

(19) World Intellectual Property Organization
International Bureau(43) International Publication Date
17 October 2002 (17.10.2002)

PCT

(10) International Publication Number
WO 02/081677 A2(51) International Patent Classification⁷: **C12N 15/00**

(21) International Application Number: PCT/GB02/01479

(22) International Filing Date: 5 April 2002 (05.04.2002)

(25) Filing Language: English

(26) Publication Language: English

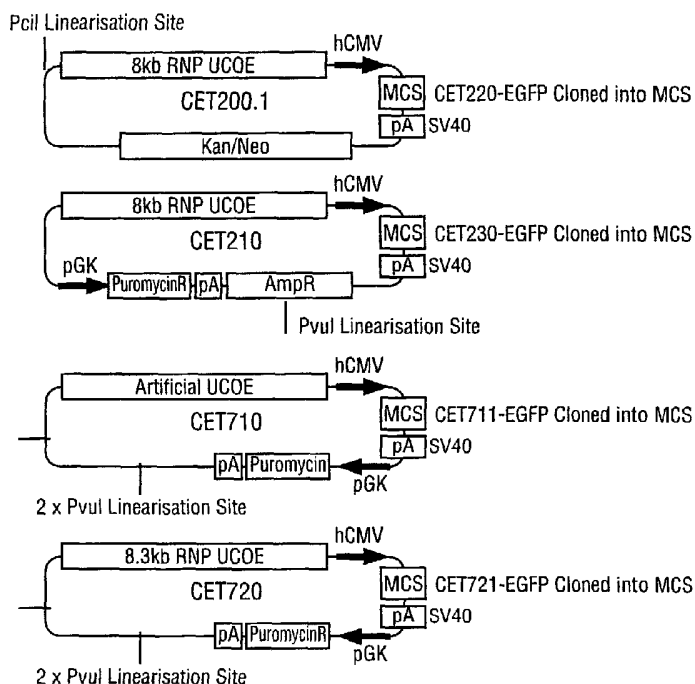
(30) Priority Data:

0108605.7	5 April 2001 (05.04.2001)	GB
60/281,605	5 April 2001 (05.04.2001)	US
0109335.0	17 April 2001 (17.04.2001)	GB
60/298,675	15 June 2001 (15.06.2001)	US

(71) Applicant: **COBRA THERAPEUTICS LIMITED**
[GB/GB]; Stephenson Building, Keele University Science Park, Keele, Staffordshire ST5 5SP (GB).(72) Inventors: **WILLIAMS, Stephen, Geraint**; Cobra Therapeutics Limited, Stephenson Building, Keele University Science Park, Keele, Staffordshire ST5 5SP (GB).**CROMBIE, Robert, Lachlan**; Cobra Therapeutics Limited, Stephenson Building, Keele University Science Park, Keele, Staffordshire ST5 5SP (GB).(74) Agents: **JONES, Keith, William**; Harrison Goddard Foote, Orlando House, 11c Compstall Road, Marple Bridge, Stockport SK6 5HH et al. (GB).(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent

[Continued on next page]

(54) Title: IMPROVED GENE EXPRESSION



(57) Abstract: Polynucleotides and vectors comprising an expressible nucleic acid flanked by a 5' extended methylation-free CpG island and a 3' selectable marker element are disclosed. Such polynucleotides and vectors provide a means for obtaining high levels of expression of the flanked expressible nucleic acid. Preferred embodiments include combinations of 5' extended methylation-free CpG islands and 3' antibiotic resistance genes.



WO 02/081677 A2



(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *without international search report and to be republished upon receipt of that report*

IMPROVED GENE EXPRESSION

FIELD OF THE INVENTION

The present invention relates to a polynucleotide comprising a ubiquitous
5 chromatin-opening element (UCOE) together with a selectable marker
element. When operably linked to, and flanking, an expressible nucleic acid
sequence, the combination of elements provides high and reproducible
levels of gene expression. The present invention also relates to a vector
comprising the polynucleotide sequence, a host cell comprising the vector
10 and use of the polynucleotide, vector or host cell in therapy, or for
applications involving protein expression in cell culture.

BACKGROUND OF THE INVENTION

The current model of chromatin structure in higher eukaryotes postulates
15 that genes are organised in "domains" (Dillon, N. & Grosveld, F. *Chromatin*
domains as potential units of eukaryotic gene function. Curr. Opin. Genet.
Dev. 4, 260-264 (1994); Higgs, D.R. Do LCRs open chromatin domains? *Cell*
95, 299-302 (1998)) Chromatin domains are envisaged to exist in either a
condensed, "closed", transcriptionally silent state, or in a de-condensed,
20 "open" and transcriptionally competent configuration. The establishment of
an open chromatin structure characterised by increased DNaseI sensitivity,
DNA hypomethylation and histone hyperacetylation, is considered a pre-
requisite to the commencement of gene expression.

25 The open and closed nature of chromatin regions is reflected in the
behaviour of transgenes that are randomly integrated into the host cell
genome. Identical constructs give different patterns of tissue-specific and
development stage-specific expression when integrated at different locations
in the mouse genome (Palmiter, R.D. & Brinster, R.L. *Ann. Ref. Genet.* 20,
30 465-499 (1986); Allen, N.D. *et al. Nature* 333, 852-855 (1988); Bonnerot, C.,
Grimber, G., Briand, P. & Nicolas, J.F. *Proc. Natl. Acad. Sci. USA* 87:6331-
6335 (1990)).

A variegated expression pattern within a given transgenic mouse tissue, known as position effect variegation (PEV), is also frequently observed (Kioussis, D. & Festenstein, R. *Curr. Opin. Genet. Dev.* 7, 614-619 (1997)).

5 When exogenous genes are integrated into the chromosome of mammalian cells cultures *in vitro*, many of the integration events result in rapid silencing of the transgene and the remainder give large variability in expression levels (Pikaart, , M.J., Recillas-Targa, F. & Felsenfield, G. *Genes Dev.* 12, 2852-2862 (1998); Fussenegger, M., Bailey, J.E., Hauser, H. & Mueller, P.P

10 *Trends Biotech.* 17, 35-42 (1999)). These position effects render transgene expression inefficient, with implication for both basic research and biotechnology applications.

The chromatin domain model of gene organisation suggests that genetic

15 control elements that are able to establish and maintain a transcriptionally competent open chromatin structure should be associated with active regions of the genome.

Locus Control Regions (LCRs) are a class of transcriptional regulatory

20 elements with long-range chromatin remodelling capability. LCRs are functionally defined in transgenic mice by their ability to confer site-of-integration independent, transgene copy number-dependent, physiological levels of expression on a gene linked in *cis*., especially single copy transgenes Fraser, P. & Grosveld, F. *Curr. Opin. Cell Biol.* 10, 361-365

25 (1998); Li, Q., Harju, S. & Peterson, K.R. *Trends Genet.* 15: 403-408 (1999). Crucially, such expression is tissue-specific. LCRs are able to obstruct the spread of heterochromatin, prevent PEV (Kioussis, D. & Festenstein, R. *Curr. Opin. Genet. Dev.* 7, 614-619 (1997)) and consist of a series of DNase I hypersensitive (HS) sites which can be located either 5' or 3' of the genes

30 that they regulate (Li, Q., Harju, S. & Peterson, K.R. *Trends Genet.* 15: 403-408 (1999)).

LCRs appear to be comprised of two separate, although not necessarily independent components. First, the establishment of an 'open chromatin domain', and second a dominant transcriptional activation capacity to confer transgene copy number dependent expression (Fraser, P. & Grosveld, F. *Curr. Opin. Cell Biol.* 10, 361-365 (1998). The molecular mechanisms by which LCRs exert their function remain a point of contention (Higgs, D.R. *Cell* 95, 299-302 (1998); Bulger, M. & Groudine, M. *Genes Dev.* 13, 2465-2477 (1999); Grosveld, F. *Curr. Opin. Genet. Dev.* 9 152-157 (1999); Bender, M.A., Bulger, M., Close, J. & Groudine, M.. *Mol. Cell* 5, 387-393 (2000).

The generation of cultured mammalian cell lines producing high levels of a therapeutic protein product is a major developing industry. Chromatin position effects make it a difficult, time consuming and expensive process. The most commonly used approach to the production of such mammalian "cell factories" relies on gene amplification induced by a combination of a drug resistance gene (e.g., DHFR, glutamine synthetase (Kaufman RJ. *Methods Enzymol* 185, 537-566 (1990)). and the maintenance of stringent selective pressure. The use of vectors containing LCRs from highly expressed gene domains, using cells derived from the appropriate tissue, greatly simplifies the procedure, giving a large proportion of clonal cell lines showing stable high levels of expression (Needham M, Gooding C, Hudson K, Antoniou M, Grosfeld F and Hollis M. *Nucleic Acids Res* 20, 997-1003 (1992); Needham M, Egerton M, Millest A, Evans S, Popplewell M, Cerillo G, McPheat J, Monk A, Jack A, Johnstone D and Hollis M. *Protein Expr Purif* 6, 124-131 (1995).

However, the tissue-specificity of LCRs, although useful in some circumstances, is also a major limitation for many applications, for instance where no LCR is known for the tissue in which expression is required, or where expression in many, or all, tissues is required.

Our co-pending patent applications PCT/GB99/02357 (WO 00/05393), US 09/358082, GB 0022995.5 and US 60/252,048 incorporated by reference herein, describe elements that are responsible, in their natural chromosomal context, for establishing an open chromatin structure across a locus that
5 consists exclusively of ubiquitously expressed, housekeeping genes. These elements are not derived from an LCR and comprise extended methylation-free CpG islands. We have used the term Ubiquitous Chromatin Opening Element (UCOE) to describe such elements.

10 In mammalian DNA, the dinucleotide CpG is recognised by a DNA methyltransferase enzyme that methylates cytosine to 5-methylcytosine. However, 5-methylcytosine is unstable and is converted to thymine. As a result, CpG dinucleotides occur far less frequently than one would expect by chance. Some sections of genomic DNA nevertheless do have a frequency
15 of CpG that is closer to that expected, and these sequences are known as "CpG islands". As used herein a "CpG island" is defined as a sequence of DNA, of at least 200bp, that has a GC content of at least 50% and an observed / expected CpG content ratio of at least 0.6 (i.e. a CpG dinucleotide content of at least 60% of that which would be expected by
20 chance) (Gardiner-Green M and Frommer M. *J Mol Biol* 196, 261–282 (1987); Rice P, Longden I and Bleasby A *Trends Genet* 16, 276–277 (2000).

Methylation-free CpG islands are well-known in the art (Bird *et al* (1985) *Cell* 40: 91-99, Tazi and Bird (1990) *Cell* 60: 909-920) and may be defined as
25 CpG islands where a substantial proportion of the cytosine residues are not methylated and which usually extend over the 5' ends of two closely spaced (0.1-3 kb) divergently transcribed genes. These regions of DNA are reported to remain hypomethylated in all tissues throughout development (Wise and Pravtcheva (1999) *Genomics* 60: 258-271). They are often associated with
30 the 5' ends of ubiquitously expressed genes, as well as an estimated 40% of genes showing a tissue-restricted expression profile (Antequera, F. & Bird, A. *Proc. Natl. Acad. Sci. USA* 90, 1195-11999 (1993); Cross, S.H. & Bird, A.P.

Curr. Opin. Genet. Dev. 5, 309-314 (1995) and are known to be localised regions of active chromatin (Tazi, J. & Bird, A. *Cell* 60, 909-920 (1990).

An 'extended' methylation-free CpG island is a methylation-free CpG island
5 that extends across a region encompassing more than one transcriptional
start site and/or extends for more than 300bp and preferably more than
500bp. The borders of the extended methylation-free CpG island are
functionally defined through the use of PCR over the region in combination
10 with restriction endonuclease enzymes whose ability to digest (cut) DNA at
their recognition sequence is sensitive to the methylation status of any CpG
residues that are present. One such enzyme is *HpaII*, which recognises and
digests at the site CCGG, which is commonly found within CpG islands, but
only if the central CG residues are *not* methylated. Therefore, PCR
conducted with *HpaII*-digested DNA and over a region harbouring *HpaII*
15 sites, does *not* give an amplification product due to *HpaII* digestion if the
DNA is *unmethylated*. The PCR will *only* give an amplified product if the DNA
is *methylated*. Therefore, beyond the methylation-free region *HpaII* will not
digest the DNA a PCR amplified product will be observed thereby defining
the boundaries of the "extended methylation-free CpG island".

20

We have demonstrated (WO 00/05393) that regions spanning methylation-free CpG islands encompassing dual, divergently transcribed promoters from the human TATA binding protein (*TBP*)/proteasome component-B1 (*PSMB1*) and heterogeneous nuclear ribonucleoprotein A2/B1
25 (*hnRNPA2*)/heterochromatin protein 1Hs γ (*HP1*^{Hs γ}) gene loci give
reproducible, physiological levels of gene expression and that they are able
to prevent a variegated expression pattern and silencing that normally occurs
with transgene integration within centromeric heterochromatin.

30 As used herein, the term "reproducible expression" means that the polynucleotide of the invention will direct expression of the expressible gene at substantially the same level of expression irrespective of its chromatin

environment and preferably irrespective of the cell type or tissue type in which the polynucleotide of the invention may be. Those of skill in the art will recognize that substantially the same level of expression of the operably-linked expressible gene is achieved, irrespective of the chromatin
5 environment of the claimed polynucleotide, and preferably irrespective of the cell type, assuming that the cell is capable of active gene expression.

We have shown (WO 00/05393) that methylation-free CpG islands associated with actively transcribing promoters possess the ability to remodel
10 chromatin and are thus thought to be a prime determinant in establishing and maintaining an open domain at housekeeping gene loci.

UCOEs confer an increased proportion of productive gene delivery events with improvements in the level and stability of transgene expression. This
15 has important research and biotechnological applications including the generation of transgenic animals and recombinant protein products in cultured cells. We have shown (WO 00/05393) beneficial effects of UCOEs on expression of the CMV-EGFP reporter construct and with the secreted, pharmaceutically valuable protein erythropoietin. The properties of UCOEs
20 also suggest utility in gene therapy, the effectiveness of which is often limited by a low frequency of productive gene delivery events and an inadequate level and duration of expression (Verma, I.M. & Somia, N. *Nature* 389: 239-242 (1997).

25 Given these significant implications and wide ranging applications, there is a desire to further optimise transgene expression levels. There is a need to further increase the levels of expression obtainable by the use of a UCOE alone, particularly in the fields of *in vivo* gene therapy and for *in vitro* production of recombinant proteins.

30

The expression of a nucleic acid operably linked to a 5' UCOE may surprisingly be further increased by the presence of a selectable element 3'

to the expressed nucleic acid, so that the expressible nucleic acid sequence is flanked by a 5' UCOE and a 3' selectable marker .

5 A selectable element that performs more than one function in a vector, such as providing a selectable marker as well as increasing expression of an operably linked gene, allows construction of more compact and efficient expression vectors.

10 Mei, Kothary and Wall (Mei, Q, Kothary R and Wall L. *Exp Cell Research* 260, 304–312 (2000) disclose constructs comprising an expressible gene (β -globin) operably linked to an LCR and a *pgk* / puromycin resistance element. However, this work teaches that it is the combination of an expressible gene, and LCR and a *tk* / neomycin resistance element that is important in imposing position effects on gene expression, with the *pgk* / puromycin
15 resistance element being used as a negative control. This paper teaches away from any beneficial effect being gained from the use of a *pgk* / puromycin resistance element. The paper does not disclose constructs comprising an extended unmethylated CpG island (or UCOE), an expressible gene and a *pgk* / puromycin resistance element, since the constructs
20 comprise LCRs. Similarly, the paper does not disclose an expressible gene operably linked to a promoter with which it is not naturally linked, also operably linked to a *pgk* / puromycin resistance element, since in each case the β -globin gene is expressed under control of its endogenous promoter.

25 Artelt *et al* compare the influence of neomycin and puromycin resistance genes on *cis*-linked genes in eukaryotic expression vectors (Artelt P, Grannemann R, Stocking C, Friel J, Bartsch J and Hauser H *Gene* 99, 249–254 (1991). They conclude that neomycin resistance genes may have a silencing effect on linked genes, but that “the gene conferring resistance to
30 puromycin from *Streptomyces alboniger* does not influence adjacent promoters”. Accordingly, there is nothing in this paper that discloses or

suggests the importance of the position or spacing use of resistance genes as disclosed in the present application.

Our co-pending patent applications PCT/GB99/02357 (WO 00/05393), US
5 09/358082, GB 0022995.5 and US 60/252,048 disclose polynucleotides and
vectors comprising extended, methylation-free CpG islands operably linked
to expressible nucleic acids with antibiotic resistance genes. However, in the
examples disclosed, the antibiotic gene is not adjacent and 3' to the
expressible nucleic acid. The surprising contribution of such an adjacent
10 selectable marker is likewise not disclosed or implied.

STATEMENTS OF THE INVENTION

15 The present invention discloses that the influence of extended, unmethylated
CpG islands (UCOEs) to upregulate expression of operably linked nucleic
acid sequences may be further increased by the presence of a selectable
element providing that said selectable marker is situated 3' of the expressible
nucleic acid sequence and adjacent to it.

20

The terms 5' and 3' are herein used with respect to the sense strand of the
expressible nucleic acid sequence. Hence the 5' end of said sequence
corresponds to the start of transcription, which proceeds in a 3' direction.

25 As used herein, the term "operably linked" refers to a relationship of
operability between elements in the polynucleotides of the invention.
"Operably linked" is a term, well known to those of skill in the art, that
describes a functional relationship between *cis*-acting DNA sequences. The
exact structural relationship may or may not be relevant and differs for
30 different types of elements. For a promoter, it implies an essentially adjacent
(usually within less than 100bp) position 5' to the open reading frame that it
drives. In the case of extended methylation-free CpG islands, it appears that
a regional effect on chromatin structure is responsible for increasing the level

and consistency of gene expression. By way of example, the element comprising an extended methylation-free CpG-island is positioned immediately 5' of the expressible gene. However, "operably-linked" embraces the possibility of being positioned elsewhere, as long as a clear
5 functional effect can be demonstrated.

In particular, the flanking of an expressible gene with a UCOE at the 5' end and a selectable element at the other results in an increase in expression of approximately two-fold. In some cases the increase is greater than five-fold
10 over that obtained with a single UCOE alone.

According to the present invention, there is provided an isolated polynucleotide that enables increased levels of expression of an operably linked gene to be obtained as compared to those obtainable using an
15 operably-linked UCOE or extended methylation-free CpG island alone.

The isolated polynucleotide comprises: an extended methylation-free CpG island, an expressible nucleic acid terminated by a polyadenylation signal and a selectable marker operably linked to a promoter, wherein both the
20 CpG island and the selectable marker are operably-linked to the expressible nucleic acid, and the components are positioned in the order: extended methylation-free CpG island, expressible nucleic acid, selectable marker, in the 5' to 3' orientation with respect to the sense strand of the expressible nucleic acid, and the polyadenylation signal at the 3' end of the expressible
25 nucleic acid is within 2000bp of the proximal end of the selectable marker.

As used herein, "proximal end" means the end of the selectable marker gene (including its promoter) that is closest to the 3' end of the expressible nucleic acid, as marked by its polyadenylation signal. It is envisaged that the
30 selectable marker might be in either orientation, so that the proximal end relative to the expressible nucleic acid might be at either the 5' promoter end

of the selectable marker or the 3', termination of transcription end, taking 5' and 3' as being according to the sense strand of the selectable marker.

Preferably, the transcriptional start of the selectable marker is within 1500bp
5 of the 3' end of the expressible nucleic acid sequence, as marked by its polyadenylation signal of the latter. More preferably, it is within 1000bp. Most preferably it is within 500bp.

In one aspect of the invention, the selectable element is an antibiotic resistance gene. Preferably it is an antibiotic resistance gene obtained from
10 a *Streptomyces* species. More preferably, said antibiotic resistance gene is operably linked to a promoter of the phosphoglycerate kinase (*pgk*) gene. Most preferably, it is the promoter of the murine *pgk* gene (Adra, CN, Boer PH and McBurney, MW. *Gene* 60, 65–74 (1987). Alternatively, it may be another mammalian *pgk* promoter.

15 In a preferred embodiment, the antibiotic resistance gene is the puromycin resistance gene from a *Streptomyces* species. Most preferably, it is the puromycin N-acetyl transferase gene from *Streptomyces alboniger* (Vara JA, Portela A, Ortin J, Jimenez A. *Nucleic Acids Res* 14, 4617–4624 (1986)

20 Alternatively, the antibiotic resistance gene is a modified form of the puromycin N-acetyl transferase gene from *Streptomyces alboniger*. Preferably this gene has been modified by manipulation of its codon usage, in a manner commonly done to adapt bacterial genes for expression in
25 mammalian host cells. Such codon modification leaves the encoded amino acid sequence unchanged, with the result that the expressed enzyme is unchanged from the wild-type puromycin N-acetyl transferase. Most preferably, the modified gene has the sequence shown in Figure 15.

30 Alternatively, the antibiotic resistance gene is a neomycin resistance gene derived from a *Streptomyces* species. Preferably it is the aminoglycoside

phosphotransferase gene from *Streptomyces fradiae* (Thompson CJ and Gray GS. *Proc Natl Acad Sci USA* 80, 5190–5194 (1983).

5 In an alternative embodiment, the antibiotic resistance gene is a hygromycin resistance gene. Preferably, it is the hygromycin phosphotransferase gene from *Streptomyces hygrosopicus*.

10 In a further alternative embodiment, the antibiotic resistance gene is a bleomycin resistance gene. Preferably, it is the bleomycin binding protein from *Streptomyces verticillus*. Alternatively, it is the bleomycin N-acetyltransferase from *Streptomyces verticillus*.

15 In another embodiment, the antibiotic resistance gene is a blasticidin resistance gene. Preferably, it is the blasticidin S-acetyltransferase gene from *Streptomyces verticillum*.

20 In another aspect of the invention, the antibiotic resistance gene is not obtained from a *Streptomyces* species. In one preferred embodiment it is the hygromycin resistance gene encoding aminocyclitol phosphotransferase from *Escherichia coli*.

In another preferred embodiment, it is the neomycin phosphotransferase gene from transposon Tn5, originally derived from *Klebsiella pneumoniae*.

25 In an alternative aspect of the invention, the selectable marker is not an antibiotic resistance gene. Alternative selection mechanisms involve using genes encoding thymidylate synthase, thymidine kinase or dihydrofolate reductase. Such selection mechanisms are well-known to those of appropriate skill in the art. In a medium lacking methionine, a gene encoding
30 glutamine synthetase may be used as a means of selection either in cells lacking an endogenous glutamine synthetase, or where use of an inhibitor, such as methionine sulfoxamine, has rendered it inactive (Kaufman RJ.

Selection and coamplification of heterologous genes in mammalian cells.
Methods Enzymol 185, 537–566 (1990).

5 In a further aspect, a screenable marker can be used. For instance, a gene
encoding a fluorescent protein, such as the *Aequoria victoria* green
fluorescent protein (GFP), or enhanced variants of it (EGFP), may be used
as a selectable marker. Transfectants containing a polynucleotide according
to the current invention, wherein the selectable marker encodes GFP, may
10 be sorted by brightness of fluorescence on a FACS, by a process well-known
in the art. Using the polynucleotide of the invention, and comparing with
expressible constructs with the selectable marker situated either 5' to the
UCOE, or 3' but remotely from the transgene (expressible nucleic acid),
higher levels of expression of the transgene will be found for comparable
levels of brightness. Selection of the brightest cells will, therefore, allow
15 selection of cells with the highest level of transgene expression.

In one aspect of the invention, the extended methylation-free CpG island
comprises a 16kb DNA fragment spanning the human hnRNP A2 gene with
5kb 5' and 1.5kb 3' flanking sequence. Preferably, the extended
20 methylation-free CpG island comprises an 8kb DNA fragment spanning the
human hnRNP A2 gene (WO 00/05393).

Alternatively, the extended methylation-free CpG island of the disclosed
polynucleotide is an 'artificial UCOE' as disclosed in our co-pending
25 applications GB 0022995.5 and US 60/252,048, comprising the human β -
actin CpG island/ promoter region or a fragment thereof. Preferably this
fragment is within the size range of 100bp to 3.0 kb and spans the human β -
actin CpG island/ promoter region or a fragment thereof. Preferably the
artificial UCOE also comprises the human PDCD2 CpG island/ promoter
30 region or a fragment thereof. More preferably the human PDCD2 CpG
island/ promoter region comprises a fragment within the size range of 100bp
to 3.0 kb. Further preferably, the extended methylation-free CpG island

comprises a DNA fragment within the size range of 100bp to 3.0 kb spanning the human β -actin CpG island/ promoter region and a DNA fragment within the size range of 100bp to 3.0 kb spanning the human PDCCD2 CpG island/ promoter region.

5

Most preferably the claimed polynucleotide of this embodiment of the invention comprises an artificial UCOE comprising a 2.0 kb DNA fragment spanning the human β -actin CpG island/ promoter region and a 1.8 kb DNA fragment spanning the human PDCCD2 CpG island/ promoter region.

10

Also provided is a vector comprising the polynucleotide of any one of the previous embodiments. This vector may alternatively be either an episomal or an integrating vector. Depending on the intended use, episomal vectors may be desirable since they are self-replicating and so persist without the need for integration. Episomal vectors of this type are described in WO98/07876. Also preferred are non-replicating, non-integrating vectors.

15

Also provided is a vector so constructed as to deliver, when linearised and integrated into a chromosome, a polynucleotide comprising an extended methylation-free CpG island, an expressible nucleic acid terminated by a polyadenylation signal and a selectable marker operably linked to a promoter, wherein both the CpG island and the selectable marker are operably-linked to the expressible nucleic acid, and the components are positioned in the order: extended methylation-free CpG island, expressible nucleic acid, selectable marker, in the 5' to 3' orientation with respect to the sense strand of the expressible nucleic acid, and the polyadenylation signal at the 3' end of the expressible nucleic acid is within 2000bp of the proximal end of the selectable marker.

20

25

Preferably the vector is a plasmid. Alternatively, the vector may be a virus, such as an adenovirus, adeno-associated virus, a herpesvirus, vaccinia virus, lentivirus or other retrovirus.

30

- Preferably said vector is an expression vector adapted for eukaryotic gene expression. Typically said adaptation includes, by example and not by way of limitation, the provision of transcription control sequences (promoter sequences) which mediate cell/tissue specific expression. Promoter and enhancer are terms well-known in the art and include the following features which are provided by example only, and not by way of limitation. Promoters are 5', *cis*-acting regulatory sequences directly linked to the initiation of transcription. Promoter elements include so-called TATA box and RNA polymerase initiation selection (RIS) sequences that function to select a site of transcription initiation. These sequences also bind polypeptides that function, *inter alia*, to facilitate transcription initiation selection by RNA polymerase.
- Enhancer elements are *cis* acting nucleic acid sequences often found 5' to the transcription initiation site of a gene (enhancers can also be found 3' to a gene sequence or even located in intronic sequences and are therefore position independent). Enhancers function to increase the rate of transcription of the gene to which the enhancer is linked. Enhancer activity is responsive to *trans* acting transcription factors (polypeptides) which have been shown to bind specifically to enhancer elements. The binding/activity of transcription factors is responsive to a number of environmental cues which include, by way of example and not by way of limitation, intermediary metabolites (e.g. glucose), environmental effectors (e.g. heat,). (See Eukaryotic Transcription Factors, by David S Latchman, Academic Press Ltd, San Diego)

Adaptations also include the provision of selectable markers and autonomous replication sequences which both facilitate the maintenance of said vector in either the eukaryotic cell or prokaryotic host. Vectors that are maintained autonomously in eukaryotic cells are referred to as episomal vectors. Other adaptations which facilitate the expression of vector encoded

genes include the provision of transcription termination/polyadenylation sequences. This also includes the provision of internal ribosome entry sites (IRES) which function to maximise expression of vector encoded genes arranged in bicistronic or multi-cistronic expression cassettes. These adaptations are well-known in the art. There is a significant amount of published literature with respect to expression vector construction and recombinant DNA techniques in general. Please see, Sambrook et al (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbour Laboratory, Cold Spring Harbour, NY and references therein; Marston, F (1987) DNA Cloning Techniques: A Practical Approach Vol. III IRL Press, Oxford UK; DNA Cloning: F M Ausubel et al, Current Protocols in Molecular Biology, John Wiley & Sons, Inc.(1994).

In a preferred method of the invention said vector encodes, and thus said polypeptide is provided with, a secretion signal to facilitate purification of said polypeptide.

Alternatively, other preferred embodiments may include further refinements to facilitate purification of expressed recombinant protein, such as affinity tags or epitopes, or enzymatic cleavage sites.

Preferably the expressible nucleic acid is a therapeutic nucleic acid.

Alternatively, the expressible nucleic acid encodes a recombinant protein for expression in an *in vitro* cell culture system.

Alternatively, the expressible gene encodes a non-polypeptide product, such as RNA. Such RNA may be an antisense RNA capable of inhibiting expression of a particular gene at a post-transcriptional level, or may have an enzymatic (ribozyme) or other function, such as a ribosomal RNA.

One preferred embodiment is a vector comprising: an extended methylation-free CpG island, an expressible nucleic acid terminated by a polyadenylation

signal and a selectable marker operably linked to a promoter, wherein both the CpG island and the selectable marker are operably-linked to the expressible nucleic acid, and the components are positioned in the order: extended methylation-free CpG island, expressible nucleic acid, selectable
5 marker, in the 5' to 3' orientation with respect to the sense strand of the expressible nucleic acid, and the polyadenylation signal at the 3' end of the expressible nucleic acid is within 2000bp of the proximal end of the selectable marker. Preferably, the the polyadenylation signal at the 3' end of the
10 the expressible nucleic acid is within 1500bp of the proximal end of the selectable marker. More preferably it is within 1000bp, most preferably, 500bp.

A preferred embodiment is a vector comprising: an extended methylation-free CpG island, a multiple cloning site, an antibiotic resistance gene
15 obtained from a *Streptomyces* species, wherein both the CpG island and the selectable marker are operably-linked to the multiple cloning site, and the components are positioned in the order: extended methylation-free CpG island, multiple cloning site, selectable marker, in the 5' to 3' orientation with respect to the sense strand of the expressible nucleic acid, and the multiple
20 cloning site is within 2000bp of the proximal end of the selectable marker.

More preferably, the multiple cloning site is further operably linked to a promoter. Further preferably the promoter is selected from CMV, EF-1 α , RSV LTR or HIV2 LTR or combinations of sequences derived therefrom.
25 More preferably the promoter is a CMV immediate/early promoter. Most preferably it is the mouse CMV immediate/early promoter. In a preferred embodiment, the vector comprises a CMV promoter, a multiple cloning site, a polyadenylation sequence and genes encoding selectable markers under suitable control elements.

30

A preferred embodiment of the vector comprises nucleotides 1–10551 of the sequence of Figure 9. A most preferred embodiment is vector CET 710.

Alternatively, the vector comprises nucleotides 1–13545 of the sequence of Figure 10, and is preferably vector CET 720.

5 Further preferred embodiment of vectors are:

CET 740 in which the puromycin resistance gene of CET 720 is replaced with the aminoglycoside phosphotransferase gene from *Streptomyces fradiae* (as listed in Figure 15). Also preferred are vectors having expressible
10 nucleic acid sequences inserted into the multiple cloning site of CET 740, such as CET 741.

CET 760 in which the puromycin resistance gene of CET 720 is replaced with the aminocyclitol phosphotransferase from *Escherichia coli* (as listed in
15 Figure 17). Also preferred are vectors having expressible nucleic acid sequences inserted into the multiple cloning site of CET 760, such as CET 761.

CET 780 in which the puromycin resistance gene of CET 720 is replaced
20 with the modified form of the puromycin N-acetyl transferase gene from *Streptomyces alboniger* (as listed in Figure 14). Also preferred are vectors having expressible nucleic acid sequences inserted into the multiple cloning site of CET 780, such as CET 781.

25 CET 820 in which the human IE CMV promoter, operably linked to the multicloning site in order to drive expression of expressible nucleic acid sequences inserted there, has been replaced with the murine IE CMV promoter. Also preferred are vectors having expressible nucleic acid sequences inserted into the multiple cloning site of CET 820, such as CET
30 821.

CET 823 in which the extended methylation-free CpG island comprising an 8kb DNA fragment spanning the human hnRNP A2 gene is replaced with the extended methylation-free CpG island comprising an 8kb fragment spanning the murine hnRNP A2 gene (as shown in the sequence of Figure 19). Also
5 preferred are vectors having expressible nucleic acid sequences inserted into the multiple cloning site of CET 823, such as CET 824.

Also provided is host cell transfected with any of the embodiments of the disclosed vectors.

10

Alternatively said polynucleotide, vector or the host cell may be used in a cell culture system to obtain expression of a desired gene product. Suitable cell culture systems are well known in the art and are fully described in the body of literature known to those skilled in the art. There is provided a method for
15 the production of a polypeptide according to the invention comprising:

- i) providing a cell transformed/transfected with a nucleic acid molecule according to the invention;
- ii) growing said cell in conditions conducive to the manufacture of said polypeptide; and
- 20 iii) purifying said polypeptide from said cell, or its growth environment.

In a preferred embodiment of the invention said nucleic acid molecule is the vector according to the invention.

The present invention also provides the polynucleotide, vector or the host
25 cell for use in therapy.

The present invention also provides use of the polynucleotide, vector or host cell in the manufacture of a composition for use in gene therapy.

The present invention also provides a method of treatment, comprising administering to a patient in need of such treatment a pharmaceutically effective amount of the polynucleotide, vector or host cell of the present invention. Preferably the patient is suffering from a disease treatable by gene therapy.

The present invention also provides a pharmaceutical composition comprising the polynucleotide and/or the vector and/or host cell, optionally in admixture with a pharmaceutically acceptable carrier or diluent, for therapy to treat a disease or provide the cells of a particular tissue with an advantageous protein or function.

The polynucleotide, vector or host cell of the invention or the pharmaceutical composition may be administered via a route which includes systemic intramuscular, intravenous, aerosol, oral (solid or liquid form), topical, ocular, rectal, intraperitoneal and/or intrathecal and local direct injection.

The exact dosage regime will, of course, need to be determined by individual clinicians for individual patients and this, in turn, will be controlled by the exact nature of the protein expressed by the gene of interest and the type of tissue that is being targeted for treatment.

The dosage also will depend upon the disease indication and the route of administration. The number of doses will depend upon the disease, and the efficacy data from clinical trials.

The amount of polynucleotide or vector DNA delivered for effective gene therapy according to the invention will preferably be in the range of between 50 ng -1000 µg of vector DNA/kg body weight; and more preferably in the range of between about 1-100 µg vector DNA/kg.

Although it is preferred according to the invention to administer the polynucleotide, vector or host cell to a mammal for *in vivo* cell uptake, an *ex vivo* approach may be utilised whereby cells are removed from an animal, transduced with the polynucleotide or vector, and then re-implanted into the animal. The liver, for example, can be accessed by an *ex vivo* approach by removing hepatocytes from an animal, transducing the hepatocytes *in vitro* and re-implanting the transduced hepatocytes into the animal (e.g., as described for rabbits by Chowdhury *et al.*, *Science* 254:1802-1805, 1991, or in humans by Wilson, *Hum. Gene Ther.* 3:179-222, 1992). Such methods also may be effective for delivery to various populations of cells in the circulatory or lymphatic systems, such as erythrocytes, T cells, B cells and haematopoietic stem cells.

Another aspect of the invention provides an isolated polynucleotide comprising a first promoter operably linked to an expressible gene to which it is not naturally operably linked and a selectable element, also operably linked and 3' to the expressible gene, comprising a *pgk* promoter and a puromycin resistance gene. The use of such a polynucleotide to obtain reproducible expression of said expressible gene in at least two tissue or cell types is also provided.

In another embodiment of the invention there is provided a non-human transgenic animal comprising an artificially introduced extended methylation-free CpG island element and an artificially introduced selectable marker element wherein both elements are operably-linked to an expressible gene situated between them and wherein reproducible expression of said expressible gene occurs in at least two tissue or cell types. Methods of making transgenic mice (Gordon *et al.*, *Proc. Natl. Acad. Sci. USA* 77:7380 (1980); Harbers *et al.*, *Nature* 293:540 (1981); Wagner *et al.*, *Proc. Natl. Acad. Sci. USA* 78:5016 (1981); and Wagner *et al.*, *Proc. Natl. Acad. Sci. USA* 78:6376 (1981), sheep pigs, chickens (see Hammer *et al.*, *Nature*

315:680 (1985)), etc., are well-known in the art and are contemplated for use according to the invention.

Such transgenic animals containing the polynucleotide of the invention also
5 may be used for long-term production of a protein of interest.

There is also provided a mammalian model for determining the efficacy of gene therapy using the polynucleotide, vector or host cell of the invention. The mammalian model comprises a transgenic animal whose cells contain
10 the vector of the present invention. Such animals permit testing prior to clinical trials in humans.

The present invention also provides the use of the polynucleotide of the present invention in producing transgenic plants.
15

The generation of transgenic plants that have increased yield, or increased resistance to disease, pests, drought or salt are well known to those skilled in the art. The present invention also provides for transgenic plant containing cells that contain the polynucleotide of the present invention. Some or all of
20 the cells comprising the artificial UCOE may originate from plants.

The present invention also relates to the use of polynucleotide of the present invention in functional genomics applications. Functional genomics relates principally to the identification of genes specifically expressed in particular
25 cell types or disease states and now provides thousands of novel gene sequences of potential interest for drug discovery or gene therapy purposes. The major problem in using this information for the development of novel therapies lies in how to determine the functions of these genes. The polypeptides of the invention can be used in a number of functional genomic
30 applications in order to determine the function of gene sequences. The functional genomic applications of the present invention include, but are not limited to:

- 5 (1) Using the polynucleotide of the present invention to achieve sustained expression of anti-sense versions of the gene sequences or ribozyme knockdown libraries, thereby determining the effects of inactivating the gene on cell phenotype.
- 10 (2) Using the polynucleotide of the present invention to prepare expression libraries for the gene sequences, such that delivery into cells will result in reliable, reproducible, sustained expression of the gene sequences. The resulting cells, expressing the gene sequences can be used in a variety of approaches to function determination and drug discovery. For example, raising neutralising antibodies to the gene product; rapid purification of the protein product of the gene itself for use in structural, functional or drug screening studies; or in cell-based drug screening.
- 15 (3) Using the polynucleotide of the present invention in approaches involving mouse embryonic stem (ES) cells and transgenic mice. One of the most powerful functional genomics approaches involves random insertion into genes in mouse ES cells of constructs which only allow drug selection following insertion into expressed genes, and which can readily be rescued for sequencing (G.Hicks *et al.*, Nature Genetics, 16, 338-334). Transgenic mice with knockout mutations in genes with novel sequences can then readily be made to probe their function. At present this technology works well for the 10% of mouse genes which are well expressed in mouse ES cells. Incorporation of the polynucleotides of the present invention into the integrating constructs will enable this technique to be extended to identify all genes expressed in mice.

DETAILED DESCRIPTION OF THE INVENTION

The invention will now be described by way of example only and with reference to the accompanying figures wherein;

5

Figure 1 shows maps of 'empty' vectors CET 200.1, 210, 710 and 720.

Insertion of the Enhanced Green Fluorescent Protein (EGFP) gene into the multicloning site results in CET 230, 711 and 721, respectively. All vectors contain a CMV promoter from which inserted genes are expressed.

10 However, in the case of CET 210 (and its EGFP-expressing derivative, CET230) although such an inserted gene would be flanked by a UCOE and a *pgk* / puromycin resistance element in the plasmid, the latter is not immediately adjacent. More importantly, it is separated by a *Pvu* I site used to linearise the plasmid before transfection. After integration into the host
15 cell chromosome, this results in the gene no longer being flanked, since both the UCOE and the *pgk* / puromycin resistance element will integrate in the same side of the gene. In the case of CET 710 (and its EGFP-expressing derivative, CET 711) and CET 720 (and its EGFP-expressing derivative, CET 721), *Pvu* I linearisation results in the integration of the gene closely flanked
20 by the UCOE on one side and the *pgk* / puromycin resistance element on the other. CET 210 (and CET 230) and CET 720 (and CET 721) carry hnRNP-derived UCOEs, while CET 710 (and CET 711) carry an 'artificial' β -actin / PDCD2 –derived UCOE.

25 **Figure 2** shows expression of EGFP from various vectors transfected into CHO-K1 cells as measured by median fluorescence on FACS analysis measured on the indicated days post-transfection. 'EGFP' depicts cells transfected with a control (pEGFP) non-UCOE containing plasmid. CET220 shows cells transfected with a plasmid where the EGFP expression unit is
30 operably linked to a hnRNP-derived UCOE but not to a *pgk* / puromycin resistance element. Instead a SV40 /neomycin resistance element is used.

The remaining cells are transfected with CET 230, 711 or 721, the structures of which are shown in Figure 1.

Figure 3 shows the proportion of the populations of cells shown in Figure 2
5 judged to be positive for expression on the indicated days post-transfection.

Figure 4 shows the expression of EGFP in CHO-K1 cells transfected with
vectors CET 220, 230, 721 and 711 as measured by median fluorescence
corrected to allow comparison without exceeding the detection capacity of
10 the FACScan. This clearly shows the comparative effect of placing the
selectable marker (puro^r) either 5' (CET230) or 3' (CET721) to the
expressible transgene (EGFP).

Figure 5 shows the expression of EGFP in CHO-K1 cells transfected with
15 vectors CET 701, 721, 704, 741, 705, 751, 706, 761, 708 and 781 as
measured by median fluorescence corrected to allow comparison without
exceeding the detection capacity of the FACScan.

Figure 6 shows the expression levels of EGFP in CHO-K1 cells transfected
20 with vectors comparing 5' human and murine hnRNP UCOEs with a 3'
puromycin resistance gene.

Figure 7 shows the effect of position of the Streptomyces neomycin
resistance gene on EGFP expression. CET741 has the selectable marker 3'
25 of the transgene, CET 745 has the marker 5' of the transgene and UCOE.
The UCOE is the human RNP UCOE in both cases.

Figure 8 shows a map of plasmid CET 700

30 **Figure 9** shows a map of plasmid CET 710

Figure 10 shows the nucleotide sequence of CET710

35 **Figure 11** shows a map of plasmid CET 720

Figure 12 shows the nucleotide sequence of CET720

Figure 13 shows the nucleotide sequence of the wild-type *S. alboniger* puromycin N-acetyl transferase gene.

5

Figure 14 shows the nucleotide sequence of the modified *S. alboniger* puromycin N-acetyl transferase gene.

Figure 15 shows the nucleotide sequence of the *S. fradiae* aminoglycoside
10 phosphotransferase gene.

Figure 16 shows the nucleotide sequence of the *S. hygroscopicus* hygromycin phosphotransferase gene.

15 **Figure 17** shows the nucleotide sequence of the *E.coli* aminocyclitol phosphotransferase (hygro^r) gene.

Figure 18 shows the nucleotide sequence of the transposon Tn5 (*Klebsiella pneumoniae*) neomycin phosphotransferase gene.

20

Figure 19 shows the nucleotide sequence of the mouse hnRNP A2 *HindIII* fragment.

Figure 20 shows a map of plasmid CET 1010

25

Figure 21 shows the nucleotide sequence of CET1010

Figure 22 shows a map of plasmid CET 1020

30 **Figure 23** shows the nucleotide sequence of CET1020

Figure 24 shows a map of plasmid CET 1030

Figure 25 shows the nucleotide sequence of CET1030

Figure 26 shows a map of plasmid CET 1110

5 **Figure 27** shows the nucleotide sequence of CET1110

Figure 28 shows a map of plasmid CET 1120

10 **Figure 29** shows the nucleotide sequence of CET1120

Figure 30 shows a map of plasmid CET 1130

Figure 31 shows the nucleotide sequence of CET1130

15

EXAMPLES

Example 1 Flanking of an expressible gene with UCOEs and selectable elements

20

Materials and Methods

Construction of PGK-Puro CET expression Vectors

25 **CET700**

The CMV-MCS-SV40pA cassette was removed from CET31 (A CMV MCS pA SV40Neo based plasmid) as an *Asel/Afl* fragment, blunt end filled with T4 DNA polymerase and ligated into pPGK-Puro (mPGK promoter, Puromycin resistance gene, bGHpA in pBluescript) that had been digested
30 with *EcoRV*.

CET 720

CET20 (8.3kb hnRNPA2 fragment in pBluescript) was digested with *HindIII* to obtain the 8kb RNP UCOE and this was then ligated into CET700 that had also been cut with *HindIII*.

5

CET710

The Artificial UCOE was removed from CET 21 (Artificial UCOE in pBluescript) as an *XbaI/ClaI* fragment, blunt end filled with T4 DNA polymerase and ligated into CET700 that had been digested with *HindIII* and again blunt end filled with T4 DNA polymerase.

10

CET 230

This vector was constructed by digesting pUC19 with *NarI* and *EcoRI* to remove approximately 160bp, followed by blunting and religation. This removed one of the two *PvuI* and *PvuII* sites in the vector backbone. The CMV-EGFP-SV40pA cassette (with its MCS deleted) was excised from pEGFPN-1 (Clontech), as an *Asel/AflI* digest followed by blunt end filling, and then inserted into the pUC19 vector backbone that had been digested with *NdeI* and *Eco109 I* and again blunt end filled.

15

20

The PGK-Puro-bGpA cassette was then removed from pPGK-Puro as an *EcoRI/XhoI* blunt end filled fragment and inserted into the unique *PvuII* site of the above vector. Finally the 8.3kb hnRNPA2 fragment was inserted into the unique *HindIII* site of this vector as a *HindIII* fragment derived from CET20.

25

For clarity:

CET230 is the EGFP-expressing version of the 'empty' vector CET210

CET711 is the EGFP-expressing version of the 'empty' vector CET710

30

CET721 is the EGFP-expressing version of the 'empty' vector CET720

Vectors based on CET 720 with different antibiotic resistant genes and with alternative promoters or UCOEs can be constructed in the following manner. The PGK promoter (bp11384-11894) and the bghpA (bp 12567-12893) can be removed from CET 720 by restriction digestion. These elements can be
5 inserted into the pBluescript backbone such that restriction sites are available for the insertion of any resistance gene sequences (derived by PCR or restriction digestion) between the PGK promoter and the bghpA in such a manner as to allow expression of that gene. The CMV-MCS-SV40pA expression cassette can also be removed from CET 720 (bp 10533-11380)
10 and inserted 5' to the PGK promoter in the above vector; alternatively the mCMV-MCS-SV40pA expression cassette can be placed in the same position (CET 801, 821, 824-EGFP expression versions). The hnRNPA2 UCOE can be removed from CET 720 (bp 2240-10525) by restriction digestion and inserted 5' to the CMV expression cassette in the above
15 vectors, alternatively other UCOEs (e.g. murine hnRNPA2) can be inserted into the same position (CET 824-EGFP expression version).

For clarity:

CET741 is the EGFP-expressing version of the 'empty' vector CET740 and
20 comprises a 5' human RNP UCOE and a 3' *S. fradiae* neo^r gene.

CET761 is the EGFP-expressing version of the 'empty' vector CET760 and comprises a 5' human RNP UCOE and a 3' *E. coli* aminocyclitol phosphotransferase (hygro^r) gene.
25

CET781 is the EGFP-expressing version of the 'empty' vector CET780 and comprises a 5' human RNP UCOE and a 3' modified *S. alboniger* puromycin N-acetyl transferase gene.

30 CET821 is the EGFP-expressing version of the 'empty' vector CET820 and comprises a 5' human RNP UCOE and a 3' wild-type *S. alboniger* puromycin

N-acetyl transferase gene. Expression of the EGFP transgene is driven by the murine (rather than human) CMV IE promoter.

- 5 CET824 is the EGFP-expressing version of the 'empty' vector CET823 and comprises a 5' murine (rather than human) RNP UCOE and a 3' wild-type *S. alboniger* puromycin N-acetyl transferase gene.

pCIA vectors

- 10 This is a series of vectors that easily allow the construction of UCOE expression vectors with the final optimal configuration (UCOE-expression cassette-resistance cassette) when integrated into the chromosome.

- 15 CET 900 is an empty cloning vector in which pairs of rare restriction sites flank the MCS. CET 901 and CET 902 contain the hCMV and mCMV promoters respectively, an MCS and the SV40pA. The same pairs of rare restriction sites also flank these cassettes.

- 20 The CET 1000 series of vectors contain various combinations of UCOEs and resistance expression cassettes. They also contain the same rare restriction sites as the CET 900 series at a position 3' to the UCOE and 5' to the resistance cassette. The vectors also contain linearisation sites 5' to the UCOE and 3' to the resistance cassette.

- 25 Expression cassettes for any transgene can therefore be constructed in the CET 900 series and then easily be transferred into the CET 1000 series such that the ultimate configuration when integrated into the chromosome is the desired UCOE-expression cassette-resistance cassette.

- 30 As described above the antibiotic gene can be exchanged within the CET 1000 series by restriction digestion or PCR.

Transfection

CHO K1 cells were transfected and selected according to standard methods and as described in the co-pending applications incorporated by reference.

5 Results

With particular reference to Figure 2, comparison of cells transfected with CET 721 and CET 230 shows a consistently higher level of expression obtained with CET 721. These two vectors are similar in that both carry an 8kb hnRNP-derived UCOE operably-linked to the CMV promoter driven
10 EGFP gene and both carry the *pgk*/puromycin resistance gene element. However, following linearisation with *Pvu* I, integration of CET 230 into the host cell chromosome results in the elements being positioned in the order: *pgk*/Puro, hnRNP UCOE, EGFP gene. The same process with CET 721 results in the EGFP gene being flanked by the UCOE and the *pgk*/Puro. The
15 levels of expression obtained with CET 230 are not significantly higher than those obtained with CET 220, a vector carrying no *pgk*/Puro element but with the same UCOE and promoter driving EGFP expression. All UCOE carrying vectors show greatly increased expression compared with the basic EGFP expression plasmid.

20

Figure 3 shows that increased expression as expressed by median fluorescence is also reflected in an increased proportion of cells within the transfected population judged to be positive, in terms of expression, at all time points following transfection. This is a measure of the lack of position
25 effects, since random integration of the construct would normally result in a range of expression levels within the (non-clonal) population of transfected cells. This is overcome by the combination of 5' UCOE and 3' selectable element, resulting in a homogenous, highly-expressing population.

30 The levels of expression in some of the pools of cells in Figure 2 are so high that the fluorescence produced has exceeded the capacity of the detector.

In Figure 4, measurements have been corrected to the linear region of the detector's response to allow comparison between constructs. This shows that the combination of UCOE and 3' flanking selectable element used in CET 721 has produced an approximately 7-fold increase in levels of expression of EGFP as compared with that obtained with the UCOE alone (CET220) or that obtained with the selectable element (puro^r) placed 5' to the UCOE. It is clear that flanking the expressed transgene with the UCOE and selectable marker is required to obtain the boost in expression.

- 10 This effect is not restricted to a particular selectable marker. Figure 7 compares expression of EGFP operable linked to a 5' human RNP UCOE and either a 5' (CET745) or 3' (CET741) placed *S. fradiae* neomycin resistance gene. There is almost a doubling of the already high expression level.

15

Example 2 Effectiveness of other 3' flanking selectable markers

Results

- Figure 5 shows the effect of flanking the EGFP transgene with a 5' human RNP UCOE and various 3' flanking antibiotic resistance genes. CET701 is a control containing no UCOE, but with the wild-type *S. alboniger* puro^r. CET 721 has both the 5' UCOE and 3' puro^r. CET704 contains the *S. fradiae* neo^r but no UCOE, CET741 has both. CET705 contains the *S. hygrosopicus* hygro^r but no UCOE, CET751 has both. CET706 has the *E. coli* hygro^r but no UCOE, CET761 has both. CET708 has the codon-modified puro^r but no UCOE, CET781 has both. In all cases the boosting effect of the 3' flanking resistance gene is evident.

30

Example 3 Combination of other UCOEs and Puro selectable element

Results

As shown in Figures 2 and 3, expression from a comparable plasmid carrying an artificially constructed UCOE (CET 711) was comparable to that obtained with the RNP UCOE both in terms of median fluorescence and proportion of positive cells. This demonstrates that the phenomenon of amplification of the effect of a UCOE by a second flanking CpG-rich element is a general one, not confined to a particular combination of the RNP UCOE and the *pgk*/Puro element. The comparison of CET 711 and CET 721 expression in Figure 4 indicates a slightly lower level of expression was obtained with CET 711, but this was still at least 6-fold higher than that obtained with a UCOE alone.

Figure 6 shows the comparable effect obtained with either a human hnRNP UCOE using the murine CMV promoter to drive expression (CET821) and the murine equivalent (CET824). CET721 comprises the human hnRNP UCOE and uses the human CMV promoter.

CLAIMS

1. An isolated polynucleotide comprising:
 - a. an extended methylation-free CpG island
 - 5 b. an expressible nucleic acid terminated by a polyadenylation signal
 - c. a selectable marker operably linked to a promoterwherein both the CpG island and the selectable marker are operably-linked to the expressible nucleic acid, and the components are positioned in the order: extended methylation-free CpG island, expressible nucleic
10 acid, selectable marker, in the 5' to 3' orientation with respect to the sense strand of the expressible nucleic acid, and the polyadenylation signal at the 3' end of the expressible nucleic acid is within 2000bp of the proximal end of the selectable marker.
- 15 2. An isolated polynucleotide comprising:
 - a. an extended methylation-free CpG island
 - b. an expressible nucleic acid terminated by a polyadenylation signal
 - c. a selectable marker operably linked to a promoterwherein both the CpG island and the selectable marker are operably-
20 linked to the expressible nucleic acid, and the components are positioned in the order: extended methylation-free CpG island, expressible nucleic acid, selectable marker, in the 5' to 3' orientation with respect to the sense strand of the expressible nucleic acid, and the polyadenylation signal at the 3' end of the expressible nucleic acid is within 1500bp of the
25 proximal end of the selectable marker.
3. An isolated polynucleotide comprising:
 - a. an extended methylation-free CpG island
 - b. an expressible nucleic acid terminated by a polyadenylation signal
 - 30 c. a selectable marker operably linked to a promoterwherein both the CpG island and the selectable marker are operably-linked to the expressible nucleic acid, and the components are positioned

in the order: extended methylation-free CpG island, expressible nucleic acid, selectable marker, in the 5' to 3' orientation with respect to the sense strand of the expressible nucleic acid, and the polyadenylation signal at the 3' end of the expressible nucleic acid is within 1000bp of the proximal end of the selectable marker.

4. An isolated polynucleotide comprising:
 - a. an extended methylation-free CpG island
 - b. an expressible nucleic acid terminated by a polyadenylation signal
 - c. a selectable marker operably linked to a promoterwherein both the CpG island and the selectable marker are operably-linked to the expressible nucleic acid, and the components are positioned in the order: extended methylation-free CpG island, expressible nucleic acid, selectable marker, in the 5' to 3' orientation with respect to the sense strand of the expressible nucleic acid, and the polyadenylation signal at the 3' end of the expressible nucleic acid is within 500 bp of the proximal end of the selectable marker.
5. The isolated polynucleotide of any one of the above claims, wherein the selectable marker is an antibiotic resistance gene.
6. The isolated polynucleotide of claim 5, wherein the antibiotic resistance gene is obtained from a *Streptomyces* species.
7. The isolated polynucleotide of either of claim 6 wherein the antibiotic resistance gene is a puromycin resistance gene.
8. The isolated polynucleotide of claim 7 wherein the puromycin resistance gene is the puromycin N-acetyl transferase gene from *Streptomyces alboniger*.

