCCCCCGTG
       GCTGACACGG AAGATCAACG GTCGGTAGAC AACAAACGGG GAGGGGGCAC
9701   CCTTCCTTGA CCCTGGAAGG TGCCACTCCC ACTGTCCTTT CCTAATAAAA
       GGAAGGAACT GGGACCTTCC ACGGTGAGGG TGACAGGAAA GGATTATTTT
9751   TGAGGAAATT GCATCGCATT GTCTGAGTAG GTGTCATTCT ATTCTGGGGG
       ACTCCTTTAA CGTAGCGTAA CAGACTCATC CACAGTAAGA TAAGACCCCC
9801   GTGGGGTGGG GCAGGACAGC AAGGGGGGGG ATTGGGRAGA CAATAGCAGG
       CACCCCACCC CGTCCTGTCG TTCCCCCCCC TAACCCYTCT GTTATCGTCC
9851   CATGCTGGGG GGGCGGTGGG GGCTATGGCT TCTGAGGCGG AAAGAACCAG
       GTACGACCCC CCCGCCACCC CCGATACCGA AGACTCCGCC TTTCTTGGTC
9901   CTGGGGCTCG AGGGCCGCCA CCGCGGTGGA GCTCCAGCTT TTGTTCCCTT
       GACCCCGAGC TCCCGGCGGT GGCGCCACCT CGAGGTCGAA AACAAGGGAA
9951   TAGTGAGGGT TAATTTCGAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC
       ATCACTCCCA ATTAAAGCTC GAACCGCATT AGTACCAGTA TCGACAAAGG
10001  TGTGTGAAAT TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA
       ACACACTTTA ACAATAGGCG AGTGTTAAGG TGTGTTGTAT GCTCGGCCTT
```

FIG. 21(contd.) 48/93

```
10051   GCATAAAGTG TAAAGCCTGG GGTGCCTAAT GAGTGAGCTA ACTCACATTA
        CGTATTTCAC ATTTCGGACC CCACGGATTA CTCACTCGAT TGAGTGTAAT
10101   ATTGCGTTGC GCTCACTGCC CGCTTTCCAG TCGGGAAACC TGTCGTGCCA
        TAACGCAACG CGAGTGACGG GCGAAAGGTC AGCCCTTTGG ACAGCACGGT
10151   GCATCGCGAG CACTTTTCGG GGAAATGTGC GCGGAACCCC TATTTGTTTA
        CGTAGCGCTC GTGAAAAGCC CCTTTACACG CGCCTTGGGG ATAAACAAAT
10201   TTTTTCTAAA TACATTCAAA TATGTATCCG CTCATGAGAC AATAACCCTG
        AAAAAGATTT ATGTAAGTTT ATACATAGGC GAGTACTCTG TTATTGGGAC
10251   ATAAATGCTT CAATAATATT GAAAAAGGAA GAGTATGAGT ATTCAACATT
        TATTTACGAA GTTATTATAA CTTTTTCCTT CTCATACTCA TAAGTTGTAA
10301   TCCGTGTCGC CCTTATTCCC TTTTTTGCGG CATTTTGCCT TCCTGTTTTT
        AGGCACAGCG GGAATAAGGG AAAAAACGCC GTAAAACGGA AGGACAAAAA
10351   GCTCACCCAG AAACGCTGGT GAAAGTAAAA GATGCTGAAG ATCAGTTGGG
        CGAGTGGGTC TTTGCGACCA CTTTCATTTT CTACGACTTC TAGTCAACCC
10401   TGCACGAGTG GGTTACATCG AACTGGATCT CAACAGCGGT AAGATCCTTG
        ACGTGCTCAC CCAATGTAGC TTGACCTAGA GTTGTCGCCA TTCTAGGAAC
10451   AGAGTTTTCG CCCCGAAGAA CGTTTTCCAA TGATGAGCAC TTTTAAAGTT
        TCTCAAAAGC GGGGCTTCTT GCAAAAGGTT ACTACTCGTG AAAATTTCAA
10501   CTGCTATGTG GCGCGGTATT ATCCCGTATT GACGCCGGGC AAGAGCAACT
        GACGATACAC CGCGCCATAA TAGGGCATAA CTGCGGCCCG TTCTCGTTGA
10551   CGGTCGCCGC ATACACTATT CTCAGAATGA CTTGGTTGAG TACTCACCAG
        GCCAGCGGCG TATGTGATAA GAGTCTTACT GAACCAACTC ATGAGTGGTC
10601   TCACAGAAAA GCATCTTACG GATGGCATGA CAGTAAGAGA ATTATGCAGT
        AGTGTCTTTT CGTAGAATGC CTACCGTACT GTCATTCTCT TAATACGTCA
10651   GCTGCCATAA CCATGAGTGA TAACACTGCG GCCAACTTAC TTCTGACAAC
        CGACGGTATT GGTACTCACT ATTGTGACGC CGGTTGAATG AAGACTGTTG
10701   GATCGGAGGA CCGAAGGAGC TAACCGCTTT TTTGCACAAC ATGGGGGATC
        CTAGCCTCCT GGCTTCCTCG ATTGGCGAAA AAACGTGTTG TACCCCCTAG
10751   ATGTAACTCG CCTTGATCGT TGGGAACCGG AGCTGAATGA AGCCATACCA
        TACATTGAGC GGAACTAGCA ACCCTTGGCC TCGACTTACT TCGGTATGGT
10801   AACGACGAGC GTGACACCAC GATGCCTGTA GCAATGGCAA CAACGTTGCG
        TTGCTGCTCG CACTGTGGTG CTACGGACAT CGTTACCGTT GTTGCAACGC
10851   CAAACTATTA ACTGGCGAAC TACTTACTCT AGCTTCCCGG CAACAATTAA
        GTTTGATAAT TGACCGCTTG ATGAATGAGA TCGAAGGGCC GTTGTTAATT
10901   TAGACTGGAT GGAGGCGGAT AAAGTTGCAG GACCACTTCT GCGCTCGGCC
        ATCTGACCTA CCTCCGCCTA TTTCAACGTC CTGGTGAAGA CGCGAGCCGG
10951   CTTCCGGCTG GCTGGTTTAT TGCTGATAAA TCTGGAGCCG GTGAGCGTGG
        GAAGGCCGAC CGACCAAATA ACGACTATTT AGACCTCGGC CACTCGCACC
11001   GTCTCGCGGT ATCATTGCAG CACTGGGGCC AGATGGTAAG CCCTCCCGTA
        CAGAGCGCCA TAGTAACGTC GTGACCCCGG TCTACCATTC GGGAGGGCAT
11051   TCGTAGTTAT CTACACGACG GGGAGTCAGG CAACTATGGA TGAACGAAAT
        AGCATCAATA GATGTGCTGC CCCTCAGTCC GTTGATACCT ACTTGCTTTA
11101   AGACAGATCG CTGAGATAGG TGCCTCACTG ATTAAGCATT GGTAACTGTC
        TCTGTCTAGC GACTCTATCC ACGGAGTGAC TAATTCGTAA CCATTGACAG
11151   AGACTCGCGA CACTGCATTA ATGAATCGGC CAACGCGCGG GGAGAGGCGG
        TCTGAGCGCT GTGACGTAAT TACTTAGCCG GTTGCGCGCC CCTCTCCGCC
11201   TTTGCGTATT GGGCGCTCTT CCGCTTCCTC GCTCACTGAC TCGCTGCGCT
        AAACGCATAA CCCGCGAGAA GGCGAAGGAG CGAGTGACTG AGCGACGCGA
11251   CGGTCGTTCG GCTGCGGCGA GCGGTATCAG CTCACTCAAA GGCGGTAATA
        GCCAGCAAGC CGACGCCGCT CGCCATAGTC GAGTGAGTTT CCGCCATTAT
11301   CGGTTATCCA CAGAATCAGG GGATAACGCA GGAAAGAACA TGTGAGCAAA
        GCCAATAGGT GTCTTAGTCC CCTATTGCGT CCTTTCTTGT ACACTCGTTT
11351   AGGCCAGCAA AAGGCCAGGA ACCGTAAAAA GGCCGCGTTG CTGGCGTTTT
        TCCGGTCGTT TTCCGGTCCT TGGCATTTTT CCGGCGCAAC GACCGCAAAA
11401   TCCATAGGCT CCGCCCCCCT GACGAGCATC ACAAAAATCG ACGCTCAAGT
        AGGTATCCGA GGCGGGGGGA CTGCTCGTAG TGTTTTTAGC TGCGAGTTCA
11451   CAGAGGTGGC GAAACCCGAC AGGACTATAA AGATACCAGG CGTTTCCCCC
        GTCTCCACCG CTTTGGGCTG TCCTGATATT TCTATGGTCC GCAAAGGGGG
```

```
11501   TGGAAGCTCC CTCGTGCGCT CTCCTGTTCC GACCCTGCCG CTTACCGGAT
        ACCTTCGAGG GAGCACGCGA GAGGACAAGG CTGGGACGGC GAATGGCCTA
11551   ACCTGTCCGC CTTTCTCCCT TCGGGAAGCG TGGCGCTTTC TCATAGCTCA
        TGGACAGGCG GAAAGAGGGA AGCCCTTCGC ACCGCGAAAG AGTATCGAGT
11601   CGCTGTAGGT ATCTCAGTTC GGTGTAGGTC GTTCGCTCCA AGCTGGGCTG
        GCGACATCCA TAGAGTCAAG CCACATCCAG CAAGCGAGGT TCGACCCGAC
11651   TGTGCACGAA CCCCCCGTTC AGCCCGACCG CTGCGCCTTA TCCGGTAACT
        ACACGTGCTT GGGGGGCAAG TCGGGCTGGC GACGCGGAAT AGGCCATTGA
11701   ATCGTCTTGA GTCCAACCCG GTAAGACACG ACTTATCGCC ACTGGCAGCA
        TAGCAGAACT CAGGTTGGGC CATTCTGTGC TGAATAGCGG TGACCGTCGT
11751   GCCACTGGTA ACAGGATTAG CAGAGCGAGG TATGTAGGCG GTGCTACAGA
        CGGTGACCAT TGTCCTAATC GTCTCGCTCC ATACATCCGC CACGATGTCT
11801   GTTCTTGAAG TGGTGGCCTA ACTACGGCTA CACTAGAAGG ACAGTATTTG
        CAAGAACTTC ACCACCGGAT TGATGCCGAT GTGATCTTCC TGTCATAAAC
11851   GTATCTGCGC TCTGCTGAAG CCAGTTACCT TCGGAAAAAG AGTTGGTAGC
        CATAGACGCG AGACGACTTC GGTCAATGGA AGCCTTTTTC TCAACCATCG
11901   TCTTGATCCG GCAAACAAAC CACCGCTGGT AGCGGTGGTT TTTTTGTTTG
        AGAACTAGGC CGTTTGTTTG GTGGCGACCA TCGCCACCAA AAAAACAAAC
11951   CAAGCAGCAG ATTACGCGCA GAAAAAAAGG ATCTCAAGAA GATCCTTTGA
        GTTCGTCGTC TAATGCGCGT CTTTTTTTCC TAGAGTTCTT CTAGGAAACT
12001   TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAACTC A
        AGAAAAGATG CCCCAGACTG CGAGTCACCT TGCTTTTGAG T
```

FIG. 21(contd.)

FIG. 23 CET1020 nucleotide sequence

```
   1  CGTTGTAAAA CGACGGCCAG TGAATTGTAA TACGACTCAC TATAGGGCGA
      GCAACATTTT GCTGCCGGTC ACTTAACATT ATGCTGAGTG ATATCCCGCT
  51  ATTGGGTACC GGGCCCCCCC TCGAAGTTTA AACATTTAAA TCTAGAAGCT
      TAACCCATGG CCCGGGGGGG AGCTTCAAAT TTGTAAATTT AGATCTTCGA
 101  TTTAACCCTC TATCCCTTTA AACTTCCTTG ATCCAGTGTA AGCACCTCCT
      AAATTGGGAG ATAGGGAAAT TTGAAGGAAC TAGGTCACAT TCGTGGAGGA
 151  AGAAAGTCAG TAGACAATAA AACAAAAGTT CTGCTTCACC GATTTACATT
      TCTTTCAGTC ATCTGTTATT TTGTTTTCAA GACGAAGTGG CTAAATGTAA
 201  TATAACCAAA TACCCTTCAC CAATACAATA AAAAAACAAA ACAACAAAAA
      ATATTGGTTT ATGGGAAGTG GTTATGTTAT TTTTTTGTTT TGTTGTTTTT
 251  ACCCCAACCA TCTGAGAAAT AATCTTCTCC TTTCCCAGCT TTATTCCCAG
      TGGGGTTGGT AGACTCTTTA TTAGAAGAGG AAAGGGTCGA AATAAGGGTC
 301  GATTCTACAT GACCAAATTA CCAGAGTCAC CACTCATTTT AATCACAACA
      CTAAGATGTA CTGGTTTAAT GGTCTCAGTG GTGAGTAAAA TTAGTGTTGT
 351  TAGTGTCAAA TAACTAGAAA ACATGAGACA ACAATGGAGA GCTGAGTAAC
      ATCACAGTTT ATTGATCTTT TGTACTCTGT TGTTACCTCT CGACTCATTG
 401  TATTAGTAGT AGTACTTTAC CAGAGAATGG CCTCTATAGG CTCACATGTA
      ATAATCATCA TCATGAAATG GTCTCTTACC GGAGATATCC GAGTGTACAT
 451  GGAATGGTTG GTCCCCAGGT GGTAGGTAGA GCTGTTTGAG GATTACGTGG
      CCTTACCAAC CAGGGGTCCA CCATCCATCT CGACAAACTC CTAATGCACC
 501  CCTTCTTGGA TGGGGGGTGG GGGTGGGGTG GGAGGGTTGG GTGGTGGGTA
      GGAAGAACCT ACCCCCCACC CCCACCCCAC CCTCCCAACC CACCACCCAT
 551  CTTAAGAGGT TTCAAAAGTC AATATTGTTT GCATTTAGCT CTTCCTTGTA
      GAATTCTCCA AAGTTTTCAG TTATAACAAA CGTAAATCGA GAAGGAACAT
 601  CTTGTGGATC AAACACAACC TGTCAGCTAC TGCTTCAAAT GTCATGCCTG
      GAACACCTAG TTTGTGTTGG ACAGTCGATG ACGAAGTTTA CAGTACGGAC
 651  CTGCCATCTT CTCAGCAGGA TGGTCATGGC CTCACCCTCT TCAACTGTAA
      GACGGTAGAA GAGTCGTCCT ACCAGTACCG GAGTGGGAGA AGTTGACATT
 701  ATCTTTCTTT CTTTTCTTCT TTTTCTTTTG GTTTCGAGAC AGGGTTTCTC
      TAGAAAGAAA GAAAAGAAGA AAAAGAAAAC CAAAGCTCTG TCCCAAAGAG
 751  TGTATAGTCC TGGCTGTCCT GGAACTCACT TTGTAGACCA GGCTGGCCTT
      ACATATCAGG ACCGACAGGA CCTTGAGTGA AACATCTGGT CCGACCGGAA
 801  GAACTCAGAA ATCCGCCTGC CTCTGCCTCC CTAGCACTGG GATTAAAGGC
      CTTGAGTCTT TAGGCGGACG GAGACGGAGG GATCGTGACC CTAATTTCCG
 851  GTGCGCCACC ACGCCCAGCT TTCAACTGGA AATCTTAATA AACTTTCCTA
      CACGCGGTGG TGCGGGTCGA AAGTTGACCT TTAGAATTAT TTGAAAGGAT
 901  GAAGTGGCCT TGGTTATGGG AGCTTATCAC AGCAATAGAA CAGCAATTAT
      CTTCACCGGA ACCAATACCC TCGAATAGTG TCGTTATCTT GTCGTTAATA
 951  GACTGGAGTA TGATAGTTAA AAACAAGCAA GCAAGCAAGC AAACACACAC
      CTGACCTCAT ACTATCAATT TTTGTTCGTT CGTTCGTTCG TTTGTGTGTG
1001  ACCAAAACAA CAAAACCCCA AGACAGAGTC ACATGTAGCC CAGGCTAGCC
      TGGTTTTGTT GTTTTGGGGT TCTGTCTCAG TGTACATCGG GTCCGATCGG
1051  TCCAAATTCA CTATATAACT GAAGAAGACC CCTAATTCCC ATTCCTCTAG
      AGGTTTAAGT GATATATTGA CTTCTTCTGG GGATTAAGGG TAAGGAGATC
1101  AATCTATACC TCAAGTACTG AATGGCTTGG TTCACAATAC CCCACTAAAT
      TTAGATATGG AGTTCATGAC TTACCGAACC AAGTGTTATG GGGTGATTTA
1151  GATTGGTCTT ACTAAGTGCA ACAAGGTAAA CCTAAAACTT CAGCCCTCAG
      CTAACCAGAA TGATTCACGT TGTTCCATTT GGATTTTGAA GTCGGGAGTC
1201  ACATCCCTTT TCCAGTATCA ATTTATAAAA TTAGATCCCA AGGATAAAAA
      TGTAGGGAAA AGGTCATAGT TAAATATTTT AATCTAGGGT TCCTATTTTT
1251  TTAATTGTAA AGTAAAATCA GAGTTCTAGC ATCAACTACA GGCTCAACCA
      AATTAACATT TCATTTTAGT CTCAAGATCG TAGTTGATGT CCGAGTTGGT
1301  TGGGGACCAC AAATAAACTA AAAGGGATAA GACTGGCTTC CCCATAATTA
      ACCCCTGGTG TTTATTTGAT TTTCCCTATT CTGACCGAAG GGGTATTAAT
```

FIG. 23(contd.) 51/93

```
1351  TTACATTTAG ATAATTTTCC TGACTACTCA ACAAAGCTAA AATATCACCA
      AATGTAAATC TATTAAAAGG ACTGATGAGT TGTTTCGATT TTATAGTGGT
1401  CTGGTTTATT TTCTCCTTCT AGGGTTTAAG CTCACTCTGA GGAGGGGCAT
      GACCAAATAA AAGAGGAAGA TCCCAAATTC GAGTGAGACT CCTCCCCGTA
1451  GCGGCACACA CTCATAGCAT CCAGGAAATA GAAATATGGT GACTATCATG
      CGCCGTGTGT GAGTATCGTA GGTCCTTTAT CTTTATACCA CTGATAGTAC
1501  GGTTCAGGGC CAACCTAGGC TTTAGAGAAA AACCTTGTCC CACAAACCAA
      CCAAGTCCCG GTTGGATCCG AAATCTCTTT TTGGAACAGG GTGTTTGGTT
1551  AAATGTCTCT TTTTTATTCT ATCAGGGGTG GATGGATTTG TTAAAGAAGT
      TTTACAGAGA AAAAATAAGA TAGTCCCCAC CTACCTAAAC AATTTCTTCA
1601  GCTTTTAAAA ACCTTGAGAT GGTTATTTAG AAGTCCCCAT GGGATACCAA
      CGAAAATTTT TGGAACTCTA CCAATAAATC TTCAGGGGTA CCCTATGGTT
1651  AATAACCCAC TATTTATATG CCCAAGCATT TCACCTCCAC AACAGTGCTA
      TTATTGGGTG ATAAATATAC GGGTTCGTAA AGTGGAGGTG TTGTCACGAT
1701  TGCACCCTTT AACATTTTTG AGACAGTAGC CCAGTCTAGT CTTTAACTTG
      ACGTGGGAAA TTGTAAAAAC TCTGTCATCG GGTCAGATCA GAAATTGAAC
1751  CAGTGATTTT TCCTGATTCA GCTTCTCCCA GTGCTGGAAT TATAGGTATG
      GTCACTAAAA AGGACTAAGT CGAAGAGGGT CACGACCTTA ATATCCATAC
1801  CACCACCATG TGTAACTACA GATGCTACTT AAAAATTTTT TAAAGGAATC
      GTGGTGGTAC ACATTGATGT CTACGATGAA TTTTTAAAAA ATTTCCTTAG
1851  ACAAAAATAA CCCCCTATCA AATGCCTAGT CCCTCTAACC ATCACCAAGT
      TGTTTTTATT GGGGGATAGT TTACGGATCA GGGAGATTGG TAGTGGTTCA
1901  GAAGGATCAC GCAGGAAAAA AAAAATCACC AGCAGCACCT CAGAACCAGG
      CTTCCTAGTG CGTCCTTTTT TTTTTAGTGG TCGTCGTGGA GTCTTGGTCC
1951  ATACTCAGTC CATCAGCATC CAGGGCCATA CCCACACTCA CAGCATCTCC
      TATGAGTCAG GTAGTCGTAG GTCCCGGTAT GGGTGTGAGT GTCGTAGAGG
2001  ACAGTTTACC AGATGATTCA TGCTTATCAC TGTATTGGGT CATCTAAGAG
      TGTCAAATGG TCTACTAAGT ACGAATAGTG ACATAACCCA GTAGATTCTC
2051  TGACCATCAG GGCTTCTGAT CACAGAATCT AGTCCACTTT GCAGACCAGT
      ACTGGTAGTC CCGAAGACTA GTGTCTTAGA TCAGGTGAAA CGTCTGGTCA
2101  TGAAGTCATG CACTATATGA GATAGAAATA CCCTCTTGCT CATTTTGGTC
      ACTTCAGTAC GTGATATACT CTATCTTTAT GGGAGAACGA GTAAAACCAG
2151  AGAAATTCAA GGATAAAAAC CCATGTTTTG TTAATGCACA CCTCCATATG
      TCTTTAAGTT CCTATTTTTG GGTACAAAAC AATTACGTGT GGAGGTATAC
2201  ATTGAGATCA ATGTGTCCTA ATTAATGTAG AAACCACAAC TGTAAATTTC
      TAACTCTAGT TACACAGGAT TAATTACATC TTTGGTGTTG ACATTTAAAG
2251  ACTCTTTTGA CATGAATCTT TTTCTAGACA GGGTCTTGGA TGCAGCCCCG
      TGAGAAAACT GTACTTAGAA AAAGATCTGT CCCAGAACCT ACGTCGGGGC
2301  ACTACCCAGA ATTTTGGAAT CCAGGCTAGC CTCAAACTCA AGGCAATCTG
      TGATGGGTCT TAAAACCTTA GGTCCGATCG GAGTTTGAGT TCCGTTAGAC
2351  CTTGCTTCAG CTTCTCACAG GCTGGATCAC AAACATACAC CTTCAGACCC
      GAACGAAGTC GAAGAGTGTC CGACCTAGTG TTTGTATGTG GAAGTCTGGG
2401  ATTTTTTTTT CCTCCCTCCG TTTTTGGTTT CTCTGTGTAG CCCTGGGTGT
      TAAAAAAAAA GGAGGGAGGC AAAAACCAAA GAGACACATC GGGACCCACA
2451  CCGTGGACTC GCTGTGTAGA TCTATCTACC AGCCTCTGTC TTGGAGTACT
      GGCACCTGAG CGACACATCT AGATAGATGG TCGGAGACAG AACCTCATGA
2501  GGGATTAAAG TTGTGGGCTA CCACTGCCTG GCTGACCCAG TTTTATTTAT
      CCCTAATTTC AACACCCGAT GGTGACGGAC CGACTGGGTC AAAATAAATA
2551  TTTAAATATA ACTTGACAAA AATAAATTTG TCTAACTTAC TAGAAATCCC
      AAATTTATAT TGAACTGTTT TTATTTAAAC AGATTGAATG ATCTTTAGGG
2601  AAGAAAACTA ACACTGGATT TAGCAACAGT CAGAAATCGC TGAAAAGAAA
      TTCTTTTGAT TGTGACCTAA ATCGTTGTCA GTCTTTAGCG ACTTTTCTTT
2651  CAGAATTGAT CTAACAGTCT TAGATCACTC CTAGACAGTT TGTAATTCTT
      GTCTTAACTA GATTGTCAGA ATCTAGTGAG GATCTGTCAA ACATTAAGAA
2701  GCTCATGGCA ACGTGAGCTC TATCTAACTC ACTCTCTGTG CACTAATGAA
      CGAGTACCGT TGCACTCGAG ATAGATTGAG TGAGAGACAC GTGATTACTT
2751  TGCTCAGTGT CTCCAGAACA GCACAGCTTC CAGGGTAATC ATGCCAACCC
      ACGAGTCACA GAGGTCTTGT CGTGTCGAAG GTCCCATTAG TACGGTTGGG
```

FIG. 23(contd.) 52/93

```
2801  ACAAGACTTT TATAGAGCTG TCCACGACTC TTCCCCCATT CAGCTCATTA
      TGTTCTGAAA ATATCTCGAC AGGTGCTGAG AAGGGGGTAA GTCGAGTAAT
2851  ACAATATGAT GGAGCTCCTG TGTGGAAATC AAGGCACACT CTGGTAGAAA
      TGTTATACTA CCTCGAGGAC ACACCTTTAG TTCCGTGTGA GACCATCTTT
2901  CTTGTTTTTT CTTTCCACTT TTCCTTGGGC TCTGAAGATT GAGCTGTTTT
      GAACAAAAAA GAAAGGTGAA AAGGAACCCG AGACTTCTAA CTCGACAAAA
2951  ATAACCCACA AACATGCATT TTTTACCTCA AAAGCATCCA GCAAAAACTG
      TATTGGGTGT TTGTACGTAA AAAATGGAGT TTTCGTAGGT CGTTTTTGAC
3001  TACAACGCTT TTTCAAAAAA ATGTATTGTG ATCCTCCTTA AGAAAAGCCT
      ATGTTGCGAA AAAGTTTTTT TACATAACAC TAGGAGGAAT TCTTTTCGGA
3051  TACTTAGTGT TAATTCCTTT TTCTTTAGAA TGCTGGTAAA TACAAGGACT
      ATGAATCACA ATTAAGGAAA AAGAAATCTT ACGACCATTT ATGTTCCTGA
3101  TAGGTAGGCT GGCTTCTAAC AGCAATTCAC CCACTTATGA TGGGATTAAA
      ATCCATCCGA CCGAAGATTG TCGTTAAGTG GGTGAATACT ACCCTAATTT
3151  GGAAGGCACA ACCATGTCCA CCACAGGTTC TAGCTCCCCC ACCCACACGC
      CCTTCCGTGT TGGTACAGGT GGTGTCCAAG ATCGAGGGGG TGGGTGTGCG
3201  CCAGAGAGGG TTTTTCTGTG TAGCTCTGAC TATTCTGGAA TTCACACTGC
      GGTCTCTCCC AAAAAGACAC ATCGAGACTG ATAAGACCTT AAGTGTGACG
3251  AGACCAGGCT GGTCTCGAAC TCAGAGATCC ACCACCACAT GGTTTCTTAA
      TCTGGTCCGA CCAGAGCTTG AGTCTCTAGG TGGTGGTGTA CCAAAGAATT
3301  TTGTAATTTT AAAGAAAAAA AAAAATCCTT CAGTTAAGAT TCTTATGTTC
      AACATTAAAA TTTCTTTTTT TTTTTAGGAA GTCAATTCTA AGAATACAAG
3351  TAGGTTTTCA CAAACTTACC AATGTAGTTT TATTGGAGGC CATTTTTTAA
      ATCCAAAAGT GTTTGAATGG TTACATCAAA ATAACCTCCG GTAAAAAATT
3401  ATTTAATCGG AGACTTGAAG AGCTATTGCA AGAAAAAAAA TGTAGGACAG
      TAAATTAGCC TCTGAACTTC TCGATAACGT TCTTTTTTTT ACATCCTGTC
3451  TTAAAATTTC ATGACACACA AAAGGCAGCT ACAAGTTTTG TGTGGATTTC
      AATTTTAAAG TACTGTGTGT TTTCCGTCGA TGTTCAAAAC ACACCTAAAG
3501  AACATGTAAA TTTCGGGTAA AAATGCAGGA AAACAGTTGA GTTCCCGTGT
      TTGTACATTT AAAGCCCATT TTTACGTCCT TTTGTCAACT CAAGGGCACA
3551  TATTAGTATG TTACTAATAA TTTCAGTATG TTAGTGAAAA TAATCTTACT
      ATAATCATAC AATGATTATT AAAGTCATAC AATCACTTTT ATTAGAATGA
3601  AAAACACTGG TACCTCAGAC AACTTTACAT GGTGAGGATT GTTACTTTCC
      TTTTGTGACC ATGGAGTCTG TTGAAATGTA CCACTCCTAA CAATGAAAGG
3651  CAATCCATAT AGAATTTTAA CAATTTTAGT GTTTATTTTG GATGAAAGGA
      GTTAGGTATA TCTTAAAATT GTTAAAATCA CAAATAAAAC CTACTTTCCT
3701  AATGACTATC TTTTGTTAGC AAATTACCAT AAGATCTTTT TCTTTAGATT
      TTACTGATAG AAAACAATCG TTTAATGGTA TTCTAGAAAA AGAAATCTAA
3751  TCTGAATACT CCAAGGAGCT CATATAATTC CATCCTTATT TTTTCAGAGG
      AGACTTATGA GGTTCCTCGA GTATATTAAG GTAGGAATAA AAAAGTCTCC
3801  CCCTCCCTGT TCAATCACGG TATAAAAAAA GGAACACATT AAGATGTCCC
      GGGAGGGACA AGTTAGTGCC ATATTTTTTT CCTTGTGTAA TTCTACAGGG
3851  AGTCCTATTT TCTGGCTTTT TTTTTCCGGG GGTGGTGGTG CGGTAATCAC
      TCAGGATAAA AGACCGAAAA AAAAAGGCCC CCACCACCAC GCCATTAGTG
3901  TCTCTATAGT CCAGTCTGGG CTTCAACGCC TGGCAATCCC CAGCCTCAAG
      AGAGATATCA GGTCAGACCC GAAGTTGCGG ACCGTTAGGG GTCGGAGTTC
3951  CTCCCAAGTA CTGTCCTGAT AAGGATAGAA GGAGTCGACC TCCTTCACGC
      GAGGGTTCAT GACAGGACTA TTCCTATCTT CCTCAGCTGG AGGAAGTGCG
4001  TCCCCTCCGA GGAGGGCTCC TTCCCAGCTC CATTCCCCGG TCGGGAGCCC
      AGGGGAGGCT CCTCCCGAGG AAGGGTCGAG GTAAGGGGCC AGCCCTCGGG
4051  GTCCCCCACC CGAGAGCGCG GGCCTCGTGG TCAGCGCCTC CGCGGGGAGA
      CAGGGGGTGG GCTCTCGCGC CCGGAGCACC AGTCGCGGAG GCGCCCCTCT
4101  AACAAAGGCG GCGGCGGGGG CTCAAGGGCA CTGCGCCACG GGCCCGCGCC
      TTGTTTCCGC CGCCGCCCCC GAGTTCCCGT GACGCGGTGC CCGGGCGCGG
4151  TCCCCCATCC GGCGGCGGCC ACGTAGCCGG GAGCGCGCCG CAGCCCGGAG
      AGGGGGTAGG CCGCCGCCGG TGCATCGGCC CTCGCGCGGC GTCGGGCCTC
4201  CCTCGGGCCT CGCAGCTGCA GAGCCTGAAC CGCTCTCTCC CTGCGGGCCT
      GGAGCCCGGA GCGTCGACGT CTCGGACTTG GCGAGAGAGG GACGCCCGGA
```

FIG. 23(contd.) 53/93

```
4251  GCGACGAGGC TGGGGGAGGG GAGGCCCGCG CTTTGTCTGG AGTCTCGGTA
      CGCTGCTCCG ACCCCCTCCC CTCCGGGCGC GAAACAGACC TCAGAGCCAT
4301  GCTGTCATCC GGCTCCCACC CTCATGCACA ATTGTCCCAT CTCCCCCACG
      CGACAGTAGG CCGAGGGTGG GAGTACGTGT TAACAGGGTA GAGGGGGTGC
4351  CACCGGCGCG GCGCCCGCCT CAGCGAGGCC CCAGCCGGTT TCCCGCAGCC
      GTGGCCGCGC CGCGGGCGGA GTCGCTCCGG GGTCGGCCAA AGGGCGTCGG
4401  CGCGGCCCAC GGGGCTCGCA GCCTCCCCGC AAGCTCGGAC GCACGGAGCA
      GCGCCGGGTG CCCCGAGCGT CGGAGGGGCG TTCGAGCCTG CGTGCCTCGT
4451  TCCTAAACCC CACCACACGC AAGATCGAAA AAAAGCAAAG GCACGAACTT
      AGGATTTGGG GTGGTGTGCG TTCTAGCTTT TTTTCGTTTC CGTGCTTGAA
4501  CACCGCTCCG ATGCTCAGGG CCGCGGATCC TGCAGAGTCT CCCGCCTGCG
      GTGGCGAGGC TACGAGTCCC GGCGCCTAGG ACGTCTCAGA GGGCGGACGC
4551  CGCTTCGGTT CAGCCACATC CGAGGGGAGG GGGCGCGGGC AGCTCCGCCG
      GCGAAGCCAA GTCGGTGTAG GCTCCCCTCC CCCGCGCCCG TCGAGGCGGC
4601  GGGGGGAGGG GGAGCACCGC CCACGCCCTG GCCGCGCGGG GCCCGCCGGG
      CCCCCCTCCC CCTCGTGGCG GGTGCGGGAC CGGCGCGCCC CGGGCGGCCC
4651  AACGCGTCCT GCGGGGGGCG GCGCGCGCAA TGCTCACCGT CCGCGGCGTG
      TTGCGCAGGA CGCCCCCCGC CGCGCGCGTT ACGAGTGGCA GGCGCCGCAC
4701  GCGCCCAGGG GGTCTCCTGG CTGGGGGGAG GGGGGGGAAG GCGGGCAGGA
      CCCGGGTCCC CCAGAGGACC GACCCCCCTC CCCCCCCTTC CGCCCGTCCT
4751  AGGACCGCGG AGGCCTCTCT GCGTCTCGGA GCGCGCCAAA GCGGGGCTCC
      TCCTGGCGCC TCCGGAGAGA CGCAGAGCCT CGCGCGGTTT CGCCCCGAGG
4801  ACCCACCTCC TTGCCCGGAT CTTGAAGGCC GGGGAGATAA ACAGCGGGGT
      TGGGTGGAGG AACGGGCCTA GAACTTCCGG CCCCTCTATT TGTCGCCCCA
4851  TCTTTAAGCA CCACCTCTCA CTAGGCGCGG GATCCCAAGG CTTGTGGCAT
      AGAAATTCGT GGTGGAGAGT GATCCGCGCC CTAGGGTTCC GAACACCGTA
4901  CCGGGGTGGT ACTTGGACTA AAAGTCCTTC TGGGAGGGAC CGAGTGAGAA
      GGCCCCACCA TGAACCTGAT TTTCAGGAAG ACCCTCCCTG GCTCACTCTT
4951  CCCCTTTGGG ACGTGTAGAA ATATTTGTGT GGTTCGAGAA TATTTGTGCG
      GGGGAAACCC TGCACATCTT TATAAACACA CCAAGCTCTT ATAAACACGC
5001  GACGGGCTTG GCAAAGGCGT AGCTGCAGAG AGCACGCTTG GGTGGAGAGG
      CTGCCCGAAC CGTTTCCGCA TCGACGTCTC TCGTGCGAAC CCACCTCTCC
5051  GCCGCACGCC CCAGCGCCGG CCTAAGCCCC TCCCGACGGC GTTATTTCAA
      CGGCGTGCGG GGTCGCGGCC GGATTCGGGG AGGGCTGCCG CAATAAAGTT
5101  ACTGCGCGAC CGTTTCTCCG CTCCCTACGC GGAGGTGGGG GCCGGACCTA
      TGACGCGCTG GCAAAGAGGC GAGGGATGCG CCTCCACCCC CGGCCTGGAT
5151  GTTCCGGACG TAGTAACACG CCGAGCGCGA GCCTTCCGCA ATTCACGGAA
      CAAGGCCTGC ATCATTGTGC GGCTCGCGCT CGGAAGGCGT TAAGTGCCTT
5201  CACAGTTGCG CAAGTGATGT AAAGCAGTCC CGCTGTACCT AAAGGGGGAG
      GTGTCAACGC GTTCACTACA TTTCGTCAGG GCGACATGGA TTTCCCCCTC
5251  TGTCACGTAC TTGGCGTAAG GAGAGTGTAG GCCCTTCCCG CCATTGGCGG
      ACAGTGCATG AACCGCATTC CTCTCACATC CGGGAAGGGC GGTAACCGCC
5301  CGGTTAGGGC GTTTACGTAA CGGCGTGACG TAAGCGGAGA CGCGTTAGTG
      GCCAATCCCG CAAATGCATT GCCGCACTGC ATTCGCCTCT GCGCAATCAC
5351  GGGGGAAGGT TCTAGAAAAG CGGCGGTCTC GGCTCCAGCG GCAGTAGCAG
      CCCCCTTCCA AGATCTTTTC GCCGCCAGAG CCGAGGTCGC CGTCATCGTC
5401  CGGCGCCGGT CCCGTGTGCA GGAGCTCCTT TGCGGCCCAG TTTCTTGGCC
      GCCGCGGCCA GGGCACACGT CCTCGAGGAA ACGCCGGGTC AAAGAACCGG
5451  ATCGCCTGCT CTCCCCACAG CGCCAGGACG AGTCCCGTGC GCGTCCGTCC
      TAGCGGACGA GAGGGGTGTC GCGGTCCTGC TCAGGGCACG CGCAGGCAGG
5501  GCGGAGGTCT TTCTCATCTC GCTCGGCTGC GGGAAATCGG GCTGAAGCGA
      CGCCTCCAGA AAGAGTAGAG CGAGCCGACG CCCTTTAGCC CGACTTCGCT
5551  CTGAGTCCGC GATGGAGGTA ACGGGTTTGA AATCAATGAG TTATTAAAAA
      GACTCAGGCG CTACCTCCAT TGCCCAAACT TTAGTTACTC AATAATTTTT
5601  TGGCATGGCG AGGCCGTAGG CACCGCAATG GAAACCGGCC ACCCGCCTCC
      ACCGTACCGC TCCGGCATCC GTGGCGTTAC CTTTGGCCGG TGGGCGGAGG
5651  GTGGTCCGGC GGAGGGGATG CGGCCACTCG AGTGGCGGTT GGCCTTGGCG
      CACCAGGCCG CCTCCCCTAC GCCGGTGAGC TCACCGCCAA CCGGAACCGC
```

FIG. 23(contd.) 54/93

