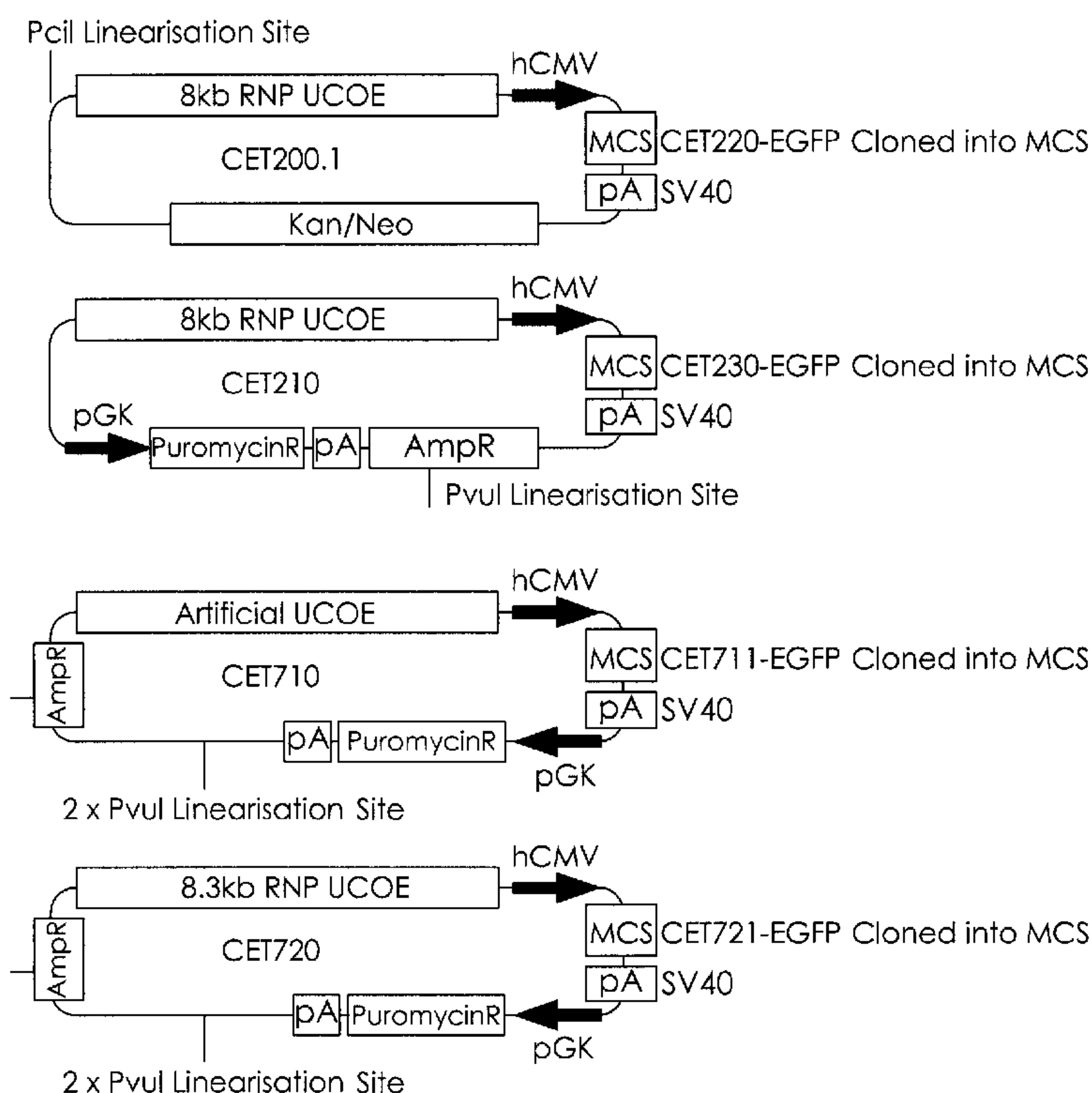




(86) Date de dépôt PCT/PCT Filing Date: 2002/04/05
(87) Date publication PCT/PCT Publication Date: 2002/10/17
(45) Date de délivrance/Issue Date: 2013/02/05
(85) Entrée phase nationale/National Entry: 2003/10/02
(86) N° demande PCT/PCT Application No.: GB 2002/001479
(87) N° publication PCT/PCT Publication No.: 2002/081677
(30) Priorités/Priorities: 2001/04/05 (US60/281,605);
2001/04/05 (GB0108605.7); 2001/04/17 (GB0109335.0);
2001/06/15 (US60/298,675)

(51) Cl.Int./Int.Cl. *C12N 15/85* (2006.01),
C12N 15/11 (2006.01)
(72) Inventeurs/Inventors:
WILLIAMS, STEVEN GERAINT, GB;
CROMBIE, ROBERT LACHLAN, GB
(73) Propriétaire/Owner:
EMD MILLIPORE CORPORATION, US
(74) Agent: CAMERON IP

(54) Titre : EXPRESSION GENIQUE AMELIOREE
(54) Title: IMPROVED GENE EXPRESSION



(57) **Abrégé/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.



(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(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 A3

(51) International Patent Classification⁷: **C12N 15/85**,
205/00, 219/00, 217/00, 15/11, 203/00

CROMBIE, Robert, Lachlan; Cobra Therapeutics Limited, Stephenson Building, Keele University Science Park, Keele, Staffordshire ST5 5SP (GB).

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

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

(74) Agents: **JONES, Keith, William**; Harrison Goddard Foote, Orlando House, 11c Compstall Road, Marple Bridge, Stockport SK6 5HH et al. (GB).

(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

(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.