9. The isolated polynucleotide of claim 8 wherein the puromycin resistance gene is a modified puromycin N-acetyl transferase gene from *Streptomyces alboniger*.
- 5 10. The isolated polynucleotide of claim 9 comprising the sequence of Fig 14.
11. The isolated polynucleotide of claims 6 wherein the antibiotic resistance gene is a neomycin resistance gene.
- 10 12. The isolated polynucleotide of claim 11 wherein the neomycin resistance gene is the aminoglycoside phosphotransferase gene from *Streptomyces fradiae*.
13. The isolated polynucleotide of claim 6 wherein the antibiotic resistance
15 gene is a hygromycin resistance gene.
14. The isolated polynucleotide of claim 13 wherein the antibiotic resistance gene is the hygromycin phosphotransferase gene from *Streptomyces hygroscopicus*.
- 20 15. The isolated polynucleotide of claim 6 wherein the antibiotic resistance gene is a bleomycin resistance gene.
16. The isolated polynucleotide of claim 15 wherein the bleomycin resistance
25 gene is the bleomycin binding protein from *Streptomyces verticillus*.
17. The isolated polynucleotide of claim 15 wherein the bleomycin resistance gene is the bleomycin N-acetyltransferase from *Streptomyces verticillus*.
- 30 18. The isolated polynucleotide of claim 6 wherein the antibiotic resistance gene is a blasticidin resistance gene.

19. The isolated polynucleotide of claim 18 wherein the blasticidin resistance gene is the blasticidin S-acetyltransferase gene from *Streptomyces verticillum*.
- 5 20. The isolated polynucleotide of claim 5, wherein the antibiotic resistance gene is the aminocyclitol phosphotransferase from *Escherichia coli*.
21. The isolated polynucleotide of claim 5, wherein the antibiotic resistance gene is the neomycin phosphotransferase gene from transposon Tn5.
- 10 22. The isolated polynucleotide of any one of the above claims wherein the extended methylation-free CpG island comprises an 8kb DNA fragment spanning the human hnRNP A2 gene.
- 15 23. The isolated polynucleotide of any one of claims 1–21 wherein the extended methylation-free CpG island comprises an 8kb DNA fragment spanning the murine hnRNP A2 gene.
- 20 24. The isolated polynucleotide of claim 23 wherein the extended methylation-free CpG island comprises nucleotides 1–7898 of the sequence of Figure 19.
- 25 25. The isolated polynucleotide of any one of claims 1–21, wherein the extended methylation-free CpG island comprises a 2.0 kb DNA fragment spanning the human β -actin CpG island/ promoter region and a 1.8 kb DNA fragment spanning the human PDCD2 CpG island/ promoter region.
26. A vector comprising the polynucleotide of any one of the previous claims.
- 30 27. A vector so constructed as to deliver, when linearised and integrated into a chromosome, the polynucleotide of any one of claims 1–25.

28. The vector of claim 26, wherein the vector is an episomal vector.
29. The vector of claim 26, wherein the vector is an integrating vector.
- 5 30. The vector of claim 26, wherein the vector is a plasmid.
31. The vector of any one of claims 26–30, wherein the expressible nucleic acid is a therapeutic nucleic acid.
- 10 32. The vector of any one of claims 26–30, wherein the expressible nucleic acid encodes a recombinant protein for expression in an *in vitro* cell culture system.
33. A vector comprising:
- 15 a. an extended methylation-free CpG island
b. a multiple cloning site
c. an antibiotic resistance gene obtained from a *Streptomyces* species
- wherein both the CpG island and the selectable marker are operably-
- 20 linked to the expressible nucleic acid, and the components are positioned in the order: extended methylation-free CpG island, multiple cloning site, selectable marker, in the 5' to 3' orientation with respect to the sense strand of the expressible nucleic acid, and the multiple cloning site is within 2000bp of the proximal end of the selectable marker.
- 25
34. The vector of claim 33 wherein the multiple cloning site is further operably linked to a promoter.
35. The vector of claim 34 wherein said promoter is a cytomegalovirus
- 30 immediate/early promoter.

36. The vector of any one of claims 26–31 comprising nucleotides 1–10551 of the sequence of Figure 10.
37. The vector CET 710.
5. 38. The vector of any one of claims 26–31 comprising nucleotides 1–13545 of the sequence of Figure 12.
39. The vector CET 720.
- 10 40. The vector CET 740.
41. The vector CET 760.
- 15 42. The vector CET 780.
43. The vector CET 820.
44. The vector CET 823.
- 20 45. The vector comprising nucleotides 1–12039 of the sequence of Figure 21.
46. The vector CET 1010
- 25 47. The vector comprising nucleotides 1–11646 of the sequence of Figure 23.
48. The vector CET 1020.
- 30 49. The vector comprising nucleotides 1–9027 of the sequence of Figure 25.

50. The vector CET 1030
51. The vector comprising nucleotides 1–12221 of the sequence of Figure 27.
- 5 52. The vector CET 1110
53. The vector comprising nucleotides 1–11828 of the sequence of Figure 29.
- 10 54. The vector CET 1120
55. The vector comprising nucleotides 1–9209 of the sequence of Figure 31.
56. The vector CET 1130
- 15 57. A host cell transfected with the vector of any of claims 26–56.
58. Use of the polynucleotide of any of claims 1–25, the vector of claims 26–56, or the host cell of claim 57 to obtain expression of an expressible
20 nucleic acid.
59. Use of the polynucleotide of any of claims 1–25, the vector of claims 26–56, or the host cell of claim 57 in cell culture system to obtain expression of a desired gene product.
- 25 60. Use of the polynucleotide of any of claims 1–25, the vector of claims 26–56, or the host cell of claim 57 for therapy
61. Use of the polynucleotide of any of claims 1–25, the vector of claims 26–56, or the host cell of claim 57 for the manufacture of a composition for
30 use in gene therapy.

62. A method of treatment, comprising the administering to a patient in need of such treatment a pharmaceutically effective amount of the polynucleotide of any of claims 1–25, the vector of claims 26–56, or the host cell of claim 57.

5

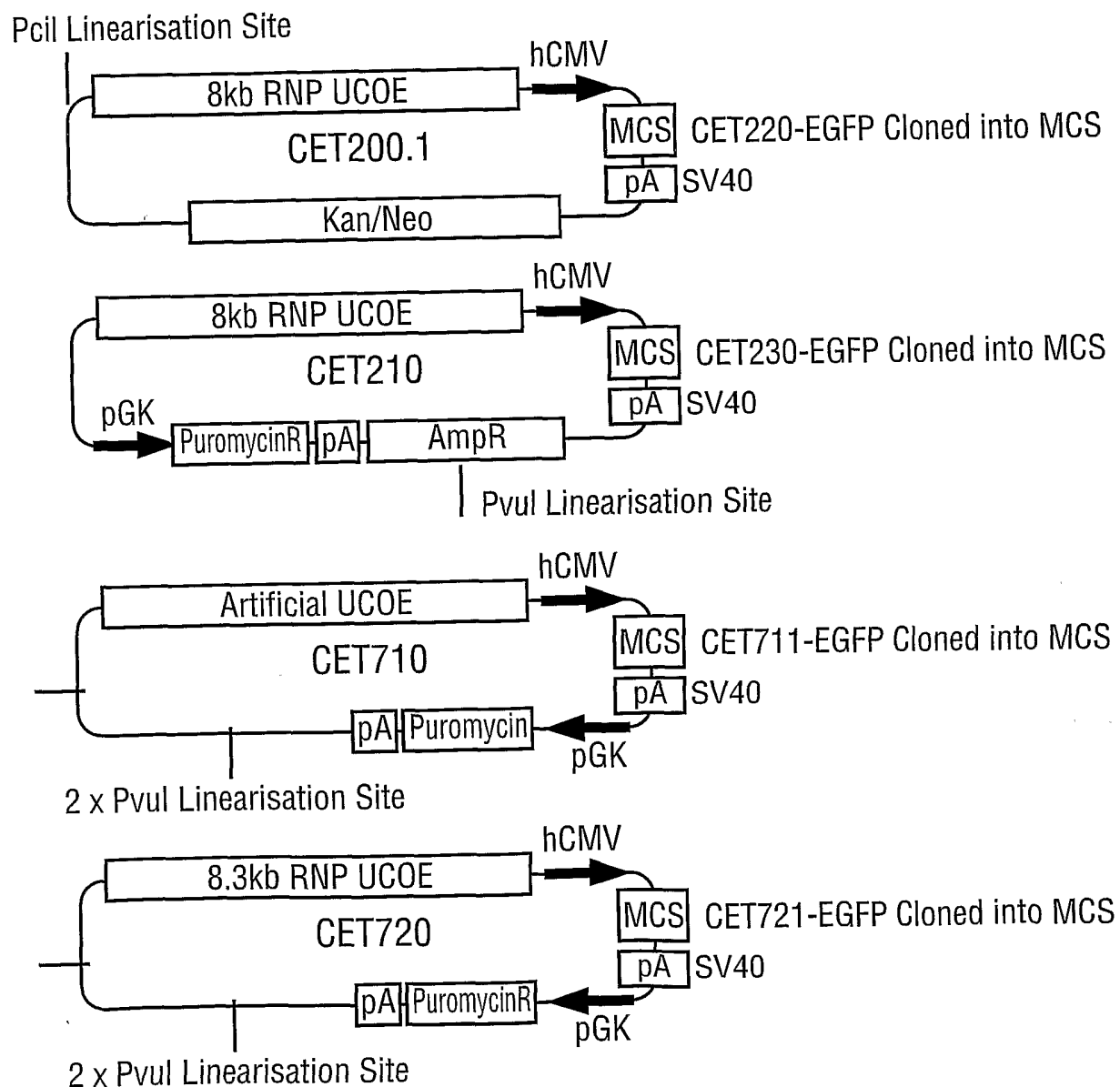
63. A pharmaceutical composition comprising the polynucleotide of any of claims 1–25, the vector of claims 26–56, or the host cell of claim 57 in combination with a pharmaceutically acceptable excipient.

10 64. A non-human transgenic animal comprising an artificially introduced extended methylation-free CpG island element and an artificially introduced selectable marker, wherein both the CpG island and the selectable marker are operably-linked to the expressible nucleic acid, and the components are positioned in the order: extended methylation-free
15 CpG island, expressible nucleic acid, selectable marker, in the 5' to 3' orientation with respect to the sense strand of the expressible nucleic acid, and the polyadenylation signal at the 3' end of the expressible nucleic acid is within 2000bp of the proximal end of the selectable marker.

20

1/93

FIG. 1



2/93

FIG. 2

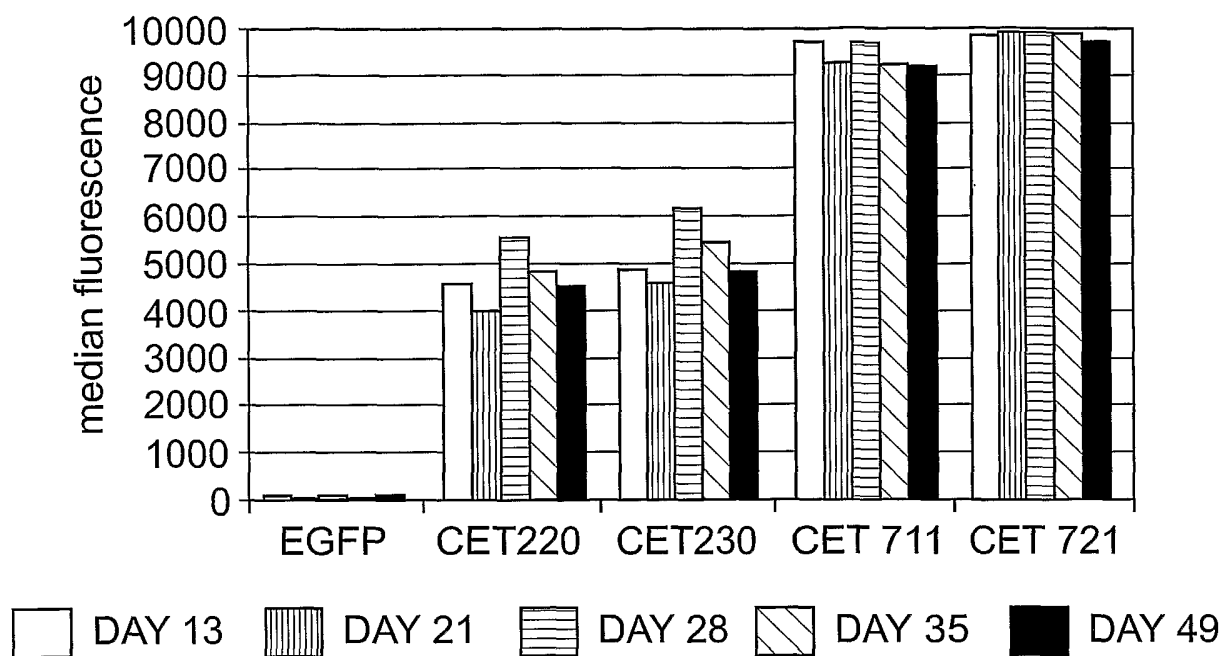


FIG. 4

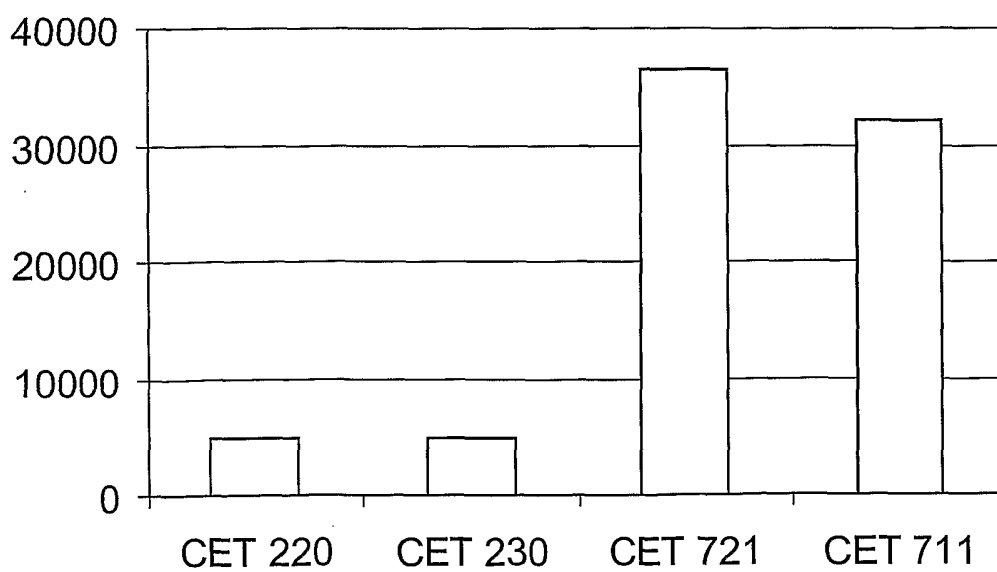


FIG. 3

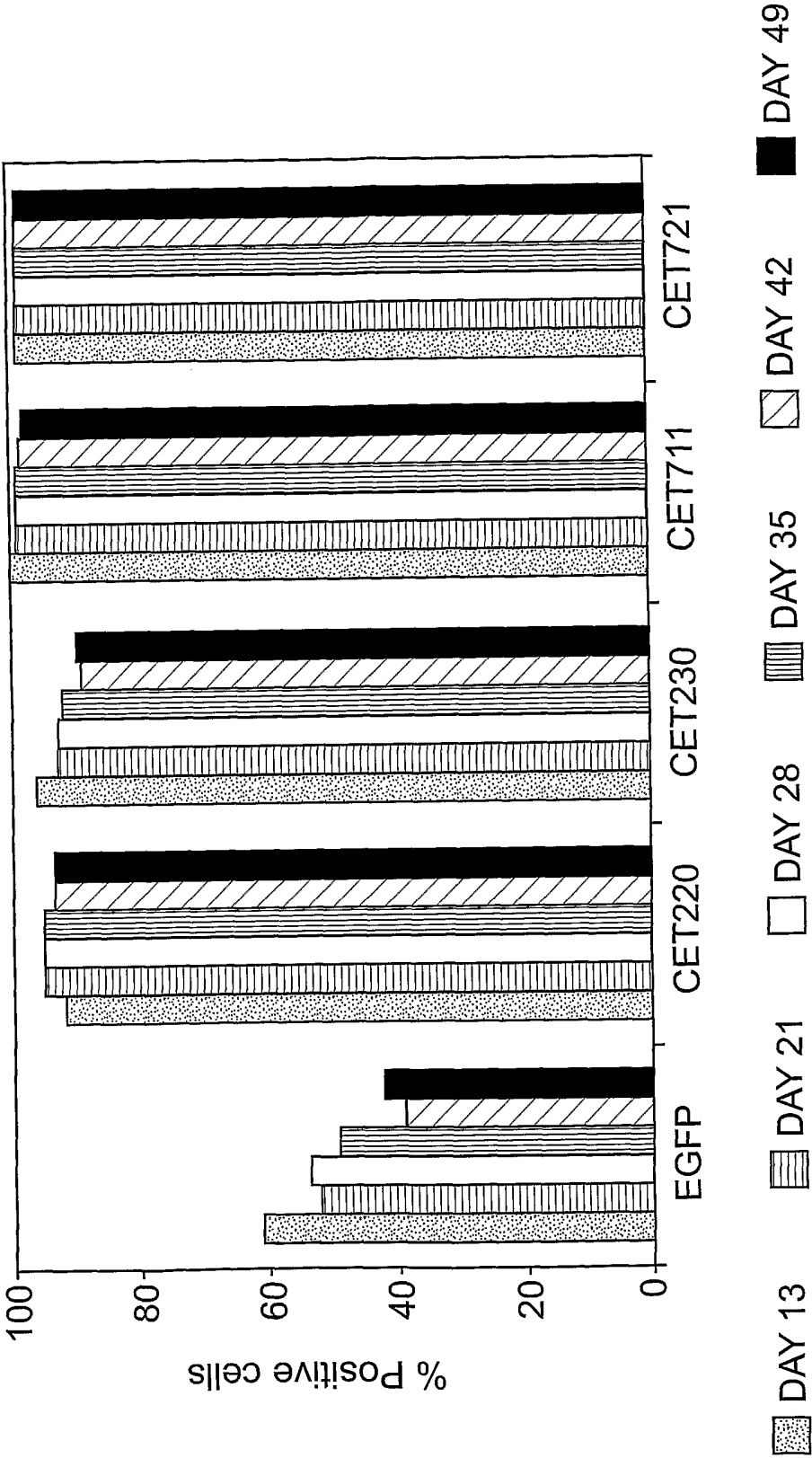
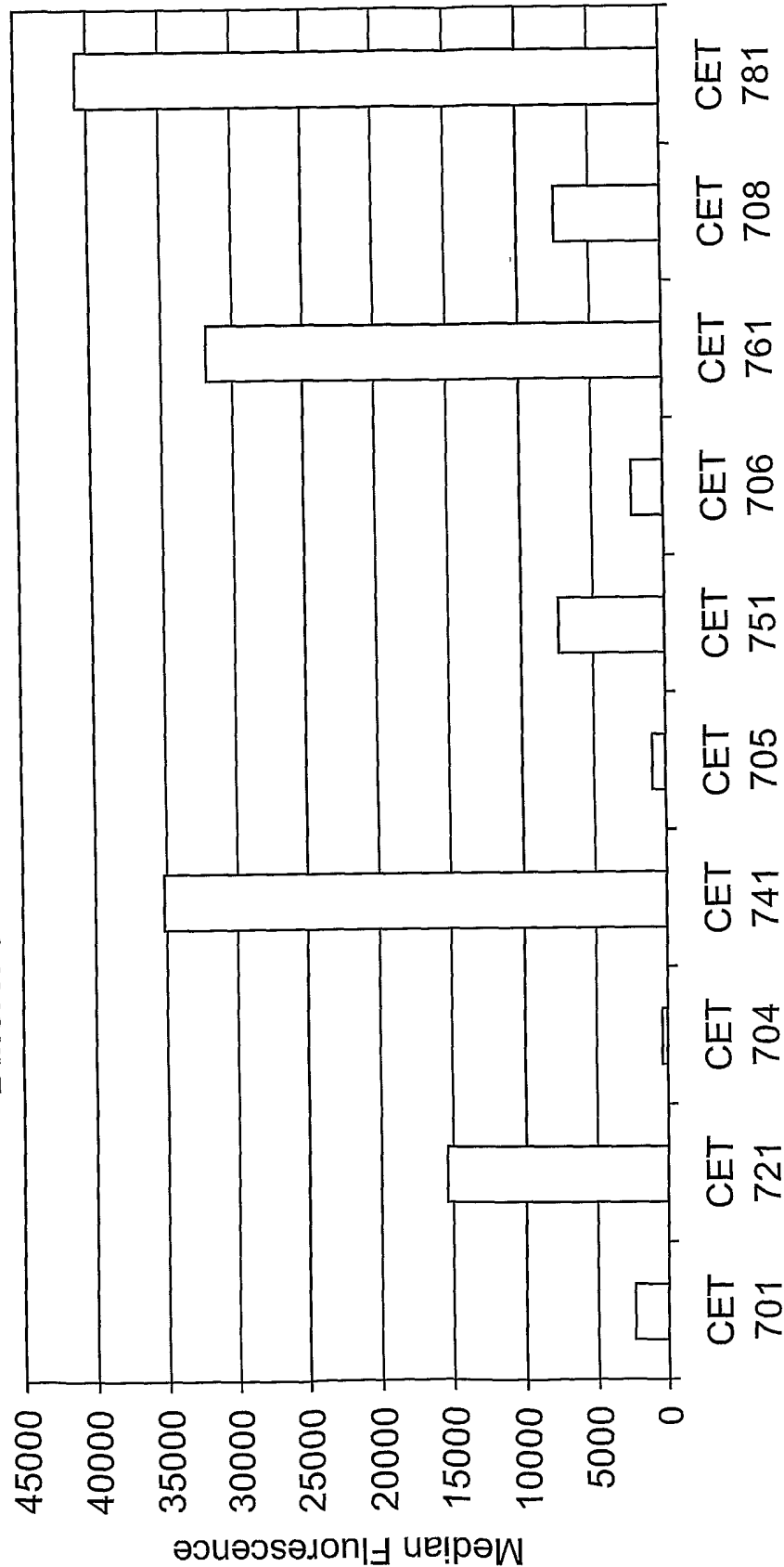


FIG. 5
Expression of EGFP in CHOK-1 Cells using Vectors with
Different Antibiotic Resistance Genes



5/93

FIG. 6
Expression Levels of EGFP in CHOK-1 cells transfected
with vectors encoding the Puromycin Gene

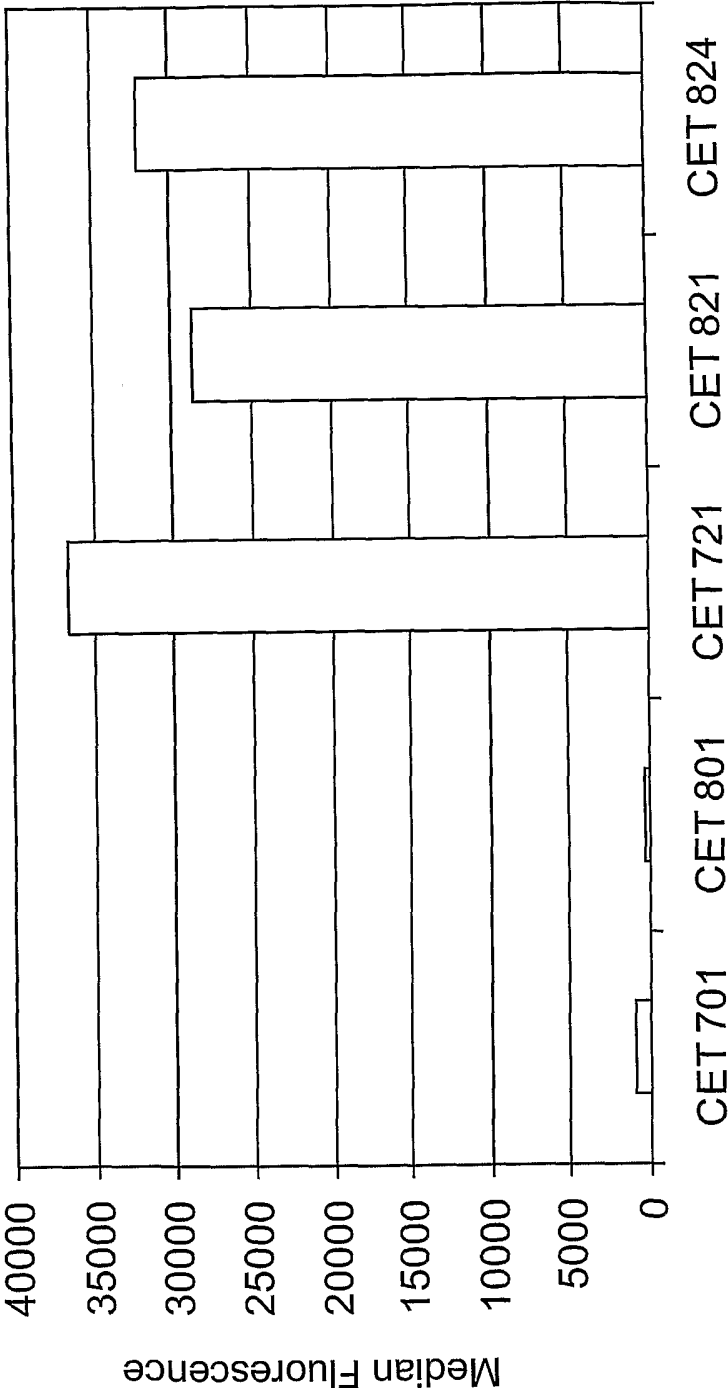
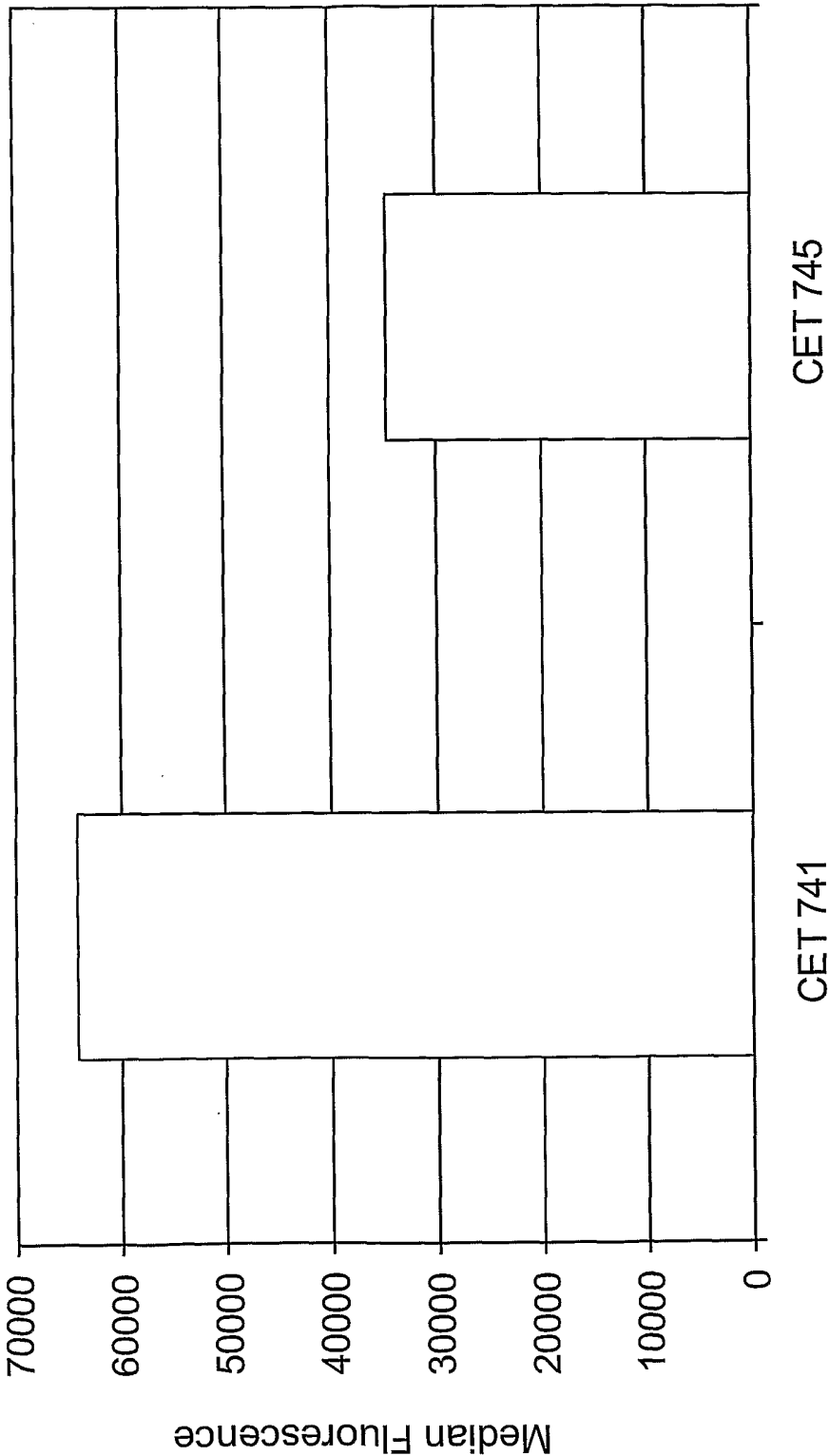


FIG. 7
Effect of Position of the Streptomyces Neomycin
Resistance Gene on EGFP expression



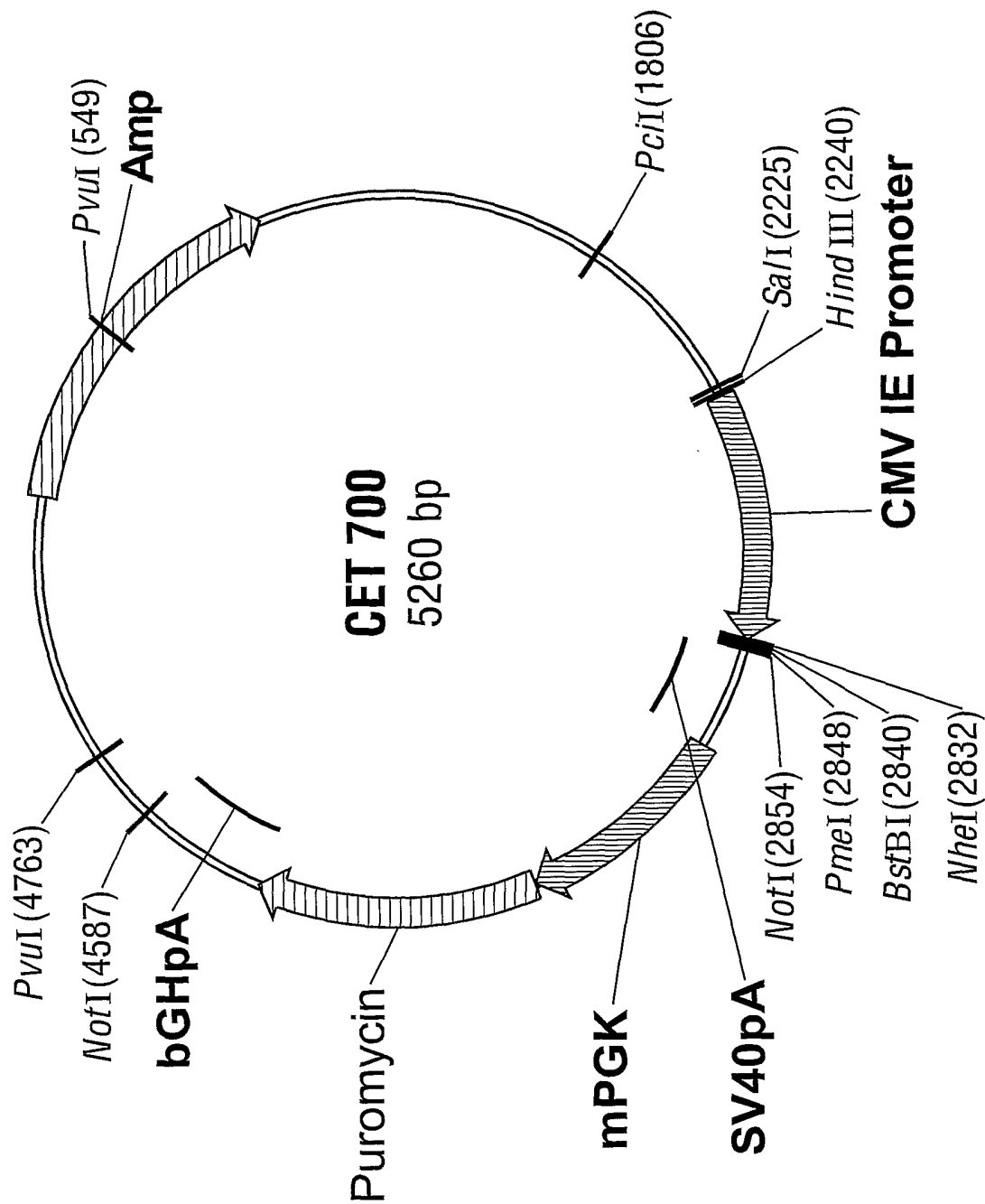
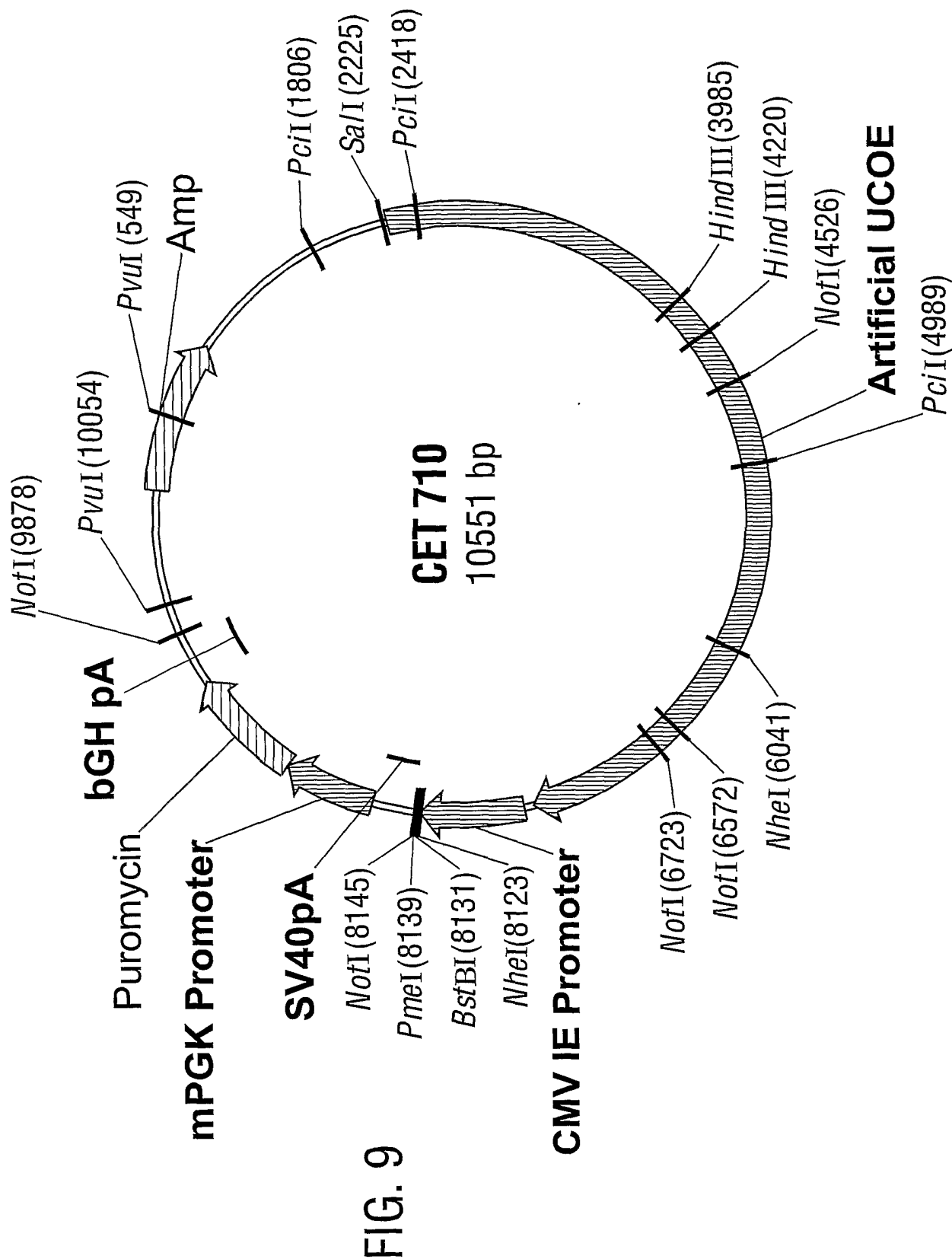


FIG. 8



9/93

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 TGTAAGGCA
151 GTCGCCCTTA TTCCCTTTTT TGCGGCATTT TGCCTTCCTG TTTTGTCTCA
   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 GGAAGTCTCA
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 TGTGCTAGC
551 GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG GGATCATGTA
   CTCCTGGCTT CCTCGATTGG CGAAAAAACG TGTGTACCC CCTAGTACAT
601 ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA
   TGAGCGGAAC TAGCAACCCT TGGCCTCGAC TTACTTCGGT ATGGTTTGCT
651 CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAAC
   GCTCGCACTG TGGTGCTACG GACATCGTTA CCGTTGTTGC AACGCGTTTG
701 TATTAACCTG 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 GCTGCCCTC AGTCCGTTGA TACCTACTTG CTTTATCTGT
951 GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC
   CTAGCGACTC TATCCACGGA GTGACTAATT CGTAACCATT GACAGTCTGG
1001 AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA TTTTAAATTT
   TTCAAATGAG TATATATGAA ATCTAACTAA ATTTTGAAGT AAAAATTAAA
1051 AAAAGGATCT AGGTGAAGAT CCTTTTTGAT AATCTCATGA CCAAATCCC
   TTTTCCTAGA TCCACTTCTA GGAAAACTA TTAGAGTACT GGTTTTAGGG
1101 TTAACGTGAG TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA
   AATTGCACTC AAAAGCAAGG TGACTCGCAG TCTGGGGCAT CTTTCTAGT
1151 AAGGATCTTC TTGAGATCCT TTTTCTGTC GCGTAATCTG CTGCTTGCAA
   TTCTTAGAAG AACTCTAGGA AAAAAGACG CGCATTAGAC GACGAACGTT
1201 ACAAAAAAAC CACCGCTACC AGCGGTGGTT TGTGTCGGG ATCAAGAGCT
   TGTTTTTTTT GTGGCGATGG TCGCCACCAA ACAAACGGCC TAGTCTCGA
1251 ACCAACTCTT TTTCCGAAGG TAACTGGCTT CAGCAGAGCG CAGATACCAA
   TGGTTGAGAA AAAGGCTTCC ATTGACCGAA GTCGTCTCGC GTCTATGGTT
1301 ATACTGTCCT TCTAGTGTAG CCGTAGTTAG GCCACCACTT CAAGAACTCT
   TATGACAGGA AGATCACATC GGCATCAATC CGGTGGTGAA GTTCTTGAGA

```

10/93
FIG. 10(contd.)

1351	GTAGCACCGC	CTACATACCT	CGCTCTGCTA	ATCCTGTTAC	CAGTGGCTGC
	CATCGTGGCG	GATGTATGGA	GCGAGACGAT	TAGGACAATG	GTCACCGACG
1401	TGCCAGTGGC	GATAAGTCGT	GTCTTACCGG	GTTGGACTCA	AGACGATAGT
	ACGGTCACCG	CTATTTCAGCA	CAGAATGGCC	CAACCTGAGT	TCTGCTATCA
1451	TACCGGATAA	GGCGCAGCGG	TCGGGCTGAA	CGGGGGGTTT	GTGCACACAG
	ATGGCCTATT	CCGCGTCGCC	AGCCCCGACTT	GCCCCCAAG	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
	GCCATTTCGCC	GTCCCAGCCT	TGTCTCTCTG	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	TGGCCTTTT
	TGCGGTCGTT	GCGCCGGAAA	AATGCCAAGG	ACCGGAAAAA	GACCGGAAAA
1801	GCTCACATGT	TCTTTCCTGC	GTTATCCCCT	GATTCTGTGG	ATAACCGTAT
	CGAGTGTACA	AGAAAGGACG	CAATAGGGGA	CTAAGACACC	TATTGGCATA
1851	TACCGCCTTT	GAGTGAGCTG	ATACGCTCG	CCGCAGCCGA	ACGACCGAGC
	ATGGCGGAAA	CTCACTCGAC	TATGCGGAGC	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	TTTCGCCCCG	CACTCGCGTT	GCGTTAATTA	CACTCAATCG
2051	TCACTCATT	GGCACCCAG	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	CGGTTTCGCG	GTAAATTGGG	AGTGATTTCC	CTTGTTTTTCG
2201	TGGGTACCGG	GCCCCCCTC	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	CTTCCCATT	TATCCCTGTT	ATCAAGCACA	AGGTGAGGCA
	ACCTCACAAAC	GAAGGGTAAT	ATAGGGACAA	TAGTTCGTGT	TCCAGTCCGT
2401	CAGAGTAAGA	CTCAAAACAT	GTTTTGGAAT	GTATGACTGG	TATGAACCTAC
	GTCTCATTCT	GAGTTTTGTA	CAAAACCTTA	CATACTGACC	ATACTTGATG
2451	AAACCAGTAA	GCTGATGTTT	TCATTTTGAG	TCTATAAATC	TAATTTTGTG
	TTTGGTTCATT	CGACTACAAA	AGTAAACTC	AGATATTTAG	ATTAAACAC
2501	GTGGTTTTGT	GTATGGCTCA	AGGCTCAAAT	TGTAAAATTT	AATATTATGT
	CACCAAAACA	CATACCGAGT	TCCGAGTTTA	ACATTTTAAA	TTATAATACA
2551	GACCAAAGAA	AGTTATACCC	AGAACCTCAA	TTTCCTCACC	TTCAAAATGG
	CTGGTTTCTT	TCAATATGGG	TCTTGAGGTT	AAAGGAGTGG	AAGTTTTACC
2601	GGCAGTTTCT	CACTCATTGG	TCTGCTGTCA	CGATTTTAAAT	GAGCTCATGC
	CCGTCAAAGA	GTGAGTAACC	AGACGACAGT	GCTAAAATTA	CTCGAGTACG
2651	ACAAACAGCC	CTTTATATAA	GGTAAGTGCT	GGATAAATGT	TGGCTACTAT
	TGTTTGTTCG	GAAATATATT	CCATTACGGA	CCTATTTACA	ACCGATGATA
2701	AATAAAATAA	GCCTCTAAGA	TACTTGGTCA	GCACAAGTAC	TACCCAAGAG
	TTATTTTATT	CGGAGATTCT	ATGAACCAGT	CGTGTTTCATG	ATGGGTTCTC

11/93
FIG. 10(contd.)

2751	TATGCACTGT	AAGTAACTG	ACAAAATTGT	GTATCTAAAA	CTGGCCAGAT
	ATACGTGACA	TTCATTTGAC	TGTTTTAACA	CATAGATTTT	GACCGGTCTA
2801	GAAAGAGAAA	CTTTTAAGGG	GCCCTTCTGC	GTGCCCAGACA	CTGTGCTAGG
	CTTTCTCTTT	GAAAATTCCC	CGGGAAGACG	CACGGGCTGT	GACACGATCC
2851	CACTCACACT	ATCCCGACCC	GAGAAACCGA	TCTGCGACCC	AGAGGAACTT
	GTGAGTGTGA	TAGGGCTGGG	CTCTTTGGCT	AGACGCTGGG	TCTCCTTGAA
2901	ACCAAGCCTC	CAGCATCTTG	TGCAGCCCTA	CTCATGGGAC	CATCTGGATA
	TGGTTTCGGAG	GTCGTAGAAC	ACGTCGGGAT	GAGTACCCTG	GTAGACCTAT
2951	CCCACCCCTG	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	TTCTTCAAC	ACCAGACCTT	AATATACCAG	GTCTACTAGA
3301	GAAACAAAAA	GGACAGCACT	ATTAGTAATC	ATTTAGTTTT	GAAGACAGTC
	CTTTGTTTTT	CCTGTCGTGA	TAATCATTAG	TAAATCAAAA	CTTCTGTCAG
3351	TAATAATTTG	CTGTCTCTAA	AGTACTATAT	TCCCTATAGT	TCTGGCATT
	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	ACCCAACGTG	TTCCGCTACT	TTGACTTGAC	TCTGAATTTT
3601	AAAAGATTAC	CCACCTGGTT	GTGCACAAGC	CTGCTTATGT	CCCAATCTCC
	TTTTCTAATG	GGTGGACCAA	CACGTGTTCT	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	TTTTCTGGG	TAGTTGATT
	GTCTTCCACC	GAGTATACTC	ATTTTTAGCA	AAAAGGACCC	ATCAACTAAG
3851	CTAAAAACTA	AAAAAGAATA	CAGAGAAAAG	TTTTATCTTC	AAACAAAACA
	GATTTTTTGAT	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
4051	TTGAGTTAAT	CAAATTAATA	AGACTGATCG	TTAAGAACGA	CTGCCAAAAA
	AACTCAATTA	GTTTAATTAT	TCTGACTAGC	AATTCCTGCT	GACGGTTTTT
4101	TACGAAAAAG	CTACTGGGAT	CCATCTTTCC	AAGACAATTT	CTATTATCTG
	ATGCTTTTTT	GATGACCCTA	GGTAGAAAGG	TTCTGTAA	GATAATAGAC

12/93
FIG. 10(contd.)

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	GCGCCTCTC	CCTCGAGACT
4301	GGCTGGGGCG	GCAGCGGTGC	TTGCCGCCGT	CCCCCTGGTC	GCTCCCGGAA
	CCGACCCCGC	CGTCGCCACG	AACGGCGGCA	GGGGGACCAG	CGAGGGCCTT
4351	TTAACGCCGC	GCACGCGTCG	GAGGCATGGC	CCCGTCCCGA	CCCGTTTGG
	AATTGCGGCG	CGTGCGCAGC	CTCCGTACCG	GGGCAGGGCT	GGGGCAAACC
4401	CGGCTCACCT	CGCAGGCCCG	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
	TGGACGTCTG	CCTTCCTCTC	GCCCCGCCGG	GTGTCGAGCG	TCCGGTCCCG
4551	CTGGGGCCCC	GGCAGCCCCG	CCGCGCCCGA	CCATGCCGGC	CGCCCCGCCA
	GACCCCGGGG	CCGTCCGGCC	GGCGCGGGTC	GGTACGGCCG	GCGGGCGGGT
4601	CCTTGCTGGG	GAAGTGTCTG	CTGCGCAGTC	GCCACGCCGG	CGCCGACTCG
	GGAACGACCC	CTTGACGAGC	GACGCGTCAG	CGGTGCGGCC	GCGGCTGAGC
4651	GCGAAGCCCA	GCTCCACAGG	CCTGGCCCCG	GCGGCAGCCA	TGCGGGGCGC
	CGCTTCGGGT	CGAGGTGTCC	GGACCGGGGC	CGCCGTCGGT	ACGCCCCGCG
4701	GGGCTGGCGT	GGGGCGCAGC	CCACAGCTGG	GTGGAAGGC	GGAAATCGGG
	CCCGACCGCA	CCCCGCGTCG	GGTGTGACCC	CAGCCTTCCG	CCTTTAGCCC
4751	CGCCGGGCGG	GAAGGCAAGA	GGCGGGCACC	TTTCCGGAGG	ACAGGAGGCG
	GCGGCCCGGC	CTTCCGTTCT	CCGCCCCTGG	AAAGGCCTCC	TGTCCTCCGC
4801	GAAACGCGTC	TGACGGGAGC	GGTTGCAGGA	CCAATGCGAG	GGAACGGGGC
	CTTTGCGCAG	ACTGCCCTCG	CCAACGTCTT	GGTTACGCTC	CCTTGCCCCG
4851	AGAGGAAACC	TCTCGGCATC	AGCCCCGCC	CTGGCGCCTC	TGCCTCCGAG
	TCTCCTTTGG	AGAGCCGTAG	TCGGGGCGGG	GACCGCGGAG	ACGGAGGCTC
4901	CCGCTTTTCT	GGTGCTCCCG	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	TTCCAATTG	TCCATTTTAC
5151	AGGACCTTAC	AGCGGAAACC	AAGAATCCTG	TGTGCCTGAG	AGTAATGAAG
	TCCTGGGATG	TCGCCTTTGG	TTCTTAGGAC	ACACGGACTC	TCATTACTTC
5201	AAGCCTCTGC	AGAAGAGTCT	TTTCTGTCAG	TCTTAAGGTC	TCTGTTTTAA
	TTCGGAGACG	TCTTCTCAGA	AAAGACAGTC	AGAATTCAG	AGACAAAATT
5251	TGTTAGTGCT	GGCTTGCTGT	ACCTGAATTC	CAAGGGAGGA	GTGTATAATG
	ACAATCACGA	CCGAACGACA	TGGACTTAAG	GTTCCCTCCT	CACATATTAC
5301	AGGCATGGCC	AACCCCCACT	TCCCATCATT	GCCTGAACCT	GTTTTTCAGG
	TCCGTACCGG	TTGGGGGTGA	AGGGTAGTAA	CGGACTTGAT	CAAAAAGTCC
5351	TTAACTTCAG	AATGCCCTTG	GTACCGCGGG	CCCCCTCTGT	GGTCCCACGC
	AATTGAAGTC	TTACGGGAAC	CATGGCGCCC	GGGGGAGACA	CCAGGGTGCG
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
	CTGGGTTCGG	TTCTTTCCCA	CTGTTCTCTG	CCCAGAAGGG	TCCGACCGAA

13/93
FIG. 10(contd.)

5551	TGAGTTCCTA	GCACCGCCCC	GCCCCCAATC	CTCTGTGGCA	CATGGAGTCT
	ACTCAAGGAT	CGTGGCGGGG	CGGGGGTTAG	GAGACACCGT	GTACCTCAGA
5601	TGGTCCCCAG	AGTCCCCCAG	CGGCCTCCAG	ATGGTCTGGG	AGGGCAGTTC
	ACCAGGGGGT	TCAGGGGGT	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	AGTTTTCTGC
5801	GCAGCTCCAG	GGTAAAGGT	GCCCTGCCCT	GTAGAGCCCA	CTTCCTTCCC
	CGTCGAGGTC	CCATTTTCCA	CGGGACGGGA	CATCTCGGGT	GAAGGAAGGG
5851	AGGGCTGCGG	CTGGGTAGGT	TTGTAGCCTT	CATCACGGGC	CACCTCCAGC
	TCCCGACGCC	GACCCATCCA	AACATCGGAA	GTAGTGCCCC	GTGGAGGTCT
5901	CACTGGACCG	CTGGCCCCCT	CCCTGTCTCT	GGGAGTGTGG	TCCTGCGACT
	GTGACCTGGC	GACCGGGGAC	GGGACAGGAC	CCCTCACACC	AGGACGCTGA
5951	CTAATGGCCG	CAAGCCACCT	GACTCCCCCA	ACACCACACT	CTACCTCTCA
	GATTACCGGC	GTTTCGGTGA	CTGAGGGGGT	TGTGGTGTGA	GATGGAGAGT
6001	AGCCAGGTC	TCTCCCTAGT	GACCCACCCA	GCACATTTAG	CTAGCTGAGC
	TCGGGTCCAG	AGAGGGATCA	CTGGGTGGGT	CGTGTAATC	GATCGACTCG
6051	CCCACAGCCA	GAGGTCTCTA	GGCCCTGCTT	TCAGGGCAGT	TGCTCTGAAG
	GGGTGTCTGGT	CTCCAGGAGT	CCGGGACGAA	AGTCCCGTCA	ACGAGACTTC
6101	TCGGCAAGGG	GGAGTGACTG	CCTGGCCACT	CCATGCCCTC	CAAGAGCTCC
	AGCCGTTCCC	CCTCACTGAC	GGACCGGTGA	GGTACGGGAG	GTTCTCGAGG
6151	TTCTGCAGGA	GCGTACAGAA	CCCAGGGCCC	TGGCACCCGT	GCAGACCCCTG
	AAGACGTCTT	CGCATGTCTT	GGGTCCCGGG	ACCGTGGGCA	CGTCTGGGAC
6201	GCCCACCCCA	CCTGGGCGCT	CAGTGCCCAA	GAGATGTCCA	CACCTAGGAT
	CGGGTGGGGT	GGACCCGCGA	GTACAGGGTT	CTCTACAGGT	GTGGATCCTA
6251	GTCCCGCGGT	GGGTGGGGGG	CCCGAGAGAC	GGGCAGGCCG	GGGGCAGGCC
	CAGGGCGCCA	CCCACCCCCC	GGGCTCTCTG	CCCGTCCGGC	CCCCGTCCGG
6301	TGGCCATGCG	GGGCCGAACC	GGGCACTGCC	CAGCGTGGGG	CGCGGGGGCC
	ACCGGTACGC	CCCGGCTTGG	CCCGTGACGG	GTGCGACCCC	GCGCCCCCGG
6351	ACGGCGCGCG	CCCCCAGCCC	CCGGGCCAG	CACCCCAAGG	CGGCCAACGC
	TGCGCGCGCG	GGGGGTCTGG	GGCCCGGGT	GTGGGGTTCC	GCCGGTTGCG
6401	CAAAACTCTC	CCTCCTCCTC	TTCCTCAATC	TCGCTCTCGC	TCTTTTTTTTT
	GTTTTGAGAG	GGAGGAGGAG	AAGGAGTTAG	AGCGAGAGCG	AGAAAAAAAAA
6451	TTTCGCAAAA	GGAGGGGAGA	GGGGGTAAAA	AAATGCTGCA	CTGTGCGGCG
	AAAGCGTTTT	CCTCCCCTCT	CCCCCATTTT	TTTACGACGT	GACACGCCGC
6501	AAGCCGGTGA	GTGAGCGGCG	CGGGGCCAAT	CAGCGTGCGC	CGTTCGGA
	TTCGGCCACT	CACTCGCCGC	GCCCCGGTTA	GTGCGACGCG	GCAAGGCTTT
6551	GTTGCCTTTT	ATGGCTCGAG	CGGCCGCGGC	GGCGCCCTAT	AAAACCCAGC
	CAACGGAAAA	TACCGAGCTC	GCCGGCGCCG	CCGCGGGATA	TTTTGGGTCTG
6601	GGCGCGACGC	GCCACCACCG	CCGAGACCGC	GTCCGCCCGC	GAGCACAGAG
	CCGCGCTGCG	CGGTGGTGGC	GGCTCTGGCG	CAGGCGGGCG	CTCGTGTCTC
6651	CCTCGCCTTT	GCCGATCCGC	CGCCCGTCCA	CACCCGCCGC	CAGGTAAGCC
	GGAGCGGAAA	CGGCTAGGCG	GCGGGCAGGT	GTGGGCGGGC	GTCCATTCTG
6701	CGGCCAGCCG	ACCGGGGCAT	GCGGCCGCGG	CCCTTCGCCC	GTGCAGAGCC
	GCCGGTCTGG	TGGCCCCGTA	CGCCGGCGCC	GGGAAGCGGG	CACGTCTCGG
6751	GCCGTCTGGG	CCGCAGCGGG	GGGCGCATGG	GGCGGAACCG	GACCGCCGTG
	CGGCAGACCC	GGCGTCGCCC	CCCGCGTACC	CCGCCTTGGC	CTGGCGGCAC
6801	GGGGGCGCGG	GAGAAGCCCC	TGGGCCCTCC	GAGATGGGGG	ACACCCACAG
	CCCCCGCGCC	CTCTTCGGGG	ACCCGAGGGC	CTCTACCCCC	TGTGGGGTGC
6851	CCAGTTCGCA	GGCGCGAGGC	CGCGCTCGGG	CGGGCGCGCT	CCGGGGGTGC
	GGTCAAGCGT	CCGCGCTCCG	GCGCGAGCCC	GCCCGCGCGA	GGCCCCACAG
6901	CGCTCTCGGG	GCGGGGGCAA	CCGGCGGGGT	CTTTGTCTGA	GCCGGGCTCT
	GCGAGAGCCC	CGCCCCCGTT	GGCCGCCCCA	GAAACAGACT	CGGCCCGAGA

14/93
FIG. 10(contd.)

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	TGCCCCGAGCC	GCTGGCGCCC	GAGGGTGTGG	CCGCTGCGTG	CGCGCGCGCG
	ACGGGCTCGG	CGACCGCGGG	CTCCACACAC	GGCGACGCAC	GCGCGCGCGC
7201	ACCCGGTCGC	TGTTTGAACC	GGGCGGAGGC	GGGGCTGGCG	CCCGGTTGGG
	TGGGCCAGCG	ACAAACTTGG	CCCGCTCCG	CCCCGACCGC	GGGCCAACCC
7251	AGGGGGTTGG	GGCCTGGCTT	CCTGCCGCGC	GCCGCGGGGA	CGCCTCCGAC
	TCCCCCAACC	CCGGACCGAA	GGACGGCGCG	CGGCGCCCCT	GCGGAGGCTG
7301	CAGTGTTTGC	CTTTTATGGT	AATAACGCGG	CCGGCCCCGG	TTCCTTTGTC
	GTCACAAACG	GAAAATACCA	TTATTGCGCC	GGCCGGGCCG	AAGGAAACAG
7351	CCCAATCTGG	GCGCGCGCCG	GCGCCCCCTG	GCGGCCTAAG	GACTCGGCGC
	GGGTTAGACC	CGCGCGCGGC	CGCGGGGGAC	CGCCGGATTG	CTGAGCCGCG
7401	GCCGGAAGTG	GCCAGGGCGG	GGGCGACTTC	GGCTCACAGC	GCGCCCCGCT
	CGGCCTTCAC	CGGTCCCGCC	CCCGCTGAAG	CCGAGTGTCG	CGCGGGCCGA
7451	ATTCTCGCAG	CTCACCATGC	CGGTGCGCCAC	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
	GTAAGTGCAG	TTACCCACCT	CATAAATGCC	ATTTGACGGG	TGAACCGTCA
7751	ACATCAAGTG	TATCATATGC	CAAGTACGCC	CCCTATTGAC	GTCAATGACG
	TGTAGTTCAC	ATAGTATACG	GTTTCATGCG	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	GCGGTTTTTG	CAGTACATCA	ATGGGCGTGG	ATAGCGGTTT	GACTCACGGG
	CGCCAAAACC	GTCATGTAGT	TACCCGCACC	TATCGCCAAA	CTGAGTGCCC
7951	GATTTCCAAG	TCTCCACCCC	ATTGACGTCA	ATGGGAGTTT	GTTTTGGCAC
	CTAAAGGTTT	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
8101	GTTTAGTGAA	CCGTCAGATC	CGTAGCGTT	CGAAGTTTAA	ACGCGGCCGC
	CAAATCACTT	GGCAGTCTAG	GCGATCGCAA	GCTTCAAATT	TGCGCCGGCG
8151	GACTCTAGAT	CATAATCAGC	CATACCACAT	TTGTAGAGGT	TTACTTGCT
	CTGAGATCTA	GTATTAGTCG	GTATGGTGTA	AACATCTCCA	AAATGAACGA
8201	TTAAAAAACC	TCCCACACCT	CCCCCTGAAC	CTGAAACATA	AAATGAATGC
	AATTTTTTGG	AGGGTGTGGA	GGGGGACTTG	GACTTTGTAT	TTACTTACG
8251	AATTGTTGTT	GTAACTTGT	TTATTGCAGC	TTATAATGGT	TACAAATAAA
	TTAACAACAA	CAATTGAACA	AATAACGTCG	AATATTACCA	ATGTTTATTT
8301	GCAATAGCAT	CACAAATTTT	ACAAATAAAG	CATTTTTTTC	ACTGCATTCT
	CGTTATCGTA	GTGTTTAAAG	TGTTTATTTT	GTAAAAAAG	TGACGTAAGA

15/93

FIG. 10(contd.)

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	ATGTGTTTAC	CGGAGACCGG	AGCGTGTGTA
8501	TCCACATCCA	CCGGTAGGCG	CCAACCGGCT	CCGTTCTTTG	GTGGCCCCCTT
	AGGTGTAGGT	GGCCATCCGC	GGTTGGCCGA	GGCAAGAAAC	CACCGGGGAA
8551	CGCGCCACCT	TCTACTCCTC	CCCTAGTCAG	GAAGTTCCCC	CCCGCCCCCG
	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	ACAAGCCAC	GGTGCGCCTC	GCCACCCGCG
	GGTCGAATGG	TACTGGCTCA	TGTTTCGGGTG	CCACGCGGAG	CGGTGGGCGC
8951	ACGACGTCCC	CAGGGCCGTA	CGCACCTCG	CCGCCGCGTT	CGCCGACTAC
	TGCTGCAGGG	GTCCC GG CAT	GCGTGGGAGC	GGCGGCGCAA	GCGGCTGATG
9001	CCCGCCACGC	GCCACACCGT	CGATCCGGAC	CGCCACATCG	AGCGGGTCAC
	GGGCGGTGCG	CGGTGTGGCA	GCTAGGCCTG	GCGGTGTAGC	TCGCCCAGTG
9051	CGAGCTGCAA	GAACCTCTTC	TCACGCGCGT	CGGGCTCGAC	ATCGGCAAGG
	GCTCGACGTT	CTTGAGAAGG	AGTGCGCGCA	GCCCGAGCTG	TAGCCGTTCC
9101	TGTGGGTGCG	GGACGACGGC	GCCGCGGTGG	CGGTCTGGAC	CACGCCGGAG
	ACACCCAGCG	CCTGCTGCCG	CGGCGCCACC	GCCAGACCTG	GTGCGGCCCTC
9151	AGCGTCGAAG	CGGGGGCGGT	GTTGCGCCGAG	ATCGGCCCGC	GCATGGCCGA
	TCGCAGCTTC	GCCCCCGCCA	CAAGCGGCTC	TAGCCGGGCG	CGTACCGGCT
9201	GTTGAGCGGT	TCCCGGCTGG	CCGCGCAGCA	ACAGATGGAA	GGCCTCCTGG
	CAACTCGCCA	AGGGCCGACC	GGCGCGTCGT	TGTCTACCTT	CCGGAGGACC
9251	CGCCGCACCG	GCCCAAGGAG	CCCGCGTGGT	TCCTGGCCAC	CGTCGGCGTC
	GCGGCGTGCG	CGGGTTCCTC	GGGCGCACCA	AGGACCGGTG	GCAGCCGCAG
9301	TCGCCCCGACC	ACCAGGGCAA	GGGTCTGGGC	AGCGCCGTG	TGCTCCCCGG
	AGCGGGCTGG	TGGTCCCGTT	CCCAGACCCG	TCGCGGCAGC	ACGAGGGGCC
9351	AGTGAGGCG	GCCGAGCGCG	CCGGGGTGCC	CGCCTTCCTG	GAGACCTCCG
	TCACCTCCGC	CGGCTCGCGC	GGCCCCACGG	GCGGAAGGAC	CTCTGGAGGC
9401	CGCCCCGCAA	CCTCCCCTTC	TACGAGCGGC	TCGGCTTCAC	CGTCACCGCC
	GCGGGGCGTT	GGAGGGGAAG	ATGCTCGCCG	AGCCGAAGTG	GCAGTGGCGG
9451	GACGTCGAGG	TGCCCCGAAG	ACCGCGCACC	TGGTGCATGA	CCCGCAAGCC
	CTGCAGCTCC	ACGGGCTTCC	TGGCGCGTGG	ACCACGTACT	GGGCGTTCCG
9501	CGGTGCCTGA	CGCCCGCCCC	ACGACCCGCA	GCGCCCGACC	GAAAGGAGCG
	GCCACGGACT	GCGGGCGGGG	TGCTGGGCGT	CGCGGGCTGG	CTTTCCTCGC
9551	CACGACCCCA	TGCATCGTAG	AGCTCGCTGA	TCAGCCTCGA	CTGTGCCTTC
	GTGCTGGGGT	ACGTAGCATC	TCGAGCGACT	AGTCGGAGCT	GACACGGAAG
9601	TAGTTGCCAG	CCATCTGTTG	TTTGCCCCCTC	CCCCGTGCCT	TCCTTGACCC
	ATCAACGGTC	GGTAGACAAC	AAACGGGGAG	GGGGCACGGA	AGGAACTGGG
9651	TGGAAGGTGC	CACTCCCACT	GTCCTTTCCT	AATAAAATGA	GGAAATTGCA
	ACCTTCCACG	GTGAGGGTGA	CAGGAAAGGA	TTATTTTACT	CCTTTAACGT
9701	TCGCATTGTC	TGAGTAGGTG	TCATTCTATT	CTGGGGGGTG	GGGTGGGGCA
	AGCGTAACAG	ACTCATCCAC	AGTAAGATAA	GACCCCCCAC	CCCACCCCGT

16/93

9751 GGACAGCAAG GGGGGGGATT GGGRAGACAA TAGCAGGCAT GCTGGGGGGG
CCTGTCGTTC CCCCCCTAA CCCYTCTGTT ATCGTCCGTA CGACCCCCCC
9801 CGGTGGGGGC TATGGCTTCT GAGGCGGAAA GAACCAGCTG GGGCTCGAGA
GCCACCCCGG ATACCGAAGA CTCCGCCTTT CTTGGTCGAC CCCGAGCTCT
9851 TCCACTAGTT CTAGCCTCGA GGCTAGAGCG GCCGCCACCG CGGTGGAGCT
AGGTGATCAA GATCGGAGCT CCGATCTCGC CGGCGGTGGC GCCACCTCGA
9901 CCAATTCGCC CTATAGTGAG TCGTATTACG CGCGCTCACT GGCCGTCGTT
GGTTAAGCGG GATATCACTC AGCATAATGC GCGCGAGTGA CCGGCAGCAA
9951 TTACAACGTC GTGACTGGGA AAACCCTGGC GTTACCCAAC TTAATCGCCT
AATGTTGCAG CACTGACCCT TTTGGGACCG CAATGGGTTG AATTAGCGGA
10001 TGCAGCACAT CCCCCTTTCG CCAGCTGGCG TAATAGCGAA GAGGCCCGCA
ACGTCGTGTA GGGGGAAAGC GGTGACCGC ATTATCGCTT CTCCGGGCGT
10051 CCGATCGCCC TTCCCAACAG TTGCGCAGCC TGAATGGCGA ATGGAAATTG
GGCTAGCGGG AAGGGTTGTC AACGCGTCGG ACTTACCGCT TACCTTTAAC
10101 TAAGCGTTAA TATTTTGTGA AAATTCGCGT TAAATTTTTG TTAATCAGC
ATTCGCAATT ATAAAACAAT TTTAAGCGCA ATTTAAAAAC AATTTAGTCG
10151 TCATTTTTTA ACCAATAGGC CGAAATCGGC AAAATCCCTT ATAAATCAAA
AGTAAAAAAT TGGTTATCCG GCTTTAGCCG TTTTAGGGAA TATTTAGTTT
10201 AGAATAGACC GAGATAGGGT TGAGTGTTGT TCCAGTTTGG AACAGAGTC
TCTTATCTGG CTCTATCCCA ACTCACAACA AGGTCAAACC TTGTTCTCAG
10251 CACTATTAAA GAACGTGGAC TCCAACGTCA AAGGGCGAAA AACCGTCTAT
GTGATAATTT CTTGCACCTG AGGTTGCAGT TTCCCGCTTT TTGGCAGATA
10301 CAGGGCGATG GCCCACTACG TGAACCATCA CCCTAATCAA GTTTTTTGGG
GTCCCGCTAC CGGGTGATGC ACTTGGTAGT GGGATTAGTT CAAAAAACC
10351 GTCGAGGTGC CGTAAAGCAC TAAATCGGAA CCCTAAAGGG AGCCCCGAT
CAGCTCCACG GCATTTTCGTG ATTTAGCCTT GGGATTTCCC TCGGGGGCTA
10401 TTAGAGCTTG ACGGGGAAAG CCGGCGAACG TGGCGAGAAA GGAAGGGAAG
AATCTCGAAC TGCCCCTTTC GGCCGCTTGC ACCGCTCTTT CCTTCCCTTC
10451 AAAGCGAAAG GAGCGGGCGC TAGGGCGCTG GCAAGTGTAG CGGTACGCT
TTTCGCTTTC CTCGCCCGCG ATCCCGCGAC CGTTCACATC GCCAGTGCGA
10501 GCGCGTAACC ACCACACCCG CCGCGCTTAA TGCGCCGCTA CAGGGCGCGT
CGCGCATTGG TGGTGTGGGC GGC CGAATT AC GCGCGAT GTCCCGCGCA
10551 CAG
GTC

FIG. 10(contd.)

17/93

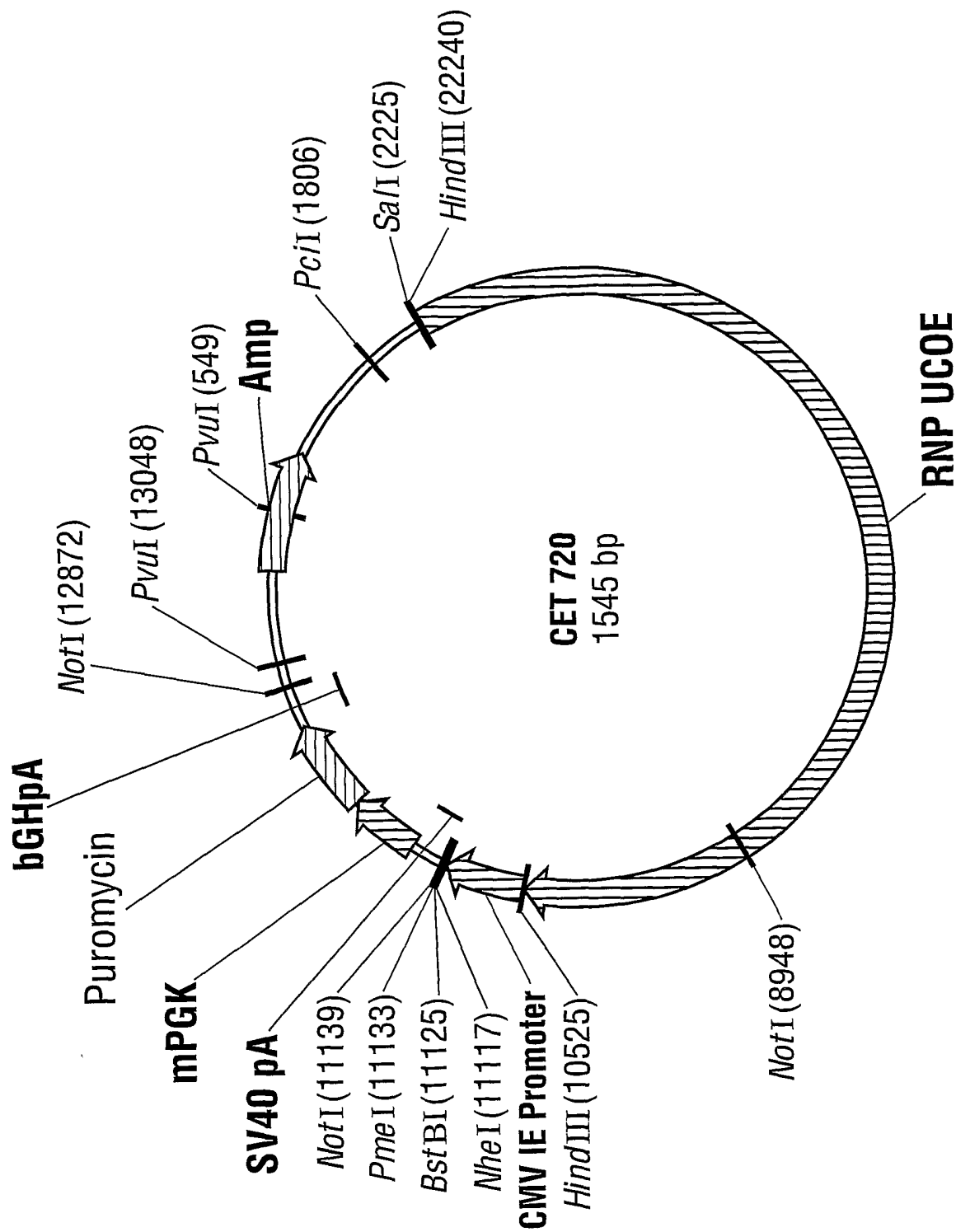


FIG. 11

18/93

FIG. 12 CET720 nucleotide sequence