```
5701  AGTTTCTGAG GGGTCGTTGG AGGAGGCCTC TGATTGTCCG ACCGCCTTCC
      TCAAAGACTC CCCAGCAACC TCCTCCGGAG ACTAACAGGC TGGCGGAAGG
5751  CCGCCCTCAG CCGCCCGGCG CCATTTCCCT CAGTTGGGGT GGGGGATGGG
      GGCGGGAGTC GGCGGGCCGC GGTAAAGGGA GTCAACCCCA CCCCCTACCC
5801  AAGTGCCCGC CGCGACCGGG CTGGACCGCT AAAGTAGCGC GTGAGCGGGC
      TTCACGGGCG GCGCTGGCCC GACCTGGCGA TTTCATCGCG CACTCGCCCG
5851  CATCGCTGGC CTTTCGATGT GCGCGGGCCT AGGGGCTCGG TTGTGTTCGC
      GTAGCGACCG GAAAGCTACA CGCGCCCGGA TCCCCGAGCC AACACAAGCG
5901  GGCGGAACGT TTCTGGGGCC CCCCCGGCTT CCCGGAGCGA GTCTGCGAAG
      CCGCCTTGCA AAGACCCCGG GGGGGCCGAA GGGCCTCGCT CAGACGCTTC
5951  CTAGCTTCCC CTCCCCCCTC TCCCGGGAAC CGGATTTGGC GGCCGCCATT
      GATCGAAGGG GAGGGGGGAG AGGGCCCTTG GCCTAAACCG CCGGCGGTAA
6001  TTCCCGTCTC CTTCCTCGCC ACGATTTTGC TTTCAACGCT TTAGGTTTAC
      AAGGGCAGAG GAAGGAGCGG TGCTAAAACG AAAGTTGCGA AATCCAAATG
6051  TAGTTTGGTT TTCTTTTTTC ACCACTGCGT AGACGTGTTT AGCGATTTTC
      ATCAAACCAA AAGAAAAAAG TGGTGACGCA TCTGCACAAA TCGCTAAAAG
6101  CTTTCTTTTG GAAGTCTTCA TACCGTTTCG AGGTGGATTT AGCGTTTTGA
      GAAAGAAAAC CTTCAGAAGT ATGGCAAAGC TCCACCTAAA TCGCAAAACT
6151  GCTTGGGTCT TCAGCGTCCT GCGCACCTCG CTAAAGGCTC TCTGCCTTCC
      CGAACCCAGA AGTCGCAGGA CGCGTGGAGC GATTTCCGAG AGACGGAAGG
6201  CCTCGACGAA ATGGCGCCAT TGCTTTCTGA AGCCACCGAG GCGCGGGGTG
      GGAGCTGCTT TACCGCGGTA ACGAAAGACT TCGGTGGCTC CGCGCCCCAC
6251  GGGGCGGGGT GGCGGCGCTC CACGAGCTTT ACTGGAACAG GCAGAGAGAA
      CCCCGCCCCA CCGCCGCGAG GTGCTCGAAA TGACCTTGTC CGTCTCTCTT
6301  CGTAGTACAA CCGAGGCCTG GGCGGGTGGC TGAAGGCAGC GTCGCTGCAA
      GCATCATGTT GGCTCCGGAC CCGCCCACCG ACTTCCGTCG CAGCGACGTT
6351  AGAGACCGTT TTATTTTTCA TAATACGTAA GATTACGGGT GCTGTAGTAA
      TCTCTGGCAA AATAAAAAGT ATTATGCATT CTAATGCCCA CGACATCATT
6401  AGCACTTGAG CATTAGTATA GTAGGAGGAA GTCAAAGTGG AAAAAATGGG
      TCGTGAACTC GTAATCATAT CATCCTCCTT CAGTTTCACC TTTTTTACCC
6451  AGCGCTCATC AGGAAGCTAG GGAGGCTATG TTGAGTGCAG GGTTACTTTC
      TCGCGAGTAG TCCTTCGATC CCTCCGATAC AACTCACGTC CCAATGAAAG
6501  CTTTTATTGC AGAACTTTTA TCTGCTTAAA GGATCCTCGG ATCGAAATAA
      GAAAATAACG TCTTGAAAAT AGACGAATTT CCTAGGAGCC TAGCTTTATT
6551  TTCAAATTAT AAGCATTTTT AAGGGAATCT TCGAATTTGT TGGTAAAGTC
      AAGTTTAATA TTCGTAAAAA TTCCCTTAGA AGCTTAAACA ACCATTTCAG
6601  AACGGATCCT TAGCACGTGG TGTTCACTTT AAGGAAGTGA AATAGCTGAC
      TTGCCTAGGA ATCGTGCACC ACAAGTGAAA TTCCTTCACT TTATCGACTG
6651  TTTTCATAGT TAGCCTTCGC TTAAAGCCTG GTTCAGTGGA CGAAAATCCA
      AAAAGTATCA ATCGGAAGCG AATTTCGGAC CAAGTCACCT GCTTTTAGGT
6701  CGTCCTGGCT ATATAAAAAC TTAGTTTGGG GTCACAGTGT TTGAGCGTGG
      GCAGGACCGA TATATTTTTG AATCAAACCC CAGTGTCACA AACTCGCACC
6751  TCATTCGGTT TTTTTATTTT TTATTTGTTT GAAATTATGA TGCATCATTA
      AGTAAGCCAA AAAAATAAAA AATAAACAAA CTTTAATACT ACGTAGTAAT
6801  CACTGATAAG CATTAGCTTT CGAATTGAAA GGGGTCTCCT TGGTTATTTT
      GTGACTATTC GTAATCGAAA GCTTAACTTT CCCCAGAGGA ACCAATAAAA
6851  CTTTGACTCT AAGCACACTT ATAAATAAAA TAACCTTGTT TATAATCGAT
      GAAACTGAGA TTCGTGTGAA TATTTATTTT ATTGGAACAA ATATTAGCTA
6901  AGTGGACGTC TGGTAAGTTT GGAAAAAACC CGAGGTAAGT AAAGAGCTTT
      TCACCTGCAG ACCATTCAAA CCTTTTTTGG GCTCCATTCA TTTCTCGAAA
6951  TGCTTTCGTT AGTGATATGA AAAAACAAGG TGTATTTAAT ACTTGCAACT
      ACGAAAGCAA TCACTATACT TTTTTGTTCC ACATAAATTA TGAACGTTGA
7001  TAGTTTAAGG AAAGCCAATT TACTGACATT TTAGTAGAGC TACCAGAAAC
      ATCAAATTCC TTTCGGTTAA ATGACTGTAA AATCATCTCG ATGGTCTTTG
7051  ACTATTTGGA GTCCTGATTA AGGCTTTTGT AACTATTTTG ACTATTTAAA
      TGATAAACCT CAGGACTAAT TCCGAAAACA TTGATAAAAC TGATAAATTT
7101  ACAATTTTGG TCGTTTTTAT TAAACATTTC AAAACCTAAA AATTGTAAAC
      TGTTAAAACC AGCAAAAATA ATTTGTAAAG TTTTGGATTT TTAACATTTG
```

FIG. 23(contd.) 55/93

```
7151  ATTGGCTTTT TGAGCACATT TTGGAGAAAC TTACAAATTT AGGCTATACA
      TAACCGAAAA ACTCGTGTAA AACCTCTTTG AATGTTTAAA TCCGATATGT
7201  GTAAAATAAC GGATTTGTTT TATAATTTTG CTTTTTCATT TCGTTGTGCA
      CATTTTATTG CCTAAACAAA ATATTAAAAC GAAAAAGTAA AGCAACACGT
7251  GTCATAGGTC CTGGATAGTA TGACCTAATT TATGAACATC TTGATAAGTT
      CAGTATCCAG GACCTATCAT ACTGGATTAA ATACTTGTAG AACTATTCAA
7301  TTTGTACTTA GCTATTGGAA AGCCAGTATT AAGTGCCTGA CAAAACCAGA
      AAACATGAAT CGATAACCTT TCGGTCATAA TTCACGGACT GTTTTGGTCT
7351  TTTAAGGTGA TATCTGGAGT TTCAGCATTC TTCATGGAGC TTGTTTCAGA
      AAATTCCACT ATAGACCTCA AAGTCGTAAG AAGTACCTCG AACAAAGTCT
7401  GTTGCAGGAT TTTTTTTTTT CATCTTGAGA TACTTACAAT TAACACCAGA
      CAACGTCCTA AAAAAAAAAA GTAGAACTCT ATGAATGTTA ATTGTGGTCT
7451  GGGGGCAGCT CAGGGAAAAG CAAATATGCC ACTTTTCAGA AACTGAATCT
      CCCCCGTCGA GTCCCTTTTC GTTTATACGG TGAAAAGTCT TTGACTTAGA
7501  TGGAAGTGGT GAATTTGGAA ACAGGTTTTT TAAATTTTTT TTAAATCTAA
      ACCTTCACCA CTTAAACCTT TGTCCAAAAA ATTTAAAAAA AATTTAGATT
7551  AAAGTAGTAA ATTTTGGACT TGGGTTGTAG AATTTAATGA ATTACAAAAG
      TTTCATCATT TAAAACCTGA ACCCAACATC TTAAATTACT TAATGTTTTC
7601  AATTCTTTAA TACCCTTTAA ATGACCTAAG AGCTGGGTAT GGTTTTTCTG
      TTAAGAAATT ATGGGAAATT TACTGGATTC TCGACCCATA CCAAAAAGAC
7651  AATTTTTTTG AAGAAAATCT AAGAAAGTTT ACGTGAATTA CAAGTTAGAT
      TTAAAAAAAC TTCTTTTAGA TTCTTTCAAA TGCACTTAAT CTTCAATCTA
7701  CGAATATTAG TGACTTTGAA ACTTGTATAG CTCAGGCAAT TTTTGGTGTA
      GCTTATAATC ACTGAAACTT TGAACATATC GAGTCCGTTA AAAACCACAT
7751  ACACAACTAA TATGCAGTTT AACATATGGT TTAAATTTGA TGTAAGTTTT
      TGTGTTGATT ATACGTCAAA TTGTATACCA AATTTAAACT ACATTCAAAA
7801  TTTTCTCCCC CCCAGAAAAC TTTAGAAACT GTTCCTTTGG AGAGGAAAAA
      AAAAGAGGGG GGGTCTTTTG AAATCTTTGA CAAGGAAACC TCTCCTTTTT
7851  GGTACTCTAC CAGCAGGTCA CCTCATATTT AAGAATTTAA TTTCCTGCAT
      CCATGAGATG GTCGTCCAGT GGAGTATAAA TTCTTAAATT AAAGGACGTA
7901  ACAAAGAAAG TGTAAATAAA AATTGAAATG GTATTTCCCT TTGCAGAGAG
      TGTTTCTTTC ACATTTATTT TTAACTTTAC CATAAAGGGA AACGTCTCTC
7951  AAAAGGAACA GTTCCGAAAG CTCTTTATTG GTGGCTTAAG CTTATCGATA
      TTTTCCTTGT CAAGGCTTTC GAGAAATAAC CACCGAATTC GAATAGCTAT
8001  CCGGTGGCGC GCCAATTGTT AATTAAGATC TGGCCCAATG GGCCGTACGA
      GGCCACCGCG CGGTTAACAA TTAATTCTAG ACCGGGTTAC CCGGCATGCT
8051  ATTCCTTAGG CTACCGGGTA GGGGAGGCGC TTTTCCCAAG GCAGTCTGGA
      TAAGGAATCC GATGGCCCAT CCCCTCCGCG AAAAGGGTTC CGTCAGACCT
8101  GCATGCGCTT TAGCAGCCCC GCTGGGCACT TGGCGCTACA CAAGTGGCCT
      CGTACGCGAA ATCGTCGGGG CGACCCGTGA ACCGCGATGT GTTCACCGGA
8151  CTGGCCTCGC ACACATTCCA CATCCACCGG CCGGTAGGCG CCAACCGGCT
      GACCGGAGCG TGTGTAAGGT GTAGGTGGCC GGCCATCCGC GGTTGGCCGA
8201  CCGTTCTTTG GTGGCCCCTT CGCGCCACCT TCTACTCCTC CCCTAGTCAG
      GGCAAGAAAC CACCGGGGAA GCGCGGTGGA AGATGAGGAG GGGATCAGTC
8251  GAAGTTCCCC CCCGCCCCGC AGCTCGCGTC GTGCAGGACG TGACAAATGG
      CTTCAAGGGG GGGCGGGGCG TCGAGCGCAG CACGTCCTGC ACTGTTTACC
8301  AAGTAGCACG TCTCACTAGT CTCGTGCAGA TGGACAGCAC CGCTGAGCAA
      TTCATCGTGC AGAGTGATCA GAGCACGTCT ACCTGTCGTG GCGACTCGTT
8351  TGGAAGCGGG TAGGCCTTTG GGGCAGCGGC CAATAGCAGC TTTGCTCCTT
      ACCTTCGCCC ATCCGGAAAC CCCGTCGCCG GTTATCGTCG AAACGAGGAA
8401  CGCTTTCTGG GCTCAGAGGC TGGGAAGGGG TGGGTCCGGG GGCGGGCTCA
      GCGAAAGACC CGAGTCTCCG ACCCTTCCCC ACCCAGGCCC CCGCCCGAGT
8451  GGGGCGGGCT CAGGGGCGGG GCGGGCGCCC GAAGGTCCTC CGGAGGCCCG
      CCCCGCCCGA GTCCCCGCCC CGCCCGCGGG CTTCCAGGAG GCCTCCGGGC
8501  GCATTCTGCA CGCTTCAAAA GCGCACGTCT GCCGCGCTGT TCTCCTCTTC
      CGTAAGACGT GCGAAGTTTT CGCGTGCAGA CGGCGCGACA AGAGGAGAAG
8551  CTCATCTCCG GGCCTTTCGA CCAGCTTACC ATGACCGAGT ACAAGCCCAC
      GAGTAGAGGC CCGGAAAGCT GGTCGAATGG TACTGGCTCA TGTTCGGGTG
```

FIG. 23(contd.) 56/93

```
8601  GGTGCGCCTC GCCACCCGCG ACGACGTCCC CAGGGCCGTA CGCACCCTCG
      CCACGCGGAG CGGTGGGCGC TGCTGCAGGG GTCCCGGCAT GCGTGGGAGC
8651  CCGCCGCGTT CGCCGACTAC CCCGCCACGC GCCACACCGT CGATCCGGAC
      GGCGGCGCAA GCGGCTGATG GGGCGGTGCG CGGTGTGGCA GCTAGGCCTG
8701  CGCCACATCG AGCGGGTCAC CGAGCTGCAA GAACTCTTCC TCACGCGCGT
      GCGGTGTAGC TCGCCCAGTG GCTCGACGTT CTTGAGAAGG AGTGCGCGCA
8751  CGGGCTCGAC ATCGGCAAGG TGTGGGTCGC GGACGACGGC GCCGCGGTGG
      GCCCGAGCTG TAGCCGTTCC ACACCCAGCG CCTGCTGCCG CGGCGCCACC
8801  CGGTCTGGAC CACGCCGGAG AGCGTCGAAG CGGGGGCGGT GTTCGCCGAG
      GCCAGACCTG GTGCGGCCTC TCGCAGCTTC GCCCCCGCCA CAAGCGGCTC
8851  ATCGGCCCGC GCATGGCCGA GTTGAGCGGT TCCCGGCTGG CCGCGCAGAA
      TAGCCGGGCG CGTACCGGCT CAACTCGCCA AGGGCCGACC GGCGCGTCTT
8901  CAGATGGAAG GCCTCCTGGC GCCGCACCGG CCCAAGGAGC CCGCGTGGTT
      GTCTACCTTC CGGAGGACCG CGGCGTGGCC GGGTTCCTCG GGCGCACCAA
8951  CCTGGCCACC GTCGCGTCTC GCCCGACCAC CAGGGCAAGG GTCTGGGCAG
      GGACCGGTGG CAGCGCAGAG CGGGCTGGTG GTCCCGTTCC CAGACCCGTC
9001  CGCCGTCGTG CTCCCCGGAG TGGAGGCGGC CGAGCGCGCC GGGGTGCCCG
      GCGGCAGCAC GAGGGGCCTC ACCTCCGCCG GCTCGCGCGG CCCCACGGGC
9051  CCTTCCTGGA GACCTCCGCG CCCCGCAACC TCCCCTTCTA CGAGCGGCTC
      GGAAGGACCT CTGGAGGCGC GGGGCGTTGG AGGGGAAGAT GCTCGCCGAG
9101  GGCTTCACCG TCACCGCCGA CGTCGAGGTG CCCGAAGGAC CGCGCACCTG
      CCGAAGTGGC AGTGGCGGCT GCAGCTCCAC GGGCTTCCTG GCGCGTGGAC
9151  GTGCATGACC CGCAAGCCCG GTGCCTGACG CCCGCCCCAC GACCCGCAGC
      CACGTACTGG GCGTTCGGGC CACGGACTGC GGGCGGGGTG CTGGGCGTCG
9201  GCCCGACCGA AAGGAGCGCA CGACCCCATG CATCGTAGAG CTCGCTGATC
      CGGGCTGGCT TTCCTCGCGT GCTGGGGTAC GTAGCATCTC GAGCGACTAG
9251  AGCCTCGACT GTGCCTTCTA GTTGCCAGCC ATCTGTTGTT TGCCCCTCCC
      TCGGAGCTGA CACGGAAGAT CAACGGTCGG TAGACAACAA ACGGGGAGGG
9301  CCGTGCCTTC CTTGACCCTG GAAGGTGCCA CTCCCACTGT CCTTTCCTAA
      GGCACGGAAG GAACTGGGAC CTTCCACGGT GAGGGTGACA GGAAAGGATT
9351  TAAAATGAGG AAATTGCATC GCATTGTCTG AGTAGGTGTC ATTCTATTCT
      ATTTTACTCC TTTAACGTAG CGTAACAGAC TCATCCACAG TAAGATAAGA
9401  GGGGGGTGGG GTGGGGCAGG ACAGCAAGGG GGGGGATTGG GRAGACAATA
      CCCCCCACCC CACCCCGTCC TGTCGTTCCC CCCCCTAACC CYTCTGTTAT
9451  GCAGGCATGC TGGGGGGGCG GTGGGGGCTA TGGCTTCTGA GGCGGAAAGA
      CGTCCGTACG ACCCCCCCGC CACCCCCGAT ACCGAAGACT CCGCCTTTCT
9501  ACCAGCTGGG GCTCGAGGGC CGCCACCGCG GTGGAGCTCC AGCTTTTGTT
      TGGTCGACCC CGAGCTCCCG GCGGTGGCGC CACCTCGAGG TCGAAAACAA
9551  CCCTTTAGTG AGGGTTAATT TCGAGCTTGG CGTAATCATG GTCATAGCTG
      GGGAAATCAC TCCCAATTAA AGCTCGAACC GCATTAGTAC CAGTATCGAC
9601  TTTCCTGTGT GAAATTGTTA TCCGCTCACA ATTCCACACA ACATACGAGC
      AAAGGACACA CTTTAACAAT AGGCGAGTGT TAAGGTGTGT TGTATGCTCG
9651  CGGAAGCATA AAGTGTAAAG CCTGGGGTGC CTAATGAGTG AGCTAACTCA
      GCCTTCGTAT TTCACATTTC GGACCCCACG GATTACTCAC TCGATTGAGT
9701  CATTAATTGC GTTGCGCTCA CTGCCCGCTT TCCAGTCGGG AAACCTGTCG
      GTAATTAACG CAACGCGAGT GACGGGCGAA AGGTCAGCCC TTTGGACAGC
9751  TGCCAGCATC GCGAGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT
      ACGGTCGTAG CGCTCGTGAA AAGCCCCTTT ACACGCGCCT TGGGGATAAA
9801  GTTTATTTTT CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA
      CAAATAAAAA GATTTATGTA AGTTTATACA TAGGCGAGTA CTCTGTTATT
9851  CCCTGATAAA TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA
      GGGACTATTT ACGAAGTTAT TATAACTTTT TCCTTCTCAT ACTCATAAGT
9901  ACATTTCCGT GTCGCCCTTA TTCCCTTTTT TGCGGCATTT TGCCTTCCTG
      TGTAAAGGCA CAGCGGGAAT AAGGGAAAAA ACGCCGTAAA ACGGAAGGAC
9951  TTTTTGCTCA CCCAGAAACG CTGGTGAAAG TAAAAGATGC TGAAGATCAG
      AAAAACGAGT GGGTCTTTGC GACCACTTTC ATTTTCTACG ACTTCTAGTC
10001 TTGGGTGCAC GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT
      AACCCACGTG CTCACCCAAT GTAGCTTGAC CTAGAGTTGT CGCCATTCTA
```

FIG. 23(contd.) 57/93

```
10051   CCTTGAGAGT TTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA
        GGAACTCTCA AAAGCGGGGC TTCTTGCAAA AGGTTACTAC TCGTGAAAAT
10101   AAGTTCTGCT ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG
        TTCAAGACGA TACACCGCGC CATAATAGGG CATAACTGCG GCCCGTTCTC
10151   CAACTCGGTC GCCGCATACA CTATTCTCAG AATGACTTGG TTGAGTACTC
        GTTGAGCCAG CGGCGTATGT GATAAGAGTC TTACTGAACC AACTCATGAG
10201   ACCAGTCACA GAAAAGCATC TTACGGATGG CATGACAGTA AGAGAATTAT
        TGGTCAGTGT CTTTTCGTAG AATGCCTACC GTACTGTCAT TCTCTTAATA
10251   GCAGTGCTGC CATAACCATG AGTGATAACA CTGCGGCCAA CTTACTTCTG
        CGTCACGACG GTATTGGTAC TCACTATTGT GACGCCGGTT GAATGAAGAC
10301   ACAACGATCG GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG
        TGTTGCTAGC CTCCTGGCTT CCTCGATTGG CGAAAAAACG TGTTGTACCC
10351   GGATCATGTA ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA
        CCTAGTACAT TGAGCGGAAC TAGCAACCCT TGGCCTCGAC TTACTTCGGT
10401   TACCAAACGA CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG
        ATGGTTTGCT GCTCGCACTG TGGTGCTACG GACATCGTTA CCGTTGTTGC
10451   TTGCGCAAAC TATTAACTGG CGAACTACTT ACTCTAGCTT CCCGGCAACA
        AACGCGTTTG ATAATTGACC GCTTGATGAA TGAGATCGAA GGGCCGTTGT
10501   ATTAATAGAC TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT
        TAATTATCTG ACCTACCTCC GCCTATTTCA ACGTCCTGGT GAAGACGCGA
10551   CGGCCCTTCC GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG
        GCCGGGAAGG CCGACCGACC AAATAACGAC TATTTAGACC TCGGCCACTC
10601   CGTGGGTCTC GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC
        GCACCCAGAG CGCCATAGTA ACGTCGTGAC CCCGGTCTAC CATTCGGGAG
10651   CCGTATCGTA GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC
        GGCATAGCAT CAATAGATGT GCTGCCCCTC AGTCCGTTGA TACCTACTTG
10701   GAAATAGACA GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA
        CTTTATCTGT CTAGCGACTC TATCCACGGA GTGACTAATT CGTAACCATT
10751   CTGTCAGACT CGCGACACTG CATTAATGAA TCGGCCAACG CGCGGGGAGA
        GACAGTCTGA GCGCTGTGAC GTAATTACTT AGCCGGTTGC GCGCCCCTCT
10801   GGCGGTTTGC GTATTGGGCG CTCTTCCGCT TCCTCGCTCA CTGACTCGCT
        CCGCCAAACG CATAACCCGC GAGAAGGCGA AGGAGCGAGT GACTGAGCGA
10851   GCGCTCGGTC GTTCGGCTGC GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG
        CGCGAGCCAG CAAGCCGACG CCGCTCGCCA TAGTCGAGTG AGTTTCCGCC
10901   TAATACGGTT ATCCACAGAA TCAGGGGATA ACGCAGGAAA GAACATGTGA
        ATTATGCCAA TAGGTGTCTT AGTCCCCTAT TGCGTCCTTT CTTGTACACT
10951   GCAAAAGGCC AGCAAAAGGC CAGGAACCGT AAAAAGGCCG CGTTGCTGGC
        CGTTTTCCGG TCGTTTTCCG GTCCTTGGCA TTTTTCCGGC GCAACGACCG
11001   GTTTTTCCAT AGGCTCCGCC CCCCTGACGA GCATCACAAA AATCGACGCT
        CAAAAAGGTA TCCGAGGCGG GGGGACTGCT CGTAGTGTTT TTAGCTGCGA
11051   CAAGTCAGAG GTGGCGAAAC CCGACAGGAC TATAAAGATA CCAGGCGTTT
        GTTCAGTCTC CACCGCTTTG GGCTGTCCTG ATATTTCTAT GGTCCGCAAA
11101   CCCCCTGGAA GCTCCCTCGT GCGCTCTCCT GTTCCGACCC TGCCGCTTAC
        GGGGGACCTT CGAGGGAGCA CGCGAGAGGA CAAGGCTGGG ACGGCGAATG
11151   CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCATA
        GCCTATGGAC AGGCGGAAAG AGGGAAGCCC TTCGCACCGC GAAAGAGTAT
11201   GCTCACGCTG TAGGTATCTC AGTTCGGTGT AGGTCGTTCG CTCCAAGCTG
        CGAGTGCGAC ATCCATAGAG TCAAGCCACA TCCAGCAAGC GAGGTTCGAC
11251   GGCTGTGTGC ACGAACCCCC CGTTCAGCCC GACCGCTGCG CCTTATCCGG
        CCGACACACG TGCTTGGGGG GCAAGTCGGG CTGGCGACGC GGAATAGGCC
11301   TAACTATCGT CTTGAGTCCA ACCCGGTAAG ACACGACTTA TCGCCACTGG
        ATTGATAGCA GAACTCAGGT TGGGCCATTC TGTGCTGAAT AGCGGTGACC
11351   CAGCAGCCAC TGGTAACAGG ATTAGCAGAG CGAGGTATGT AGGCGGTGCT
        GTCGTCGGTG ACCATTGTCC TAATCGTCTC GCTCCATACA TCCGCCACGA
11401   ACAGAGTTCT TGAAGTGGTG GCCTAACTAC GGCTACACTA GAAGGACAGT
        TGTCTCAAGA ACTTCACCAC CGGATTGATG CCGATGTGAT CTTCCTGTCA
11451   ATTTGGTATC TGCGCTCTGC TGAAGCCAGT TACCTTCGGA AAAAGAGTTG
        TAAACCATAG ACGCGAGACG ACTTCGGTCA ATGGAAGCCT TTTTCTCAAC
```

```
11501   GTAGCTCTTG ATCCGGCAAA CAAACCACCG CTGGTAGCGG TGGTTTTTTT
        CATCGAGAAC TAGGCCGTTT GTTTGGTGGC GACCATCGCC ACCAAAAAAA
11551   GTTTGCAAGC AGCAGATTAC GCGCAGAAAA AAAGGATCTC AAGAAGATCC
        CAAACGTTCG TCGTCTAATG CGCGTCTTTT TTTCCTAGAG TTCTTCTAGG
11601   TTTGATCTTT TCTACGGGGT CTGACGCTCA GTGGAACGAA AACTCA
        AAACTAGAAA AGATGCCCCA GACTGCGAGT CACCTTGCTT TTGAGT
```

FIG. 23(contd.)

FIG. 24

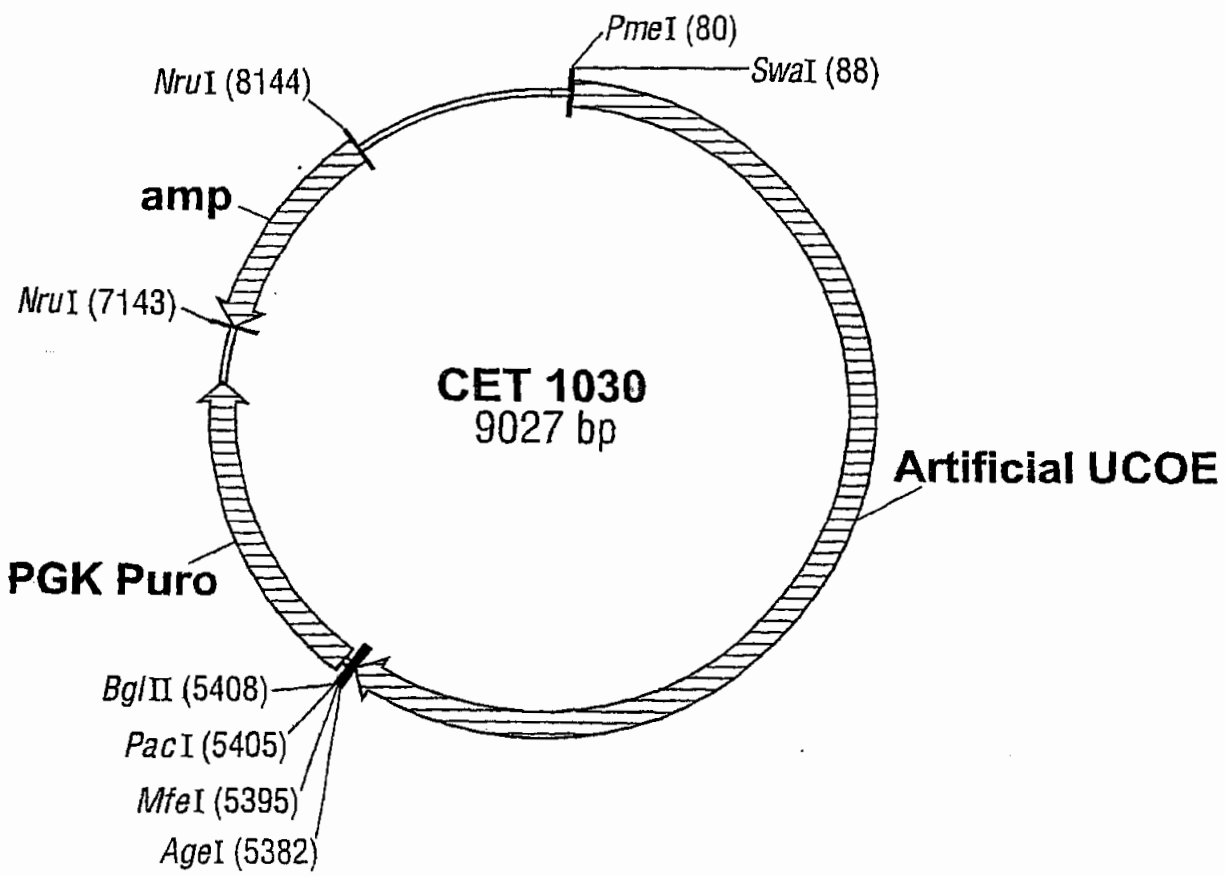


FIG. 25 CET1030 nucleotide sequence

```
   1  CGTTGTAAAA CGACGGCCAG TGAATTGTAA TACGACTCAC TATAGGGCGA
      GCAACATTTT GCTGCCGGTC ACTTAACATT ATGCTGAGTG ATATCCCGCT
  51  ATTGGGTACC GGGCCCCCCC TCGAAGTTTA AACATTTAAA TCTAGAACTA
      TAACCCATGG CCCGGGGGGG AGCTTCAAAT TTGTAAATTT AGATCTTGAT
 101  GTGGATCCCC CGGGCTGCAG GAATTCGATA TCAAGCTCAT GGCACCTGTA
      CACCTAGGGG GCCCGACGTC CTTAAGCTAT AGTTCGAGTA CCGTGGACAT
 151  TTGTACTCTT ATCAGTCATT ATATGGACTT TAACTTCCCC AGATATTATT
      AACATGAGAA TAGTCAGTAA TATACCTGAA ATTGAAGGGG TCTATAATAA
 201  TGGGCTCCTC CATAAGACTG TGAGCATCTG ACCACTGGAG TGTTGCTTCC
      ACCCGAGGAG GTATTCTGAC ACTCGTAGAC TGGTGACCTC ACAACGAAGG
 251  CATTATATCC CTGTTATCAA GCACAAGGTC AGGCACAGAG TAAGACTCAA
      GTAATATAGG GACAATAGTT CGTGTTCCAG TCCGTGTCTC ATTCTGAGTT
 301  AACATGTTTT GGAATGTATG ACTGGTATGA ACTACAAACC AGTAAGCTGA
      TTGTACAAAA CCTTACATAC TGACCATACT TGATGTTTGG TCATTCGACT
 351  TGTTTTCATT TTGAGTCTAT AAATCTAATT TTGTGGTGGT TTTGTGTATG
      ACAAAAGTAA AACTCAGATA TTTAGATTAA AACACCACCA AAACACATAC
 401  GCTCAAGGCT CAAATTGTAA AATTTAATAT TATGTGACCA AAGAAAGTTA
      CGAGTTCCGA GTTTAACATT TTAAATTATA ATACACTGGT TTCTTTCAAT
 451  TACCCAGAAC CTCAATTTCC TCACCTTCAA AATGGGGCAG TTTCTCACTC
      ATGGGTCTTG GAGTTAAAGG AGTGGAAGTT TTACCCCGTC AAAGAGTGAG
 501  ATTGGTCTGC TGTCACGATT TTAATGAGCT CATGCACAAA CAGCCCTTTA
      TAACCAGACG ACAGTGCTAA AATTACTCGA GTACGTGTTT GTCGGGAAAT
 551  TATAAGGTAA GTGCTGGATA AATGTTGGCT ACTATAATAA AATAAGCCTC
      ATATTCCATT CACGACCTAT TTACAACCGA TGATATTATT TTATTCGGAG
 601  TAAGATACTT GGTCAGCACA AGTACTACCC AAGAGTATGC ACTGTAAGTA
      ATTCTATGAA CCAGTCGTGT TCATGATGGG TTCTCATACG TGACATTCAT
 651  AACTGACAAA ATTGTGTATC TAAAACTGGC CAGATGAAAG AGAAACTTTT
      TTGACTGTTT TAACACATAG ATTTTGACCG GTCTACTTTC TCTTTGAAAA
 701  AAGGGGCCCT TCTGCGTGCC CGACACTGTG CTAGGCACTC ACACTATCCC
      TTCCCCGGGA AGACGCACGG GCTGTGACAC GATCCGTGAG TGTGATAGGG
 751  GACCCGAGAA ACCGATCTGC GACCCAGAGG AACTTACCAA GCCTCCAGCA
      CTGGGCTCTT TGGCTAGACG CTGGGTCTCC TTGAATGGTT CGGAGGTCGT
 801  TCTTGTGCAG CCCTACTCAT GGGACCATCT GGATACCCAC CCTTGTCTTT
      AGAACACGTC GGGATGAGTA CCCTGGTAGA CCTATGGGTG GGAACAGAAA
 851  ACAGGGAGCA GAACACACCT CTTATGTGTC AGAAAACAAA GTCCAGGAAG
      TGTCCCTCGT CTTGTGTGGA GAATACACAG TCTTTTGTTT CAGGTCCTTC
 901  TATATTTTTA CCTGAGGCAA TATCTGAAAA TTGTATGCTA CAGCCTCCAA
      ATATAAAAAT GGACTCCGTT ATAGACTTTT AACATACGAT GTCGGAGGTT
 951  AGTGAGTCTT CCTCTCAGTA CCTCTCTTCT AGGCACATGG AGCCCTTTCT
      TCACTCAGAA GGAGAGTCAT GGAGAGAAGA TCCGTGTACC TCGGGAAAGA
1001  TCCAAGTATT ATGTTTAACC ACTTAATGAA TGAAGTCCTG AAACTGCTTA
      AGGTTCATAA TACAAATTGG TGAATTACTT ACTTCAGGAC TTTGACGAAT
1051  CCCATGCTCC CTATAATCTC TGAGTAATCT TCCTTTTCCA CAACCTCAGG
      GGGTACGAGG GATATTAGAG ACTCATTAGA AGGAAAAGGT GTTGGAGTCC
1101  CATAATCTCA TCTTCTGTTT CTATTACAAT TTCAAATTCT GGAAAAAGGA
      GTATTAGAGT AGAAGACAAA GATAATGTTA AAGTTTAAGA CCTTTTTCCT
1151  AGTTGTGGTC TGGAATTATA TGGTCCAGAT GATCTGAAAC AAAAAGGACA
      TCAACACCAG ACCTTAATAT ACCAGGTCTA CTAGACTTTG TTTTTCCTGT
1201  GCACTATTAG TAATCATTTA GTTTTGAAGA CAGTCTAATA ATTTGCTGTC
      CGTGATAATC ATTAGTAAAT CAAAACTTCT GTCAGATTAT TAAACGACAG
1251  TCTAAAGTAC TATATTCCCT ATAGTTCTGG CATTTTAGAT AAAGGGTCAT
      AGATTTCATG ATATAAGGGA TATCAAGACC GTAAAATCTA TTTCCCAGTA
1301  AAATTAAATG CCTATATGGT GACATTATTC AGTGATTCAG ACTTCACAGC
      TTTAATTTAC GGATATACCA CTGTAATAAG TCACTAAGTC TGAAGTGTCG
```

FIG. 25(contd.)