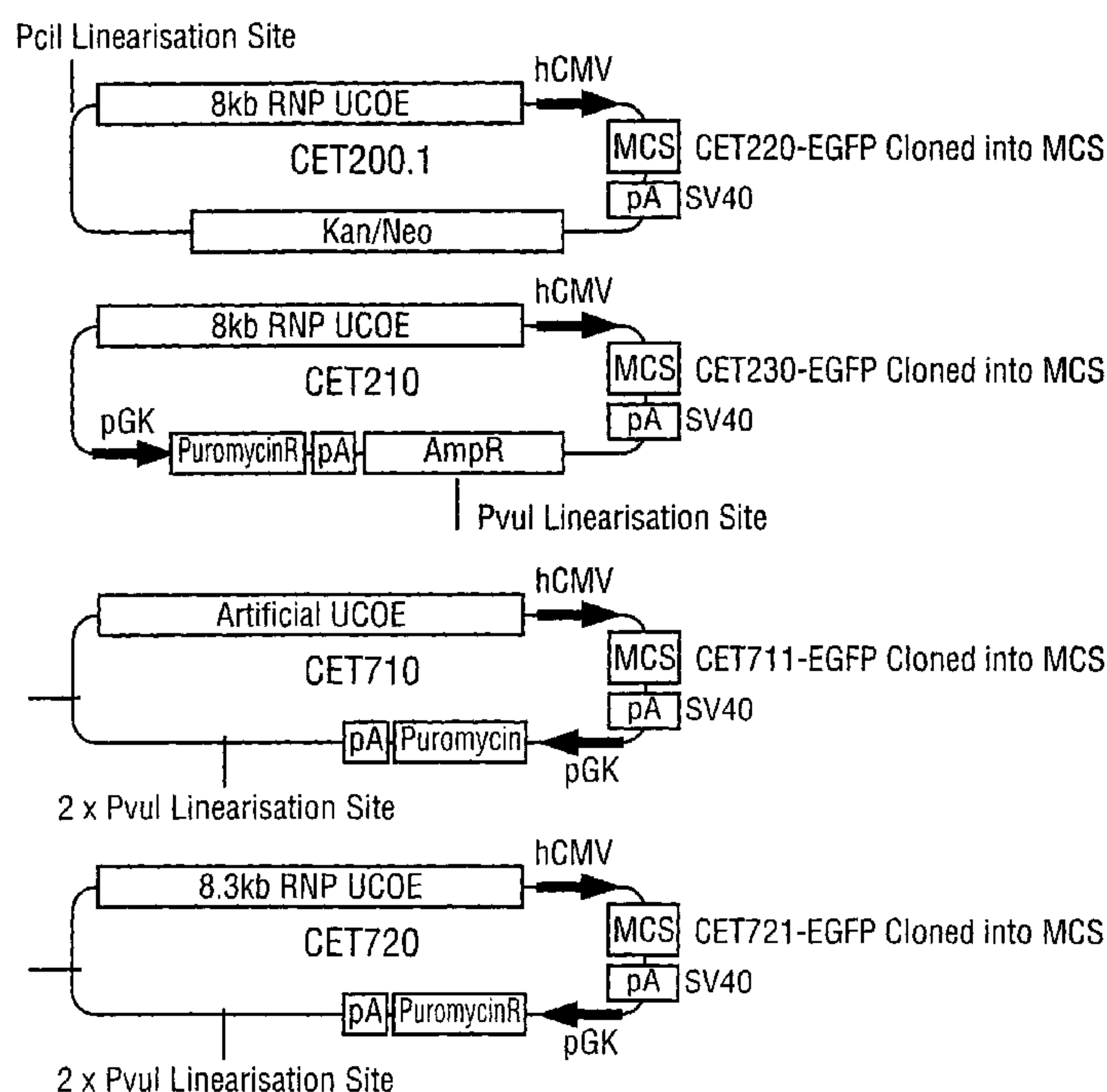
(71) Applicant: **ML LABORATORIES PLC** [GB/GB]; 17 Hanover Square, London W1S 1HU (GB).

(72) Inventors: **WILLIAMS, Stephen, Geraint**; Cobra Therapeutics Limited, Stephenson Building, Keele University Science Park, Keele, Staffordshire ST5 5SP (GB).

(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 A3

WO 02/081677 A3



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

(88) Date of publication of the international search report:
19 December 2002

Published:

- *with international search report*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

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.

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. Rev. 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, Millett 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.

5 Our co-pending patent applications PCT/GB1999/002357 published as WO 00/05393 and
US 09/358082 published as US 2002/0106789, describe elements that are responsible, in
their natural chromosomal context, for establishing an open chromatin structure across a
locus that consists exclusively of ubiquitously expressed, housekeeping genes. These
elements are not derived from an LCR and comprise extended methylationfree CpG
10 islands. We have used the term Ubiquitous Chromatin Opening Element (UCOE) to
describe such elements.

In mammalian DNA, the dinucleotide CpG is recognised by a DNA methyltransferase
enzyme that methylates cytosine to 5-methylcytosine. However, 5-methylcytosine is
15 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 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
20 ratio of at least 0.6 (i.e. a CpG dinucleotide content of at least 60% of that which would be
expected by 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,
25 Tazi and Bird (1990) *Cell* 60: 909-920) and may be defined as 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
30 associated with 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 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 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* 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".

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 (*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.

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 published as WO 00/005393 and
US 09/358082 published as US 2002/0106789, 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 selectable marker is likewise not disclosed or implied.

STATEMENTS OF THE INVENTION

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.

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.

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
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 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 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 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 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.

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 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

WO 02/081677

10

PCT/GB02/01479

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 of the 3' end of the expressible nucleic acid sequence, as marked by its
5 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 a *Streptomyces* species. More preferably, said antibiotic resistance gene is
10 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), (SEQ ID NO: 5)

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 14 (SEQ ID NO: 3).

30

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

WO 02/081677

11

PCT/GB02/01479

phosphotransferase gene from *Streptomyces fradiae* (Thompson CJ and Gray GS. *Proc Natl Acad Sci USA* 80, 5190-5194 (1983) (SEQ ID NO: 4).

In an alternative embodiment, the antibiotic resistance gene is a hygromycin resistance gene. Preferably, it is the hygromycin phosphotransferase gene
5 from *Streptomyces hygrosopicus* (SEQ ID NO: 6).

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-
10 acetyltransferase from *Streptomyces verticillus*.

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

15

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* (SEQ ID NO: 7).

20

In another preferred embodiment, it is the neomycin phosphotransferase gene from transposon Tn5, originally derived from *Klebsiella pneumoniae* (SEQ ID NO: 8).

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).

In a further aspect, a screenable marker can be used. For instance, a gene
5 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
be sorted by brightness of fluorescence on a FACS, by a process well-known
10 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 PD CD2 CpG island/ promoter
30 region or a fragment thereof. More preferably the human PD CD2 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 PDCD2 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 PDCD2 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
5 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
10 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.

15 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
20 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
25 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
30 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 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 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 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 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. 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.

A preferred embodiment of the vector comprises nucleotides 1-10551 of the sequence of Figure 10 (SEQ ID NO: 1). A most preferred embodiment is vector CET 710.

Alternatively, the vector comprises nucleotides 1-13545 of the sequence of Figure 12, and is preferably vector CET 720 (SEQ ID NO: 2).

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, SEQ ID NO: 4) Also preferred are vectors
10 having expressible 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
15 with the aminocyclitol phosphotransferase from *Escherichia coli* (as listed in Figure 17, SEQ ID NO: 7). Also preferred are vectors having expressible nucleic acid sequences inserted into the multiple cloning site of CET 760, such as CET 761.

20 CET 780 in which the puromycin resistance gene of CET 720 is replaced with the modified form of the puromycin N-acetyl transferase gene from *Streptomyces alboniger* (as listed in Figure 14, SEQ ID NO: 3). 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
30 sequences inserted into the multiple cloning site of CET 820, such as CET 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
5 extended methylation-free CpG island comprising an 8kb fragment spanning the murine hnRNP A2 gene (as shown in the sequence of Figure 19, SEQ ID NO: 15). Also preferred are vectors having expressible nucleic acid sequences inserted into the multiple cloning site of CET 823, such as CET 824.