```

1   GGTGGCACTT TTCGGGGAAA TGTGCGCGGA ACCCCTATTT GTTTATTTTT
    CCACCGTGAA AAGCCCCTTT ACACGCGCCT TGGGGATAAA CAAATAAAAA
51  CTAAATACAT TCAAATATGT ATCCGCTCAT GAGACAATAA CCTTGATAAA
    GATTTATGTA AGTTTATACA TAGGCGAGTA CTCTGTTATT GGGACTATTT
101 TGCTTCAATA ATATTGAAAA AGGAAGAGTA TGAGTATTCA ACATTTCCGT
    ACGAAGTTAT TATAACTTTT TCCTTCTCAT ACTCATAAGT TGTAAGGCA
151 GTCGCCCTTA TTCCCTTTTT TCGGCGCATTT TGCCCTTCCTG TTTTGTCTCA
    CAGCGGGAAT AAGGGAAAAA ACGCCGTAAA ACGGAAGGAC AAAAACGAGT
201 CCCAGAAACG CTGGTGAAAAG TAAAAGATGC TGAAGATCAG TTGGGTGCAC
    GGGTCTTTGC GACCACTTTC ATTTTCTACG ACTTCTAGTC AACCACGTG
251 GAGTGGGTTA CATCGAACTG GATCTCAACA GCGGTAAGAT CCTTGAGAGT
    CTCACCCAAT GTAGCTTGAC CTAGAGTTGT CGCCATTCTA GGAACCTCTA
301 TTTTCGCCCCG AAGAACGTTT TCCAATGATG AGCACTTTTA AAGTTCTGCT
    AAAGCGGGGC TTCTTGCAAA AGGTTACTAC TCGTGAAAAT TTCAAGACGA
351 ATGTGGCGCG GTATTATCCC GTATTGACGC CGGGCAAGAG CAACTCGGTC
    TACACGCGCG 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 TGTGTCTAGC
551 GAGGACCGAA GGAGCTAACC GCTTTTTTGC ACAACATGGG GGATCATGTA
    CTCCTGGCTT CCTCGATTGG CGAAAAACG TGTGTATCCC CCTAGTACAT
601 ACTCGCCTTG ATCGTTGGGA ACCGGAGCTG AATGAAGCCA TACCAAACGA
    TGAGCGGAAC TAGCAACCCT TGGCCTCGAC TTACTTCGGT ATGGTTTGCT
651 CGAGCGTGAC ACCACGATGC CTGTAGCAAT GGCAACAACG TTGCGCAAAC
    GCTCGCACTG TGGTGCTACG GACATCGTTA CCGTTGTTGC AACGCGTTTG
701 TATTAACCTG CGAACTACTT ACTCTAGCTT CCCGGCAACA ATTAATAGAC
    ATAATTGACC GCTTGATGAA TGAGATCGAA GGGCCGTTGT TAATTATCTG
751 TGGATGGAGG CGGATAAAGT TGCAGGACCA CTTCTGCGCT CGGCCCTTCC
    ACCTACCTCC GCCTATTTCA ACGTCTGGT GAAGACGCGA GCCGGGAAGG
801 GGCTGGCTGG TTTATTGCTG ATAAATCTGG AGCCGGTGAG CGTGGGTCTC
    CCGACCGACC AAATAACGAC TATTTAGACC TCGGCCACTC GCACCCAGAG
851 GCGGTATCAT TGCAGCACTG GGGCCAGATG GTAAGCCCTC CCGTATCGTA
    CGCCATAGTA ACGTCGTGAC CCCGGTCTAC CATTGCGGAG GGCATAGCAT
901 GTTATCTACA CGACGGGGAG TCAGGCAACT ATGGATGAAC GAAATAGACA
    CAATAGATGT GCTGCCCCCTC AGTCCGTTGA TACCTACTTG CTTTATCTGT
951 GATCGCTGAG ATAGGTGCCT CACTGATTAA GCATTGGTAA CTGTCAGACC
    CTAGCGACTC TATCCACGGA GTGACTAATT CGTAACCATT GACAGTCTGG
1001 AAGTTTACTC ATATATACTT TAGATTGATT TAAAACTTCA TTTTAAATTT
    TTCAAATGAG TATATATGAA ATCTAACTAA ATTTTGAAGT AAAAATTTAA
1051 AAAAGGATCT AGGTGAAGAT CCTTTTGTAT AATCTCATGA CCAAATCCC
    TTTTCCTAGA TCCACTTCTA GGAAAACTA TTAGAGTACT GGTTTTAGGG
1101 TTAACGTGAG TTTTCGTTCC ACTGAGCGTC AGACCCCGTA GAAAAGATCA
    AATTGCACTC AAAAGCAAGG TGA CTGCGAG TCTGGGGCAT CTTTTCTAGT
1151 AAGGATCTTC TTGAGATCCT TTTTTTCTGC GCGTAATCTG CTGCTTGCAA
    TTCTAGAAAG 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 ATACTGTCTT 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	GTGGAAGTCA	AGACGATAGT
	ACGGTCACCG	CTATTCAGCA	CAGAATGGCC	CAACCTGAGT	TCTGCTATCA
1451	TACCGGATAA	GGCGCAGCGG	TCGGGCTGAA	CGGGGGGTTC	GTGCACACAG
	ATGGCCTATT	CCGCGTCGCC	AGCCCGACTT	GGGGGGCAAG	CACGTGTGTC
1501	CCCAGCTTGG	AGCGAACGAC	CTACACCGAA	CTGAGATACC	TACAGCGTGA
	GGGTCTGAACC	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	TTTTTTGTGAT	GCTCGTCAGG	GGGGCGGAGC	CTATGGAAAA
	ACTCGCAGCT	AAAAACACTA	CGAGCAGTCC	CCCCGCCTCG	GATACCTTTT
1751	ACGCCAGCAA	CGCGGCCTTT	TTACGGTTCC	TGGCCTTTTG	CTGGCCTTTT
	TGCGGTCTGT	GCGCCGAAAA	AATGCCAAGG	ACCGGAAAAAC	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	TTTCGCCCCG	CACTCGCGTT	GCGTTAATTA	CACTCAATCG
2051	TCACTCATT	GGCACCCAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG
	AGTGAGTAAT	CCGTGGGGTC	CGAAATGTGA	AATACGAAGG	CCGAGCATAAC
2101	TTGTGTGGAA	TTGTGAGCGG	ATAACAATTT	CACACAGGAA	ACAGCTATGA
	AACACACCTT	AACACTCGCC	TATTGTATAA	GTGTGTCCTT	TGTCGATACT
2151	CCATGATTAC	GCCAAGCGCG	CAATTAACCC	TCACTAAAGG	GAACAAAAGC
	GGTACTAATG	CGGTTCGCGC	GTTAATTGGG	AGTGATTTC	CTTGTTTTTCG
2201	TGGGTACCGG	GGGGGGGCTC	GAGGTCGACG	GTATCGATAA	GCTTCAATGT
	ACCCATGGCC	CGGGGGGGAG	CTCCAGCTGC	CATAGCTATT	CGAAGTTACA
2251	TTTTAGCACC	CTCTGTGTGG	AGGAAAATAA	TGCAGATTAT	TCTAATTAGT
	AAAATCGTGG	GAGACACACC	TCCTTTTATT	ACGTCTAATA	AGATTAATCA
2301	GTAATATCTA	ACCACATTAA	AATATATTAC	ATAGTAAACT	AACTCCATA
	CATTATAGAT	TGGTGTAATT	TTATATAATG	TATCATTTGA	TGTGAGGTAT
2351	ATTTTATAAA	TTTGACTCCC	CAGGGTAATA	AACTAGTCTC	TAGTCTGCTC
	TAAAATATTT	AAACTGAGGG	GTCCCATTTAT	TTGATCAGAG	ATCAGACGAG
2401	ACCTTCAACT	GTACAATAAA	GTCTTGGTTC	TTTTGAAATA	GACCTCAAAT
	TGGAAGTTGA	CATGTTATTT	CAGAACCAAG	AAAACTTTAT	CTGGAGTTTA
2451	GAGACACCTA	AAATTCAAAG	TGTCTTTACA	TTTAAAGACA	CCTACAGGAA
	CTCTGTGGAT	TTTAAGTTTC	ACAGAAATGT	AAATTTCTGT	GGATGCTCCTT
2501	AGCAGGTAAA	AGAGCCAGGT	TAAAAACAAA	TTCTAAAACC	ACTTAGCTGC
	TCGTCCATTT	TCTCGGTCCA	ATTTTGTGTT	AAGATTTTGG	TGAATCGACG
2551	AGTTAAACAT	ATAGTAAAGA	TGCTACTAAG	TTTCTTACTC	TGTAAATCCC
	TCAATTTGTA	TATCATTTCT	ACGTGATTTT	AAAGAATGAG	ACATTTAGGG
2601	TTCCACTTCA	GGAAATATTC	CACTTTCCCA	TTCACTACAC	GTCGATCTAG
	AAGGTGAAGT	CCTTTATAAG	GTGAAAGGGT	AAGTGATGTG	CAGCTAGATC
2651	TACTTTTTTC	ACGACAAATT	CTTCAGGCTC	TGCCTCTTCA	ACTTTTTTTAC
	ATGAAAAAGG	TGCTGTTTAA	GAAGTCCGAG	ACGGAGAAGT	TGAAAAAATG
2701	TCTTTCCATT	CTGTTTTTTT	CCCATTTTTT	GCTAAAATAA	AACAAAAGAG
	AGAAAGGTAA	GACAAAAAAA	GGGTAAAAAA	CGATTTTATT	TTGTTTTTCTC
2751	AAATTAAGAA	ATATTCCTCT	TGAATTTTGA	GCACATTTTC	AAGGCTCAAT
	TTTAATTCTT	TATAAGGAGA	ACTTAAAAC	CGTGTAAGA	TTCCGAGTTA

FIG. 12(contd.) 20/93

2801	TGCTTATATT	ATTATCACAT	TCGACATAAA	TTTTTACTTC	TATATCCCAG
	ACGAATATAA	TAATAGTGTA	AGCTGTATTT	AAAAATGAAG	ATATAGGGTC
2851	GGCAGACACC	TTCTGGAAAG	ATTAAAGTC	AACAGACAAT	AAAATAAAAG
	CCGTCTGTGG	AAGACCTTTC	TAATTTTTCAG	TTGTCTGTTA	TTTTATTTTC
2901	AATGCTTTAT	CTTGTTTATT	TAGTTCAAAC	TTACAACCCA	CCACCAAAT
	TTACGAAATA	GAACAAGTAA	ATCAAGTTTG	AATGTTGGGT	GGTGGTTTTA
2951	AATACAATAA	AAAAACACTA	TCTGGAAACA	GTTATTTTTT	TCCAGTCTTT
	TTATGTTATT	TTTTTGTGAT	AGACCTTTGT	CAATAAAAAA	AGGTCAGAAA
3001	TTTTTTGAGA	CAGGGTCTCA	CACCTCTGTC	GCCCAGGCTG	GAGTGCAGTG
	AAAAAACTCT	GTCCCAGAGT	GTGAGAACAG	CGGGTCCGAC	CTCACGTCAC
3051	GCGTGATCTC	AGCTCACTGC	AACCTCCGCC	TCCCCAGGTT	CAAGCAGTTC
	CGCACTAGAG	TCGAGTGACG	TTGGAGGCGG	AGGGGTCCAA	GTTCGTCAAG
3101	TCATGCCTCA	GCCCTCCAGAG	TAGCTGGGAT	TATAGGCGGA	TGCCACCATG
	AGTACGGAGT	CGGAGGTCTC	ATCGACCCTA	ATATCCGCCT	ACGGTGGTAC
3151	CCGGGCTAAT	TTTTTTTGTG	TTTTTATTAG	AAACAGGGTT	TCACCATGTT
	GGCCCGATTA	AAAAAAACAC	AAAAATAATC	TTTGTCCCAA	AGTGGTACAA
3201	GACCAGGCTG	GTCTCAAACCT	CCTGACCTGA	AGTGATTACAC	CAGCCTGGGC
	CTGGTCCGAC	CAGAGTTTGA	GGACTGGACT	TCACTAAGTG	GTCCGACCCG
3251	CTCCCAAAGT	GCTGGCATTA	CAGGCGTGAG	CCACTGCGCC	CGGCCCTGTA
	GAGGGTTTCA	CGACCGTAAT	GTCCGCACTC	GGTGACGCGG	GCCGGGACAT
3301	GTCTTAAAG	ACCAAGTTTA	CTAATTTTCA	CTCATTTTAA	CAACACTGCA
	CAGAATTTTC	TGGTTCAAAT	GATTAAAAGT	GAGTAAAATT	GTTGTGACGT
3351	ACAAACAAC	ATGCAGGAAG	TACCTAAAGG	GTGATCCAGA	GAAGCAAGTA
	TGTTTGTGTA	TACGTCCTTC	ATGGATTTC	CACTAGGTCT	CTTCGTTTCA
3401	GTAGTGACAG	GTCTTAGGTG	AACCTATGAC	AGACCTTGTA	TCCACCCCA
	CATCACTGTC	CAGAATCCAC	TTGGATACTG	TCTGGAACAT	AGGTGGGGGT
3451	GATGGTAAAA	GCCCCAGCCC	CCTTCTCAAT	TCAAATATTA	ATGTCAAAAG
	CTACCATTTT	CGGGGTCGGG	GGAAGAGTTA	AGTTTATAAT	TACAGTTTTT
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	TATTTCTTAC
	AATTTCTGT	TCTAATAAAA	AGGTATTAGA	GACATCTATT	ATAAAGGATG
3651	CACCTAGAAC	AAAACATAAA	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	TCAAGATTTT	CTATGAGAGG	ACTAAAAATG	GGGAAAATTT
	AAGGTTACCG	AGTTCTAAAG	GATACTCTCC	TGATTTTTTAC	CCCTTTTAAA
3901	TTATGAGAGG	ATTAAAAATG	GGGGAAAAAA	AACCCTGAAA	TGGTTAATCA
	AATACTCTCC	TAATTTTAC	CCCCTTTTTT	TTGGGACTTT	ACCAATTAGT
3951	GAAGATCCTA	TGGGCTGAGA	AGGAATCCAT	CTTAACATTT	CATCTTAAAG
	CTTCTAGGAT	ACCCGACTCT	TCCTTAGGTA	GAATTGTAAA	GTAAGATTTT
4001	CAAATGCTAT	TGCCGGGGGC	AGTGGCTCAT	GCCTGTAATC	CCAGCACTTT
	GTTTACGATA	ACGGCCCCCG	TCACCGAGTA	CGGACATTAG	GGTCGTGAAA
4051	GGGAGGCCGA	GGTGGGCAGA	TCATCTGAGG	TCAGGAGTTT	GAGACCAGCC
	CCCTCCGGCT	CCACCCGTCT	AGTAGACTCC	AGTCCTCAA	CTCTGGTCGG
4101	TGACCAACAT	GGAGAAACCC	CGTTTCTACT	AAAAATACAA	AATTAGCCAG
	ACTGGTTGTA	CCTCTTTGGG	GCAAAGATGA	TTTTTATGTT	TAAATCGGTC
4151	GCATAGTGGT	GCATGCCTGT	AATCCAGCT	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
	TGTTTTTGT	TTTTGGGTTT	ATGATAAATT	TTTTCTATTT	GGAATTAACG
4351	TCAATCATT	AAGCCATCCC	ACAAGTAAAG	CAGCAAGCAG	AAAAAAGTTA
	AGTTAGTAAT	TTCGGTAGGG	TGTTCATTTT	GTCGTTTCGT	TTTTTTCAAT
4401	AGAACACCTC	AAGGCTACAG	AAGGACATTT	CAAGCTATGC	AGGCATATGA
	TCTTGTGGAG	TTCCGATGTC	TTCCTGTAAA	GTTTCGATACG	TCCGTATACT
4451	AGTGTGCAGA	CAGATATGTA	AGAAAGGCCT	CAAGACTGCA	AAAGGGCATT
	TCACACGTCT	GTCTATACAT	TCTTTCGGGA	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	ATTCTCATT	GTCAAGATAC	TTGAGTACTG
	AAAGATTCAG	TCCACAATAC	TAAGAGTAAT	CAGTTCATATG	AACTCATGAC
4651	GGCCCAAACA	GCTTCTGCC	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	GTGTTCCCT	CTTCCGATTG
4801	AAAATAAAGT	CAACAATAAA	TACAGAAAAT	GAAAAGGATA	CACACTAAGG
	TTTTATTTCA	GTTGTTATTT	ATGTCTTTTA	CTTTTCCTAT	GTGTGATTCC
4851	TTTACAAAAA	AAAAAAGGCA	GACAAAATGC	CATACAGTAT	TCATTCACTA
	AAATGTTTTT	TTTTTCCGT	CTGTTTTACG	GTATGTCATA	AGTAAGTGAT
4901	CTATGGCATT	CATAAGCTAG	TTTCAAATGC	TCACTATTTT	CTTTTATAGT
	GATACCGTAA	GTATTCGATC	AAAGTTTACG	AGTGATAAAA	GAAAATATCA
4951	ATATATTTGC	CTTAACCCAG	CACTTTTTTC	CAAAAGTGGA	TGAGTCAAAA
	TATATAAACG	GAATTGGGTC	GTGAAAAAAG	GTTTTACCT	ACTCAGTTTT
5001	TAAATTTCCC	ATTATTTAAG	TGAAATTAAC	AGCACACATA	TCTCACAACA
	ATTTAAAGGG	TAATAAATTC	ACTTTAATTG	TCGTGTGTAT	AGAGTGTTGT
5051	CTAATGAATT	TTTAAAATGG	AAAGTTAAGA	ACTTTTAAAG	TGGCCAACCT
	GATTACTTAA	AAATTTTACC	TTTCAATTCT	TGAAAATTTT	ACCGGTTGGA
5101	GTGATCCTTC	ACAAAATAAA	CTAAATACAA	TAACAGACCC	CAAAGGCTAT
	CACTAGGAAG	TGTTTTATTT	GATTTATGTT	ATTGTCTGGG	GTTTCCGATA
5151	CAATTGCGTG	CAAAAACAAC	TTCTGTTTTT	CAGGGTAAAC	AGAATCTAAT
	GTTAACGCAC	GTTTTTGTTG	AAGACAAAAG	GTCCCATTTG	TCTTAGATTA
5201	GCAGAACTA	ATGCAGGGTA	AACAGACTTA	ATGCAGAATC	TAATGATGGC
	CGTCTTĀGAT	TACGTCCCAT	TTGTCTGAAT	TACGTCTTAG	ATTACTACCG
5251	ACAAATTAAA	AATCACTAAC	GTGCCCTTTT	TAGTGTGAAA	CCCAGAGAGA
	TGTTTAATTT	TTAGTGATTG	CACGGGAAAA	ATCACACTTT	GGGTCTCTCT
5301	GCACATACAA	GCCAAAACA	AATGCTTTAT	TTTACCTAGG	AGACATTAAC
	CGTGTATGTT	CGGTTTTTGT	TTACGAAATA	AAATGGATCC	TCTGTAATTG
5351	ATTCACCTTT	ACGTGTTTAA	GATTAATGCA	ATGTTAAATA	TTGTGAAAAC
	TAAGTGGAAG	TGCACAAATT	CTAATTACGT	TACAATTTAT	AACACTTTTG
5401	TGTAACCTTG	AATTCATGA	TTTTTATGTG	AATATTCAG	GGTTTAAAAA
	ACATTGAAAC	TTAAAGTACT	AAAAATACAC	TTATAAGGTC	CCAAATTTTT
5451	AACTTGĀAAC	ATGACATGGC	TGAATAAGAT	AAAAAAAAAA	TCTAGCCTTT
	TTGAACATTG	TACTGTACCG	ACTTATTCTA	TTTTTTTTTT	AGATCGGAAA
5501	TCTCCCTTCT	GGCTCATATT	TGCGATTTCG	ATCATTTTGT	TTAAAAAACA
	AGAGGGAAGA	CCGAGTATAA	ACGCTAAAGC	TAGTAAAACA	AATTTTTTGT
5551	AAACACTGCA	ATGAATTAAG	CTTAATATTC	TTCTATGTTT	TAGAGTAAGT
	TTTGTGACGT	TACTTAATTT	GAATTATAAG	AAGATACAAA	ATCTCATTCA
5601	TAAAACAAGA	TAAAGTGACC	AAAGTAATTT	GAAAGATTCA	ATGACTTTTG
	ATTTTGTTCT	ATTTCACTGG	TTTCATTAAA	CTTTCTAAGT	TACTGAAAAC
5651	CTCCAACCTA	GGTGACAAAG	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 TTTTACACA
5801 GTATTTCTAT TTGAAGATAA AAATCCTCAG CAAAATCCAC ATTGCACTGT
      CATAAAGATA AACTTCTATT TTTAGGAGTC GTTTTAGGTG TAACGTGACA
5851 CCTTCAAAGA TTAGCCTTCT TTGAAGTAGT TAAGACACTA TTAAGCCAAG
      GGAAGTTTCT AATCGGAAGA AACTTGATCA ATTCTGTGAT AATTCGGTTC
5901 CCAGTATCTC CCTGTAATGA ATTCTGTTTT CTCTTAATTT TCCCCTGTAA
      GGTACATAGAG GGACATTACT TAAGCAAAAA GAGAATTAAA AGGGGACATT
5951 TTTACACTGG GAGAGCTGGG AAATATGTGG ATGTAAATTT CTCAGCCACA
      AAATGTGACC CTCTCGACCC TTTATACACC TACATTTAAA GAGTCGGTGT
6001 GAGATGCAAA GTTATACTGT GGGGAAAAAA AACTTGAGTT AAATCCTTAC
      CTCTACGTTT CAATATGACA CCCCTTTTTT TTGAAGTCAA TTTAGGAATG
6051 ATATTTTAGG TTTTCATTAA CTTACCAATG TAGTTTGTGTT GGAGGCCATT
      TATAAAATCC AAAAGTAATT GAATGGTTAC ATCAAAACAA CCTCCGGTAA
6101 TTTTTTATTG CAGACTTGAA GAGCTATTAC TAGAAAAATG CATGACAGTT
      AAAAAATAAC GTCTGAAGTT CTCGATAATG ATCTTTTTTAC GTACTGTCAA
6151 AAGGTAAGTT TGCATGACAC AAAAAAGGTA ACTAAATACA AATTCTGTTT
      TTCCATTCAA ACGTACTGTG TTTTTTCCAT TGATTTATGT TTAAGACAAA
6201 GGATTCCAAC CCCCAAGTAG AGAGCGCACA CTTTCAAACG TGAATACAAA
      CCTAAGGTTG GGGGTTTCATC TCTCGCGTGT GAAAGTTTGC ACTTATGTTT
6251 TCCAGAGTAG ATCTGCGCTC CTACCTACAT TGCTTATGAT GTACTTAAGT
      AGGTCTCATC TAGACGCGAG GATGGATGTA ACGAATACTA CATGAATTCA
6301 ACGTGTCTTA ACCATGTGAG TCTAGAAAGA CTTTACTGGG GATCCTGGTA
      TGCACAGGAT TGGTACACTC AGATCTTTCT GAAATGACCC CTAGGACCAT
6351 CCTAAAACAG CTTACACTGG CTTAAAATAG GGGACCAATG TCTTTTCCAA
      GGATTTTGTG GAAGTGTACC GAATTTTATC CCCTGGTTAC AGAAAAGGTT
6401 TCTAAGTCCC ATTTATAATA AAGTCCATGT TCCATTTTTA AAGGACAATC
      AGATTCAGGG TAAATATTAT TTCAGGTACA AGGTAAAAAT TTCCTGTTAG
6451 CTTTCGGTTT AAAACCAGGC ACGATTACCC AAACAACCTCA CAACGGTAAA
      GAAAGCCAAA TTTTGGTCCG TGCTAATGGG TTTGTTGAGT GTTGCCATTT
6501 GCACTGTGAA TCTTCTCTGT TCTGCAATCC CAACTTGTTT TCTGCTCAGA
      CGTGACACTT AGAAGAGACA AGACGTTAGG GTTGAACCAA AGACGAGTCT
6551 AACCTCCCT CTTTCCAATC GGTAATTAAA TAACAAAAGG AAAAACTTA
      TTGGGAGGGA GAAAGGTTAG CCATTAATTT ATTGTTTTC TTTTGTGAAT
6601 AGATGCTTCA ACCCGTTTC GTGACACTTT GAAAAAAGAA TCACCTCTTG
      TCTACGAAGT TGGGGCAAAG CACTGTGAAA CTTTTTTCTT AGTGGAGAAC
6651 CAAACACCCG CTCCCGACCC CCGCCGCTGA AGCCCGGCGT CCAGAGGCCT
      GTTTGTGGGC GAGGGCTGGG GCGGCGGACT 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 CACGTTAAC TGGCGTGGCG
      GGTGTGGCGT GCGCGGATGG GGGCGCGCCG GTGCAATTGA CCGCCAGCGG
6851 GCAGCCTCGG GACAGCCGGC CGCGCGCCGC CAGGCTCGCG GACGCGGGAC
      CGTCGGAGCC CTGTGCGCCG GCGCGCGGCG GTCCGAGCGC CTGCGCCCTG
6901 CACGCGCCGC CCTCCGGGAG GCCCAAGTCT CGACCCAGCC CGCGTGGCG
      GTGCGCGGCG GGAGGCCCTC CGGGTTTACA GCTGGGTGCG GCGCACCGC
6951 CTGGGGGAGG GGGCGCCTCC GCCGGAACGC GGGTGGGGGA GGGGAGGGGG
      GACCCCTCC CCCGCGGAGG CGGCCTTGCG CCCACCCCT CCCCTCCCCC
7001 AAATGCGCTT TGTCTCGAAA TGGGGCAACC GTCGCCACAG CTCCCTACCC
      TTTACGCGAA ACAGAGCTTT ACCCGTTGG CAGCGGTGTC GAGGGATGGG
7051 CCTCGAGGGC AGAGCAGTCC CCCCACTAAC TACCGGGCTG GCCGCGCGCC
      GGAGCTCCCG TCTCGTCAGG GGGGTGATTG ATGGCCCGAC CGGCGCGCGG
7101 AGGCCAGCCG CGAGGCCACC GCCCGACCCT CCACTCCTTC CCGCAGCTCC
      TCCGGTCGGC GCTCCGGTGG CGGGCTGGGA GGTGAGGAAG GCGTCGAGG

```

FIG. 12(contd.) 23/93

7151	CGGCGCGGGG	TCCGGCGAGA	AGGGGAGGGG	AGGGGAGCGG	AGAACCGGGC
	GCCGCGCCCC	AGGCCGCTCT	TCCCCTCCCC	TCCCCTCGCC	TCTTGGCCCCG
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
	CAGACGTGCG	AGAGGCCGGG	AGGCGCGGAT	GTCGAGTTCG	GTGTAGGCTT
7351	GGGGGAGGGA	GCCGGGAGCT	GCGCGCGGGG	CCGCCGGGGG	GAGGGGTGGC
	CCCCCTCCCT	CGGCCCTCGA	GCGCGCGCCC	GGCGGCCCCC	CTCCCCACCG
7401	ACCGCCACG	CCGGGCGGCC	ACGAAGGGCG	GGGCAGCGGG	CGCGCGCGCG
	TGGCGGGTGC	GGCCCGCCGG	TGCTTCCCGC	CCCGTCGCCC	GCGCGCGCGC
7451	GCGGGGGGAG	GGGCCGGCGC	GCGCCCCGCT	GGGAATTGGG	GCCCTAGGGG
	CGCCCCCTC	CCCGGCCGCG	GCGCGGGCGA	CCCTTAACCC	CGGGATCCCC
7501	GAGGGCGGAG	GCGCCGACGA	CCGCGGCACT	TACCGTTCGC	GGCGTGGCGC
	CTCCCGCCTC	CGCGGCTGCT	GGCGCCGTGA	ATGGCAAGCG	CCGCACCGCG
7551	CCGGTGGTCC	CCAAGGGGAG	GGAAGGGGGA	GGCGGGGCGA	GGACAGTGAC
	GGCCACCAGG	GGTTCCCTC	CCTTCCCCCT	CCGCCCCGCT	CCTGTCACTG
7601	CGGAGTCTCC	TCAGCGGTGG	CTTTTCTGCT	TGGCAGCCTC	AGCGGCTGGC
	GCCTCAGAGG	AGTCGCCACC	GAAAAGACGA	ACCGTCGGAG	TCGCCGACCG
7651	GCCAAAACCG	GACTCCGCC	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
	GGGAAACCTT	GGTGAAAGTC	GTCGCTTGAG	AGGACATGTG	GTCCCCAGTC
7801	TTCCACAGAC	GCGGGCCAGG	GGTGGGTCAT	TGCGGCGTGA	ACAATAATTT
	AAGGTGTCTG	CGCCCCGTCC	CCACCCAGTA	ACGCCGCACT	TGTTATTAAA
7851	GACTAGAAGT	TGATTGCGGT	GTTTCCGGAA	GGGGCCGAGT	CAATCCGCCG
	CTGATCTTCA	ACTAAGCCCA	CAAAGGCCCT	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	CCGTCCGGCT	TAATTTCAAA
	ACCTAGAGGT	CCGACGCCTA	ATTCCGGGAG	GGCAGCCGCA	ATTAAAGTTT
8051	CTGCGCGACG	TTTCTCACCT	GCCTTCGCCA	AGGCAGGGGC	CGGGACCCTA
	GACGCGCTGC	AAAGAGTGGA	CGGAAGCGGT	TCCGTCCCCG	GCCCTGGGAT
8101	TTCCAAGAGG	TAGTAACTAG	CAGGACTCTA	GCCTTCGCCA	ATTCATTGAG
	AAGGTTCTCC	ATCATTGATC	GTCCTGAGAT	CGGAAGGCGT	TAAGTAACTC
8151	CGCATTTACG	GAAGTAACGT	CGGGTACTGT	CTCTGGCCGC	AAGGGTGGGA
	GCGTAAATGC	CTTCATTGCA	GCCCATTGACA	GAGACCGGCG	TTCCCAACCT
8201	GGAGTACGCA	TTTGGCGTAA	GGTGGGGCGT	AGAGCCTTCC	CGCCATTGGC
	CCTCATGCGT	AAACCGCATT	CCACCCCGCA	TCTCGGAAGG	GCGGTAACCG
8251	GGCGGATAGG	GCGTTTACGC	GACGGCCTGA	CGTAGCGGAA	GACGCGTTAG
	CCGCCTATCC	CGCAAATGCG	CTGCCGGA	GCATCGCCTT	CTGCGCAATC
8301	TGGGGGGGAA	GGTTCTAGAA	AAGCGGCGGC	AGCGGCTCTA	GCGGCAGTAG
	ACCCCCCTT	CCAAGATCTT	TTCCGCCGCG	TCGCCGAGAT	CGCCGTCATC
8351	CAGCAGCGCC	GGGTCCCGTG	CGGAGGTGCT	CCTCGCAGAG	TTGTTTCTCG
	GTCGTGCGCG	CCCAGGGCAC	GCCTCCACGA	GGAGCGTCTC	AACAAAGAGC
8401	AGCAGCGGCA	GTTCTCACTA	CAGCGCCAGG	ACGAGTCCGG	TTCGTGTTCCG
	TGCTCGCCGT	CAAGAGTGAT	GTCGCGGTCC	TGCTCAGGCC	AAGCACAAGC
8451	TCCGCGGAGA	TCTCTCTCAT	CTCGCTCGGC	TGCGGGAAAT	CGGGCTGAAG
	AGGCGCCTCT	AGAGAGAGTA	GAGCGAGCCG	ACGCCCTTTA	GCCCCGACTTC
8501	CGACTGAGTC	CGCGATGGAG	GTAACGGGTT	TGAAATCAAT	GAGTTATTGA
	GCTGACTCAG	GCGCTACCTC	CATTGCCCAA	ACTTTAGTTA	CTCAATAACT
8551	AAAGGGCATG	GCGAGGCCGT	TGGCGCCTCA	GTGGAAAGTCG	GCCAGCCGCC
	TTTCCCGTAC	CGCTCCGGCA	ACCGCGGAGT	CACCTTCAGC	CGGTCCGGCG

FIG. 12(contd.) 24/93

8601	TCCGTGGGAG	AGAGGCAGGA	AATCGGACCA	ATTCAGTAGC	AGTGGGGGCTT
	AGGCACCCTC	TCTCCGTCCT	TTAGCCTGGT	TAAGTCATCG	TCACCCCGAA
8651	AAGGTTTATG	AACGGGGTCT	TGAGCGGAGG	CCTGAGCGTA	CAAACAGCTT
	TTCCAAATAC	TTGCCCCAGA	ACTCGCCTCC	GGA CT CGCAT	GTTTGT CGAA
8701	CCCCACCCTC	AGCCTCCCGG	CGCCATTTCC	CTTCACTGGG	GGTGGGGGAT
	GGGGTGGGAG	TCGGAGGGCC	GCGGTAAAGG	GAAGTGACCC	CCACCCCTA
8751	GGGGAGCTTT	CACATGGCGG	ACGCTGCCCC	GCTGGGGTGA	AAGTGGGGCG
	CCCCTCGAAA	GTGTACCGCC	TGCGACGGGG	CGACCCCACT	TTACCCCGC
8801	CGGAGGCGGG	AATTCTTATT	CCCTTTCTAA	AGCACGCTGC	TTCGGGGGCC
	GCCTCCGCCC	TTAAGAATAA	GGGAAAGATT	TCGTGCGACG	AAGCCCCCGG
8851	ACGGCGTCTC	CTCGGCGAGC	GTTTCGGCGG	GCAGCAGGTG	CTCGTAGCGG
	TGCCGCAGAG	GAGCCGCTCG	CAAAGCCGCC	CGTCGTCCAC	GAGCACTCGC
8901	AGGCTGCGGA	GCTTCCCCTC	CCCCTCTCTC	CCGGGAACCG	ATTTGGCGGC
	TCCGACGCCT	CGAAGGGGAG	GGGGAGAGAG	GGCCCTTGGC	TAAACCGCCG
8951	CGCCATTTTC	ATGGCTCGCC	TTCTCTCAG	CGTTTTCTCT	ATAACTCTTT
	GCGGTAAAAG	TACCGAGCGG	AAGGAGAGTC	GCAAAAGGAA	TATTGAGAAA
9001	TATTTTCTTA	GTGTGCTTTC	TCTATCAAGA	AGTAGAAGTG	GTAACTATT
	ATAAAAGAAT	CACACGAAAG	AGATAGTTCT	TCATCTTCAC	CAATTGATAA
9051	TTTTTTTTCT	TCTCGGGCTG	TTTTCATATC	GTTTCGAGGT	GGATTTGGAG
	AAAAAAAAGA	AGAGCCCGAC	AAAAGTATAG	CAAAGCTCCA	CCTAAACCTC
9101	TGTTTTGTGA	GCTTGATCT	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	GAGGCCCAA	CGCCCGGGCC	AAAGTCTCTT
9251	GATATCACCA	CCCAGGGCGT	CGGGCCGGGT	TCAATGCGAG	CCGTAGGACA
	CTATAGTGGT	GGGTCCCGCA	GCCCCGCCCA	AGTTACGCTC	GGCATCCTGT
9301	AAGAAACCAT	TTTATGTTTT	TCCTGTCTTT	TTTTTCCTTT	GAGTAACGGT
	TTCTTTGGTA	AAATACAAA	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	GAAGTTTAA	CCTTACGGGA	CAGATCCATG	GACCTTTATA	GATGATTACG
	CTTCAAATTG	GGAATGCCCT	GTCTAGGTAC	CTGGAATAT	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	AAAAGAACT	GGGATTCGTG	CAAATATTTT
9851	TAACATTGTT	TAAAATCGAT	AGTGGACATC	GGGTAAAGTT	GGATAAATTG
	ATTGTAACAA	ATTTTAGCTA	TCACCTGTAG	CCCATTCAAA	CCTATTAAAC
9901	TGAGGTAAGT	AATGAGTTTT	TGCTTTTTGT	TAGTGATTGT	TAAAACCTGT
	ACTCCATTCA	TTACTCAAAA	ACGAAAAACA	ATCACTAAAC	ATTTTGAACA
9951	TATAAATGTA	CATTATCCGT	AATTCAGTT	TAGAGATAAC	CTATGTGCTG
	ATATTTACAT	GTAATAGGCA	TTAAAGTCAA	ATCTCTATTG	GATACACGAC
10001	ACGACAATTA	AGAATAAAAA	CTAGCTGAAA	AAATGAAAAT	AACTATCGTG
	TGCTGTAAAT	TCTTATTTTT	GATCGACTTT	TTTACTTTTA	TTGATAGCAC

FIG. 12(contd.) 25/93

10051	ACAAGTAACC	ATTTCAAAAG	ACTGCTTTGT	GTCTCATAGG	AGCTAGTTTG
	TGTTTCATTGG	TAAAGTTTTT	TGACGAAACA	CAGAGTATCC	TCGATCAAAC
10101	ATCATTTTCAG	TTAATTTTTT	CTTTAATTTT	TACGAGTCAT	GAAAACTACA
	TAGTAAAGTC	AATTAAAAAA	GAAATTAAAA	ATGCTCAGTA	CTTTTGATGT
10151	GGAAAAAAA	TCTGAACTGG	GTTTTACCAC	TACTTTTTTAG	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	GGTTTTTTTAA	CCTAGTTAAC	TTTTAAGACA	GGTTTCCTTA
	AGAGCAAAC	CCAAAAAATT	GGATCAATTG	AAAATTCTGT	CCAAAGGAAT
10401	ACATTCATAA	GTGCCCAGAA	TACAGCTGTG	TAGTACAGCA	TATAAAGATT
	TGTAAGTATT	CACGGGTCTT	ATGTCGACAC	ATCATGTCTG	ATATTTCTAA
10451	TCAGCTCTGA	GGTTTTTTCCT	ATTGACTTGG	AAAATTGTTT	TGTGCCTGTC
	AGTCGAGACT	CCAAAAAGGA	TAAGTGAACC	TTTTAACAAA	ACACGGACAG
10501	GCTTGCCACA	TGGCCAATCA	AGTAAGCTTG	ATTAATAGTA	ATCAATTACG
	CGAACGGTGT	ACCGGTTAGT	TCATTCTGAAC	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	GTGATGTAGT
10751	AGTGTATCAT	ATGCCAAGTA	CGCCCCCTAT	TGACGTCAAT	GACGGTAAAT
	TCACATAGTA	TACGGTTTCAT	GCGGGGGATA	ACTGCAGTTA	CTGCCATTTA
10801	GGCCCGCCTG	GCATTATGCC	CAGTACATGA	CCTTATGGGA	CTTTCCTACT
	CCGGGCGGAC	CGTAATACGG	GTGATGTACT	GGAATACCCT	GAAAGGATGA
10851	TGGCAGTACA	TCTACGTATT	AGTCATCGCT	ATTACCATGG	TGATGCGGTT
	ACCGTCATGT	AGATGCATAA	TCAGTAGCGA	TAATGGTACC	ACTACGCCAA
10901	TTGGCAGTAC	ATCAATGGGC	GTGGATAGCG	GTTTGACTCA	CGGGGATTTT
	AACCGTCATG	TAGTTACCCG	CACCTATCGC	CAAACAGAGT	GCCCCTAAAG
10951	CAAGTCTCCA	CCCCATTGAC	GTCAATGGGA	GTTTGTTTTG	GCACCAAAAT
	GTTTCAGAGGT	GGGGTAACTG	CAGTTACCCT	CAAACAAAAC	CGTGGTTTTA
11001	CAACGGGACT	TTCCAAAATG	TCGTAACAAC	TCCGCCCCAT	TGACGCAAAT
	GTTGCCCTGA	AAGGTTTAC	AGCATTTGTT	AGGCGGGGTA	ACTGCGTTTA
11051	GGGCGGTAGG	CGTGTACGGT	GGGAGGTCTA	TATAAGCAGA	GCTGGTTTAG
	CCCGCCATCC	GCACATGCCA	CCCTCCAGAT	ATATTCGTCT	CGACCAAATC
11101	TGAACCGTCA	GATCCGCTAG	CGTTGGAAGT	TTAAACGCGG	CCGCGACTCT
	ACTTGGCAGT	CTAGGCGATC	GCAAGCTTCA	AATTTGCGCC	GGCGCTGAGA
11151	AGATCATAAT	CAGCCATACC	ACATTTGTAG	AGGTTTTACT	TGCTTTAAAA
	TCTAGTATTA	GTGCGGTATG	TGTAAACATC	TCCAAAATGA	ACGAAATTTT
11201	AACCTCCCAC	ACCTCCCCCT	GAACCTGAAA	CATAAAATGA	ATGCAATTGT
	TTGGAGGGTG	TGGAGGGGGA	CTTGGACTTT	GTATTTTACT	TACGTTAACA
11251	TGTTGTTAAC	TTGTTTATTG	CAGCTTATAA	TGGTTACAAA	TAAAGCAATA
	ACAACAATTG	AACAAATAAC	GTCGAATATT	ACCAATGTTT	ATTTTCGTTAT
11301	GCATCACAAA	TTTCACAAAT	AAAGCATTTT	TTTCACTGCA	TTCTAGTTGT
	CGTAGTGTTT	AAAGTGTTTA	TTTCGTAAAA	AAAGTGACGT	AAGATCAACA
11351	GGTTTGTTCCA	AACTCATCAA	TGTATCTTAA	ATCGAATTCT	ACCGGGTAGG
	CCAAACAGGT	TTGAGTAGTT	ACATAGAATT	TAGCTTAAGA	TGGCCCATCC
11401	GGAGGCGCTT	TTCCCAAGGC	AGTCTGGAGC	ATGCGCTTTA	GCAGCCCCGC
	CCTCCGCGAA	AAGGGTTCCG	TCAGACCTCG	TACGCGAAAT	CGTCGGGGCG
11451	TGGGCACTTG	GCGCTACACA	AGTGCCCTCT	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	CCTCCCCCTAG	TCAGGAAGTT	CCCCCCCCGC	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	TCCGACCCCT
11751	GGGGTGGGTC	CGGGGGCGGG	CTCAGGGGCG	GGCTCAGGGG	CGGGGCGGGC
	CCCCACCCAG	GCCCCCGCCC	GAGTCCCCGC	CCGAGTCCCC	GCCCCGCCCC
11801	GCCCCAAGGT	CCTCCCGGAG	CCCGGCATTC	TGCACGCTTC	AAAAGCGCAC
	CGGGCTTCCA	GGAGGCCCTC	GGGCCGTAAG	ACGTGCGAAG	TTTTCGCGTG
11851	GTCTGCCGCG	CTGTTCTCCT	CTTCCTCATC	TCCGGGCCTT	TCGACCAGCT
	CAGACGGCGC	GACAAGAGGA	GAAGGAGTAG	AGGCCCGGAA	AGCTGGTCGA
11901	TACCATGACC	GAGTACAAGC	CCACGGTGCG	CCTCGCCACC	CGCGACGACG
	ATGGTACTGG	CTCATGTTCT	GGTGCCACGC	GGAGCGGTGG	GCGCTGCTGC
11951	TCCCCAGGGC	CGTACGCACC	CTCGCCGCCG	CGTTTCGCCG	CTACCCCGCC
	AGGGGTCCCG	GCATGCGTGG	GAGCGGCGGC	GCAAGCGGCT	GATGGGGCGG
12001	ACGCGCCACA	CCGTGATCC	GGACCGCCAC	ATCGAGCGGG	TCACCGAGCT
	TGCGCGGTGT	GGCAGCTAGG	CCTGGCGGTG	TAGCTCGCCC	AGTGGCTCGA
12051	GCAAGAACTC	TTCTCACGC	GCGTCGGGCT	CGACATCGGC	AAGGTGTGGG
	CGTTCTTGAG	AAGGAGTGCG	CGCAGCCCGA	GCTGTAGCCG	TTCCACACCC
12101	TCGCGGACGA	CGGCGCCGCG	GTGGCGGTCT	GGACCACGCC	GGAGAGCGTC
	AGCGCCTGCT	GCCGCGGCGC	CACCGCCAGA	CCTGGTGCGG	CCTCTCGCAG
12151	GAAGCGGGGG	CGGTGTTTCG	CGAGATCGGC	CCGCGCATGG	CCGAGTTGAG
	CTTCGCCCCC	GCCACAAGCG	GCTCTAGCCG	GGCGCGTACC	GGCTCAACTC
12201	CGGTTCCCCG	CTGGCCGCGC	AGCAACAGAT	GGAAGGCCTC	CTGGCGCCGC
	GCCAAGGGCC	GACCGGCGCG	TCGTTGTCTA	CCTTCCGGAG	GACCTCGCCG
12251	ACCGGCCCAA	GGAGCCCGCG	TGGTTCCTGG	CCACCGTCGG	GCTCTCGCCC
	TGGCCGGGTT	CCTCGGGCGC	ACCAAGGACC	GGTGGCAGCC	GCAGAGCGGG
12301	GACCACCAGG	GCAAGGGTCT	GGGCAGCGCC	GTGCTGCTCC	CCGGAGTGGA
	CTGGTGGTCC	CGTTCCCAGA	CCCGTCGCGG	CAGCACGAGG	GGCCTCACCT
12351	GGCGGCCGAG	CGCGCCGGGG	TGCCCGCCTT	CCTGGAGACC	TCCGCGCCCC
	CCGCCGGCTC	GCGCGGCCCC	ACGGCGGGAA	GGACCTCTGG	AGGCGCGGGG
12401	GCAACCTCCC	CTTCTACGAG	CGGCTCGGCT	TCACCGTCAC	CGCCGACGTC
	CGTTGGAGGG	GAAGATGCTC	GCCGAGCCGA	AGTGGCAGTG	GCGGCTGCAG
12451	GAGGTGCCCC	AAGGACCGCG	CACCTGGTGC	ATGACCCGCA	AGCCCGGTGC
	CTCCACGGGC	TTCTTGGCGC	GTGGACCACG	TACTGGGCGT	TCGGGCCACG
12501	CTGACGCCCC	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	GGTGTTCATC	TATTCTGGGG	GGTGGGGTGG	GGCAGGACAG
	ACAGACTCAT	CCACAGTAAG	ATAAGACCCC	CCACCCACC	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	AGCGGCCCGC	ACCGCGGTGG	AGCTCCAATT
	TCAAGATCGG	AGCTCCGATC	TCGCCGGCGG	TGGCGCCACC	TCGAGGTTAA
12901	CGCCCTATAG	TGAGTCGTAT	TACGCGCGCT	CACTGGCCGT	CGTTTTACAA
	GCGGGATATC	ACTCAGCATA	ATGCGCGCGA	GTGACCGGCA	GCAAAATGTT

27/93

12951	CGTCGTGACT	GGGAAAACCC	TGGCGTTACC	CAACTTAATC	GCCTTGCAGC
	GCAGCACTGA	CCCTTTTGGG	ACCGCAATGG	GTTGAATTAG	CGGAACGTCC
13001	ACATCCCCCT	TTCGCCAGCT	GGCGTAATAG	CGAAGAGGCC	CGCACCGATC
	TGTAGGGGGA	AAGCGGTCGA	CCGCATTATC	GCTTCTCCGG	GCGTGGCTAG
13051	GCCCTTCCCA	ACAGTTGCGC	AGCCTGAATG	GCGAATGGAA	ATTGTAAGCG
	CGGGAAGGGT	TGTCAACGCG	TCGGACTTAC	CGCTTACCTT	TAACATTCCG
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	AACCTTGTTT	TCAGGTGATA
13251	TAAAGAACGT	GGACTCCAAC	GTCAAAGGGC	GAAAAACCGT	CTATCAGGGC
	ATTTCTTGCA	CCTGAGGTTG	CAGTTTCCCG	CTTTTTGGCA	GATAGTCCCG
13301	GATGGCCAC	TACGTGAACC	ATCACCCCTA	TCAAGTTTTT	TGGGGTCGAG
	CTACCGGGTG	ATGCACTTGG	TAGTGGGATT	AGTTCAAAAA	ACCCAGCTC
13351	GTGCCGTAAA	GCACTAAATC	GGAACCCTAA	AGGGAGCCCC	CGATTTAGAG
	CACGGCATT	CGTGATTAG	CCTTGGGATT	TCCCTCGGGG	GCTAAATCTC
13401	CTTGACGGGG	AAAGCCGGCG	AACGTGGCGA	GAAAGGAAGG	GAAGAAAGCG
	GAACTGCCCC	TTTCGGCCGC	TTGCACCGCT	CTTTCCTTCC	CTTCTTTTCG
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.)

28/93

FIG. 13

S. alboniger puromycin N-acetyl transferase
nucleotide sequence