```
1351  CTTTTTTTTT TTTTTACAAA GGTGTTCCAG GCATGAAAAA TTTTAAAGTA
      GAAAAAAAAA AAAAATGTTT CCACAAGGTC CGTACTTTTT AAAATTTCAT
1401  CTATACCTTT CCTAATTTTA CCTTTAAAGT TGTCCTGGAA ATATCTGGGT
      GATATGGAAA GGATTAAAAT GGAAATTTCA ACAGGACCTT TATAGACCCA
1451  TGACAAAGGC GATGAAACTG AACTGAGACT TAAAAAAAAG ATTACCCACC
      ACTGTTTCCG CTACTTTGAC TTGACTCTGA ATTTTTTTTC TAATGGGTGG
1501  TGGTTGTGCA CAAGCCTGCT TATGTCCCAA TCTCCAGTCT AGGGTCTGAT
      ACCAACACGT GTTCGGACGA ATACAGGGTT AGAGGTCAGA TCCCAGACTA
1551  GCTCCTTGCT GCAGTAATAT GCTTTGTGGC ATCTGGAGCA CGTTTTGGGG
      CGAGGAACGA CGTCATTATA CGAAACACCG TAGACCTCGT GCAAAACCCC
1601  CCTAAACAGC CACAAACCCT GCAGAGATGA GCACCAGACT TAAGCTGGAG
      GGATTTGTCG GTGTTTGGGA CGTCTCTACT CGTGGTCTGA ATTCGACCTC
1651  ACACACTGAT TCTCCTGTTT CTGGGGGAGG ATTCTCAGAA GGTGGCTCAT
      TGTGTGACTA AGAGGACAAA GACCCCCTCC TAAGAGTCTT CCACCGAGTA
1701  ATGAGTAAAA ATCGTTTTTC CTGGGTAGTT GATTCCTAAA AACTAAAAAA
      TACTCATTTT TAGCAAAAAG GACCCATCAA CTAAGGATTT TTGATTTTTT
1751  GAATACAGAG AAAAGTTTTA TCTTCAAACA AAACAGCAAT TCACATATTT
      CTTATGTCTC TTTTCAAAAT AGAAGTTTGT TTTGTCGTTA AGTGTATAAA
1801  TATCCTCTGC ACGTAAAACT GAAAATAACA ACAACAAAAA AGAAATGAAA
      ATAGGAGACG TGCATTTTGA CTTTTATTGT TGTTGTTTTT TCTTTACTTT
1851  GTTTTTGCTT TCAGGAATAA GCTTTTAAAA TCCAGAAACT AGATTTCGTC
      CAAAAACGAA AGTCCTTATT CGAAAATTTT AGGTCTTTGA TCTAAAGCAG
1901  CGGTACACGC AACTGAGTTG CCTCCTAGAG GTGGTTTGAG TTAATCAAAT
      GCCATGTGCG TTGACTCAAC GGAGGATCTC CACCAAACTC AATTAGTTTA
1951  TAATAAGACT GATCGTTAAG AACGACTGCC AAAAATACGA AAAAGCTACT
      ATTATTCTGA CTAGCAATTC TTGCTGACGG TTTTTATGCT TTTTCGATGA
2001  GGGATCCATC TTTCCAAGAC AATTTCTATT ATCTGAATTA ACACCATACC
      CCCTAGGTAG AAAGGTTCTG TTAAAGATAA TAGACTTAAT TGTGGTATGG
2051  TGGTACCCAC TGATTAAAAG CTGGGGGTTA CCAATGCGCG TGGGCACAGT
      ACCATGGGTG ACTAATTTTC GACCCCCAAT GGTTACGCGC ACCCGTGTCA
2101  TAGAAGCTTA TGTAGCAAAA ATGAGCACAT CCTGGAAGGG CCCGGGAGAA
      ATCTTCGAAT ACATCGTTTT TACTCGTGTA GGACCTTCCC GGGCCCTCTT
2151  GGTGCTCCTG GGGCAGCGCG GAGAGGGAGC TCTGAGGCTG GGGCGGCAGC
      CCACGAGGAC CCCGTCGCGC CTCTCCCTCG AGACTCCGAC CCCGCCGTCG
2201  GGTGCTTGCC GCCGTCCCCC TGGTCGCTCC CGGAATTAAC GCCGCGCACG
      CCACGAACGG CGGCAGGGGG ACCAGCGAGG GCCTTAATTG CGGCGCGTGC
2251  CGTCGGAGGC ATGGCCCCGT CCCGACCCCG TTTGGCGGCT CACCTCGCAG
      GCAGCCTCCG TACCGGGGCA GGGCTGGGGC AAACCGCCGA GTGGAGCGTC
2301  GCCGGCACAG CACGGCTGCT CGCGGCAGCA GAAGAGGAAG ATGCAGCGGT
      CGGCCGTGTC GTGCCGACGA GCGCCGTCGT CTTCTCCTTC TACGTCGCCA
2351  GGAAGGCGTC CGGGCGGCCA GGCAGCGGCG CATACACCTG CAGCAGGAAG
      CCTTCCGCAG GCCCGCCGGT CCGTCGCCGC GTATGTGGAC GTCGTCCTTC
2401  GAGAGCGGGC GGCCGCACAG CTCGCAGGCC AGGGCCTGGG GCCCCGGCAG
      CTCTCGCCCG CCGGCGTGTC GAGCGTCCGG TCCCGGACCC CGGGGCCGTC
2451  CCCGGCCGCG CCCAGCCATG CCGGCCGCCC GCCCACCTTG CTGGGGAACT
      GGGCCGGCGC GGGTCGGTAC GGCCGGCGGG CGGGTGGAAC GACCCCTTGA
2501  GCTCGCTGCG CAGTCGCCAC GCCGGCGCCG ACTCGGCGAA GCCCAGCTCC
      CGAGCGACGC GTCAGCGGTG CGGCCGCGGC TGAGCCGCTT CGGGTCGAGG
2551  ACAGGCCTGG CCCCGGCGGC AGCCATGCGG GGCGCGGGCT GGCGTGGGGC
      TGTCCGGACC GGGGCCGCCG TCGGTACGCC CCGCGCCCGA CCGCACCCCG
2601  GCAGCCCACA GCTGGGTCGG AAGGCGGAAA TCGGGCGCCG GGCCGGAAGG
      CGTCGGGTGT CGACCCAGCC TTCCGCCTTT AGCCCGCGGC CCGGCCTTCC
2651  CAAGAGGCGG GCACCTTTCC GGAGGACAGG AGGCGGAAAC GCGTCTGACG
      GTTCTCCGCC CGTGGAAAGG CCTCCTGTCC TCCGCCTTTG CGCAGACTGC
2701  GGAGCGGTTG CAGGACCAAT GCGAGGGAAC GGGGCAGAGG AAACCTCTCG
      CCTCGCCAAC GTCCTGGTTA CGCTCCCTTG CCCCGTCTCC TTTGGAGAGC
2751  GCATCAGCCC CGCCCCTGGC GCCTCTGCCT CCGAGCCGCT TTCCTGGTGC
      CGTAGTCGGG GCGGGGACCG CGGAGACGGA GGCTCGGCGA AAGGACCACG
```

FIG. 25(contd.) 61/93

```
2801  CTCCGGGTGC TCTGGGATGG TTCTGGTCTT TGGGAGAGTG GCAGCTGGTG
      GAGGCCCACG AGACCCTACC AAGACCAGAA ACCCTCTCAC CGTCGACCAC
2851  ACGGCGCTCC GCTCACCTCT GCACATGTCT TGCTGTGGGC CTGCGGGTGG
      TGCCGCGAGG CGAGTGGAGA CGTGTACAGA ACGACACCCG GACGCCCACC
2901  CCGCCAGGGA GGCAGAGCCC TCCCGCAAAC CTTCCCTGCT GGTGTCCACC
      GGCGGTCCCT CCGTCTCGGG AGGGCGTTTG GAAGGGACGA CCACAGGTGG
2951  TCAGGGTGTG GGAAACCTGT GCGCTGGCCG AGTGCTAACC AAGAGTAGGC
      AGTCCCACAC CCTTTGGACA CGCGACCGGC TCACGATTGG TTCTCATCCG
3001  AGTGAAAGAC AAATGAAGGT TGAACAGGTA AAGTGAGGAC CCTACAGCGG
      TCACTTTCTG TTTACTTCCA ACTTGTCCAT TTCACTCCTG GGATGTCGCC
3051  AAACCAAGAA TCCTGTGTGC CTGAGAGTAA TGAAGAAGCC TCTGCAGAAG
      TTTGGTTCTT AGGACACACG GACTCTCATT ACTTCTTCGG AGACGTCTTC
3101  AGTCTTTTCT GTCAGTCTTA AGGTCTCTGT TTTAATGTTA GTGCTGGCTT
      TCAGAAAAGA CAGTCAGAAT TCCAGAGACA AAATTACAAT CACGACCGAA
3151  GCTGTACCTG AATTCCAAGG GAGGAGTGTA TAATGAGGCA TGGCCAACCC
      CGACATGGAC TTAAGGTTCC CTCCTCACAT ATTACTCCGT ACCGGTTGGG
3201  CCACTTCCCA TCATTGCCTG AACTAGTTTT TCAGGTTAAC TTCAGAATGC
      GGTGAAGGGT AGTAACGGAC TTGATCAAAA AGTCCAATTG AAGTCTTACG
3251  CCTTGGTACC GCGGGCCCCC TCTGTGGTCC CACGCCACTG ATCGCTGCAT
      GGAACCATGG CGCCCGGGGG AGACACCAGG GTGCGGTGAC TAGCGACGTA
3301  GCCCACCACC TGGGTACACA CAGTCTGTGA TTCCCGGAGC AGAACGGACC
      CGGGTGGTGG ACCCATGTGT GTCAGACACT AAGGGCCTCG TCTTGCCTGG
3351  CTGCCCACCC GGTCTTGTGT GCTACTCAGT GGACAGACCC AAGGCAAGAA
      GACGGGTGGG CCAGAACACA CGATGAGTCA CCTGTCTGGG TTCCGTTCTT
3401  AGGGTGACAA GGACAGGGTC TTCCCAGGCT GGCTTTGAGT TCCTAGCACC
      TCCCACTGTT CCTGTCCCAG AAGGGTCCGA CCGAAACTCA AGGATCGTGG
3451  GCCCCGCCCC CAATCCTCTG TGGCACATGG AGTCTTGGTC CCCAGAGTCC
      CGGGGCGGGG GTTAGGAGAC ACCGTGTACC TCAGAACCAG GGGTCTCAGG
3501  CCCAGCGGCC TCCAGATGGT CTGGGAGGGC AGTTCAGCTG TGGCTGCGCA
      GGGTCGCCGG AGGTCTACCA GACCCTCCCG TCAAGTCGAC ACCGACGCGT
3551  TAGCAGACAT ACAACGGACG GTGGGCCCAG ACCCAGGCTG TGTAGACCCA
      ATCGTCTGTA TGTTGCCTGC CACCCGGGTC TGGGTCCGAC ACATCTGGGT
3601  GCCCCCCCGC CCCGCAGTGC CTAGGTCACC CACTAACGCC CCAGGCCTGG
      CGGGGGGGCG GGGCGTCACG GATCCAGTGG GTGATTGCGG GGTCCGGACC
3651  TCTTGGCTGG GCGTGACTGT TACCCTCAAA AGCAGGCAGC TCCAGGGTAA
      AGAACCGACC CGCACTGACA ATGGGAGTTT TCGTCCGTCG AGGTCCCATT
3701  AAGGTGCCCT GCCCTGTAGA GCCCACTTCC TTCCCAGGGC TGCGGCTGGG
      TTCCACGGGA CGGGACATCT CGGGTGAAGG AAGGGTCCCG ACGCCGACCC
3751  TAGGTTTGTA GCCTTCATCA CGGGCCACCT CCAGCCACTG GACCGCTGGC
      ATCCAAACAT CGGAAGTAGT GCCCGGTGGA GGTCGGTGAC CTGGCGACCG
3801  CCCTGCCCTG TCCTGGGGAG TGTGGTCCTG CGACTCTAAT GGCCGCAAGC
      GGGACGGGAC AGGACCCCTC ACACCAGGAC GCTGAGATTA CCGGCGTTCG
3851  CACCTGACTC CCCCAACACC ACACTCTACC TCTCAAGCCC AGGTCTCTCC
      GTGGACTGAG GGGGTTGTGG TGTGAGATGG AGAGTTCGGG TCCAGAGAGG
3901  CTAGTGACCC ACCCAGCACA TTTAGCTAGC TGAGCCCCAC AGCCAGAGGT
      GATCACTGGG TGGGTCGTGT AAATCGATCG ACTCGGGGTG TCGGTCTCCA
3951  CCTCAGGCCC TGCTTTCAGG GCAGTTGCTC TGAAGTCGGC AAGGGGGAGT
      GGAGTCCGGG ACGAAAGTCC CGTCAACGAG ACTTCAGCCG TTCCCCCTCA
4001  GACTGCCTGG CCACTCCATG CCCTCCAAGA GCTCCTTCTG CAGGAGCGTA
      CTGACGGACC GGTGAGGTAC GGGAGGTTCT CGAGGAAGAC GTCCTCGCAT
4051  CAGAACCCAG GGCCCTGGCA CCCGTGCAGA CCCTGGCCCA CCCCACCTGG
      GTCTTGGGTC CCGGGACCGT GGGCACGTCT GGGACCGGGT GGGGTGGACC
4101  GCGCTCAGTG CCCAAGAGAT GTCCACACCT AGGATGTCCC GCGGTGGGTG
      CGCGAGTCAC GGGTTCTCTA CAGGTGTGGA TCCTACAGGG CGCCACCCAC
4151  GGGGGCCCGA GAGACGGGCA GGCCGGGGGC AGGCCTGGCC ATGCGGGGCC
      CCCCCGGGCT CTCTGCCCGT CCGGCCCCCG TCCGGACCGG TACGCCCCGG
4201  GAACCGGGCA CTGCCCAGCG TGGGGCGCGG GGGCCACGGC GCGCGCCCCC
      CTTGGCCCGT GACGGGTCGC ACCCCGCGCC CCCGGTGCCG CGCGCGGGGG
```

FIG. 25(contd.) 62/93

```
4251  AGCCCCCGGG CCCAGCACCC CAAGGCGGCC AACGCCAAAA CTCTCCCTCC
      TCGGGGGCCC GGGTCGTGGG GTTCCGCCGG TTGCGGTTTT GAGAGGGAGG
4301  TCCTCTTCCT CAATCTCGCT CTCGCTCTTT TTTTTTTCG CAAAAGGAGG
      AGGAGAAGGA GTTAGAGCGA GAGCGAGAAA AAAAAAAAGC GTTTTCCTCC
4351  GGAGAGGGGG TAAAAAAATG CTGCACTGTG CGGCGAAGCC GGTGAGTGAG
      CCTCTCCCCC ATTTTTTTAC GACGTGACAC GCCGCTTCGG CCACTCACTC
4401  CGGCGCGGGG CCAATCAGCG TGCGCCGTTC CGAAAGTTGC CTTTTATGGC
      GCCGCGCCCC GGTTAGTCGC ACGCGGCAAG GCTTTCAACG GAAAATACCG
4451  TCGAGCGGCC GCGGCGGCGC CCTATAAAAC CCAGCGGCGC GACGCGCCAC
      AGCTCGCCGG CGCCGCCGCG GGATATTTTG GGTCGCCGCG CTGCGCGGTG
4501  CACCGCCGAG ACCGCGTCCG CCCGCGAGCA CAGAGCCTCG CCTTTGCCGA
      GTGGCGGCTC TGGCGCAGGC GGGCGCTCGT GTCTCGGAGC GGAAACGGCT
4551  TCCGCCGCCC GTCCACACCC GCCGCCAGGT AAGCCCGGCC AGCCGACCGG
      AGGCGGCGGG CAGGTGTGGG CGGCGGTCCA TTCGGGCCGG TCGGCTGGCC
4601  GGCATGCGGC CGCGGCCCTT CGCCCGTGCA GAGCCGCCGT CTGGGCCGCA
      CCGTACGCCG GCGCCGGGAA GCGGGCACGT CTCGGCGGCA GACCCGGCGT
4651  GCGGGGGGCG CATGGGGCGG AACCGGACCG CCGTGGGGGG CGCGGGAGAA
      CGCCCCCCGC GTACCCCGCC TTGGCCTGGC GGCACCCCCC GCGCCCTCTT
4701  GCCCCTGGGC CTCCGGAGAT GGGGGACACC CCACGCCAGT TCGCAGGCGC
      CGGGGACCCG GAGGCCTCTA CCCCCTGTGG GGTGCGGTCA AGCGTCCGCG
4751  GAGGCCGCGC TCGGGCGGGC GCGCTCCGGG GGTGCCGCTC TCGGGGCGGG
      CTCCGGCGCG AGCCCGCCCG CGCGAGGCCC CCACGGCGAG AGCCCCGCCC
4801  GGCAACCGGC GGGGTCTTTG TCTGAGCCGG GCTCTTGCCA ATGGGGATCG
      CCGTTGGCCG CCCCAGAAAC AGACTCGGCC CGAGAACGGT TACCCCTAGC
4851  CACGGTGGGC GCGGCGTAGC CCCCGTCAGG CCCGGTGGGG GCTGGGGCGC
      GTGCCACCCG CGCCGCATCG GGGGCAGTCC GGGCCACCCC CGACCCCGCG
4901  CATGCGCGTG CGCGCTGGTC CTTTGGGCGC TAACTGCGTG CGCGCTGGGA
      GTACGCGCAC GCGCGACCAG GAAACCCGCG ATTGACGCAC GCGCGACCCT
4951  ATTGGCGCTA ATTGCGCGTG CGCGCTGGGA CTCAATGGCG CTAATCGCGC
      TAACCGCGAT TAACGCGCAC GCGCGACCCT GAGTTACCGC GATTAGCGCG
5001  GTGCGTTCTG GGGCCCGGGC GCTTGCGCCA CTTCCTGCCC GAGCCGCTGG
      CACGCAAGAC CCCGGGCCCG CGAACGCGGT GAAGGACGGG CTCGGCGACC
5051  CGCCCGAGGG TGTGGCCGCT GCGTGCGCGC GCGCGACCCG GTCGCTGTTT
      GCGGGCTCCC ACACCGGCGA CGCACGCGCG CGCGCTGGGC CAGCGACAAA
5101  GAACCGGGCG GAGGCGGGGC TGGCGCCCGG TTGGGAGGGG GTTGGGGCCT
      CTTGGCCCGC CTCCGCCCCG ACCGCGGGCC AACCCTCCCC CAACCCCGGA
5151  GGCTTCCTGC CGCGCGCCGC GGGGACGCCT CCGACCAGTG TTTGCCTTTT
      CCGAAGGACG GCGCGCGGCG CCCCTGCGGA GGCTGGTCAC AAACGGAAAA
5201  ATGGTAATAA CGCGGCCGGC CCGGCTTCCT TTGTCCCCAA TCTGGGCGCG
      TACCATTATT GCGCCGGCCG GGCCGAAGGA AACAGGGGTT AGACCCGCGC
5251  CGCCGGCGCC CCCTGGCGGC CTAAGGACTC GGCGCGCCGG AAGTGGCCAG
      GCGGCCGCGG GGGACCGCCG GATTCCTGAG CCGCGCGGCC TTCACCGGTC
5301  GGCGGGGGCG ACTTCGGCTC ACAGCGCGCC CGGCTATTCT CGCAGCTCAC
      CCGCCCCCGC TGAAGCCGAG TGTCGCGCGG GCCGATAAGA GCGTCGAGTG
5351  CATGCCGGTC GCCACCATGA GCTTATCGAT ACCGGTGGCG CGCCAATTGT
      GTACGGCCAG CGGTGGTACT CGAATAGCTA TGGCCACCGC GCGGTTAACA
5401  TAATTAAGAT CTGGCCCAAT GGGCCGTACG AATTCCTTAG GCTACCGGGT
      ATTAATTCTA GACCGGGTTA CCCGGCATGC TTAAGGAATC CGATGGCCCA
5451  AGGGGAGGCG CTTTTCCCAA GGCAGTCTGG AGCATGCGCT TTAGCAGCCC
      TCCCCTCCGC GAAAAGGGTT CCGTCAGACC TCGTACGCGA AATCGTCGGG
5501  CGCTGGGCAC TTGGCGCTAC ACAAGTGGCC TCTGGCCTCG CACACATTCC
      GCGACCCGTG AACCGCGATG TGTTCACCGG AGACCGGAGC GTGTGTAAGG
5551  ACATCCACCG GCCGGTAGGC GCCAACCGGC TCCGTTCTTT GGTGGCCCCT
      TGTAGGTGGC CGGCCATCCG CGGTTGGCCG AGGCAAGAAA CCACCGGGGA
5601  TCGCGCCACC TTCTACTCCT CCCCTAGTCA GGAAGTTCCC CCCCGCCCCG
      AGCGCGGTGG AAGATGAGGA GGGGATCAGT CCTTCAAGGG GGGGCGGGGC
5651  CAGCTCGCGT CGTGCAGGAC GTGACAAATG GAAGTAGCAC GTCTCACTAG
      GTCGAGCGCA GCACGTCCTG CACTGTTTAC CTTCATCGTG CAGAGTGATC
```

FIG. 25(contd.) 63/93

```
5701  TCTCGTGCAG ATGGACAGCA CCGCTGAGCA ATGGAAGCGG GTAGGCCTTT
      AGAGCACGTC TACCTGTCGT GGCGACTCGT TACCTTCGCC CATCCGGAAA
5751  GGGGCAGCGG CCAATAGCAG CTTTGCTCCT TCGCTTTCTG GGCTCAGAGG
      CCCCGTCGCC GGTTATCGTC GAAACGAGGA AGCGAAAGAC CCGAGTCTCC
5801  CTGGGAAGGG GTGGGTCCGG GGGCGGGCTC AGGGGCGGGC TCAGGGGCGG
      GACCCTTCCC CACCCAGGCC CCCGCCCGAG TCCCCGCCCG AGTCCCCGCC
5851  GGCGGGCGCC CGAAGGTCCT CCGGAGGCCC GGCATTCTGC ACGCTTCAAA
      CCGCCCGCGG GCTTCCAGGA GGCCTCCGGG CCGTAAGACG TGCGAAGTTT
5901  AGCGCACGTC TGCCGCGCTG TTCTCCTCTT CCTCATCTCC GGGCCTTTCG
      TCGCGTGCAG ACGGCGCGAC AAGAGGAGAA GGAGTAGAGG CCCGGAAAGC
5951  ACCAGCTTAC CATGACCGAG TACAAGCCCA CGGTGCGCCT CGCCACCCGC
      TGGTCGAATG GTACTGGCTC ATGTTCGGGT GCCACGCGGA GCGGTGGGCG
6001  GACGACGTCC CCAGGGCCGT ACGCACCCTC GCCGCCGCGT TCGCCGACTA
      CTGCTGCAGG GGTCCCGGCA TGCGTGGGAG CGGCGGCGCA AGCGGCTGAT
6051  CCCCGCCACG CGCCACACCG TCGATCCGGA CCGCCACATC GAGCGGGTCA
      GGGGCGGTGC GCGGTGTGGC AGCTAGGCCT GGCGGTGTAG CTCGCCCAGT
6101  CCGAGCTGCA AGAACTCTTC CTCACGCGCG TCGGGCTCGA CATCGGCAAG
      GGCTCGACGT TCTTGAGAAG GAGTGCGCGC AGCCCGAGCT GTAGCCGTTC
6151  GTGTGGGTCG CGGACGACGG CGCCGCGGTG GCGGTCTGGA CCACGCCGGA
      CACACCCAGC GCCTGCTGCC GCGGCGCCAC CGCCAGACCT GGTGCGGCCT
6201  GAGCGTCGAA GCGGGGGCGG TGTTCGCCGA GATCGGCCCG CGCATGGCCG
      CTCGCAGCTT CGCCCCCGCC ACAAGCGGCT CTAGCCGGGC GCGTACCGGC
6251  AGTTGAGCGG TTCCCGGCTG GCCGCGCAGA ACAGATGGAA GGCCTCCTGG
      TCAACTCGCC AAGGGCCGAC CGGCGCGTCT TGTCTACCTT CCGGAGGACC
6301  CGCCGCACCG GCCCAAGGAG CCCGCGTGGT TCCTGGCCAC CGTCGCGTCT
      GCGGCGTGGC CGGGTTCCTC GGGCGCACCA AGGACCGGTG GCAGCGCAGA
6351  CGCCCGACCA CCAGGGCAAG GGTCTGGGCA GCGCCGTCGT GCTCCCCGGA
      GCGGGCTGGT GGTCCCGTTC CCAGACCCGT CGCGGCAGCA CGAGGGGCCT
6401  GTGGAGGCGG CCGAGCGCGC CGGGGTGCCC GCCTTCCTGG AGACCTCCGC
      CACCTCCGCC GGCTCGCGCG GCCCCACGGG CGGAAGGACC TCTGGAGGCG
6451  GCCCCGCAAC CTCCCCTTCT ACGAGCGGCT CGGCTTCACC GTCACCGCCG
      CGGGGCGTTG GAGGGGAAGA TGCTCGCCGA GCCGAAGTGG CAGTGGCGGC
6501  ACGTCGAGGT GCCCGAAGGA CCGCGCACCT GGTGCATGAC CCGCAAGCCC
      TGCAGCTCCA CGGGCTTCCT GGCGCGTGGA CCACGTACTG GGCGTTCGGG
6551  GGTGCCTGAC GCCCGCCCCA CGACCCGCAG CGCCCGACCG AAAGGAGCGC
      CCACGGACTG CGGGCGGGGT GCTGGGCGTC GCGGGCTGGC TTTCCTCGCG
6601  ACGACCCCAT GCATCGTAGA GCTCGCTGAT CAGCCTCGAC TGTGCCTTCT
      TGCTGGGGTA CGTAGCATCT CGAGCGACTA GTCGGAGCTG ACACGGAAGA
6651  AGTTGCCAGC CATCTGTTGT TTGCCCCTCC CCCGTGCCTT CCTTGACCCT
      TCAACGGTCG GTAGACAACA AACGGGGAGG GGGCACGGAA GGAACTGGGA
6701  GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAATGAG GAAATTGCAT
      CCTTCCACGG TGAGGGTGAC AGGAAAGGAT TATTTTACTC CTTTAACGTA
6751  CGCATTGTCT GAGTAGGTGT CATTCTATTC TGGGGGGTGG GGTGGGGCAG
      GCGTAACAGA CTCATCCACA GTAAGATAAG ACCCCCCACC CCACCCCGTC
6801  GACAGCAAGG GGGGGGATTG GGRAGACAAT AGCAGGCATG CTGGGGGGGC
      CTGTCGTTCC CCCCCCTAAC CCYTCTGTTA TCGTCCGTAC GACCCCCCCG
6851  GGTGGGGGCT ATGGCTTCTG AGGCGGAAAG AACCAGCTGG GGCTCGAGGG
      CCACCCCCGA TACCGAAGAC TCCGCCTTTC TTGGTCGACC CCGAGCTCCC
6901  CCGCCACCGC GGTGGAGCTC CAGCTTTTGT TCCCTTTAGT GAGGGTTAAT
      GGCGGTGGCG CCACCTCGAG GTCGAAAACA AGGGAAATCA CTCCCAATTA
6951  TTCGAGCTTG GCGTAATCAT GGTCATAGCT GTTTCCTGTG TGAAATTGTT
      AAGCTCGAAC CGCATTAGTA CCAGTATCGA CAAAGGACAC ACTTTAACAA
7001  ATCCGCTCAC AATTCCACAC AACATACGAG CCGGAAGCAT AAAGTGTAAA
      TAGGCGAGTG TTAAGGTGTG TTGTATGCTC GGCCTTCGTA TTTCACATTT
7051  GCCTGGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG CGTTGCGCTC
      CGGACCCCAC GGATTACTCA CTCGATTGAG TGTAATTAAC GCAACGCGAG
7101  ACTGCCCGCT TTCCAGTCGG GAAACCTGTC GTGCCAGCAT CGCGAGCACT
      TGACGGGCGA AAGGTCAGCC CTTTGGACAG CACGGTCGTA GCGCTCGTGA
```

FIG. 25(contd.) 64/93