10

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

Alternatively said polynucleotide, vector or the host cell may be used in a cell
15 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 the production of a polypeptide according to the invention comprising:

- 20 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
- iii) purifying said polypeptide from said cell, or its growth environment.

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

The present invention also provides the polynucleotide, vector or the host cell for use in 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 μ g of vector DNA/kg body weight; and more preferably in the range of between about 1-100 μ g 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-
- 15 based drug screening.
- 20 (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
- 25 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.
- 30

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 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 the FACScan. This clearly shows the comparative effect of placing the selectable marker (puro') either 5' (CET230) or 3' (CET721) to the expressible transgene (EGFP).

Figure 5 shows the expression of EGFP in CHO-K1 cells transfected with 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 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' 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

Figure 9 shows a map of plasmid CET 710

Figure 10 shows the nucleotide sequence of CET710 (SEQ ID NO: 1)

Figure 11 shows a map of plasmid CET 720

WO 02/081677

25

PCT/GB02/01479

Figure 12 shows the nucleotide sequence of CET720 (SEQ ID NO: 2)

Figure 13 shows the nucleotide sequence of the wild-type *S. alboniger* puromycin N-acetyl transferase gene (SEQ ID NO: 5).

5

Figure 14 shows the nucleotide sequence of the modified *S. alboniger* puromycin N-acetyl transferase gene (SEQ ID NO: 3).

Figure 15 shows the nucleotide sequence of the *S. fradiae* aminoglycoside phosphotransferase gene (SEQ ID NO: 4).

10

Figure 16 shows the nucleotide sequence of the *S. hygroscopicus* hygromycin phosphotransferase gene (SEQ ID NO: 6).

Figure 17 shows the nucleotide sequence of the *E. coli* aminocyclitol phosphotransferase (hygro^r) gene (SEQ ID NO: 7).

15

Figure 18 shows the nucleotide sequence of the transposon Tn5 (*Klebsiella pneumoniae*) neomycin phosphotransferase gene (SEQ ID NO: 8).

20

Figure 19 shows the nucleotide sequence of the mouse hnRNP A2 *Hind*III fragment (SEQ ID NO: 15).

Figure 20 shows a map of plasmid CET 1010

25

Figure 21 shows the nucleotide sequence of CET1010 (SEQ ID NO: 9)

Figure 22 shows a map of plasmid CET 1020

Figure 23 shows the nucleotide sequence of CET1020 (SEQ ID NO: 10)

30

Figure 24 shows a map of plasmid CET 1030

WO 02/081677

26

PCT/GB02/01479

Figure 25 shows the nucleotide sequence of CET1030 (SEQ ID NO: 11)

5 Figure 26 shows a map of plasmid CET 1110

Figure 27 shows the nucleotide sequence of CET1110 (SEQ ID NO: 12)

Figure 28 shows a map of plasmid CET 1120

10

Figure 29 shows the nucleotide sequence of CET1120 (SEQ ID NO: 13)

Figure 30 shows a map of plasmid CET 1130

15 Figure 31 shows the nucleotide sequence of CET1130 (SEQ ID NO: 14)

EXAMPLES

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

Materials and Methods

25 **Construction of PGK-Puro CET expression Vectors**

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
30 T4 DNA polymerase and ligated into pPGK-Puro (mPGK promoter, Puromycin resistance gene, bGHpA in pBluescript) that had been digested 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/AflII* 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 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 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 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.

25

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.

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 vector comprising an isolated polynucleotide, said polynucleotide comprising:
 - a. an extended methylation-free CpG island encompassing dual, divergently transcribed promoters;
 - b. an expressible nucleic acid terminated by a polyadenylation signal;
 - c. a selectable marker gene operably linked to a promoter;wherein the vector is capable of linearization and integration into a chromosome such that 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 gene, 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 2000 base pairs of the proximal end of the selectable marker.
2. The vector of claim 1, wherein the polyadenylation signal at the 3' end of the expressible nucleic acid is within 1500 base pairs of the proximal end of the selectable marker.
3. The vector of claim 2, wherein the polyadenylation signal at the 3' end of the expressible nucleic acid is within 1000 base pairs of the proximal end of the selectable marker.
4. The vector of claim 3, wherein the polyadenylation signal at the 3' end of the expressible nucleic acid is within 500 base pairs of the proximal end of the selectable marker.
5. The vector of any one of claims 1 to 4, wherein the selectable marker is an antibiotic resistance gene.
6. The vector of claim 5, wherein the antibiotic resistance gene is obtained from a *Streptomyces* species.

7. The vector of claim 5, wherein the antibiotic resistance gene is selected from the group consisting of: a puromycin resistance gene; a neomycin resistance gene; a hygromycin resistance gene; a bleomycin resistance gene; and a blasticidin resistance gene.
8. The vector of claim 7, wherein the puromycin resistance gene is a puromycin N-acetyl transferase gene from *Streptomyces alboniger*.
9. The vector of claim 7, wherein the neomycin resistance gene is the aminoglycoside phosphotransferase gene from *Streptomyces fradiae*.
10. The vector of claim 7, wherein the hygromycin resistance gene is the hygromycin phosphotransferase gene from *Streptomyces hygrosopicus*.
11. The vector of claim 7, wherein the bleomycin resistance gene is the bleomycin binding protein from *Streptomyces verticillus*.
12. The vector of claim 7, wherein the blasticidin resistance gene is the blasticidin S-acetyltransferase gene from *Streptomyces verticillum*.
13. The vector of any one of claims 1 to 5, wherein the selectable marker is the antibiotic resistance gene aminocyclitol phosphotransferase from *Escherichia coli*.
14. The vector of any one of claims 1 to 5, wherein the selectable marker is the antibiotic resistance gene neomycin phosphotransferase from transposon Tn5.
15. The vector of any one of claims 1 to 14, wherein the extended methylation-free CpG island comprises an 8 kilobase DNA fragment spanning the human heterogeneous nuclear ribonucleoprotein A2/B1 gene.
16. The vector of any one of claims 1 to 14, wherein the extended methylation-free CpG island comprises an 8 kilobase DNA fragment spanning the murine heterogeneous nuclear ribonucleoprotein A2/B1 gene.