```
1  ATGACCGAGT ACAAGCCCAC GGTGCGCCTC GCCACCCGCG ACGACGTCCC
   TACTGGCTCA TGTTCGGGTG CCACGCGGAG CGGTGGGCGC TGCTGCAGGG
51  CAGGGCCGTA CGCACCCCTCG CCGCCGCGTT CGCCGACTAC CCCGCCACGC
   GTCCCGGCAT GCGTGGGAGC GCGGCGCGAA GCGGCTGATG GGGCGGTGCG
101  GCCACACCGT CGATCCGGAC CGCCACATCG AGCGGGTCAC CGAGCTGCAA
   CGGTGTGGCA GCTAGGCCTG GCGGTGTAGC TCGCCCAGTG GCTCGACGTT
151  GAACTCTTCC TCACGCGCGT CGGGCTCGAC ATCGGCAAGG TGTGGGTGCG
   CTTGAGAAGG AGTGCGCGCA GCCCGAGCTG TAGCCGTTCC ACACCCAGCG
201  GGACGACGGC GCCGCGGTGG CGGTCTGGAC CACGCCGGAG AGCGTCGAAG
   CCTGCTGCCG CGGCGCCACC GCCAGACCTG GTGCGGCCTC TCGCAGCTTC
251  CGGGGGCGGT GTTCGCCGAG ATCGGCCCCG GCATGGCCGA GTTGAGCGGT
   GCCCCCGCCA CAAGCGGCTC TAGCCGGGCG CGTACCGGCT CAACTCGCCA
301  TCCCGGCTGG CCGCGCAGCA ACAGATGGAA GGCTCCTTGG CGCCGCACCG
   AGGGCCGACC GGCGCGTCGT TGTCTACCTT CCGGAGGACC GCGGCGTGCG
351  GCCCAAGGAG CCCGCGTGTT 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  TGCCCCAAGG ACCGCGCACC TGGTGCATGA CCCGCAAGCC CGGTGCCTGA
   ACGGGCTTCC TGGCGCGTGG ACCACGTA CTGGCGTTCGG GCCACGGACT
```

29/93

FIG. 14

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

```
1  ATGACTGAAT ACAAACCAAC TGTTGCGCTG GCAACTCGTG ATGATGTTCC
   TACTGACTTA TGTTTGTTG ACAAGCGGAC CGTTGAGCAC TACTACAAGG
51  ACGTGCAGTT CGCACCCTGG CTGCTGCATT TGCTGACTAC CCTGCAACCC
   TGCACGTCAA GCGTGGGACC GACGACGTAA ACGACTGATG GGACGTTGGG
101 GTCACACTGT GGACCCAGAC CGCCACATTG AACGTGTGAC TGAAGTGCAG
   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 ATTGGCCAC 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 ACGTACGA CTACACCTTC
551 TGCCAGAAGG CCCACGCACT TGGTGCATGA CTCGCAAACC AGGTGCTTAA
   ACGGTCTTCC GGGTGCCTGA ACCACGTACT GAGCGTTTGG TCCACGAATT
```

30/93

FIG. 15 *S. fradiae* aminoglycoside
phosphotransferase nucleotide sequence

```

1  ATGGACGACA GCACGTTGCG CCGGAAGTAC CCGCACCACG AGTGGCACGC
   TACCTGCTGT CGTGCAACGC GGCCTTCATG GCGTGGGTGC 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 GCGGGCCGAG
   GGCGGGCGCAC CGAGCAGTGC CTCCGGCAGG GGCCGCAGCG CCGCCGGCTC
301 GAGTGGCCCC AGCACCAGCG GTTCGCCGTG GTCGAGGCGA TGGCGGAGCT
   CTCACCGGGC TCGTGGTCGC CAAGCGGCAC CAGCTCCGCT ACCGCCTCGA
351 GGCCCGCGCC CTCCACGAGC TGCCCGTGGA GGAAGTCCCC TCCGACCGGC
   CCGGGCGCGG GAGGTGCTCG ACGGGCACCT CCTGACGGGG AGGCTGGCCG
401 GCCTCGACGC GCGGGTCGCC GAGGCCCGGC GGAACGTTCG CGAGGGCTTG
   CGGAGCTGCG CCGCCAGCGG CTCCGGGCCG CCTTGACAGC GCTCCCGAAC
451 GTGGACCTCG ACGACCTGCA GGAGGAGCGG GCCGGGTGGA CCGGCGACCA
   CACCTGGAGC TGCTGGACGT CCTCCTCGCC CGGCCACCT GGCCGCTGGT
501 GCTCCTGGCG GAGCTCGACC GCACCCGTCC CGAGAAGGAG GACCTGGTCG
   CGAGGACCGC CTCGAGCTGG CGTGGGCAGG GCTCTTCCTC CTGGACCAGC
551 TCTGCCATGG CGACCTGTGC CCCAACAACG TCCTGCTCGA CCGCGGGACC
   AGACGGTACC GCTGGACACG GGGTTGTTGC AGGACGAGCT GGGGCCCTGG
601 TGCCGGGTCA CCGGCGTGAT CGACGTCGGC CGCCTCGGGG TCGCCGACCG
   ACGGCCAGT GGCCGCACTA GCTGCAGCCG GCGGAGCCCC AGCGGCTGGC
651 CCACGCCGAC ATCGCCTTGG CCGCCCGCGA GCTGGAGATC GACGAGGACC
   GGTGCGGCTG TAGCGGAACC GCGGGGCGCT CGACCTCTAG CTGCTCCTGG
701 CCTGGTTCGG CCGCGCCTAC GCCGAGCGGT TCCTGGAGCG GTACGGCGCC
   GGACCAAGCC GGGGCGGATG CGGCTCGCCA AGGACCTCGC CATGCCGCGG
751 CACCGCGTCG ACAAGGAGAA GCTGGCCTTC TACCAGCTTC TCGACGAGTT
   GTGGCGCAGC TGTTCCCTCT CGACCGGAAG ATGGTCGAAG AGCTGCTCAA
801 CTTCTAG
   GAAGATC

```

31/93

FIG. 16 *S. hygroscopicus* hygromycin
phosphotransferase nucleotide sequence

```

1  ATGACACAAG AATCCCTGTT ACTTCTCGAC CGTATTGATT CGGATGATTG
   TACTGTGTTC TTAGGGACAA TGAAGAGCTG GCATAACTAA GCCTACTAAG
51  CTACGCGAGC CTGCGGAACG ACCAGGAATT CTGGGAGCCG CTGGCCCCGCC
   GATGCGCTCG GACGCCTTGC TGGTCCTTAA GACCCTCGGC GACCGGGCGG
101 GAGCCCTGGA GGAGCTCGGG CTGCCGGTGC CGCCGGTGCT GCGGGTGCCC
   CTCGGGACCT CCTCGAGCCC GACGGCCACG GCGGCCACGA CGCCACGGG
151 GGCAGAGACA CCAACCCCGT ACTGGTCGGC GAGCCCGACC CGGTGATCAA
   CCGCTCTCGT GGTGCGGGCA TGACCAGCCG CTCGGGCTGG GCCACTAGTT
201 GCTGTTTCGC GAGCACTGGT GCGGTCCGGA GAGCCTCGCG TCGGAGTCGG
   CGACAAGCCG CTCGTGACCA CGCCAGGCCT CTCGGAGCGC AGCCTCAGCC
251 AGGCGTACGC GGTCTGGCG GACGCCCCGG TGCCGGTGCC CCGCCTCCTC
   TCCGCATGCG CCAGGACCGC CTGCGGGGCC ACGGCCACGG GCGCGAGGAG
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 GCCGCGCCAC
451 CTCGGCCGCG TGCACAGGGT GCCGCTGACC GGAACACCG TGCTCACCCC
   GAGCCGGCCG ACGTGTCCCA CGGCGACTGG CCCTTGTGGC ACGAGTGGGG
501 CCATTCCGAG GTCTTCCCGG AACTGCTGCG GGAACGCCGC GCGGCGACCG
   GGTAAAGGCTC CAGAAGGGCC TTGACGACGC CCTTGCGGCG CGCCGCTGGC
551 TCGAGGACCA CCGCGGGTGG GGCTACCTCT CGCCCCGGCT GCTGGACCGC
   AGCTCCTGGT GCGGCCACCG CCGATGGAGA GCGGGGCCGA CGACCTGGCG
601 CTGGAGGACT GGCTGCCGGA CGTGGACACG CTGCTGGCCG GCCGCGAACC
   GACCTCCTGA CCGACGGCCT GCACCTGTGC GACGACCGGC CGGCGCTTGG
651 CCGGTTCTGC CACGGCGACC TGCACGGGAC CAACATCTTC GTGGACCTGG
   GGCCAAGCAG GTGCCGCTGG ACGTGCCCTG GTTGTAGAAG CACCTGGACC
701 CCGCGACCGA GGTACCCGGG ATCGTCGACT TCACCGACGT CTATGCGGGA
   GCGGCTGGCT CCAGTGGCCC TAGCAGCTGA AGTGGCTGCA GATACGCCCT
751 GACTCCCGCT ACAGCCTGGT GCAACTGCAT CTCAACGCCT TCCGGGGCGA
   CTGAGGGCGA TGTCGGACCA CGTTGACGTA GAGTTGCGGA AGGCCCCGCT
801 CCGCGAGATC CTGGCCGCGC TGCTCGACGG GCGCAGTGG AAGCGGACCG
   GCGGCTCTAG GACCGGCGCG ACGAGCTGCC CCGCGTCACC TTCGCTGGC
851 AGGACTTCGC CCGCGAACTG CTCGCCTTCA CCTTCCTGCA CGACTTCGAG
   TCCTGAAGCG GCGGCTTGAC GAGCGGAAGT GGAAGGACGT GCTGAAGCTC
901 GTGTTTCGAG AGACCCCGCT GGATCTCTCC GGCTTCACCG ATCCGGAGGA
   CACAAGCTCC TCTGGGGCGA CCTAGAGAGG CCGAAGTGGC TAGGCCTCCT
951 ACTGGCGCAG TTCCTCTGGG GGCCGCCGGA CACCGCCCCC GCGGCTGA
   TGACCGCGTC AAGGAGACCC CCGGCGGCCT GTGGCGGGGG CCGCGGACT

```

32/93

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

```

1  ATGAAAAAGC CTGAACTCAC CGCGACGTCT GTCGCGAAGT TTCTGATCGA
   TACTTTTTTCG GACTTGAGTG GCGCTGCAGA CAGCGCTTCA AAGACTAGCT
51  AAAGTTCGAC AGCGTCTCCG ACCTGATGCA GCTCTCGGAG GCGGAAGAAT
   TTTCAAGCTG TCGCAGAGGC TGGACTACGT CGAGAGCCTC CCGCTTCTTA
101 CTCGTGCTTT CAGCTTCGAT GTAGGAGGGC GTGGATATGT CCTGCGGGTA
   GAGCACGAAA GTCGAAGCTA CATCCTCCCG CACCTATACA GGACGCCCAT
151 AATAGCTGCG CCGATGGTTT CTACAAAGAT CGTTATGTTT ATCGGCACTT
   TTATCGACGC GGCTACCAA GATGTTTCTA GCAATACAAA TAGCCGTGAA
201 TGCATCGGCC GCGCTCCCGA TTCCGGGAAGT 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 AGCGGGTTCTG
   CGAGTACCTA CGCTAGCGAC GCCGGCTAGA ATCGGTCTGC TCGCCCAAGC
401 GCCCATTCGG ACCGCAAGGA ATCGGTCAAT AACTACATG 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 GGATTTTCGGC
   CCCGGCTCCT GACGGGGCTT CAGGCCGTGG AGCACGTGCG CCTAAAGCCG
601 TCCAACAATG TCCTGACGGA CAATGGCCGC ATAACAGCGG TCATTGACTG
   AGGTTGTTAC AGGACTGCCT GTTACCGGCG TATTGTCGCC AGTAACTGAC
651 GAGCGAGGCG ATGTTCTGGG 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 GCCGAGGCC GCATATACGA
801 CCGCATTTGGT CTTGACCAAC TCTATCAGAG CTTGGTTGAC GGCAATTTTCG
   GGC GTAACCA GAAGTGGTTG AGATAGTCTC GAACCAACTG CCGTTAAAGC
851 ATGATGCAGC TTGGGCGCAG GGTGCATGCG ACGCAATCGT CCGATCCGGA
   TACTACGTCG AACCCGCGTC CCAGCTACGC TCGGTTAGCA GGCTAGGCCT
901 GCCGGGACTG TCGGGCGTAC ACAAATCGCC CGCAGAAGCG CGGCCGTCTG
   CGGCCCTGAC AGCCCGCATG TGTTTAGCGG GCGTCTTCGC GCCGGCAGAC
951 GACCGATGGC TGTGTAGAAG TACTCGCCGA TAGTGGAAC CGACGCCCCA
   CTGGCTACCG ACACATCTTC ATGAGCGGCT ATCACCTTTG GCTGCGGGGT
1001 GCACTCGTCC GAGGGCAAAG GAATGA
     CGTGAGCAGG CTCCCGTTTC CTTACT

```

33/93

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

```

1  ATGATTGAAC AAGATGGATT GCACGCAGGT TCTCCGGCCG CTTGGGTGGA
   TACTAACTTG TTCTACCTAA CGTGCCTCCA 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 GAATGAAGTG CAAGACGAGG CAGCGCGGCT
   TGGCTGGACA GGCCACGGGA CTTACTTGAC GTTCTGCTCC GTCGCGCCGA
201 ATCGTGGCTG GCCACGACGG GCGTTCCTTG CGCAGCTGTG CTCGACGTTG
   TAGCACCGAC CCGTGCTGCC CGCAAGGAAC GCGTCGACAC GAGCTGCAAC
251 TCACTGAAGC GGGAAGGGAC TGGCTGCTAT TGGGCGAAGT GCCGGGGCAG
   AGTGAAGTTC 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 CATGCCCCGAC GGCGAGGATC
   TCGGCTTGAC AAGCGGTCCG AGTTCGCTC GTACGGGCTG CCGCTCCTAG
551 TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT GGTGAAAAT
   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

```

34/93
FIG. 19 Mouse hnRNP *Hind*III fragment
 nucleotide sequence

```

1  AAGCTTTTAA CCCTCTATCC CTTTAAACTT CCTTGATCCA GTGTAAGCAC
   TTCGAAAATT GGGAGATAGG GAAATTTGAA GGAAGTAGGT CACATTCGTG
51  CTCCTAGAAA GTCAGTAGAC AATAAAACAA AAGTTCTGCT TCACCGATTT
   GAGGATCTTT CAGTCATCTG TTATTTTGTT TTCAAGACGA AGTGGCTAAA
101 ACATTTATAA CCAAATACCC TTCACCAATA CAATAAAAAA ACAAACAAC
   TGTAATATTT GGTATATGGG AAGTGGTTAT GTTATTTTTT TGTTTTGTG
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 TCTTTTGTA TCTGTTGTTA CCTCTCGACT
301 GTAACATTA GTAGTAGTAC TTTACCAGAG AATGGCCTCT ATAGGCTCAC
   CATTGATAAT CATCATCATG AAATGGTCTC TTACCGGAGA TATCCGAGTG
351 ATGTAGGAAT GGTGGGTCCC CAGGTGGTAG GTAGAGCTGT TTGAGGATTA
   TACATCCTTA CCAACCAGGG GTCCACCATC CATCTCGACA AACTCCTAAT
401 CGTGGCCTTC TTGGATGGGG GGTGGGGGTG GGGTGGGAGG GTTGGGTGGT
   GCACCGGAAG AACCTACCCC CCACCCCCAC CCCACCCTCC CAACCCACCA
451 GGGTACTTAA GAGGTTTCAA AAGTCAATAT TGTTTGCAAT 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 TACCGGACGT GGAGAAGTTG
601 TGTAATCTT TCTTCTTTT CTTCTTTTTC TTTTGGTTTC GAGACAGGGT
   ACATTTAGAA AGAAAGAAAA GAAGAAAAAG AAAACCAAAG CTCTGTCCCA
651 TTCTCTGTAT AGTCCTGGCT GTCCTGGAAC TCACCTTGTA GACCAGGCTG
   AAGAGACATA TCAGGACCGA CAGGACCTTG AGTGAAACAT CTGGTCCGAC
701 GCCTTGAACT CAGAAATCCG CCTGCCTCTG CCTCCCTAGC ACTGGGATTA
   CGGAACCTGA GTCTTTAGGC GGACGGAGAC GGAGGGATCG TGACCCTAAT
751 AAGGCGTGCG CCACCACGCC CAGCTTTCAA CTGGAAATCT TAATAAACTT
   TTCCGCACGC GGTGGTGCGG GTCGAAAGTT GACCTTTAGA ATTATTTGAA
801 TCCTAGAAGT GGCTTGGTT ATGGGAGCTT ATCACAGCAA TAGAACAGCA
   AGGATCTTCA CCGGAACCAA TACCCTCGAA TAGTGTGCTT ATCTTGTCGT
851 ATTATGACTG GAGTATGATA GTTAAAAACA AGCAAGCAAG CAAGCAAACA
   TAATACTGAC CTCATACTAT CAATTTTTGT TCGTTCGTTT GTTCGTTTGT
901 CACACACCAA AACAACAAAA CCCCAGACA GAGTCACATG TAGCCCAGGC
   GTGTGTGGTT TTGTTGTTTT GGGGTCTCTG CTCAGTGTA ATCGGGTCCG
951 TAGCCTCCAA ATTCACTATA TAACTGAAGA AGACCCCTAA TTCCCATTC
   ATCGGAGGTT TAAGTGATAT ATTGACTTCT TCTGGGGATT AAGGGTAAGG
1001 TCTAGAATCT ATACCTCAAG TACTGAATGG CTTGGTTCAC AATACCCAC
   AGATCTTAGA TATGGAGTTC ATGACTTACC GAACCAAGTG TTATGGGGTG
1051 TAAATGATTG GTCTTACTAA GTGCAACAAG GTAAACCTAA AACTTCAGCC
   ATTTACTAAC CAGAATGATT CACGTTGTTC CATTTGGATT TTGAAGTCGG
1101 CTCAGACATC CCTTTTCCAG TATCAATTTA TAAAATTAGA TCCAAGGAT
   GAGTCTGTAG GGAAAAGGTC ATAGTTAAAT ATTTTAATCT AGGGTTCCTA
1151 AAAAAATTAAT TGTAAGTAA AATCAGAGT CTAGCATCAA CTACAGGCTC
   TTTTAAATTA ACATTTTATT TTAGTCTCAA GATCGTAGTT GATGTCCGAG
1201 AACCATGGGG ACCACAAATA AACTAAAAGG GATAAGACTG GCTTCCCAT
   TTGGTACCCC TGGTGTTTAT TTGATTTTCC CTATTCTGAC CGAAGGGGTA
1251 AATTATTACA TTTAGATAAT TTTCTGACT ACTCAACAAA GCTAAAATAT
   TTAATAATGT AAATCTATTA AAAGGACTGA TGAGTTGTTT CGATTTTATA
1301 CACCACTGGT TTATTTTCTC CTTCTAGGGT TTAAGCTCAC TCTGAGGAGG
   GTGGTGAQCA AATAAAAGAG GAAGATCCCA AATTCGAGTG AGACTCCTCC

```


35/93

FIG. 19(contd.)

```

1351  GGCATGCGGC ACACACTCAT AGCATCCAGG AAATAGAAAT ATGGTGACTA
      CCGTACGCCG TGTGTGAGTA TCGTAGGTCC TTTATCTTTA TACCACTGAT
1401  TCATGGGTTC AGGGCCAACC TAGGCCTTAG AGAAAAACCT TGTCCCACAA
      AGTACCCAAG TCCCGGTTGG ATCCGAAATC TCTTTTTTGA ACAGGGTGTT
1451  ACCAAAAATG TCTCTTTTTT ATTCTATCAG GGGTGGATGG ATTTGTTAAA
      TGGTTTTTAC AGAGAAAAAA TAAGATAGTC CCCACCTACC TAAACAATTT
1501  GAAGTGCTTT TAAAAACCTT GAGATGGTTA TTTAGAAGTC CCCATGGGAT
      CTTACGAAA ATTTTTGGAA CTCTACCAAT AAATCTTCAG GGGTACCCTA
1551  ACCAAAAATA CCCACTATTT ATATGCCCAA GCATTTCCACC TCCACAACAG
      TGGTTTTTATT GGGTGATAAA TATACGGGTT CGTAAAGTGG AGGTGTTGTC
1601  TGCTATGCAC CCTTTAACAT TTTTGAGACA GTAGCCCAGT CTAGTCTTTA
      ACGATACGTG GGAAATTGTA AAAACTCTGT CATCGGGTCA GATCAGAAAT
1651  ACTTGCAGTG ATTTTTCTCTG ATTCAGCTTC TCCCAGTGCT GGAATTATAG
      TGAACGTCAC TAAAAAGGAC TAAGTCGAAG AGGGTCACGA CCTTAATATC
1701  GTATGCACCA CCATGTGTAA CTACAGATGC TACTTAAAAA TTTTTTAAAG
      CATACGTGGT GGTACACATT GATGTCTACG ATGAATTTTT AAAAAATTTT
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
      GGTCCATGA GTCAGGTAGT CGTAGGTCCC GGTATGGGTG TGAGTGTCTG
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 TGTGACATT
2151  ATTTCACTCT TTTGACATGA ATCTTTTCT AGACAGGGTC TTGGATGCAG
      TAAAGTGAGA AAACTGTACT TAGAAAAAGA TCTGTCCCAG AACCTACGTC
2201  CCCCAGCTAC CCAGAATTTT GGAATCCAGG CTAGCCTCAA ACTCAAGGCA
      GGGGCTGATG GGTCTTAAAA CCTTAGGTCC GATCGGAGTT TGAGTTCCGT
2251  ATCTGCTTGC TTCAGCTTCT CACAGGCTGG ATCACAAACA TACACCTTCA
      TAGACGAACG AAGTCGAAGA GTGTCCGACC TAGTGTTTGT ATGTGGAAGT
2301  GACCCATTTT TTTTTCCTCC CTCCGTTTTT GGTTCCTCTG 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
      AAATAAAAT TATATTGAAC TGTTTTTATT TAAACAGATT GAATGATCTT
2501  ATCCCAAGAA AACTAACACT GGATTTAGCA ACAGTCAGAA ATCGCTGAAA
      TAGGGTTCTT TTGATTGTGA CCTAAATCGT TGTCAGTCTT TAGCGACTTT
2551  AGAAACAGAA TTGATCTAAC AGTCTTAGAT CACTCCTAGA CAGTTTGTA
      TCTTTGTCTT AACTAGATTG TCAGAAATCTA GTGAGGATCT GTCAAACATT
2601  TTCTTGCTCA TGGCAACGTG AGCTCTATCT AACTCACTCT CTGTGCACTA
      AAGAACGAGT ACCGTTGCAC TCGAGATAGA TTGAGTGAGA GACACGTGAT
2651  ATGAATGCTC AGTGTCTCCA GAACAGCACA GCTTCCAGGG TAATCATGCC
      TACTTACGAG TCACAGAGGT CTTGTCTGTG CGAAGGTCCC ATTAGTACGG
2701  AACCACAAAG ACTTTTATAG AGCTGTCCAC GACTCTTCCC CCATTCAGCT
      TTGGGTGTTC TGAAAATATC TCGACAGGTG CTGAGAAGGG GGTAAGTCGA
2751  CATTAACAAT ATGATGGAGC TCCTGTGTGG AAATCAAGGC AACTCTGGT
      GTAATTGTTA TACTACCTCG AGGACACACC TTTAGTTCCG TGTGAGACCA

```

FIG. 19(contd.)

36/93

2801	AGAAACTTGT	TTTTTCTTTC	CACTTTTCCT	TGGGCTCTGA	AGATTGAGCT
	TCTTTGAACA	AAAAAGAAAG	GTGAAAAGGA	ACCCGAGACT	TCTAACTCGA
2851	GTTTTATAAC	CCACAAACAT	GCATTTTFTA	CCTCAAAAGC	ATCCAGCAAA
	CAAAATATTG	GGTGTTTGTA	CGTAAAAAAT	GGAGTTTTCG	TAGGTCGTTT
2901	AACTGTACAA	CGCTTTTTC	AAAAAATGTA	TTGTGATCCT	CCTTAAGAAA
	TTGACATGTT	GCGAAAAAGT	TTTTTTACAT	AACACTAGGA	GGAATTCTTT
2951	AGCCTTACTT	AGTGTTAATT	CCTTTTCTT	TAGAATGCTG	GTAAATACAA
	TCGGAATGAA	TCACAATTAA	GGAAAAAGAA	ATCTTACGAC	CATTTATGTT
3001	GGACTTAGGT	AGGCTGGCTT	CTAACAGCAA	TTCACCCACT	TATGATGGGA
	CCTGAATCCA	TCCGACCGAA	GATTGTCGTT	AAGTGGGTGA	ATACTACCCT
3051	TTAAAGGAAG	GCACAACCAT	GTCCACCACA	GGTTCTAGCT	CCCCACCCA
	AATTTCCTTC	CGTGTGCGTA	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	AAAAAATAAA	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	ATTTTCATGAC	ACACAAAAGG	CAGCTACAAG	TTTTGTGTGG
	CTGTCAATTT	TAAAGTACTG	TGTGTTTTCC	GTCGATGTTT	AAAACACACC
3401	ATTTCAACAT	GTAAATTTTC	GGTAAAAATG	CAGGAAAACA	GTTGAGTTCC
	TAAAGTTGTA	CATTTAAGC	CCATTTTTTAC	GTCCTTTTGT	CAACTCAAGG
3451	CGTGTTATTA	GTATGTTACT	AATAATTTCA	GTATGTTAGT	GAAAATAATC
	GCACAATAAT	CATACAATGA	TTATTAAAGT	CATACAATCA	CTTTTATTAG
3501	TTACTAAAAC	ACTGGTACCT	CAGACAACTT	TACATGGTGA	GGATTGTTAC
	AATGATTTTG	TGACCATGGA	GTCTGTTGAA	ATGTACCACT	CTTAACAATG
3551	TTTCCCAATC	CATATAGAAT	TTTAACAATT	TTAGTGTTTA	TTTGGATGA
	AAAGGGTTAG	GTATATCTTA	AAATTGTTAA	AATCACAAT	AAAACCTACT
3601	AAGGAAATGA	CTATCTTTTG	TTAGCAAATT	ACCATAAGAT	CTTTTCTTTT
	TTCCTTTACT	GATAGAAAAC	AATCGTTTAA	TGGTATTCTA	GAAAAAGAAA
3651	AGATTTCTGA	ATACTCCAAG	GAGCTCATAT	AATTCCATCC	TTATTTTTTTC
	TCTAAAGACT	TATGAGGTTT	CTCGAGTATA	TTAAGGTAGG	AATAAAAAAG
3701	AGAGGCCCTC	CCTGTTCAAT	CACGGTATAA	AAAAAGGAAC	ACATTAAGAT
	TCTCCGGGAG	GGACAAGTTA	GTGCCATATT	TTTTTCCTTG	TGTAATTCTA
3751	GTCCCAGTCC	TATTTTCTGG	CTTTTTTTTT	CCGGGGGTGG	TGGTGCGGTA
	CAGGGTCAGG	ATAAAAGACC	GAAAAAATAA	GGCCCCCACC	ACCACGCCAT
3801	ATCACTCTCT	ATAGTCCAGT	CTGGGCTTCA	ACGCCTGGCA	ATCCCCAGCC
	TAGTGAGAGA	TATCAGGTCA	GACCCGAAGT	TGCGGACCGT	TAGGGGTCGG
3851	TCAAGCTCCC	AAGTACTGTC	CTGATAAGGA	TAGAAGGAGT	CGACCTCCTT
	AGTTTCGAGG	TTCATGACAG	GACTATTCTT	ATCTTCCTCA	GCTGGAGGAA
3901	CACGCTCCCC	TCCGAGGAGG	GCTCCTTCCC	AGCTCCATTC	CCCGGTCGGG
	GTGCGAGGGG	AGGCTCCTCC	CGAGGAAGGG	TCGAGGTAAG	GGGCCAGCCC
3951	AGCCCGTCCT	CCACCCGAGA	GCGCGGGCCT	CGTGGTCAGC	GCCTCCGCGG
	TCGGGCAGGG	GGTGGGCTCT	CGCGCCCGGA	GCACCAAGTC	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	ACTTGCGCAG	AGAGGAGCGC
4151	GGCCTGCGAC	GAGGCTGGGG	GAGGGGAGGC	CCGCGCTTTG	TCTGGGACTCT
	CCGGACGCTG	CTCCGACCCC	CTCCCCCTCC	GGCGCGAAAC	AGACCTCAGA
4201	CGGTAGCTGT	CATCCGGCTC	CCACCCTCAT	GCACAATTGT	CCCATCTCCC
	GCCATCGACA	GTAGGCCGAG	GGTGGGAGTA	CGTGTTAACA	GGGTAGAGGG

FIG. 19(contd.)

37/93

```

4251 CCACGCACCG GCGCGGCGCC CGCCTCAGCG AGGCCCCAGC CGGTTTCCCG
      GGTGCGTGGC CGCGCCGCGG GCGGAGTCGC TCCGGGGTGC GCCAAAGGGC
4301 CAGCCCCGCG CCCACGGGGC TCGCAGCCTC CCCGCAAGCT CGGACGCACG
      GTCGGGCGCC GGGTGCCCCG AGCGTCGGAG GGGCGTTCGA GCCTGCGTGC
4351 GAGCATCCTA AACCCACCA CACGCAAGAT CGAAAAAAG CAAAGGCACG
      CTCGTAGGAT TTGGGGTGGT GTGCGTTCTA GCTTTTTTTC GTTTCGGTGC
4401 AACTTCACCG CTCCGATGCT CAGGGCCGCG GATCCTGCAG AGTCTCCCGC
      TTGAAGTGGC GAGGCTACGA GTCCCGGCGC CTAGGACGTC TCAGAGGGCG
4451 CTGCGCGCTT CGGTTACGCC ACATCCGAGG GGAGGGGGCG CGGGCAGCTC
      GACGCGCGAA GCCAAGTCGG TGTAGGCTCC CCTCCCCGCG GCCCGTCGAG
4501 CGCCGGGGGG GAGGGGGAGC ACCGCCACG CCCTGGCCGC GCGGGGCCCG
      GCGGCCCCC CTCCCCCTCG TGGCGGGTGC GGGACCGGCG CGCCCCGGGC
4551 CCGGGAACGC GTCCTGCGGG GGGCGGCGCG CGCAATGCTC ACCGTCCGCG
      GGGCCTTGCG CAGGACGCCC CCCGCCGCGC GCGTTACGAG TGGCAGGCGC
4601 GCGTGGCGCC CAGGGGGTCT CCTGGCTGGG GGGAGGGGGG GGAAGGCGGG
      CGCACCGCGG GTCCCCCAGA GGACCGACCC CCCTCCCCC CCTTCCGCCC
4651 CAGGAAGGAC CGCGGAGGCC TCTCTGCGTC TCGGAGCGCG CCAAAGCGGG
      GTCCTTCCTG GCGCCTCCGG AGAGACGCAG AGCCTCGCGC GGTTCGCCCC
4701 GCTCCACCCA CCTCCTTGCC CGGATCTTGA AGGCCGGGGA GATAAACAGC
      CGAGGTGGGT GGAGGAACGG GCCTAGAACT TCCGGCCCCCT CTATTTGTGC
4751 GGGGTTCTTT AAGCACCACC TCTCACTAGG CGCGGGATCC CAAGGCTTGT
      CCCCAGAAA TTCGTGGTGG AGAGTGATCC GCGCCCTAGG GTTCCGAACA
4801 GGCATCCGGG GTGGTACTTG GACTAAAAGT CTTTCTGGGA GGGACCGAGT
      CCGTAGGCC CACCATGAAC CTGATTTTCA GGAAGACCCT CCCTGGCTCA
4851 GAGAACCCCT TTGGGACGTG TAGAAATATT TGTGTGGTTC GAGAATATTT
      CTCTTGGGGA AACCCTGCAC ATCTTTATAA ACACACCAAG CTCTTATAAA
4901 GTGCGGACGG GCTTGGCAAA GGCCTAGCTG CAGAGAGCAC GCTTGGGTGG
      CACGCTTGCC CGAACCGTTT CCGCATCGAC GTCTCTCGTG CGAACCCACC
4951 AGAGGGCCGC ACGCCCCAGC GCCGGCCTAA GCCCCTCCCG ACGGCGTTAT
      TCTCCCGGCG TGCGGGGTCT GCGCCGGATT CGGGGAGGGC TGCCGCAATA
5001 TTCAAAGTGC GCGACCGTTT CTCCGCTCCC TACGCGGAGG TGGGGGCCGG
      AAGTTTGACG CGCTGGCAAA GAGGCGAGGG ATGCGCCTCC ACCCCCGGCC
5051 ACCTAGTTCC GGACGTAGTA ACACGCCGAG CGCGAGCCTT CCGCAATTCA
      TGGATCAAGG CCTGCATCAT TGTGCGGCTC GCGCTCGGAA GGCGTAAAGT
5101 CGGAACACAG TTGCGCAAGT GATGTAAAGC AGTCCCGCTG TACCTAAAGG
      GCCTTGTTGTC AACCGTTTCA CTACATTTTCG TCAGGGCGAC ATGGATTTCC
5151 GGGAGTGTCA CGTACTTGGC GTAAGGAGAG TGTAGGCCCT TCCCGCCATT
      CCCTCACAGT GCATGAACCG CATTCCTCTC ACATCCGGGA AGGGCGGTAA
5201 GGC GGCGGTT AGGGCGTTTA CGTAACGGCG TGACGTAAGC GGAGACGCGT
      CCGCCGCCAA TCCCGCAAAT GCATTGCCGC ACTGCATTCT CCTCTGCGCA
5251 TAGTGGGGGG AAGGTTCTAG AAAAGCGGCG GTCTCGGCTC CAGCGGCAGT
      ATCACCCCCC TTCCAAGATC TTTTCGCCGC CAGAGCCGAG GTCGCCGTCA
5301 AGCAGCGGCG CCGGTCCCGT GTGCAGGAGC TCCTTTGCGG CCCAGTTTCT
      TCGTCGCCGC GGCCAGGGCA CACGTCTCTG AGGAAACGCC GGGTCAAAGA
5351 TGGCCATCGC CTGCTCTCCC CACAGCGCCA GGACGAGTCC CGTGCGCGTC
      ACCGGTAGCG GACGAGAGGG GTGTGCGGGT CCTGCTCAGG GCACGCGCAG
5401 CGTCCGCGGA GGTCTTTCTC ATCTCGCTCG GCTGCGGGAA ATCGGGCTGA
      GCAGGCGCCT CCAGAAAGAG TAGAGCGAGC CGACGCCCTT TAGCCCGACT
5451 AGCGACTGAG TCCGCGATGG AGGTAACGGG TTTGAAATCA ATGAGTTATT
      TCGCTGACTC AGGCGCTACC TCCATTGCC AAACTTTAGT TACTCAATAA
5501 AAAAATGGCA TGGCGAGGCC GTAGGCACCG CAATGGAAC CGGCCACCCG
      TTTTACCCT ACCGCTCCGG CATCCGTGGC GTTACCTTTG CCGCGTGGGC
5551 CCTCCGTGGT CCGGCGGAGG GGTGCGGCC ACTCGAGTGG CGGTTGGCCT
      GGAGGCACCA GGCCGCCTCC CCTACGCCGG TGAGCTCACC GCCAACCGBA
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.)

38/93

```

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  TTCCCCCTCC  CCCTCTCCCG  GGAACCGGAT  TTGGCGGCCG
      GCTTCGATCG  AAGGGGAGGG  GGGAGAGGGC  CCTTGCCCTA  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  TTTGAGGTG  GATTTAGCGT
      AAAAGGAAAG  AAAACCTTCA  GAAGTATGGC  AAAGCTCCAC  CTAAATCGCA
6051  TTTGAGCTTG  GGTCTTCAGC  GTCCTGCGCA  CCTCGCTAAA  GGCTCTCTGC
      AAACTCGAAC  CCAGAAAGTC  CAGGACGCGT  GGAGCGATTT  CCGAGAGACG
6101  CTTCCCTCG  ACGAAATGGC  GCCATTGCTT  TCTGAAGCCA  CCGAGGCGCG
      GAAGGGGAGC  TGCTTTACCG  CGGTAACGAA  AGACTTCGGT  GGCTCCGCGC
6151  GGGTGGGGGC  GGGGTGGCGG  CGCTCCACGA  GCTTTACTGG  AACAGGCAGA
      CCCACCCCG  CCCACCGCC  GCGAGGTGCT  CGAAATGACC  TTGTCCGTCT
6201  GAGAACGTAG  TACAACCGAG  GCCTGGGCGG  GTGGCTGAAG  GCAGCGTCGC
      CTCTTGCAATC  ATGTTGGCTC  CGGACCGCGC  CACCGACTTC  CGTCGCAGCG
6251  TGCAAAGAGA  CCGTTTTATT  TTTTATAATA  CGTAAGATTA  CCGGTGCTGT
      ACGTTTCTCT  GGCAAAATAA  AAAGTATTAT  GCATTCTAAT  GCCCACGACA
6301  AGTAAAGCAC  TTGAGCATTA  GTATAGTAGG  AGGAAGTCAA  AGTGGAAGAA
      TCATTTCTGT  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  TTTTAAAGGG  AATCTTCGAA  TTTGTTGGTA
      TTATTAAGTT  TAATATTCTG  AAAAATTCCC  TTAGAAGCTT  AAACAACCAT
6501  AAGTCAACGG  ATCCTTAGCA  CGTGGTGTTT  ACTTTAAGGA  AGTGAAATAG
      TTCAGTTGCC  TAGGAATCGT  GCACCACAAG  TGAAATTCCCT  TCACTTTATC
6551  CTGACTTTTC  ATAGTTAGCG  TTCGCTTAAA  GCCTGGTTCA  GTGGACGAAA
      GACTGAAAAG  TATCAATCGG  AAGCGAATTT  CGGACCAAGT  CACCTGCTTT
6601  ATCCACGTCC  TGGCTATATA  AAAACTTAGT  TTGGGGTCAC  AGTGTGTTGAG
      TAGGTGCGAG  ACCGATATAT  TTTTGAATCA  AACCCAGTG  TCACAACTC
6651  CGTGGTCAAT  CGGTTTTTTT  ATTTTTTATT  TGTTTGAAAT  TATGATGCAT
      GCACCAGTAA  GCCAAAAAAA  TAAAAAATAA  ACAAACTTTA  ATACTACGTA
6701  CATTACACTG  ATAAGCATTA  GCTTTCGAAT  TGAAAGGGGT  CTCCTTGTTT
      GTAATGTGAC  TATTCGTAAT  CGAAAGCTTA  ACTTTCCCCA  GAGGAACCAA
6751  ATTTTCTTTG  ACTCTAAGCA  CACTTATAAA  TAAAATAACC  TTGTTTATAA
      TAAAAGAAAC  TGAGATTCTG  GTGAATATTT  ATTTTATTGG  AACAAATATT
6801  TCGATAGTGG  ACGTCTGGTA  AGTTTGGAAA  AAACCCGAGG  TAAGTAAAGA
      AGCTATCACC  TGCAGACCAT  TCAAACCTTT  TTTGGGCTCC  ATTCATTTCT
6851  GCTTTTGCTT  TCGTTAGTGA  TATGAAAAAA  CAAGGTGTAT  TTAATACTTG
      CGAAAAAGAA  AGCAATCACT  ATACTTTTTT  GTTCCACATA  AATTATGAAC
6901  CAACCTTAGT  TAAGGAAAGC  CAATTTACTG  ACATTTTAGT  AGAGCTACCA
      GTTGAATCAA  ATTCCTTTTC  GTTAAATGAC  TGTAATAATCA  TCTCGATGGT
6951  GAAACACTAT  TTGGAGTCCT  GATTAAGGCT  TTTGTAACCT  TTTTGACTAT
      CTTTGTGATA  AACCTCAGGA  CTAATTCCGA  AAACATTGAT  AAAACTGATA
7001  TTAAAACAAT  TTTGGTCGTT  TTTATTAAAC  ATTTCAAAAC  CTAAAAATTG
      AATTTTGTTA  AAACCAGCAA  AAATAATTGG  TAAAGTTTGG  GATTTTAAAC
7051  TAAACATTGG  CTTTTTGAGC  ACATTTTGGG  GAACTTTACA  AATTTAGGCT
      ATTTGTAAAC  GAAAAACTCG  TGTAATAACCT  CTTTGAATGT  TTAAATCCGA
7101  ATACAGTAAA  ATAACGGAAT  TGTTTTATAA  TTTTGCTTTT  TCATTTCTGT
      TATGTCATTT  TATTGCCTAA  ACAAATATT  AAAACGAAAA  AGTAAAGCAA

```

39/93

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	AATCTTGGA	GTGGTGAATT	TGGAAACAGG	TTTTTTAAAT	TTTTTTTAAA
	TTAGAACCTT	CACCACTTAA	ACCTTTGTCC	AAAAAATTTA	AAAAAATTTT
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	AAAAC TTCTT	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	GTTTTTTTTT	TCCCCCCCAG	AAAAC TTTAG	AAACTGTTCC	TTTGGAGAGG
	CAAAAAAAG	AGGGGGGGTC	TTTTGAAATC	TTTGACAAGG	AAACCTCTCC
7751	AAAAAGGTAC	TCTACCAGCA	GGTCACCTCA	TATTTAAGAA	TTTAATTTCC
	TTTTTCCATG	AGATGGTTCG	CCAGTGGAGT	ATAAATTCTT	AAATTAAGG
7801	TGCATACAAA	GAAAGTGTA	ATAAAAATTG	AAATGGTATT	TCCCTTTGCA
	ACGTATGTTT	CTTTCACATT	TATTTTTAAC	TTTACCATAA	AGGGAAACGT
7851	GAGAGAAAAG	GAACAGTTCC	GAAAGCTCTT	TATTGGTGGC	TTAAGCTT
	CTCTCTTTT	CTTGTCAGAG	CTTTCGAGAA	ATAACCACCG	AATTCGAA

FIG. 19(contd.)