```
7151  TTTCGGGGAA ATGTGCGCGG AACCCCTATT TGTTTATTTT TCTAAATACA
      AAAGCCCCTT TACACGCGCC TTGGGGATAA ACAAATAAAA AGATTTATGT
7201  TTCAAATATG TATCCGCTCA TGAGACAATA ACCCTGATAA ATGCTTCAAT
      AAGTTTATAC ATAGGCGAGT ACTCTGTTAT TGGGACTATT TACGAAGTTA
7251  AATATTGAAA AAGGAAGAGT ATGAGTATTC AACATTTCCG TGTCGCCCTT
      TTATAACTTT TTCCTTCTCA TACTCATAAG TTGTAAAGGC ACAGCGGGAA
7301  ATTCCCTTTT TTGCGGCATT TTGCCTTCCT GTTTTTGCTC ACCCAGAAAC
      TAAGGGAAAA AACGCCGTAA AACGGAAGGA CAAAAACGAG TGGGTCTTTG
7351  GCTGGTGAAA GTAAAAGATG CTGAAGATCA GTTGGGTGCA CGAGTGGGTT
      CGACCACTTT CATTTTCTAC GACTTCTAGT CAACCCACGT GCTCACCCAA
7401  ACATCGAACT GGATCTCAAC AGCGGTAAGA TCCTTGAGAG TTTTCGCCCC
      TGTAGCTTGA CCTAGAGTTG TCGCCATTCT AGGAACTCTC AAAAGCGGGG
7451  GAAGAACGTT TTCCAATGAT GAGCACTTTT AAAGTTCTGC TATGTGGCGC
      CTTCTTGCAA AAGGTTACTA CTCGTGAAAA TTTCAAGACG ATACACCGCG
7501  GGTATTATCC CGTATTGACG CCGGGCAAGA GCAACTCGGT CGCCGCATAC
      CCATAATAGG GCATAACTGC GGCCCGTTCT CGTTGAGCCA GCGGCGTATG
7551  ACTATTCTCA GAATGACTTG GTTGAGTACT CACCAGTCAC AGAAAAGCAT
      TGATAAGAGT CTTACTGAAC CAACTCATGA GTGGTCAGTG TCTTTTCGTA
7601  CTTACGGATG GCATGACAGT AAGAGAATTA TGCAGTGCTG CCATAACCAT
      GAATGCCTAC CGTACTGTCA TTCTCTTAAT ACGTCACGAC GGTATTGGTA
7651  GAGTGATAAC ACTGCGGCCA ACTTACTTCT GACAACGATC GGAGGACCGA
      CTCACTATTG TGACGCCGGT TGAATGAAGA CTGTTGCTAG CCTCCTGGCT
7701  AGGAGCTAAC CGCTTTTTTG CACAACATGG GGGATCATGT AACTCGCCTT
      TCCTCGATTG GCGAAAAAAC GTGTTGTACC CCCTAGTACA TTGAGCGGAA
7751  GATCGTTGGG AACCGGAGCT GAATGAAGCC ATACCAAACG ACGAGCGTGA
      CTAGCAACCC TTGGCCTCGA CTTACTTCGG TATGGTTTGC TGCTCGCACT
7801  CACCACGATG CCTGTAGCAA TGGCAACAAC GTTGCGCAAA CTATTAACTG
      GTGGTGCTAC GGACATCGTT ACCGTTGTTG CAACGCGTTT GATAATTGAC
7851  GCGAACTACT TACTCTAGCT TCCCGGCAAC AATTAATAGA CTGGATGGAG
      CGCTTGATGA ATGAGATCGA AGGGCCGTTG TTAATTATCT GACCTACCTC
7901  GCGGATAAAG TTGCAGGACC ACTTCTGCGC TCGGCCCTTC CGGCTGGCTG
      CGCCTATTTC AACGTCCTGG TGAAGACGCG AGCCGGGAAG GCCGACCGAC
7951  GTTTATTGCT GATAAATCTG GAGCCGGTGA GCGTGGGTCT CGCGGTATCA
      CAAATAACGA CTATTTAGAC CTCGGCCACT CGCACCCAGA GCGCCATAGT
8001  TTGCAGCACT GGGGCCAGAT GGTAAGCCCT CCCGTATCGT AGTTATCTAC
      AACGTCGTGA CCCCGGTCTA CCATTCGGGA GGGCATAGCA TCAATAGATG
8051  ACGACGGGGA GTCAGGCAAC TATGGATGAA CGAAATAGAC AGATCGCTGA
      TGCTGCCCCT CAGTCCGTTG ATACCTACTT GCTTTATCTG TCTAGCGACT
8101  GATAGGTGCC TCACTGATTA AGCATTGGTA ACTGTCAGAC TCGCGACACT
      CTATCCACGG AGTGACTAAT TCGTAACCAT TGACAGTCTG AGCGCTGTGA
8151  GCATTAATGA ATCGGCCAAC GCGCGGGGAG AGGCGGTTTG CGTATTGGGC
      CGTAATTACT TAGCCGGTTG CGCGCCCCTC TCCGCCAAAC GCATAACCCG
8201  GCTCTTCCGC TTCCTCGCTC ACTGACTCGC TGCGCTCGGT CGTTCGGCTG
      CGAGAAGGCG AAGGAGCGAG TGACTGAGCG ACGCGAGCCA GCAAGCCGAC
8251  CGGCGAGCGG TATCAGCTCA CTCAAAGGCG GTAATACGGT TATCCACAGA
      GCCGCTCGCC ATAGTCGAGT GAGTTTCCGC CATTATGCCA ATAGGTGTCT
8301  ATCAGGGGAT AACGCAGGAA AGAACATGTG AGCAAAAGGC CAGCAAAAGG
      TAGTCCCCTA TTGCGTCCTT TCTTGTACAC TCGTTTTCCG GTCGTTTTCC
8351  CCAGGAACCG TAAAAAGGCC GCGTTGCTGG CGTTTTTCCA TAGGCTCCGC
      GGTCCTTGGC ATTTTTCCGG CGCAACGACC GCAAAAAGGT ATCCGAGGCG
8401  CCCCCTGACG AGCATCACAA AAATCGACGC TCAAGTCAGA GGTGGCGAAA
      GGGGGACTGC TCGTAGTGTT TTTAGCTGCG AGTTCAGTCT CCACCGCTTT
8451  CCCGACAGGA CTATAAAGAT ACCAGGCGTT TCCCCCTGGA AGCTCCCTCG
      GGGCTGTCCT GATATTTCTA TGGTCCGCAA AGGGGGACCT TCGAGGGAGC
8501  TGCGCTCTCC TGTTCCGACC CTGCCGCTTA CCGGATACCT GTCCGCCTTT
      ACGCGAGAGG ACAAGGCTGG GACGGCGAAT GGCCTATGGA CAGGCGGAAA
8551  CTCCCTTCGG GAAGCGTGGC GCTTTCTCAT AGCTCACGCT GTAGGTATCT
      GAGGGAAGCC CTTCGCACCG CGAAAGAGTA TCGAGTGCGA CATCCATAGA
```

```
8601    CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG CACGAACCCC
        GTCAAGCCAC ATCCAGCAAG CGAGGTTCGA CCCGACACAC GTGCTTGGGG
8651    CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACTATCG TCTTGAGTCC
        GGCAAGTCGG GCTGGCGACG CGGAATAGGC CATTGATAGC AGAACTCAGG
8701    AACCCGGTAA GACACGACTT ATCGCCACTG GCAGCAGCCA CTGGTAACAG
        TTGGGCCATT CTGTGCTGAA TAGCGGTGAC CGTCGTCGGT GACCATTGTC
8751    GATTAGCAGA GCGAGGTATG TAGGCGGTGC TACAGAGTTC TTGAAGTGGT
        CTAATCGTCT CGCTCCATAC ATCCGCCACG ATGTCTCAAG AACTTCACCA
8801    GGCCTAACTA CGGCTACACT AGAAGGACAG TATTTGGTAT CTGCGCTCTG
        CCGGATTGAT GCCGATGTGA TCTTCCTGTC ATAAACCATA GACGCGAGAC
8851    CTGAAGCCAG TTACCTTCGG AAAAAGAGTT GGTAGCTCTT GATCCGGCAA
        GACTTCGGTC AATGGAAGCC TTTTTCTCAA CCATCGAGAA CTAGGCCGTT
8901    ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTTTGCAAG CAGCAGATTA
        TGTTTGGTGG CGACCATCGC CACCAAAAAA ACAAACGTTC GTCGTCTAAT
8951    CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
        GCGCGTCTTT TTTTCCTAGA GTTCTTCTAG GAAACTAGAA AAGATGCCCC
9001    TCTGACGCTC AGTGGAACGA AAACTCA
        AGACTGCGAG TCACCTTGCT TTTGAGT
```

FIG. 25(contd.)

FIG. 26

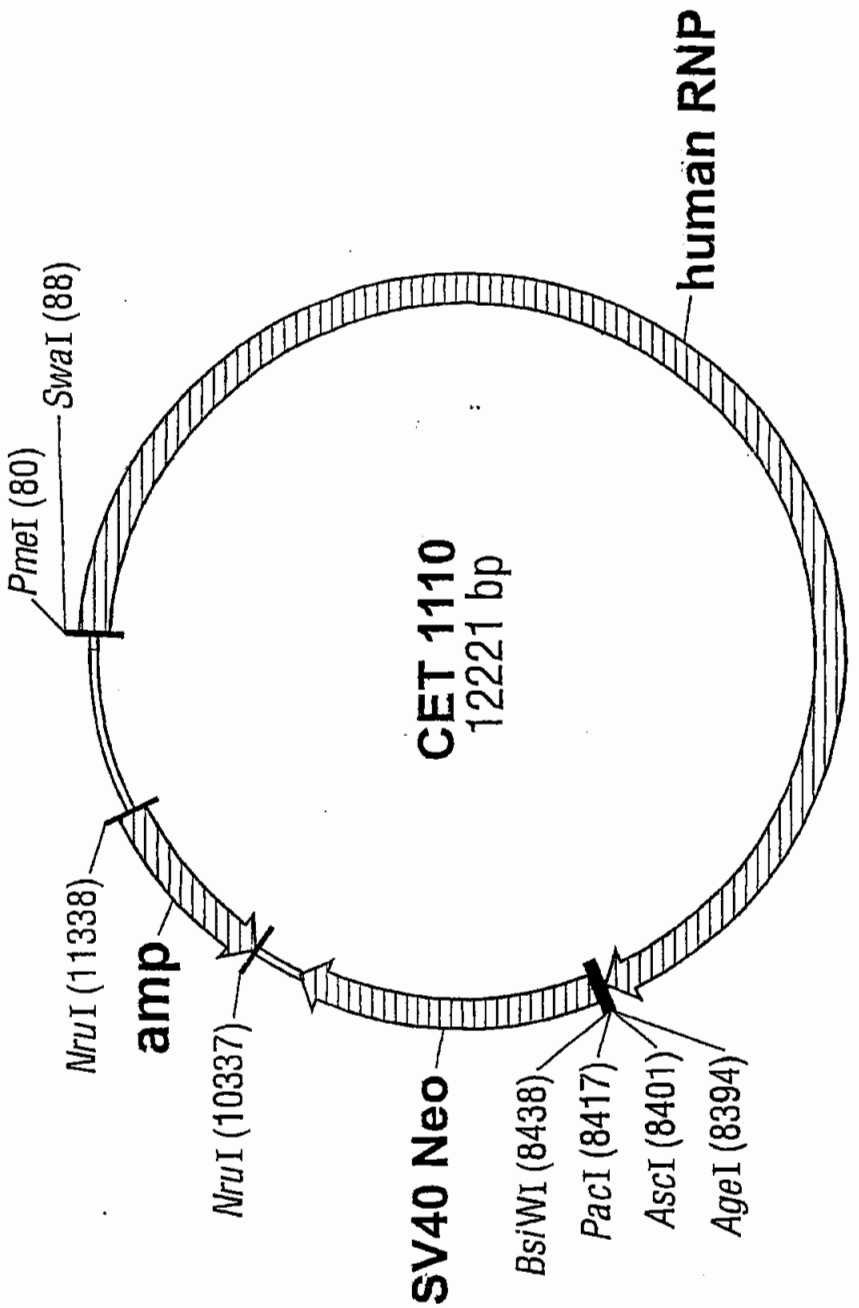

human RNP
SwaI (88)
PmeI (80)
CET 1110
12221 bp
NruI (11338)
amp
NruI (10337)
SV40 Neo
BsiWI (8438)
PacI (8417)
AscI (8401)
AgeI (8394)

FIG. 27 CET1110 nucleotide sequence

```
   1  CGTTGTAAAA CGACGGCCAG TGAATTGTAA TACGACTCAC TATAGGGCGA
      GCAACATTTT GCTGCCGGTC ACTTAACATT ATGCTGAGTG ATATCCCGCT
  51  ATTGGGTACC GGGCCCCCCC TCGAAGTTTA AACATTTAAA TCTAGAAGCT
      TAACCCATGG CCCGGGGGGG AGCTTCAAAT TTGTAAATTT AGATCTTCGA
 101  TCAATGTTTT TAGCACCCTC TGTGTGGAGG AAAATAATGC AGATTATTCT
      AGTTACAAAA ATCGTGGGAG ACACACCTCC TTTTATTACG TCTAATAAGA
 151  AATTAGTGTA ATATCTAACC ACATTAAAAT ATATTACATA GTAAACTACA
      TTAATCACAT TATAGATTGG TGTAATTTTA TATAATGTAT CATTTGATGT
 201  CTCCATAATT TTATAAATTT GACTCCCCAG GGTAATAAAC TAGTCTCTAG
      GAGGTATTAA AATATTTAAA CTGAGGGGTC CCATTATTTG ATCAGAGATC
 251  TCTGCTCACC TTCAACTGTA CAATAAAGTC TTGGTTCTTT TGAAATAGAC
      AGACGAGTGG AAGTTGACAT GTTATTTCAG AACCAAGAAA ACTTTATCTG
 301  CTCAAATGAG ACACCTAAAA TTCAAAGTGT CTTTACATTT AAAGACACCT
      GAGTTTACTC TGTGGATTTT AAGTTTCACA GAAATGTAAA TTTCTGTGGA
 351  ACAGGAAAGC AGGTAAAAGA GCCAGGTTAA AAACAAATTC TAAAACCACT
      TGTCCTTTCG TCCATTTTCT CGGTCCAATT TTTGTTTAAG ATTTTGGTGA
 401  TAGCTGCAGT TAAACATATA GTAAAGATGC ACTAAAGTTT CTTACTCTGT
      ATCGACGTCA ATTTGTATAT CATTTCTACG TGATTTCAAA GAATGAGACA
 451  AAATCCCTTC CACTTCAGGA AATATTCCAC TTTCCCATTC ACTACACGTC
      TTTAGGGAAG GTGAAGTCCT TTATAAGGTG AAAGGGTAAG TGATGTGCAG
 501  GATCTAGTAC TTTTTCCACG ACAAATTCTT CAGGCTCTGC CTCTTCAACT
      CTAGATCATG AAAAAGGTGC TGTTTAAGAA GTCCGAGACG GAGAAGTTGA
 551  TTTTTACTCT TTCCATTCTG TTTTTTTCCC ATTTTTTGCT AAAATAAAAC
      AAAAATGAGA AAGGTAAGAC AAAAAAAGGG TAAAAAACGA TTTTATTTTG
 601  AAAAGAGAAA TTAAGAAATA TTCCTCTTGA ATTTTGAGCA CATTTTCAAG
      TTTTCTCTTT AATTCTTTAT AAGGAGAACT TAAAACTCGT GTAAAAGTTC
 651  GCTCAATTGC TTATATTATT ATCACATTCG ACATAAATTT TTACTTCTAT
      CGAGTTAACG AATATAATAA TAGTGTAAGC TGTATTTAAA AATGAAGATA
 701  ATCCCAGGGC AGACACCTTC TGGAAAGATT AAAAGTCAAC AGACAATAAA
      TAGGGTCCCG TCTGTGGAAG ACCTTTCTAA TTTTCAGTTG TCTGTTATTT
 751  ATAAAAGAAT GCTTTATCTT GTTCATTTAG TTCAAACTTA CAACCCACCA
      TATTTTCTTA CGAAATAGAA CAAGTAAATC AAGTTTGAAT GTTGGGTGGT
 801  CCAAAATAAT ACAATAAAAA AACACTATCT GGAAACAGTT ATTTTTTTCC
      GGTTTTATTA TGTTATTTTT TTGTGATAGA CCTTTGTCAA TAAAAAAAGG
 851  AGTCTTTTTT TTTGAGACAG GGTCTCACAC TCTTGTCGCC CAGGCTGGAG
      TCAGAAAAAA AAACTCTGTC CCAGAGTGTG AGAACAGCGG GTCCGACCTC
 901  TGCAGTGGCG TGATCTCAGC TCACTGCAAC CTCCGCCTCC CCAGGTTCAA
      ACGTCACCGC ACTAGAGTCG AGTGACGTTG GAGGCGGAGG GGTCCAAGTT
 951  GCAGTTCTCA TGCCTCAGCC TCCAGAGTAG CTGGGATTAT AGGCGGATGC
      CGTCAAGAGT ACGGAGTCGG AGGTCTCATC GACCCTAATA TCCGCCTACG
1001  CACCATGCCG GGCTAATTTT TTTTGTGTTT TTATTAGAAA CAGGGTTTCA
      GTGGTACGGC CCGATTAAAA AAAACACAAA AATAATCTTT GTCCCAAAGT
1051  CCATGTTGAC CAGGCTGGTC TCAAACTCCT GACCTGAAGT GATTCACCAG
      GGTACAACTG GTCCGACCAG AGTTTGAGGA CTGGACTTCA CTAAGTGGTC
1101  CCTGGGCCTC CCAAAGTGCT GGCATTACAG GCGTGAGCCA CTGCGCCCGG
      GGACCCGGAG GGTTTCACGA CCGTAATGTC CGCACTCGGT GACGCGGGCC
1151  CCCTGTAGTC TTAAAAGACC AAGTTTACTA ATTTTCACTC ATTTTAACAA
      GGGACATCAG AATTTTCTGG TTCAAATGAT TAAAAGTGAG TAAAATTGTT
1201  CACTGCAACA AACAACTATG CAGGAAGTAC CTAAAGGGTG ATCCAGAGAA
      GTGACGTTGT TTGTTGATAC GTCCTTCATG GATTTCCCAC TAGGTCTCTT
1251  GCAAGTAGTA GTGACAGGTC TTAGGTGAAC CTATGACAGA CCTTGTATCC
      CGTTCATCAT CACTGTCCAG AATCCACTTG GATACTGTCT GGAACATAGG
1301  ACCCCCAGAT GGTAAAAGCC CCAGCCCCCT TCTCAATTCA AATATTAATG
      TGGGGGTCTA CCATTTTCGG GGTCGGGGGA AGAGTTAAGT TTATAATTAC
```

FIG. 27(contd.)

```
1351  TCAAAAGCAT CAATGATACA GAGAAAAGAT AAATGCAGAA TGAAAACATG
      AGTTTTCGTA GTTACTATGT CTCTTTTCTA TTTACGTCTT ACTTTTGTAC
1401  GTTCAAAATC CTGATACCAA CTGCAGGGTC AACTATAGAG ACCACTAGGA
      CAAGTTTTAG GACTATGGTT GACGTCCCAG TTGATATCTC TGGTGATCCT
1451  GGTTCAATTA AAGGACAAGA TTATTTTTCC ATAATCTCTG TAGATAATAT
      CCAAGTTAAT TTCCTGTTCT AATAAAAAGG TATTAGAGAC ATCTATTATA
1501  TTCCTACCAC TTAGAACAAA ACTATAAAGC TATCACTTCA AGAGACCAAC
      AAGGATGGTG AATCTTGTTT TGATATTTCG ATAGTGAAGT TCTCTGGTTG
1551  ATTACAAATT TATTTTAATT CCCTAAGGTG AAAAAAATCC TTCCTTCCTG
      TAATGTTTAA ATAAAATTAA GGGATTCCAC TTTTTTTAGG AAGGAAGGAC
1601  GTTTCTCAAG AGAAAGTCTA TACTGGTAAC CAAATTCACT TTAAACAGGC
      CAAAGAGTTC TCTTTCAGAT ATGACCATTG GTTTAAGTGA AATTTGTCCG
1651  ATTTTCTTTG GTATGACACT ATTTAAGAGA AGCAGGAAAC CAACGTGAAC
      TAAAAGAAAC CATACTGTGA TAAATTCTCT TCGTCCTTTG GTTGCACTTG
1701  CAGCTCTTTC CAATGGCTCA AGATTTCCTA TGAGAGGACT AAAAATGGGG
      GTCGAGAAAG GTTACCGAGT TCTAAAGGAT ACTCTCCTGA TTTTTACCCC
1751  AAAATTTTTA TGAGAGGATT AAAAATGGGG GAAAAAAAAC CCTGAAATGG
      TTTTAAAAAT ACTCTCCTAA TTTTTACCCC CTTTTTTTTG GGACTTTACC
1801  TTAATCAGAA GATCCTATGG GCTGAGAAGG AATCCATCTT AACATTTCAT
      AATTAGTCTT CTAGGATACC CGACTCTTCC TTAGGTAGAA TTGTAAAGTA
1851  CTTAAAGCAA ATGCTATTGC CGGGGGCAGT GGCTCATGCC TGTAATCCCA
      GAATTTCGTT TACGATAACG GCCCCCGTCA CCGAGTACGG ACATTAGGGT
1901  GCACTTTGGG AGGCCGAGGT GGGCAGATCA TCTGAGGTCA GGAGTTTGAG
      CGTGAAACCC TCCGGCTCCA CCCGTCTAGT AGACTCCAGT CCTCAAACTC
1951  ACCAGCCTGA CCAACATGGA GAAACCCCGT TTCTACTAAA AATACAAAAT
      TGGTCGGACT GGTTGTACCT CTTTGGGGCA AAGATGATTT TTATGTTTTA
2001  TAGCCAGGCA TAGTGGTGCA TGCCTGTAAT CCCAGCTACT TGGGAGGCTG
      ATCGGTCCGT ATCACCACGT ACGGACATTA GGGTCGATGA ACCCTCCGAC
2051  AGGCAGGAGA ACTGCTTGAA CCCAGGAGGC TTAAGTTGCG GTGAGCCAAG
      TCCGTCCTCT TGACGAACTT GGGTCCTCCG AATTCAACGC CACTCGGTTC
2101  ATCACGCCAT TGCACTCTAG CCTGGACAAC AAGAGAAAAA CTCTGTCTCA
      TAGTGCGGTA ACGTGAGATC GGACCTGTTG TTCTCTTTTT GAGACAGAGT
2151  AAAAAACACA AAAACAAAAA ACCCAAATAC TATTTAAAAA AGATAAACCT
      TTTTTTGTGT TTTTGTTTTT TGGGTTTATG ATAAATTTTT TCTATTTGGA
2201  TAATTGCTCA ATCATTAAAG CCATCCCACA AGTAAAGCAG CAAGCAGAAA
      ATTAACGAGT TAGTAATTTC GGTAGGGTGT TCATTTCGTC GTTCGTCTTT
2251  AAAGTTAAGA ACACCTCAAG GCTACAGAAG GACATTTCAA GCTATGCAGG
      TTTCAATTCT TGTGGAGTTC CGATGTCTTC CTGTAAAGTT CGATACGTCC
2301  CATATGAAGT GTGCAGACAG ATATGTAAGA AAGGCCTCAA GACTGCAAAA
      GTATACTTCA CACGTCTGTC TATACATTCT TTCCGGAGTT CTGACGTTTT
2351  GGGCATTTCA AGCTATGCAA GCATATAGGT AACACATACA CACACACAAA
      CCCGTAAAGT TCGATACGTT CGTATATCCA TTGTGTATGT GTGTGTGTTT
2401  ATAAAATCCC CTGAAATACA AAAACATGCA GCAAACACCT GACGTTTTTG
      TATTTTAGGG GACTTTATGT TTTTGTACGT CGTTTGTGGA CTGCAAAAAC
2451  GATACCATTT CTAAGTCAGG TGTTATGATT CTCATTAGTC AAGATACTTG
      CTATGGTAAA GATTCAGTCC ACAATACTAA GAGTAATCAG TTCTATGAAC
2501  AGTACTGGGC CCAAACAGCT TTCTGCCACT GTACAGTACA AGAAGGTAGG
      TCATGACCCG GGTTTGTCGA AAGACGGTGA CATGTCATGT TCTTCCATCC
2551  AATAATGGTG GGAGGAGCAA AGACAAACTG TAATAGACAG AAGTGTATCA
      TTATTACCAC CCTCCTCGTT TCTGTTTGAC ATTATCTGTC TTCACATAGT
2601  GATACCTATA CTACATGAAA AACAAAACAG CTACTGCCAC AAAGGGAGAA
      CTATGGATAT GATGTACTTT TTGTTTTGTC GATGACGGTG TTTCCCTCTT
2651  GGCTAACAAA ATAAAGTCAA CAATAAATAC AGAAAATGAA AAGGATACAC
      CCGATTGTTT TATTTCAGTT GTTATTTATG TCTTTTACTT TTCCTATGTG
2701  ACTAAGGTTT ACAAAAAAAA AAAGGCAGAC AAAATGCCAT ACAGTATTCA
      TGATTCCAAA TGTTTTTTTT TTTCCGTCTG TTTTACGGTA TGTCATAAGT
2751  TTCACTACTA TGGCATTCAT AAGCTAGTTT CAAATGCTCA CTATTTTCTT
      AAGTGATGAT ACCGTAAGTA TTCGATCAAA GTTTACGAGT GATAAAAGAA
```

FIG. 27(contd.)

```
2801    TTATAGTATA TATTTGCCTT AACCCAGCAC TTTTTTCCAA AAGTGGATGA
        AATATCATAT ATAAACGGAA TTGGGTCGTG AAAAAAGGTT TTCACCTACT
2851    GTCAAAATAA ATTTCCCATT ATTTAAGTGA AATTAACAGC ACACATATCT
        CAGTTTTATT TAAAGGGTAA TAAATTCACT TTAATTGTCG TGTGTATAGA
2901    CACAACACTA ATGAATTTTT AAAATGGAAA GTTAAGAACT TTTAAAGTGG
        GTGTTGTGAT TACTTAAAAA TTTTACCTTT CAATTCTTGA AAATTTCACC
2951    CCAACCTGTG ATCCTTCACA AAATAAACTA AATACAATAA CAGACCCCAA
        GGTTGGACAC TAGGAAGTGT TTTATTTGAT TTATGTTATT GTCTGGGGTT
3001    AGGCTATCAA TTGCGTGCAA AAACAACTTC TGTTTTCCAG GGTAAACAGA
        TCCGATAGTT AACGCACGTT TTTGTTGAAG ACAAAAGGTC CCATTTGTCT
3051    ATCTAATGCA GAATCTAATG CAGGGTAAAC AGACTTAATG CAGAATCTAA
        TAGATTACGT CTTAGATTAC GTCCCATTTG TCTGAATTAC GTCTTAGATT
3101    TGATGGCACA AATTAAAAAT CACTAACGTG CCCTTTTTAG TGTGAAACCC
        ACTACCGTGT TTAATTTTTA GTGATTGCAC GGGAAAAATC ACACTTTGGG
3151    AGAGAGAGCA CATACAAGCC AAAAACAAAT GCTTTATTTT ACCTAGGAGA
        TCTCTCTCGT GTATGTTCGG TTTTTGTTTA CGAAATAAAA TGGATCCTCT
3201    CATTAACATT CACCTTTACG TGTTTAAGAT TAATGCAATG TTAAATATTG
        GTAATTGTAA GTGGAAATGC ACAAATTCTA ATTACGTTAC AATTTATAAC
3251    TGAAAACTGT AACTTTGAAT TTCATGATTT TTATGTGAAT ATTCCAGGGT
        ACTTTTGACA TTGAAACTTA AAGTACTAAA AATACACTTA TAAGGTCCCA
3301    TTAAAAAAAC TTGTAACATG ACATGGCTGA ATAAGATAAA AAAAAAATCT
        AATTTTTTTG AACATTGTAC TGTACCGACT TATTCTATTT TTTTTTAGA
3351    AGCCTTTTCT CCCTTCTGGC TCATATTTGC GATTTCGATC ATTTTGTTTA
        TCGGAAAAGA GGGAAGACCG AGTATAAACG CTAAAGCTAG TAAAACAAAT
3401    AAAAACAAAA CACTGCAATG AATTAAACTT AATATTCTTC TATGTTTTAG
        TTTTTGTTTT GTGACGTTAC TTAATTTGAA TTATAAGAAG ATACAAAATC
3451    AGTAAGTTAA AACAAGATAA AGTGACCAAA GTAATTTGAA AGATTCAATG
        TCATTCAATT TTGTTCTATT TCACTGGTTT CATTAAACTT TCTAAGTTAC
3501    ACTTTTGCTC CAACCTAGGT GCACAAGGTA CCTTGTTCTT TAAATTGGGC
        TGAAAACGAG GTTGGATCCA CGTGTTCCAT GGAACAAGAA ATTTAACCCG
3551    TTTAATGAAA ATACTTCTCC AGAATTCTGG GGATTTAAGA AAAATTATGC
        AAATTACTTT TATGAAGAGG TCTTAAGACC CCTAAATTCT TTTTAATACG
3601    CAACCAACAA GGGCTTTACC ATTTTATGTA ACATTTTTCA ACGCTGCAAA
        GTTGGTTGTT CCCGAAATGG TAAAATACAT TGTAAAAAGT TGCGACGTTT
3651    AATGTGTGTA TTTCTATTTG AAGATAAAAA TCCTCAGCAA AATCCACATT
        TTACACACAT AAAGATAAAC TTCTATTTTT AGGAGTCGTT TTAGGTGTAA
3701    GCACTGTCCT TCAAAGATTA GCCTTCTTTG AACTAGTTAA GACACTATTA
        CGTGACAGGA AGTTTCTAAT CGGAAGAAAC TTGATCAATT CTGTGATAAT
3751    AGCCAAGCCA GTATCTCCCT GTAATGAATT CGTTTTTCTC TTAATTTTCC
        TCGGTTCGGT CATAGAGGGA CATTACTTAA GCAAAAAGAG AATTAAAAGG
3801    CCTGTAATTT ACACTGGGAG AGCTGGGAAA TATGTGGATG TAAATTTCTC
        GGACATTAAA TGTGACCCTC TCGACCCTTT ATACACCTAC ATTTAAAGAG
3851    AGCCACAGAG ATGCAAAGTT ATACTGTGGG GAAAAAAAAC TTGAGTTAAA
        TCGGTGTCTC TACGTTTCAA TATGACACCC CTTTTTTTTG AACTCAATTT
3901    TCCTTACATA TTTTAGGTTT TCATTAACTT ACCAATGTAG TTTTGTTGGA
        AGGAATGTAT AAAATCCAAA AGTAATTGAA TGGTTACATC AAAACAACCT
3951    GGCCATTTTT TTTATTGCAG ACTTGAAGAG CTATTACTAG AAAAATGCAT
        CCGGTAAAAA AAATAACGTC TGAACTTCTC GATAATGATC TTTTTACGTA
4001    GACAGTTAAG GTAAGTTTGC ATGACACAAA AAAGGTAACT AAATACAAAT
        CTGTCAATTC CATTCAAACG TACTGTGTTT TTTCCATTGA TTTATGTTTA
4051    TCTGTTTGGA TTCCAACCCC CAAGTAGAGA GCGCACACTT TCAAACGTGA
        AGACAAACCT AAGGTTGGGG GTTCATCTCT CGCGTGTGAA AGTTTGCACT
4101    ATACAAATCC AGAGTAGATC TGCGCTCCTA CCTACATTGC TTATGATGTA
        TATGTTTAGG TCTCATCTAG ACGCGAGGAT GGATGTAACG AATACTACAT
4151    CTTAAGTACG TGTCCTAACC ATGTGAGTCT AGAAAGACTT TACTGGGGAT
        GAATTCATGC ACAGGATTGG TACACTCAGA TCTTTCTGAA ATGACCCCTA
4201    CCTGGTACCT AAAACAGCTT CACATGGCTT AAAATAGGGG ACCAATGTCT
        GGACCATGGA TTTTGTCGAA GTGTACCGAA TTTTATCCCC TGGTTACAGA
```

FIG. 27(contd.)


```
4251  TTTCCAATCT AAGTCCCATT TATAATAAAG TCCATGTTCC ATTTTTAAAG
      AAAGGTTAGA TTCAGGGTAA ATATTATTTC AGGTACAAGG TAAAAATTTC
4301  GACAATCCTT TCGGTTTAAA ACCAGGCACG ATTACCCAAA CAACTCACAA
      CTGTTAGGAA AGCCAAATTT TGGTCCGTGC TAATGGGTTT GTTGAGTGTT
4351  CGGTAAAGCA CTGTGAATCT TCTCTGTTCT GCAATCCCAA CTTGGTTTCT
      GCCATTTCGT GACACTTAGA AGAGACAAGA CGTTAGGGTT GAACCAAAGA
4401  GCTCAGAAAC CCTCCCTCTT TCCAATCGGT AATTAAATAA CAAAAGGAAA
      CGAGTCTTTG GGAGGGAGAA AGGTTAGCCA TTAATTTATT GTTTTCCTTT
4451  AAACTTAAGA TGCTTCAACC CCGTTTCGTG ACACTTTGAA AAAAGAATCA
      TTTGAATTCT ACGAAGTTGG GGCAAAGCAC TGTGAAACTT TTTTCTTAGT
4501  CCTCTTGCAA ACACCCGCTC CCGACCCCCG CCGCTGAAGC CCGGCGTCCA
      GGAGAACGTT TGTGGGCGAG GGCTGGGGGC GGCGACTTCG GGCCGCAGGT
4551  GAGGCCTAAG CGCGGGTGCC CGCCCCCACC CGGGAGCGCG GGCCTCGTGG
      CTCCGGATTC GCGCCCACGG GCGGGGGTGG GCCCTCGCGC CCGGAGCACC
4601  TCAGCGCATC CGCGGGGAGA AACAAAGGCC GCGGCACGGG GGCTCAAGGG
      AGTCGCGTAG GCGCCCCTCT TTGTTTCCGG CGCCGTGCCC CCGAGTTCCC
4651  CACTGCGCCA CACCGCACGC GCCTACCCCC GCGCGGCCAC GTTAACTGGC
      GTGACGCGGT GTGGCGTGCG CGGATGGGGG CGCGCCGGTG CAATTGACCG
4701  GGTCGCCGCA GCCTCGGGAC AGCCGGCCGC GCGCCGCCAG GCTCGCGGAC
      CCAGCGGCGT CGGAGCCCTG TCGGCCGGCG CGCGGCGGTC CGAGCGCCTG
4751  GCGGGACCAC GCGCCGCCCT CCGGGAGGCC CAAGTCTCGA CCCAGCCCCG
      CGCCCTGGTG CGCGGCGGGA GGCCCTCCGG GTTCAGAGCT GGGTCGGGGC
4801  CGTGGCGCTG GGGGAGGGGG CGCCTCCGCC GGAACGCGGG TGGGGGAGGG
      GCACCGCGAC CCCCTCCCCC GCGGAGGCGG CCTTGCGCCC ACCCCCTCCC
4851  GAGGGGGAAA TGCGCTTTGT CTCGAAATGG GGCAACCGTC GCCACAGCTC
      CTCCCCCTTT ACGCGAAACA GAGCTTTACC CCGTTGGCAG CGGTGTCGAG
4901  CCTACCCCCT CGAGGGCAGA GCAGTCCCCC CACTAACTAC CGGGCTGGCC
      GGATGGGGGA GCTCCCGTCT CGTCAGGGGG GTGATTGATG GCCCGACCGG
4951  GCGCCCCAGG CCAGCCGCGA GGCCACCGCC CGACCCTCCA CTCCTTCCCG
      CGCGCGGTCC GGTCGGCGCT CCGGTGGCGG GCTGGGAGGT GAGGAAGGGC
5001  CAGCTCCCGG CGCGGGGTCC GGCGAGAAGG GGAGGGGAGG GGAGCGGAGA
      GTCGAGGGCC GCGCCCCAGG CCGCTCTTCC CCTCCCCTCC CCTCGCCTCT
5051  ACCGGGCCCC CGGGACGCGT GTGGCATCTG AAGCACCACC AGCGAGCGAG
      TGGCCCGGGG GCCCTGCGCA CACCGTAGAC TTCGTGGTGG TCGCTCGCTC
5101  AGCTAGAGAG AAGGAAAGCC ACCGACTTCA CCGCCTCCGA GCTGCTCCGG
      TCGATCTCTC TTCCTTTCGG TGGCTGAAGT GGCGGAGGCT CGACGAGGCC
5151  GTCGCGGGTC TGCAGCGTCT CCGGCCCTCC GCGCCTACAG CTCAAGCCAC
      CAGCGCCCAG ACGTCGCAGA GGCCGGGAGG CGCGGATGTC GAGTTCGGTG
5201  ATCCGAAGGG GGAGGGAGCC GGGAGCTGCG CGCGGGGCCG CCGGGGGGAG
      TAGGCTTCCC CCTCCCTCGG CCCTCGACGC GCGCCCCGGC GGCCCCCCTC
5251  GGGTGGCACC GCCCACGCCG GGCGGCCACG AAGGGCGGGG CAGCGGGCGC
      CCCACCGTGG CGGGTGCGGC CCGCCGGTGC TTCCCGCCCC GTCGCCCGCG
5301  GCGCGCGGCG GGGGGAGGGG CCGGCGCCGC GCCCGCTGGG AATTGGGGCC
      CGCGCGCCGC CCCCCTCCCC GGCCGCGGCG CGGGCGACCC TTAACCCCGG
5351  CTAGGGGGAG GGCGGAGGCG CCGACGACCG CGGCACTTAC CGTTCGCGGC
      GATCCCCCTC CCGCCTCCGC GGCTGCTGGC GCCGTGAATG GCAAGCGCCG
5401  GTGGCGCCCG GTGGTCCCCA AGGGGAGGGA AGGGGGAGGC GGGGCGAGGA
      CACCGCGGGC CACCAGGGGT TCCCCTCCCT TCCCCCTCCG CCCCGCTCCT
5451  CAGTGACCGG AGTCTCCTCA GCGGTGGCTT TTCTGCTTGG CAGCCTCAGC
      GTCACTGGCC TCAGAGGAGT CGCCACCGAA AAGACGAACC GTCGGAGTCG
5501  GGCTGGCGCC AAAACCGGAC TCCGCCCACT TCCTCGCCCG CCGGTGCGAG
      CCGACCGCGG TTTTGGCCTG AGGCGGGTGA AGGAGCGGGC GGCCACGCTC
5551  GGTGTGGAAT CCTCCAGACG CTGGGGGAGG GGGAGTTGGG AGCTTAAAAA
      CCACACCTTA GGAGGTCTGC GACCCCCTCC CCCTCAACCC TCGAATTTTT
5601  CTAGTACCCC TTTGGGACCA CTTTCAGCAG CGAACTCTCC TGTACACCAG
      GATCATGGGG AAACCCTGGT GAAAGTCGTC GCTTGAGAGG ACATGTGGTC
5651  GGGTCAGTTC CACAGACGCG GGCCAGGGGT GGGTCATTGC GGCGTGAACA
      CCCAGTCAAG GTGTCTGCGC CCGGTCCCCA CCCAGTAACG CCGCACTTGT
```

FIG. 27(contd.)