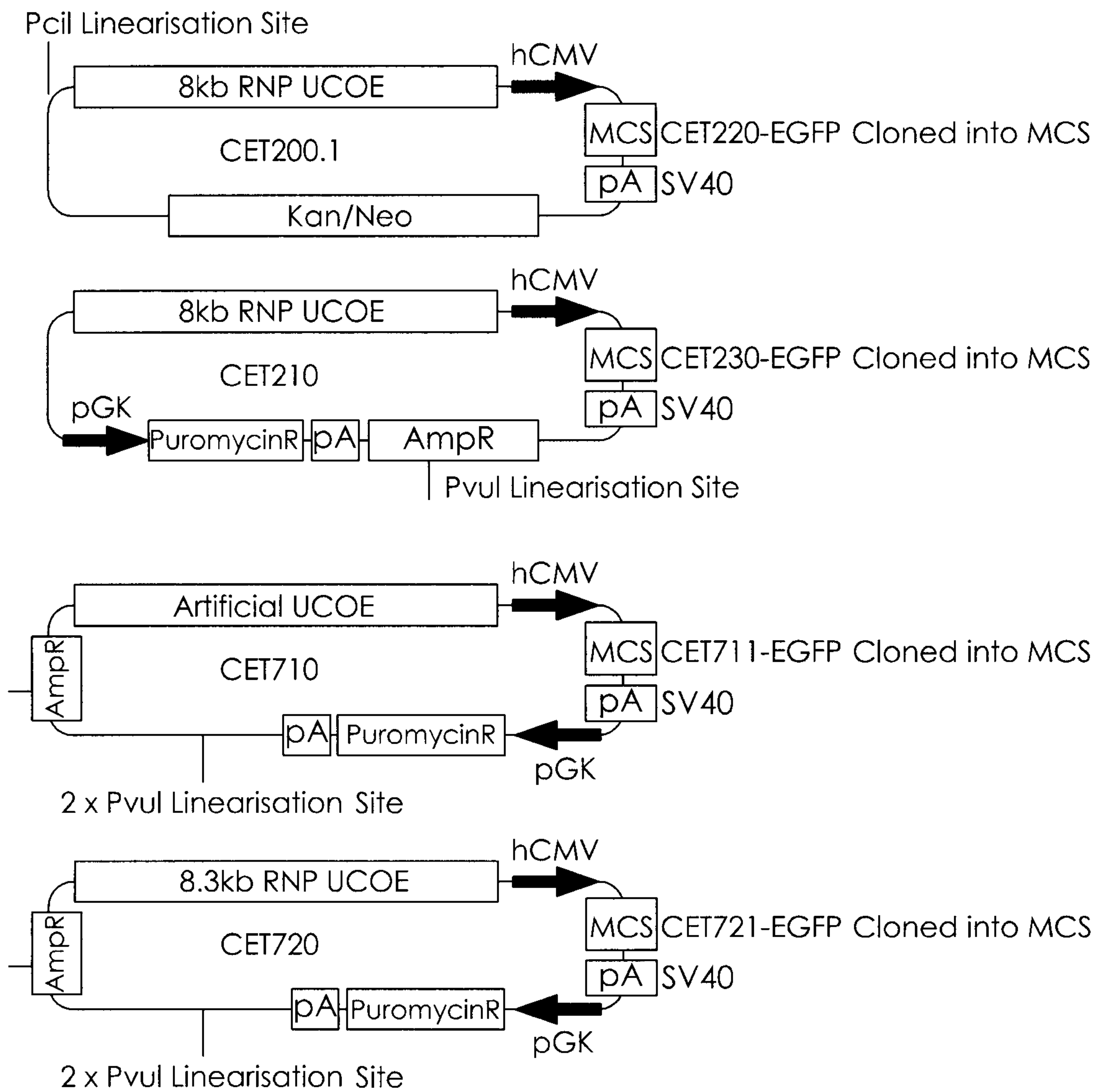
17. The vector of claim 16, wherein the extended methylation-free CpG island comprises nucleotides 1-7898 of the sequence of SEQ ID NO: 15.
18. The vector of any one of claims 1 to 14, wherein the extended methylation-free CpG island comprises a 2.0 kilobase DNA fragment spanning the human β -actin CpG island / promoter region and a 1.8 kilobase DNA fragment spanning the human PDCD2 CpG island/promoter region.
19. The vector of any one of claims 1 to 18, wherein the expressible nucleic acid is a nucleic acid for use as a therapeutic agent.
20. The vector of any one of claims 1 to 18, wherein the expressible nucleic acid encodes a recombinant protein for expression in an *in vitro* cell culture system.
21. The vector of any one of claims 1 to 18, wherein the expressible nucleic acid is contained within a multiple cloning site and the multiple cloning site is further operably linked to a promoter.
22. The vector of claim 21, wherein said promoter is a cytomegalovirus immediate/early promoter.
23. The vector of claim 1 comprising nucleotides 1-10551 of the sequence of SEQ ID NO: 1.
24. The vector of claim 1 comprising nucleotides 1-13545 of the sequence of SEQ ID NO: 2 including an extended methylation-free CpG island comprising an 8 kilobase fragment spanning the human heterogeneous nuclear ribonucleoprotein A2/B1 gene, a puromycin resistance gene and a human IE CMV promoter.
25. The vector of claim 24, in which the puromycin resistance gene is replaced with the aminoglycoside phosphotransferase gene from *Streptomyces fradiae* of the sequence of SEQ ID NO: 4.

26. The vector of claim 24, in which the puromycin resistance gene is replaced with the aminocyclitol phosphotransferase gene from *Escherichia coli* of the sequence of SEQ ID NO: 7.
27. The vector of claim 24, in which the puromycin resistance gene is replaced with a modified form of the puromycin N-acetyl transferase gene from *Streptomyces alboniger* of the sequence of SEQ ID NO: 3.
28. The vector of claim 24, in which the human IE CMV promoter is replaced with a murine IE CMV promoter.
29. The vector of claim 24, in which the extended methylation-free CpG island comprising an 8 kilobase fragment spanning the human heterogeneous nuclear ribonucleoprotein A2/B1 gene is replaced with the extended methylation-free CpG island comprising an 8 kilobase fragment spanning the murine heterogeneous nuclear ribonucleoprotein A2/B1 gene of the sequence of SEQ ID NO: 15.
30. The vector of claim 1 comprising nucleotides 1-12039 of the sequence of SEQ ID NO: 9.
31. The vector of claim 1 comprising nucleotides 1-11646 of the sequence of SEQ ID NO: 10.
32. The vector of claim 1 comprising nucleotides 1-9027 of the sequence of SEQ ID NO: 11.
33. The vector of claim 1 comprising nucleotides 1-12221 of the sequence of SEQ ID NO: 12.
34. The vector of claim 1 comprising nucleotides 1-11828 of the sequence of SEQ. ID. NO:13.

35. The vector of claim 1 comprising nucleotides 1-9209 of the sequence of SEQ. ID NO: 14.
36. A host cell transfected with the vector of any of claims 1 to 35.
37. Use of the vector of any one of claims 1 to 35 to obtain expression of the expressible nucleic acid.
38. Use of the vector of any one of claims 1 to 35 in a cell culture system to obtain expression of a desired gene product.
39. Use of the vector of any one of claims 1 to 35 as a medicament.
40. A pharmaceutical composition comprising the vector of any one of claims 1 to 35 in combination with a pharmaceutically acceptable excipient.
41. The vector of claim 1, wherein linearization is accomplished with a *Pvu I* site.
42. The vector of claim 11, wherein the bleomycin binding protein *Streptomyces verticillus* is a bleomycin N-acetyltransferase from *Streptomyces verticillus*.
43. Use of the host cell of claim 36 to obtain expression of the expressible nucleic acid.
44. Use of the host cell of claim 36 in a cell culture system to obtain expression of a desired gene product.
45. Use of the host cell of claim 36 as a medicament.
46. A pharmaceutical composition comprising the host cell of claim 36 in combination with a pharmaceutically acceptable excipient.