40/93

FIG. 20

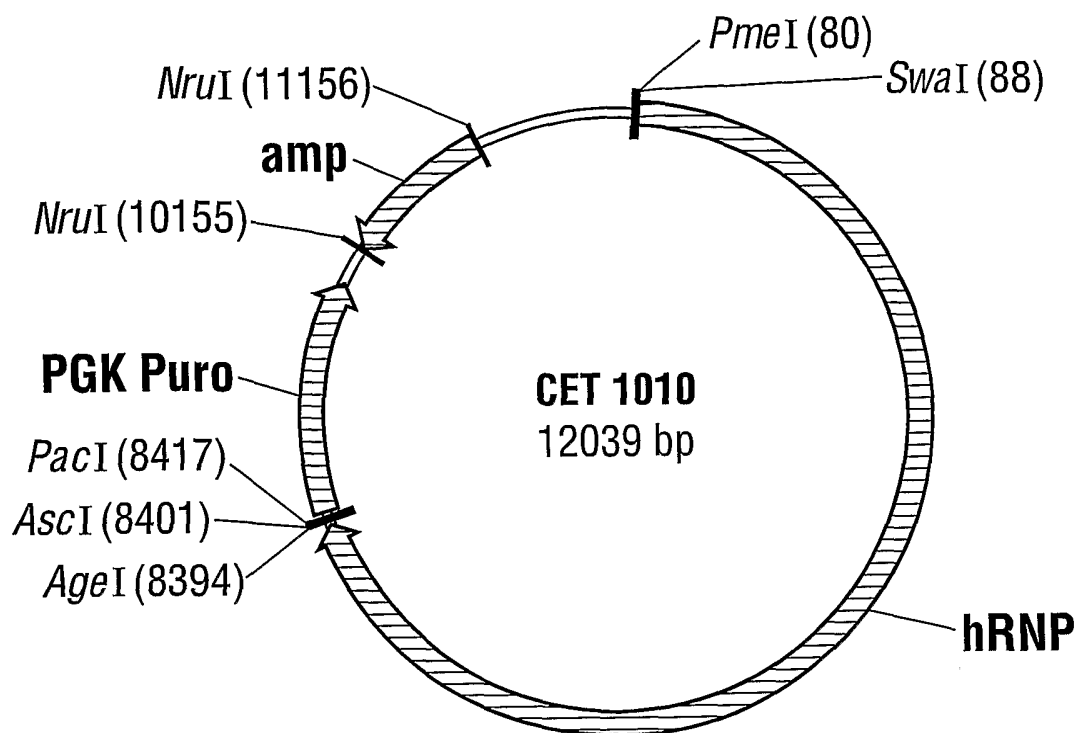
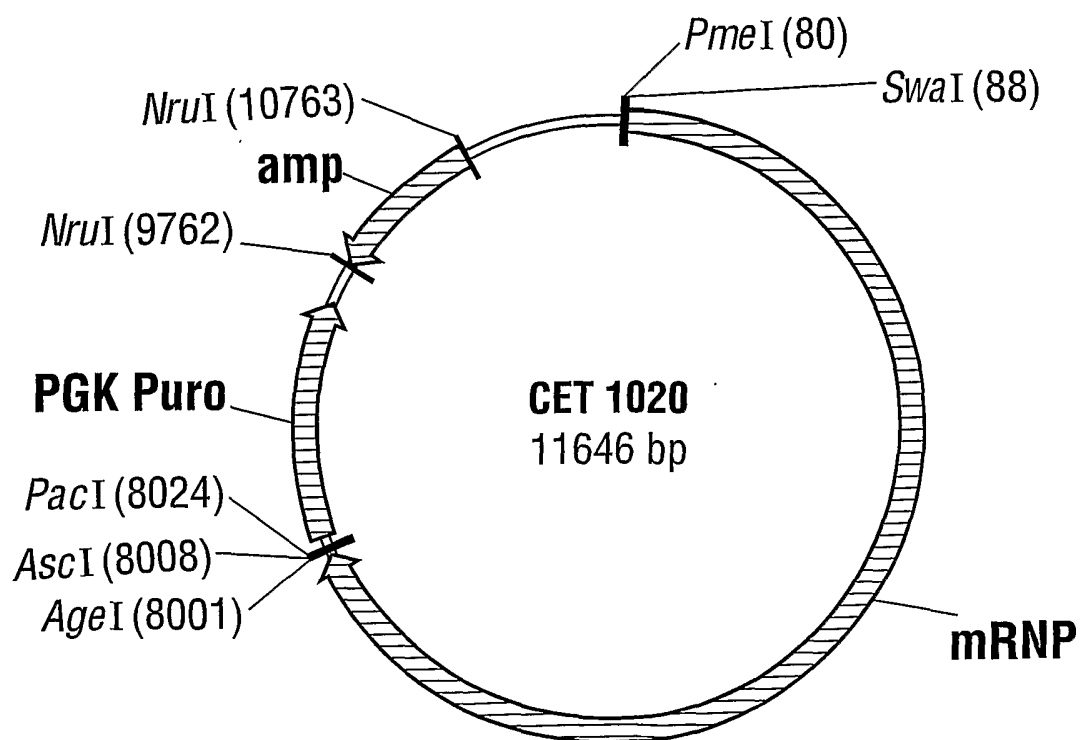


FIG. 22



41/93

FIG. 21 CET1010 nucleotide sequence

1	CGTTGTAAAA	CGACGGCCAG	TGAATTGTAA	TACGACTCAC	TATAGGGCGA
	GCAACATTTT	GCTGCCGGTC	ACTTAACATT	ATGCTGAGTG	ATATCCCGCT
51	ATTGGGTACC	GGGCCCCCCC	TCGAAGTTTA	AACATTTTAA	TCTAGAAGCT
	TAACCCATGG	CCCGGGGGGG	AGCTTCAAAT	TTGTAAATTT	AGATCTTCGA
101	TCAATGTTTT	TAGCACCCCTC	TGTGTGGAGG	AAAATAATGC	AGATTATTCT
	AGTTACAAAA	ATCGTGGGAG	ACACACCTCC	TTTTATTACG	TCTAATAAGA
151	AATTAGTGTA	ATATCTAACC	ACATTAAAAAT	ATATTACATA	GTAAACTACA
	TTAATCACAT	TATAGATTGG	TGTAATTTTA	TATAATGTAT	CATTTGATGT
201	CTCCATAATT	TTATAAATTT	GACTCCCCAG	GGTAATAAAC	TAGTCTCTAG
	GAGGTATTAA	AATATTTTAA	CTGAGGGGTC	CCATTATTTG	ATCAGAGATC
251	TCTGCTCACC	TTCAACTGTA	CAATAAAGTC	TTGGTTCCTT	TGAAATAGAC
	AGACGAGTGG	AAGTTGACAT	GTTATTTTCAG	AACCAAGAAA	ACTTTATCTG
301	CTCAAATGAG	ACACCTAAAA	TTCAAAGTGT	CTTTACATTT	AAAGACACCT
	GAGTTTACTC	TGTGGATTTT	AAGTTTCACA	GAAATGTAAA	TTTCTGTGGA
351	ACAGGAAAGC	AGGTAAAAGA	GCCAGGTTAA	AAACAAATTC	TAAAACCACT
	TGTCCTTTTCG	TCCATTTTCT	CGGTCCAATT	TTTGTTTAAG	ATTTTGGTGA
401	TAGCTGCAGT	TAAACATATA	GTAAAGATGC	ACTAAAGTTT	CTTACTCTGT
	ATCGACGTCA	ATTTGTATAT	CATTTCTACG	TGATTTCAAA	GAATGAGACA
451	AAATCCCTTC	CACTTCAGGA	AATATTCAC	TTTCCCATTTC	ACTACACGTC
	TTTAGGGAAG	GTGAAGTCCT	TTATAAGGTG	AAAGGGTAAG	TGATGTGCAG
501	GATCTAGTAC	TTTTTCCACG	ACAAATTCCT	CAGGCTCTGC	CTCTTCAACT
	CTAGATCATG	AAAAAGGTGC	TGTTTAAGAA	GTCCGAGACG	GAGAAGTTGA
551	TTTTTACTCT	TTCCATTCTG	TTTTTTTCCC	ATTTTTTGCT	AAAATAAAAC
	AAAAATGAGA	AAGGTAAGAC	AAAAAAAGGG	TAAAAAACGA	TTTTATTTTG
601	AAAAGAGAAA	TTAAGAAATA	TTCTCTTGA	ATTTTGAGCA	CATTTTCAAG
	TTTTCTCTTT	AATTCTTTAT	AAGGAGAACT	TAAAACTCGT	GTAAAAGTTC
651	GCTCAATTGC	TTATATTATT	ATCACATTCTG	ACATAAATTT	TTACTTCTAT
	CGAGTTAACG	AATATAATAA	TAGTGTAAGC	TGTATTTAAA	AATGAAGATA
701	ATCCCAGGGC	AGACACCTTC	TGGAAAGATT	AAAAGTCAAC	AGACAATAAA
	TAGGGTCCCG	TCTGTGGAAG	ACCTTTCTAA	TTTTTCAGTTG	TCTGTTATTT
751	ATAAAAGAAT	GCTTTATCTT	GTTCATTTAG	TTCAAACCTTA	CAACCCACCA
	TATTTTCTTA	CGAAATAGAA	CAAGTAAATC	AAGTTTGAAT	GTTGGGTGGT
801	CCAAAATAAT	ACAATAAAAA	AACACTATCT	GGAAACAGTT	ATTTTTTTCC
	GGTTTTATTA	TGTTATTTTTT	TTGTGATAGA	CCTTTGTCAA	TAAAAAAGG
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	TCAAACCTCCT	GACCTGAAGT	GATTCACCAG
	GGTACAACCTG	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	AACAACATATG	CAGGAAGTAC	CTAAAGGGTG	ATCCAGAGAA
	GTGACGTTGT	TTGTTGATAC	GTCTTCATG	GATTTCCAC	TAGGTCTCTT
1251	GCAAGTAGTA	GTGACAGGTC	TTAGGTGAAC	CTATGACAGA	CCTTGATATCC
	CGTTCATCAT	CACTGTCCAG	AATCCACTTG	GATACTGTCT	GGAACATAGG
1301	ACCCCAGAT	GGTAAAAGCC	CCAGCCCCCT	TCTCAATTCA	AATATTAATG
	TGGGGGTCTA	CCATTTTCGG	GGTCGGGGGA	AGAGTTAAGT	TTATAATTAC

FIG. 21(contd.) 42/93

1351	TCAAAAGCAT	CAATGATACA	GAGAAAAGAT	AAATGCAGAA	TGAAAACATG
	AGTTTTTCGTA	GTTACTATGT	CTCTTTTCTA	TTTACGTCTT	ACTTTTGTAC
1401	GTTCAAAATC	CTGATACCAA	CTGCAGGGTC	AACTATAGAG	ACCACTAGGA
	CAAGTTTTTAG	GACTATGGTT	GACGTCCCAG	TTGATATCTC	TGGTGATCCT
1451	GGTTCAATTA	AAGGACAAGA	TTATTTTTTCC	ATAATCTCTG	TAGATAATAT
	CCAAGTTAAT	TTCCTGTTCT	AATAAAAAGG	TATTAGAGAC	ATCTATTATA
1501	TTCCTACCAC	TTAGAACAAA	ACTATAAAGC	TATCACTTCA	AGAGACCAAC
	AAGGATGGTG	AATCTTGTTT	TGATATTTTCG	ATAGTGAAGT	TCTCTGGTTG
1551	ATTACAAATT	TATTTTAATT	CCCTAAGGTG	AAAAAATCC	TTCCTTCCTG
	TAATGTTTAA	ATAAAATTAA	GGGATTCCAC	TTTTTTTtagg	AAGGAAGGAC
1601	GTTTCTCAAG	AGAAAGTCTA	TACTGGTAAC	CAAATTCAC	TTAAACAGGC
	CAAAGAGTTC	TCTTTCAGAT	ATGACCATTG	GTTTAAGTGA	AATTTGTCCG
1651	ATTTTCTTTG	GTATGACACT	ATTTAAGAGA	AGCAGGAAAC	CAACGTGAAC
	TAAAAGAAAC	CATACTGTGA	TAAATTCTCT	TCGTCCTTTG	GTTGCACTTG
1701	CAGCTCTTTC	CAATGGCTCA	AGATTTCCCTA	TGAGAGGACT	AAAAATGGGG
	GTCGAGAAAG	GTTACCGAGT	TCTAAAGGAT	ACTCTCCTGA	TTTTTACCCC
1751	AAAATTTTTTA	TGAGAGGATT	AAAAATGGGG	GAAAAAAAC	CCTGAAATGG
	TTTTTAAAAAT	ACTCTCCTAA	TTTTTACCCC	CTTTTTTTTG	GGACTTTACC
1801	TTAATCAGAA	GATCCTATGG	GCTGAGAAGG	AATCCATCTT	ACATTTTCAT
	AATTAGTCTT	CTAGGATACC	CGACTCTTCC	TTAGGTAGAA	TTGTAAAGTA
1851	CTTAAAGCAA	ATGCTATTGC	CGGGGGCAGT	GGCTCATGCC	TGTAATCCCA
	GAATTTTCGTT	TACGATAACG	GCCCCCGTCA	CCGAGTACGG	ACATTAGGGT
1901	GCACTTTGGG	AGGCCGAGGT	GGGCAGATCA	TCTGAGGTCA	GGAGTTTGAG
	CGTGAAACCC	TCCGGCTCCA	CCCGTCTAGT	AGACTCCAGT	CCTCAAACCTC
1951	ACCAGCCTGA	CCAACATGGA	GAAACCCCGT	TTCTACTAAA	AATACAAAAT
	TGGTCGGACT	GGTTGTACCT	CTTTGGGGCA	AAGATGATTT	TTATGTTTTTA
2001	TAGCCAGGCA	TAGTGGTGCA	TGCCTGTAAT	CCCAGCTACT	TGGGAGGCTG
	ATCGGTCCGT	ATCACCACGT	ACGGACATTA	GGGTCGATGA	ACCCTCCGAC
2051	AGGCAGGAGA	ACTGCTTGAA	CCCAGGAGGC	TTAAGTTGCG	GTGAGCCAAG
	TCCGTCCTCT	TGACGAACTT	GGGTCTCCG	AATTCAACGC	CACTCGGTTC
2101	ATCACGCCAT	TGCACTCTAG	CCTGGACAAC	AAGAGAAAAA	CTCTGTCTCA
	TAGTGCGGTA	ACGTGAGATC	GGACCTGTTG	TTCTCTTTTT	GAGACAGAGT
2151	AAAAAACACA	AAAACAAAAA	ACCCAAATAC	TATTTAAAAA	AGATAAACCT
	TTTTTTTGTGT	TTTTGTTTTT	TGGGTTTATG	ATAAATTTTT	TCTATTTGGA
2201	TAATTGCTCA	ATCATTAAAG	CCATCCCACA	AGTAAAGCAG	CAAGCAGAAA
	ATTAACGAGT	TAGTAATTTT	GGTAGGGTGT	TCATTTTCGT	GTTCTGCTTT
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	TTACATAGT
2601	GATACCTATA	CTACATGAAA	AACAAAACAG	CTACTGCCAC	AAAGGGAGAA
	CTATGGATAT	GATGTACTTT	TTGTTTTGTC	GATGACGGTG	TTTCCCTCTT
2651	GGCTAACAAA	ATAAAGTCAA	CAATAAATAC	AGAAAATGAA	AAGGATACAC
	CCGATTGTTT	TATTTTCAGTT	GTTATTTATG	TCTTTTACTT	TTCTATGTG
2701	ACTAAGGTTT	ACAAAAAAA	AAAGGCAGAC	AAAATGCCAT	ACAGTATTCA
	TGATTCCAAA	TGTTTTTTTT	TTTCCGCTCTG	TTTTACGGTA	TGTCATAAGT
2751	TTCACTACTA	TGGCATTCAT	AAGCTAGTTT	CAAATGCTCA	CTATTTTCTT
	AAGTGATGAT	ACCGTAAGTA	TTCGATCAAA	GTTTACGAGT	GATAAAAGAA

FIG. 21(contd.) 43/93

2801	TTATAGTATA	TATTTGCCTT	AACCCAGCAC	TTTTTTCCAA	AAGTGGATGA
	AATATCATAT	ATAAACGGAA	TTGGGTCGTG	AAAAAAGGTT	TTCACCTACT
2851	GTCAAAATAA	ATTTCCCATT	ATTTAAGTGA	AATTAACAGC	ACACATATCT
	CAGTTTTTATT	TAAAGGGTAA	TAAATTCACT	TTAATTGTCTG	TGTGTATAGA
2901	CACAACACTA	ATGAATTTTT	AAAATGGAAA	GTTAAGAAGT	TTTAAAGTGG
	GTGTTGTGAT	TACTTAAAAA	TTTTACCTTT	CAATTCCTGA	AAATTTTACC
2951	CCAACCTGTG	ATCCTTCACA	AAATAAACTA	AATACAATAA	CAGACCCCAA
	GGTTGGACAC	TAGGAAGTGT	TTTATTTGAT	TTATGTTATT	GTCTGGGGTT
3001	AGGCTATCAA	TTGCGTGCAA	AAACAACCTC	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	GTATGTTCCG	TTTTTGTTTA	CGAAATAAAA	TGGATCCTCT
3201	CATTAACATT	CACCTTTACG	TGTTTAAGAT	TAATGCAATG	TTAAATATTG
	GTAATTGTAA	GTGGAAATGC	ACAAATTCCTA	ATTACGTTAC	AATTTATAAC
3251	TGAAAACGTG	AACTTTGAAT	TTCATGATTT	TTATGTGAAT	ATTCCAGGGT
	ACTTTTGACA	TTGAAACTTA	AAGTACTAAA	AATACACTTA	TAAGGTCCCA
3301	TTAAAAAAAC	TTGTAACATG	ACATGGCTGA	ATAAGATAAA	AAAAAAATCT
	AATTTTTTTG	AACATTGTAT	TGTACCGACT	TATTCTATTT	TTTTTTTAGA
3351	AGCCTTTTCT	CCCTTCTGGC	TCATATTTGC	GATTTTCGATC	ATTTTGTTTA
	TCGGAAGAGA	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	CTTTTTTTTT	AACTCAATTT
3901	TCCTTACATA	TTTTAGGTTT	TCATTAACTT	ACCAATGTAG	TTTTGTTGGA
	AGGAATGTAT	AAAATCCAAA	AGTAATTGAA	TGGTTACATC	AAAACAACCT
3951	GGCCATTTTT	TTTATTGCAG	ACTTGAAGAG	CTATTACTAG	AAAAATGCAT
	CCGGTAAAAA	AAATAACGTC	TGAACCTCTC	GATAATGATC	TTTTTACGTA
4001	GACAGTTAAG	GTAAGTTTGC	ATGACACAAA	AAAGGTAAC	AAATACAAAT
	CTGTCAATTC	CATTCAAACG	TACTGTGTTT	TTTCCATTGA	TTTATGTTTA
4051	TCTGTTTGGA	TTCCAACCCC	CAAGTAGAGA	GCGCACACTT	TCAAACGTGA
	AGACAAACCT	AAGGTTGGGG	GTTCATCTCT	CGCGTGTGAA	AGTTTGCAC
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	TTTTGTGCGA	GTGTACCGAA	TTTTATCCCC	TGGTTACAGA

FIG. 21(contd.) 44/93

4251	TTTCCAATCT	AAGTCCCATT	TATAATAAAG	TCCATGTTCC	ATTTTAAAG
	AAAGGTTAGA	TTCAGGGTAA	ATATTATTTC	AGGTACAAGG	TAAAAATTTC
4301	GACAATCCTT	TCGGTTTAAA	ACCAGGCACG	ATTACCCAAA	CAACTCACAA
	CTGTTAGGAA	AGCCAAATTT	TGGTCCGTGC	TAATGGGTTT	GTTGAGTGTT
4351	CGGTAAAGCA	CTGTGAATCT	TCTCTGTTCT	GCAATCCCAA	CTTGGTTTCT
	GCCATTTTCGT	GACACTTAGA	AGAGACAAGA	CGTTAGGGTT	GAACCAAAGA
4401	GCTCAGAAAC	CCTCCCTCTT	TCCAATCGGT	AATTAAATAA	CAAAAGGAAA
	CGAGTCTTTG	GGAGGGAGAA	AGGTTAGCCA	TTAATTTTAT	GTTTTCTTTT
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
	CTCCGGATTCT	GCGCCACGCG	GCGGGGGTGG	GCCCTCGCGC	CCGGAGCACC
4601	TCAGCGCATC	GCGGGGAGAA	AACAAAGGCC	GCGGCACGGG	GGCTCAAGGG
	AGTCGCGTAG	GCGCCCCCTCT	TTGTTTCCGG	CGCCGTGCC	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	GTTGAGAGCT	GGGTCGGGGC
4801	CGTGGCGCTG	GGGGAGGGGG	CGCCTCCGCC	GGAACGCGGG	TGGGGGAGGG
	GCACCGCGAC	CCCCTCCCCC	GCGGAGGCGG	CCTTGCGCCC	ACCCCTCCC
4851	GAGGGGGAAA	TGCGCTTTGT	CTCGAAATGG	GGCAACCGTC	GCCACAGCTC
	CTCCCCCTTT	ACGCGAAACA	GAGCTTTACC	CCGTTGGCAG	CGGTGTCGAG
4901	CCTACCCCTT	CGAGGGCAGA	GCAGTCCCCC	CACTAACTAC	CGGGCTGGCC
	GGATGGGGGA	GCTCCCGTCT	CGTCAGGGGG	GTGATTGATG	GCCCCACCGG
4951	GCGCGCCAGG	CCAGCCGCGA	GGCCACCGCC	CGACCCTCCA	CTCCTTCCCG
	CGCGCGGTCC	GGTCGGCGCT	CCGGTGGCAG	GCTGGGAGGT	GAGGAAGGGC
5001	CAGCTCCCGG	CGCGGGGTCC	GGCGAGAAGG	GGAGGGGAGG	GGAGCGGAGA
	GTGAGGGGCC	GCGCCCCAGG	CCGCTCTTCC	CCTCCCCCTC	CCTCGCCTCT
5051	ACCGGGCCCC	CGGGACGCGT	GTGGCATCTG	AAGCACCACC	AGCGAGCGAG
	TGGCCCAGGG	GCCCTGCGCA	CACCGTAGAC	TTGCTGGTGG	TCGCTCGCTC
5101	AGCTAGAGAG	AAGGAAAGCC	ACCGACTTCA	CCGCCTCCGA	GCTGCTCCGG
	TCGATCTCTC	TTCCTTTCCG	TGGCTGAAGT	GGCGGAGGCT	CGACGAGGCC
5151	GTCGCGGGTC	TGCAGCGTCT	CCGGCCCTCC	GCGCCTACAG	CTCAAGCCAC
	CAGCGCCCA	ACGTGCGAGA	GGCCGGGAGG	CGCGGATGTC	GAGTTCCGGT
5201	ATCCGAAGGG	GGAGGGAGCC	GGGAGCTGCG	CGCGGGGCCG	CCGGGGGGAG
	TAGGCTTCCC	CCTCCCTCGG	CCCTCGACGC	GCGCCCCGGC	GGCCCCCTC
5251	GGGTGGCACC	GCCCACGCCG	GGCGGCCACG	AAGGGCGGGG	CAGCGGGCGC
	CCCACCGTGG	CGGGTGCGGC	CCGCCGGTGC	TTCCCGCCCC	GTGCGCCGCG
5301	GCGCGCGGCG	GGGGGAGGGG	CCGGCGCCGC	GCCCCTGGG	AATTGGGGCC
	CGCGCGCCGC	CCCCCTCCCC	GGCCGCGGCG	CGGGCGACCC	TTAACCCCGG
5351	CTAGGGGGAG	GGCGGAGGCG	CCGACGACCG	CGGCACTTAC	CGTTGCGGGC
	GATCCCCCTC	CCGCCCTCCG	GGCTGCTGGC	GCCGTGAATG	GCAAGCGCCG
5401	GTGGCGCCCG	GTGGTCCCCA	AGGGGAGGGA	AGGGGGAGGC	GGGGCGAGGA
	CACCGCGGGC	CACCGAGGGT	TCCCTCCCTT	TCCCTCCCTC	CCCCGCTCCT
5451	CAGTGACCGG	AGTCTCCTCA	GCGGTGGCTT	TTCTGCTTGG	CAGCCTCAGC
	GTCACTGGCC	TCAGAGGAGT	CGCCACCGAA	AAGACGAACC	GTCGGAGTCG
5501	GGCTGGCGCC	AAAACCGGAC	TCCGCCCACT	TCCTCGCCCC	CCGGTGCGAG
	CCGACCGCGG	TTTTGGCCTG	AGCGGGGTGA	AGGAGCGGGC	GGCCACGCTC
5551	GGTGTGGAAT	CCTCCAGACG	CTGGGGGAGG	GGGAGTTGGG	AGCTTAAAAA
	CCACACCTTA	GGAGGTCTGC	GACCCCCCTC	CCCTCAACCC	TCGAATTTTT
5601	CTAGTACCC	TTTGGGACCA	CTTTCAGCAG	CGAACTCTCC	TGTACACCAG
	GATCATGGGG	AAACCCTGGT	GAAAGTCGTC	GCTTGAGAGG	ACATGTGGTC
5651	GGGTCAAGTC	CACAGACGCG	GGCCAGGGGT	GGGTCAATTG	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 AGGGCCCCAAC
5851  AGGGGGCTGG ATCTCCAGGC TCGCGATTAA GCCCTCCCCG TCGGCGTTAA
      TCCCCCGACC TAGAGGTCCG ACGCCTAATT CGGGGAGGGC AGCCGCAATT
5901  TTTCAAACCTG CGCGACGTTT CTCACCTGCC TTCGCCAAGG CAGGGGCCGG
      AAAGTTTGAC GCGCTGCAAA GAGTGGACGG AAGCGGTTCC GTCCCCGGCC
5951  GACCCTATTC CAAGAGGTAG TAAC TAGCAG GACTCTAGCC TTCCGCAATT
      CTGGGATAAG GTTCTCCATC ATTGATCGTC CTGAGATCGG AAGGCGTTAA
6001  CATTGAGCGC ATTTACGGAA GTAACGTCGG GTACTGTCTC TGGCCGCAAG
      GTAACGCGC TAAATGCCTT CATTGCAGCC CATGACAGAG ACCGGCGTTC
6051  GGTGGGAGGA GTACGCATTT GCGTAAGGT GGGCGTAGA CCCTTCCCGC
      CCACCTCCT CATGCGTAAA CCGCATTCCA CCCC GCATCT CGGAAGGGCG
6101  CATTGGCGGC GGATAGGGCG TTTACGCGAC GGCTGACGT AGCGGAAGAC
      GTAACGCGC CCTATCCCGC AAATGCGCTG CCGGACTGCA TCGCCTTCTG
6151  GCGTTAGTGG GGGGGAAGGT TCTAGAAAAG CGGCGGCAGC GGCTCTAGCG
      CGCAATCACC CCCCTTCCA 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  GTGTTCTGTC GCGGAGATCT CTCTCATCTC GCTCGGCTGC GGGAAATCGG
      CACAAGCAGG CGCCTCTAGA GAGAGTAGAG CGAGCCGACG CCCTTTAGCC
6351  GCTGAAGCGA CTGAGTCCGC GATGGAGGTA ACGGGTTTGA AATCAATGAG
      CGACTTCGCT GACTCAGGCG CTACCTCCAT TGCCCAAAC TTAGTTACTC
6401  TTATTGAAAA GGGCATGGCG AGGCCGTTGG CGCCTCAGTG GAAGTCGGCC
      AATAACTTTT CCCGTACCGC TCCGGAACCG GCGGAGTCAC CTTCAGCCGG
6451  AGCCGCCTCC GTGGGAGAGA GGCAGGAAAT CGGACCAATT CAGTAGCAGT
      TCGGCGGAGG CACCCTCTCT CCGTCTTTTA GCCTGGTTAA GTCATCGTCA
6501  GGGGCTTAAG GTTTATGAAC GGGGTCTTGA GCGGAGGCCT GAGCGTACAA
      CCCC GAATTC 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 CCCC ACTTTC
6651  TGGGGCGCGG AGGCGGGAAT TCTTATTTCC TTTCTAAAGC ACGCTGCTTC
      ACCCGCGGCC TCCGCCCTTA AGAATAAGGG AAAGATTTTCG TGCGACGAAG
6701  GGGGGCCACG GCGTCTCCTC GGCGAGCGTT TCGGCGGGCA GCAGGTCTCTC
      CCCC CGGTGC CGCAGAGGAG CCGCTCGCAA AGCCGCCCGT CGTCCAGGAG
6751  GTGAGCGAGG CTGCGGAGCT TCCCCTCCCC CTCTCTCCCC GGAACCGATT
      CACTCGCTCC GACGCCTCGA AGGGGAGGGG GAGAGAGGGC CCTTGGCTAA
6801  TGGCGGCCGC CATTTTCATG GCTCGCCTTC CTCTCAGCGT TTTCTTATA
      ACCGCCGGCG GTAAAAGTAC CGAGCGGAAG GAGAGTCGCA AAAGGAATAT
6851  ACTCTTTTAT TTTCTTAGTG TGCTTTCTCT ATCAAGAAGT AGAAGTGGTT
      TGAGAAAATA AAAGAATCAC ACGAAAGAGA TAGTTCTTCA TCTTCACCAA
6901  AACTATTTTT TTTTCTTCTT CGGGCTGTTT TCATATCGTT TCGAGGTGGA
      TTGATAAAAA AAAAGAAGA GCCCGACAAA AGTATAGCAA AGCTCCACCT
6951  TTTGGAGTGT TTTGTGAGCT TGGATCTTTA GAGTCCTGCG CACCTCATTA
      AAACCTCA CA AAACACTCGA ACCTAGAAAT CTCAGGACGC GTTGAGTAAT
7001  AAGGCGCTCA GCCTTCCCCT CGATGAAATG GCGCCATTGC GTTCGGAAGC
      TTCCGCGAGT CGGAAGGGGA GCTACTTTAC CGCGGTATACG CAAGCCTTCG
7051  CACACCGAAG AGCGGGGAGG GGGGGTGCTC CGGGTTTGCG GGCCCGGTTT
      GTGTGGCTTC TCGCCCCTCC CCCCACGAG GCCCAAACGC CCGGGCCAAA
7101  CAGAGAAGAT ATCACCACCC AGGGCGTCGG GCCGGGTTCA ATGCGAGCCG
      GTCTCTTCTA TAGTGGTGGG TCCCGCAGCC CGGCCCAAGT TACGCTCGGC

```

FIG. 21(contd.) 46/93

```

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 TGAGTAAAT TCATTACAAC
7401 TTAAGTGAAG GTTTAACCTT TACGGGACAG ATCCATGGAC CTTTATAGAT
    AATGACTCTT CAAATTGGGA ATGCCCTGTC TAGGTACCTG GAAATATCTA
7451 GATTACGAGG AAAGTGAAAT AACGATTTTG TCCTTAGTTA TACTTCGATT
    CTAATGCTCC TTTCACCTTA TTGCTAAAAC AGGAATCAAT ATGAAGCTAA
7501 AAAACATGGC TTCAGAGGCT CCTTCCTGTA ATGCGTATGG ATTGATGTGC
    TTTTGTACCG AAGTCTCCGA GGAAGGACAT TACGCATACC TAACTACACG
7551 AAAACTGTTT TGGGCCTGGG CCGCTCTGTA TTTGAACTTT GTTACTTTTC
    TTTTGACAAA ACCCGGACCC GGCGAGACAT AAACCTTGAA CAATGAAAAG
7601 TCATTTTGTT TGCAATCTTG GTTGAACATT ACATTGATAA GCATAAGGTC
    AGTAAAACAA ACGTTAGAAC CAACTTGTA TGTAACTATT CGTATTCCAG
7651 TCAAGCGAAG GGGGTCTACC TGGTTATTTT TCTTTGACCC TAAGCACGTT
    AGTTCGCTTC CCCAGATGG ACCAATAAAA AGAACTGGG ATTCGTGCAA
7701 TATAAAATAA CATTGTTTAA AATCGATAGT GGACATCGGG TAAGTTTGGA
    ATATTTTATT GTAACAAATT TTAGCTATCA CCTGTAGCCC ATTCAAACCT
7751 TAAATTGTGA GGTAAGTAAT GAGTTTTTGC TTTTGTAG TAGATTGTAA
    ATTTAACACT CCATTCATTA CTCAAAACG AAAACAATC ACTAAACATT
7801 AACTTGTTAT AAATGTACAT TATCCGTAAT TTCAGTTTAG AGATAACCTA
    TTGAACAATA TTTACATGTA ATAGGCATTA AAGTCAAATC TCTATTGGAT
7851 TGTGCTGACG ACAATTAAGA ATAAAACTA GCTGAAAAAA TGAAAATAAC
    ACACGACTGC TGTTAATTCT TATTTTTGAT CGACTTTTTT ACTTTTATG
7901 TATCGTGACA AGTAACCATT TCAAAAGACT GCTTTGTGTC TCATAGGAGC
    ATAGCACTGT TCATTGGTAA AGTTTTCTGA CGAAACACAG AGTATCCTCG
7951 TAGTTTGATC ATTTTCAGTTA ATTTTTTCTT TAATTTTTTAC GAGTCATGAA
    ATCAAACTAG TAAAGTCAAT TAAAAAAGAA ATTAATAATG CTCAGTACTT
8001 AACTACAGGA AAAAAAATCT GAACTGGGTT TTACCACTAC TTTTATAGGAG
    TTGATGTCCT TTTTTTTAGA CTTGACCCAA AATGGTGATG AAAAATCCTC
8051 TTGGGAGCAT GCGAATGGAG GGAGAGCTCC GTAGAACTGG GATGAGAGCA
    AACCTCGTA CGCTTACCTC CCTCTCGAGG CATCTTGACC CTACTCTCGT
8101 GCAATTAATG CTGCTTGCTA GGAACAAAAA ATAATTGATT GAAAATTACG
    CGTTAATTAC GACGAACGAT CCTTGTTTTT TATTAATACTA CTTTAAATGC
8151 TGTGACTTTT TAGTTTGCAT TATGCGTTTG TAGCAGTTGG TCCTGGATAT
    AACTGAAAA ATCAAACGTA ATACGCAAAC ATCGTCAACC AGGACCTATA
8201 CACTTTCTCT CGTTTGAGGT TTTTAAACCT AGTTAACTTT TAAGACAGGT
    GTGAAAGAGA GCAAACCTCA AAAAATTGGA TCAATTGAAA ATTCTGTCCA
8251 TTCCTTAACA TTCATAAGTG CCCAGAATAC AGCTGTGTAG TACAGCATAT
    AAGGAATTGT AAGTATTCAC GGGTCTTATG TCGACACATC ATGTCGTATA
8301 AAAGATTTCA GCTCTGAGGT TTTTCTTATT GACTTGGAAT 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 CGTTTTCCC AAGGCAGTCT GGAGCATGCG
    TCCGATGGCC CATCCCCTCC GCGAAAAGGG TTCCGTGAGA CCTCTGACGC
8501 CTTTAGCAGC CCCGCTGGGC ACTTGCGGCT ACACAAGTGG CCTCTGGCCT
    GAAATCGTGG GGGCGACCCG TGAACCGCGA TGTGTTTACC GGAGACCGGA
8551 CGCACACATT CCACATCCAC CGGCCGGTAG GCGCCAACCG GCTCCGTTCT
    GCGTGTGTAA GGTGTAGGTG GCCGGCCATC CGCGGTTGGC CGAGGCAAGA

```

FIG. 21(contd.) 47/93

8601	TTGGTGGCCC	CTTCGCGCCA	CCTTCTACTC	CTCCCCTAGT	CAGGAAGTTC
	AACCACCGGG	GAAGCGCGGT	GGAAGATGAG	GAGGGGATCA	GTCTTCAAG
8651	CCCCCGCCC	CGCAGCTCGC	GTCTGTCAGG	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	TTCTCATCT
	CGTGCGAAGT	TTTCGCGTGC	AGACGGCGCG	ACAAGAGGAG	AAGGAGTAGA
8951	CCGGGCCTTT	CGACCAGCTT	ACCATGACCG	AGTACAAGCC	CACGGTGCGC
	GGCCCCGAAA	GCTGGTCGAA	TGGTACTGGC	TCATGTTCCG	GTGCCACGCG
9001	CTCGCCACCC	GCGACGACGT	CCCCAGGGCC	GTACGCACCC	TCGCCGCCGC
	GAGCGGTGGG	CGCTGCTGCA	GGGGTCCCGG	CATGCGTGGG	AGCGGCGGGC
9051	GTTCCGCCGAC	TACCCCGCCA	CGCGCCACAC	CGTCGATCCG	GACCGCCACA
	CAAGCGGCTG	ATGGGGCGGT	GCGCGGTGTG	GCAGCTAGGC	CTGGCGGTGT
9101	TCGAGCGGGT	CACCGAGCTG	CAAGAACTCT	TCCTCACGCG	CGTCGGGCTC
	AGCTCGCCCA	GTGGCTCGAC	GTTCTTGAGA	AGGAGTGCGC	GCAGCCCGAG
9151	GACATCGGCA	AGGTGTGGGT	GCGCGACGAC	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	GAAGGCCCTCC	TGGCGCCGCA	CCGGCCCAAG	GAGCCCGCGT	GGTTCCTGGC
	CTTCCGGAGG	ACCGCGGCGT	GGCCGGGTTC	CTCGGGCGCA	CCAAGGACCG
9351	CACCGTCGGC	GTCTCGCCCG	ACCACCAGGG	CAAGGGTCTG	GGCAGCGCCG
	GTGGCAGCCG	CAGAGCGGGC	TGGTGGTCCC	GTTCACAGAC	CCGTGCGGGC
9401	TCGTGCTCCC	CGGAGTGAG	GCGGCCGAGC	GCGCCGGGGT	GCCCGCCTTC
	AGCACGAGGG	GCCTCACCTC	CGCCGGCTCG	CGCGGCCCCA	CGGGCGGAAG
9451	CTGGAGACCT	CCGCGCCCCG	CAACCTCCCC	TTCTACGAGC	GGCTCGGCTT
	GACCTCTGGA	GGCGCGGGGC	GTTGGAGGGG	AAGATGCTCG	CCGAGCCGAA
9501	CACCGTCACC	GCCGACGTCG	AGGTGCCCCA	AGGACCGCGC	ACCTGGTGCA
	GTGGCAGTGG	CGGCTGCAGC	TCCACGGGCT	TCCTGGCGCG	TGGACCACGT
9551	TGACCCGCAA	GCCCCGTGCC	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
	CACCCACCCC	CGTCCTGTGC	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	TAATTTTCAG	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
	CGTATTTTCAC	ATTTCCGACC	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	TCCGTGTGCG	CCTTATTCCC	TTTTTTGCGG	CATTTTGCCT	TCCTGTTTTT
	AGGCACAGCG	GGAATAAGGG	AAAAAACGCC	GTAAAAACGA	AGGACAAAAA
10351	GCTCACCCAG	AAACGCTGGT	GAAAGTAAAA	GATGCTGAAG	ATCAGTTGGG
	CGAGTGGGTC	TTTGCGACCA	CTTTCATTTT	CTACGACTTC	TAGTCAACCC
10401	TGCACGAGTG	GTTTACATCG	AACTGGATCT	CAACAGCGGT	AAGATCCTTG
	ACGTGCTCAC	CCAATGTAGC	TTGACCTAGA	GTTGTGCGCA	TTCTAGGAAC
10451	AGAGTTTTTCG	CCCCGAAGAA	CGTTTTTCAA	TGATGAGCAC	TTTTAAAGTT
	TCTCAAAAGC	GGGGCTTCTT	GCAAAAGGTT	ACTACTCGTG	AAAATTTCAA
10501	CTGCTATGTG	GCGCGGTATT	ATCCCGTATT	GACGCCGGGC	AAGAGCAACT
	GACGATACAC	CGCGCCATAA	TAGGGCATAA	CTGCGGCCCG	TTCTCGTTGA
10551	CGGTGCGCGC	ATACACTATT	CTCAGAATGA	CTTGTTGAG	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	TACCCCTAG
10751	ATGTAAGTCG	CCTTGATCGT	TGGGAACCGG	AGCTGAATGA	AGCCATACCA
	TACATTGAGC	GGAAGTAGCA	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	TAATTGTAAC	CCATTGACAG
11151	AGACTCGCGA	CACTGCATTA	ATGAATCGGC	CAACGCGCGG	GGAGAGGCGG
	TCTGAGCGCT	GTGACGTAAT	TACTTAGCCG	GTTGCGCGCC	CCTCTCCGCC
11201	TTTGCGTATT	GGGCGCTCTT	CCGCTTCCTC	GCTCACTGAC	TCGCTGCGCT
	AAACGCATAA	CCCGCGAGAA	GGCGAAGGAG	CGAGTGACTG	AGCGACGCGA
11251	CGGTGCTTCG	GCTGCGGCGA	GCGGTATCAG	CTCACTCAA	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	TGGCATTTTTT	CCGGCGCAAC	GACCGCAAAA
11401	TCCATAGGCT	CCGCCCCCCT	GACGAGCATC	ACAAAAATCG	ACGCTCAAGT
	AGGTATCCGA	GGCGGGGGGA	CTGCTCGTAG	TGTTTTTAGC	TGCGAGTTCA
11451	CAGAGGTGQC	GAAACCCGAC	AGGACTATAA	AGATACCAGG	CGTTTCCCCC
	GTCTCCACCG	CTTTGGGCTG	TCCTGATATT	TCTATGGTCC	GCAAAGGGGG

49/93

```
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
      CCGTGACCAT TGTCCTAATC GTCTCGCTCC ATACATCCGC CACGATGTCT
11801  GTTCTTGAAG TGGTGGCCTA ACTACGGCTA CACTAGAAGG ACAGTATTTG
      CAAGAACTTC ACCACCGGAT TGATGCCGAT GTGATCTTCC TGTCATAAAC
11851  GTATCTGCGC TCTGCTGAAG CCAGTTACCT TCGGAAAAAG AGTTGGTAGC
      CATAGACGCG AGACGACTTC GGTCAATGGA AGCCTTTTTT TCAACCATCG
11901  TCTTGATCCG GCAAACAAAC CACCGCTGGT AGCGGTGGTT TTTTGTGTTG
      AGAACTAGGC CGTTTGTTTG GTGGCGACCA TCGCCACCAA AAAAACAAC
11951  CAAGCAGCAG ATTACGCGCA GAAAAAAGG ATCTCAAGAA GATCCTTTGA
      GTTCGTGCTC TAATGCGCGT CTTTTTTTCC TAGAGTTCTT CTAGGAAACT
12001  TCTTTTCTAC GGGGTCTGAC GCTCAGTGGA ACGAAAAC TC A
      AGAAAAGATG CCCAGACTG CGAGTCACCT TGCTTTTGAG T
```

FIG. 21(contd.)

50/93

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 TTTTTTGT TTGTTGTTTT
251 ACCCCAACCA TCTGAGAAAT AATCTTCTCC TTTCCAGCT TTATTCCAG
   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 GAGTGACAT
451 GGAATGGTTG GTCCCCAGGT GGTAGGTAGA GCTGTTTGAG GATTACGTGG
   CCTTACCAAC CAGGGGTCCA CCATCCATCT CGACAACTC CTAATGCACC
501 CCTTCTTGGA TGGGGGGTGG GGGTGGGGTG GGAGGGTTGG GTGGTGGGTA
   GGAAGAACCT ACCCCCCACC CCCACCCAC CCTCCCAACC CACCACCCAT
551 CTTAAGAGGT TTCAAAAGTC AATATTGTTT GCATTTAGCT CTTCTTGTA
   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 GGAACCTACT TTGTAGACCA GGCTGGCCTT
   ACATATCAGG ACCGACAGGA CCTTGAGTGA AACATCTGGT CCGACCGGAA
801 GAACTCAGAA ATCCGCCTGC CTCTGCCTCC CTAGCACTGG GATTAAAGGC
   CTTGAGTCTT TAGGCGGACG GAGACGGAGG GATCGTGACC CTAATTTCCG
851 GTGCGCCACC ACGCCAGCT TTCAACTGGA AATCTTAATA AACTTTCCTA
   CACGCGGTGG TCGGGGTCGA AAGTTGACCT TTAGAATTAT TTGAAAGGAT
901 GAAGTGGCCT TGGTTATGGG AGCTTATCAC AGCAATAGAA CAGCAATTAT
   CTTACCGGA ACCAATACCC TCGAATAGTG TCGTTATCTT GTCGTTAATA
951 GACTGGAGTA TGATAGTTAA AAACAAGCAA GCAAGCAAGC AAACACACAC
   CTGACCTCAT ACTATCAATT TTTGTTCGTT CGTTCGTTCT TTTGTGTGTG
1001 ACCAAAACAA CAAAACCCCA AGACAGAGTC ACATGTAGCC CAGGCTAGCC
   TGGTTTTGTT GTTTTGGGGT TCTGTCTCAG GTTACATCGG GTCCGATCGG
1051 TCCAAATTCA CTATATAACT GAAGAAGACC CTAATTCCC ATTCCTCTAG
   AGGTTTAAGT GATATATTGA CTTCTTCTGG GGATTAAGGG TAAGGAGATC
1101 AATCTATACC TCAAGTACTG AATGGCTTGG TTCACAATAC CCCACTAAAT
   TTAGATATGG AGTTCATGAC TTACCGAACC AAGTGTTATG GGGTGATTTA
1151 GATTGGTCTT ACTAAGTGCA ACAAGGTAAA CCTAAACTT CAGCCCTCAG
   CTAACCAGAA TGATTCACGT TGTTCCATTT GGATTTTGAA GTCGGGAGTC
1201 ACATCCCTTT TCCAGTATCA ATTTATAAAA TTAGATCCCA AGGATAAAAA
   TGTAGGGAAA AGGTCATAGT TAAATATTTT AATCTAGGGT TCCTATTTTT
1251 TTAATTGTAA AGTAAAATCA GAGTTCAGC ATCAACTACA GGCTCAACCA
   AATTAACATT TCATTTTAGT CTCAAGATCG TAGTTGATGT CCGAGTTGGT
1301 TGGGGACCAC AAATAAATAA 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	TGGAACCTCTA	CCAATAAATC	TTCAGGGGTA	CCCTATGGTT
1651	AATAACCCAC	TATTTATATG	CCCAAGCATT	TCACCTCCAC	AACAGTGCTA
	TTATTGGGTG	ATAAATATAC	GGGTTTCGTAA	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	ATTTCTTAG
1851	ACAAAAATAA	CCCCCTATCA	AATGCCTAGT	CCCTCTAACC	ATCACCAAGT
	TGTTTTTTATT	GGGGGATAGT	TTACGGATCA	GGGAGATTGG	TAGTGGTTCA
1901	GAAGGATCAC	GCAGGAAAAA	AAAAATCACC	AGCAGCACCT	CAGAACCAAG
	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	TGTAAATTTT
	TAACCTTAGT	TACACAGGAT	TAATTACATC	TTTGGTGTTG	ACATTTAAAG
2251	ACTCTTTTGA	CATGAATCTT	TTTCTAGACA	GGGTCTTGGA	TGCAGCCCCG
	TGAGAAAACT	GTACTIONGAA	AAAGATCTGT	CCCAGAACCT	ACGTCGGGGC
2301	ACTACCCAGA	ATTTTGGAAT	CCAGGCTAGC	CTCAAACCTCA	AGGCAATCTG
	TGATGGGTCT	TAAAACCTTA	GGTCCGATCG	GAGTTTGAGT	TCCGTTAGAC
2351	CTTGCTTCAG	CTTCTCACAG	GCTGGATCAC	AAACATACAC	CTTCAGACCC
	GAACGAAGTC	GAAGAGTGTC	CGACCTAGTG	TTTGTATGTG	GAAGTCTGGG
2401	ATTTTTTTTT	CCTCCCTCCG	TTTTTGGTTT	CTCTGTGTAG	CCCTGGGTGT
	TAAAAAATAA	GGAGGGAGGC	AAAAACCAAA	GAGACACATC	GGGACCCACA
2451	CCGTGGACTC	GCTGTGTAGA	TCTATCTACC	AGCCTCTGTC	TTGGAGTACT
	GGCACCTGAG	CGACACATCT	AGATAGATGG	TCGGAGACAG	AACCTCATGA
2501	GGGATTAAAG	TTGTGGGCTA	CCACTGCCTG	GCTGACCCAG	TTTTATTTAT
	CCCTAATTTT	AACACCCGAT	GGTGACGGAC	CGACTGGGTC	AAAATAAATA
2551	TTTAAATATA	ACTTGACAAA	AATAAATTTG	TCTAACTTAC	TAGAAATCCC
	AAATTTATAT	TGAACTGTTT	TTATTTAAAC	AGATTGAATG	ATCTTTTAGG
2601	AAGAAACTA	ACACTGGATT	TAGCAACAGT	CAGAAATCGC	TGAAAAGAAA
	TTCTTTTGAT	TGTGACCTAA	ATCGTTGTCA	GTCTTTAGCG	ACTTTTCTTT
2651	CAGAAATTGAT	CTAACAGTCT	TAGATCACTC	CTAGACAGTT	TGTAATTCCT
	GTCTTAACTA	GATTGTCAGA	ATCTAGTGAG	GATCTGTCAA	ACATTAAGAA
2701	GCTCATGGCA	ACGTGAGCTC	TATCTAATCT	ACTCTCTGTG	CACATAATGAA
	CGAGTACCGT	TGCACTCGAG	ATAGATTGAG	TGAGAGACAC	GTGATTACTT
2751	TGCTCAGTGT	CTCCAGAACA	GCACAGCTTC	CAGGGTAATC	ATGCCAACCC
	ACGAGTCACA	GAGGTCTTGT	CGTGTGGAAG	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	TTCCCTGGGC	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	ATGTTCCCTGA
3101	TAGGTAGGCT	GGCTTCTAAC	AGCAATTCAC	CCACTTATGA	TGGGATTAAA
	ATCCATCCGA	CCGAAGATTG	TCGTAAAGTG	GGTGAATACT	ACCCTAATTT
3151	GGAAGGCACA	ACCATGTCCA	CCACAGGTC	TAGCTCCCC	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	CAAACCTACC	AATGTAGTTT	TATTGGAGGC	CATTTTTTTAA
	ATCCAAAAGT	GTTTGAATGG	TTACATCAAA	ATAACCTCCG	GTAAAAAATT
3401	ATTTAATCGG	AGACTTGAAG	AGCTATTGCA	AGAAAAAAA	TGTAGGACAG
	TAAATTAGCC	TCTGAACTTC	TCGATAACGT	TCTTTTTTTTT	ACATCCTGTC
3451	TTAAAATTTT	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	CTACTTTCCCT
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
	GGGAGGGGA	AGTTAGTGCC	ATATTTTTTT	CCTTGTGTAA	TTCTACAGGG
3851	AGTCCTATTT	TCTGGCTTTT	TTTTTCCGGG	GGTGGTGGTG	CGGTAATCAC
	TCAGGATAAA	AGACCGAAAA	AAAAAGGCC	CCACCACCAC	GCCATTAGTG
3901	TCTCTATAGT	CCAGTCTGGG	CTTCAACGCC	TGGCAATCCC	GACCTCAAG
	AGAGATATCA	GGTCAGACCC	GAAGTTGCGG	ACCGTTAGGG	GTCGGAGTTC
3951	CTCCCAAGTA	CTGTCCTGAT	AAGGATAGAA	GGAGTCGACC	TCCTTCACGC
	GAGGGTTCAT	GACAGGACTA	TTCTATCTT	CCTCAGCTGG	AGGAAGTGCG
4001	TCCCCTCCGA	GGAGGGCTCC	TTCCCAGCTC	CATTCCCCGG	TCGGGAGCCC
	AGGGGAGGCT	CCTCCGAGG	AAGGGTCGAG	GTAAGGGGCC	AGCCCTCGGG
4051	GTCCTCCAC	CGAGAGCGCG	GGCCTCGTGG	TCAGCGCCTC	CGCGGGGAGA
	CAGGGGGTGG	GCTCTCGCGC	CCGGAGCACC	AGTCGCGGAG	GCGCCCTCT
4101	AACAAAGGCG	GCGGCGGGGG	CTCAAGGGCA	CTGCGCCACG	GGCCGCGCGC
	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	GAGGCCCCGC	CTTTGTCTGG	AGTCTCGGTA
	CGCTGCTCCG	ACCCCTCC	CTCCGGGCGC	GAAACAGACC	TCAGAGCCAT
4301	GCTGTCATCC	GGCTCCCACC	CTCATGCACA	ATTGTCCCAT	CTCCCCCAGC
	CGACAGTAGG	CCGAGGGTGG	GAGTACGTGT	TAACAGGGTA	GAGGGGGTGC
4351	CACCGGCGCG	GCGCCCGCCT	CAGCGAGGCC	CCAGCCGGTT	TCCCCGAGCC
	GTGGCCGCGC	CGCGGGCGGA	GTCGCTCCGG	GGTCGGCCAA	AGGGCGTCGG
4401	CGCGGCCAC	GGGGCTCGCA	GCCTCCCCGC	AAGCTCGGAC	GCACGGAGCA
	GCGCCGGGTG	CCCCGAGCGT	CGGAGGGGCG	TTCGAGCCTG	CGTGCTCTCGT
4451	TCCTAAACCC	CACCACACGC	AAGATCGAAA	AAAAGCAAAG	GCACGAACTT
	AGGATTTGGG	GTGGTGTGCG	TTCTAGCTTT	TTTTCTGTTTC	CGTGCTTGAA
4501	CACCGCTCCG	ATGCTCAGGG	CCGCGGATCC	TGCAGAGTCT	CCCGCCTGCG
	GTGGCGAGGC	TACGAGTCCC	GGCGCCTAGG	ACGTCTCAGA	GGGCGGACGC
4551	CGCTTCGGTT	CAGCCACATC	CGAGGGGAGG	GGGCGCGGGC	AGCTCCGCCG
	GCGAAGCCAA	GTCGGTGTAG	GCTCCCCCTC	CCGCGCGCCG	TGAGGCGGCG
4601	GGGGGGAGGG	GGAGCACCGC	CCACGCCCTG	GCCGCGCGGG	GCCCCGCCGG
	CCCCCTCCC	CCTCGTGGCG	GGTGCGGGAC	CGGCGCGCCC	CGGGCGGCCC
4651	AACGCTCCT	GCGGGGGGCG	GCGCGCGCAA	TGCTCACCGT	CCGCGGCGTG
	TTGCGCAGGA	CGCCCCCGCG	CGCGCGCGTT	ACGAGTGGCA	GGCGCCGCAC
4701	CGCCCCAGGG	GGTCTCCTGG	CTGGGGGGAG	GGGGGGGAAG	GCGGGCAGGA
	CGCGGGTCCC	CCAGAGGACC	GACCCCCCTC	CCCCCCTTC	CGCCCGTCTT
4751	AGGACCGCGG	AGGCCTCTCT	GCGTCTCGGA	GCGCGCCAAA	GCGGGGCTCC
	TCCTGGCGCC	TCCGGAGAGA	CGCAGAGCCT	CGCGCGGTTT	CGCCCCGAGG
4801	ACCCACCTCC	TTGCCCCGAT	CTTGAAGGCC	GGGGAGATAA	ACAGCGGGGT
	TGGGTGGAGG	AACGGGCCTA	GAACTTCCGG	CCCCTCTATT	TGTCGCCCCA
4851	TCTTTAAGCA	CCACCTCTCA	CTAGGCGCGG	GATCCCAAGG	CTTGTGGCAT
	AGAAATTTCGT	GGTGGAGAGT	GATCCGCGCC	CTAGGGTTCC	GAACACCGTA
4901	CCGGGGTGGT	ACTTGGACTA	AAAGTCCTTC	TGGGAGGGAC	CGAGTGAGAA
	GGCCCCACCA	TGAACCTGAT	TTTCAGGAAG	ACCCTCCCTG	GCTCACTCTT
4951	CCCCTTTGGG	ACGTGTAGAA	ATATTTGTGT	GGTTCGAGAA	TATTTGTGCG
	GGGGAACCC	TGCACATCTT	TATAAACACA	CCAAGCTCTT	ATAAACACGC
5001	GACGGGCTTG	GCAAAGGCGT	AGCTGCAGAG	AGCACGCTTG	GGTGGAGAGG
	CTGCCCCGAAC	CGTTTTCCGCA	TCGACGTCTC	TCGTGCGAAC	CCACCTCTCC
5051	GCCGCACGCC	CCAGCGCCGG	CCTAAGCCCC	TCCCGACGGC	GTTATTTCAA
	CGGCGTGCGG	GGTCGCGGCC	GGATTGCGGG	AGGGCTGCCG	CAATAAAGTT
5101	ACTGCGCGAC	CGTTTTCTCCG	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	TTTCGTGAGG	GCGACATGGA	TTTCCCCCTC
5251	TGTACGTAC	TTGGCGTAAG	GAGAGTGTAG	GCCCTTCCCG	CCATTGGCGG
	ACAGTGCATG	AACCGCATTC	CTCTCACATC	CGGGAAGGGC	GGTAACCGCC
5301	CGGTTAGGGC	GTTTACGTAA	CGGCGTGACG	TAAGCGGAGA	CGCGTTAGTG
	GCCAATCCCG	CAAATGCATT	GCCGCACTGC	ATTGCGCTCT	GCGCAATCAC
5351	GGGGGAAGGT	TCTAGAAAAG	CGGCGGTCTC	GGCTCCAGCG	GCAGTAGCAG
	CCCCCTTCCA	AGATCTTTTC	GCCGCCAGAG	CCGAGGTCGC	CGTCATCGTC
5401	CGGCGCCGCT	CCCGTGTGCA	GGAGTCTCCT	TGCGGCCAG	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	TGCCCAAAC	TTAGTTACTC	AATAATTTT
5601	TGGCATGGCG	AGGCCGTAGG	CACCGCAATG	GAAACCGGCC	ACCCGCTCC
	ACCGTACCGC	TCCGGCATCC	GTGGCGTTAC	CTTTGGCCGG	TGGGCGGAGG
5651	GTGGTCCGGC	GGAGGGGATG	CGGCCACTCG	AGTGGCGGTT	GGCCTTGGCG
	CACCAGGCCG	CCTCCCCCTAC	GCCGGTGAGC	TCACCGCCAA	CCGGAACCGC

FIG. 23(contd.) 54/93

5701	AGTTTCTGAG	GGGTCGTTGG	AGGAGGCCTC	TGATTGTCCG	ACCGCCTTCC
	TCAAAGACTC	CCCAGCAACC	TCTCCGGAG	ACTAACAGGC	TGGCGGAAGG
5751	CCGCCCTCAG	CCGCCCGGCG	CCATTTCCCT	CAGTTGGGGT	GGGGGATGGG
	GGCGGGAGTC	GGCGGGCCGC	GGTAAAGGGA	GTCAACCCCA	CCCCCTACCC
5801	AAGTGCCCGC	CGCGACCGGG	CTGGACCGCT	AAAGTAGCGC	GTGAGCGGGC
	TTCACGGGCG	GCGCTGGCCC	GACCTGGCGA	TTTCATCGCG	CACTCGCCCCG
5851	CATCGCTGGC	CTTTCGATGT	GCGCGGGCCT	AGGGGCTCGG	TTGTGTTTCG
	GTAGCGACCG	GAAAGCTACA	CGCGCCCGGA	TCCCCGAGCC	AACACAAGCG
5901	GGCGGAACGT	TTCTGGGGCC	CCCCCGGCTT	CCCGGAGCGA	GTCTGCGAAG
	CCGCCTTGCA	AAGACCCCGG	GGGGGCGGAA	GGGCCTCGCT	CAGACGCTTC
5951	CTAGCTTCCC	CTCCCCCCTC	TCCCGGGAAC	CGGATTGGC	GGCCGCCATT
	GATCGAAGGG	GAGGGGGGAG	AGGGCCCTTG	GCCTAAACCG	CCGGCGGTAA
6001	TTCCCGTCTC	CTTCCTCGCC	ACGATTTTGC	TTTCAACGCT	TTAGGTTTAC
	AAGGGCAGAG	GAAGGAGCGG	TGCTAAAACG	AAAGTTGCGA	AATCCAAATG
6051	TAGTTTGTTT	TTCTTTTTTC	ACCACTGCGT	AGACGTGTTT	AGCGATTTTC
	ATCAAACCAA	AAGAAAAAAG	TGGTGACGCA	TCTGCACAAA	TCGCTAAAAG
6101	CTTTCTTTTG	GAAGTCTTCA	TACCGTTTCG	AGGTGGATTT	AGCGTTTTGA
	GAAAGAAAAC	CTTCAGAAAT	ATGGCAAAGC	TCCACCTAAA	TCGCAAAACT
6151	GCTTGGGTCT	TCAGCGTCCT	GCGCACCTCG	CTAAAGGCTC	TCTGCCTTCC
	CGAACCCAGA	AGTCGCAGGA	CGCGTGAGC	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	GTGCTGCAAA
	GCATCATGTT	GGCTCCGGAC	CCGCCACCG	ACTTCCGTCT	CAGCGACGTT
6351	AGAGACCGTT	TTATTTTTCA	TAATACGTAA	GATTACGGGT	GCTGTAGTAA
	TCTCTGGCAA	AATAAAAAGT	ATTATGCATT	CTAATGCCCA	CGACATCATT
6401	AGCACTTGAG	CATTAGTATA	GTAGGAGGAA	GTCAAAGTGG	AAAAAATGGG
	TCGTGAACTC	GTAATCATAT	CATCCTCCTT	CAGTTTCACC	TTTTTTTACC
6451	AGCGCTCATC	AGGAAGCTAG	GGAGGCTATG	TTGAGTGCAG	GGTTACTTTC
	TCGCGAGTAG	TCCTTCGATC	CCTCCGATAC	AACTCACGTC	CCAATGAAAG
6501	CTTTTATTGC	AGAACTTTTA	TCTGCTTAAA	GGATCCTCGG	ATCGAAATAA
	GAAAATAAAG	TCTTGAAAAT	AGACGAATTT	CCTAGGAGCC	TAGCTTTATT
6551	TTCAAATTAT	AAGCATTTTT	AAGGGAATCT	TCGAATTTGT	TGGTAAAGTC
	AAGTTTAATA	TTCTGAAAAA	TTCCCTTAGA	AGCTTAAACA	ACCATTTTCA
6601	AACGGATCCT	TAGCACGTGG	TGTTCACTTT	AAGGAAGTGA	AATAGCTGAC
	TTGCCTAGGA	ATCGTGCACC	ACAAGTGAAA	TTCCCTTACT	TTATCGACTG
6651	TTTTTCATAG	TAGCCTTCGC	TAAAGCCCTG	GTTCACTGGA	CGAAAATCCA
	AAAAGTATCA	ATCGGAAGCG	AATTTCCGAC	CAAGTCACCT	GCTTTTAGGT
6701	CGTCCTGGCT	ATATAAAAAC	TTAGTTTGGG	GTCACAGTGT	TTGAGCGTGG
	GCAGGACCGA	TATATTTTTG	AATCAAACCC	CAGTGTACCA	AACTCGCACC
6751	TCATTTCGGT	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	TTCTGTGTGA	TATTTATTTT	ATTGGAACAA	ATATTAGCTA
6901	AGTGGACGTC	TGGTAAGTTT	GGAAAAAACC	CGAGGTAAGT	AAAGAGCTTT
	TCACCTGCAG	ACCATTCAAA	CCTTTTTTGG	GCTCCATTCA	TTTCTCGAAA
6951	TGCTTTCGTT	AGTGATATGA	AAAAACAAGG	TGTATTTAAT	ACTTGCAACT
	ACGAAAGCAA	TCACTATACT	TTTTTGTTC	ACATAAATTA	TGAACGTTGA
7001	TAGTTTAAGG	AAAGCCAATT	TACTGACATT	TTAGTAGAGC	TACCAGAAAC
	ATCAAATTCC	TTTCGGTTAA	ATGACTGTAA	AATCATCTCG	ATGGTCTTTG
7051	ACTATTGGGA	GTCCTGATTA	AGGCTTTTGT	AACTATTTTG	ACTATTTTAA
	TGATAAACCT	CAGGACTAAT	TCCGAAAACA	TTGATAAAAC	TGATAAATTT
7101	ACAATTTTGG	TCGTTTTTAT	TAAACATTTT	AAAACCTAAA	AATTGTAAAC
	TGTTAAACC	AGCAAAAATA	ATTTGTAAAG	TTTTGGATTT	TTAACATTTG

FIG. 23(contd.) 55/93