```
5701    ATAATTTGAC TAGAAGTTGA TTCGGGTGTT TCCGGAAGGG GCCGAGTCAA
        TATTAAACTG ATCTTCAACT AAGCCCACAA AGGCCTTCCC CGGCTCAGTT
5751    TCCGCCGAGT TGGGGCACGG AAAACAAAAA GGGAAGGCTA CTAAGATTTT
        AGGCGGCTCA ACCCCGTGCC TTTTGTTTTT CCCTTCCGAT GATTCTAAAA
5801    TCTGGCGGGG GTTATCATTG GCGTAACTGC AGGGACCACC TCCCGGGTTG
        AGACCGCCCC CAATAGTAAC CGCATTGACG TCCCTGGTGG AGGGCCCAAC
5851    AGGGGGCTGG ATCTCCAGGC TGCGGATTAA GCCCCTCCCG TCGGCGTTAA
        TCCCCCCGACC TAGAGGTCCG ACGCCTAATT CGGGGAGGGC AGCCGCAATT
5901    TTTCAAACTG CGCGACGTTT CTCACCTGCC TTCGCCAAGG CAGGGGCCGG
        AAAGTTTGAC GCGCTGCAAA GAGTGGACGG AAGCGGTTCC GTCCCCGGCC
5951    GACCCTATTC CAAGAGGTAG TAACTAGCAG GACTCTAGCC TTCCGCAATT
        CTGGGATAAG GTTCTCCATC ATTGATCGTC CTGAGATCGG AAGGCGTTAA
6001    CATTGAGCGC ATTTACGGAA GTAACGTCGG GTACTGTCTC TGGCCGCAAG
        GTAACTCGCG TAAATGCCTT CATTGCAGCC CATGACAGAG ACCGGCGTTC
6051    GGTGGGAGGA GTACGCATTT GGCGTAAGGT GGGGCGTAGA GCCTTCCCGC
        CCACCCTCCT CATGCGTAAA CCGCATTCCA CCCCGCATCT CGGAAGGGCG
6101    CATTGGCGGC GGATAGGGCG TTTACGCGAC GGCCTGACGT AGCGGAAGAC
        GTAACCGCCG CCTATCCCGC AAATGCGCTG CCGGACTGCA TCGCCTTCTG
6151    GCGTTAGTGG GGGGGAAGGT TCTAGAAAAG CGGCGGCAGC GGCTCTAGCG
        CGCAATCACC CCCCCTTCCA AGATCTTTTC GCCGCCGTCG CCGAGATCGC
6201    GCAGTAGCAG CAGCGCCGGG TCCCGTGCGG AGGTGCTCCT CGCAGAGTTG
        CGTCATCGTC GTCGCGGCCC AGGGCACGCC TCCACGAGGA GCGTCTCAAC
6251    TTTCTCGAGC AGCGGCAGTT CTCACTACAG CGCCAGGACG AGTCCGGTTC
        AAAGAGCTCG TCGCCGTCAA GAGTGATGTC GCGGTCCTGC TCAGGCCAAG
6301    GTGTTCGTCC GCGGAGATCT CTCTCATCTC GCTCGGCTGC GGGAAATCGG
        CACAAGCAGG CGCCTCTAGA GAGAGTAGAG CGAGCCGACG CCCTTTAGCC
6351    GCTGAAGCGA CTGAGTCCGC GATGGAGGTA ACGGGTTTGA AATCAATGAG
        CGACTTCGCT GACTCAGGCG CTACCTCCAT TGCCCAAACT TTAGTTACTC
6401    TTATTGAAAA GGGCATGGCG AGGCCGTTGG CGCCTCAGTG GAAGTCGGCC
        AATAACTTTT CCCGTACCGC TCCGGCAACC GCGGAGTCAC CTTCAGCCGG
6451    AGCCGCCTCC GTGGGAGAGA GGCAGGAAAT CGGACCAATT CAGTAGCAGT
        TCGGCGGAGG CACCCTCTCT CCGTCCTTTA GCCTGGTTAA GTCATCGTCA
6501    GGGGCTTAAG GTTTATGAAC GGGGTCTTGA GCGGAGGCCT GAGCGTACAA
        CCCCGAATTC CAAATACTTG CCCCAGAACT CGCCTCCGGA CTCGCATGTT
6551    ACAGCTTCCC CACCCTCAGC CTCCCGGCGC CATTTCCCTT CACTGGGGGT
        TGTCGAAGGG GTGGGAGTCG GAGGGCCGCG GTAAAGGGAA GTGACCCCCA
6601    GGGGGATGGG GAGCTTTCAC ATGGCGGACG CTGCCCCGCT GGGGTGAAAG
        CCCCCTACCC CTCGAAAGTG TACCGCCTGC GACGGGGCGA CCCCACTTTC
6651    TGGGGCGCGG AGGCGGGAAT TCTTATTCCC TTTCTAAAGC ACGCTGCTTC
        ACCCCGCGCC TCCGCCCTTA AGAATAAGGG AAAGATTTCG TGCGACGAAG
6701    GGGGGCCACG GCGTCTCCTC GGCGAGCGTT TCGGCGGGCA GCAGGTCCTC
        CCCCCGGTGC CGCAGAGGAG CCGCTCGCAA AGCCGCCCGT CGTCCAGGAG
6751    GTGAGCGAGG CTGCGGAGCT TCCCCTCCCC CTCTCTCCCG GGAACCGATT
        CACTCGCTCC GACGCCTCGA AGGGGAGGGG GAGAGAGGGC CCTTGGCTAA
6801    TGGCGGCCGC CATTTTCATG GCTCGCCTTC CTCTCAGCGT TTTCCTTATA
        ACCGCCGGCG GTAAAAGTAC CGAGCGGAAG GAGAGTCGCA AAAGGAATAT
6851    ACTCTTTTAT TTTCTTAGTG TGCTTTCTCT ATCAAGAAGT AGAAGTGGTT
        TGAGAAAATA AAAGAATCAC ACGAAAGAGA TAGTTCTTCA TCTTCACCAA
6901    AACTATTTTT TTTTTCTTCT CGGGCTGTTT TCATATCGTT TCGAGGTGGA
        TTGATAAAAA AAAAAGAAGA GCCCGACAAA AGTATAGCAA AGCTCCACCT
6951    TTTGGAGTGT TTTGTGAGCT TGGATCTTTA GAGTCCTGCG CACCTCATTA
        AAACCTCACA AAACACTCGA ACCTAGAAAT CTCAGGACGC GTGGAGTAAT
7001    AAGGCGCTCA GCCTTCCCCT CGATGAAATG GCGCCATTGC GTTCGGAAGC
        TTCCGCGAGT CGGAAGGGGA GCTACTTTAC CGCGGTAACG CAAGCCTTCG
7051    CACACCGAAG AGCGGGGAGG GGGGGTGCTC CGGGTTTGCG GGCCCGGTTT
        GTGTGGCTTC TCGCCCCTCC CCCCCACGAG GCCCAAACGC CCGGGCCAAA
7101    CAGAGAAGAT ATCACCACCC AGGGCGTCGG GCCGGGTTCA ATGCGAGCCG
        GTCTCTTCTA TAGTGGTGGG TCCCGCAGCC CGGCCCAAGT TACGCTCGGC
```

FIG. 27(contd.)

```
7151    TAGGACAAAG AAACCATTTT ATGTTTTTCC TGTCTTTTTT TTCCTTTGAG
        ATCCTGTTTC TTTGGTAAAA TACAAAAAGG ACAGAAAAAA AAGGAAACTC
7201    TAACGGTTTT ATCTGGGTCT GCAGTCAGTA AAACGACAGA TGAACCGCGG
        ATTGCCAAAA TAGACCCAGA CGTCAGTCAT TTTGCTGTCT ACTTGGCGCC
7251    CAAAATAAAC ATAAATTGGA AGCCATCGGC CACGAGGGGC AGGGACGAAG
        GTTTTATTTG TATTTAACCT TCGGTAGCCG GTGCTCCCCG TCCCTGCTTC
7301    GTGGTTTTCT GGGCGGGGGA GGGATATTCG CGTCAGAATC CTTTACTGTT
        CACCAAAAGA CCCGCCCCCT CCCTATAAGC GCAGTCTTAG GAAATGACAA
7351    CTTAAGGATT CCGTTTAAGT TGTAGAGCTG ACTCATTTTA AGTAATGTTG
        GAATTCCTAA GGCAAATTCA ACATCTCGAC TGAGTAAAAT TCATTACAAC
7401    TTACTGAGAA GTTTAACCCT TACGGGACAG ATCCATGGAC CTTTATAGAT
        AATGACTCTT CAAATTGGGA ATGCCCTGTC TAGGTACCTG GAAATATCTA
7451    GATTACGAGG AAAGTGAAAT AACGATTTTG TCCTTAGTTA TACTTCGATT
        CTAATGCTCC TTTCACTTTA TTGCTAAAAC AGGAATCAAT ATGAAGCTAA
7501    AAAACATGGC TTCAGAGGCT CCTTCCTGTA ATGCGTATGG ATTGATGTGC
        TTTTGTACCG AAGTCTCCGA GGAAGGACAT TACGCATACC TAACTACACG
7551    AAAACTGTTT TGGGCCTGGG CCGCTCTGTA TTTGAACTTT GTTACTTTTC
        TTTTGACAAA ACCCGGACCC GGCGAGACAT AAACTTGAAA CAATGAAAAG
7601    TCATTTTGTT TGCAATCTTG GTTGAACATT ACATTGATAA GCATAAGGTC
        AGTAAAACAA ACGTTAGAAC CAACTTGTAA TGTAACTATT CGTATTCCAG
7651    TCAAGCGAAG GGGGTCTACC TGGTTATTTT TCTTTGACCC TAAGCACGTT
        AGTTCGCTTC CCCCAGATGG ACCAATAAAA AGAAACTGGG ATTCGTGCAA
7701    TATAAAATAA CATTGTTTAA AATCGATAGT GGACATCGGG TAAGTTTGGA
        ATATTTTATT GTAACAAATT TTAGCTATCA CCTGTAGCCC ATTCAAACCT
7751    TAAATTGTGA GGTAAGTAAT GAGTTTTTGC TTTTTGTTAG TGATTTGTAA
        ATTTAACACT CCATTCATTA CTCAAAAACG AAAAACAATC ACTAAACATT
7801    AACTTGTTAT AAATGTACAT TATCCGTAAT TTCAGTTTAG AGATAACCTA
        TTGAACAATA TTTACATGTA ATAGGCATTA AAGTCAAATC TCTATTGGAT
7851    TGTGCTGACG ACAATTAAGA ATAAAAACTA GCTGAAAAAA TGAAAATAAC
        ACACGACTGC TGTTAATTCT TATTTTTGAT CGACTTTTTT ACTTTTATTG
7901    TATCGTGACA AGTAACCATT TCAAAAGACT GCTTTGTGTC TCATAGGAGC
        ATAGCACTGT TCATTGGTAA AGTTTTCTGA CGAAACACAG AGTATCCTCG
7951    TAGTTTGATC ATTTCAGTTA ATTTTTTCTT TAATTTTTAC GAGTCATGAA
        ATCAAACTAG TAAAGTCAAT TAAAAAAGAA ATTAAAAATG CTCAGTACTT
8001    AACTACAGGA AAAAAAATCT GAACTGGGTT TTACCACTAC TTTTTAGGAG
        TTGATGTCCT TTTTTTTAGA CTTGACCCAA AATGGTGATG AAAAATCCTC
8051    TTGGGAGCAT GCGAATGGAG GGAGAGCTCC GTAGAACTGG GATGAGAGCA
        AACCCTCGTA CGCTTACCTC CCTCTCGAGG CATCTTGACC CTACTCTCGT
8101    GCAATTAATG CTGCTTGCTA GGAACAAAAA ATAATTGATT GAAAATTACG
        CGTTAATTAC GACGAACGAT CCTTGTTTTT TATTAACTAA CTTTTAATGC
8151    TGTGACTTTT TAGTTTGCAT TATGCGTTTG TAGCAGTTGG TCCTGGATAT
        ACACTGAAAA ATCAAACGTA ATACGCAAAC ATCGTCAACC AGGACCTATA
8201    CACTTTCTCT CGTTTGAGGT TTTTTAACCT AGTTAACTTT TAAGACAGGT
        GTGAAAGAGA GCAAACTCCA AAAAATTGGA TCAATTGAAA ATTCTGTCCA
8251    TTCCTTAACA TTCATAAGTG CCCAGAATAC AGCTGTGTAG TACAGCATAT
        AAGGAATTGT AAGTATTCAC GGGTCTTATG TCGACACATC ATGTCGTATA
8301    AAAGATTTCA GCTCTGAGGT TTTTCCTATT GACTTGGAAA ATTGTTTTGT
        TTTCTAAAGT CGAGACTCCA AAAAGGATAA CTGAACCTTT TAACAAAACA
8351    GCCTGTCGCT TGCCACATGG CCAATCAAGT AAGCTTATCG ATACCGGTGG
        CGGACAGCGA ACGGTGTACC GGTTAGTTCA TTCGAATAGC TATGGCCACC
8401    CGCGCCAATT GTTAATTAAG ATCTGGCCCA ATGGGCCGTA CGAATTTGAG
        GCGCGGTTAA CAATTAATTC TAGACCGGGT TACCCGGCAT GCTTAAACTC
8451    GCGGAAAGAA CCAGCTGTGG AATGTGTGTC AGTTAGGGTG TGGAAAGTCC
        CGCCTTTCTT GGTCGACACC TTACACACAG TCAATCCCAC ACCTTTCAGG
8501    CCAGGCTCCC CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC
        GGTCCGAGGG GTCGTCCGTC TTCATACGTT TCGTACGTAG AGTTAATCAG
8551    AGCAACCAGG TGTGGAAAGT CCCCAGGCTC CCCAGCAGGC AGAAGTATGC
        TCGTTGGTCC ACACCTTTCA GGGGTCCGAG GGGTCGTCCG TCTTCATACG
```

FIG. 27(contd.)

```
8601   AAAGCATGCA TCTCAATTAG TCAGCAACCA TAGTCCCGCC CCTAACTCCG
       TTTCGTACGT AGAGTTAATC AGTCGTTGGT ATCAGGGCGG GGATTGAGGC
8651   CCCATCCCGC CCCTAACTCC GCCCAGTTCC GCCCATTCTC CGCCCCATGG
       GGGTAGGGCG GGGATTGAGG CGGGTCAAGG CGGGTAAGAG GCGGGGTACC
8701   CTGACTAATT TTTTTTATTT ATGCAGAGGC CGAGGCCGCC TCGGCCTCTG
       GACTGATTAA AAAAAATAAA TACGTCTCCG GCTCCGGCGG AGCCGGAGAC
8751   AGCTATTCCA GAAGTAGTGA GGAGGCTTTT TTGGAGGCCT AGGCTTTTGC
       TCGATAAGGT CTTCATCACT CCTCCGAAAA AACCTCCGGA TCCGAAAACG
8801   AAAGATCGAT CAAGAGACAG GATGAGGATC GTTTCGCATG ATTGAACAAG
       TTTCTAGCTA GTTCTCTGTC CTACTCCTAG CAAAGCGTAC TAACTTGTTC
8851   ATGGATTGCA CGCAGGTTCT CCGGCCGCTT GGGTGGAGAG GCTATTCGGC
       TACCTAACGT GCGTCCAAGA GGCCGGCGAA CCCACCTCTC CGATAAGCCG
8901   TATGACTGGG CACAACAGAC AATCGGCTGC TCTGATGCCG CCGTGTTCCG
       ATACTGACCC GTGTTGTCTG TTAGCCGACG AGACTACGGC GGCACAAGGC
8951   GCTGTCAGCG CAGGGGCGCC CGGTTCTTTT TGTCAAGACC GACCTGTCCG
       CGACAGTCGC GTCCCCGCGG GCCAAGAAAA ACAGTTCTGG CTGGACAGGC
9001   GTGCCCTGAA TGAACTGCAA GACGAGGCAG CGCGGCTATC GTGGCTGGCC
       CACGGGACTT ACTTGACGTT CTGCTCCGTC GCGCCGATAG CACCGACCGG
9051   ACGACGGGCG TTCCTTGCGC AGCTGTGCTC GACGTTGTCA CTGAAGCGGG
       TGCTGCCCGC AAGGAACGCG TCGACACGAG CTGCAACAGT GACTTCGCCC
9101   AAGGGACTGG CTGCTATTGG GCGAAGTGCC GGGGCAGGAT CTCCTGTCAT
       TTCCCTGACC GACGATAACC CGCTTCACGG CCCCGTCCTA GAGGACAGTA
9151   CTCACCTTGC TCCTGCCGAG AAAGTATCCA TCATGGCTGA TGCAATGCGG
       GAGTGGAACG AGGACGGCTC TTTCATAGGT AGTACCGACT ACGTTACGCC
9201   CGGCTGCATA CGCTTGATCC GGCTACCTGC CCATTCGACC ACCAAGCGAA
       GCCGACGTAT GCGAACTAGG CCGATGGACG GGTAAGCTGG TGGTTCGCTT
9251   ACATCGCATC GAGCGAGCAC GTACTCGGAT GGAAGCCGGT CTTGTCGATC
       TGTAGCGTAG CTCGCTCGTG CATGAGCCTA CCTTCGGCCA GAACAGCTAG
9301   AGGATGATCA AGAGCATCAG GGGCTCGCGC CAGCCGAACT GTTCGCCAGG
       TCCTACTAGT TCTCGTAGTC CCCGAGCGCG GTCGGCTTGA CAAGCGGTCC
9351   CTCAAGGCGA GCATGCCCGA CGGCGAGGAT CTCGTCGTGA CCCATGGCGA
       GAGTTCCGCT CGTACGGGCT GCCGCTCCTA GAGCAGCACT GGGTACCGCT
9401   TGCCTGCTTG CCGAATATCA TGGTGGAAAA TGGCCGCTTT TCTGGATTCA
       ACGGACGAAC GGCTTATAGT ACCACCTTTT ACCGGCGAAA AGACCTAAGT
9451   TCGACTGTGG CCGGCTGGGT GTGGCGGACC GCTATCAGGA CATAGCGTTG
       AGCTGACACC GGCCGACCCA CACCGCCTGG CGATAGTCCT GTATCGCAAC
9501   GCTACCCGTG ATATTGCTGA AGAGCTTGGC GGCGAATGGG CTGACCGCTT
       CGATGGGCAC TATAACGACT TCTCGAACCG CCGCTTACCC GACTGGCGAA
9551   CCTCGTGCTT TACGGTATCG CCGCTCCCGA TTCGCAGCGC ATCGCCTTCT
       GGAGCACGAA ATGCCATAGC GGCGAGGGCT AAGCGTCGCG TAGCGGAAGA
9601   ATCGCCTTCT TGACGAGTTC TTCTGAGCGG GACTCTGGGG TTCGAAATGA
       TAGCGGAAGA ACTGCTCAAG AAGACTCGCC CTGAGACCCC AAGCTTTACT
9651   CCGACCAAGC GACGCCCAAC CTGCCATCAC GAGATTTCGA TTCCACCGCC
       GGCTGGTTCG CTGCGGGTTG GACGGTAGTG CTCTAAAGCT AAGGTGGCGG
9701   GCCTTCTATG AAAGGTTGGG CTTCGGAATC GTTTTCCGGG ACGCCGGCTG
       CGGAAGATAC TTTCCAACCC GAAGCCTTAG CAAAAGGCCC TGCGGCCGAC
9751   GATGATCCTC CAGCGCGGGG ATCTCATGCT GGAGTTCTTC GCCCACCCTA
       CTACTAGGAG GTCGCGCCCC TAGAGTACGA CCTCAAGAAG CGGGTGGGAT
9801   GGGGGAGGCT AACTGAAACA CGGAAGGAGA CAATACCGGA AGGAACCCGC
       CCCCCTCCGA TTGACTTTGT GCCTTCCTCT GTTATGGCCT TCCTTGGGCG
9851   GCTATGACGG CAATAAAAAG ACAGAATAAA ACGCACGGTG TTGGGTCGTT
       CGATACTGCC GTTATTTTTC TGTCTTATTT TGCGTGCCAC AACCCAGCAA
9901   TGTTCATAAA CGCGGGGTTC GGTCCCAGGG CTGGCACTCT GTCGATACCC
       ACAAGTATTT GCGCCCCAAG CCAGGGTCCC GACCGTGAGA CAGCTATGGG
9951   CACCGAGACC CCATTGGGGC CAATACGCCC GCGTTTCTTC CTTTTCCCCA
       GTGGCTCTGG GGTAACCCCG GTTATGCGGG CGCAAAGAAG GAAAAGGGGT
10001  CCCCACCCCC CAAGTTCGGG TGAAGGCCCA GGGCTCGCAG CCAACGTCGG
       GGGGTGGGGG GTTCAAGCCC ACTTCCGGGT CCCGAGCGTC GGTTGCAGCC
```

FIG. 27(contd.)

```
10051   GGCGGCAGGC CCTGCCATAG CCTCAAATTC CTTAGGCTCG AGGGCCGCCA
        CCGCCGTCCG GGACGGTATC GGAGTTTAAG GAATCCGAGC TCCCGGCGGT
10101   CCGCGGTGGA GCTCCAGCTT TTGTTCCCTT TAGTGAGGGT TAATTTCGAG
        GGCGCCACCT CGAGGTCGAA AACAAGGGAA ATCACTCCCA ATTAAAGCTC
10151   CTTGGCGTAA TCATGGTCAT AGCTGTTTCC TGTGTGAAAT TGTTATCCGC
        GAACCGCATT AGTACCAGTA TCGACAAAGG ACACACTTTA ACAATAGGCG
10201   TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG TAAAGCCTGG
        AGTGTTAAGG TGTGTTGTAT GCTCGGCCTT CGTATTTCAC ATTTCGGACC
10251   GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC
        CCACGGATTA CTCACTCGAT TGAGTGTAAT TAACGCAACG CGAGTGACGG
10301   CGCTTTCCAG TCGGGAAACC TGTCGTGCCA GCATCGCGAG CACTTTTCGG
        GCGAAAGGTC AGCCCTTTGG ACAGCACGGT CGTAGCGCTC GTGAAAAGCC
10351   GGAAATGTGC GCGGAACCCC TATTTGTTTA TTTTTCTAAA TACATTCAAA
        CCTTTACACG CGCCTTGGGG ATAAACAAAT AAAAAGATTT ATGTAAGTTT
10401   TATGTATCCG CTCATGAGAC AATAACCCTG ATAAATGCTT CAATAATATT
        ATACATAGGC GAGTACTCTG TTATTGGGAC TATTTACGAA GTTATTATAA
10451   GAAAAAGGAA GAGTATGAGT ATTCAACATT TCCGTGTCGC CCTTATTCCC
        CTTTTTCCTT CTCATACTCA TAAGTTGTAA AGGCACAGCG GGAATAAGGG
10501   TTTTTTGCGG CATTTTGCCT TCCTGTTTTT GCTCACCCAG AAACGCTGGT
        AAAAAACGCC GTAAAACGGA AGGACAAAAA CGAGTGGGTC TTTGCGACCA
10551   GAAAGTAAAA GATGCTGAAG ATCAGTTGGG TGCACGAGTG GGTTACATCG
        CTTTCATTTT CTACGACTTC TAGTCAACCC ACGTGCTCAC CCAATGTAGC
10601   AACTGGATCT CAACAGCGGT AAGATCCTTG AGAGTTTTCG CCCCGAAGAA
        TTGACCTAGA GTTGTCGCCA TTCTAGGAAC TCTCAAAAGC GGGGCTTCTT
10651   CGTTTTCCAA TGATGAGCAC TTTTAAAGTT CTGCTATGTG GCGCGGTATT
        GCAAAAGGTT ACTACTCGTG AAAATTTCAA GACGATACAC CGCGCCATAA
10701   ATCCCGTATT GACGCCGGGC AAGAGCAACT CGGTCGCCGC ATACACTATT
        TAGGGCATAA CTGCGGCCCG TTCTCGTTGA GCCAGCGGCG TATGTGATAA
10751   CTCAGAATGA CTTGGTTGAG TACTCACCAG TCACAGAAAA GCATCTTACG
        GAGTCTTACT GAACCAACTC ATGAGTGGTC AGTGTCTTTT CGTAGAATGC
10801   GATGGCATGA CAGTAAGAGA ATTATGCAGT GCTGCCATAA CCATGAGTGA
        CTACCGTACT GTCATTCTCT TAATACGTCA CGACGGTATT GGTACTCACT
10851   TAACACTGCG GCCAACTTAC TTCTGACAAC GATCGGAGGA CCGAAGGAGC
        ATTGTGACGC CGGTTGAATG AAGACTGTTG CTAGCCTCCT GGCTTCCTCG
10901   TAACCGCTTT TTTGCACAAC ATGGGGGATC ATGTAACTCG CCTTGATCGT
        ATTGGCGAAA AAACGTGTTG TACCCCCTAG TACATTGAGC GGAACTAGCA
10951   TGGGAACCGG AGCTGAATGA AGCCATACCA AACGACGAGC GTGACACCAC
        ACCCTTGGCC TCGACTTACT TCGGTATGGT TTGCTGCTCG CACTGTGGTG
11001   GATGCCTGTA GCAATGGCAA CAACGTTGCG CAAACTATTA ACTGGCGAAC
        CTACGGACAT CGTTACCGTT GTTGCAACGC GTTTGATAAT TGACCGCTTG
11051   TACTTACTCT AGCTTCCCGG CAACAATTAA TAGACTGGAT GGAGGCGGAT
        ATGAATGAGA TCGAAGGGCC GTTGTTAATT ATCTGACCTA CCTCCGCCTA
11101   AAAGTTGCAG GACCACTTCT GCGCTCGGCC CTTCCGGCTG GCTGGTTTAT
        TTTCAACGTC CTGGTGAAGA CGCGAGCCGG GAAGGCCGAC CGACCAAATA
11151   TGCTGATAAA TCTGGAGCCG GTGAGCGTGG GTCTCGCGGT ATCATTGCAG
        ACGACTATTT AGACCTCGGC CACTCGCACC CAGAGCGCCA TAGTAACGTC
11201   CACTGGGGCC AGATGGTAAG CCCTCCCGTA TCGTAGTTAT CTACACGACG
        GTGACCCCGG TCTACCATTC GGGAGGGCAT AGCATCAATA GATGTGCTGC
11251   GGGAGTCAGG CAACTATGGA TGAACGAAAT AGACAGATCG CTGAGATAGG
        CCCTCAGTCC GTTGATACCT ACTTGCTTTA TCTGTCTAGC GACTCTATCC
11301   TGCCTCACTG ATTAAGCATT GGTAACTGTC AGACTCGCGA CACTGCATTA
        ACGGAGTGAC TAATTCGTAA CCATTGACAG TCTGAGCGCT GTGACGTAAT
11351   ATGAATCGGC CAACGCGCGG GGAGAGGCGG TTTGCGTATT GGGCGCTCTT
        TACTTAGCCG GTTGCGCGCC CCTCTCCGCC AAACGCATAA CCCGCGAGAA
11401   CCGCTTCCTC GCTCACTGAC TCGCTGCGCT CGGTCGTTCG GCTGCGGCGA
        GGCGAAGGAG CGAGTGACTG AGCGACGCGA GCCAGCAAGC CGACGCCGCT
11451   GCGGTATCAG CTCACTCAAA GGCGGTAATA CGGTTATCCA CAGAATCAGG
        CGCCATAGTC GAGTGAGTTT CCGCCATTAT GCCAATAGGT GTCTTAGTCC
```

```
11501   GGATAACGCA GGAAAGAACA TGTGAGCAAA AGGCCAGCAA AAGGCCAGGA
        CCTATTGCGT CCTTTCTTGT ACACTCGTTT TCCGGTCGTT TTCCGGTCCT
11551   ACCGTAAAAA GGCCGCGTTG CTGGCGTTTT TCCATAGGCT CCGCCCCCCT
        TGGCATTTTT CCGGCGCAAC GACCGCAAAA AGGTATCCGA GGCGGGGGGA
11601   GACGAGCATC ACAAAAATCG ACGCTCAAGT CAGAGGTGGC GAAACCCGAC
        CTGCTCGTAG TGTTTTTAGC TGCGAGTTCA GTCTCCACCG CTTTGGGCTG
11651   AGGACTATAA AGATACCAGG CGTTTCCCCC TGGAAGCTCC CTCGTGCGCT
        TCCTGATATT TCTATGGTCC GCAAAGGGGG ACCTTCGAGG GAGCACGCGA
11701   CTCCTGTTCC GACCCTGCCG CTTACCGGAT ACCTGTCCGC CTTTCTCCCT
        GAGGACAAGG CTGGGACGGC GAATGGCCTA TGGACAGGCG GAAAGAGGGA
11751   TCGGGAAGCG TGGCGCTTTC TCATAGCTCA CGCTGTAGGT ATCTCAGTTC
        AGCCCTTCGC ACCGCGAAAG AGTATCGAGT GCGACATCCA TAGAGTCAAG
11801   GGTGTAGGTC GTTCGCTCCA AGCTGGGCTG TGTGCACGAA CCCCCCGTTC
        CCACATCCAG CAAGCGAGGT TCGACCCGAC ACACGTGCTT GGGGGGCAAG
11851   AGCCCGACCG CTGCGCCTTA TCCGGTAACT ATCGTCTTGA GTCCAACCCG
        TCGGGCTGGC GACGCGGAAT AGGCCATTGA TAGCAGAACT CAGGTTGGGC
11901   GTAAGACACG ACTTATCGCC ACTGGCAGCA GCCACTGGTA ACAGGATTAG
        CATTCTGTGC TGAATAGCGG TGACCGTCGT CGGTGACCAT TGTCCTAATC
11951   CAGAGCGAGG TATGTAGGCG GTGCTACAGA GTTCTTGAAG TGGTGGCCTA
        GTCTCGCTCC ATACATCCGC CACGATGTCT CAAGAACTTC ACCACCGGAT
12001   ACTACGGCTA CACTAGAAGG ACAGTATTTG GTATCTGCGC TCTGCTGAAG
        TGATGCCGAT GTGATCTTCC TGTCATAAAC CATAGACGCG AGACGACTTC
12051   CCAGTTACCT TCGGAAAAAG AGTTGGTAGC TCTTGATCCG GCAAACAAAC
        GGTCAATGGA AGCCTTTTTC TCAACCATCG AGAACTAGGC CGTTTGTTTG
12101   CACCGCTGGT AGCGGTGGTT TTTTTGTTTG CAAGCAGCAG ATTACGCGCA
        GTGGCGACCA TCGCCACCAA AAAAACAAAC GTTCGTCGTC TAATGCGCGT
12151   GAAAAAAAGG ATCTCAAGAA GATCCTTTGA TCTTTTCTAC GGGGTCTGAC
        CTTTTTTTCC TAGAGTTCTT CTAGGAAACT AGAAAAGATG CCCCAGACTG
12201   GCTCAGTGGA ACGAAAACTC A
        CGAGTCACCT TGCTTTTGAG T
```

FIG. 27(contd.)

FIG. 28

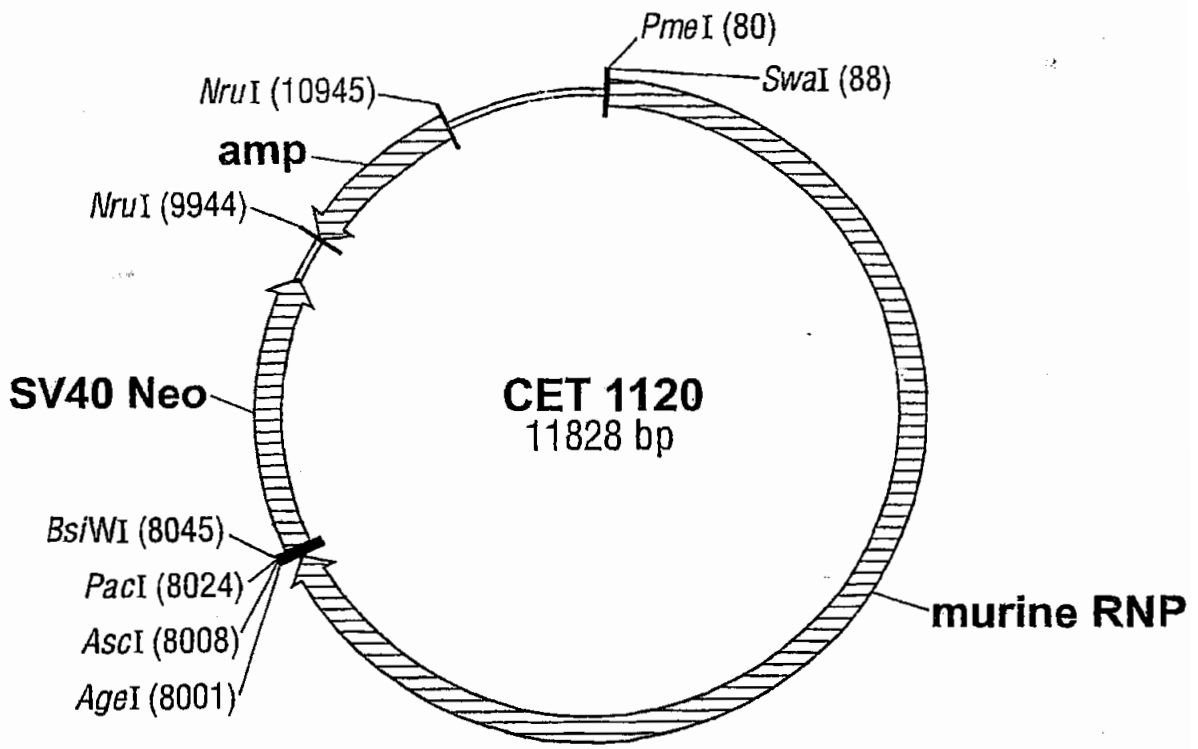

PmeI (80)
SwaI (88)
NruI (10945)
amp
NruI (9944)
SV40 Neo
CET 1120
11828 bp
BsiWI (8045)
PacI (8024)
AscI (8008)
AgeI (8001)
murine RNP

FIG. 29 CET1120 nucleotide sequence