1/93

FIG. 1



2/93

FIG. 2

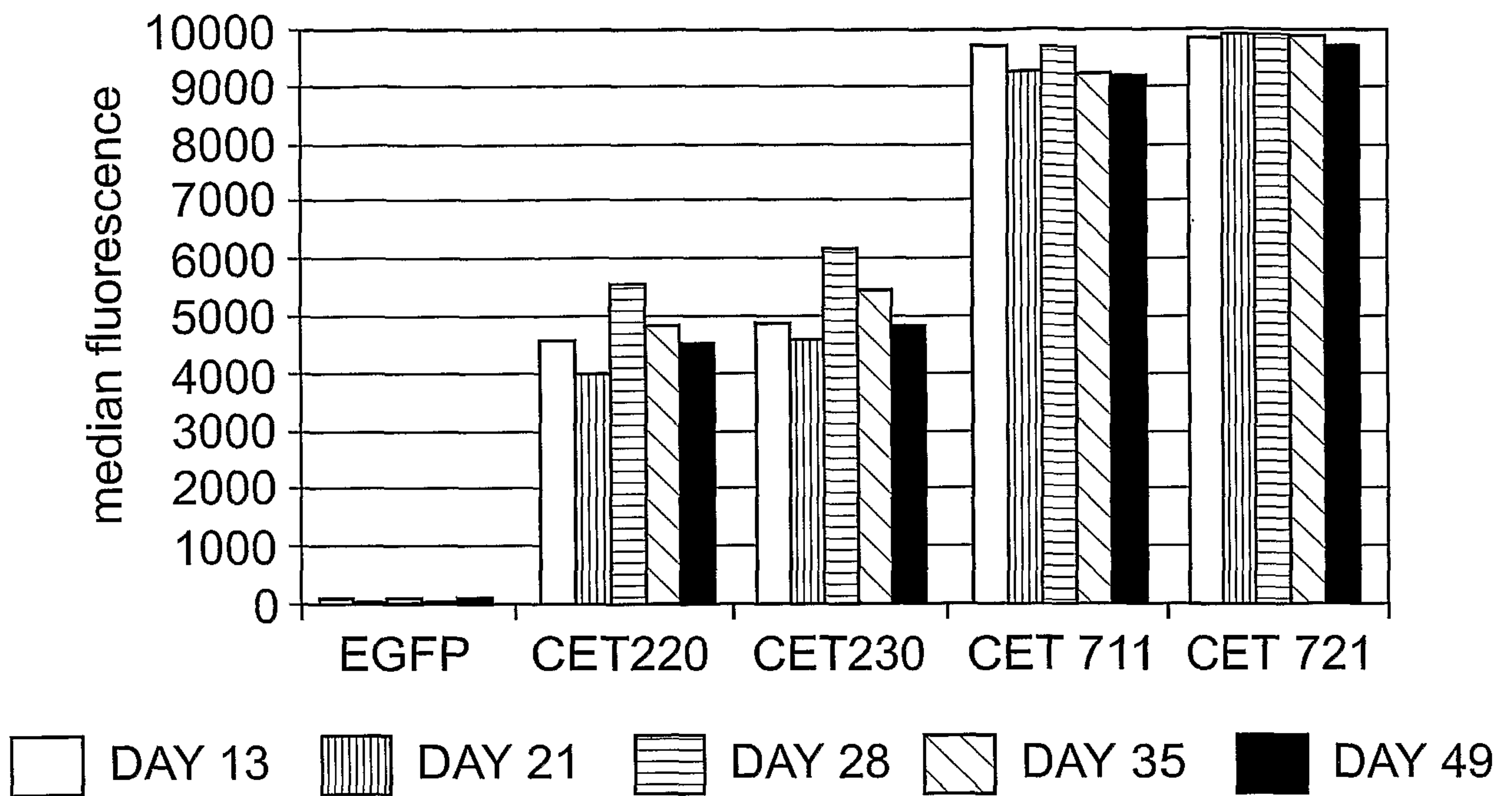
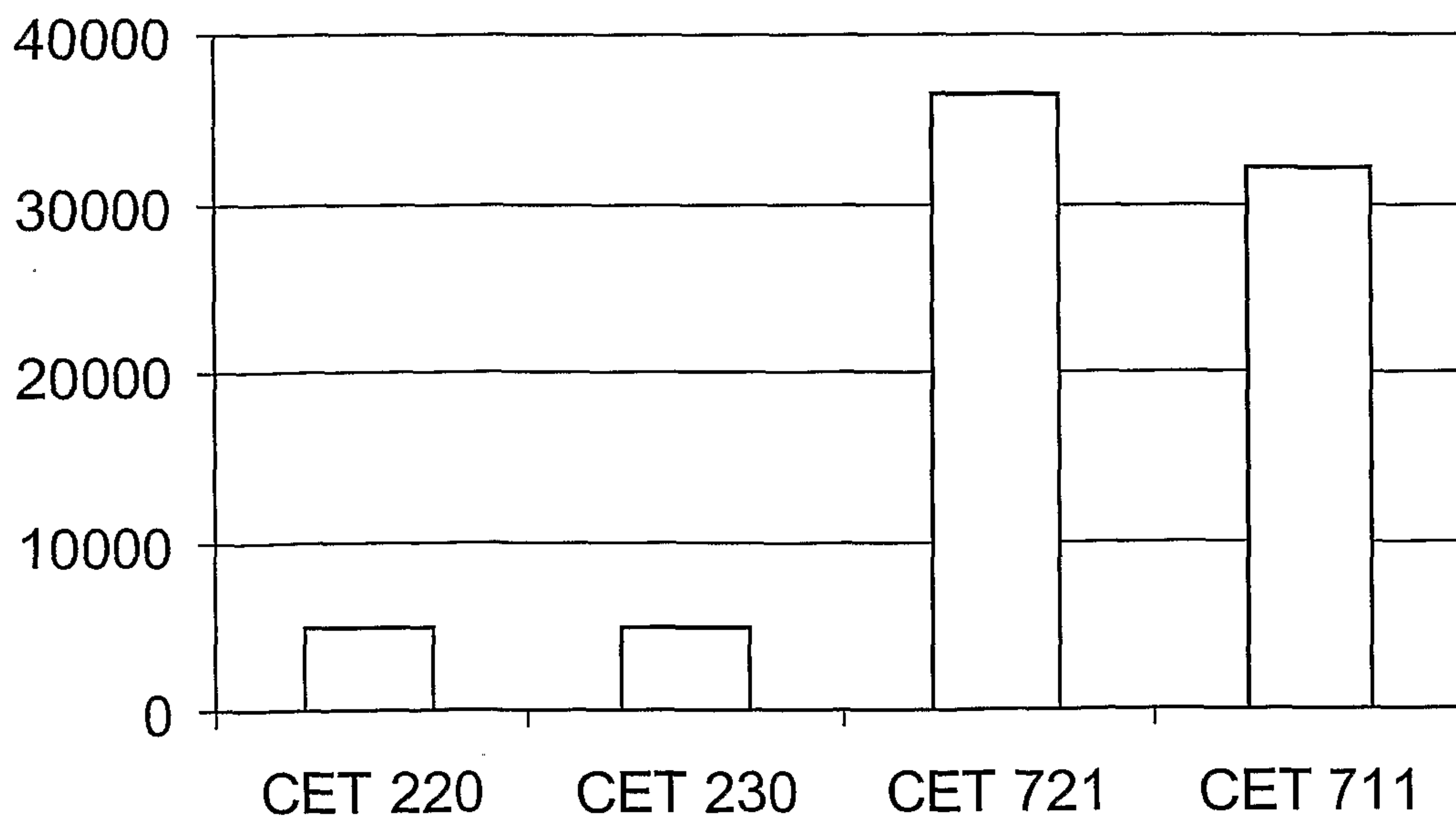