```

7151  ATTGGCTTTT  TGAGCACATT  TTGGAGAAAC  TTACAAATTT  AGGCTATACA
      TAACCGAAAA  ACTCGTGTA  AACCTCTTTG  AATGTTTAAA  TCCGATATGT
7201  GTAAATAAAC  GGATTTGTTT  TATAATTTTG  CTTTTTCATT  TCGTTGTGCA
      CATTTTATTG  CCTAAACAAA  ATATTAAAC  GAAAAAGTAA  AGCAACACGT
7251  GTCATAGGTC  CTGGATAGTA  TGACCTAATT  TATGAACATC  TTGATAAGTT
      CAGTATCCAG  GACCTATCAT  ACTGGATTAA  ATACTTGTAG  AACTATTCAA
7301  TTTGTACTTA  GCTATTGGAA  AGCCAGTATT  AAGTGCCTGA  CAAAACCAGA
      AAACATGAAT  CGATAACCTT  TCGGTCATA  TTCACGGACT  GTTTTGGTCT
7351  TTTAAGGTGA  TATCTGGAGT  TTCAGCATT  TTCATGGAGC  TGTTCAGCA
      AAATTCCACT  ATAGACCTCA  AAGTCGTAAG  AAGTACCTCG  AACAAAGTCT
7401  GTTGCAAGAT  TTTTTTTTTT  CATCTTGAGA  TACTTACAAT  TAACACCAGA
      CAACGTCCTA  AAAAAAAAAA  GTAGAACTCT  ATGAATGTTA  ATTGTGGTCT
7451  GGGGGCAGCT  CAGGGAAAAG  CAAATATGCC  ACTTTTCAGA  AACTGAATCT
      CCCCCGTCGA  GTCCCTTTTC  GTTTATACGG  TGAAAAGTCT  TTGACTTAGA
7501  TGGAAGTGGT  GAATTTGGAA  ACAGTTTTTT  TAAATTTTTT  TTAAATCTAA
      ACCTTCACCA  CTTAAACCTT  TGTCCAAAAA  ATTTAAAAAA  AATTTAGATT
7551  AAAGTAGTAA  ATTTTGGACT  TGGGTGTAG  AATTTAATGA  ATTACAAAAG
      TTTTCATCATT  TAAAACCTGA  ACCCAACATC  TTAAATTACT  TAATGTTTTT
7601  AATTCTTTAA  TACCCTTTAA  ATGACCTAAG  AGCTGGGTAT  GGTTTTCTG
      TTAAGAAATT  ATGGGAAATT  TACTGGATTC  TCGACCCATA  CAAAAAGAC
7651  AATTTTTTTG  AAGAAAATCT  AAGAAAGTTT  ACGTGAATTA  GAAGTTAGAT
      TTAAAAAAC  TTCTTTTAGA  TTCTTTCAA  TGCCTTAAT  CTTCAATCTA
7701  CGAATATTAG  TGACTTTGAA  ACTTGTATAG  CTCAGGCAAT  TTTTGGTGTA
      GCTTATAATC  ACTGAAACTT  TGAACATATC  GAGTCCGTTA  AAAACCACAT
7751  ACACAATAA  TATGCAGTTT  AACATATGGT  TTAAATTTGA  TGTAAGTTTT
      TGTGTTGATT  ATACGTCAA  TTGTATACCA  AATTTAAACT  ACATTCAAAA
7801  TTTTCTCCCC  CCCAGAAAAC  TTTAGAAACT  GTTCCTTTGG  AGAGGAAAAA
      AAAAGAGGGG  GGGTCTTTTG  AAATCTTTGA  CAAGGAAACC  TCTCCTTTTT
7851  GGTACTCTAC  CAGCAGGTCA  CCTCATATTT  AAGAATTTAA  TTCTTAAATT  AAAGGACGTA
      CCATGAGATG  GTCGTCCAGT  GGAGTATAAA  TTCTTAAATT  AAAGGACGTA
7901  ACAAAGAAAG  TGTAATAAAA  AATTGAAATG  GTATTTCCCT  TTGCAGAGAG
      TGTTCCTTTC  ACATTTATTT  TTAACCTTAC  CATAAAGGGA  AACGTCTCTC
7951  AAAAGGAACA  GTTCCGAAAG  CTCTTTATTG  GTGGCTTAAG  CTTATCGATA
      TTTTCCTTGT  CAAGGCTTTC  GAGAAATAAC  CACCGAATTC  GAATAGCTAT
8001  CCGGTGGCGC  GCCAATTGTT  AATTAAGATC  TGGCCCAATG  GGCCGTACGA
      GGCCACCGCG  CGGTAAACAA  TTAATTCTAG  ACCGGGTAC  CCGGCATGCT
8051  ATTCTTAGG  CTACCGGTA  GGGGAGGCG  TTTCCCAAG  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  CCGTCTTTT  GTGGCCCTT  CGCGCCACCT  TCTACTCCTC  CCCTAGTCAG
      GGCAAGAAAC  CACCGGGGAA  GCGCGGTGGA  AGATGAGGAG  GGGATCAGTC
8251  GAAGTTCCCC  CCCGCCCGC  AGCTCGCGTC  GTGCAGGACG  TGACAAATGG
      CTTCAAGGGG  GGGCGGGGCG  TCGAGCGCAG  CACGTCCTGC  ACTGTTTACC
8301  AAGTAGCACG  TCTCACTAGT  CTCGTGCAGA  TGGACAGCAC  CGCTGAGCAA
      TTCATCGTGC  AGAGTGATCA  GAGCACGTCT  ACCTGTCTGT  GCGACTCGTT
8351  TGGAAGCGGG  TAGGCCTTTG  GGGCAGCGGC  CAATAGCAGC  TTTGCTCCTT
      ACCTTCGCCC  ATCCGGAAC  CCCGTCGCCG  GTTATCGTCG  AAACGAGGAA
8401  CGCTTTCTGG  GCTCAGAGGC  TGGGAAGGGG  TGGGTCCGGG  GGCGGGCTCA
      GCGAAAGACC  CGAGTCTCCG  ACCCTTCCCC  ACCCAGGCCC  CCGCCGAGT
8451  GGGGCGGGCT  CAGGGGCGGG  GCGGGCGCCC  GAAGGTCCCTC  CGGAGGCCCG
      CCGCGCCCGA  GTCCCCGCC  CGCCCGCGGG  CTTCAGGAG  GCCTCCGGGC
8501  GCATTCTGCA  CGCTTCAAAA  GCGCACGTCT  GCCGCGCTGT  TCTCCTCTTC
      CGTAAGACGT  GCGAAGTTTT  CGCGTGCAGA  CGGCGCGACA  AGAGGAGAAG
8551  CTCATCTCCG  GGCCTTTCGA  CCAGCTTACC  ATGACCGAGT  ACAAGCCCAC
      GAGTAGAGGC  CCGGAAAGCT  GGTGCAATGG  TACTGGCTCA  TGTTCCGGGTG

```

FIG. 23(contd.) 56/93

8601	GGTGCGCCTC	GCCACCCGCG	ACGACGTCCC	CAGGGCCGTA	CGCACCCCTCG
	CCACGCGGAG	CGGTGGGCGC	TGCTGCAGGG	GTCCCCGCAT	GCGTGGGAGC
8651	CCGCCGCGTT	CGCCGACTAC	CCCGCCACGC	GCCACACCGT	CGATCCGGAC
	GGCGGCGCAA	GCGGCTGATG	GGGCGGTGCG	CGGTGTGGCA	GCTAGGCCTG
8701	CGCCACATCG	AGCGGGTCAC	CGAGCTGCAA	GAACCTCTTC	TCACGCGCGT
	GCGGTGTAGC	TCGCCCAGTG	GCTCGACGTT	CTTGAGAAGG	AGTGCGCGCA
8751	CGGGCTCGAC	ATCGGCAAGG	TGTGGGTGCG	GGACGACGGC	GCCGCGGTGG
	GCCCGAGCTG	TAGCCGTTCC	ACACCCAGCG	CCTGCTGCCG	CGGCGCCACC
8801	CGGTCTGGAC	CACGCCGAG	AGCGTCGAAG	CGGGGGCGGT	GTTCCGCCGAG
	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	GTGCGGTCTC	GCCCGACCAC	CAGGGCAAGG	GTCTGGGCAG
	GGACCGGTGG	CAGCGCAGAG	CGGGCTGGTG	GTCCCGTTCC	CAGACCCGTC
9001	CGCCGTCGTG	CTCCCCGAG	TGGAGGCGGC	CGAGCGCGCC	GGGGTGCCCC
	GCGGCAGCAC	GAGGGGCCTC	ACCTCCGCCG	GCTCGCGCGG	CCCCACGGGC
9051	CCTTCCTGGA	GACCTCCGCG	CCCCGCAACC	TCCCTTCTA	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	CTGGGCGTCC
9201	GCCCGACCGA	AAGGAGCGCA	CGACCCCATG	CATCGTAGAG	CTCGCTGATC
	CGGGCTGGCT	TTCCTCGCGT	GCTGGGGTAC	GTAGCATCTC	GAGCGACTAG
9251	AGCCTCGACT	GTGCCTTCTA	GTTGCCAGCC	ATCTGTTGTT	TGCCCCCTCC
	TCGGAGCTGA	CACGGAAGAT	CAACGGTCGG	TAGACAACAA	ACGGGGAGGG
9301	CCGTGCCTTC	CTTGACCCTG	GAAGGTGCCA	CTCCCCTGT	CCTTTCCTAA
	GGCACGGAAG	GAACGTGGAC	CTTCCACGGT	GAGGGTGACA	GGAAAGGATT
9351	TAAATGAGG	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	ACCCCCCGCG	CACCCCGGAT	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	TGATAGCTCG
9651	CGGAAGCATA	AAGTGTAAG	CCTGGGGTGC	CTAATGAGTG	AGTAACTCA
	GCCTTCGTAT	TTCACATTTT	GGACCCACG	GATTACTCAC	TCGATTGAGT
9701	CATTAATTGC	GTTGCGCTCA	CTGCCCGCTT	TCCAGTCGGG	AAACCTGTCTG
	GTAATTAACG	CAACGCGAGT	GACGGGCGAA	AGGTCAGCCC	TTTGGACAGC
9751	TGCCAGCATC	CGGAGCACTT	TTCGGGGAAA	TGTGCGCGGA	ACCCCTATTT
	ACGGTCGTAG	CGCTCGTGAA	AAGCCCTTT	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
	GGAACCTCTCA	AAAGCGGGGC	TTCTTGCAAA	AGGTTACTAC	TCGTGAAAAAT
10101	AAGTTCTGCT	ATGTGGCGCG	GTATTATCCC	GTATTGACGC	CGGGCAAGAG
	TTCAAGACGA	TACACCGCGC	CATAATAGGG	CATAACTGCG	GCCCGTTCTC
10151	CAACTCGGTC	GCCGCATACA	CTATTCTCAG	AATGACTTGG	TTGAGTACTC
	GTTGAGCCAG	CGGCGTATGT	GATAAGAGTC	TTACTGAACC	AATCATGAG
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
	TGTTTGCTAGC	CTCCTGGCTT	CCTCGATTGG	CGAAAAAACG	TGTTGTACCC
10351	GGATCATGTA	ACTCGCCTTG	ATCGTTGGGA	ACCGGAGCTG	AATGAAGCCA
	CCTAGTACAT	TGAGCGGAAC	TAGCAACCCT	TGGCCTCGAC	TTACTTCCGT
10401	TACCAAACGA	CGAGCGTGAC	ACCACGATGC	CTGTAGCAAT	GGCAACAACG
	ATGGTTTGCT	GCTCGCACTG	TGGTGCTACG	GACATCGTTA	CCGTTGTTGC
10451	TTGCGCAAAAC	TATTAACCTG	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	CATTCCGGAG
10651	CCGTATCGTA	GTTATCTACA	CGACGGGGAG	TCAGGCAACT	ATGGATGAAC
	GGCATAGCAT	CAATAGATGT	GCTGCCCCCTC	AGTCCGTTGA	TACCTACTTG
10701	GAAATAGACA	GATCGCTGAG	ATAGGTGCCT	CACTGATTAA	GCATTGGTAA
	CTTTATCTGT	CTAGCGACTC	TATCCACGGA	GTGACTAATT	CGTAACCATT
10751	CTGTCAGACT	CGCGACACTG	CATTAATGAA	TCGGCCAACG	CGCGGGGAGA
	GACAGTCTGA	GCGCTGTGAC	GTAATTACTT	AGCCGGTTGC	GCGCCCCCTC
10801	GGCGGTTTGC	GTATTGGGCG	CTCTTCCGCT	TCCTCGCTCA	CTGACTCGCT
	CCGCCAAACG	CATAACCCGC	GAGAAGGCGA	AGGAGCGAGT	GACTGAGCGA
10851	GCGCTCGGTC	GTTGCGGCTG	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	GTCCTTGCGA	TTTTTCCGGC	GCAACGACCG
11001	GTTTTTCCAT	AGGCTCCGCC	CCCCTGACGA	GCATCACAAA	AATCGACGCT
	CAAAAAGGTA	TCCGAGGCGG	GGGGACTGCT	CGTAGTGTTT	TTAGCTGCGA
11051	CAAGTCAGAG	GTGGCGAAAC	CCGACAGGAC	TATAAAGATA	CCAGGCGTTT
	GTTCACTCTC	CACCGCTTTG	GGCTGTCTCTG	ATATTTCTAT	GGTCCGCAAA
11101	CCCCCTGGAA	GCTCCCTCGT	GCGCTCTCCT	GTTCCGACCC	TGCCGCTTAC
	GGGGGACCTT	CGAGGGAGCA	CGCGAGAGGA	CAAGGCTGGG	ACGGCGAATG
11151	CGGATACCTG	TCCGCCTTTC	TCCCTTCGGG	AAGCGTGCGG	CTTTCTCATA
	GCCTATGGAC	AGGCGGAAAG	AGGGAAGCCC	TTCGCACCGC	GAAAGAGTAT
11201	GCTCACGCTG	TAGGTATCTC	AGTTCGGTGT	AGGTCGTTTCG	CTCCAAGCTG
	CGAGTGCGAC	ATCCATAGAG	TCAAGCCACA	TCCAGCAAGC	GAGGTTTCGAC
11251	GGCTGTGTGC	ACGAACCCCC	CGTTCAGCCC	GACCGCTGCG	CCTTATCCGG
	CCGACACACG	TGCTTGGGGG	GCAAGTCGGG	CTGGCGACGC	GGAATAGGCC
11301	TAACTATCGT	CTTGAGTCCA	ACCCGGTAAG	ACACGACTTA	TCGCCACTGG
	ATTGATAGCA	GAATCAGGT	TGGGCCATTC	TGTGCTGAAT	AGCGGTGACC
11351	CAGCAGCCAC	TGGTAACAGG	ATTAGCAGAG	CGAGGTATGT	AGGCGGTGCT
	GTCGTCGGTG	ACCATTTGTCC	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

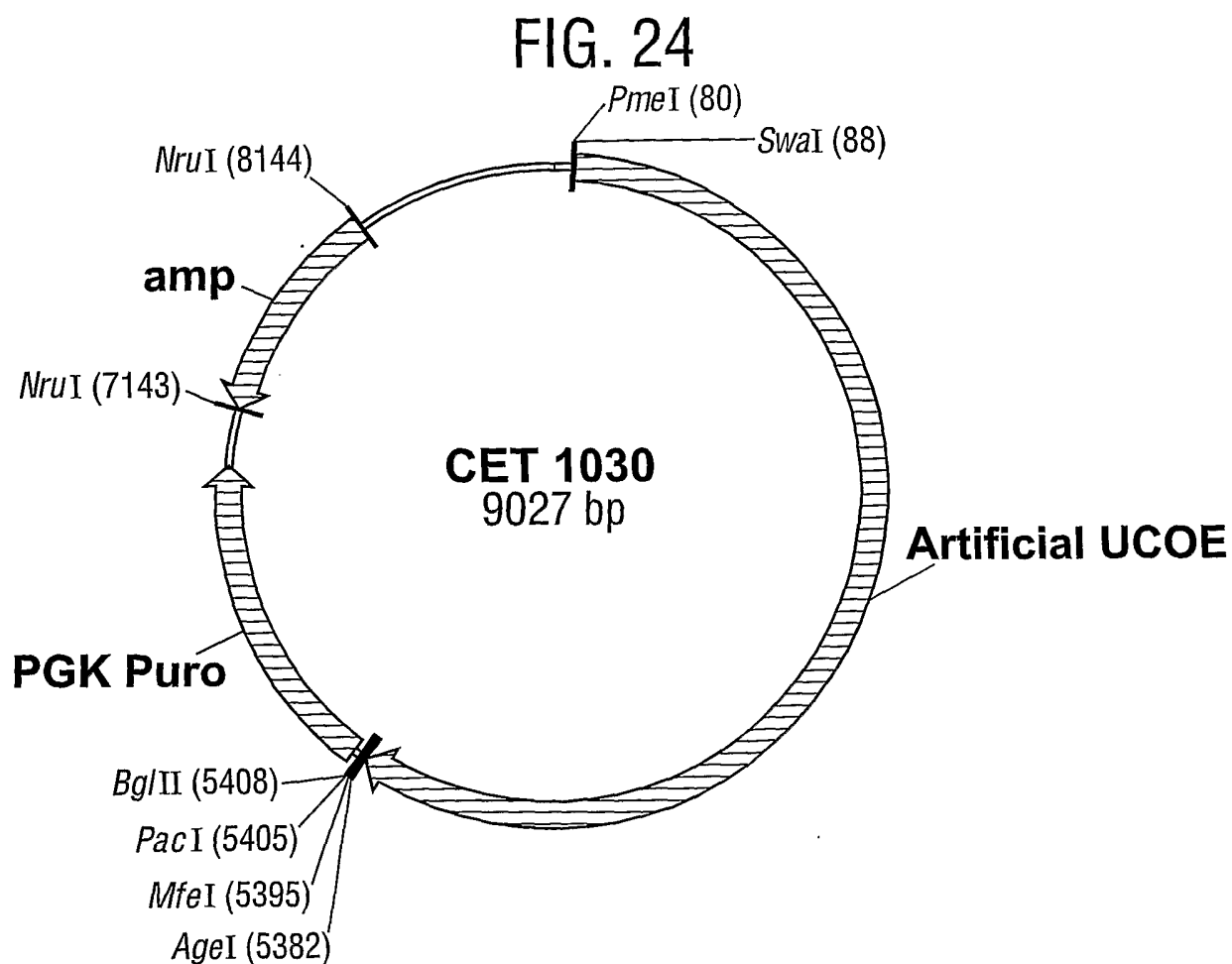
58/93

```

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
        AAAC TAGAAA AGATGCCCCA GACTGCGAGT CACCTTGCTT TTGAGT

```

FIG. 23(contd.)



59/93

FIG. 25 **cET1030 nucleotide sequence**

```

1   CGTTGTAAAA CGACGGCCAG TGAATTGTAA TACGACTCAC TATAGGGCGA
   GCAACATTTT GCTGCCGGTC ACTTAACATT ATGCTGAGTG ATATCCCCT
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 TGTTCCTTCC
   ACCCGAGGAG GTATTCTGAC ACTCGTAGAC TGGTGACCTC ACAACGAAGG
251 CATTATATCC CTGTTATCAA GCACAAGGTC AGGCACAGAG TAAGACTCAA
   GTAATATAGG GACAATAGTT CGTGTTCAG TCCGTGTCTC ATTCTGAGTT
301 AACATGTTTT GGAATGTATG ACTGGTATGA ACTACAAACC AGTAAGCTGA
   TTGTACAAAA CCTTACATAC TGACCATACT TGATGTTTGG TCATTCGACT
351 TGTTTTCATT TTGAGTCTAT AAATCTAATT TTGTGGTGGT TTTGTGTATG
   ACAAAGTAA 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 TAAACTGGC CAGATGAAAG AGAACTTTT
   TTGACTGTTT TAACACATAG ATTTTGACCG GTCTACTTTC TCTTTGAAAA
701 AAGGGGCCCT TCTGCGTGCC CGACACTGTG CTAGGCACTC ACASTATCCC
   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
   TGTCCCTCCT CTTGTGTGGA GAATACACAG TCTTTTGTTC CAGGTCCTTC
901 TATATTTTTTA CCTGAGGCAA TATCTGAAAA TTGTATGCTA CAGCCTCCAA
   ATATAAAAT GGAATCCGTT ATAGACTTTT AACATACGAT GTCGGAGGTT
951 AGTGAGTCTT CCTCTCAGTA CCTCTCTTCT AGGCACATGG AGCCCTTTCT
   TCACTCAGAA GGAGAGTCAT GGAGAGAAGA TCCGTGTACC TCGGGAAAGA
1001 TCCAAGTATT ATGTTTAACC ACTTAATGAA TGAAGTCCTG AAAGTCTTA
   AGGTTCTATA 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 TGGAAATATA TGGTCCAGAT GATCTGAAAC AAAAAGGACA
   TCAACACCAG ACCTTAATAT ACCAGGTCTA CTAGACTTTG TTTTTCCTGT
1201 GCACTATTAG TAATCATTTA GTTTTGAAGA CAGTCTAATA ATTTGCTGTC
   CGTGATAATC ATTAGTAAAT CAAAACCTCT GTCAGATTAT TAAACGACAG
1251 TCTAAAGTAC TATATTCCCT ATAGTTCTGG CATTTTAGAT AAAGGGTCAT
   AGATTTTCAT ATATAAGGGA TATCAAGACC GTAAAATCTA TTTCCAGTA
1301 AAATTAATG CCTATATGGT GACATTATTC AGTGATTCAG ACTTCACAGC
   TTTAATTTAC GGATATACCA CTGTAATAAG TCACTAAGTC TGAAGTGTCTG

```

FIG. 25(contd.) 60/93

1351	CTTTTTTTTT	TTTTTACAAA	GGTGTTCAG	GCATGAAAAA	TTTTAAAGTA
	GAAAAAAAAA	AAAAATGTTT	CCACAAGGTC	CGTACTTTTT	AAAATTTTCAT
1401	CTATACCTTT	CCTAATTTTA	CCTTTAAAGT	TGTCCTGGAA	ATATCTGGGT
	GATATGGAAA	GGATTAAAAT	GGAAATTTCA	ACAGGACCTT	TATAGACCCA
1451	TGACAAAGGC	GATGAAACTG	AACTGAGACT	TAAAAAAAAG	ATTACCCACC
	ACTGTTTCCG	CTACTTTGAC	TTGACTCTGA	ATTTTTTTTT	TAATGGGTGG
1501	TGGTTGTGCA	CAAGCCTGCT	TATGTCCCAA	TCTCCAGTCT	AGGGTCTGAT
	ACCAACACGT	GTTCCGACGA	ATACAGGGTT	AGAGGTCAGA	TCCCAGACTA
1551	GCTCCTTGCT	GCAGTAATAT	GCTTTGTGGC	ATCTGGAGCA	CGTTTTGGGG
	CGAGGAACGA	CGTCATTATA	CGAAACACCG	TAGACCTCGT	GCAAAACCCC
1601	CCTAAACAGC	CACAAACCCT	GCAGAGATGA	GCACCAGACT	TAAGCTGGAG
	GGATTTGTCT	GTGTTTGGGA	CGTCTCTACT	CGTGGTCTGA	ATTGCACCTC
1651	ACACACTGAT	TCTCCTGTTT	CTGGGGGAGG	ATTCTCAGAA	GGTGGCTCAT
	TGTGTGACTA	AGAGGACAAA	GACCCCTCC	TAAGAGTCTT	CCACCGAGTA
1701	ATGAGTAAAA	ATCGTTTTTC	CTGGGTAGTT	GATTCCTAAA	AACTAAAAAA
	TACTCATTTT	TAGCAAAAAG	GACCCATCAA	CTAAGGATTT	TTGATTTTTT
1751	GAATACAGAG	AAAAGTTTTA	TCTTCAAACA	AAACAGCAAT	TCACATATTT
	CTTATGTCTC	TTTTCAAAAT	AGAAGTTTGT	TTTGTCTGTA	AGTGATATAA
1801	TATCCTCTGC	ACGTAAAACT	GAAAATAACA	ACAACAAAAA	AGAAATGAAA
	ATAGGAGACG	TGCATTTTGA	CTTTTATTGT	TGTTGTTTTT	TCTTTACTTT
1851	GTTTTTGCTT	TCAGGAATAA	GCTTTTAAAA	TCCAGAAACT	AGATTTTCGTC
	CAAAAACGAA	AGTCCTTATT	CGAAAATTTT	AGGTCTTTGA	TCTAAAGCAG
1901	CGGTACACGC	AACTGAGTTG	CCTCCTAGAG	GTGGTTTGAG	TTAATCAAAT
	GCCATGTGCG	TTGACTCAAC	GGAGGATCTC	CACCAAACCTC	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	CCCGACCCCC	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	GTCTCCTTC
2401	GAGAGCGGGC	GGCCGCACAG	CTCGCAGGCC	AGGGCCTGGG	GCCCCGCCAG
	CTCTCGCCCC	CCGGCGTGTC	GAGCGTCCGG	TCCCGGACCC	CGGGGCCGTC
2451	CCCGGCCGCG	CCCAGCCATG	CCGGCCGCCC	GCCCACCTTG	CTGGGGAAC
	GGGCCGCGCG	GGGTCGGTAC	GGCCGGCGGG	CGGGTGGAAC	GACCCCTTGA
2501	GCTCGCTGCG	CAGTCGCCAC	GCCGGCGCCG	ACTCGGCGAA	GCCCAGCTCC
	CGAGCGACGC	GTCAGCGGTG	CGGCCGCGGC	TGAGCCGCTT	CGGGTCGAGG
2551	ACAGGCCTGG	CCCCGGCGGC	AGCCATGCGG	GGCGCGGGCT	GGCGTGGGGC
	TGTCGGGACC	GGGGCCGCCG	TCGGTACGCC	CCGCGCCCCG	CCGCACCCCC
2601	GCAGCCCACA	GCTGGGTCCG	AAGGCGGAAA	TCGGGCGCCG	GGCCGGAAGG
	CGTCGGGTGT	CGACCCAGCC	TTCCGCTTTT	AGCCCGCGGC	CCGGCCTTCC
2651	CAAGAGGCGG	GCACCTTTCC	GGAGGACAGG	AGGCGGAAAC	GCGTCTGACG
	GTTCTCCGCC	CGTGGAAGG	CCTCCTGTCC	TCCGCCTTTG	CGCAGACTGC
2701	GGAGCGGTTG	CAGGACCAAT	GCGAGGGAAC	GGGGCAGAGG	AAACCTCTCG
	CTTCGCCAAC	GTCCTGGTTA	CGCTCCCTTG	CCCCGTCTCC	TTTGGAGAGC
2751	GCATCAGCCC	CGCCCCGTGG	GCCTCTGCCT	CCGAGCCGCT	TTCTTGGTGC
	CGTAGTCGGG	GCGGGGACCG	CGGAGACGGA	GGCTCGGCGA	AAGGACCACG

FIG. 25(contd.) 61/93

2801	CTCCGGGTGC	TCTGGGATGG	TTCTGGTCTT	TGGGAGAGTG	GCAGCTGGTG
	GAGGCCACG	AGACCCTACC	AAGACCAGAA	ACCCTCTCAC	CGTCGACCAC
2851	ACGGCGCTCC	GCTCACCTCT	GCACATGTCT	TGCTGTGGGC	CTGCGGGTGG
	TGCCGCGAGG	CGAGTGGAGA	CGTGTACAGA	ACGACACCCG	GACGCCACCC
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	GGATGTGCGC
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	AAC TAGTTT	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	TCTTGCTTGG
3351	GTGCCCCACC	GGTCTTGTGT	GCTACTCAGT	GGACAGACCC	AAGGCAAGAA
	GACGGGTGGG	CCAGAACACA	CGATGAGTCA	CCTGTCTGGG	TTCCGTTCTT
3401	AGGGTGACAA	GGACAGGGTC	TTCCCAGGCT	GGCTTTGAGT	TCCTAGCACC
	TCCCCTGTGT	CCTGTCCCAG	AAGGGTCCGA	CCGAAACTCA	AGGATCGTGG
3451	GCCCCGCCCC	CAATCCTCTG	TGGCACATGG	AGTCTTGCTC	CCCAGAGTCC
	CGGGGCGGGG	GTTAGGAGAC	ACCGTGTACC	TCAGAACCAG	GGGTCTCAGG
3501	CCCAGCGGCC	TCCAGATGGT	CTGGGAGGGC	AGTTCAGCTG	TGGCTGCGCA
	GGGTGCGCCG	AGGTCTACCA	GACCCTCCCG	TCAAGTCGAC	ACCGACGCGT
3551	TAGCAGACAT	ACAACGGACG	GTGGGCCCCAG	ACCCAGGCTG	TGTAGACCCA
	ATCGTCTGTA	TGTTGCCTGC	CACCCGGGTC	TGGGTCCGAC	ACATCTGGGT
3601	GCCCCCCCCG	CCCGCAGTGC	CTAGGTCACC	CACTAACGCC	CCAGGCCTGG
	CGGGGGGGCG	GGGCGTCACG	GATCCAGTGG	GTGATTGCGG	GGTCCGGACC
3651	TCTTGGGCTG	GCGTGACTGT	TACCCTCAAA	AGCAGGCAGC	TCCAGGGTAA
	AGAACCGACC	CGCACTGACA	ATGGGAGTTT	TCGTCCGTCG	AGGTCCCATT
3701	AAGGTGCCCT	GCCCTGTAGA	GCCCCTTCC	TTCCCAGGGC	TGCGGCTGGG
	TTCCACGGGA	CGGGACATCT	CGGGTGAAGG	AAGGGTCCCG	ACGCCGACCC
3751	TAGGTTTGTA	GCCTTCATCA	CGGGCCACCT	CCAGCCACTG	GACCGCTGGC
	ATCCAAACAT	CGGAAGTAGT	GCCCAGGTGA	GGTCGGTGAC	CTGGCGACCC
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	TGCGTCTCCA
3951	CCTCAGGCCC	TGCTTTTCAGG	GCAGTTGCTC	TGAAGTCGGC	AAGGGGGAGT
	GGAGTCCGGG	ACGAAAGTCC	CGTCAACGAG	ACTTCAGCCG	TTCCCCCTCA
4001	GACTGCCTGG	CCACTCCATG	CCCTCCAAGA	GCTCCTTCTG	CAGGAGCGTA
	CTGACGGACC	GGTGAGGTAC	GGGAGGTTCT	CGAGGAAGAC	GTCTCTGCAT
4051	CAGAACCAG	GGCCCTGGCA	CCCGTGACAG	CCCTGGCCCA	CCCCACCTGG
	GTCTTGGGTC	CCGGGACCGT	GGGCACGTCT	GGGACCGGGT	GGGGTGGACC
4101	GCGCTCAGTG	CCCAAGAGAT	GTCCACACCT	AGGATGTCCC	GCGGTGGGTG
	CGCGAGTCAC	GGGTTCTCTA	CAGGTGTGGA	TCCTACAGGG	CGCCACCCAC
4151	GGGGGGCCCG	GAGACGGGCA	GGCCGGGGGC	AGGCCTGGCC	ATGCGGGGCC
	CCCCCGGGCT	CTCTGCCCCG	CCGGCCCCCG	TCCGGACCGG	TACGCCCCGG
4201	GAACCGGGCA	CTGCCCAGCG	TGGGGCGCGG	GGGCCACGGC	GCGCGCCCCC
	CTTGGCCCGT	GACGGGTGCG	ACCCCGCGCC	CCCGGTGCCG	CGCGCGGGGG

FIG. 25(contd.) 62/93

```

4251  AGCCCCCGGG CCCAGCACCC CAAGGCGGCC AACGCCAAAA CTCTCCCTCC
      TCGGGGGGCC GGGTCGTGGG GTTCCGCCGG TTGCGGTTTT GAGAGGGAGG
4301  TCCTCTTCCT CAATCTCGCT CTCGCTCTTT TTTTTTTTCG CAAAAGGAGG
      AGGAGAAGGA GTTAGAGCGA GAGCGAGAAA AAAAAAAGC GTTTTCCTCC
4351  GGAGAGGGGG TAAAAAATG CTGCACTGTG CGGCGAAGCC GGTGAGTGAG
      CCTCTCCCCC ATTTTTTTTAC GACGTGACAC GCCGCTTCGG CCACTCACTC
4401  CGGCGCGGGG CCAATCAGCG TCGCCCGTTC CGAAAGTTGC CTTTTATGGC
      GCCGCGCCCC GGTAGTTCG ACGCGGCAAG GCTTTCACG GAAAATACCG
4451  TCGAGCGGCC GCGGCGGCGC CCTATAAAAC CCAGCGGCGC GACGCGCCAC
      AGCTCGCCGG CGCCGCGCGG GGATATTTTG GGTGCGCCGG CTGCGCGGTG
4501  CACCGCCGAG ACCGCGTCCG CCCGCGAGCA CAGAGCCTCG CCTTTGCCGA
      GTGGCGGCTC TGGCGCAGGC GGGCGCTCGT GTCTCGGAGC GGAAACGGCT
4551  TCCGCCGCCG GTCCACACCC GCCGCCAGGT AAGCCCGGCC AGCCGACCCG
      AGGCGGCGGG CAGGTGTGGG CGGCGGTCCA TTCGGGCCGG TCGGCTGGCC
4601  GGCATGCGGC CGCGGCCCTT CGCCCGTGCA GAGCCGCGT CTGGGCCGCA
      CCGTACGCCG GCGCCGGGAA GCGGCGACGT CTCGGCGGCA GACCCGGCGT
4651  GCGGGGGGCG CATGGGGCGG AACCAGACCG CCGTGGGGGG CGCGGGAGAA
      CGCCCCCGCG GTACCCCGCC TTGGCCTGGC GGCACCCCCC GCGCCCTCTT
4701  GCCCCCTGGG CTCCGGAGAT GGGGGACACC CCACGCCAGT TCGCAGGCGC
      CGGGGACCCG GAGGCCTCTA CCCCCTGTGG GGTGCGGTCA AGCGTCCGCG
4751  GAGGCCGCGC TCGGGCGGGC GCGCTCCGGG GGTGCGGCTC TCGGGGCGGG
      CTCCGCGCGC AGCCCGCCCG CGCGAGGCCC CCACGGCGAG AGCCCGGCCC
4801  GGCAACCGGC GGGGTCTTTG TCTGAGCCGG GCTCTTGCCA ATGGGGATCG
      CCGTTGGCCG CCCAGAAAC AGACTCGGCC CGAGAACGGT TACCCCTAGC
4851  CACGGTGGGC GCGGCGTAGC CCCCCTCAGG 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 CTTCTGCCC 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 GCGCCGCGCG GGCCGAAGGA AACAGGGGTT AGACCCGCGC
5251  CGCCGGCGCC CCCTGGCGGC CTAAGGACTC GGCGCGCCGG AAGTGGCCAG
      GCGGCCGCGG GGGACCGCCG GATTCCCTGAG CCGCGCGGCC TTCACCGGTC
5301  GCGGGGGGCG 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
      TCCCCCTCCG GAAAAGGGTT CCGTCAGACC TCGTACGCGA AATCGTCGGG
5501  CGCTGGGCAC TTGGCGCTAC ACAAGTGGCC TCTGGCCTCG CACACATTCC
      GCGACCCGTG AACCGCGATG GTTTCACCGG AGACCGGAGC GTGTGTAAGG
5551  ACATCCACCG GCCGGTAGGC GCCAACCGGC TCCGTTCTTT GGTGGCCCCT
      TGTAGGTGGC CGGCCATCCG CGGTTGGCCG AGGCAAGAAA CCACCGGGGA
5601  TCGCGCCACC TTCTACTCCT CCCCTAGTCA GGAAGTTCCC CCCCGCCCGG
      AGCGCGGTGG AAGATGAGGA GGGGATCAGT CCTTCAAGGG GGGGCGGGGC
5651  CAGCTCGCGT CGTGCAAGGAC GTGACAAATG GAAGTAGCAC GTCTCACTAG
      GTCGAGCGCA GCACGTCCTG CACTGTTTAC CTTTCATCGT 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	CCGGAGGCC	GGCATTCTGC	ACGCTTCAAA
	CCGCCCCGCG	GCTTCCAGGA	GGCCTCCGGG	CCGTAAGACG	TGCGAAGTTT
5901	AGCGCACGTC	TGCCGCGCTG	TTCTCCTCTT	CCTCATCTCC	GGGCCTTTTCG
	TCGCGTGCAG	ACGGCAGCGAC	AAGAGGAGAA	GGAGTAGAGG	CCCGGAAAGC
5951	ACCAGCTTAC	CATGACCGAG	TACAAGCCCA	CGGTGCGCCT	CGCCACCCGC
	TGGTCGAATG	GTACTGGCTC	ATGTTCCGGT	GCCACGCGGA	GCGGTGGGCG
6001	GACGACGTCC	CCAGGGCCGT	ACGCACCCTC	GCCGCCGCGT	TCGCCGACTA
	CTGCTGCAGG	GGTCCCAGCA	TGCGTGGGAG	CGGCGGCGCA	AGCGGCTGAT
6051	CCCCGCCACG	CGCCACACCG	TCGATCCGGA	CCGCCACATC	GAGCGGGTCA
	GGGGCGGTGC	GCGGTGTGGC	AGCTAGGCCT	GGCGGTGTAG	CTCGCCAGT
6101	CCGAGCTGCA	AGAACTCTTC	CTCACGCGCG	TCGGGCTCGA	CATCGGCAAG
	GGCTCGACGT	TCTTGAGAAG	GAGTGCGCGC	AGCCCCAGCT	GTAGCCGTTC
6151	GTGTGGGTGC	CGGACGACGG	CGCCGCGGTG	GCGGTCTGGA	CCACGCCGGA
	CACACCCAGC	GCCTGCTGCC	GCGGCGCCAC	CGCCAGACCT	GGTGCGGCCT
6201	GAGCGTCGAA	GCGGGGGCGG	TGTTCCGCCA	GATCGGCCCG	CGCATGCGCCG
	CTCGCAGCTT	CGCCCCCGCC	ACAAGCGGCT	CTAGCCGGGC	CGGTACCGGC
6251	AGTTGAGCGG	TTCCCCGGCTG	GCCGCGCAGA	ACAGATGGAA	GGCCTCCTGG
	TCAACTCGCC	AAGGGCCGAC	CGGCGCGTCT	TGTCTACCTT	CCGGAGGACC
6301	CGCCGCACCG	GCCCAAGGAG	CCCGCGTGGT	TCCTGGCCAC	CGTCGCGTCT
	GCGGCGTGCC	CGGGTTCCTC	GGGCGCACCA	AGGACCGGTG	GCAGCGCAGA
6351	CGCCCGACCA	CCAGGGCAAG	GGTCTGGGCA	GCGCCGTCGT	GCTCCCCGGA
	GCGGGCTGGT	GGTCCCGTTC	CCAGACCCGT	CGCGGCAGCA	CGAGGGGCCT
6401	GTGGAGGCGG	CCGAGCGCGC	CGGGGTGCC	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	CCACGTA CTG	GGCGTTCGGG
6551	GGTGCCTGAC	GCCCCGCCCA	CGACCCGCAG	CGCCCGACCG	AAAGGAGCGC
	CCACGGA CTG	CGGGCGGGGT	GCTGGGCGTC	GCGGGCTGGC	TTTCCTCGCG
6601	ACGACCCCAT	GCATCGTAGA	GCTCGCTGAT	CAGCCTCGAC	TGTGCCTTCT
	TGCTGGGGTA	CGTAGCATCT	CGAGCGACTA	GTGCGAGCTG	ACACGGAAGA
6651	AGTTGCCAGC	CATCTGTTGT	TTGCCCTCC	CCCGTGCCTT	CCTTGACCCT
	TCAACGGTCG	G TAGACAACA	AACGGGGAGG	GGGCACGGAA	GGA ACTGGGA
6701	GGAAGGTGCC	ACTCCCACTG	TCCTTTCCTA	ATAAAATGAG	GAAATTGCAT
	CCTTCCACGG	TGAGGGTGAC	AGGAAAGGAT	TATTTTACTC	CTTTAACGTA
6751	CGCATTTGCT	GAGTAGGTGT	CATTCTATTC	TGGGGGGTGG	GGTGGGGCAG
	GCGTAACAGA	CTCATCCACA	GTAAGATAAG	ACCCCCCACC	CCACCCCGTC
6801	GACAGCAAGG	GGGGGGATTG	GGRAGACAAT	AGCAGGCATG	CTGGGGGGGC
	CTGTGCTTCC	CCCCCCTAAC	CCYTCTGTTA	TCGTCCGTAC	GACCCCCCGG
6851	GGTGGGGGCT	ATGGCTTCTG	AGGCGGAAAG	AACCAGCTGG	GGCTCGAGGG
	CCACCCCCGA	TACCGAAGAC	TCCGCCTTTC	TTGGTCGACC	CCGAGCTCCC
6901	CCGCCACCGC	GGTGGAGCTC	CAGCTTTTGT	TCCCTTTAGT	GAGGGTTAAT
	GGCGGTGGCG	CCACCTCGAG	GTCGAAAACA	AGGGAAATCA	CTCCCAATTA
6951	TTGAGCTTG	GCGTAATCAT	GGTCATAGCT	GTTTCCTGTG	TGAAATTGTT
	AAGCTCGAAC	CGCATTAGTA	CCAGTATCGA	CAAAGGACAC	ACTTTAACAA
7001	ATCCGCTCAC	AATTCCACAC	AACATACGAG	CCGGAAGCAT	AAAGTGTA AA
	TAGCGGAGTG	TTAAGGTGTG	TTGTATGCTC	GGCCTTCGTA	TTTCACATTT
7051	GCCTGGGGTG	CCTAATGAGT	GAGCTAACTC	ACATTAATTG	CGTTGCGCTC
	CGGACCCAC	GGATTACTCA	CTCGATTGAG	TGTAATTAAC	GCAACGCGAG
7101	ACTGCCCGCT	TTCCAGTCGG	GAAACCTGTC	GTGCCAGCAT	CGCGAGCACT
	TGACGGGCGA	AAGGTCAGCC	CTTTGGACAG	CACGGTCGTA	GCGCTCGTGA

FIG. 25(contd.) 64/93

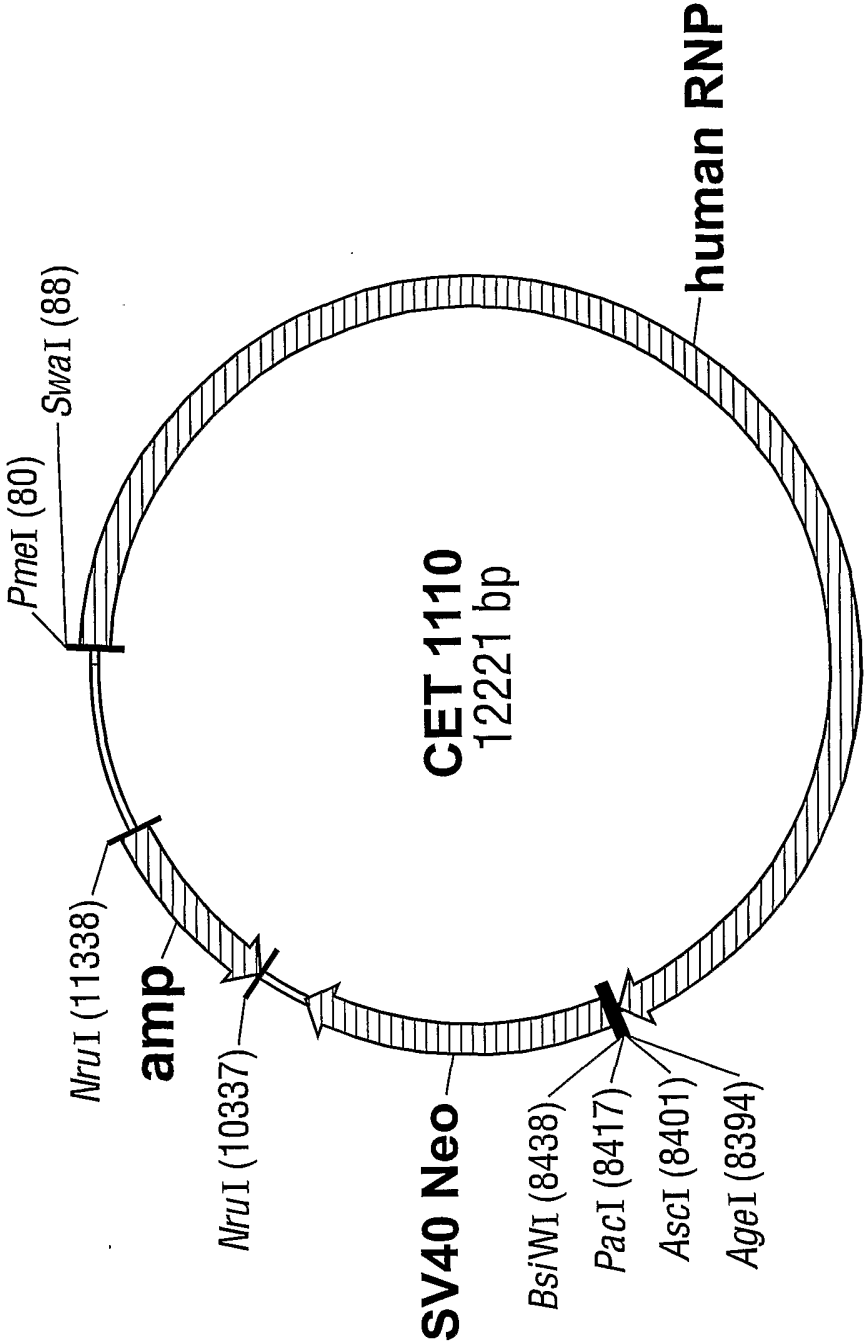
7151	TTTCGGGGAA	ATGTGCGCGG	AACCCCTATT	TGTTTATTTT	TCTAAATACA
	AAAGCCCCTT	TACACGCGCC	TTGGGGATAA	ACAAATAAAA	AGATTTATGT
7201	TTCAAATATG	TATCCGCTCA	TGAGACAATA	ACCCTGATAA	ATGCTTCAAT
	AAGTTTATAC	ATAGGCGAGT	ACTCTGTTAT	TGGGACTATT	TACGAAGTTA
7251	AATATTGAAA	AAGGAAGAGT	ATGAGTATTC	AACATTTCCG	TGTCGCCCTT
	TTATAACTTT	TTCCTTCTCA	TACTCATAAG	TTGTAAAGGC	ACAGCGGGAA
7301	ATTCCCTTTT	TTGCGGCATT	TTGCCTTCCT	GTTTTTGCTC	ACCCAGAAAC
	TAAGGGAAAA	AACGCCGTAA	AACGGAAGGA	CAAAAACGAG	TGGGTCTTTG
7351	GCTGGTGAAA	GTAAAAGATG	CTGAAGATCA	GTTGGGTGCA	CGATGGGTT
	CGACCACTTT	CATTTTCTAC	GACTTCTAGT	CAACCCAGT	GCTACCCCAA
7401	ACATCGAACT	GGATCTCAAC	AGCGGTAAGA	TCCTTGAGAG	TTTTCGCCCC
	TGTAGCTTGA	CCTAGAGTTG	TCGCCATTCT	AGGAACCTCT	AAAAGCGGGG
7451	GAAGAACGTT	TTCCAATGAT	GAGCATTCTT	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
	GAATGCCCTAC	CGTACTGTCA	TTCTCTTAAT	ACGTCACGAC	GGTATTGGTA
7651	GAGTGATAAC	ACTGCGGCCA	ACTTACTTCT	GACAACGATC	GGAGGACCGA
	CTCACTATTG	TGACGCCGGT	TGAATGAAGA	CTGTTGCTAG	CCTCCTGGCT
7701	AGGAGCTAAC	CGCTTTTTTTG	CACAACATGG	GGGATCATGT	AACTCGCCTT
	TCCTCGATTG	GCGAAAAAAC	GTGTTGTACC	CCCTAGTACA	TTGAGCGGAA
7751	GATCGTTGGG	AACCGGAGCT	GAATGAAGCC	ATACCAAACG	ACGAGCGTGA
	CTAGCAACCC	TTGGCCTCGA	CTTACTTCGG	TATGGTTTGC	TGCTCGCACT
7801	CACCACGATG	CCTGTAGCAA	TGGCAACAAC	GTTGCGCAAA	CTATTAAC TG
	GTGGTGCTAC	GGACATCGTT	ACCGTTGTTG	CAACGCGTTT	GATAATTGAC
7851	GCGAACTACT	TACTCTAGCT	TCCCGGCAAC	AATTAATAGA	CTGGATGGAG
	CGCTTGATGA	ATGAGATCGA	AGGGCCGTTG	TTAATTATCT	GACCTACCTC
7901	GCGGATAAAG	TTGCAGGACC	ACTTCTGCGC	TCGGCCCTTC	CGGCTGGCTG
	CGCCTATTTT	AACGTCCTGG	TGAAGACGCG	AGCCGGGAAG	GCCGACCGAC
7951	GTTTATTGCT	GATAAATCTG	GAGCCGGTGA	GCGTGGGTCT	CGCGGTATCA
	CAAATAACGA	CTATTTAGAC	CTCGGCCACT	CGCACCCAGA	GCGCCATAGT
8001	TTGCAGCACT	GGGGCCAGAT	GGTAAGCCCT	CCCGTATCGT	AGTTATCTAC
	AACGTCGTGA	CCCCGGTCTA	CCATTTCGGG	GGGCATAGCA	TCAATAGATG
8051	ACGACGGGGA	GTCAGGCAAC	TATGGATGAA	CGAAATAGAC	AGATCGCTGA
	TGCTGCCCCCT	CAGTCCGTTG	ATACCTACTT	GCTTTATCTG	TCTAGCGACT
8101	GATAGGTGCC	TCACTGATTA	AGCATTTGGT	ACTGTCAGAC	TCGCGACACT
	CTATCCACGG	AGTGACTAAT	TCGTAACCAT	TGACAGTCTG	AGCGCTGTGA
8151	GCATTAATGA	ATCGGCCAAC	GCGCGGGGAG	AGGCGGTTTG	CGTATTGGGC
	CGTAATTACT	TAGCCGGTTG	CGCGCCCCCT	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	TCTTGTTACAC	TCGTTTTCCG	GTCGTTTTCC
8351	CCAGGAACCG	TAAAAAGGCC	GCGTTGCTGG	CGTTTTTCCA	TAGGCTCCGC
	GGTCCTTGGC	ATTTTTCCGG	CGCAACGACC	GCAAAAAGGT	ATCCGAGGCG
8401	CCCCCTGACG	AGCATCACAA	AAATCGACGC	TCAAGTCAGA	GGTGGCGAAA
	GGGGGACTGC	TCGTAGTGTT	TTAGCTGCG	AGTTCAGTCT	CCACCGCTTT
8451	CCCGACAGGA	CTATAAAGAT	ACCAGGCGTT	TCCCCCTGGA	AGTCCCTCG
	GGGCTGTCCT	GATATTTCTA	TGGTCCGCAA	AGGGGGACCT	TCGAGGGAGC
8501	TGCGCTCTCC	TGTTCCGACC	CTGCCGCTTA	CCGGATACCT	GTCCGCCTTT
	ACGCGAGAGG	ACAAGGCTGG	GACGGCGAAT	GGCCTATGGA	CAGGCGGAAA
8551	CTCCCTTCGG	GAAGCGTGCC	GCTTTCTCAT	AGCTCACGCT	GTAGGTATCT
	GAGGGAAGCC	CTTCGCACCG	CGAAAGAGTA	TCGAGTGCGA	CATCCATAGA

65/93

```
8601 CAGTTCGGTG TAGGTCGTTC GCTCCAAGCT GGGCTGTGTG CACGAACCCC
      GTCAAGCCAC ATCCAGCAAG CGAGGTTCTGA CCCGACACAC GTGCTTGGGG
8651 CCGTTCAGCC CGACCGCTGC GCCTTATCCG GTAACATATCG 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 TTTTCTCAA CCATCGAGAA CTAGGCCGTT
8901 ACAAACCACC GCTGGTAGCG GTGGTTTTTT TGTGTGCAAG CAGCAGATTA
      TGTTTGGTGG CGACCATCGC CACCAAAAAA ACAAACGTTC GTCGTCTAAT
8951 CGCGCAGAAA AAAAGGATCT CAAGAAGATC CTTTGATCTT TTCTACGGGG
      GCGCGTCTTT TTTTCCTAGA GTTCTTCTAG GAAACTAGAA AAGATGCCCC
9001 TCTGACGCTC AGTGGAACGA AAACCTCA
      AGACTGCGAG TCACCTTGCT TTTGAGT
```

FIG. 25(contd.)