```
   1  CGTTGTAAAA CGACGGCCAG TGAATTGTAA TACGACTCAC TATAGGGCGA
      GCAACATTTT GCTGCCGGTC ACTTAACATT ATGCTGAGTG ATATCCCGCT
  51  ATTGGGTACC GGGCCCCCCC TCGAAGTTTA AACATTTAAA TCTAGAAGCT
      TAACCCATGG CCCGGGGGGG AGCTTCAAAT TTGTAAATTT AGATCTTCGA
 101  TTTAACCCTC TATCCCTTTA AACTTCCTTG ATCCAGTGTA AGCACCTCCT
      AAATTGGGAG ATAGGGAAAT TTGAAGGAAC TAGGTCACAT TCGTGGAGGA
 151  AGAAAGTCAG TAGACAATAA AACAAAAGTT CTGCTTCACC GATTTACATT
      TCTTTCAGTC ATCTGTTATT TTGTTTTCAA GACGAAGTGG CTAAATGTAA
 201  TATAACCAAA TACCCTTCAC CAATACAATA AAAAAACAAA ACAACAAAAA
      ATATTGGTTT ATGGGAAGTG GTTATGTTAT TTTTTTGTTT TGTTGTTTTT
 251  ACCCCAACCA TCTGAGAAAT AATCTTCTCC TTTCCCAGCT TTATTCCCAG
      TGGGGTTGGT AGACTCTTTA TTAGAAGAGG AAAGGGTCGA AATAAGGGTC
 301  GATTCTACAT GACCAAATTA CCAGAGTCAC CACTCATTTT AATCACAACA
      CTAAGATGTA CTGGTTTAAT GGTCTCAGTG GTGAGTAAAA TTAGTGTTGT
 351  TAGTGTCAAA TAACTAGAAA ACATGAGACA ACAATGGAGA GCTGAGTAAC
      ATCACAGTTT ATTGATCTTT TGTACTCTGT TGTTACCTCT CGACTCATTG
 401  TATTAGTAGT AGTACTTTAC CAGAGAATGG CCTCTATAGG CTCACATGTA
      ATAATCATCA TCATGAAATG GTCTCTTACC GGAGATATCC GAGTGTACAT
 451  GGAATGGTTG GTCCCCAGGT GGTAGGTAGA GCTGTTTGAG GATTACGTGG
      CCTTACCAAC CAGGGGTCCA CCATCCATCT CGACAAACTC CTAATGCACC
 501  CCTTCTTGGA TGGGGGGTGG GGGTGGGGTG GGAGGGTTGG GTGGTGGGTA
      GGAAGAACCT ACCCCCCACC CCCACCCCAC CCTCCCAACC CACCACCCAT
 551  CTTAAGAGGT TTCAAAAGTC AATATTGTTT GCATTTAGCT CTTCCTTGTA
      GAATTCTCCA AAGTTTTCAG TTATAACAAA CGTAAATCGA GAAGGAACAT
 601  CTTGTGGATC AAACACAACC TGTCAGCTAC TGCTTCAAAT GTCATGCCTG
      GAACACCTAG TTTGTGTTGG ACAGTCGATG ACGAAGTTTA CAGTACGGAC
 651  CTGCCATCTT CTCAGCAGGA TGGTCATGGC CTCACCCTCT TCAACTGTAA
      GACGGTAGAA GAGTCGTCCT ACCAGTACCG GAGTGGGAGA AGTTGACATT
 701  ATCTTTCTTT CTTTTCTTCT TTTTCTTTTG GTTTCGAGAC AGGGTTTCTC
      TAGAAAGAAA GAAAAGAAGA AAAAGAAAAC CAAAGCTCTG TCCCAAAGAG
 751  TGTATAGTCC TGGCTGTCCT GGAACTCACT TTGTAGACCA GGCTGGCCTT
      ACATATCAGG ACCGACAGGA CCTTGAGTGA AACATCTGGT CCGACCGGAA
 801  GAACTCAGAA ATCCGCCTGC CTCTGCCTCC CTAGCACTGG GATTAAAGGC
      CTTGAGTCTT TAGGCGGACG GAGACGGAGG GATCGTGACC CTAATTTCCG
 851  GTGCGCCACC ACGCCCAGCT TTCAACTGGA AATCTTAATA AACTTTCCTA
      CACGCGGTGG TGCGGGTCGA AAGTTGACCT TTAGAATTAT TTGAAAGGAT
 901  GAAGTGGCCT TGGTTATGGG AGCTTATCAC AGCAATAGAA CAGCAATTAT
      CTTCACCGGA ACCAATACCC TCGAATAGTG TCGTTATCTT GTCGTTAATA
 951  GACTGGAGTA TGATAGTTAA AAACAAGCAA GCAAGCAAGC AAACACACAC
      CTGACCTCAT ACTATCAATT TTTGTTCGTT CGTTCGTTCG TTTGTGTGTG
1001  ACCAAAACAA CAAAACCCCA AGACAGAGTC ACATGTAGCC CAGGCTAGCC
      TGGTTTTGTT GTTTTGGGGT TCTGTCTCAG TGTACATCGG GTCCGATCGG
1051  TCCAAATTCA CTATATAACT GAAGAAGACC CCTAATTCCC ATTCCTCTAG
      AGGTTTAAGT GATATATTGA CTTCTTCTGG GGATTAAGGG TAAGGAGATC
1101  AATCTATACC TCAAGTACTG AATGGCTTGG TTCACAATAC CCCACTAAAT
      TTAGATATGG AGTTCATGAC TTACCGAACC AAGTGTTATG GGGTGATTTA
1151  GATTGGTCTT ACTAAGTGCA ACAAGGTAAA CCTAAAACTT CAGCCCTCAG
      CTAACCAGAA TGATTCACGT TGTTCCATTT GGATTTTGAA GTCGGGAGTC
1201  ACATCCCTTT TCCAGTATCA ATTTATAAAA TTAGATCCCA AGGATAAAAA
      TGTAGGGAAA AGGTCATAGT TAAATATTTT AATCTAGGGT TCCTATTTTT
1251  TTAATTGTAA AGTAAAATCA GAGTTCTAGC ATCAACTACA GGCTCAACCA
      AATTAACATT TCATTTTAGT CTCAAGATCG TAGTTGATGT CCGAGTTGGT
1301  TGGGGACCAC AAATAAACTA AAAGGGATAA GACTGGCTTC CCCATAATTA
      ACCCCTGGTG TTTATTTGAT TTTCCCTATT CTGACCGAAG GGGTATTAAT
```

FIG. 29(contd.) 78/93

```
1351  TTACATTTAG ATAATTTTCC TGACTACTCA ACAAAGCTAA AATATCACCA
      AATGTAAATC TATTAAAAGG ACTGATGAGT TGTTTCGATT TTATAGTGGT
1401  CTGGTTTATT TTCTCCTTCT AGGGTTTAAG CTCACTCTGA GGAGGGGCAT
      GACCAAATAA AAGAGGAAGA TCCCAAATTC GAGTGAGACT CCTCCCCGTA
1451  GCGGCACACA CTCATAGCAT CCAGGAAATA GAAATATGGT GACTATCATG
      CGCCGTGTGT GAGTATCGTA GGTCCTTTAT CTTTATACCA CTGATAGTAC
1501  GGTTCAGGGC CAACCTAGGC TTTAGAGAAA AACCTTGTCC CACAAACCAA
      CCAAGTCCCG GTTGGATCCG AAATCTCTTT TTGGAACAGG GTGTTTGGTT
1551  AAATGTCTCT TTTTTATTCT ATCAGGGGTG GATGGATTTG TTAAAGAAGT
      TTTACAGAGA AAAAATAAGA TAGTCCCCAC CTACCTAAAC AATTTCTTCA
1601  GCTTTTAAAA ACCTTGAGAT GGTTATTTAG AAGTCCCCAT GGGATACCAA
      CGAAAATTTT TGGAACTCTA CCAATAAATC TTCAGGGGTA CCCTATGGTT
1651  AATAACCCAC TATTTATATG CCCAAGCATT TCACCTCCAC AACAGTGCTA
      TTATTGGGTG ATAAATATAC GGGTTCGTAA AGTGGAGGTG TTGTCACGAT
1701  TGCACCCTTT AACATTTTTG AGACAGTAGC CCAGTCTAGT CTTTAACTTG
      ACGTGGGAAA TTGTAAAAAC TCTGTCATCG GGTCAGATCA GAAATTGAAC
1751  CAGTGATTTT TCCTGATTCA GCTTCTCCCA GTGCTGGAAT TATAGGTATG
      GTCACTAAAA AGGACTAAGT CGAAGAGGGT CACGACCTTA ATATCCATAC
1801  CACCACCATG TGTAACTACA GATGCTACTT AAAAATTTTT TAAAGGAATC
      GTGGTGGTAC ACATTGATGT CTACGATGAA TTTTTAAAAA ATTTCCTTAG
1851  ACAAAAATAA CCCCCTATCA AATGCCTAGT CCCTCTAACC ATCACCAAGT
      TGTTTTTATT GGGGGATAGT TTACGGATCA GGGAGATTGG TAGTGGTTCA
1901  GAAGGATCAC GCAGGAAAAA AAAAATCACC AGCAGCACCT CAGAACCAGG
      CTTCCTAGTG CGTCCTTTTT TTTTTAGTGG TCGTCGTGGA GTCTTGGTCC
1951  ATACTCAGTC CATCAGCATC CAGGGCCATA CCCACACTCA CAGCATCTCC
      TATGAGTCAG GTAGTCGTAG GTCCCGGTAT GGGTGTGAGT GTCGTAGAGG
2001  ACAGTTTACC AGATGATTCA TGCTTATCAC TGTATTGGGT CATCTAAGAG
      TGTCAAATGG TCTACTAAGT ACGAATAGTG ACATAACCCA GTAGATTCTC
2051  TGACCATCAG GGCTTCTGAT CACAGAATCT AGTCCACTTT GCAGACCAGT
      ACTGGTAGTC CCGAAGACTA GTGTCTTAGA TCAGGTGAAA CGTCTGGTCA
2101  TGAAGTCATG CACTATATGA GATAGAAATA CCCTCTTGCT CATTTTGGTC
      ACTTCAGTAC GTGATATACT CTATCTTTAT GGGAGAACGA GTAAAACCAG
2151  AGAAATTCAA GGATAAAAAC CCATGTTTTG TTAATGCACA CCTCCATATG
      TCTTTAAGTT CCTATTTTTG GGTACAAAAC AATTACGTGT GGAGGTATAC
2201  ATTGAGATCA ATGTGTCCTA ATTAATGTAG AAACCACAAC TGTAAATTTC
      TAACTCTAGT TACACAGGAT TAATTACATC TTTGGTGTTG ACATTTAAAG
2251  ACTCTTTTGA CATGAATCTT TTTCTAGACA GGGTCTTGGA TGCAGCCCCG
      TGAGAAAACT GTACTTAGAA AAAGATCTGT CCCAGAACCT ACGTCGGGGC
2301  ACTACCCAGA ATTTTGGAAT CCAGGCTAGC CTCAAACTCA AGGCAATCTG
      TGATGGGTCT TAAAACCTTA GGTCCGATCG GAGTTTGAGT TCCGTTAGAC
2351  CTTGCTTCAG CTTCTCACAG GCTGGATCAC AAACATACAC CTTCAGACCC
      GAACGAAGTC GAAGAGTGTC CGACCTAGTG TTTGTATGTG GAAGTCTGGG
2401  ATTTTTTTTT CCTCCCTCCG TTTTTGGTTT CTCTGTGTAG CCCTGGGTGT
      TAAAAAAAAA GGAGGGAGGC AAAAACCAAA GAGACACATC GGGACCCACA
2451  CCGTGGACTC GCTGTGTAGA TCTATCTACC AGCCTCTGTC TTGGAGTACT
      GGCACCTGAG CGACACATCT AGATAGATGG TCGGAGACAG AACCTCATGA
2501  GGGATTAAAG TTGTGGGCTA CCACTGCCTG GCTGACCCAG TTTTATTTAT
      CCCTAATTTC AACACCCGAT GGTGACGGAC CGACTGGGTC AAAATAAATA
2551  TTTAAATATA ACTTGACAAA AATAAATTTG TCTAACTTAC TAGAAATCCC
      AAATTTATAT TGAACTGTTT TTATTTAAAC AGATTGAATG ATCTTTAGGG
2601  AAGAAAACTA ACACTGGATT TAGCAACAGT CAGAAATCGC TGAAAAGAAA
      TTCTTTTGAT TGTGACCTAA ATCGTTGTCA GTCTTTAGCG ACTTTTCTTT
2651  CAGAATTGAT CTAACAGTCT TAGATCACTC CTAGACAGTT TGTAATTCTT
      GTCTTAACTA GATTGTCAGA ATCTAGTGAG GATCTGTCAA ACATTAAGAA
2701  GCTCATGGCA ACGTGAGCTC TATCTAACTC ACTCTCTGTG CACTAATGAA
      CGAGTACCGT TGCACTCGAG ATAGATTGAG TGAGAGACAC GTGATTACTT
2751  TGCTCAGTGT CTCCAGAACA GCACAGCTTC CAGGGTAATC ATGCCAACCC
      ACGAGTCACA GAGGTCTTGT CGTGTCGAAG GTCCCATTAG TACGGTTGGG
```

FIG. 29(contd.)


```
2801  ACAAGACTTT TATAGAGCTG TCCACGACTC TTCCCCCATT CAGCTCATTA
      TGTTCTGAAA ATATCTCGAC AGGTGCTGAG AAGGGGGTAA GTCGAGTAAT
2851  ACAATATGAT GGAGCTCCTG TGTGGAAATC AAGGCACACT CTGGTAGAAA
      TGTTATACTA CCTCGAGGAC ACACCTTTAG TTCCGTGTGA GACCATCTTT
2901  CTTGTTTTTT CTTTCCACTT TTCCTTGGGC TCTGAAGATT GAGCTGTTTT
      GAACAAAAAA GAAAGGTGAA AAGGAACCCG AGACTTCTAA CTCGACAAAA
2951  ATAACCCACA AACATGCATT TTTTACCTCA AAAGCATCCA GCAAAAACTG
      TATTGGGTGT TTGTACGTAA AAAATGGAGT TTTCGTAGGT CGTTTTTGAC
3001  TACAACGCTT TTTCAAAAAA ATGTATTGTG ATCCTCCTTA AGAAAAGCCT
      ATGTTGCGAA AAAGTTTTTT TACATAACAC TAGGAGGAAT TCTTTTCGGA
3051  TACTTAGTGT TAATTCCTTT TTCTTTAGAA TGCTGGTAAA TACAAGGACT
      ATGAATCACA ATTAAGGAAA AAGAAATCTT ACGACCATTT ATGTTCCTGA
3101  TAGGTAGGCT GGCTTCTAAC AGCAATTCAC CCACTTATGA TGGGATTAAA
      ATCCATCCGA CCGAAGATTG TCGTTAAGTG GGTGAATACT ACCCTAATTT
3151  GGAAGGCACA ACCATGTCCA CCACAGGTTC TAGCTCCCCC ACCCACACGC
      CCTTCCGTGT TGGTACAGGT GGTGTCCAAG ATCGAGGGGG TGGGTGTGCG
3201  CCAGAGAGGG TTTTTCTGTG TAGCTCTGAC TATTCTGGAA TTCACACTGC
      GGTCTCTCCC AAAAAGACAC ATCGAGACTG ATAAGACCTT AAGTGTGACG
3251  AGACCAGGCT GGTCTCGAAC TCAGAGATCC ACCACCACAT GGTTTCTTAA
      TCTGGTCCGA CCAGAGCTTG AGTCTCTAGG TGGTGGTGTA CCAAAGAATT
3301  TTGTAATTTT AAAGAAAAAA AAAAATCCTT CAGTTAAGAT TCTTATGTTC
      AACATTAAAA TTTCTTTTTT TTTTTAGGAA GTCAATTCTA AGAATACAAG
3351  TAGGTTTTCA CAAACTTACC AATGTAGTTT TATTGGAGGC CATTTTTTAA
      ATCCAAAAGT GTTTGAATGG TTACATCAAA ATAACCTCCG GTAAAAAATT
3401  ATTTAATCGG AGACTTGAAG AGCTATTGCA AGAAAAAAAA TGTAGGACAG
      TAAATTAGCC TCTGAACTTC TCGATAACGT TCTTTTTTTT ACATCCTGTC
3451  TTAAAATTTC ATGACACACA AAAGGCAGCT ACAAGTTTTG TGTGGATTTC
      AATTTTAAAG TACTGTGTGT TTTCCGTCGA TGTTCAAAAC ACACCTAAAG
3501  AACATGTAAA TTTCGGGTAA AAATGCAGGA AAACAGTTGA GTTCCCGTGT
      TTGTACATTT AAAGCCCATT TTTACGTCCT TTTGTCAACT CAAGGGCACA
3551  TATTAGTATG TTACTAATAA TTTCAGTATG TTAGTGAAAA TAATCTTACT
      ATAATCATAC AATGATTATT AAAGTCATAC AATCACTTTT ATTAGAATGA
3601  AAAACACTGG TACCTCAGAC AACTTTACAT GGTGAGGATT GTTACTTTCC
      TTTTGTGACC ATGGAGTCTG TTGAAATGTA CCACTCCTAA CAATGAAAGG
3651  CAATCCATAT AGAATTTTAA CAATTTTAGT GTTTATTTTG GATGAAAGGA
      GTTAGGTATA TCTTAAAATT GTTAAAATCA CAAATAAAAC CTACTTTCCT
3701  AATGACTATC TTTTGTTAGC AAATTACCAT AAGATCTTTT TCTTTAGATT
      TTACTGATAG AAAACAATCG TTTAATGGTA TTCTAGAAAA AGAAATCTAA
3751  TCTGAATACT CCAAGGAGCT CATATAATTC CATCCTTATT TTTTCAGAGG
      AGACTTATGA GGTTCCTCGA GTATATTAAG GTAGGAATAA AAAAGTCTCC
3801  CCCTCCCTGT TCAATCACGG TATAAAAAAA GGAACACATT AAGATGTCCC
      GGGAGGGACA AGTTAGTGCC ATATTTTTTT CCTTGTGTAA TTCTACAGGG
3851  AGTCCTATTT TCTGGCTTTT TTTTTCCGGG GGTGGTGGTG CGGTAATCAC
      TCAGGATAAA AGACCGAAAA AAAAAGGCCC CCACCACCAC GCCATTAGTG
3901  TCTCTATAGT CCAGTCTGGG CTTCAACGCC TGGCAATCCC CAGCCTCAAG
      AGAGATATCA GGTCAGACCC GAAGTTGCGG ACCGTTAGGG GTCGGAGTTC
3951  CTCCCAAGTA CTGTCCTGAT AAGGATAGAA GGAGTCGACC TCCTTCACGC
      GAGGGTTCAT GACAGGACTA TTCCTATCTT CCTCAGCTGG AGGAAGTGCG
4001  TCCCCTCCGA GGAGGGCTCC TTCCCAGCTC CATTCCCCGG TCGGGAGCCC
      AGGGGAGGCT CCTCCCGAGG AAGGGTCGAG GTAAGGGGCC AGCCCTCGGG
4051  GTCCCCCACC CGAGAGCGCG GGCCTCGTGG TCAGCGCCTC CGCGGGGAGA
      CAGGGGGTGG GCTCTCGCGC CCGGAGCACC AGTCGCGGAG GCGCCCCTCT
4101  AACAAAGGCG GCGGCGGGGG CTCAAGGGCA CTGCGCCACG GGCCCGCGCC
      TTGTTTCCGC CGCCGCCCCC GAGTTCCCGT GACGCGGTGC CCGGGCGCGG
4151  TCCCCCATCC GGCGGCGGCC ACGTAGCCGG GAGCGCGCCG CAGCCCGGAG
      AGGGGGTAGG CCGCCGCCGG TGCATCGGCC CTCGCGCGGC GTCGGGCCTC
4201  CCTCGGGCCT CGCAGCTGCA GAGCCTGAAC CGCTCTCTCC CTGCGGGCCT
      GGAGCCCGGA GCGTCGACGT CTCGGACTTG GCGAGAGAGG GACGCCCGGA
```

FIG. 29(contd.) 80/93

```
4251  GCGACGAGGC TGGGGGAGGG GAGGCCCGCG CTTTGTCTGG AGTCTCGGTA
      CGCTGCTCCG ACCCCCTCCC CTCCGGGCGC GAAACAGACC TCAGAGCCAT
4301  GCTGTCATCC GGCTCCCACC CTCATGCACA ATTGTCCCAT CTCCCCCACG
      CGACAGTAGG CCGAGGGTGG GAGTACGTGT TAACAGGGTA GAGGGGGTGC
4351  CACCGGCGCG GCGCCCGCCT CAGCGAGGCC CCAGCCGGTT TCCCGCAGCC
      GTGGCCGCGC CGCGGGCGGA GTCGCTCCGG GGTCGGCCAA AGGGCGTCGG
4401  CGCGGCCCAC GGGGCTCGCA GCCTCCCCGC AAGCTCGGAC GCACGGAGCA
      GCGCCGGGTG CCCCGAGCGT CGGAGGGGCG TTCGAGCCTG CGTGCCTCGT
4451  TCCTAAACCC CACCACACGC AAGATCGAAA AAAAGCAAAG GCACGAACTT
      AGGATTTGGG GTGGTGTGCG TTCTAGCTTT TTTTCGTTTC CGTGCTTGAA
4501  CACCGCTCCG ATGCTCAGGG CCGCGGATCC TGCAGAGTCT CCCGCCTGCG
      GTGGCGAGGC TACGAGTCCC GGCGCCTAGG ACGTCTCAGA GGGCGGACGC
4551  CGCTTCGGTT CAGCCACATC CGAGGGGAGG GGGCGCGGGC AGCTCCGCCG
      GCGAAGCCAA GTCGGTGTAG GCTCCCCTCC CCCGCGCCCG TCGAGGCGGC
4601  GGGGGGAGGG GGAGCACCGC CCACGCCCTG GCCGCGCGGG GCCCGCCGGG
      CCCCCCTCCC CCTCGTGGCG GGTGCGGGAC CGGCGCGCCC CGGGCGGCCC
4651  AACGCGTCCT GCGGGGGGCG GCGCGCGCAA TGCTCACCGT CCGCGGCGTG
      TTGCGCAGGA CGCCCCCCGC CGCGCGCGTT ACGAGTGGCA GGCGCCGCAC
4701  GCGCCCAGGG GGTCTCCTGG CTGGGGGGAG GGGGGGGAAG GCGGGCAGGA
      CGCGGGTCCC CCAGAGGACC GACCCCCCTC CCCCCCCTTC CGCCCGTCCT
4751  AGGACCGCGG AGGCCTCTCT GCGTCTCGGA GCGCGCCAAA GCGGGGCTCC
      TCCTGGCGCC TCCGGAGAGA CGCAGAGCCT CGCGCGGTTT CGCCCCGAGG
4801  ACCCACCTCC TTGCCCGGAT CTTGAAGGCC GGGGAGATAA ACAGCGGGGT
      TGGGTGGAGG AACGGGCCTA GAACTTCCGG CCCCTCTATT TGTCGCCCCA
4851  TCTTTAAGCA CCACCTCTCA CTAGGCGCGG GATCCCAAGG CTTGTGGCAT
      AGAAATTCGT GGTGGAGAGT GATCCGCGCC CTAGGGTTCC GAACACCGTA
4901  CCGGGGTGGT ACTTGGACTA AAAGTCCTTC TGGGAGGGAC CGAGTGAGAA
      GGCCCCACCA TGAACCTGAT TTTCAGGAAG ACCCTCCCTG GCTCACTCTT
4951  CCCCTTTGGG ACGTGTAGAA ATATTTGTGT GGTTCGAGAA TATTTGTGCG
      GGGGAAACCC TGCACATCTT TATAAACACA CCAAGCTCTT ATAAACACGC
5001  GACGGGCTTG GCAAAGGCGT AGCTGCAGAG AGCACGCTTG GGTGGAGAGG
      CTGCCCGAAC CGTTTCCGCA TCGACGTCTC TCGTGCGAAC CCACCTCTCC
5051  GCCGCACGCC CCAGCGCCGG CCTAAGCCCC TCCCGACGGC GTTATTTCAA
      CGGCGTGCGG GGTCGCGGCC GGATTCGGGG AGGGCTGCCG CAATAAAGTT
5101  ACTGCGCGAC CGTTTCTCCG CTCCCTACGC GGAGGTGGGG GCCGGACCTA
      TGACGCGCTG GCAAAGAGGC GAGGGATGCG CCTCCACCCC CGGCCTGGAT
5151  GTTCCGGACG TAGTAACACG CCGAGCGCGA GCCTTCCGCA ATTCACGGAA
      CAAGGCCTGC ATCATTGTGC GGCTCGCGCT CGGAAGGCGT TAAGTGCCTT
5201  CACAGTTGCG CAAGTGATGT AAAGCAGTCC CGCTGTACCT AAAGGGGGAG
      GTGTCAACGC GTTCACTACA TTTCGTCAGG GCGACATGGA TTTCCCCCTC
5251  TGTCACGTAC TTGGCGTAAG GAGAGTGTAG GCCCTTCCCG CCATTGGCGG
      ACAGTGCATG AACCGCATTC CTCTCACATC CGGGAAGGGC GGTAACCGCC
5301  CGGTTAGGGC GTTTACGTAA CGGCGTGACG TAAGCGGAGA CGCGTTAGTG
      GCCAATCCCG CAAATGCATT GCCGCACTGC ATTCGCCTCT GCGCAATCAC
5351  GGGGGAAGGT TCTAGAAAAG CGGCGGTCTC GGCTCCAGCG GCAGTAGCAG
      CCCCCTTCCA AGATCTTTTC GCCGCCAGAG CCGAGGTCGC CGTCATCGTC
5401  CGGCGCCGGT CCCGTGTGCA GGAGCTCCTT TGCGGCCCAG TTTCTTGGCC
      GCCGCGGCCA GGGCACACGT CCTCGAGGAA ACGCCGGGTC AAAGAACCGG
5451  ATCGCCTGCT CTCCCCACAG CGCCAGGACG AGTCCCGTGC GCGTCCGTCC
      TAGCGGACGA GAGGGGTGTC GCGGTCCTGC TCAGGGCACG CGCAGGCAGG
5501  GCGGAGGTCT TTCTCATCTC GCTCGGCTGC GGGAAATCGG GCTGAAGCGA
      CGCCTCCAGA AAGAGTAGAG CGAGCCGACG CCCTTTAGCC CGACTTCGCT
5551  CTGAGTCCGC GATGGAGGTA ACGGGTTTGA AATCAATGAG TTATTAAAAA
      GACTCAGGCG CTACCTCCAT TGCCCAAACT TTAGTTACTC AATAATTTTT
5601  TGGCATGGCG AGGCCGTAGG CACCGCAATG GAAACCGGCC ACCCGCCTCC
      ACCGTACCGC TCCGGCATCC GTGGCGTTAC CTTTGGCCGG TGGGCGGAGG
5651  GTGGTCCGGC GGAGGGGATG CGGCCACTCG AGTGGCGGTT GGCCTTGGCG
      CACCAGGCCG CCTCCCCTAC GCCGGTGAGC TCACCGCCAA CCGGAACCGC
```

FIG. 29(contd.) 81/93

```
5701  AGTTTCTGAG GGGTCGTTGG AGGAGGCCTC TGATTGTCCG ACCGCCTTCC
      TCAAAGACTC CCCAGCAACC TCCTCCGGAG ACTAACAGGC TGGCGGAAGG
5751  CCGCCCTCAG CCGCCCGGCG CCATTTCCCT CAGTTGGGGT GGGGGATGGG
      GGCGGGAGTC GGCGGGCCGC GGTAAAGGGA GTCAACCCCA CCCCCTACCC
5801  AAGTGCCCGC CGCGACCGGG CTGGACCGCT AAAGTAGCGC GTGAGCGGGC
      TTCACGGGCG GCGCTGGCCC GACCTGGCGA TTTCATCGCG CACTCGCCCG
5851  CATCGCTGGC CTTTCGATGT GCGCGGGCCT AGGGGCTCGG TTGTGTTCGC
      GTAGCGACCG GAAAGCTACA CGCGCCCGGA TCCCCGAGCC AACACAAGCG
5901  GGCGGAACGT TTCTGGGGCC CCCCCGGCTT CCCGGAGCGA GTCTGCGAAG
      CCGCCTTGCA AAGACCCCGG GGGGGCCGAA GGGCCTCGCT CAGACGCTTC
5951  CTAGCTTCCC CTCCCCCCTC TCCCGGGAAC CGGATTTGGC GGCCGCCATT
      GATCGAAGGG GAGGGGGGAG AGGGCCCTTG GCCTAAACCG CCGGCGGTAA
6001  TTCCCGTCTC CTTCCTCGCC ACGATTTTGC TTTCAACGCT TTAGGTTTAC
      AAGGGCAGAG GAAGGAGCGG TGCTAAAACG AAAGTTGCGA AATCCAAATG
6051  TAGTTTGGTT TTCTTTTTTC ACCACTGCGT AGACGTGTTT AGCGATTTTC
      ATCAAACCAA AAGAAAAAAG TGGTGACGCA TCTGCACAAA TCGCTAAAAG
6101  CTTTCTTTTG GAAGTCTTCA TACCGTTTCG AGGTGGATTT AGCGTTTTGA
      GAAAGAAAAC CTTCAGAAGT ATGGCAAAGC TCCACCTAAA TCGCAAAACT
6151  GCTTGGGTCT TCAGCGTCCT GCGCACCTCG CTAAAGGCTC TCTGCCTTCC
      CGAACCCAGA AGTCGCAGGA CGCGTGGAGC GATTTCCGAG AGACGGAAGG
6201  CCTCGACGAA ATGGCGCCAT TGCTTTCTGA AGCCACCGAG GCGCGGGGTG
      GGAGCTGCTT TACCGCGGTA ACGAAAGACT TCGGTGGCTC CGCGCCCCAC
6251  GGGGCGGGGT GGCGGCGCTC CACGAGCTTT ACTGGAACAG GCAGAGAGAA
      CCCCGCCCCA CCGCCGCGAG GTGCTCGAAA TGACCTTGTC CGTCTCTCTT
6301  CGTAGTACAA CCGAGGCCTG GGCGGGTGGC TGAAGGCAGC GTCGCTGCAA
      GCATCATGTT GGCTCCGGAC CCGCCCACCG ACTTCCGTCG CAGCGACGTT
6351  AGAGACCGTT TTATTTTTCA TAATACGTAA GATTACGGGT GCTGTAGTAA
      TCTCTGGCAA AATAAAAAGT ATTATGCATT CTAATGCCCA CGACATCATT
6401  AGCACTTGAG CATTAGTATA GTAGGAGGAA GTCAAAGTGG AAAAAATGGG
      TCGTGAACTC GTAATCATAT CATCCTCCTT CAGTTTCACC TTTTTTACCC
6451  AGCGCTCATC AGGAAGCTAG GGAGGCTATG TTGAGTGCAG GGTTACTTTC
      TCGCGAGTAG TCCTTCGATC CCTCCGATAC AACTCACGTC CCAATGAAAG
6501  CTTTTATTGC AGAACTTTTA TCTGCTTAAA GGATCCTCGG ATCGAAATAA
      GAAAATAACG TCTTGAAAAT AGACGAATTT CCTAGGAGCC TAGCTTTATT
6551  TTCAAATTAT AAGCATTTTT AAGGGAATCT TCGAATTTGT TGGTAAAGTC
      AAGTTTAATA TTCGTAAAAA TTCCCTTAGA AGCTTAAACA ACCATTTCAG
6601  AACGGATCCT TAGCACGTGG TGTTCACTTT AAGGAAGTGA AATAGCTGAC
      TTGCCTAGGA ATCGTGCACC ACAAGTGAAA TTCCTTCACT TTATCGACTG
6651  TTTTCATAGT TAGCCTTCGC TTAAAGCCTG GTTCAGTGGA CGAAAATCCA
      AAAAGTATCA ATCGGAAGCG AATTTCGGAC CAAGTCACCT GCTTTTAGGT
6701  CGTCCTGGCT ATATAAAAAC TTAGTTTGGG GTCACAGTGT TTGAGCGTGG
      GCAGGACCGA TATATTTTTG AATCAAACCC CAGTGTCACA AACTCGCACC
6751  TCATTCGGTT TTTTTATTTT TTATTTGTTT GAAATTATGA TGCATCATTA
      AGTAAGCCAA AAAAATAAAA AATAAACAAA CTTTAATACT ACGTAGTAAT
6801  CACTGATAAG CATTAGCTTT CGAATTGAAA GGGGTCTCCT TGGTTATTTT
      GTGACTATTC GTAATCGAAA GCTTAACTTT CCCCAGAGGA ACCAATAAAA
6851  CTTTGACTCT AAGCACACTT ATAAATAAAA TAACCTTGTT TATAATCGAT
      GAAACTGAGA TTCGTGTGAA TATTTATTTT ATTGGAACAA ATATTAGCTA
6901  AGTGGACGTC TGGTAAGTTT GGAAAAAACC CGAGGTAAGT AAAGAGCTTT
      TCACCTGCAG ACCATTCAAA CCTTTTTTGG GCTCCATTCA TTTCTCGAAA
6951  TGCTTTCGTT AGTGATATGA AAAAACAAGG TGTATTTAAT ACTTGCAACT
      ACGAAAGCAA TCACTATACT TTTTTGTTCC ACATAAATTA TGAACGTTGA
7001  TAGTTTAAGG AAAGCCAATT TACTGACATT TTAGTAGAGC TACCAGAAAC
      ATCAAATTCC TTTCGGTTAA ATGACTGTAA AATCATCTCG ATGGTCTTTG
7051  ACTATTTGGA GTCCTGATTA AGGCTTTTGT AACTATTTTG ACTATTTAAA
      TGATAAACCT CAGGACTAAT TCCGAAAACA TTGATAAAAC TGATAAATTT
7101  ACAATTTTGG TCGTTTTTAT TAAACATTTC AAAACCTAAA AATTGTAAAC
      TGTTAAAACC AGCAAAAATA ATTTGTAAAG TTTTGGATTT TTAACATTTG
```

FIG. 29(contd.)