FIG. 4



3/93

FIG. 3

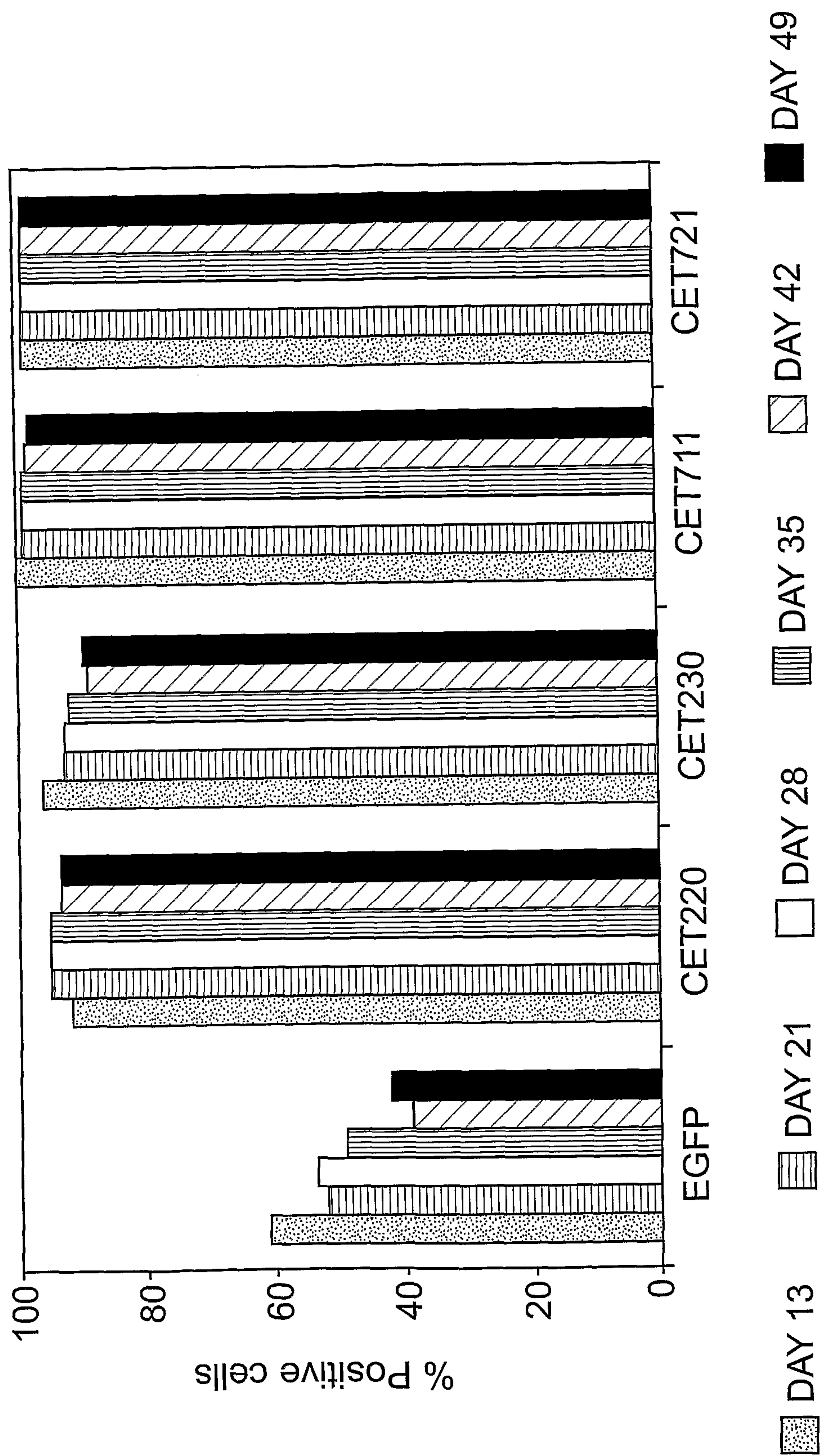


FIG. 5
Expression of EGFP in CHOK-1 Cells using Vectors with
Different Antibiotic Resistance Genes