FIG. 26



67/93

FIG. 27 CET1110 nucleotide sequence

1	CGTTGTAAAA	CGACGGCCAG	TGAATTGTAA	TACGACTCAC	TATAGGGCGA
	GCAACATTTT	GCTGCCGGTC	ACTTAACATT	ATGCTGAGTG	ATATCCCGCT
51	ATTGGGTACC	GGGCCCCCCC	TCGAAGTTTA	AACATTTTAA	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	AATATTTTAA	CTGAGGGGTC	CCATTATTTG	ATCAGAGATC
251	TCTGCTCACC	TTCAACTGTA	CAATAAAGTC	TTGGTTCTTT	TGAAATAGAC
	AGACGAGTGG	AAGTTGACAT	GTTATTTTCAG	AACCAAGAAA	ACTTTATCTG
301	CTCAAATGAG	ACACCTAAAA	TTCAAAGTGT	CTTTACATTT	AAAGACACCT
	GAGTTTACTC	TGTGGATTTT	AAGTTTCACA	GAAATGTAAA	TTTCTGTGGA
351	ACAGGAAAGC	AGGTAAAAGA	GCCAGGTTAA	AAACAAATTC	TAAAACCACT
	TGTCCTTTTC	TCCATTTTCT	CGGTCCAATT	TTTGTTTAAG	ATTTTGGTGA
401	TAGCTGCAGT	TAAACATATA	GTAAAGATGC	ACTAAAGTTT	CTTACTCTGT
	ATCGACGTCA	ATTTGTATAT	CATTTCTACG	TGATTTCAAA	GAATGAGACA
451	AAATCCCTTC	CAC TTCAGGA	AATATTCCAC	TTTCCCATTC	ACTACACGTC
	TTTAGGGGAAG	GTGAAGTCCT	TTATAAGGTG	AAAGGGTAAG	TGATGTGCAG
501	GATCTAGTAC	TTTTTCCACG	ACAAATTCTT	CAGGCTCTGC	CTCTTCAACT
	CTAGATCATG	AAAAAGGTGC	TGTTTAAGAA	GTCCGAGACG	GAGAAGTTGA
551	TTTTTACTCT	TTCCATTCTG	TTTTTTTCCC	ATTTTTTTGCT	AAAATAAAAC
	AAAAATGAGA	AAGGTAAGAC	AAAAAAAGGG	TAAAAAACGA	TTTTATTTTG
601	AAAAGAGAAA	TTAAGAAATA	TTCTCTTGA	ATTTTGAGCA	CATTTTCAAG
	TTTTTCTCTT	AATTCTTTAT	AAGGAGAACT	TAAAACTCGT	GTAAAAGTTC
651	GCTCAATTGC	TTATATTATT	ATCACATTCT	ACATAAATTT	TTACTTCTAT
	CGAGTTAACG	AATATAATAA	TAGTGTAAGC	TGTATTTTAA	AATGAAGATA
701	ATCCCAGGGC	AGACACCTTC	TGGAAAGATT	AAAAGTCAAC	AGACAATAAA
	TAGGGTCCCG	TCTGTGGAAG	ACCTTTCTAA	TTTTCAGTTG	TCTGTTATTT
751	ATAAAAGAAT	GCTTTATCTT	GTTCAATTAG	TTCAAACCTA	CAACCCACCA
	TATTTTCTTA	CGAAATAGAA	CAAGTAAATC	AAGTTTGAAT	GTTGGGTGGT
801	CCAAAATAAT	ACAATAAAAA	AACACTATCT	GGAAACAGTT	ATTTTTTTTCC
	GGTTTTATTA	TGTTATTTTT	TTGTGATAGA	CCTTTGTCAA	TAAAAAAAGG
851	AGTCTTTTTT	TTTGAGACAG	GGTCTCACAC	TCTTGTCCGC	CAGGCTGGAG
	TCAGAAAAAA	AAACTCTGTC	CCAGAGTGTG	AGAACAGCGG	GTCCGACCTC
901	TGCAGTGGCG	TGATCTCAGC	TCAGTCAAC	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	TCAAACCTCT	GACCTGAAGT	GATTACACAG
	GGTACAACCTG	GTCCGACCAG	AGTTTGAGGA	CTGGACTTCA	CTAAGTGGTC
1101	CCTGGGCCTC	CCAAAGTGCT	GGCATTACAG	GCGTGAGCCA	CTGCGCCCGG
	GGACCCGGAG	GGTTTCACGA	CCGTAATGTC	CGCACTCGGT	GACGCGGGCC
1151	CCCTGTAGTC	TTAAAAGACC	AAGTTTACTA	ATTTTCACTC	ATTTTAAACA
	GGGACATCAG	AATTTTCTGG	TTCAAATGAT	TAAAAGTGAG	TAAAATTGTT
1201	CACTGCAACA	AACAACATAT	CAGGAAGTAC	CTAAAGGGTG	ATCCAGAGAA
	GTGACGTTGT	TTGTTGATAC	GTCTTCATG	GATTTCCAC	TAGGTCCTCT
1251	GCAAGTAGTA	GTGACAGGTC	TTAGGTGAAC	CTATGACAGA	CCTTGTATCC
	CGTTCATCAT	CACTGTCCAG	AATCCACTTG	GATACTGTCT	GGAACATAGG
1301	ACCCCCAGAT	GGTAAAAGCC	CCAGCCCCCT	TCTCAATTCA	AATATTAATG
	TGGGGGTCTA	CCATTTTTCG	GGTCGGGGGA	AGAGTTAAGT	TTATAATTAC

FIG. 27(contd.)

68/93

1351	TCAAAAGCAT	CAATGATACA	GAGAAAAGAT	AAATGCAGAA	TGAAAACATG
	AGTTTTTCGTA	GTTACTATGT	CTCTTTTCTA	TTTACGTCTT	ACTTTTGTAC
1401	GTTCAAAATC	CTGATACCAA	CTGCAGGGTC	AACTATAGAG	ACCACTAGGA
	CAAGTTTTTAG	GACTATGGTT	GACGTCCCAG	TTGATATCTC	TGGTGATCCT
1451	GGTTC AATTA	AAGGACAAGA	TTATTTTTC	ATAATCTCTG	TAGATAATAT
	CCAAGTTAAT	TTCCTGTTCT	AATAAAAAGG	TATTAGAGAC	ATCTATTATA
1501	TTCTTACCAC	TTAGAACAAA	ACTATAAAGC	TATCACTTCA	AGAGACCAAC
	AAGGATGGTG	AATCTTGTTT	TGATATTTTCG	ATAGTGAAGT	TCTCTGGTTG
1551	ATTACAAATT	TATTTTAATT	CCCTAAGGTG	AAAAAATCC	TTCTTTCCTG
	TAATGTTTAA	ATAAAATTAA	GGGATTCAC	TTTTTTTAGG	AAGGAAGGAC
1601	GTTTCTCAAG	AGAAAGTCTA	TACTGGTAAC	CAAATTCACT	TTAAACAGGC
	CAAAGAGTTC	TCTTTCAGAT	ATGACCATTG	GTTTAAAGTGA	AATTTGTCCG
1651	ATTTTCTTTG	GTATGACACT	ATTTAAGAGA	AGCAGGAAAC	CAACGTGAAC
	TAAAAGAAAC	CATACTGTGA	TAAATTCTCT	TCGTCTTTTG	GTGCACTTG
1701	CAGCTCTTTC	CAATGGCTCA	AGATTTCCCTA	TGAGAGGACT	AAAAATGGGG
	GTCGAGAAAG	GTTACCGAGT	TCTAAAGGAT	ACTCTCCTGA	TTTTTACCCC
1751	AAAATTTTTTA	TGAGAGGATT	AAAAATGGGG	GAAAAAAAC	CCTGAAATGG
	TTTTAAAAAT	ACTCTCCTAA	TTTTTACCCC	CTTTTTTTTG	GGACTTTACC
1801	TTAATCAGAA	GATCCTATGG	GCTGAGAAGG	AATCCATCTT	AACATTTTAT
	AATTAGTCTT	CTAGGATACC	CGACTCTTCC	TTAGGTAGAA	TTGTAAAGTA
1851	CTTAAAGCAA	ATGCTATTGC	CGGGGCGAGT	GGCTCATGCC	TGTAATCCCA
	GAATTTTCGTT	TACGATAACG	GCCCCCGTCA	CCGAGTACGG	ACATTAGGGT
1901	GCACCTTTGGG	AGGCCGAGGT	GGGCAGATCA	TCTGAGGTCA	GGAGTTTGAG
	CGTGAAACCC	TCCGGCTCCA	CCCGTCTAGT	AGACTCCAGT	CCTCAAACTC
1951	ACCAGCCTGA	CCAACATGGA	GAAACCCCGT	TTCTACTAAA	AATACAAAAT
	TGGTCGGACT	GGTTGTACCT	CTTTGGGGCA	AAGATGATTT	TTATGTTTTTA
2001	TAGCCAGGCA	TAGTGGTGCA	TGCCTGTAAT	CCCAGCTACT	TGGGAGGCTG
	ATCGGTCCGT	ATCACCACGT	ACGGACATTA	GGGTCGATGA	ACCCTCCGAC
2051	AGGCAGGAGA	ACTGCTTGAA	CCCAGGAGGC	TTAAGTTGCG	GTGAGCCAAG
	TCCGTCCTCT	TGACGAACTT	GGGTCCTCCG	AATTCAACGC	CACTCGGTTT
2101	ATCACGCCAT	TGCACTCTAG	CCTGGACAAC	AAGAGAAAAA	CTCTGTCTCA
	TAGTGCGGTA	ACGTGAGATC	GGACCTGTTG	TTCTCTTTTT	GAGACAGAGT
2151	AAAAAACACA	AAAACAAAAA	ACCCAAATAC	TATTTAAAAA	AGATAAACCT
	TTTTTTGTGT	TTTTGTTTTT	TGGGTTTATG	ATAAATTTTT	TCTATTTGGA
2201	TAATTGCTCA	ATCATTAAG	CCATCCACA	AGTAAAGCAG	CAAGCAGAAA
	ATTAACGAGT	TAGTAATTTT	GGTAGGGTGT	TCATTTTCGTC	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	TTACATAGT
2601	GATACCTATA	CTACATGAAA	AACAAAACAG	CTACTGCCAC	AAAGGGAGAA
	CTATGGATAT	GATGTACTTT	TTGTTTGTGTC	GATGACGGTG	TTTCCCTCTT
2651	GGCTAACAAA	ATAAAGTCAA	CAATAAATAC	AGAAAATGAA	AAGGATACAC
	CCGATTGTTT	TATTTTCAGTT	GTTATTTTATG	TCTTTTACTT	TTCTATGTG
2701	ACTAAGGTTT	ACAAAAAATA	AAAGGCAGAC	AAAATGCCAT	ACAGTATTCA
	TGATTCCAAA	TGTTTTTTTTT	TTTCCGTCTG	TTTTACGGTA	TGTCATAAGT
2751	TTCACTACTA	TGGCATTCAT	AAGCTAGTTT	CAAATGCTCA	CTATTTTCTT
	AAGTGATGAT	ACCGTAAGTA	TTGATCAAAA	GTTTACGAGT	GATAAAAGAA

FIG. 27(contd.)

69/93

2801	TTATAGTATA	TATTTGCCTT	AACCCAGCAC	TTTTTTCCAA	AAGTGGATGA
	AATATCATAT	ATAAACGGAA	TTGGGTCGTG	AAAAAAGGTT	TTCACCTACT
2851	GTCAAAATAA	ATTTCCCATT	ATTTAAGTGA	AATTAACAGC	ACACATATCT
	CAGTTTTTATT	TAAAGGGTAA	TAAATTCACT	TTAATTGTCG	TGTGTATAGA
2901	CACAACACTA	ATGAATTTTT	AAAATGGAAA	GTTAAGAACT	TTTAAAGTGG
	GTGTTGTGAT	TACTTAAAAA	TTTTACCTTT	CAATTCTTGA	AAATTTTCACC
2951	CCAACCTGTG	ATCCTTCACA	AAATAAACTA	AATACAATAA	CAGACCCCAA
	GGTTGGACAC	TAGGAAGTGT	TTTATTTGAT	TTATGTTATT	GTCTGGGGTT
3001	AGGCTATCAA	TTGCGTGCAA	AAACAACCTC	TGTTTTCCAG	GGTAAACAGA
	TCCGATAGTT	AACGCACGTT	TTTGTTGAAG	ACAAAAGGTC	CCATTTGTCT
3051	ATCTAATGCA	GAATCTAATG	CAGGGTAAAC	AGACTTAATG	CAGAATCTAA
	TAGATTACGT	CTTAGATTAC	GTCCCATTTC	TCTGAATTAC	GTCTTAGATT
3101	TGATGGCACA	AATTAAAAAT	CACTAACGTG	CCCTTTTTAG	TGTGAAACCC
	ACTACCGTGT	TTAATTTTTA	GTGATTGCAC	GGGAAAAATC	ACACTTTGGG
3151	AGAGAGAGCA	CATACAAGCC	AAAAACAAAT	GCTTTATTTT	ACCTAGGAGA
	TCTCTCTCGT	GTATGTTCCG	TTTTTGTTTA	CGAAATAAAA	TGGATCCTCT
3201	CATTAAACATT	CACCTTTACG	TGTTTAAGAT	TAATGCAATG	TTAAATATTG
	GTAATTGTAA	GTGGAAATGC	ACAAATTCTA	ATTACGTTAC	AATTTATAAC
3251	TGAAAACGTG	AACCTTTGAAT	TTCATGATTT	TTATGTGAAT	ATTCCAGGGT
	ACTTTTGACA	TTGAAACTTA	AAGTACTAAA	AATACACTTA	TAAGGTCCCA
3301	TTAAAAAAAC	TTGTAACATG	ACATGGCTGA	ATAAGATAAA	AAAAAAATCT
	AATTTTTTTT	AACATTGTAC	TGTACCGACT	TATTCATATT	TTTTTTTTAG
3351	AGCCTTTTCT	CCCTTCTGGC	TCATATTTCG	GATTTTCGAT	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	AATTTAAAGG
3801	CCTGTAATTT	ACACTGGGAG	AGCTGGGAAA	TATGTGGATG	TAAATTTCTC
	GGACATTAAA	TGTGACCCTC	TCGACCCTTT	ATACACCTAC	ATTTAAAGAG
3851	AGCCACAGAG	ATGCAAAGTT	ATACTGTGGG	GAAAAAAAC	TTGAGTTAAA
	TCGGTGTCTC	TACGTTTCAA	TATGACACCC	CTTTTTTTTT	AACTCAATTT
3901	TCCTTACATA	TTTTAGGTTT	TCATTAACCT	ACCAATGTAG	TTTTGTTGGA
	AGGAATGTAT	AAAATCCAAA	AGTAATTGAA	TGGTTACATC	AAAACAACCT
3951	GGCCATTTTT	TTTATTGCAG	ACTTGAAGAG	CTATTACTAG	AAAAATGCAT
	CCGGTAAAAA	AAATAACGTC	TGAACCTCTC	GATAATGATC	TTTTTACGTA
4001	GACAGTTAAG	GTAAGTTTGC	ATGACACAAA	AAAGGTAAC	AAATACAAAT
	CTGTCAATTC	CATTCAAACG	TACTGTGTTT	TTTCCATTGA	TTTATGTTTA
4051	TCTGTTTGGA	TTCCAACCCC	CAAGTAGAGA	GCGCACACTT	TCAAACGTGA
	AGACAAACCT	AAGGTTGGGG	GTTTCATCTC	CGCGTGTGAA	AGTTTGCAC
4101	ATACAAATCC	AGAGTAGATC	TGCGCTCCTA	CCTACATTGC	TTATGATGTA
	TATGTTTAGG	TCTCATCTAG	ACGCGAGGAT	GGATGTAACG	AATACTACAT
4151	CTTAAGTACG	TGTCCTAACC	ATGTGAGTCT	AGAAAGACTT	TACTGGGGAT
	GAATTCATGC	ACAGGATTGG	TACATCAGA	TCTTTCTGAA	ATGACCCCTA
4201	CCTGGTACCT	AAAACAGCTT	CACATGGCTT	AAAATAGGGG	ACCAATGTCT
	GGACCATGGA	TTTTGTGCGA	GTGTACCGAA	TTTTATCCCC	TGGTTACAGA

FIG. 27(contd.) 70/93

```

4251   TTTCCAATCT AAGTCCCATF TATAATAAAG TCCATGTTCC ATTTTAAAG
      AAAGGTTAGA TTCAGGGTAA ATATTATTTT AGGTACAAGG TAAAAATTTT
4301   GACAATCCTT TCGGTTTAAA ACCAGGCACG ATTACCCAAA CAACTCACAA
      CTGTTAGGAA AGCCAAATTT TGGTCCGTGC TAATGGGTTT GTTGAGTGTT
4351   CCGTAAAGCA CTGTGAATCT TCTCTGTTCT GCAATCCCAA CTGGGTTTCT
      GCCATTTCTG GACACTTAGA AGAGACAAGA CGTTAGGGTT GAACCAAAGA
4401   GCTCAGAAAC CCTCCCTCTT TCCAATCGGT AATTAAATAA CAAAAGGAAA
      CGAGTCTTTG GGAGGGAGAA AGGTAGCCA TTAATTTATT GTTTTCCTTT
4451   AAACCTAAGA TGCTTCAACC CCGTTTCGTG ACACTTGAA AAAAGAATCA
      TTTGAATTCT ACGAAGTTGG GGCAAAGCAC TGTGAAACTT TTTTCTTAGT
4501   CCTCTTGCAA ACACCCGCTC CCGACCCCGG CCGCTGAAGC CCGGCGTCCA
      GGAGAACGTT TGTGGGCGAG GGCTGGGGGC GCGGACTTCG GGCCGCAAGT
4551   GAGGCCTAAG CGCGGGTGCC CGCCCCCACC CGGGAGCGCG GGCCTCGTGG
      CTCCGGATTG GCGCCACGCG GCGGGGGTGG GCCCTCGCGC CCGGAGCACC
4601   TCAGCGCATC CGCGGGGAGA AACAAAGGCC GCGGCACGGG GGCTCAAGGG
      AGTCGCGTAG GCGCCCTCTT TTGTTTCCGG CGCCGTGCCC CCGAGTTCCC
4651   CACTGCGCCA CACCGCACGC GCGTACCCCG GCGCGGCCAC GTTAACTGGC
      GTGACGCGGT GTGGCGTGCG CGGATGGGGG CGCGCCGGTG CAATTGACCG
4701   GGTCGCGCGA GCCTCGGGAC AGCCGGCCCG 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   GAGGGGAAA TGCGCTTTGT CTCGAAATGG GGCAACCGTC GCCACAGCTC
      CTCCCCCTTT ACGCGAAACA GAGCTTTACC CCGTTGGCAG CGGTGTGAG
4901   CCTACCCCTT CGAGGGCAGA GCAGTCCCCC CACTAACTAC CGGGCTGGCC
      GGATGGGGGA GCTCCCGTCT CGTCAGGGGG GTGATTGATG GCCCGACCGG
4951   GCGCGCCAGG CCAGCCGCGA GGCCACCGCC CGACCCTCCA CTCCTTCCCC
      CGCGCGGTCC GGTGCGCGCT CCGGTGGCGG GCTGGGAGGT GAGGAAGGGC
5001   CAGCTCCCGG CGCGGGGTCC GCGGAGAAGG GGAGGGGAGG GGAGCGGAGA
      GTCGAGGGGC GCGCCCCAGG CCGCTCTTCC CCTCCCCTCC CTCGCTCT
5051   ACCGGGCCCC CGGGACGCGT GTGGCATCTG AAGCACCACC AGCGAGCGAG
      TGGCCCGGGG GCCCTGCGCA CACCGTAGAC TTCGTGGTGG TCGCTCGCTC
5101   AGCTAGAGAG AAGGAAAGCC ACCGACTTCA CCGCTCCGA GCTGCTCCGG
      TCGATCTCTC TTCCTTTCGG TGGCTGAAGT GGCGGAGGCT CGACGAGGCC
5151   GTCGCGGGTC TGCAGCGTCT CCGGCCCTCC GCGCTACAG CTCAAGCCAC
      CAGCGCCAG ACGTCGCAGA GGCCGGGAGG CGCGGATGTC GAGTTCGGTG
5201   ATCCGAAGGG GGAGGGAGCC GGGAGCTGCG CGCGGGGCCG CCGGGGGGAG
      TAGGCTTCCC CCTCCCTCGG CCTTCGACGC GCGCCCCGGC GGCCCCCTC
5251   GGGTGGCACC GCCACGCGG GCGGGCCACG AAGGGCGGGG CAGCGGGCGC
      CCCACCGTGG CGGGTGCGGC CCGCCGGTGC TTCCCGCCCC GTCGCCCCGC
5301   GCGCGCGGCG GGGGGAGGGG CCGGCGCCGC GCCCGCTGGG AATTGGGGCC
      CGCGCGCCGC CCCCCTCCCC GGCCGCGGCG CGGGCGACCC TTAACCCCGG
5351   CTAGGGGGAG GCGGAGGCG CCGACGACCG CGGCACTTAC CGTTCGCGGC
      GATCCCCCTC CCGCCTCCGC GGCTGCTGGC GCCGTGAATG GCAAGCGCCG
5401   GTGGCGCCCG GTGGTCCCCA AGGGGAGGGA AGGGGGAGGC GGGGCGAGGA
      CACCGCGGGC CACAGGGGT TCCCCTCCCT TCCCCCTCCG CCGGCTCCT
5451   CAGTGACCGG AGTCTCCTCA GCGGTGGCTT TTCTGCTTGG CAGCCTCAGC
      GTCAGTGGCC TCAGAGGAGT CGCCACCGAA AAGACGAACC GTCGGAGTCG
5501   GGCTGGCGCC AAAACCGGAC TCCGCCACT TCCTCGCCCC CCGGTGCGAG
      CCGACCGCGG TTTTGGCCTG AGGCGGGTGA AGGAGCGGGC GGCCACGCTC
5551   GGTGTGGAAT CCTCCAGACG CTGGGGGAGG GGGAGTTGGG AGCTTAAAAA
      CCACACCTTA GGAGGTCTGC GACCCCTCC CCCTCAACCC TCGAATTTTT
5601   CTAGTACCCC TTTGGGACCA CTTTCAGCAG CGAACTCTCC TGTACACCAG
      GATCATGGGG AAACCCTGGT GAAAGTCGTC GCTTGAGAGG ACATGTGGTC
5651   GGGTCAGTTC CACAGACGCG GGCCAGGGGT GGGTCATTGC GGCCTGAACA
      CCCAGTCAAG GTGTCTGCGC CCGGTCCCCA CCCAGTAACG CCGCACTTGT

```

FIG. 27(contd.)

71/93

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	GCCCCCTCCG	TCGGCGTTAA
	TCCCCCGACC	TAGAGGTCCG	ACGCCTAATT	CGGGGAGGGC	AGCCGCAATT
5901	TTTCAAAC TG	CGCGACGTTT	CTCACCTGCC	TTCGCCAAGG	CAGGGGCCGG
	AAAGTTTGAC	GCGCTGCAAA	GAGTGGACGG	AAGCGGTTCC	GTCCCCGGCC
5951	GACCCTATTG	CAAGAGGTAG	TAAGTAGCAG	GACTCTAGCC	TTCCGCAATT
	CTGGGATAAG	GTTCTCCATC	ATTGATCGTC	CTGAGATCGG	AAGGCGTTAA
6001	CATTGAGCGC	ATTTACGGAA	GTAACGTCGG	GTACTGTCTC	TGGCCGCAAG
	GTAACGTCGC	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	GTCGCGGCC	AGGGCACGCC	TCCACGAGGA	GCGTCTCAAC
6251	TTTCTCGAGC	AGCGGCAGTT	CTCACTACAG	CGCCAGGACG	AGTCCGGTTC
	AAAGAGCTCG	TCGCCGTCAA	GAGTGATGTC	GCGGTCCTGC	TCAGGCCAAG
6301	GTGTTTCGTCC	GCGGAGATCT	CTCTCATCTC	GCTCGGCTGC	GGGAAATCGG
	CACAAGCAGG	CGCCTCTAGA	GAGAGTAGAG	CGAGCCGACG	CCCTTTAGCC
6351	GCTGAAGCGA	CTGAGTCCGC	GATGGAGGTA	ACGGGTTTGA	AATCAATGAG
	CGACTTCGCT	GACTCAGGCG	CTACCTCCAT	TGCCCAAAC	TTAGTTACTC
6401	TTATTGAAAA	GGGCATGGCG	AGGCCGTTGG	CGCCTCAGTG	GAAGTCGGCC
	AATAACTTTT	CCCGTACCGC	TCCGGCAACC	GCGGAGTCAC	CTTCAGCCGG
6451	AGCCGCCTCC	GTGGGAGAGA	GGCAGGAAAT	CGGACCAATT	CAGTAGCAGT
	TCGGCGGAGG	CACCCTCTCT	CCGTCCCTTA	GCCTGGTTAA	GTCATCGTCA
6501	GGGGCTTAAG	GTTTATGAAC	GGGGTCTTGA	GCGGAGGCCT	GAGCGTACAA
	CCCCGAATTC	CAAATACTTG	CCCCAGAACT	CGCCTCCGGA	CTCGCATGTT
6551	ACAGCTTCCC	CACCCTCAGC	CTCCCGGCGC	CATTTCCTTT	CACTGGGGGT
	TGTCGAAGGG	GTGGGAGTCG	GAGGGCCGCG	GTAAAGGGAA	GTGACCCCCA
6601	GGGGGATGGG	GAGCTTTTAC	ATGGCGGACG	CTGCCCCGCT	GGGGTGAAAG
	CCCCCTACCC	CTCGAAAGTG	TACCGCCTGC	GACGGGGCGA	CCCCACTTTC
6651	TGGGGCGCGG	AGGCGGGAAT	TCTTATTCCC	TTTCTAAAGC	ACGCTGCTTC
	ACCCCGCGCC	TCCGCCCTTA	AGAATAAGGG	AAAGATTTTC	TGCGACGAAG
6701	GGGGGCCACG	GCGTCTCCTC	GGCGAGCGTT	TCGGCGGGCA	GCAGGTCCTC
	CCCCCGGTGC	CGCAGAGGAG	CCGCTCGCAA	AGCCGCCCGT	CGTCCAGGAG
6751	GTGAGCGAGG	CTGCGGAGCT	TCCCTTCCCC	CTCTCTCCCG	GGAACCGATT
	CACTCGCTCC	GACGCCTCGA	AGGGGAGGGG	GAGAGAGGGC	CCTTGGCTAA
6801	TGGCGGCCCG	CATTTTCATG	GCTCGCCTTC	CTCTCAGCGT	TTTCTTATA
	ACCGCCGGCG	GTAAAAGTAC	CGAGCGGAAG	GAGAGTCGCA	AAAGGAATAT
6851	ACTCTTTTAT	TTTCTTAGTG	TGCTTTCTCT	ATCAAGAAGT	AGAAGTGGTT
	TGAGAAAATA	AAAGAATCAC	ACGAAAGAGA	TAGTTCCTCA	TCTTCACCAA
6901	AACTATTTT	TTTTTCTTCT	CGGGCTGTTT	TCATATCGTT	TCGAGGTGGA
	TTGATAAAAA	AAAAAGAAGA	GCCCGACAAA	AGTATAGCAA	AGCTCCACCT
6951	TTTGAGAGTG	TTTGTGAGCT	TGGATCTTTA	GAGTCCTGCG	CACCTCATTA
	AAACCTCACA	AAACACTCGA	ACCTAGAAAT	CTCAGGACGC	GTGGAGTAAT
7001	AAGGCGCTCA	GCCTTCCCC	CGATGAAATG	GCGCCATTGC	GTTTCGGAAGC
	TTCCGCGAGT	CGGAAGGGGA	GCTACTTTAC	CGCGGTAACG	CAAGCCTTCG
7051	CACACCGAAG	AGCGGGGAGG	GGGGGTGCTC	CGGGTTTGCG	GGCCCGGTTT
	GTGTGGCTTC	TCGCCCCCTC	CCCCACGAG	GCCCAAACGC	CCGGGCCAAA
7101	CAGAGAAGAT	ATCACCACCC	AGGGCGTCGG	GCCGGGTTCA	ATGCGAGCCG
	GTCTCTTCTA	TAGTGGTGGG	TCCCGCAGCC	CGGCCCAAGT	TACGCTCGGC

FIG. 27(contd.)

72/93

7151	TAGGACAAAG	AAACCATTTT	ATGTTTTTTC	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	GGGATATTTC	CGTCAGAATC	CTTTACTGTT
	CACCAAAAGA	CCCGCCCCCT	CCCTATAAGC	GCAGTCTTAG	GAAATGACAA
7351	CTTAAGGATT	CCGTTTAAGT	TGTAGAGCTG	ACTCATTTTA	AGTAATGTTG
	GAATTCTTAA	GGCAAATTCA	ACATCTCGAC	TGAGTAAAT	TCATTACAAC
7401	TTACTGAGAA	GTTTAACCTT	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	TAACCTACACG
7551	AAAACGTGTT	TGGGCCTGGG	CCGCTCTGTA	TTTGAACCTT	GTTACTTTTC
	TTTTTGACAAA	ACCCGGACCC	GGCGAGACAT	AAACTTGAAA	CAATGAAAAG
7601	TCATTTTGTT	TGCAATCTTG	GTTGAACATT	ACATTGATAA	GCATAAGGTC
	AGTAAACAA	ACGTTAGAAC	CAACTTGTA	TGTAACCTATT	CGTATTCCAG
7651	TCAAGCGAAG	GGGGTCTACC	TGGTTATTTT	TCTTTGACCC	TAAGCACGTT
	AGTTTCGCTTC	CCCCAGATGG	ACCAATAAAA	AGAACTGGG	ATTTCGTGCAA
7701	TATAAAATAA	CATTGTTTAA	AATCGATAGT	GGACATCGGG	TAAGTTTGGA
	ATATTTTATT	GTAACAAAT	TTAGCTATCA	CCTGTAGCCC	ATTCAAACCT
7751	TAAATTGTGA	GGTAAGTAAT	GAGTTTTTGC	TTTTTGTTAG	TGATTTGTAA
	ATTTAACACT	CCATTTCATTA	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	ATTTTCAGTTA	ATTTTTTCTT	TAATTTTTTAC	GAGTCATGAA
	ATCAAACCTAG	TAAAGTCAAT	TAAAAAAGAA	ATTAAAAATG	CTCAGTACTT
8001	AACTACAGGA	AAAAAAATCT	GAAGTGGGTT	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	TATTAACATA	CTTTTAATGC
8151	TGTGACTTTT	TAGTTTGCAT	TATGCGTTTG	TAGCAGTTGG	TCCTGGATAT
	ACACTGAAAA	ATCAAACGTA	ATACGCAAAC	ATCGTCAACC	AGGACCTATA
8201	CACTTTCTCT	CGTTTGAGGT	TTTTTAACCT	AGTTAACTTT	TAAGACAGGT
	GTGAAAGAGA	GCAAACCTCA	AAAAATTGGA	TCAATTGAAA	ATTCTGTCCA
8251	TTCCTTAACA	TTCATAAGTG	CCCAGAATAC	AGCTGTGTAG	TACAGCATAT
	AAGGAATTGT	AAGTATTCAC	GGGTCTTATG	TCGACACATC	ATGTCGTATA
8301	AAAGATTTC	GCTCTGAGGT	TTTTTCCTATT	GAAGTTGAAA	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	TCAATCCAC	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.)

73/93

8601	AAAGCATGCA	TCTCAATTAG	TCAGCAACCA	TAGTCCCGCC	CCTAACTCCG
	TTTCGTACGT	AGAGTTAATC	AGTCGTTGGT	ATCAGGGCGG	GGATTGAGGC
8651	CCCATCCCGC	CCCTAACTCC	GCCCAGTTCC	GCCCATTTCTC	CGCCCCATGG
	GGGTAGGGCG	GGGATTGAGG	CGGGTCAAGG	CGGGTAAGAG	GCGGGGTACC
8701	CTGACTAATT	TTTTTTTATT	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	TAAGTTGTTC
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	TGAAGTCAA	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	GACGATAAACC	CGCTTCACGG	CCCCGTCCTA	GAGGACAGTA
9151	CTCACCTTGC	TCCTGCCGAG	AAAGTATCCA	TCATGGCTGA	TGCAATGCGG
	GAGTGGAACG	AGGACGGCTC	TTTCATAGGT	AGTACCGACT	ACGTTACGCC
9201	CGGCTGCATA	CGCTTGATCC	GGCTACCTGC	CCATTCGACC	ACCAAGCGAA
	GCCGACGTAT	GCGAACTAGG	CCGATGGACG	GGTAAGCTGG	TGGTTGCTTT
9251	ACATCGCATC	GAGCGAGCAC	GTACTCGGAT	GGAAGCCGGT	CTTGTGATC
	TGTAGCGTAG	CTCGCTCGTG	CATGAGCCTA	CCTTCGGCCA	GAACAGCTAG
9301	AGGATGATCA	AGAGCATCAG	GGGCTCGCGC	CAGCCGAAC	GTTCCGCCAGG
	TCCTACTAGT	TCTCGTAGTC	CCCGAGCGCG	GTCCGCTTGA	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	TCTCGAACC	CCGCTTACCC	GACTGGCGAA
9551	CCTCGTGCTT	TACGGTATCG	CCGCTCCCGA	TTCCGAGCGC	ATCGCCTTCT
	GGAGCACGAA	ATGCCATAGC	GGCGAGGGCT	AAGCGTCGCG	TAGCGGAAGA
9601	ATCGCCTTCT	TGACGAGTTC	TTCTGAGCGG	GACTCTGGGG	TTCGAAATGA
	TAGCGGAAGA	ACTGCTCAAG	AAGACTCGCC	CTGAGACCCC	AAGCTTTACT
9651	CCGACCAAGC	GACGCCCCAAC	CTGCCATCAC	GAGATTTTGA	TTCCACCGCC
	GGCTGGTTTC	CTGCGGGTTG	GACGGTAGTG	CTCTAAAGCT	AAGGTGGCGG
9701	GCCTTCTATG	AAAGGTTGGG	CTTCGGAATC	GTTTTCCGGG	ACGCCGGCTG
	CGGAAGATAC	TTTCCAACCC	GAAGCCTTAG	CAAAAGGCCC	TGCGGCCGAC
9751	GATGATCCTC	CAGCGCGGGG	ATCTCATGCT	GGAGTTCCTC	GCCCACCCCTA
	CTACTAGGAG	GTGCGGCCCC	TAGAGTACGA	CCTCAAGAAG	CGGGTGGGAT
9801	GGGGGAGGCT	AACTGAAACA	CGGAAGGAGA	CAATACCGGA	AGGAACCCGC
	CCCCCTCCGA	TTGACTTTGT	GCCTTCCTCT	GTTATGGCCT	TCCTTGGGCG
9851	GCTATGACGG	CAATAAAAAG	ACAGAATAAA	ACGCACGGTG	TTGGGTCTGT
	CGATACTGCC	GTTATTTTTT	TGTCTTATTT	TGCGTGCCAC	AACCCAGCAA
9901	TGTTTCATAA	CGCGGGGTTT	GGTCCCAGGG	CTGGCACTCT	GTGATAACCC
	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	GTTGTCAGCC

FIG. 27(contd.)

74/93

10051	GGCGGCAGGC	CCTGCCATAG	CCTCAAATTC	CTTAGGCTCG	AGGGCCGCCA
	CCGCCGTCCG	GGACGGTATC	GGAGTTTAAG	GAATCCGAGC	TCCCGCGCGT
10101	CCGCGGTGGA	GCTCCAGCTT	TTGTTCCCTT	TAGTGAGGGT	TAATTTTCGAG
	GGCGCCACCT	CGAGGTCGAA	AACAAGGGAA	ATCACTCCCA	ATTAAAGCTC
10151	CTTGCGTAA	TCATGGTCAT	AGCTGTTTCC	TGTGTGAAAT	TGTTATCCGC
	GAACCGCATT	AGTACCAGTA	TCGACAAAGG	ACACACTTTA	ACAATAGGCG
10201	TCACAATTCC	ACACAACATA	CGAGCCGGAA	GCATAAAGTG	TAAAGCCTGG
	AGTGTTAAGG	TGTGTTGTAT	GCTCGGCCTT	CGTATTTTAC	ATTTTCGGACC
10251	GGTGCCCTAAT	GAGTGAGCTA	ACTCACATTA	ATTGCGTTGC	GCTCACTGCC
	CCACGGATTA	CTCACTCGAT	TGAGTGTAAT	TAACGCAACG	CGAGTGACGG
10301	CGCTTTCCAG	TCGGGAAACC	TGTCGTGCCA	GCATCGCGAG	CACTTTTTCGG
	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	TCCGTGTTCG	CCTTATTCCT
	CTTTTTCTCT	CTCATACTCA	TAAGTTGTAA	AGGCACAGCG	GGAATAAGGG
10501	TTTTTTGCGG	CATTTTGCTT	TCCTGTTTTT	GCTCACCCAG	AAACGCTGGT
	AAAAAACGCC	GTAAAAACGA	AGGACAAAAA	CGAGTGGGTC	TTTGCGACCA
10551	GAAAGTAAAA	GATGCTGAAG	ATCAGTTGGG	TGCACGAGTG	GGTTACATCG
	CTTTCATTTT	CTACGACTTC	TAGTCAACCC	ACGTGCTCAC	CCAATGTAGC
10601	AACTGGATCT	CAACAGCGGT	AAGATCCTTG	AGAGTTTTTC	CCCCGAAGAA
	TTGACCTAGA	GTTGTCGCCA	TTCTAGGAAC	TCTCAAAGC	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	CGGTGAATG	AAGACTGTTG	CTAGCCTCCT	GGCTTCCTCG
10901	TAACCGCTTT	TTTGACACAAC	ATGGGGGATC	ATGTAACCTG	CCTTGATCGT
	ATTGGCGAAA	AAACGTGTTG	TACCCCTAG	TACATTGAGC	GGAACCTAGCA
10951	TGGGAACCGG	AGCTGAATGA	AGCCATACCA	AACGACGAGC	GTGACACCAC
	ACCCTTGGCC	TCGACTTACT	TCGGTATGGT	TTGCTGCTCG	CACTGTGGTG
11001	GATGCCTGTA	GCAATGGCAA	CAACGTTGCG	CAAACCTATTA	ACTGGCGAAC
	CTACGGACAT	CGTTACCGTT	GTTGCAACGC	GTTTGATAAT	TGACCGCTTG
11051	TACTTACTCT	AGCTTCCCGG	CAACAATTA	TAGACTGGAT	GGAGGCGGAT
	ATGAATGAGA	TCGAAGGGCC	GTTGTTAATT	ATCTGACCTA	CCTCCGCCTA
11101	AAAGTTGCAG	GACCACTTCT	GCGCTCGGCC	CTTCCGGCTG	GCTGGTTTTAT
	TTTCAACGTC	CTGGTGAAGA	CGCGAGCCGG	GAAGGCCGAC	CGACCAAATA
11151	TGCTGATAAA	TCTGGAGCCG	GTGAGCGTGG	GTCTCGCGGT	ATCATTGACG
	ACGACTATTT	AGACCTCGGC	CACTCGCACC	CAGAGCGCCA	TAGTAACGTC
11201	CACTGGGGCC	AGATGGTAAG	CCCTCCCGTA	TCGTAGTTAT	CTACACGACG
	GTGACCCCGG	TCTACCATTTC	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	CGGTGCTTCG	GCTGCGGGCA
	GGCGAAGGAG	CGAGTGACTG	AGCGACGCGA	GCCAGCAAGC	CGACGCCGCT
11451	GCGGTATCAG	CTCACTCAAA	GGCGGTAATA	CGGTTATCCA	CAGAATCAGG
	CGCCATAGTC	GAGTGAGTTT	CCGCCATTAT	GCCAATAGGT	GTCTTAGTCC

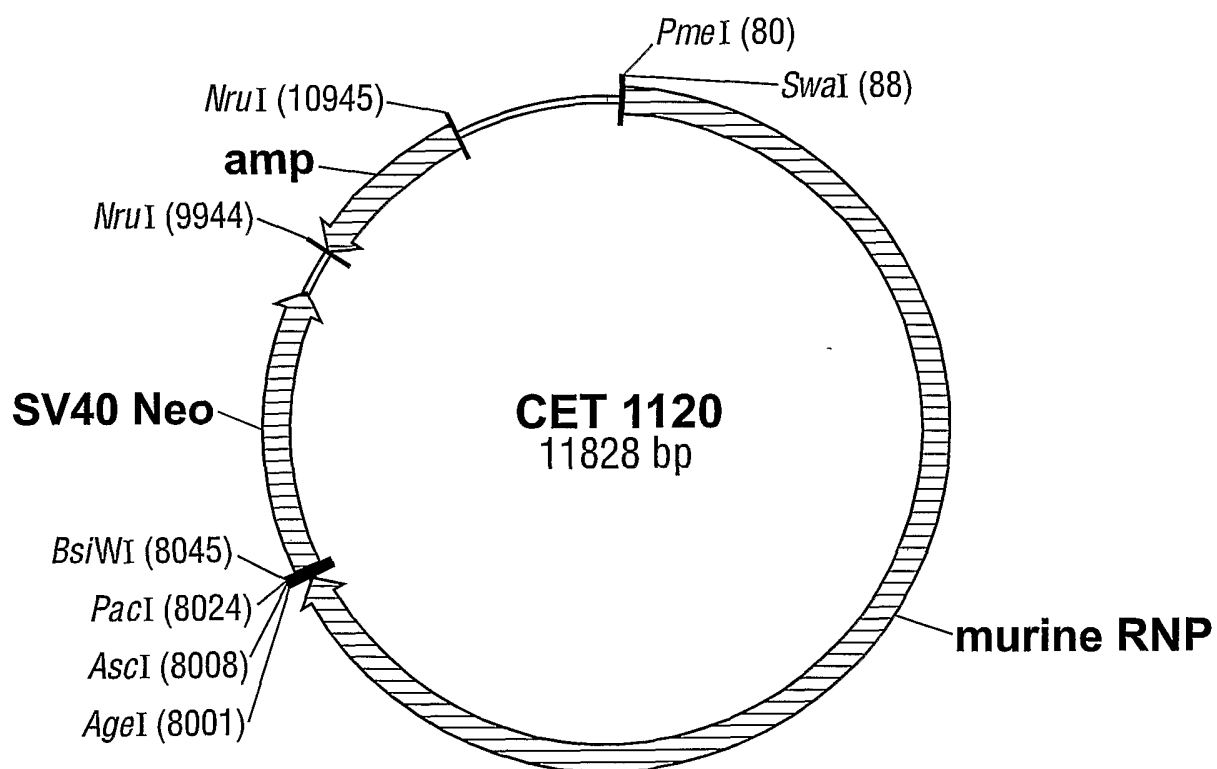
75/93

```
11501  GGATAACGCA  GGAAAGAACA  TGTGAGCAAA  AGGCCAGCAA  AAGGCCAGGA
      CCTATTGCGT  CCTTTCTTGT  ACACTCGTTT  TCCGGTCGTT  TTCCGGTCCT
11551  ACCGTAAAAA  GGCCGCGTTG  CTGGCGTTTT  TCCATAGGCT  CCGCCCCCCT
      TGGCATT TTT  CCGGCGCAAC  GACCGCAAAA  AGGTATCCGA  GGC GGGGGGA
11601  GACGAGCATC  ACAAAAATCG  ACGCTCAAGT  CAGAGGTGGC  GAAACCCGAC
      CTGCTCGTAG  TGT TTTTAGC  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
      CAT TCTGTGC  TGAATAGCGG  TGACCGTCGT  CGGTGACCAT  TGT CCTAATC
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  TTTTGT TTTG  CAAGCAGCAG  ATTACGCGCA
      GTGGCGACCA  TCGCCACCAA  AAAAACAAAC  GTTCGTCGTC  TAATGCGCGT
12151  GAAAAAAGG  ATCTCAAGAA  GATCCTTTGA  TCTTTTCTAC  GGGGTCTGAC
      CTTTTTTTCC  TAGAGTTCTT  CTAGGAAACT  AGAAAAGATG  CCCCAGACTG
12201  GCTCAGTGGA  ACGAAAAC TC  A
      CGAGTCACCT  TGCTTTTGAG  T
```

FIG. 27(contd.)