```
7151  ATTGGCTTTT TGAGCACATT TTGGAGAAAC TTACAAATTT AGGCTATACA
      TAACCGAAAA ACTCGTGTAA AACCTCTTTG AATGTTTAAA TCCGATATGT
7201  GTAAAATAAC GGATTTGTTT TATAATTTTG CTTTTTCATT TCGTTGTGCA
      CATTTTATTG CCTAAACAAA ATATTAAAAC GAAAAAGTAA AGCAACACGT
7251  GTCATAGGTC CTGGATAGTA TGACCTAATT TATGAACATC TTGATAAGTT
      CAGTATCCAG GACCTATCAT ACTGGATTAA ATACTTGTAG AACTATTCAA
7301  TTTGTACTTA GCTATTGGAA AGCCAGTATT AAGTGCCTGA CAAAACCAGA
      AAACATGAAT CGATAACCTT TCGGTCATAA TTCACGGACT GTTTTGGTCT
7351  TTTAAGGTGA TATCTGGAGT TTCAGCATTC TTCATGGAGC TTGTTTCAGA
      AAATTCCACT ATAGACCTCA AAGTCGTAAG AAGTACCTCG AACAAAGTCT
7401  GTTGCAGGAT TTTTTTTTTT CATCTTGAGA TACTTACAAT TAACACCAGA
      CAACGTCCTA AAAAAAAAAA GTAGAACTCT ATGAATGTTA ATTGTGGTCT
7451  GGGGGCAGCT CAGGGAAAAG CAAATATGCC ACTTTTCAGA AACTGAATCT
      CCCCCGTCGA GTCCCTTTTC GTTTATACGG TGAAAAGTCT TTGACTTAGA
7501  TGGAAGTGGT GAATTTGGAA ACAGGTTTTT TAAATTTTTT TTAAATCTAA
      ACCTTCACCA CTTAAACCTT TGTCCAAAAA ATTTAAAAAA AATTTAGATT
7551  AAAGTAGTAA ATTTTGGACT TGGGTTGTAG AATTTAATGA ATTACAAAAG
      TTTCATCATT TAAAACCTGA ACCCAACATC TTAAATTACT TAATGTTTTC
7601  AATTCTTTAA TACCCTTTAA ATGACCTAAG AGCTGGGTAT GGTTTTTCTG
      TTAAGAAATT ATGGGAAATT TACTGGATTC TCGACCCATA CCAAAAAGAC
7651  AATTTTTTTG AAGAAAATCT AAGAAAGTTT ACGTGAATTA GAAGTTAGAT
      TTAAAAAAAC TTCTTTTAGA TTCTTTCAAA TGCACTTAAT CTTCAATCTA
7701  CGAATATTAG TGACTTTGAA ACTTGTATAG CTCAGGCAAT TTTTGGTGTA
      GCTTATAATC ACTGAAACTT TGAACATATC GAGTCCGTTA AAAACCACAT
7751  ACACAACTAA TATGCAGTTT AACATATGGT TTAAATTTGA TGTAAGTTTT
      TGTGTTGATT ATACGTCAAA TTGTATACCA AATTTAAACT ACATTCAAAA
7801  TTTTCTCCCC CCCAGAAAAC TTTAGAAACT GTTCCTTTGG AGAGGAAAAA
      AAAAGAGGGG GGGTCTTTTG AAATCTTTGA CAAGGAAACC TCTCCTTTTT
7851  GGTACTCTAC CAGCAGGTCA CCTCATATTT AAGAATTTAA TTTCCTGCAT
      CCATGAGATG GTCGTCCAGT GGAGTATAAA TTCTTAAATT AAAGGACGTA
7901  ACAAAGAAAG TGTAAATAAA AATTGAAATG GTATTTCCCT TTGCAGAGAG
      TGTTTCTTTC ACATTTATTT TTAACTTTAC CATAAAGGGA AACGTCTCTC
7951  AAAAGGAACA GTTCCGAAAG CTCTTTATTG GTGGCTTAAG CTTATCGATA
      TTTTCCTTGT CAAGGCTTTC GAGAAATAAC CACCGAATTC GAATAGCTAT
8001  CCGGTGGCGC GCCAATTGTT AATTAAGATC TGGCCCAATG GGCCGTACGA
      GGCCACCGCG CGGTTAACAA TTAATTCTAG ACCGGGTTAC CCGGCATGCT
8051  ATTTGAGGCG GAAAGAACCA GCTGTGGAAT GTGTGTCAGT TAGGGTGTGG
      TAAACTCCGC CTTTCTTGGT CGACACCTTA CACACAGTCA ATCCCACACC
8101  AAAGTCCCCA GGCTCCCCAG CAGGCAGAAG TATGCAAAGC ATGCATCTCA
      TTTCAGGGGT CCGAGGGGTC GTCCGTCTTC ATACGTTTCG TACGTAGAGT
8151  ATTAGTCAGC AACCAGGTGT GGAAAGTCCC CAGGCTCCCC AGCAGGCAGA
      TAATCAGTCG TTGGTCCACA CCTTTCAGGG GTCCGAGGGG TCGTCCGTCT
8201  AGTATGCAAA GCATGCATCT CAATTAGTCA GCAACCATAG TCCCGCCCCT
      TCATACGTTT CGTACGTAGA GTTAATCAGT CGTTGGTATC AGGGCGGGGA
8251  AACTCCGCCC ATCCCGCCCC TAACTCCGCC CAGTTCCGCC CATTCTCCGC
      TTGAGGCGGG TAGGGCGGGG ATTGAGGCGG GTCAAGGCGG GTAAGAGGCG
8301  CCCATGGCTG ACTAATTTTT TTTATTTATG CAGAGGCCGA GGCCGCCTCG
      GGGTACCGAC TGATTAAAAA AAATAAATAC GTCTCCGGCT CCGGCGGAGC
8351  GCCTCTGAGC TATTCCAGAA GTAGTGAGGA GGCTTTTTTG GAGGCCTAGG
      CGGAGACTCG ATAAGGTCTT CATCACTCCT CCGAAAAAAC CTCCGGATCC
8401  CTTTTGCAAA GATCGATCAA GAGACAGGAT GAGGATCGTT TCGCATGATT
      GAAAACGTTT CTAGCTAGTT CTCTGTCCTA CTCCTAGCAA AGCGTACTAA
8451  GAACAAGATG GATTGCACGC AGGTTCTCCG GCCGCTTGGG TGGAGAGGCT
      CTTGTTCTAC CTAACGTGCG TCCAAGAGGC CGGCGAACCC ACCTCTCCGA
8501  ATTCGGCTAT GACTGGGCAC AACAGACAAT CGGCTGCTCT GATGCCGCCG
      TAAGCCGATA CTGACCCGTG TTGTCTGTTA GCCGACGAGA CTACGGCGGC
8551  TGTTCCGGCT GTCAGCGCAG GGGCGCCCGG TTCTTTTTGT CAAGACCGAC
      ACAAGGCCGA CAGTCGCGTC CCCGCGGGCC AAGAAAAACA GTTCTGGCTG
```

FIG. 29(contd.)


```
8601  CTGTCCGGTG CCCTGAATGA ACTGCAAGAC GAGGCAGCGC GGCTATCGTG
      GACAGGCCAC GGGACTTACT TGACGTTCTG CTCCGTCGCG CCGATAGCAC
8651  GCTGGCCACG ACGGGCGTTC CTTGCGCAGC TGTGCTCGAC GTTGTCACTG
      CGACCGGTGC TGCCCGCAAG GAACGCGTCG ACACGAGCTG CAACAGTGAC
8701  AAGCGGGAAG GGACTGGCTG CTATTGGGCG AAGTGCCGGG GCAGGATCTC
      TTCGCCCTTC CCTGACCGAC GATAACCCGC TTCACGGCCC CGTCCTAGAG
8751  CTGTCATCTC ACCTTGCTCC TGCCGAGAAA GTATCCATCA TGGCTGATGC
      GACAGTAGAG TGGAACGAGG ACGGCTCTTT CATAGGTAGT ACCGACTACG
8801  AATGCGGCGG CTGCATACGC TTGATCCGGC TACCTGCCCA TTCGACCACC
      TTACGCCGCC GACGTATGCG AACTAGGCCG ATGGACGGGT AAGCTGGTGG
8851  AAGCGAAACA TCGCATCGAG CGAGCACGTA CTCGGATGGA AGCCGGTCTT
      TTCGCTTTGT AGCGTAGCTC GCTCGTGCAT GAGCCTACCT TCGGCCAGAA
8901  GTCGATCAGG ATGATCAAGA GCATCAGGGG CTCGCGCCAG CCGAACTGTT
      CAGCTAGTCC TACTAGTTCT CGTAGTCCCC GAGCGCGGTC GGCTTGACAA
8951  CGCCAGGCTC AAGGCGAGCA TGCCCGACGG CGAGGATCTC GTCGTGACCC
      GCGGTCCGAG TTCCGCTCGT ACGGGCTGCC GCTCCTAGAG CAGCACTGGG
9001  ATGGCGATGC CTGCTTGCCG AATATCATGG TGGAAAATGG CCGCTTTTCT
      TACCGCTACG GACGAACGGC TTATAGTACC ACCTTTTACC GGCGAAAAGA
9051  GGATTCATCG ACTGTGGCCG GCTGGGTGTG GCGGACCGCT ATCAGGACAT
      CCTAAGTAGC TGACACCGGC CGACCCACAC CGCCTGGCGA TAGTCCTGTA
9101  AGCGTTGGCT ACCCGTGATA TTGCTGAAGA GCTTGGCGGC GAATGGGCTG
      TCGCAACCGA TGGGCACTAT AACGACTTCT CGAACCGCCG CTTACCCGAC
9151  ACCGCTTCCT CGTGCTTTAC GGTATCGCCG CTCCCGATTC GCAGCGCATC
      TGGCGAAGGA GCACGAAATG CCATAGCGGC GAGGGCTAAG CGTCGCGTAG
9201  GCCTTCTATC GCCTTCTTGA CGAGTTCTTC TGAGCGGGAC TCTGGGGTTC
      CGGAAGATAG CGGAAGAACT GCTCAAGAAG ACTCGCCCTG AGACCCCAAG
9251  GAAATGACCG ACCAAGCGAC GCCCAACCTG CCATCACGAG ATTTCGATTC
      CTTTACTGGC TGGTTCGCTG CGGGTTGGAC GGTAGTGCTC TAAAGCTAAG
9301  CACCGCCGCC TTCTATGAAA GGTTGGGCTT CGGAATCGTT TTCCGGGACG
      GTGGCGGCGG AAGATACTTT CCAACCCGAA GCCTTAGCAA AAGGCCCTGC
9351  CCGGCTGGAT GATCCTCCAG CGCGGGGATC TCATGCTGGA GTTCTTCGCC
      GGCCGACCTA CTAGGAGGTC GCGCCCCTAG AGTACGACCT CAAGAAGCGG
9401  CACCCTAGGG GGAGGCTAAC TGAAACACGG AAGGAGACAA TACCGGAAGG
      GTGGGATCCC CCTCCGATTG ACTTTGTGCC TTCCTCTGTT ATGGCCTTCC
9451  AACCCGCGCT ATGACGGCAA TAAAAAGACA GAATAAAACG CACGGTGTTG
      TTGGGCGCGA TACTGCCGTT ATTTTTCTGT CTTATTTTGC GTGCCACAAC
9501  GGTCGTTTGT TCATAAACGC GGGGTTCGGT CCCAGGGCTG GCACTCTGTC
      CCAGCAAACA AGTATTTGCG CCCCAAGCCA GGGTCCCGAC CGTGAGACAG
9551  GATACCCCAC CGAGACCCCA TTGGGGCCAA TACGCCCGCG TTTCTTCCTT
      CTATGGGGTG GCTCTGGGGT AACCCCGGTT ATGCGGGCGC AAAGAAGGAA
9601  TTCCCCACCC CACCCCCCAA GTTCGGGTGA AGGCCCAGGG CTCGCAGCCA
      AAGGGGTGGG GTGGGGGGTT CAAGCCCACT TCCGGGTCCC GAGCGTCGGT
9651  ACGTCGGGGC GGCAGGCCCT GCCATAGCCT CAAATTCCTT AGGCTCGAGG
      TGCAGCCCCG CCGTCCGGGA CGGTATCGGA GTTTAAGGAA TCCGAGCTCC
9701  GCCGCCACCG CGGTGGAGCT CCAGCTTTTG TTCCCTTTAG TGAGGGTTAA
      CGGCGGTGGC GCCACCTCGA GGTCGAAAAC AAGGGAAATC ACTCCCAATT
9751  TTTCGAGCTT GGCGTAATCA TGGTCATAGC TGTTTCCTGT GTGAAATTGT
      AAAGCTCGAA CCGCATTAGT ACCAGTATCG ACAAAGGACA CACTTTAACA
9801  TATCCGCTCA CAATTCCACA CAACATACGA GCCGGAAGCA TAAAGTGTAA
      ATAGGCGAGT GTTAAGGTGT GTTGTATGCT CGGCCTTCGT ATTTCACATT
9851  AGCCTGGGGT GCCTAATGAG TGAGCTAACT CACATTAATT GCGTTGCGCT
      TCGGACCCCA CGGATTACTC ACTCGATTGA GTGTAATTAA CGCAACGCGA
9901  CACTGCCCGC TTTCCAGTCG GGAAACCTGT CGTGCCAGCA TCGCGAGCAC
      GTGACGGGCG AAAGGTCAGC CCTTTGGACA GCACGGTCGT AGCGCTCGTG
9951  TTTTCGGGGA AATGTGCGCG GAACCCCTAT TTGTTTATTT TTCTAAATAC
      AAAAGCCCCT TTACACGCGC CTTGGGGATA AACAAATAAA AAGATTTATG
10001 ATTCAAATAT GTATCCGCTC ATGAGACAAT AACCCTGATA AATGCTTCAA
      TAAGTTTATA CATAGGCGAG TACTCTGTTA TTGGGACTAT TTACGAAGTT
```

FIG. 29(contd.)

```
10051   TAATATTGAA AAAGGAAGAG TATGAGTATT CAACATTTCC GTGTCGCCCT
        ATTATAACTT TTTCCTTCTC ATACTCATAA GTTGTAAAGG CACAGCGGGA
10101   TATTCCCTTT TTTGCGGCAT TTTGCCTTCC TGTTTTTGCT CACCCAGAAA
        ATAAGGGAAA AAACGCCGTA AAACGGAAGG ACAAAAACGA GTGGGTCTTT
10151   CGCTGGTGAA AGTAAAAGAT GCTGAAGATC AGTTGGGTGC ACGAGTGGGT
        GCGACCACTT TCATTTTCTA CGACTTCTAG TCAACCCACG TGCTCACCCA
10201   TACATCGAAC TGGATCTCAA CAGCGGTAAG ATCCTTGAGA GTTTTCGCCC
        ATGTAGCTTG ACCTAGAGTT GTCGCCATTC TAGGAACTCT CAAAAGCGGG
10251   CGAAGAACGT TTTCCAATGA TGAGCACTTT TAAAGTTCTG CTATGTGGCG
        GCTTCTTGCA AAAGGTTACT ACTCGTGAAA ATTTCAAGAC GATACACCGC
10301   CGGTATTATC CCGTATTGAC GCCGGGCAAG AGCAACTCGG TCGCCGCATA
        GCCATAATAG GGCATAACTG CGGCCCGTTC TCGTTGAGCC AGCGGCGTAT
10351   CACTATTCTC AGAATGACTT GGTTGAGTAC TCACCAGTCA CAGAAAAGCA
        GTGATAAGAG TCTTACTGAA CCAACTCATG AGTGGTCAGT GTCTTTTCGT
10401   TCTTACGGAT GGCATGACAG TAAGAGAATT ATGCAGTGCT GCCATAACCA
        AGAATGCCTA CCGTACTGTC ATTCTCTTAA TACGTCACGA CGGTATTGGT
10451   TGAGTGATAA CACTGCGGCC AACTTACTTC TGACAACGAT CGGAGGACCG
        ACTCACTATT GTGACGCCGG TTGAATGAAG ACTGTTGCTA GCCTCCTGGC
10501   AAGGAGCTAA CCGCTTTTTT GCACAACATG GGGGATCATG TAACTCGCCT
        TTCCTCGATT GGCGAAAAAA CGTGTTGTAC CCCCTAGTAC ATTGAGCGGA
10551   TGATCGTTGG GAACCGGAGC TGAATGAAGC CATACCAAAC GACGAGCGTG
        ACTAGCAACC CTTGGCCTCG ACTTACTTCG GTATGGTTTG CTGCTCGCAC
10601   ACACCACGAT GCCTGTAGCA ATGGCAACAA CGTTGCGCAA ACTATTAACT
        TGTGGTGCTA CGGACATCGT TACCGTTGTT GCAACGCGTT TGATAATTGA
10651   GGCGAACTAC TTACTCTAGC TTCCCGGCAA CAATTAATAG ACTGGATGGA
        CCGCTTGATG AATGAGATCG AAGGGCCGTT GTTAATTATC TGACCTACCT
10701   GGCGGATAAA GTTGCAGGAC CACTTCTGCG CTCGGCCCTT CCGGCTGGCT
        CCGCCTATTT CAACGTCCTG GTGAAGACGC GAGCCGGGAA GGCCGACCGA
10751   GGTTTATTGC TGATAAATCT GGAGCCGGTG AGCGTGGGTC TCGCGGTATC
        CCAAATAACG ACTATTTAGA CCTCGGCCAC TCGCACCCAG AGCGCCATAG
10801   ATTGCAGCAC TGGGGCCAGA TGGTAAGCCC TCCCGTATCG TAGTTATCTA
        TAACGTCGTG ACCCCGGTCT ACCATTCGGG AGGGCATAGC ATCAATAGAT
10851   CACGACGGGG AGTCAGGCAA CTATGGATGA ACGAAATAGA CAGATCGCTG
        GTGCTGCCCC TCAGTCCGTT GATACCTACT TGCTTTATCT GTCTAGCGAC
10901   AGATAGGTGC CTCACTGATT AAGCATTGGT AACTGTCAGA CTCGCGACAC
        TCTATCCACG GAGTGACTAA TTCGTAACCA TTGACAGTCT GAGCGCTGTG
10951   TGCATTAATG AATCGGCCAA CGCGCGGGGA GAGGCGGTTT GCGTATTGGG
        ACGTAATTAC TTAGCCGGTT GCGCGCCCCT CTCCGCCAAA CGCATAACCC
11001   CGCTCTTCCG CTTCCTCGCT CACTGACTCG CTGCGCTCGG TCGTTCGGCT
        GCGAGAAGGC GAAGGAGCGA GTGACTGAGC GACGCGAGCC AGCAAGCCGA
11051   GCGGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG
        CGCCGCTCGC CATAGTCGAG TGAGTTTCCG CCATTATGCC AATAGGTGTC
11101   AATCAGGGGA TAACGCAGGA AAGAACATGT GAGCAAAAGG CCAGCAAAAG
        TTAGTCCCCT ATTGCGTCCT TTCTTGTACA CTCGTTTTCC GGTCGTTTTC
11151   GCCAGGAACC GTAAAAAGGC CGCGTTGCTG GCGTTTTTCC ATAGGCTCCG
        CGGTCCTTGG CATTTTTCCG GCGCAACGAC CGCAAAAAGG TATCCGAGGC
11201   CCCCCCTGAC GAGCATCACA AAAATCGACG CTCAAGTCAG AGGTGGCGAA
        GGGGGGACTG CTCGTAGTGT TTTTAGCTGC GAGTTCAGTC TCCACCGCTT
11251   ACCCGACAGG ACTATAAAGA TACCAGGCGT TTCCCCCTGG AAGCTCCCTC
        TGGGCTGTCC TGATATTTCT ATGGTCCGCA AAGGGGGACC TTCGAGGGAG
11301   GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC TGTCCGCCTT
        CACGCGAGAG GACAAGGCTG GGACGGCGAA TGGCCTATGG ACAGGCGGAA
11351   TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC
        AGAGGGAAGC CCTTCGCACC GCGAAAGAGT ATCGAGTGCG ACATCCATAG
11401   TCAGTTCGGT GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC
        AGTCAAGCCA CATCCAGCAA GCGAGGTTCG ACCCGACACA CGTGCTTGGG
11451   CCCGTTCAGC CCGACCGCTG CGCCTTATCC GGTAACTATC GTCTTGAGTC
        GGGCAAGTCG GGCTGGCGAC GCGGAATAGG CCATTGATAG CAGAACTCAG
```

```
11501   CAACCCGGTA AGACACGACT TATCGCCACT GGCAGCAGCC ACTGGTAACA
        GTTGGGCCAT TCTGTGCTGA ATAGCGGTGA CCGTCGTCGG TGACCATTGT
11551   GGATTAGCAG AGCGAGGTAT GTAGGCGGTG CTACAGAGTT CTTGAAGTGG
        CCTAATCGTC TCGCTCCATA CATCCGCCAC GATGTCTCAA GAACTTCACC
11601   TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA TCTGCGCTCT
        ACCGGATTGA TGCCGATGTG ATCTTCCTGT CATAAACCAT AGACGCGAGA
11651   GCTGAAGCCA GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA
        CGACTTCGGT CAATGGAAGC CTTTTTCTCA ACCATCGAGA ACTAGGCCGT
11701   AACAAACCAC CGCTGGTAGC GGTGGTTTTT TTGTTTGCAA GCAGCAGATT
        TTGTTTGGTG GCGACCATCG CCACCAAAAA AACAAACGTT CGTCGTCTAA
11751   ACGCGCAGAA AAAAAGGATC TCAAGAAGAT CCTTTGATCT TTTCTACGGG
        TGCGCGTCTT TTTTTCCTAG AGTTCTTCTA GGAAACTAGA AAAGATGCCC
11801   GTCTGACGCT CAGTGGAACG AAAACTCA
        CAGACTGCGA GTCACCTTGC TTTTGAGT
```

FIG. 29(contd.)

FIG. 30

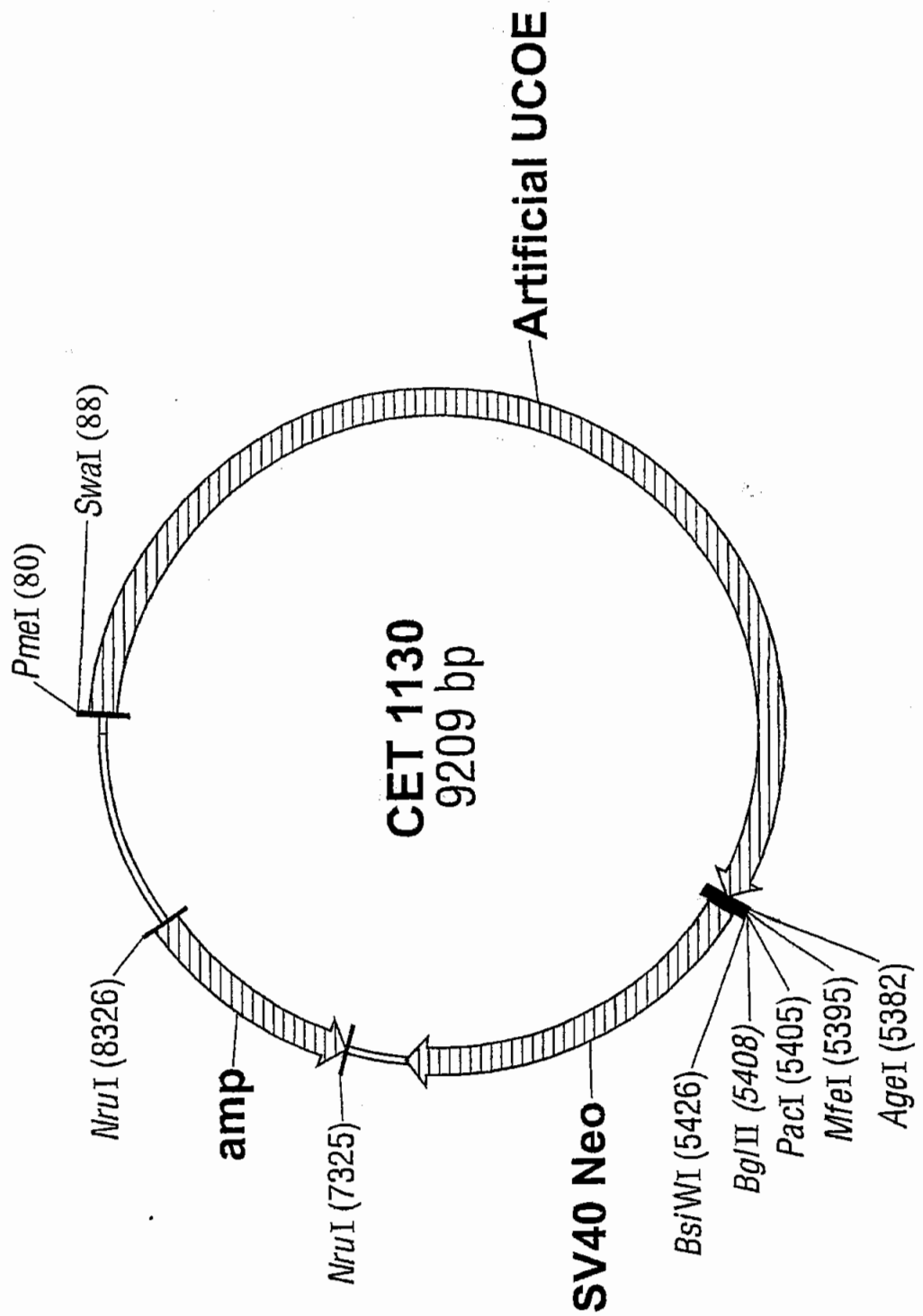

Artificial UCOE
SwaI (88)
PmeI (80)
CET 1130
9209 bp
NruI (8326)
amp
NruI (7325)
SV40 Neo
BsiWI (5426)
BglII (5408)
PacI (5405)
MfeI (5395)
AgeI (5382)

FIG. 31 CET1130 nucleotide sequence

```
   1  CGTTGTAAAA CGACGGCCAG TGAATTGTAA TACGACTCAC TATAGGGCGA
      GCAACATTTT GCTGCCGGTC ACTTAACATT ATGCTGAGTG ATATCCCGCT
  51  ATTGGGTACC GGGCCCCCCC TCGAAGTTTA AACATTTAAA TCTAGAACTA
      TAACCCATGG CCCGGGGGGG AGCTTCAAAT TTGTAAATTT AGATCTTGAT
 101  GTGGATCCCC CGGGCTGCAG GAATTCGATA TCAAGCTCAT GGCACCTGTA
      CACCTAGGGG GCCCGACGTC CTTAAGCTAT AGTTCGAGTA CCGTGGACAT
 151  TTGTACTCTT ATCAGTCATT ATATGGACTT TAACTTCCCC AGATATTATT
      AACATGAGAA TAGTCAGTAA TATACCTGAA ATTGAAGGGG TCTATAATAA
 201  TGGGCTCCTC CATAAGACTG TGAGCATCTG ACCACTGGAG TGTTGCTTCC
      ACCCGAGGAG GTATTCTGAC ACTCGTAGAC TGGTGACCTC ACAACGAAGG
 251  CATTATATCC CTGTTATCAA GCACAAGGTC AGGCACAGAG TAAGACTCAA
      GTAATATAGG GACAATAGTT CGTGTTCCAG TCCGTGTCTC ATTCTGAGTT
 301  AACATGTTTT GGAATGTATG ACTGGTATGA ACTACAAACC AGTAAGCTGA
      TTGTACAAAA CCTTACATAC TGACCATACT TGATGTTTGG TCATTCGACT
 351  TGTTTTCATT TTGAGTCTAT AAATCTAATT TTGTGGTGGT TTTGTGTATG
      ACAAAAGTAA AACTCAGATA TTTAGATTAA AACACCACCA AAACACATAC
 401  GCTCAAGGCT CAAATTGTAA AATTTAATAT TATGTGACCA AAGAAAGTTA
      CGAGTTCCGA GTTTAACATT TTAAATTATA ATACACTGGT TTCTTTCAAT
 451  TACCCAGAAC CTCAATTTCC TCACCTTCAA AATGGGGCAG TTTCTCACTC
      ATGGGTCTTG GAGTTAAAGG AGTGGAAGTT TTACCCCGTC AAAGAGTGAG
 501  ATTGGTCTGC TGTCACGATT TTAATGAGCT CATGCACAAA CAGCCCTTTA
      TAACCAGACG ACAGTGCTAA AATTACTCGA GTACGTGTTT GTCGGGAAAT
 551  TATAAGGTAA GTGCTGGATA AATGTTGGCT ACTATAATAA AATAAGCCTC
      ATATTCCATT CACGACCTAT TTACAACCGA TGATATTATT TTATTCGGAG
 601  TAAGATACTT GGTCAGCACA AGTACTACCC AAGAGTATGC ACTGTAAGTA
      ATTCTATGAA CCAGTCGTGT TCATGATGGG TTCTCATACG TGACATTCAT
 651  AACTGACAAA ATTGTGTATC TAAAACTGGC CAGATGAAAG AGAAACTTTT
      TTGACTGTTT TAACACATAG ATTTTGACCG GTCTACTTTC TCTTTGAAAA
 701  AAGGGGCCCT TCTGCGTGCC CGACACTGTG CTAGGCACTC ACACTATCCC
      TTCCCCGGGA AGACGCACGG GCTGTGACAC GATCCGTGAG TGTGATAGGG
 751  GACCCGAGAA ACCGATCTGC GACCCAGAGG AACTTACCAA GCCTCCAGCA
      CTGGGCTCTT TGGCTAGACG CTGGGTCTCC TTGAATGGTT CGGAGGTCGT
 801  TCTTGTGCAG CCCTACTCAT GGGACCATCT GGATACCCAC CCTTGTCTTT
      AGAACACGTC GGGATGAGTA CCCTGGTAGA CCTATGGGTG GGAACAGAAA
 851  ACAGGGAGCA GAACACACCT CTTATGTGTC AGAAAACAAA GTCCAGGAAG
      TGTCCCTCGT CTTGTGTGGA GAATACACAG TCTTTTGTTT CAGGTCCTTC
 901  TATATTTTTA CCTGAGGCAA TATCTGAAAA TTGTATGCTA CAGCCTCCAA
      ATATAAAAAT GGACTCCGTT ATAGACTTTT AACATACGAT GTCGGAGGTT
 951  AGTGAGTCTT CCTCTCAGTA CCTCTCTTCT AGGCACATGG AGCCCTTTCT
      TCACTCAGAA GGAGAGTCAT GGAGAGAAGA TCCGTGTACC TCGGGAAAGA
1001  TCCAAGTATT ATGTTTAACC ACTTAATGAA TGAAGTCCTG AAACTGCTTA
      AGGTTCATAA TACAAATTGG TGAATTACTT ACTTCAGGAC TTTGACGAAT
1051  CCCATGCTCC CTATAATCTC TGAGTAATCT TCCTTTTCCA CAACCTCAGG
      GGGTACGAGG GATATTAGAG ACTCATTAGA AGGAAAAGGT GTTGGAGTCC
1101  CATAATCTCA TCTTCTGTTT CTATTACAAT TTCAAATTCT GGAAAAAGGA
      GTATTAGAGT AGAAGACAAA GATAATGTTA AAGTTTAAGA CCTTTTTCCT
1151  AGTTGTGGTC TGGAATTATA TGGTCCAGAT GATCTGAAAC AAAAAGGACA
      TCAACACCAG ACCTTAATAT ACCAGGTCTA CTAGACTTTG TTTTTCCTGT
1201  GCACTATTAG TAATCATTTA GTTTTGAAGA CAGTCTAATA ATTTGCTGTC
      CGTGATAATC ATTAGTAAAT CAAAACTTCT GTCAGATTAT TAAACGACAG
1251  TCTAAAGTAC TATATTCCCT ATAGTTCTGG CATTTTAGAT AAAGGGTCAT
      AGATTTCATG ATATAAGGGA TATCAAGACC GTAAAATCTA TTTCCCAGTA
1301  AAATTAAATG CCTATATGGT GACATTATTC AGTGATTCAG ACTTCACAGC
      TTTAATTTAC GGATATACCA CTGTAATAAG TCACTAAGTC TGAAGTGTCG
```

FIG. 31(contd.) 88/93

```
1351  CTTTTTTTTT TTTTTACAAA GGTGTTCCAG GCATGAAAAA TTTTAAAGTA
      GAAAAAAAAA AAAAATGTTT CCACAAGGTC CGTACTTTTT AAAATTTCAT
1401  CTATACCTTT CCTAATTTTA CCTTTAAAGT TGTCCTGGAA ATATCTGGGT
      GATATGGAAA GGATTAAAAT GGAAATTTCA ACAGGACCTT TATAGACCCA
1451  TGACAAAGGC GATGAAACTG AACTGAGACT TAAAAAAAAG ATTACCCACC
      ACTGTTTCCG CTACTTTGAC TTGACTCTGA ATTTTTTTTC TAATGGGTGG
1501  TGGTTGTGCA CAAGCCTGCT TATGTCCCAA TCTCCAGTCT AGGGTCTGAT
      ACCAACACGT GTTCGGACGA ATACAGGGTT AGAGGTCAGA TCCCAGACTA
1551  GCTCCTTGCT GCAGTAATAT GCTTTGTGGC ATCTGGAGCA CGTTTTGGGG
      CGAGGAACGA CGTCATTATA CGAAACACCG TAGACCTCGT GCAAAACCCC
1601  CCTAAACAGC CACAAACCCT GCAGAGATGA GCACCAGACT TAAGCTGGAG
      GGATTTGTCG GTGTTTGGGA CGTCTCTACT CGTGGTCTGA ATTCGACCTC
1651  ACACACTGAT TCTCCTGTTT CTGGGGGAGG ATTCTCAGAA GGTGGCTCAT
      TGTGTGACTA AGAGGACAAA GACCCCCTCC TAAGAGTCTT CCACCGAGTA
1701  ATGAGTAAAA ATCGTTTTTC CTGGGTAGTT GATTCCTAAA AACTAAAAAA
      TACTCATTTT TAGCAAAAAG GACCCATCAA CTAAGGATTT TTGATTTTTT
1751  GAATACAGAG AAAAGTTTTA TCTTCAAACA AAACAGCAAT TCACATATTT
      CTTATGTCTC TTTTCAAAAT AGAAGTTTGT TTTGTCGTTA AGTGTATAAA
1801  TATCCTCTGC ACGTAAAACT GAAAATAACA ACAACAAAAA AGAAATGAAA
      ATAGGAGACG TGCATTTTGA CTTTTATTGT TGTTGTTTTT TCTTTACTTT
1851  GTTTTTGCTT TCAGGAATAA GCTTTTAAAA TCCAGAAACT AGATTTCGTC
      CAAAAACGAA AGTCCTTATT CGAAAATTTT AGGTCTTTGA TCTAAAGCAG
1901  CGGTACACGC AACTGAGTTG CCTCCTAGAG GTGGTTTGAG TTAATCAAAT
      GCCATGTGCG TTGACTCAAC GGAGGATCTC CACCAAACTC AATTAGTTTA
1951  TAATAAGACT GATCGTTAAG AACGACTGCC AAAAATACGA AAAAGCTACT
      ATTATTCTGA CTAGCAATTC TTGCTGACGG TTTTTATGCT TTTTCGATGA
2001  GGGATCCATC TTTCCAAGAC AATTTCTATT ATCTGAATTA ACACCATACC
      CCCTAGGTAG AAAGGTTCTG TTAAAGATAA TAGACTTAAT TGTGGTATGG
2051  TGGTACCCAC TGATTAAAAG CTGGGGGTTA CCAATGCGCG TGGGCACAGT
      ACCATGGGTG ACTAATTTTC GACCCCCAAT GGTTACGCGC ACCCGTGTCA
2101  TAGAAGCTTA TGTAGCAAAA ATGAGCACAT CCTGGAAGGG CCCGGGAGAA
      ATCTTCGAAT ACATCGTTTT TACTCGTGTA GGACCTTCCC GGGCCCTCTT
2151  GGTGCTCCTG GGGCAGCGCG GAGAGGGAGC TCTGAGGCTG GGGCGGCAGC
      CCACGAGGAC CCCGTCGCGC CTCTCCCTCG AGACTCCGAC CCCGCCGTCG
2201  GGTGCTTGCC GCCGTCCCCC TGGTCGCTCC CGGAATTAAC GCCGCGCACG
      CCACGAACGG CGGCAGGGGG ACCAGCGAGG GCCTTAATTG CGGCGCGTGC
2251  CGTCGGAGGC ATGGCCCCGT CCCGACCCCG TTTGGCGGCT CACCTCGCAG
      GCAGCCTCCG TACCGGGGCA GGGCTGGGGC AAACCGCCGA GTGGAGCGTC
2301  GCCGGCACAG CACGGCTGCT CGCGGCAGCA GAAGAGGAAG ATGCAGCGGT
      CGGCCGTGTC GTGCCGACGA GCGCCGTCGT CTTCTCCTTC TACGTCGCCA
2351  GGAAGGCGTC CGGGCGGCCA GGCAGCGGCG CATACACCTG CAGCAGGAAG
      CCTTCCGCAG GCCCGCCGGT CCGTCGCCGC GTATGTGGAC GTCGTCCTTC
2401  GAGAGCGGGC GGCCGCACAG CTCGCAGGCC AGGGCCTGGG GCCCCGGCAG
      CTCTCGCCCG CCGGCGTGTC GAGCGTCCGG TCCCGGACCC CGGGGCCGTC
2451  CCCGGCCGCG CCCAGCCATG CCGGCCGCCC GCCCACCTTG CTGGGGAACT
      GGGCCGGCGC GGGTCGGTAC GGCCGGCGGG CGGGTGGAAC GACCCCTTGA
2501  GCTCGCTGCG CAGTCGCCAC GCCGGCGCCG ACTCGGCGAA GCCCAGCTCC
      CGAGCGACGC GTCAGCGGTG CGGCCGCGGC TGAGCCGCTT CGGGTCGAGG
2551  ACAGGCCTGG CCCCGGCGGC AGCCATGCGG GGCGCGGGCT GGCGTGGGGC
      TGTCCGGACC GGGGCCGCCG TCGGTACGCC CCGCGCCCGA CCGCACCCCG
2601  GCAGCCCACA GCTGGGTCGG AAGGCGGAAA TCGGGCGCCG GGCCGGAAGG
      CGTCGGGTGT CGACCCAGCC TTCCGCCTTT AGCCCGCGGC CCGGCCTTCC
2651  CAAGAGGCGG GCACCTTTCC GGAGGACAGG AGGCGGAAAC GCGTCTGACG
      GTTCTCCGCC CGTGGAAAGG CCTCCTGTCC TCCGCCTTTG CGCAGACTGC
2701  GGAGCGGTTG CAGGACCAAT GCGAGGGAAC GGGGCAGAGG AAACCTCTCG
      CCTCGCCAAC GTCCTGGTTA CGCTCCCTTG CCCCGTCTCC TTTGGAGAGC
2751  GCATCAGCCC CGCCCCTGGC GCCTCTGCCT CCGAGCCGCT TTCCTGGTGC
      CGTAGTCGGG GCGGGGACCG CGGAGACGGA GGCTCGGCGA AAGGACCACG
```

FIG. 31(contd.)