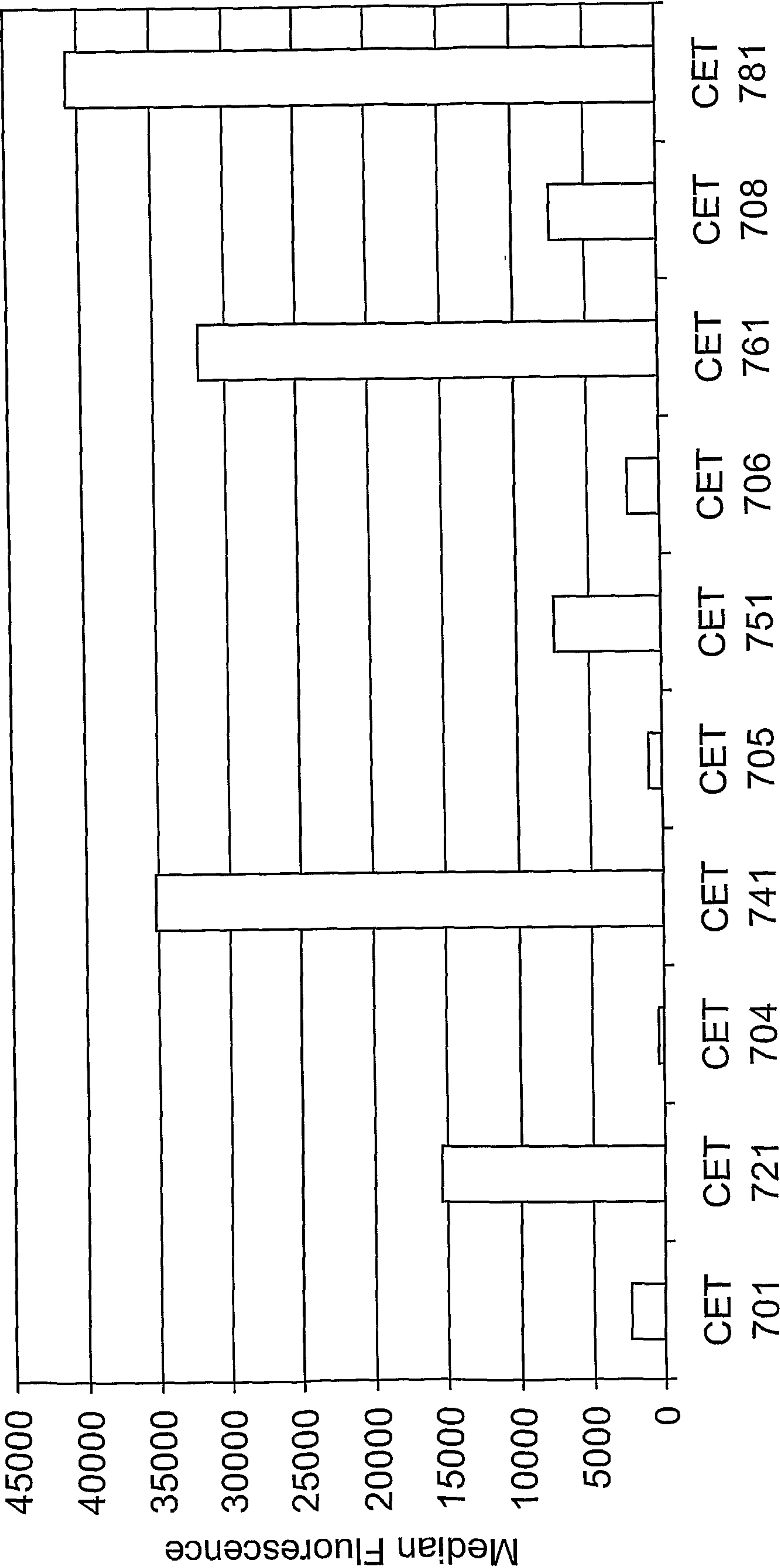
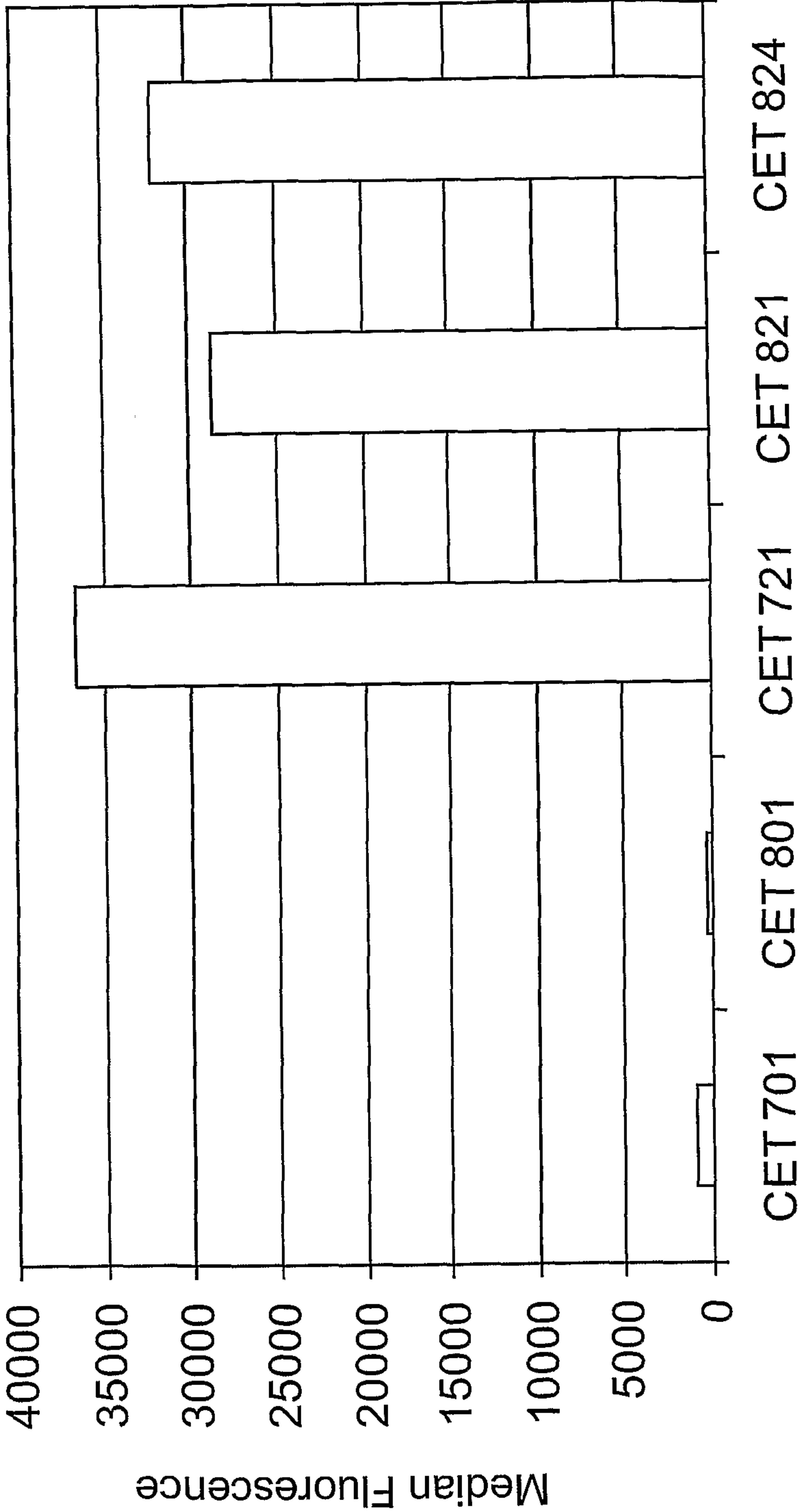
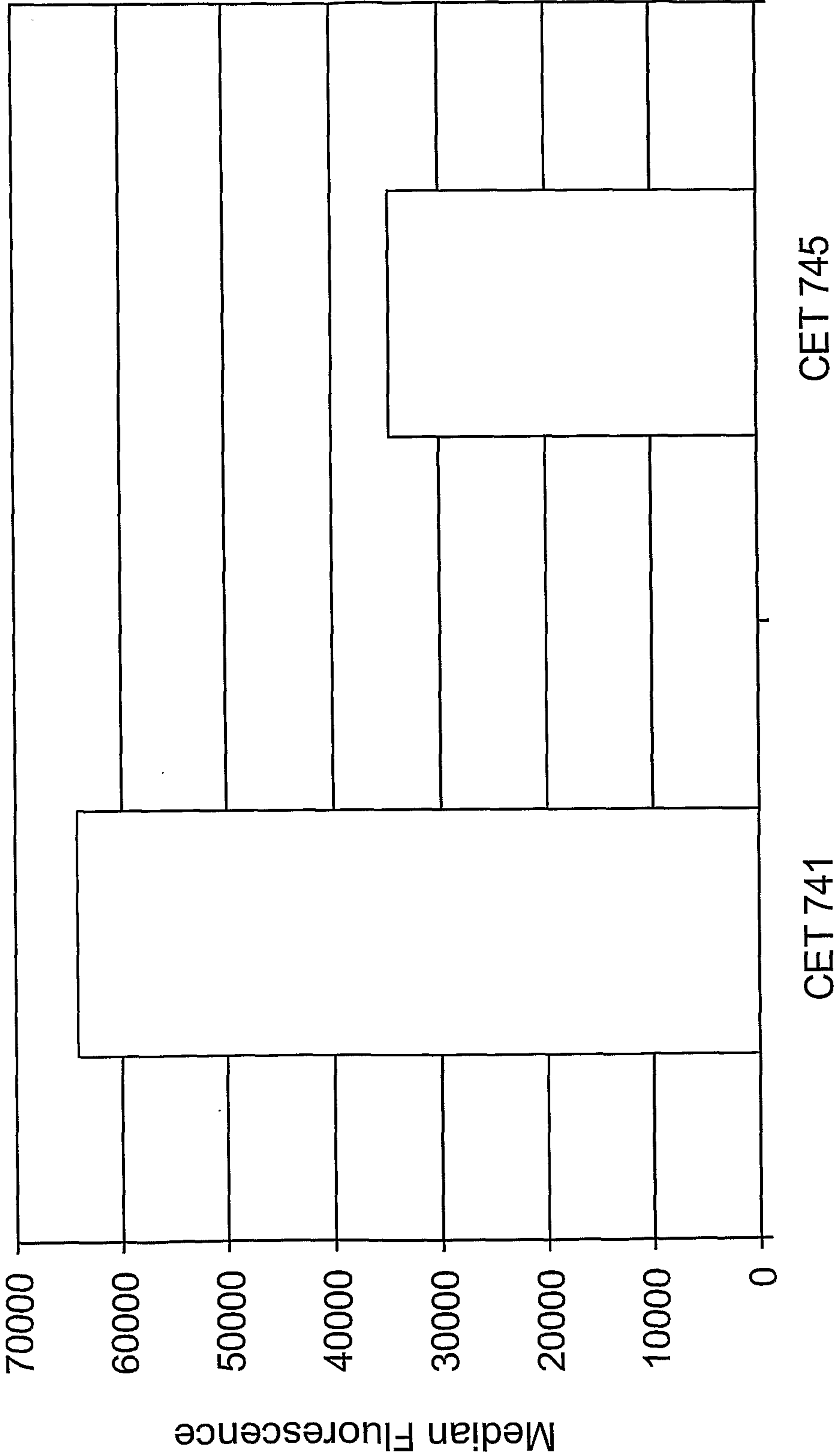