76/93

FIG. 28



77/93 **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 TTTAACCCCTC 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 TTTCCAGCT TTATCCAG
   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 CGACAACTC CTAATGCACC
501 CCTTCTTGGA TGGGGGGTGG GGGTGGGGTG GGAGGGTTGG GTGGTGGGTA
   GGAAGAACCT ACCCCCCACC CCCACCCAC CCTCCCAACC CACCACCCAT
551 CTTAAGAGGT TTCAAAAGTC AATATTGTTT GCATTTAGCT CTTCTTGTGA
   GAATTCTCCA AAGTTTTTCAG 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 ATCTTCTTTT 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
   CTTACCGGA ACCAATACCC TCGAATAGTG TCGTTATCTT GTCGTTAATA
951 GACTGGAGTA TGATAGTTAA AAACAAGCAA GCAAGCAAGC AAACACACAC
   CTGACCTCAT ACTATCAATT TTTGTTTCGT CGTTCGTTTCG TTTGTGTGTG
1001 ACCAAAACAA CAAAACCCCA AGACAGAGTC ACATGTAGCC CAGGCTAGCC
   TGGTTTTGTT GTTTTGGGGT TCTGTCTCAG GTTACATCGG GTCCGATCGG
1051 TCCAAATTCA CTATATAACT GAAGAAGACC CCTAATTCCC ATTCTCTAG
   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 AGTAAATCA GAGTTCTAGC ATCAACTACA GGCTCAACCA
   AATTAACAAT TCATTTTAGT CTCAAGATCG TAGTTGATGT CCGAGTTGGT
1301 TGGGGACCAC AAATAAACTA AAAGGGATAA GACTGGCTTC CCCATAATTA
   ACCCTGGTG TTTATTTGAT TTTCCCTATT CTGACCGAAG GGGTATTAAT

```

FIG. 29(contd.) 78/93

1351	TTACATTTAG	ATAATTTTCC	TGACTACTCA	ACAAAGCTAA	AATATCACCA
	AATGTAAATC	TATTAAAAGG	ACTGATGAGT	TGTTTCGATT	TTATAGTGGT
1401	CTGGTTTATT	TTCTCCTTCT	AGGGTTTAAAG	CTCACTCTGA	GGAGGGGCAT
	GACCAAATAA	AAGAGGAAGA	TCCCAAATTC	GAGTGAGACT	CCTCCCCGTA
1451	GCGGCACACA	CTCATAGCAT	CCAGGAAATA	GAAATATGGT	GACTATCATG
	CGCCGTGTGT	GAGTATCGTA	GGTCCTTTAT	CTTTATACCA	CTGATAGTAC
1501	GGTTCAGGGC	CAACCTAGGC	TTTAGAGAAA	AACCTTGTC	CACAAACCAA
	CCAAGTCCCG	GTTGGATCCG	AAATCTCTTT	TTGGAACAGG	GTGTTTGGTT
1551	AAATGTCTCT	TTTTTATTCT	ATCAGGGGTG	GATGGATTTG	TTAAAGAAGT
	TTTACAGAGA	AAAAATAAGA	TAGTCCCCAC	CTACCTAAAC	AATTTCTTCA
1601	GCTTTTAAAA	ACCTTGAGAT	GGTTATTTAG	AAGTCCCCAT	GGGATACCAA
	CGAAAATTTT	TGGAACCTCTA	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	TGTAACCTACA	GATGCTACTT	AAAAATTTTT	TAAAGGAATC
	GTGGTGGTAC	ACATTGATGT	CTACGATGAA	TTTTTAAAAA	ATTTCTTAG
1851	ACAAAAATAA	CCCCCTATCA	AATGCCTAGT	CCCTCTAACC	ATCACCAAGT
	TGTTTTTTATT	GGGGGATAGT	TTACGGATCA	GGGAGATTGG	TAGTGGTTCA
1901	GAAGGATCAC	GCAGGAAAAA	AAAAATCACC	AGCAGCACCT	CAGAACCCAGG
	CTTCCTAGTG	CGTCCTTTTT	TTTTTAGTGG	TCGTCGTGGA	GTCTTGGTCC
1951	ATACTCAGTC	CATCAGCATC	CAGGGCCATA	CCCACACTCA	CAGCATCTCC
	TATGAGTCAG	GTAGTCGTAG	GTCCCGGTAT	GGGTGTGAGT	GTGCTAGAGG
2001	ACAGTTTACC	AGATGATTCA	TGCTTATCAC	TGTATTGGGT	CATCTAAGAG
	TGTCAAATGG	TCTACTAAGT	ACGAATAGTG	ACATAACCCA	GTAGATTCTC
2051	TGACCATCAG	GGCTTCTGAT	CACAGAATCT	AGTCCACTTT	GCAGACCAGT
	ACTGGTAGTC	CCGAAGACTA	GTGTCTTAGA	TCAGGTGAAA	CGTCTGGTCA
2101	TGAAGTCATG	CACATATATG	GATAGAAATA	CCCTCTTGCT	CATTTTGGTC
	AGTTTCAGTAC	GTGATATACT	CTATCTTTAT	GGGAGAACGA	GTAACCACAG
2151	AGAAATTCAA	GGATAAAAAA	CCATGTTTTG	TTAATGCACA	CCTCCATATG
	TCTTTAAGTT	CCTATTTTTG	GGTACAAAAC	AATTACGTGT	GGAGCATATC
2201	ATTGAGATCA	ATGTGTCCTA	ATTAATGTAG	AAACCACAAC	TGTAAATTTT
	TAACCTCTAGT	TACACAGGAT	TAATTACATC	TTTGGTGTG	ACATTTAAAG
2251	ACTCTTTTGA	CATGAATCTT	TTTCTAGACA	GGGTCTTGGA	TGCAGCCCCG
	TGAGAAAAC	GTACTTAGAA	AAAGATCTGT	CCCAGAACCT	ACGTCGGGGC
2301	ACTACCCAGA	ATTTTGGAAT	CCAGGCTAGC	CTCAAACCTCA	AGGCAATCTG
	TGATGGGTCT	TAAAACCTTA	GGTCCGATCG	GAGTTTGAGT	TCCGTTAGAC
2351	CTTGCTTCAG	CTTCTCACAG	GCTGGATCAC	AAACATACAC	CTTCAGACCC
	GAACGAAGTC	GAAGAGTGTC	CGACCTAGTG	TTTGTATGTG	GAAGCTGGG
2401	ATTTTTTTTT	CCTCCCTCCG	TTTTTGGTTT	CTCTGTGTAG	CCCTGGGTGT
	TAAAAAATAA	GGAGGGAGGC	AAAAACCAA	GAGACACATC	GGGACCCACA
2451	CCGTGGACTC	GCTGTGTAGA	TCTATCTACC	AGCCTCTGTC	TTGGAGTACT
	GGCACCTGAG	CGACACATCT	AGATAGATGG	TCGGAGACAG	AACCTCATGA
2501	GGGATTAAAG	TTGTGGGCTA	CCACTGCCTG	GCTGACCCAG	TTTTATTTAT
	CCCTAATTTT	AACACCCGAT	GGTGACGGAC	CGACTGGGTC	AAAATAAATA
2551	TTTAAATATA	ACTTGACAAA	AATAAATTTG	TCTAACTTAC	TAGAAATCCC
	AAATTTATAT	TGAACTGTTT	TTATTTAAAC	AGATTGAATG	ATCTTTAGGG
2601	AAGAAAATA	ACACTGGATT	TAGCAACAGT	CAGAAATCGC	TGAAAAGAAA
	TTCTTTTGAT	TGTGACCTAA	ATCGTTGTCA	GTCTTTAGCG	ACTTTTCTTT
2651	CAGAATTGAT	CTAACAGTCT	TAGATCACTC	CTAGACAGTT	TGTAATTCCT
	GTCTTAACCTA	GATTGTCAGA	ATCTAGTGAG	GATCTGTCAA	ACATTAAGAA
2701	GCTCATGGCA	ACGTGAGCTC	TATCTAACTC	ACTCTCTGTG	CACATAATGAA
	CGAGTACCGT	TGCACTCGAG	ATAGATTGAG	TGAGAGACAC	GTGATTACTT
2751	TGCTCAGTGT	CTCCAGAACA	GCACAGCTTC	CAGGGTAATC	ATGCCAACCC
	ACGAGTCACA	GAGGTCTTGT	CGTGTCTGAAG	GTCCCATTAG	TACGGTTGGG

FIG. 29(contd.) 79/93

2801	ACAAGACTTT	TATAGAGCTG	TCCACGACTC	TTCCCCCATT	CAGCTCATTA
	TGTTCTGAAA	ATATCTCGAC	AGGTGCTGAG	AAGGGGGTAA	GTGAGTAAT
2851	ACAATATGAT	GGAGCTCCTG	TGTGGAAATC	AAGGCACACT	CTGGTAGAAA
	TGTTATACTA	CCTCGAGGAC	ACACCTTTAG	TTCCGTGTGA	GACCATCTTT
2901	CTTGTTTTTT	CTTTCCACTT	TTCTTGGGC	TCTGAAGATT	GAGCTGTTTT
	GAACAAAAAA	GAAAGGTGAA	AAGGAACCCG	AGACTTCTAA	CTCGACAAAA
2951	ATAACCCACA	AACATGCATT	TTTTACCTCA	AAAGCATCCA	GCAAAAACTG
	TATTGGGTGT	TTGTACGTAA	AAAATGGAGT	TTTCGTAGGT	CGTTTTTTGAC
3001	TACAACGCTT	TTTCAAAAAA	ATGTATTGTG	ATCCTCCTTA	AGAAAAGCCT
	ATGTTGCGAA	AAAGTTTTTT	TACATAACAC	TAGGAGGAAT	TCTTTTCGGA
3051	TACTTAGTGT	TAATTCCTTT	TTCTTTAGAA	TGCTGGTAAA	TACAAGGACT
	ATGAATCACA	ATTAAGGAAA	AAGAAATCTT	ACGACCATTT	ATGTTCTCTGA
3101	TAGGTAGGCT	GGCTTCTAAC	AGCAATTCAC	CCACTTATGA	TGGGATTAAA
	ATCCATCCGA	CCGAAGATTG	TCGTTAAGTG	GGTGAATACT	ACCCTAATTT
3151	GGAAGGCACA	ACCATGTCCA	CCACAGGTTT	TAGCTCCCCC	ACCCACACGC
	CCTTCCGTGT	TGGTACAGGT	GGTGTCCAAG	ATCGAGGGGG	TGGGTGTGCG
3201	CCAGAGAGGG	TTTTTCTGTG	TAGCTCTGAC	TATTCTGGAA	TTACACTGTC
	GGTCTCTCCC	AAAAAGACAC	ATCGAGACTG	ATAAGACCTT	AAGTGTGACG
3251	AGACCAGGCT	GGTCTCGAAC	TCAGAGATCC	ACCACCACAT	GGTTTCTTAA
	TCTGGTCCGA	CCAGAGCTTG	AGTCTCTAGG	TGGTGGTGTA	CCAAAGAATT
3301	TTGTAATTTT	AAAGAAAAAA	AAAAATCCTT	CAGTTAAGAT	TCTTATGTTT
	AACATTAAAA	TTTCTTTTTT	TTTTTAGGAA	GTCAATTCTA	AGAATACAAG
3351	TAGGTTTTCA	CAAACCTACC	AATGTAGTTT	TATTGGAGGC	CATTTTTTTAA
	ATCCAAAAGT	GTTTGAATGG	TTACATCAAA	ATAACCTCCG	GTAAAAAATT
3401	ATTTAATCGG	AGACTTGAAG	AGCTATTGCA	AGAAAAAATA	TGTAGGACAG
	TAAATTAGCC	TCTGAACTTC	TCGATAACGT	TCTTTTTTTT	ACATCCTGTC
3451	TTAAAATTTT	ATGACACACA	AAAGGCAGCT	ACAAGTTTTG	TGTGGATTTT
	AATTTTAAAG	TACTGTGTGT	TTTCCGTCTG	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	CTACTTTTCT
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	AAAAAGGCC	CCACCACCAC	GCCATTAGTG
3901	TCTCTATAGT	CCAGTCTGGG	CTTCAACGCC	TGGCAATCCC	CAGCCTCAAG
	AGAGATATCA	GGTCAGACCC	GAAGTTGCGG	ACCGTTAGGG	GTGCGAGTTC
3951	CTCCCAAGTA	CTGTCTTGAT	AAGGATAGAA	GGAGTCGACC	TCCTTCACGC
	GAGGGTTCAT	GACAGGACTA	TTCTATCTTT	CCTCAGCTGG	AGGAAGTGCG
4001	TCCCCCTCCG	GGAGGGCTCC	TTCCCAGCTC	CATTCCCCGG	TCGGGAGCCC
	AGGGGAGGCT	CCTCCCGAGG	AAGGGTCGAG	GTAAGGGGCC	AGCCCTCGGG
4051	GTCCCCCACC	CGAGAGCGCG	GGCCTCGTGG	TCAGCGCCTC	CGCGGGGAGA
	CAGGGGGTGG	GCTCTCGCGC	CCGAGCACC	AGTCGCGGAG	GCGCCCCCTC
4101	AACAAAGGCG	GCGGCGGGGG	CTCAAGGGCA	CTGCGCCACG	GGCCCCGCGC
	TTGTTTCCGC	CGCCGCCCCC	GAGTTCCCGT	GACGCGGTGC	CCGGGCGCGG
4151	TCCCCCATCC	GGCGGCGGCC	ACGTAGCCGG	GAGCGCGCCG	CAGCCCCGAG
	AGGGGGTAGG	CCGCCGCCGG	TGCATCGGCC	CTCGCGCGGC	GTGCGGCCCT
4201	CCTCGGGCCT	CGCAGCTGCA	GAGCCTGAAC	CGCTCTCTCC	CTGCGGGCCT
	GGAGCCCCGA	GCGTCGACGT	CTCGGACTTG	GCGAGAGAGG	GACGCCCGGA

FIG. 29(contd.) 80/93

4251	GCGACGAGGC	TGGGGGAGGG	GAGGCCCGCG	CTTTGTCTGG	AGTCTCGGTA
	CGCTGCTCCG	ACCCCTTCCC	CTCCGGGCGC	GAAACAGACC	TCAGAGCCAT
4301	GCTGTCATCC	GGCTCCCACC	CTCATGCACA	ATTGTCCCAT	CTCCCCCAGC
	CGACAGTAGG	CCGAGGGTGG	GAGTACGTGT	TAACAGGGTA	GAGGGGGTGC
4351	CACCGGCGCG	GCGCCCGCCT	CAGCGAGGCC	CCAGCCGGTT	TCCCGCAGCC
	GTGGCCGCGC	CGCGGGCGGA	GTCGCTCCGG	GGTCGGCCAA	AGGGCGTCGG
4401	CGCGGCCAC	GGGGCTCGCA	GCCTCCCCGC	AAGCTCGGAC	GCACGGAGCA
	GCGCCGGGTG	CCCCGAGCGT	CGGAGGGGCG	TTGAGCCTTG	CGTGCTCGT
4451	TCCTAAACCC	CACCACACGC	AAGATCGAAA	AAAAGCAAAG	GCACGAACCT
	AGGATTTGGG	GTGGTGTGCG	TTCTAGCTTT	TTTTCGTTTC	CGTGCTTGAA
4501	CACCGCTCCG	ATGCTCAGGG	CCGCGGATCC	TGCAGAGTCT	CCCGCTGCG
	GTGGCGAGGC	TACGAGTCCC	GGCGCGTAGG	ACGTCTCAGA	GGGCGGACGC
4551	CGCTTCGGTT	CAGCCACATC	CGAGGGGAGG	GGGCGCGGGC	AGCTCCGCCG
	GCGAAGCCAA	GTGCGGTAG	GCTCCCCCTC	CCCGCGCCCC	TCGAGGCGGC
4601	GGGGGGAGGG	GGAGCACC GC	CCACGCCCTG	GCCGCGCGGG	GCCCGCCGGG
	CCCCCTCCC	CCTCGTGGCG	GGTGCGGGAC	CGGCGCGCCC	CGGGCGGCCC
4651	AACGCGTCCT	GCGGGGGGCG	GCGCGCGCAA	TGCTCACCGT	CGCGGCGGTG
	TTGCGCAGGA	CGCCCCCGCG	CGCGCGCGTT	ACGAGTGGCA	GGCGCCGCAC
4701	GCGCCCAGGG	GGTCTCCTGG	CTGGGGGGAG	GGGGGGGAAG	GCGGGCAGGA
	CGCGGGTCCC	CCAGAGGACC	GACCCCCCTC	CCCCCCTTTC	CGCCCGTCCT
4751	AGGACCGCGG	AGGCCTCTCT	GCGTCTCGGA	GCGCGCCAAA	GCGGGGCTCC
	TCCTGGCGCC	TCCGGAGAGA	CGCAGAGCCT	CGCGCGGTTT	CGCCCCGAGG
4801	ACCCACCTCC	TTGCCCCGAT	CTTGAAGGCC	GGGGAGATAA	ACAGCGGGGT
	TGGGTGGAGG	AACGGGCCTA	GAACCTCCGG	CCCCCTCTATT	TGTGCCCCCA
4851	TCTTTAAGCA	CCACCTCTCA	CTAGGCGCGG	GATCCCAAGG	CTTGTGGCAT
	AGAAATTTCGT	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
	CTGCCCCAAC	CGTTTCCGCA	TCGACGTCTC	TCGTGCGAAC	CCACCTCTCC
5051	GCCGCACGCC	CCAGCGCCGG	CCTAAGCCCC	TCCCGACGGC	GTTATTTCAA
	CGGCGTGCGG	GGTCGCGGCC	GGATTGCGGG	AGGGCTGCCG	CAATAAAGTT
5101	ACTGCGCGAC	CGTTTCTCCG	CTCCCTACGC	GGAGGTGGGG	GCCGGACCTA
	TGACGCGCTG	GCAAAGAGGC	GAGGGATGCG	CCTCCACCCC	CGGCCTGGAT
5151	GTTCCGGACG	TAGTAACACG	CCGAGCGCGA	GCCTTCCGCA	ATTACGGAA
	CAAGGCCTGC	ATCATTGTGC	GGCTCGCGCT	CGGAAGGCGT	TAAGTGCCTT
5201	CACAGTTGCG	CAAGTGATGT	AAAGCAGTCC	CGCTGTACCT	AAAGGGGGAG
	GTGTCAACGC	GTTCACTACA	TTTCGTGAGG	GCGACATGGA	TTTCCCCCTC
5251	TGTCACGTAC	TTGGCGTAAG	GAGAGTGTAG	GCCCTTCCCC	CCATTGGCGG
	ACAGTGCATG	AACCGCATTC	CTCTCACATC	CGGGAAGGGC	GGTAACCGCC
5301	CGGTTAGGGC	GTTTACGTAA	CGGCGTGACG	TAAGCGGAGA	CGCGTTAGTG
	GCCAATCCCG	CAAATGCATT	GCCGCACTGC	ATTGCGCTCT	GCGCAATCAC
5351	GGGGGAAGGT	TCTAGAAAAG	CGGCGGTCTC	GGCTCCAGCG	GCAGTAGCAG
	CCCCCTTCCA	AGATCTTTTC	GCCGCCAGAG	CCGAGGTGCG	CGTCATCGTC
5401	CGGCGCCGGT	CCCGTGTGCA	GGAGCTCCTT	TGCGGCCAG	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
	GA CTCAGGCG	CTACCTCCAT	TGCCCCAACT	TTAGTTACTC	AATAATTTTT
5601	TGGCATGGCG	AGGCCGTAGG	CACCGCAATG	GAAACCGGCC	ACCCGCCTCC
	ACCGTACCGC	TCCGGCATCC	GTGGCGTTAC	CTTTGGCCGG	TGGGCGGAGG
5651	GTGGTCCGGC	GGAGGGGATG	CGGCCACTCG	AGTGCGGGTT	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
	TTACGGGGCG	GCGCTGGCCC	GACCTGGCGA	TTTCATCGCG	CACTCGCCCCG
5851	CATCGCTGGC	CTTTCGATGT	GCGCGGGCCT	AGGGGCTCGG	TTGTGTTCGC
	GTAGCGACCG	GAAAGCTACA	CGCGCCCCGA	TCCCCGAGCC	AACACAAGCG
5901	GGCGGAACGT	TTCTGGGGCC	CCCCCGGCTT	CCCGGAGCGA	GTCTGCGAAG
	CCGCCTTGCA	AAGACCCCGG	GGGGGCCGAA	GGGCCTCGCT	CAGACGCTTC
5951	CTAGCTTCCC	CTCCCCCCTC	TCCCCGGAAC	CGGATTTGGC	GGCCGCCATT
	GATCGAAGGG	GAGGGGGGAG	AGGGCCCTTG	GCCTAAACCG	CCGGCGGTAA
6001	TTCCCGTCTC	CTTCCTCGCC	ACGATTTTGC	TTTCAACGCT	TTAGGTTTAC
	AAGGGCAGAG	GAAGGAGCGG	TGCTAAAACG	AAAGTTGCGA	AATCCAAATG
6051	TAGTTTGGTT	TTCTTTTTTTC	ACCACTGCGT	AGACGTGTTT	AGCGATTTTC
	ATCAAACCAA	AAGAAAAAAG	TGGTGACGCA	TCTGCACAAA	TCGCTAAAAG
6101	CTTTCTTTTG	GAAGTCTTCA	TACCGTTTTCG	AGGTGGATTT	AGCGTTTTGA
	GAAAGAAAAAC	CTTCAGAAGT	ATGGCAAAGC	TCCACCTAAA	TCGCAAAACT
6151	GCTTGGGTCT	TCAGCGTCCT	GCGCACCTCG	CTAAAGGCTC	TCTGCCTTCC
	CGAACCCAGA	AGTCGCAGGA	CGCGTGAGC	GATTTCCGAG	AGACGGAAGG
6201	CCTCGACGAA	ATGGCGCCAT	TGCTTCTCTGA	AGCCACCGAG	GCGCGGGGTG
	GGAGCTGCTT	TACCGCGGTA	ACGAAAGACT	TCGGTGGCTC	CGCGCCCCAC
6251	GGGGCGGGGT	GGCGGCGCTC	CACGAGCTTT	ACTGGAACAG	GCAGAGAGAA
	CCCCGCCCCA	CCGCCGCGAG	GTGCTCGAAA	TGACCTTGTC	CGTCTCTCTT
6301	CGTAGTACAA	CCGAGGCCTG	GGCGGGTGGC	TGAAGGCAGC	GTGCTGCAA
	GCATCATGTT	GGCTCCGGAC	CCGCCCACCG	ACTTCCGTCTG	CAGCGACGTT
6351	AGAGACCGTT	TTATTTTTC	TAATACGTAA	GATTACGGGT	GCTGTAGTAA
	TCTCTGGCAA	AATAAAAAGT	ATTATGCATT	CTAATGCCCA	CGACATCATT
6401	AGCACTTGAG	CATTAGTATA	GTAGGAGGAA	GTCAAAGTGG	AAAAAATGGG
	TCGTGAACCTC	GTAATCATAT	CATCCTCCTT	CAGTTTCACC	TTTTTTACCC
6451	AGCGCTCATC	AGGAAGCTAG	GGAGGCTATG	TTGAGTGCAG	GCTTACTTTC
	TCGCGAGTAG	TCCTTCGATC	CCTCCGATAC	AACTCACGTC	CCAATGAAAG
6501	CTTTTATTGC	AGAACTTTTA	TCTGCTTAAA	GGATCCTCGG	ATCGAAATAA
	GAAAATAACG	TCTTGAAAAT	AGACGAATTT	CCTAGGAGCC	TAGCTTTATT
6551	TTCAAATTAT	AAGCATTTTT	AAGGGAATCT	TCGAATTTGT	TGGTAAAGTC
	AAGTTTAATA	TTTCGTAAAA	TTCCCTTAGA	AGCTTAAACA	ACCATTTTCAG
6601	AACGGATCCT	TAGCACGTGG	TGTTCACTTT	AAGGAAGTGA	AATAGCTGAC
	TTGCCTAGGA	ATCGTGACAC	ACAAGTGAAA	TTCTTCACT	TTATCGACTG
6651	TTTTTCATAGT	TAGCCTTCGC	TTAAAGCCTG	GTTCAGTGGA	CGAAAATCCA
	AAAAGTATCA	ATCGGAAGCG	AATTTTCGGAC	CAAGTCACCT	GCTTTTAGGT
6701	CGTCCTGGCT	ATATAAAAAC	TTAGTTTGGG	GTCACAGTGT	TTGAGCGTGG
	GCAGGACCGA	TATATTTTTG	AATCAAACCC	CAGTGTACAC	AACTCGCACC
6751	TCATTTCGGTT	TTTTTATTTT	TTATTTGTTT	GAAATTATGA	TGCATCATTA
	AGTAAGCCAA	AAAAATAAAA	AATAAACAAA	CTTTAATACT	ACGTAGTAAT
6801	CACTGATAAG	CATTAGCTTT	CGAATTGAAA	GGGGTCTCCT	TGGTTATTTT
	GTGACTATTG	GTAATCGAAA	GCTTAACTTT	CCCAGAGGA	ACCAATAAAA
6851	CTTTGACTCT	AAGCACACTT	ATAAATAAAA	TAACCTTGTT	TATAATCGAT
	GAAACTGAGA	TTTCGTGTGAA	TATTTATTTT	ATTGGAACAA	ATATTAGCTA
6901	AGTGGACGTC	TGGTAAGTTT	GGAAAAAACC	CGAGGTAAGT	AAAGAGCTTT
	TCACCTGCAG	ACCATTCAAA	CCTTTTTTGG	GCTCCATTCA	TTTCTCGAAA
6951	TGCTTTCGTT	AGTGATATGA	AAAAACAAGG	TGTATTTAAT	ACTTGCAACT
	ACGAAAGCAA	TCACTATACT	TTTTTGTTC	ACATAAATTA	TGAACGTTGA
7001	TAGTTTAAGG	AAAGCCAATT	TACTGACATT	TTAGTAGAGC	TACCAGAAAC
	ATCAAATTCC	TTTCGGTTAA	ATGACTGTAA	AATCATCTCG	ATGGTCTTTG
7051	ACTATTTGGA	GTCTTGATTA	AGGCTTTTGT	AACTATTTTG	ACTATTTAAA
	TGATAAACCT	CAGGACTAAT	TCCGAAAACA	TTGATAAAAC	TGATAAATTT
7101	ACAATTTTGG	TCGTTTTTAT	TAAACATTTT	AAAACCTAAA	AATTGTAAAC
	TGTTAAAACC	AGCAAAAATA	ATTTGTAAAG	TTTTGGATT	TTAACATTTG

FIG. 29(contd.) 82/93

7151	ATTGGCTTTT	TGAGCACAATT	TTGGAGAAAC	TTACAAATTT	AGGCTATACA
	TAACCGAAAA	ACTCGTGTA	AACCTCTTTG	AATGTTTAAA	TCCGATATGT
7201	GTAAAATAAC	GGATTTGTTT	TATAATTTTG	CTTTTTTCATT	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	TAATGTTTTTC
7601	AATTCCTTAA	TACCCCTTAA	ATGACCTAAG	AGCTGGGTAT	GGTTTTTCTG
	TTAAGAAATT	ATGGGAAATT	TACTGGATTTC	TCGACCCATA	CCAAAAAGAC
7651	AATTTTTTTT	AAGAAAATCT	AAGAAAGTTT	ACGTGAATTA	GAAGTTAGAT
	TTAAAAAAAC	TTCTTTTAGA	TTCTTTCAAA	TGCACTTAAT	CTTCAATCTA
7701	CGAATATTAG	TGACTTTGAA	ACTTGTATAG	CTCAGGCAAT	TTTTGGTGTA
	GCTTATAATC	ACTGAAACTT	TGAACATATC	GAGTCCGTTA	AAAACCACAT
7751	ACACAACATA	TATGCAGTTT	AACATATGGT	TTAAATTTGA	TGTAAGTTTTT
	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
	TTTTCTTGT	CAAGGCTTTC	GAGAAATAAC	CACCGAATTC	GAATAGCTAT
8001	CCGGTGGCGC	GCCAATTGTT	AATTAAGATC	TGGCCCAATG	GGCCGTACGA
	GGCCACCGCG	CGGTAAACAA	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	TCGTCGCTCT
8201	AGTATGCAAA	GCATGCATCT	CAATTAGTCA	GCAACCATAG	TCCCGCCCCCT
	TCATACGTTT	CGTACGTAGA	GTTAATCAGT	CGTTGGTATC	AGGGCGGGGA
8251	AACTCCGCCC	ATCCCGCCCC	TAACCTCCGC	CAGTTCGCCC	CATTCTCCGC
	TTGAGGCGGG	TAGGGCGGGG	ATTGAGGCGG	GTCAAGGCGG	GTAAGAGGCG
8301	CCCATGGCTG	ACTAATTTTT	TTTATTTATG	CAGAGGCCGA	GGCCGCCTCG
	GGGTACCGAC	TGATTAAAAA	AAATAAATAC	GTCTCCGGCT	CCGGCGGAGC
8351	GCCTCTGAGC	TATTCAGAA	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	ATTTCGGCTAT	GACTGGGCAC	AACAGACAAT	CGGCTGCTCT	GATGCCGCCG
	TAAGCCGATA	CTGACCCGTG	TTGTCTGTTA	GCCGACGAGA	CTACGGCGGC
8551	TGTTCCGGCT	GTCAGCGCAG	GGGCGCCCGG	TTCTTTTTGT	CAAGACCGAC
	ACAAGGCCGA	CAGTCGCGTC	CCCGCGGGCC	AAGAAAAACA	GTTCTGGCTG

FIG. 29(contd.) 83/93

8601	CTGTCCGGTG	CCCTGAATGA	ACTGCAAGAC	GAGGCAGCGC	GGCTATCGTG
	GACAGGCCAC	GGGACTTACT	TGACGTTCTG	CTCCGTCGCG	CCGATAGCAC
8651	GCTGGCCACG	ACGGGCGTTC	CTTGCGCAGC	TGTGCTCGAC	GTTGTCACTG
	CGACCGGTGC	TGCCCCGAAG	GAACGCGTCG	ACACGAGCTG	CAACAGTGAC
8701	AAGCGGGAAG	GGACTGGCTG	CTATTGGGCG	AAGTGCCGGG	GCAGGATCTC
	TTCGCCCTTC	CCTGACCGAC	GATAACCCGC	TTACGCGCCC	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	GCTCGTGCA	GAGCCTACCT	TCGGCCAGAA
8901	GTCGATCAGG	ATGATCAAGA	GCATCAGGGG	CTCGCGCCAG	CCGAACGTGT
	CAGCTAGTCC	TACTAGTTCT	CGTAGTCCCC	GAGCGCGGTC	GGCTTGACAA
8951	CGCCAGGCTC	AAGGCGAGCA	TGCCCCGACG	CGAGGATCTC	GTCGTGACCC
	GCGGTCCGAG	TTCCGCTCGT	ACGGGCTGCC	GCTCCTAGAG	CAGCACTGGG
9001	ATGGCGATGC	CTGCTTGCCG	AATATCATGG	TGGAATAATG	CCGCTTTTCT
	TACCGCTACG	GACGAACGGC	TTATAGTACC	ACCTTTTACC	GGCGAAAAGA
9051	GGATTTCATCG	ACTGTGGCCG	GCTGGGTGTG	GCGGACCGCT	ATCAGGACAT
	CCTAAGTAGC	TGACACCGGC	CGACCCACAC	CGCCTGGCGA	TAGTCCTGTA
9101	AGCGTTGGCT	ACCCGTGATA	TTGCTGAAGA	GCTTGGCGGC	GAATGGGCTG
	TCGCAACCGA	TGGGCACTAT	AACGACTTCT	CGAACC GCCG	CTTACCCGAC
9151	ACCGCTTCCT	CGTGCTTTAC	GGTATCGCCG	CTCCCGATTC	GCAGCGCATC
	TGGCGAAGGA	GCACGAAATG	CCATAGCGGC	GAGGGCTAAG	CGTCGCGTAG
9201	GCCTTCTATC	GCCTTCTTGA	CGAGTTCCTC	TGAGCGGGAC	TCTGGGGTTC
	CGGAAGATAG	CGGAAGAACT	GCTCAAGAAG	ACTCGCCCTG	AGACCCCAAG
9251	GAAATGACCG	ACCAAGCGAC	GCCCAACCTG	CCATCACGAG	ATTTTCGATTC
	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	TTCTCTGTG	ATGGCCTTCC
9451	AACCCGCGCT	ATGACGGCAA	TAAAAAGACA	GAATAAAACG	CACGGTGTG
	TTGGGCGCGA	TACTGCCGTT	ATTTTCTGT	CTTATTTTGC	GTGCCACAAC
9501	GGTCGTTTGT	TCATAAACGC	GGGGTTCGGT	CCCAGGGCTG	GCACTCTGTC
	CCAGCAAACA	AGTATTTGCG	CCCCAAGCCA	GGGTCCCGAC	CGTGAGACAG
9551	GATACCCAC	CGAGACCCCA	TTGGGGCCAA	TACGCCCCG	TTTCTTCCTT
	CTATGGGGTG	GCTCTGGGGT	AACCCCGGTT	ATGCGGGCGC	AAAGAAGGAA
9601	TTCCCCACCC	CACCCCCCAA	GTTTCGGGTGA	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	GGTCGAAAC	AAGGGAAATC	ACTCCCAATT
9751	TTTCGAGCTT	GGCGTAATCA	TGGTCATAGC	TGTTTCCTGT	GTGAAATTGT
	AAAGCTCGAA	CCGCATTAGT	ACCAGTATCG	ACAAAGGACA	CACTTTAACA
9801	TATCCGCTCA	CAATTCCACA	CAACATACGA	GCCGGAAGCA	TAAAGTGTA
	ATAGGCGAGT	GTAAAGGTGT	GTTGTATGCT	CGGCCTTCGT	ATTTACATT
9851	AGCCTGGGGT	GCCTAATGAG	TGAGCTAACT	CACATTAATT	CGGTGCGCT
	TCGGACCCCA	CGGATTACTC	ACTCGATTGA	GTGTAATTAA	CGCAACGCGA
9901	CACTGCCCCG	TTTCCAGTCG	GGAAACCTGT	CGTGCCAGCA	TCGCGAGCAC
	GTGACGGGCG	AAAGGTCAGC	CCTTTGGACA	GCACGGTCGT	AGCGCTCGTG
9951	TTTTTCGGGA	AATGTGCGCG	GAACCCCTAT	TTGTTTATTT	TTCTAAATAC
	AAAAGCCCCCT	TTACACGCGC	CTTGGGGATA	AACAAATAAA	AAGATTTATG
10001	ATTCAAATAT	GTATCCGCTC	ATGAGACAAT	AACCCTGATA	AATGCTTCAA
	TAAGTTTATA	CATAGGCGAG	TACTCTGTTA	TTGGGACTAT	TTACGAAGTT

84/93

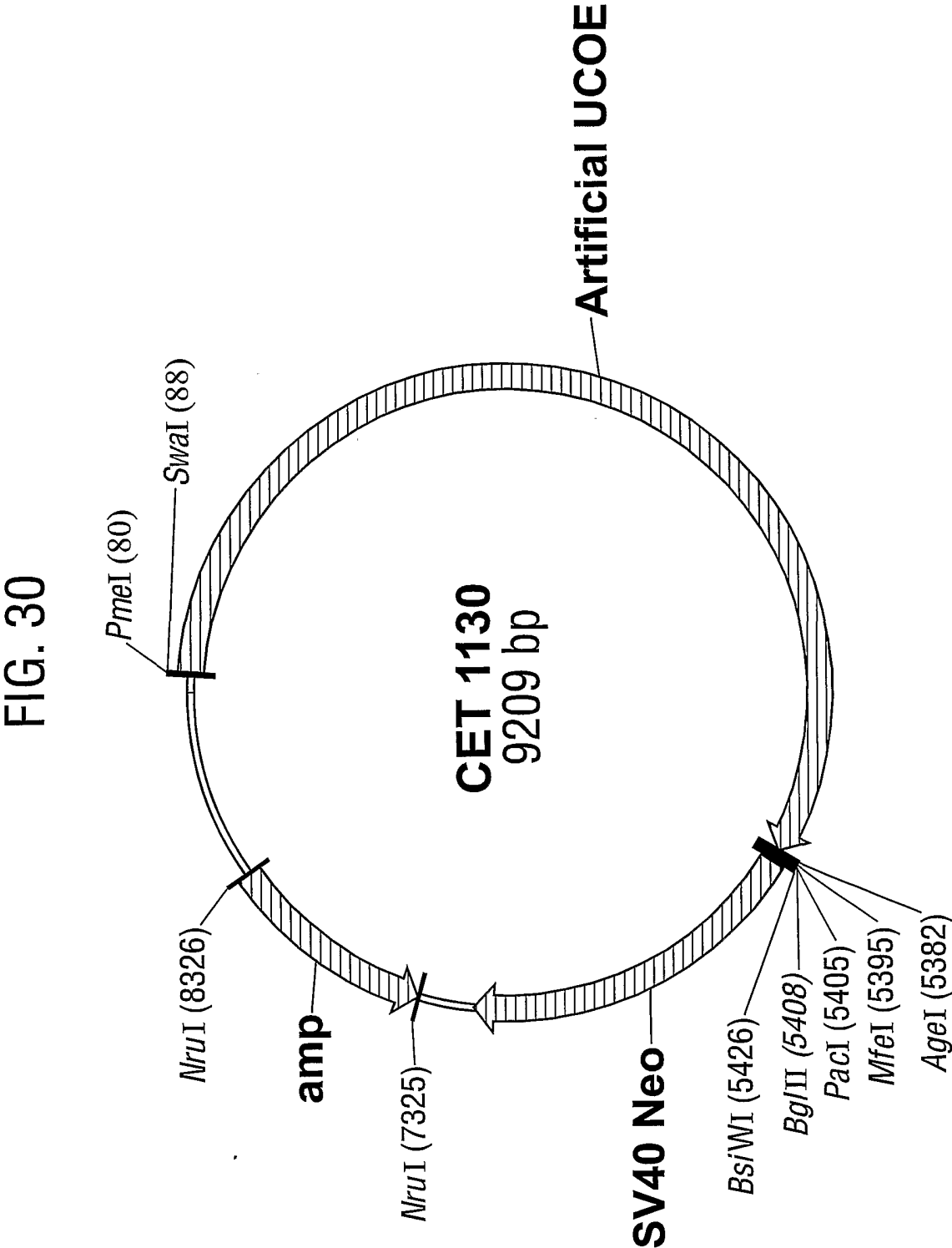
FIG. 29(contd.)

10051	TAATATTGAA	AAAGGAAGAG	TATGAGTATT	CAACATTTCC	GTGTGCCCC
	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	TAGGAAGTCT	CAAAAGCGGG
10251	CGAAGAACGT	TTTCCAATGA	TGAGCACTTT	TAAAGTTCTG	CTATGTGGCG
	GCTTCTTGCA	AAAGGTTACT	ACTCGTGAAA	ATTTCAAGAC	GATACACCGC
10301	CGGTATTATC	CCGTATTGAC	GCCGGGCAAG	AGCAACTCGG	TGCGCCGATA
	GCCATAATAG	GGCATAACTG	CGGCCCGTTC	TCGTTGAGCC	AGCGGCGTAT
10351	CACTATTCTC	AGAATGACTT	GGTTGAGTAC	TCACCAGTCA	CAGAAAAGCA
	GTGATAAGAG	TCTTACTGAA	CCAACCTCATG	AGTGGTCAGT	GTCTTTTTCGT
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	TAACCTCGCT
	TTCTCTGATT	GGCGAAAAAA	CGTGTTGTAC	CCCCTAGTAC	ATTGAGCGGA
10551	TGATCGTTGG	GAACCGGAGC	TGAATGAAGC	CATACCAAAC	GACGAGCGTG
	ACTAGCAACC	CTTGCCCTCG	ACTTACTTCG	GTATGGTTTG	CTGCTCGCAC
10601	ACACCACGAT	GCCTGTAGCA	ATGGCAACAA	CGTTGCGCAA	ACTATTAACT
	TGTGGTGCTA	CGGACATCGT	TACCGTTGTT	GCAACGCGTT	TGATAATTGA
10651	GGCGAAGTAC	TTACTCTAGC	TTCCCGGCAA	CAATTAATAG	ACTGGATGGA
	CCGCTTGATG	AATGAGATCG	AAGGGCCGTT	GTTAATTATC	TGACCTACCT
10701	GGCGGATAAA	GTTGCAGGAC	CACTTCTGCG	CTCGGCCCTT	CCGGCTGGCT
	CCGCCTATTT	CAACGTCTTG	GTGAAGACGC	GAGCCGGGAA	GGCCGACCGA
10751	GGTTTATTGC	TGATAAATCT	GGAGCCGGTG	AGCGTGGGTC	TCGCGGTATC
	CCAAATAACG	ACTATTTAGA	CCTCGGCCAC	TCGCACCCAG	AGCGCCATAG
10801	ATTGCAGCAC	TGGGGCCAGA	TGGTAAGCCC	TCCCGTATCG	TAGTTATCTA
	TAACGTCTGT	ACCCCGGTCT	ACCATTTCGG	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	GCGCGCCCCCT	CTCCGCCAAA	CGCATAACCC
11001	CGCTCTTCCG	CTTCCTCGCT	CACTGACTCG	CTGCGCTCGG	TCGTTTCGGCT
	GCGAGAAGGC	GAAGGAGCGA	GTGACTGAGC	GACGCGAGCC	AGCAAGCCGA
11051	GCGGCGAGCG	GTATCAGCTC	ACTCAAAGGC	GGTAATACGG	TTATCCACAG
	CGCCGCTCGC	CATAGTCGAG	TGAGTTTCCG	CCATTATGCC	AATAGGTGTC
11101	AATCGGCGGA	TAACGCAGGA	AAGAACATGT	GAGCAAAAGG	CCAGCAAAAG
	TTAGTCCCCT	ATTGCGTCCT	TTCTTGATACA	CTCGTTTTTC	GGTCGTTTTTC
11151	GCCAGGAACC	GTAAAAAGGC	CGCGTTGCTG	GCGTTTTTTC	ATAGGCTCCG
	CGGTCCTTGG	CATTTTTTCCG	GCGCAACGAC	CGCAAAAAGG	TATCCGAGGC
11201	CCCCCCTGAC	GAGCATCACA	AAAATCGACG	CTCAAGTCAG	AGGTGGCGAA
	GGGGGGACTG	CTCGTAGTGT	TTTTAGCTGC	GAGTTCAGTC	TCCACCGCTT
11251	ACCCGACAGG	ACTATAAAGA	TACCAGGCGT	TTCCCCCTGG	AAGCTCCCTC
	TGGGCTGTCC	TGATATTTCT	ATGGTCCGCA	AAGGGGGACC	TTGAGGGGAG
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	GCGAGGTTTC	ACCCGACACA	CGTGCTTGGG
11451	CCCGTTTCAGC	CCGACCGCTG	CGCCTTATCC	GGTAACTATC	GTCTTGAGTC
	GGGCAAGTCG	GGCTGGCGAC	GCGGAATAGG	CCATTGATAG	CAGAACTCAG

85/93

```
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.)



87/93

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 TGTGCTTCC
   ACCCGAGGAG GTATTCTGAC ACTCGTAGAC TGGTGACCTC ACAACGAAGG
251 CATTATATCC CTGTTATCAA GCACAAGGTC AGGCACAGAG TAAGACTCAA
   GTAATATAGG GACAATAGTT CGTGTTCAG TCCGTGTCTC ATTCTGAGTT
301 AACATGTTTT GGAATGTATG ACTGGTATGA ACTACAAACC AGTAAGCTGA
   TTGTACAAAA CCTTACATAC TGACCATACT TGATGTTTGG TCATTGCACT
351 TGTTTTTCATT TTGAGTCTAT AAATCTAATT TTGTGGTGGT TTTGTGTATG
   ACAAAAAGTAA 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
   ATATTCCTATT CACGACCTAT TTACAACCGA TGATATTATT TTATTCGGAG
601 TAAGATACTT GGTGAGCACA AGTACTACCC AAGAGTATGC ACTGTAAGTA
   ATTCTATGAA CCAGTCGTGT TCATGATGGG TTCTCATACG TGACATTCAT
651 AACTGACAAA ATTGTGTATC TAAAACTGGC CAGATGAAAG AGAACTTTTT
   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 TCTTTTGTTC CAGGTCCTTC
901 TATATTTTTT 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 AAAGTCTTA
   AGGTTTCATA 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 CAAAACCTCT GTCAGATTAT TAAACGACAG
1251 TCTAAAGTAC TATATTCCTT ATAGTTCTGG CATTTTAGAT AAAGGGTCAT
   AGATTTTCATG ATATAAGGGA TATCAAGACC GTAAAATCTA TTTCCCAGTA
1301 AAATTAAATG CCTATATGGT GACATTATTC AGTGATTCAG ACTTCACAGC
   TTTAATTTAC GGATATACCA CTGTAATAAG TCACTAAGTC TGAAGTGTCTG

```

FIG. 31(contd.) 88/93

```

1351 CTTTTTTTTT TTTTACAAA GGTGTTCCAG GCATGAAAAA TTTTAAAGTA
      GAAAAAAAAA AAAAATGTTT CCACAAGGTC CGTACTTTTT AAAATTTTCAT
1401 CTATACCTTT CCTAATTTTA CCTTTAAAGT TGTCTTGGA ATATCTGGGT
      GATATGGAAA GGATTAAAAT GGAAATTTCA ACAGGACCTT TATAGACCCA
1451 TGACAAAGGC GATGAAACTG AACTGAGACT TAAAAAAAAG ATTACCCACC
      ACTGTTTCCG CTACTTTGAC TTGACTCTGA ATTTTTTTTTC TAATGGGTGG
1501 TGGTTGTGCA CAAGCCTGCT TATGTCCCAA TCTCCAGTCT AGGGTCTGAT
      ACCAACACGT GTTCGACGCA ATACAGGGTT AGAGGTCAGA TCCCAGACTA
1551 GCTCCTTGCT GCAGTAATAT GCTTTGTGGC ATCTGGAGCA CGTTTTGGGG
      CGAGGAACGA CGTCATTATA CGAAACACCG TAGACCTCGT GCAAAACCCC
1601 CCTAAACAGC CACAAACCTT GCAGAGATGA GCACCAGACT TAAGCTGGAG
      GGATTTGTCT GTGTTTGGGA CGTCTCTACT CGTGGTCTGA ATTCGACCTC
1651 ACACACTGAT TCTCCTGTTT CTGGGGGAGG ATTCTCAGAA GGTGGCTCAT
      TGTGTGACTA AGAGGACAAA GACCCCTCC TAAGAGTCTT CCACCGAGTA
1701 ATGAGTAAAA ATCGTTTTTC CTGGGTAGTT GATTCTTAAA AACTAAAAAA
      TACTCATTTT TAGCAAAAAG GACCCATCAA CTAAGGATTT TTGATTTTTT
1751 GAATACAGAG AAAAGTTTTT TCTTCAAACA AAACAGCAAT TCACATATTT
      CTTATGTCTC TTTTCAAAAT AGAAGTTTGT TTTGTCGTTA AGTGTATAAA
1801 TATCCTCTGC ACGTAAACT GAAAATAACA ACAACAAAAA AGAAATGAAA
      ATAGGAGACG TGCATTTTGA CTTTTATTGT TGTGTTTTTT TCTTTACTTT
1851 GTTTTTGCTT TCAGGAATAA GCTTTTAAAA TCCAGAACT AGATTTTCGT
      CAAAAACGAA AGTCCTTATT CGAAAATTTT AGGTCTTTGA TCTAAAGCAG
1901 CGGTACACGC AACTGAGTTG CCTCCTAGAG GTGGTTTGAG TTAATCAAAT
      GCCATGTGCG TTGACTCAAC GGAGGATCTC CACCAAACCTC AATTAGTTTA
1951 TAATAAGACT GATCGTTAAG AACGACTGCC AAAAATACGA AAAAGCTACT
      ATTATTCTGA CTAGCAATTC TTGCTGACGG TTTTATGCT TTTTCGATGA
2001 GGGATCCATC TTTCCAAGAC AATTTCTATT ATCTGAATTA ACACATACC
      CCCTAGGTAG AAAGGTTCTG TTAAAGATAA TAGACTTAAT TGTGGTATGG
2051 TGGTACCCAC TGATTAAAAG CTGGGGGTTA CCAATGCGCG TGGGCACAGT
      ACCATGGGTG ACTAATTTTC GACCCCAAT GGTACGCGC ACCCGTGTCA
2101 TAGAAGCTTA TGTAGCAAAA ATGAGCACAT CCTGGAAGGG CCCGGGAGAA
      ATCTTCGAAT ACATCGTTTT TACTCGTGTA GGACCTTCCC GGGCCCTCTT
2151 GGTGCTCCTG GGCAGCGCG CTCTCCCTCG AGACTCCGAC CCCGCCGTCG
      CCACGAGGAC CCCGTCGCGC TGGTCGCTCC CGGAATTAAC GCGCGCACG
2201 GGTGCTTGCC GCCGTCCCC TGGTCGCTCC CGGAATTAAC GCGCGCACG
      CCACGAACCG CGGCAGGGGG ACCAGCGAGG GCCTTAATTG CCGCGCGTGC
2251 CGTCGGAGGC ATGGCCCCGT CCCGACCCG TTTGGCGGCT CACCTCGCAG
      GCAGCCTCCG TACCGGGGCA GGGCTGGGGC AAACCGCCGA GTGGAGCGTC
2301 GCCGGCACAG CACGGCTGCT CGCGGCAGCA GAAGAGGAAG ATGCAGCGGT
      CGGCCGTGTC GTGCCGACGA GCGCCGTCTT CTTCTCCTTC TACGTGCGCA
2351 GGAAGGCGTC CGGGCGGCCA GGCAGCGGCG CATAACCTG CAGCAGGAAG
      CCTTCCGCAG GCCCGCCGGT CCGTCGCCGC GTATGTGGAC GTCGTCTTTC
2401 GAGAGCGGGC GGCCGCACAG CTCGCAGGCC AGGGCCTGGG GCCCCGGCAG
      CTCTCGCCCG CCGGCGTGTC GAGCGTCCG TCCCGGACCC CGGGGCCGTC
2451 CCGGCGCGCG CCCAGCCATG CCGGCCGCCC GCCCACCTTG CTGGGGAAC
      GGGCCGGCGC GGGTCGGTAC GGCCGGCGGG CGGGTGGAAC GACCCCTTGA
2501 GCTCGCTGCG CAGTCGCCAC GCCGGCGCGG ACTCGGCGAA GCCAGCTCC
      CGAGCGACGC GTCAGCGGTG CGGCCGCGGC TGAGCCGCTT CGGGTCGAGG
2551 ACAGGCCTGG CCGCGGCGGC AGCCATGCGG GCGCGGGCT GCGTCTGACG
      TGTCCGGACC GGGGCCGCCG TCGGTACGCC CCGCGCCCGA CCGACCCCG
2601 GCAGCCACA GCTGGGTGCG AAGGCGGAAA TCGGGCGCCG GGCCGGAAGG
      CGTCGGGTGT CGACCCAGCC TTCCGCCTTT AGCCGCGGCG CCGGCCTTCC
2651 CAAGAGGCGG GCACCTTTCC GGAGGACAGG AGGCGGAAAC GCGTCTGACG
      GTTCTCCGCC CGTGGAAGG CCTCCTGTCC TCCGCCTTTG CGCAGACTGC
2701 GGAGCGGTTG CAGGACCAAT GCGAGGGAAC GGGGAGAGG AAACCTCTCG
      CCTCGCCAAC GTCCTGGTTA CGCTCCCTTG CCGGCTCTCC TTTGGAGAGC
2751 GCATCAGCCC CGCCCTGCG GCCTCTGCCT CCGAGCCGCT TTCCTGGTGC
      CGTAGTCGGG GCGGGGACCG CGGAGACGGA GGCTCGGCGA AAGGACCACG

```

FIG. 31(contd.) 89/93

```

2801 CTCCGGGTGC TCTGGGATGG TTCTGGTCTT TGGGAGAGTG GCAGCTGGTG
      GAGGCCACG AGACCCTACC AAGACCAGAA ACCCTCTCAC CGTCGACCAC
2851 ACGGCGCTCC GCTCACCTCT GCACATGTCT TGCTGTGGGC CTGCGGGTGG
      TGCCGCGAGG CGAGTGGAGA CGTGTACAGA ACGACACCCG GACGCCCACC
2901 CCGCCAGGGA GGCAGAGCCC TCCCGCAAAC CTTCCTTGCT GGTGTCCACC
      GCGGTCCCT CCGTCTCGGG AGGGCGTTTG GAAGGGACGA CCACAGGTGG
2951 TCAGGGTGTG GGAAACCTGT GCGCTGGCCG AGTGCTAACC AAGAGTAGGC
      AGTCCCACAC CCTTTGGACA CGCGACCGGC TCACGATTGG TTCTCATCCG
3001 AGTGAAAGAC AAATGAAGGT TGAACAGGTA AAGTGAGGAC CCTACAGCGG
      TCACTTTCTG TTTACTTCCA ACTTGTCCAT TCACTCCTG GGATGTGCGC
3051 AAACCAAGAA TCCTGTGTGC CTGAGAGTAA TGAAGAAGCC TCTGCAGAAG
      TTTGGTCTT AGGACACACG GACTCTCATT ACTTCTTCGG AGACGTCTTC
3101 AGTCTTTTCT GTCAGTCTTA AGGTCTCTGT TTTAATGTTA GTGCTGGCTT
      TCAGAAAAGA CAGTCAGAAAT TCCAGAGACA AAATTACAAT CACGACCGAA
3151 GCTGTACCTG AATTCCAAGG GAGGAGTGTA TAATGAGGCA TGGCCAACCC
      CGACATGGAC TTAAGGTTCC CTCCTCACAT ATTACTCCGT ACCGGTTGGG
3201 CCACTTCCCA TCATTGCCTG AACTAGTTTT TCAGGTAAAC TTCAGAAATGC
      GGTGAAGGGT AGTAACGGAC TTGATCAAAA AGTCCAATTG AAGTCTTACG
3251 CCTTGTTACC GCGGGCCCCC TCTGTGGTCC CACGCCACTG ATCGCTGCAT
      GGAACCATGG GCGCCGGGGG 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 TTCCAGGCT GGCTTTGAGT TCCTAGCACC
      TCCCCTGTT 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 GCCCCCCCCG CCGCAGTGC CTAGGTCACC CACTAACGCC CCAGGCCTGG
      CGGGGGGGCG GGGCGTCACG GATCCAGTGG GTGATTGCGG GGTCCGGACC
3651 TCTTGGCTGG GCGTGA CTGT TACCCTCAA AGCAGGCAGC TCCAGGGTAA
      AGAACCGACC CGCACTGACA ATGGGAGTTT TCGTCCGTCG AGGTCCCATT
3701 AAGGTGCCCT GCCCTGTAGA GCCCACTTCC TTCCAGGGC TCGCGCTGGG
      TTCCACGGGA CGGGACATCT CGGGTGAAGG AAGGGTCCCG ACGCCGACCC
3751 TAGGTTTGTG GCCTTCATCA CGGGCCACCT CCAGCCACTG GACCGCTGGC
      ATCCAAACAT CGGAAGTAGT GCCCGGTGGA GGTCCGGTGAC CTGGCGACCG
3801 CCTTGCCCTG TCCTGGGGAG TGTGGTCTTG CGACTCTAAT GGCCGCAAGC
      GGGACGGGAC AGGACCCCTC ACACCAGGAC GCTGAGATTA CCGGCGTTCG
3851 CACCTGACTC CCCCACACC ACACTCTACC TCTCAAGCCC AGGTCTCTCC
      GTGGACTGAG GGGGTGTGG TGTGAGATGG AGAGTTCGGG TCCAGAGAGG
3901 CTAGTGACCC ACCCAGCACA TTTAGCTAGC TGAGCCCCAC AGCCAGAGGT
      GATCACTGGG TGGGTCTGTG 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 GCGCCTGGCA CCGTGCAGA CCCTGGCCCA CCCACCTGG
      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 GACGGGTGCG ACCCGCGGCC CCGGTGCCG CGCGCGGGGG

```

FIG. 31(contd.) 90/93

4251	AGCCCCCGGG	CCCAGCACCC	CAAGGCGGCC	AACGCCAAAA	CTCTCCCTCC
	TCGGGGGGCCC	GGGTCTGTGG	GTTCGCCCGG	TTGCGGTTTT	GAGAGGGAGG
4301	TCCTCTTCCT	CAATCTCGCT	CTCGCTCTTT	TTTTTTTTTCG	CAAAAGGAGG
	AGGAGAAGGA	GTTAGAGCGA	GAGCGAGAAA	AAAAAAAAAGC	GTTTTCTCTC
4351	GGAGAGGGGG	TAAAAAAATG	CTGCACTGTG	CGGCGAAGCC	GGTGAGTGAG
	CCTCTCCCCC	ATTTTTTTTAC	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	CCGCGCAGCA	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	CATGGGGGCG	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	AGCCCGCCCC	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	GCGCGACACG	GAAACCCGCG	ATTGACGCAC	GCGCGACCCCT
4951	ATTGGCGCTA	ATTGCGCGTG	CGCGCTGGGA	CTCAATGGCG	CTAATCGCGC
	TAACCGCGAT	TAACGCGCAC	GCGCGACCCCT	GAGTTACCGC	GATTAGCGCG
5001	GTGCGTTCTG	GGGCCCCGGG	GCTTGCGCCA	CTTCCTGCCC	GAGCCGCTGG
	CACGCAAGAC	CCCGGGCCCC	CGAACGCGGT	GAAGGACGGG	CTCGGCGACC
5051	CGCCCGAGGG	TGTGGCCGCT	GCGTGCGCGC	GCGCGACCCG	CTCGCTGTTT
	GCGGGCTCCC	ACACCGGGCA	CGCACGCGCG	CGCGCTGGGC	CAGCGACAAA
5101	GAACCGGGCG	GAGGCGGGGC	TGGCGCCCCG	TTGGGAGGGG	GTTGGGGCCT
	CTTGGCCCGC	CTCCGCCCCG	ACCGCGGGCC	AACCCTCCCC	CAACCCCGGA
5151	GGCTTCCTGC	CGCGCGCCGC	GGGGACGCCT	CCGACCAGTG	TTTGCCTTTT
	CCGAAGGACG	GCGCGCGGGC	CCCCTGCGGA	GGCTGGTCAC	AAACGGA AAA
5201	ATGGTAATAA	CGCGGCCGCG	CCGGCTTCCT	TTGTCCCCAA	TCTGGGCGCG
	TACCATTATT	GCGCCGGCCG	GGCCGAAGGA	AACAGGGGTT	AGACCCGCGC
5251	CGCCGGCGCC	CCCTGGCGGC	CTAAGGACTC	GGCGCGCCGG	AAGTGGCCAG
	GCGGCCGCGG	GGGACCGCCG	GATTCCTGAG	CCGCGCGGCC	TTACCCGGTC
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	AATCCACAC	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.) 91/93

```

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 TGAATGGGCA
      GTCCAAGAGG CCGGCGAACC CACCTCTCCG ATAAGCCGAT ACTGACCCGT
5901  CAACAGACAA TCGGCTGCTC TGATGCCGCC GTGTTCCGGC TGTGAGCGCA
      GTTGTCTGTT AGCCGACGAG ACTACGGCGG CACAAGGCCG ACAGTCGCGT
5951  GGGGCGCCCG GTTCTTTTTG TCAAGACCGA CCTGTCCGGT GCCCTGAATG
      CCCC GCGGGC CAAGAAAAAC AGTTCTGGCT GGACAGGCCA CGGGACTTAC
6001  AACTGCAAGA CGAGGCAGCG CGGCTATCGT GGCTGGCCAC GACGGGCGTT
      TTGACGTTCT GCTCCGTCGC GCCGATAGCA CCGACCGGTG CTGCCCAGCA
6051  CCTTGCGCAG CTGTGCTCGA CGTTGTCACT GAAGCGGGAA GGGACTGGCT
      GGAACGCGTC GACACGAGCT GCAACAGTGA CTTGCGCCCT CCCTGACCGA
6101  GCTATTGGGC GAAGTGCCGG GGCAGGATCT CCTGTCATCT CACCTTGCTC
      CGATAACCCG CTTACAGGCC 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 TGGATTATC GACTGTGGCC
      CTTATAGTAC CACCTTTTAC CGGCGAAAAG ACCTAAGTAG CTGACACCGG
6451  GGCTGGGTGT GCGGACCGC TATCAGGACA TAGCGTTGGC TACCCGTGAT
      CCGACCCACA CCGCCTGGCG ATAGTCCTGT ATCGCAACCG ATGGGCACTA
6501  ATTGCTGAAG AGCTTGCGCG CGAATGGGCT GACCGCTTCC TCGTGCTTAA
      TAACGACTTC TCGAACC GCCATTACCGA 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 GATTTGATG CCACCGCCGC CTTCTATGAA
      GCGGGTTGGA CGGTAGTGCT CTAAAGCTAA GGTGGCGGCG GAAGATACTT
6701  AGGTTGGGCT TCGGAATCGT TTTCCGGGAC GCCGGCTGGA TGATCCTCCA
      TCCAACCCGA AGCCTTAGCA AAAGGCCCTG CGGCCGACCT ACTAGGAGGT
6751  GCGCGGGGAT CTCATGCTGG AGTTCTTCGC CCACCCTAGG GGGAGGCTAA
      CGCGCCCTTA GAGTACGACC TCAAGAAGCG GGTGGGATCC CCCTCCGATT
6801  CTGAAACACG GAAGGAGACA ATACCGGAAG GAACCCGCGC TATGACGGCA
      GACTTTGTGC CTTCTCTGT TATGGCCTTC CTTGGGCGCG ATACTGCCGT
6851  ATAAAAAGAC AGAATAAAAC GCACGGTGTT GGGTCGTTTG TTCATAAACG
      TATTTTCTG TCTTATTTTG CGTGCCACAA CCCAGCAAAC AAGTATTTGC
6901  CGGGGTTTGG TCCCAGGGCT GGCACCTCTG 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 GCGTCCGGG
7051  TGCCATAGCC TCAAATTCTT TAGGCTCGAG GGCCGCCACC GCGGCTGGAGC
      ACGGTATCGG AGTTTAAGGA ATCCGAGCTC CCGCGGTGG GCGCACCTCG
7101  TCCAGCTTTT GTTCCCTTTA GTGAGGGTTA ATTTGAGCT TGGCGTAATC
      AGGTCGAAA CAAGGGAAAT CACTCCCAAT TAAAGCTCGA ACCGCATTAG

```

FIG. 31(contd.) 92/93

7151	ATGGTCATAG	CTGTTTCCTG	TGTGAAATTG	TTATCCGCTC	ACAATTCCAC
	TACCAGTATC	GACAAAGGAC	ACACTTTAAC	AATAGGCGAG	TGTTAAGGTG
7201	ACAACATACG	AGCCGGAAGC	ATAAAGTGTA	AAGCCTGGGG	TGCCTAATGA
	TGTTGTATGC	TCGGCCTTCG	TATTTACAT	TTCGGACCCC	ACGATTACT
7251	GTGAGCTAAC	TCACATTAAT	TGCGTTGCGC	TCACTGCCCG	CTTCCAGTC
	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	TTTTCTTCT
7451	GTATGAGTAT	TCAACATTTT	CGTGTGCGCC	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
	TGTCCGCATT	CTAGGAACTC	TCAAAAGCGG	GGCTTCTTGC	AAAAGGTTAC
7651	ATGAGCACTT	TTAAAGTTCT	GCTATGTGGC	GCGGTATTAT	CCCGTATTGA
	TACTCGTGAA	AATTTCAAGA	CGATACACCG	CGCCATAATA	CGGCATAACT
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	TCAACGTCCCT
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	GTAAGTTATCT	ACACGACGGG	GAGTCAGGCA
	TACCATTCTG	GAGGGCATAG	CATCAATAGA	TGTGCTGCCC	CTCAGTCCGT
8251	ACTATGGATG	AACGAAATAG	ACAGATCGCT	GAGATAGGTG	CCTCACTGAT
	TGATACCTAC	TTGCTTTATC	TGTCTAGCGA	CTCTATCCAC	GGAGTGACTA
8301	TAAGCATTGG	TAAGTGTGAG	ACTCGCGACA	CTGCATTAAT	GAATCGGCCA
	ATTGCTAACC	ATTGACAGTC	TGAGCGCTGT	GACGTAATTA	CTTAGCCGGT
8351	ACGCGCGGGG	AGAGGCGGTT	TGCGTATTGG	GCGCTCTTCC	GCTTCCTCGC
	TGCGCGCCCC	TCTCCGCCAA	ACGCATAACC	CGCGAGAAGG	CGAAGGAGCG
8401	TCACTGACTC	GCTGCGCTCG	GTCGTTCCGG	TGCGGCGAGC	GGTATCAGCT
	AGTGACTGAG	CGACGCGAGC	CAGCAAGCCG	ACGCCGCTCG	CCATAGTCGA
8451	CACTCAAAGG	CGGTAATACG	GTTATCCACA	GAATCAGGGG	ATAACGCAGG
	GTGAGTTTCC	GCCATTATGC	CAATAGGTGT	CTTAGTCCCC	TATTGCGTCC
8501	AAAGAACATG	TGAGCAAAAAG	GCCAGCAAAA	GGCCAGGAAC	CGTAAAAAGG
	TTTCTTGTAC	ACTCGTTTTT	CGGTCGTTTT	CCGGTCCTTG	GCATTTTTCC
8551	CCGCGTTGCT	GGCGTTTTTC	CATAGGCTCC	GCCCCCTTGA	CGAGCATCAC
	GGCGCAACGA	CCGCAAAAAG	GTATCCGAGG	CGGGGGGACT	GCTCGTAGTG

93/93

8601 AAAAATCGAC GCTCAAGTCA GAGGTGGCGA AACCCGACAG GACTATAAAG
TTTTTAGCTG CGAGTTCAGT CTCCACCGCT TTGGGCTGTC CTGATATTTC
8651 ATACCAGGCG TTTCCCCCTG GAAGCTCCCT CGTGCGCTCT CCTGTTCCGA
TATGGTCCGC AAAGGGGGAC CTTGAGGGA GCACGCGAGA GGACAAGGCT
8701 CCCTGCCGCT TACCGGATAC CTGTCCGCCT TTCTCCCTTC GGGAAAGCGT
GGGACGGCGA ATGGCCTATG GACAGGCGGA AAGAGGGAAG CCCTTCGCAC
8751 GCGCTTTCTC ATAGCTCACG CTGTAGGTAT CTCAGTTCGG TGTAGGTCGT
CGCGAAAGAG TATCGAGTGC GACATCCATA GAGTCAAGCC ACATCCAGCA
8801 TCGCTCCAAG CTGGGCTGTG TGCACGAACC CCCCCTTCAG CCCGACCGCT
AGCGAGGTTC GACCCGACAC ACGTGCTTGG GGGGCAAGTC GGGCTGGCGA
8851 GCGCCTTATC CGGTAACAT CTCTTGAGT 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 GGAAAAGAG 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
GAGTCTTCT AGGAAACTAG AAAAGATGCC CCAGACTGCG AGTCACCTTG
9201 GAAAACTCA
CTTTTGAGT

FIG. 31(contd.)