```
2801  CTCCGGGTGC TCTGGGATGG TTCTGGTCTT TGGGAGAGTG GCAGCTGGTG
      GAGGCCCACG AGACCCTACC AAGACCAGAA ACCCTCTCAC CGTCGACCAC
2851  ACGGCGCTCC GCTCACCTCT GCACATGTCT TGCTGTGGGC CTGCGGGTGG
      TGCCGCGAGG CGAGTGGAGA CGTGTACAGA ACGACACCCG GACGCCCACC
2901  CCGCCAGGGA GGCAGAGCCC TCCCGCAAAC CTTCCCTGCT GGTGTCCACC
      GGCGGTCCCT CCGTCTCGGG AGGGCGTTTG GAAGGGACGA CCACAGGTGG
2951  TCAGGGTGTG GGAAACCTGT GCGCTGGCCG AGTGCTAACC AAGAGTAGGC
      AGTCCCACAC CCTTTGGACA CGCGACCGGC TCACGATTGG TTCTCATCCG
3001  AGTGAAAGAC AAATGAAGGT TGAACAGGTA AAGTGAGGAC CCTACAGCGG
      TCACTTTCTG TTTACTTCCA ACTTGTCCAT TTCACTCCTG GGATGTCGCC
3051  AAACCAAGAA TCCTGTGTGC CTGAGAGTAA TGAAGAAGCC TCTGCAGAAG
      TTTGGTTCTT AGGACACACG GACTCTCATT ACTTCTTCGG AGACGTCTTC
3101  AGTCTTTTCT GTCAGTCTTA AGGTCTCTGT TTTAATGTTA GTGCTGGCTT
      TCAGAAAAGA CAGTCAGAAT TCCAGAGACA AAATTACAAT CACGACCGAA
3151  GCTGTACCTG AATTCCAAGG GAGGAGTGTA TAATGAGGCA TGGCCAACCC
      CGACATGGAC TTAAGGTTCC CTCCTCACAT ATTACTCCGT ACCGGTTGGG
3201  CCACTTCCCA TCATTGCCTG AACTAGTTTT TCAGGTTAAC TTCAGAATGC
      GGTGAAGGGT AGTAACGGAC TTGATCAAAA AGTCCAATTG AAGTCTTACG
3251  CCTTGGTACC GCGGGCCCCC TCTGTGGTCC CACGCCACTG ATCGCTGCAT
      GGAACCATGG CGCCCGGGGG AGACACCAGG GTGCGGTGAC TAGCGACGTA
3301  GCCCACCACC TGGGTACACA CAGTCTGTGA TTCCCGGAGC AGAACGGACC
      CGGGTGGTGG ACCCATGTGT GTCAGACACT AAGGGCCTCG TCTTGCCTGG
3351  CTGCCCACCC GGTCTTGTGT GCTACTCAGT GGACAGACCC AAGGCAAGAA
      GACGGGTGGG CCAGAACACA CGATGAGTCA CCTGTCTGGG TTCCGTTCTT
3401  AGGGTGACAA GGACAGGGTC TTCCCAGGCT GGCTTTGAGT TCCTAGCACC
      TCCCACTGTT CCTGTCCCAG AAGGGTCCGA CCGAAACTCA AGGATCGTGG
3451  GCCCCGCCCC CAATCCTCTG TGGCACATGG AGTCTTGGTC CCCAGAGTCC
      CGGGGCGGGG GTTAGGAGAC ACCGTGTACC TCAGAACCAG GGGTCTCAGG
3501  CCCAGCGGCC TCCAGATGGT CTGGGAGGGC AGTTCAGCTG TGGCTGCGCA
      GGGTCGCCGG AGGTCTACCA GACCCTCCCG TCAAGTCGAC ACCGACGCGT
3551  TAGCAGACAT ACAACGGACG GTGGGCCCAG ACCCAGGCTG TGTAGACCCA
      ATCGTCTGTA TGTTGCCTGC CACCCGGGTC TGGGTCCGAC ACATCTGGGT
3601  GCCCCCCCGC CCCGCAGTGC CTAGGTCACC CACTAACGCC CCAGGCCTGG
      CGGGGGGGCG GGGCGTCACG GATCCAGTGG GTGATTGCGG GGTCCGGACC
3651  TCTTGGCTGG GCGTGACTGT TACCCTCAAA AGCAGGCAGC TCCAGGGTAA
      AGAACCGACC CGCACTGACA ATGGGAGTTT TCGTCCGTCG AGGTCCCATT
3701  AAGGTGCCCT GCCCTGTAGA GCCCACTTCC TTCCCAGGGC TGCGGCTGGG
      TTCCACGGGA CGGGACATCT CGGGTGAAGG AAGGGTCCCG ACGCCGACCC
3751  TAGGTTTGTA GCCTTCATCA CGGGCCACCT CCAGCCACTG GACCGCTGGC
      ATCCAAACAT CGGAAGTAGT GCCCGGTGGA GGTCGGTGAC CTGGCGACCG
3801  CCCTGCCCTG TCCTGGGGAG TGTGGTCCTG CGACTCTAAT GGCCGCAAGC
      GGGACGGGAC AGGACCCCTC ACACCAGGAC GCTGAGATTA CCGGCGTTCG
3851  CACCTGACTC CCCCAACACC ACACTCTACC TCTCAAGCCC AGGTCTCTCC
      GTGGACTGAG GGGGTTGTGG TGTGAGATGG AGAGTTCGGG TCCAGAGAGG
3901  CTAGTGACCC ACCCAGCACA TTTAGCTAGC TGAGCCCCAC AGCCAGAGGT
      GATCACTGGG TGGGTCGTGT AAATCGATCG ACTCGGGGTG TCGGTCTCCA
3951  CCTCAGGCCC TGCTTTCAGG GCAGTTGCTC TGAAGTCGGC AAGGGGGAGT
      GGAGTCCGGG ACGAAAGTCC CGTCAACGAG ACTTCAGCCG TTCCCCCTCA
4001  GACTGCCTGG CCACTCCATG CCCTCCAAGA GCTCCTTCTG CAGGAGCGTA
      CTGACGGACC GGTGAGGTAC GGGAGGTTCT CGAGGAAGAC GTCCTCGCAT
4051  CAGAACCCAG GGCCCTGGCA CCCGTGCAGA CCCTGGCCCA CCCCACCTGG
      GTCTTGGGTC CCGGGACCGT GGGCACGTCT GGGACCGGGT GGGGTGGACC
4101  GCGCTCAGTG CCCAAGAGAT GTCCACACCT AGGATGTCCC GCGGTGGGTG
      CGCGAGTCAC GGGTTCTCTA CAGGTGTGGA TCCTACAGGG CGCCACCCAC
4151  GGGGGCCCGA GAGACGGGCA GGCCGGGGGC AGGCCTGGCC ATGCGGGGCC
      CCCCCGGGCT CTCTGCCCGT CCGGCCCCCG TCCGGACCGG TACGCCCCGG
4201  GAACCGGGCA CTGCCCAGCG TGGGGCGCGG GGGCCACGGC GCGCGCCCCC
      CTTGGCCCGT GACGGGTCGC ACCCCGCGCC CCCGGTGCCG CGCGCGGGGG
```

FIG. 31(contd.)

```
4251  AGCCCCCGGG CCCAGCACCC CAAGGCGGCC AACGCCAAAA CTCTCCCTCC
      TCGGGGGCCC GGGTCGTGGG GTTCCGCCGG TTGCGGTTTT GAGAGGGAGG
4301  TCCTCTTCCT CAATCTCGCT CTCGCTCTTT TTTTTTTTCG CAAAAGGAGG
      AGGAGAAGGA GTTAGAGCGA GAGCGAGAAA AAAAAAAAGC GTTTTCCTCC
4351  GGAGAGGGGG TAAAAAAATG CTGCACTGTG CGGCGAAGCC GGTGAGTGAG
      CCTCTCCCCC ATTTTTTTAC GACGTGACAC GCCGCTTCGG CCACTCACTC
4401  CGGCGCGGGG CCAATCAGCG TGCGCCGTTC CGAAAGTTGC CTTTTATGGC
      GCCGCGCCCC GGTTAGTCGC ACGCGGCAAG GCTTTCAACG GAAAATACCG
4451  TCGAGCGGCC GCGGCGGCGC CCTATAAAAC CCAGCGGCGC GACGCGCCAC
      AGCTCGCCGG CGCCGCCGCG GGATATTTTG GGTCGCCGCG CTGCGCGGTG
4501  CACCGCCGAG ACCGCGTCCG CCCGCGAGCA CAGAGCCTCG CCTTTGCCGA
      GTGGCGGCTC TGGCGCAGGC GGGCGCTCGT GTCTCGGAGC GGAAACGGCT
4551  TCCGCCGCCC GTCCACACCC GCCGCCAGGT AAGCCCGGCC AGCCGACCGG
      AGGCGGCGGG CAGGTGTGGG CGGCGGTCCA TTCGGGCCGG TCGGCTGGCC
4601  GGCATGCGGC CGCGGCCCTT CGCCCGTGCA GAGCCGCCGT CTGGGCCGCA
      CCGTACGCCG GCGCCGGGAA GCGGGCACGT CTCGGCGGCA GACCCGGCGT
4651  GCGGGGGGCG CATGGGGCGG AACCGGACCG CCGTGGGGGG CGCGGGAGAA
      CGCCCCCCGC GTACCCCGCC TTGGCCTGGC GGCACCCCCC GCGCCCTCTT
4701  GCCCCTGGGC CTCCGGAGAT GGGGGACACC CCACGCCAGT TCGCAGGCGC
      CGGGGACCCG GAGGCCTCTA CCCCCTGTGG GGTGCGGTCA AGCGTCCGCG
4751  GAGGCCGCGC TCGGGCGGGC GCGCTCCGGG GGTGCCGCTC TCGGGGCGGG
      CTCCGGCGCG AGCCCGCCCG CGCGAGGCCC CCACGGCGAG AGCCCCGCCC
4801  GGCAACCGGC GGGGTCTTTG TCTGAGCCGG GCTCTTGCCA ATGGGGATCG
      CCGTTGGCCG CCCCAGAAAC AGACTCGGCC CGAGAACGGT TACCCCTAGC
4851  CACGGTGGGC GCGGCGTAGC CCCCGTCAGG CCCGGTGGGG GCTGGGGCGC
      GTGCCACCCG CGCCGCATCG GGGGCAGTCC GGGCCACCCC CGACCCCGCG
4901  CATGCGCGTG CGCGCTGGTC CTTTGGGCGC TAACTGCGTG CGCGCTGGGA
      GTACGCGCAC GCGCGACCAG GAAACCCGCG ATTGACGCAC GCGCGACCCT
4951  ATTGGCGCTA ATTGCGCGTG CGCGCTGGGA CTCAATGGCG CTAATCGCGC
      TAACCGCGAT TAACGCGCAC GCGCGACCCT GAGTTACCGC GATTAGCGCG
5001  GTGCGTTCTG GGGCCCGGGC GCTTGCGCCA CTTCCTGCCC GAGCCGCTGG
      CACGCAAGAC CCCGGGCCCG CGAACGCGGT GAAGGACGGG CTCGGCGACC
5051  CGCCCGAGGG TGTGGCCGCT GCGTGCGCGC GCGCGACCCG GTCGCTGTTT
      GCGGGCTCCC ACACCGGCGA CGCACGCGCG CGCGCTGGGC CAGCGACAAA
5101  GAACCGGGCG GAGGCGGGGC TGGCGCCCGG TTGGGAGGGG GTTGGGGCCT
      CTTGGCCCGC CTCCGCCCCG ACCGCGGGCC AACCCTCCCC CAACCCCGGA
5151  GGCTTCCTGC CGCGCGCCGC GGGGACGCCT CCGACCAGTG TTTGCCTTTT
      CCGAAGGACG GCGCGCGGCG CCCCTGCGGA GGCTGGTCAC AAACGGAAAA
5201  ATGGTAATAA CGCGGCCGGC CCGGCTTCCT TTGTCCCCAA TCTGGGCGCG
      TACCATTATT GCGCCGGCCG GGCCGAAGGA AACAGGGGTT AGACCCGCGC
5251  CGCCGGCGCC CCCTGGCGGC CTAAGGACTC GGCGCGCCGG AAGTGGCCAG
      GCGGCCGCGG GGGACCGCCG GATTCCTGAG CCGCGCGGCC TTCACCGGTC
5301  GGCGGGGGCG ACTTCGGCTC ACAGCGCGCC CGGCTATTCT CGCAGCTCAC
      CCGCCCCCGC TGAAGCCGAG TGTCGCGCGG GCCGATAAGA GCGTCGAGTG
5351  CATGCCGGTC GCCACCATGA GCTTATCGAT ACCGGTGGCG CGCCAATTGT
      GTACGGCCAG CGGTGGTACT CGAATAGCTA TGGCCACCGC GCGGTTAACA
5401  TAATTAAGAT CTGGCCCAAT GGGCCGTACG AATTTGAGGC GGAAAGAACC
      ATTAATTCTA GACCGGGTTA CCCGGCATGC TTAAACTCCG CCTTTCTTGG
5451  AGCTGTGGAA TGTGTGTCAG TTAGGGTGTG GAAAGTCCCC AGGCTCCCCA
      TCGACACCTT ACACACAGTC AATCCCACAC CTTTCAGGGG TCCGAGGGGT
5501  GCAGGCAGAA GTATGCAAAG CATGCATCTC AATTAGTCAG CAACCAGGTG
      CGTCCGTCTT CATACGTTTC GTACGTAGAG TTAATCAGTC GTTGGTCCAC
5551  TGGAAAGTCC CCAGGCTCCC CAGCAGGCAG AAGTATGCAA AGCATGCATC
      ACCTTTCAGG GGTCCGAGGG GTCGTCCGTC TTCATACGTT TCGTACGTAG
5601  TCAATTAGTC AGCAACCATA GTCCCGCCCC TAACTCCGCC CATCCCGCCC
      AGTTAATCAG TCGTTGGTAT CAGGGCGGGG ATTGAGGCGG GTAGGGCGGG
5651  CTAACTCCGC CCAGTTCCGC CCATTCTCCG CCCCATGGCT GACTAATTTT
      GATTGAGGCG GGTCAAGGCG GGTAAGAGGC GGGGTACCGA CTGATTAAAA
```

FIG. 31(contd.)

```
5701  TTTTATTTAT GCAGAGGCCG AGGCCGCCTC GGCCTCTGAG CTATTCCAGA
      AAAATAAATA CGTCTCCGGC TCCGGCGGAG CCGGAGACTC GATAAGGTCT
5751  AGTAGTGAGG AGGCTTTTTT GGAGGCCTAG GCTTTTGCAA AGATCGATCA
      TCATCACTCC TCCGAAAAAA CCTCCGGATC CGAAAACGTT TCTAGCTAGT
5801  AGAGACAGGA TGAGGATCGT TTCGCATGAT TGAACAAGAT GGATTGCACG
      TCTCTGTCCT ACTCCTAGCA AAGCGTACTA ACTTGTTCTA CCTAACGTGC
5851  CAGGTTCTCC GGCCGCTTGG GTGGAGAGGC TATTCGGCTA TGACTGGGCA
      GTCCAAGAGG CCGGCGAACC CACCTCTCCG ATAAGCCGAT ACTGACCCGT
5901  CAACAGACAA TCGGCTGCTC TGATGCCGCC GTGTTCCGGC TGTCAGCGCA
      GTTGTCTGTT AGCCGACGAG ACTACGGCGG CACAAGGCCG ACAGTCGCGT
5951  GGGGCGCCCG GTTCTTTTTG TCAAGACCGA CCTGTCCGGT GCCCTGAATG
      CCCCGCGGGC CAAGAAAAAC AGTTCTGGCT GGACAGGCCA CGGGACTTAC
6001  AACTGCAAGA CGAGGCAGCG CGGCTATCGT GGCTGGCCAC GACGGGCGTT
      TTGACGTTCT GCTCCGTCGC GCCGATAGCA CCGACCGGTG CTGCCCGCAA
6051  CCTTGCGCAG CTGTGCTCGA CGTTGTCACT GAAGCGGGAA GGGACTGGCT
      GGAACGCGTC GACACGAGCT GCAACAGTGA CTTCGCCCTT CCCTGACCGA
6101  GCTATTGGGC GAAGTGCCGG GGCAGGATCT CCTGTCATCT CACCTTGCTC
      CGATAACCCG CTTCACGGCC CCGTCCTAGA GGACAGTAGA GTGGAACGAG
6151  CTGCCGAGAA AGTATCCATC ATGGCTGATG CAATGCGGCG GCTGCATACG
      GACGGCTCTT TCATAGGTAG TACCGACTAC GTTACGCCGC CGACGTATGC
6201  CTTGATCCGG CTACCTGCCC ATTCGACCAC CAAGCGAAAC ATCGCATCGA
      GAACTAGGCC GATGGACGGG TAAGCTGGTG GTTCGCTTTG TAGCGTAGCT
6251  GCGAGCACGT ACTCGGATGG AAGCCGGTCT TGTCGATCAG GATGATCAAG
      CGCTCGTGCA TGAGCCTACC TTCGGCCAGA ACAGCTAGTC CTACTAGTTC
6301  AGCATCAGGG GCTCGCGCCA GCCGAACTGT TCGCCAGGCT CAAGGCGAGC
      TCGTAGTCCC CGAGCGCGGT CGGCTTGACA AGCGGTCCGA GTTCCGCTCG
6351  ATGCCCGACG GCGAGGATCT CGTCGTGACC CATGGCGATG CCTGCTTGCC
      TACGGGCTGC CGCTCCTAGA GCAGCACTGG GTACCGCTAC GGACGAACGG
6401  GAATATCATG GTGGAAAATG GCCGCTTTTC TGGATTCATC GACTGTGGCC
      CTTATAGTAC CACCTTTTAC CGGCGAAAAG ACCTAAGTAG CTGACACCGG
6451  GGCTGGGTGT GGCGGACCGC TATCAGGACA TAGCGTTGGC TACCCGTGAT
      CCGACCCACA CCGCCTGGCG ATAGTCCTGT ATCGCAACCG ATGGGCACTA
6501  ATTGCTGAAG AGCTTGGCGG CGAATGGGCT GACCGCTTCC TCGTGCTTTA
      TAACGACTTC TCGAACCGCC GCTTACCCGA CTGGCGAAGG AGCACGAAAT
6551  CGGTATCGCC GCTCCCGATT CGCAGCGCAT CGCCTTCTAT CGCCTTCTTG
      GCCATAGCGG CGAGGGCTAA GCGTCGCGTA GCGGAAGATA GCGGAAGAAC
6601  ACGAGTTCTT CTGAGCGGGA CTCTGGGGTT CGAAATGACC GACCAAGCGA
      TGCTCAAGAA GACTCGCCCT GAGACCCCAA GCTTTACTGG CTGGTTCGCT
6651  CGCCCAACCT GCCATCACGA GATTTCGATT CCACCGCCGC CTTCTATGAA
      GCGGGTTGGA CGGTAGTGCT CTAAAGCTAA GGTGGCGGCG GAAGATACTT
6701  AGGTTGGGCT TCGGAATCGT TTTCCGGGAC GCCGGCTGGA TGATCCTCCA
      TCCAACCCGA AGCCTTAGCA AAAGGCCCTG CGGCCGACCT ACTAGGAGGT
6751  GCGCGGGGAT CTCATGCTGG AGTTCTTCGC CCACCCTAGG GGGAGGCTAA
      CGCGCCCCTA GAGTACGACC TCAAGAAGCG GGTGGGATCC CCCTCCGATT
6801  CTGAAACACG GAAGGAGACA ATACCGGAAG GAACCCGCGC TATGACGGCA
      GACTTTGTGC CTTCCTCTGT TATGGCCTTC CTTGGGCGCG ATACTGCCGT
6851  ATAAAAAGAC AGAATAAAAC GCACGGTGTT GGGTCGTTTG TTCATAAACG
      TATTTTTCTG TCTTATTTTG CGTGCCACAA CCCAGCAAAC AAGTATTTGC
6901  CGGGGTTCGG TCCCAGGGCT GGCACTCTGT CGATACCCCA CCGAGACCCC
      GCCCCAAGCC AGGGTCCCGA CCGTGAGACA GCTATGGGGT GGCTCTGGGG
6951  ATTGGGGCCA ATACGCCCGC GTTTCTTCCT TTTCCCCACC CCACCCCCCA
      TAACCCCGGT TATGCGGGCG CAAAGAAGGA AAAGGGGTGG GGTGGGGGGT
7001  AGTTCGGGTG AAGGCCCAGG GCTCGCAGCC AACGTCGGGG CGGCAGGCCC
      TCAAGCCCAC TTCCGGGTCC CGAGCGTCGG TTGCAGCCCC GCCGTCCGGG
7051  TGCCATAGCC TCAAATTCCT TAGGCTCGAG GGCCGCCACC GCGGTGGAGC
      ACGGTATCGG AGTTTAAGGA ATCCGAGCTC CCGGCGGTGG CGCCACCTCG
7101  TCCAGCTTTT GTTCCCTTTA GTGAGGGTTA ATTTCGAGCT TGGCGTAATC
      AGGTCGAAAA CAAGGGAAAT CACTCCCAAT TAAAGCTCGA ACCGCATTAG
```

FIG. 31(contd.)

```
7151  ATGGTCATAG CTGTTTCCTG TGTGAAATTG TTATCCGCTC ACAATTCCAC
      TACCAGTATC GACAAAGGAC ACACTTTAAC AATAGGCGAG TGTTAAGGTG
7201  ACAACATACG AGCCGGAAGC ATAAAGTGTA AAGCCTGGGG TGCCTAATGA
      TGTTGTATGC TCGGCCTTCG TATTTCACAT TTCGGACCCC ACGGATTACT
7251  GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACTGCCCG CTTTCCAGTC
      CACTCGATTG AGTGTAATTA ACGCAACGCG AGTGACGGGC GAAAGGTCAG
7301  GGGAAACCTG TCGTGCCAGC ATCGCGAGCA CTTTTCGGGG AAATGTGCGC
      CCCTTTGGAC AGCACGGTCG TAGCGCTCGT GAAAAGCCCC TTTACACGCG
7351  GGAACCCCTA TTTGTTTATT TTTCTAAATA CATTCAAATA TGTATCCGCT
      CCTTGGGGAT AAACAAATAA AAAGATTTAT GTAAGTTTAT ACATAGGCGA
7401  CATGAGACAA TAACCCTGAT AAATGCTTCA ATAATATTGA AAAAGGAAGA
      GTACTCTGTT ATTGGGACTA TTTACGAAGT TATTATAACT TTTTCCTTCT
7451  GTATGAGTAT TCAACATTTC CGTGTCGCCC TTATTCCCTT TTTTGCGGCA
      CATACTCATA AGTTGTAAAG GCACAGCGGG AATAAGGGAA AAAACGCCGT
7501  TTTTGCCTTC CTGTTTTTGC TCACCCAGAA ACGCTGGTGA AAGTAAAAGA
      AAAACGGAAG GACAAAAACG AGTGGGTCTT TGCGACCACT TTCATTTTCT
7551  TGCTGAAGAT CAGTTGGGTG CACGAGTGGG TTACATCGAA CTGGATCTCA
      ACGACTTCTA GTCAACCCAC GTGCTCACCC AATGTAGCTT GACCTAGAGT
7601  ACAGCGGTAA GATCCTTGAG AGTTTTCGCC CCGAAGAACG TTTTCCAATG
      TGTCGCCATT CTAGGAACTC TCAAAAGCGG GGCTTCTTGC AAAAGGTTAC
7651  ATGAGCACTT TTAAAGTTCT GCTATGTGGC GCGGTATTAT CCCGTATTGA
      TACTCGTGAA AATTTCAAGA CGATACACCG CGCCATAATA GGGCATAACT
7701  CGCCGGGCAA GAGCAACTCG GTCGCCGCAT ACACTATTCT CAGAATGACT
      GCGGCCCGTT CTCGTTGAGC CAGCGGCGTA TGTGATAAGA GTCTTACTGA
7751  TGGTTGAGTA CTCACCAGTC ACAGAAAAGC ATCTTACGGA TGGCATGACA
      ACCAACTCAT GAGTGGTCAG TGTCTTTTCG TAGAATGCCT ACCGTACTGT
7801  GTAAGAGAAT TATGCAGTGC TGCCATAACC ATGAGTGATA ACACTGCGGC
      CATTCTCTTA ATACGTCACG ACGGTATTGG TACTCACTAT TGTGACGCCG
7851  CAACTTACTT CTGACAACGA TCGGAGGACC GAAGGAGCTA ACCGCTTTTT
      GTTGAATGAA GACTGTTGCT AGCCTCCTGG CTTCCTCGAT TGGCGAAAAA
7901  TGCACAACAT GGGGGATCAT GTAACTCGCC TTGATCGTTG GGAACCGGAG
      ACGTGTTGTA CCCCCTAGTA CATTGAGCGG AACTAGCAAC CCTTGGCCTC
7951  CTGAATGAAG CCATACCAAA CGACGAGCGT GACACCACGA TGCCTGTAGC
      GACTTACTTC GGTATGGTTT GCTGCTCGCA CTGTGGTGCT ACGGACATCG
8001  AATGGCAACA ACGTTGCGCA AACTATTAAC TGGCGAACTA CTTACTCTAG
      TTACCGTTGT TGCAACGCGT TTGATAATTG ACCGCTTGAT GAATGAGATC
8051  CTTCCCGGCA ACAATTAATA GACTGGATGG AGGCGGATAA AGTTGCAGGA
      GAAGGGCCGT TGTTAATTAT CTGACCTACC TCCGCCTATT TCAACGTCCT
8101  CCACTTCTGC GCTCGGCCCT TCCGGCTGGC TGGTTTATTG CTGATAAATC
      GGTGAAGACG CGAGCCGGGA AGGCCGACCG ACCAAATAAC GACTATTTAG
8151  TGGAGCCGGT GAGCGTGGGT CTCGCGGTAT CATTGCAGCA CTGGGGCCAG
      ACCTCGGCCA CTCGCACCCA GAGCGCCATA GTAACGTCGT GACCCCGGTC
8201  ATGGTAAGCC CTCCCGTATC GTAGTTATCT ACACGACGGG GAGTCAGGCA
      TACCATTCGG GAGGGCATAG CATCAATAGA TGTGCTGCCC CTCAGTCCGT
8251  ACTATGGATG AACGAAATAG ACAGATCGCT GAGATAGGTG CCTCACTGAT
      TGATACCTAC TTGCTTTATC TGTCTAGCGA CTCTATCCAC GGAGTGACTA
8301  TAAGCATTGG TAACTGTCAG ACTCGCGACA CTGCATTAAT GAATCGGCCA
      ATTCGTAACC ATTGACAGTC TGAGCGCTGT GACGTAATTA CTTAGCCGGT
8351  ACGCGCGGGG AGAGGCGGTT TGCGTATTGG GCGCTCTTCC GCTTCCTCGC
      TGCGCGCCCC TCTCCGCCAA ACGCATAACC CGCGAGAAGG CGAAGGAGCG
8401  TCACTGACTC GCTGCGCTCG GTCGTTCGGC TGCGGCGAGC GGTATCAGCT
      AGTGACTGAG CGACGCGAGC CAGCAAGCCG ACGCCGCTCG CCATAGTCGA
8451  CACTCAAAGG CGGTAATACG GTTATCCACA GAATCAGGGG ATAACGCAGG
      GTGAGTTTCC GCCATTATGC CAATAGGTGT CTTAGTCCCC TATTGCGTCC
8501  AAAGAACATG TGAGCAAAAG GCCAGCAAAA GGCCAGGAAC CGTAAAAAGG
      TTTCTTGTAC ACTCGTTTTC CGGTCGTTTT CCGGTCCTTG GCATTTTTCC
8551  CCGCGTTGCT GGCGTTTTTC CATAGGCTCC GCCCCCCTGA CGAGCATCAC
      GGCGCAACGA CCGCAAAAAG GTATCCGAGG CGGGGGGACT GCTCGTAGTG
```

```
8601   AAAAATCGAC GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAAG
       TTTTTAGCTG CGAGTTCAGT CTCCACCGCT TTGGGCTGTC CTGATATTTC
8651   ATACCAGGCG TTTCCCCCTG GAAGCTCCCT CGTGCGCTCT CCTGTTCCGA
       TATGGTCCGC AAAGGGGGAC CTTCGAGGGA GCACGCGAGA GGACAAGGCT
8701   CCCTGCCGCT TACCGGATAC CTGTCCGCCT TTCTCCCTTC GGGAAGCGTG
       GGGACGGCGA ATGGCCTATG GACAGGCGGA AAGAGGGAAG CCCTTCGCAC
8751   GCGCTTTCTC ATAGCTCACG CTGTAGGTAT CTCAGTTCGG TGTAGGTCGT
       CGCGAAAGAG TATCGAGTGC GACATCCATA GAGTCAAGCC ACATCCAGCA
8801   TCGCTCCAAG CTGGGCTGTG TGCACGAACC CCCCGTTCAG CCCGACCGCT
       AGCGAGGTTC GACCCGACAC ACGTGCTTGG GGGGCAAGTC GGGCTGGCGA
8851   GCGCCTTATC CGGTAACTAT CGTCTTGAGT CCAACCCGGT AAGACACGAC
       CGCGGAATAG GCCATTGATA GCAGAACTCA GGTTGGGCCA TTCTGTGCTG
8901   TTATCGCCAC TGGCAGCAGC CACTGGTAAC AGGATTAGCA GAGCGAGGTA
       AATAGCGGTG ACCGTCGTCG GTGACCATTG TCCTAATCGT CTCGCTCCAT
8951   TGTAGGCGGT GCTACAGAGT TCTTGAAGTG GTGGCCTAAC TACGGCTACA
       ACATCCGCCA CGATGTCTCA AGAACTTCAC CACCGGATTG ATGCCGATGT
9001   CTAGAAGGAC AGTATTTGGT ATCTGCGCTC TGCTGAAGCC AGTTACCTTC
       GATCTTCCTG TCATAAACCA TAGACGCGAG ACGACTTCGG TCAATGGAAG
9051   GGAAAAAGAG TTGGTAGCTC TTGATCCGGC AAACAAACCA CCGCTGGTAG
       CCTTTTTCTC AACCATCGAG AACTAGGCCG TTTGTTTGGT GGCGACCATC
9101   CGGTGGTTTT TTTGTTTGCA AGCAGCAGAT TACGCGCAGA AAAAAAGGAT
       GCCACCAAAA AAACAAACGT TCGTCGTCTA ATGCGCGTCT TTTTTTCCTA
9151   CTCAAGAAGA TCCTTTGATC TTTTCTACGG GGTCTGACGC TCAGTGGAAC
       GAGTTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG
9201   GAAAACTCA
       CTTTTGAGT
```

FIG. 31(contd.)