FIG. 6
Expression Levels of EGFP in CHOK-1 cells transfected
with vectors encoding the Puromycin Gene



6/93

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



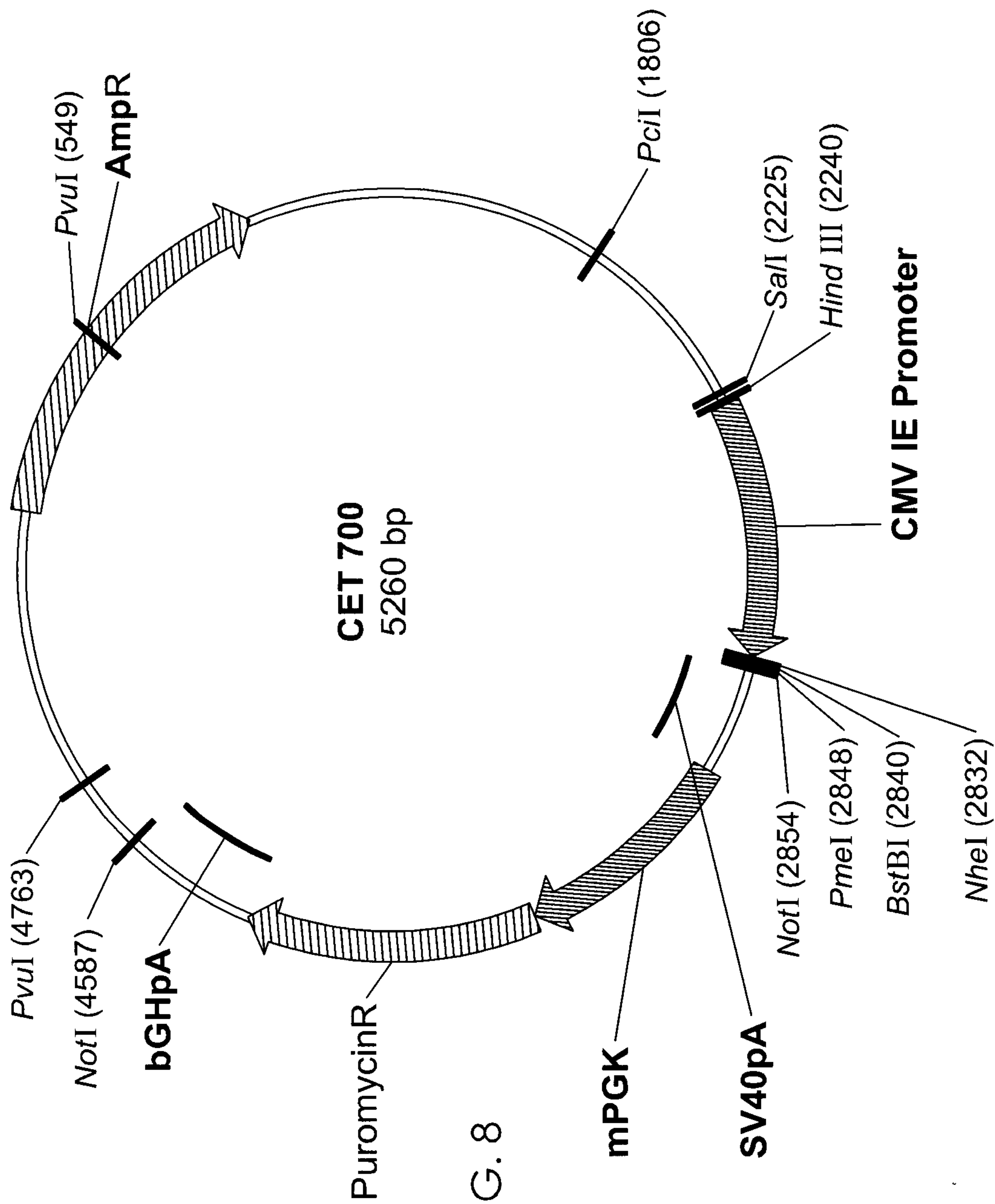


FIG. 8

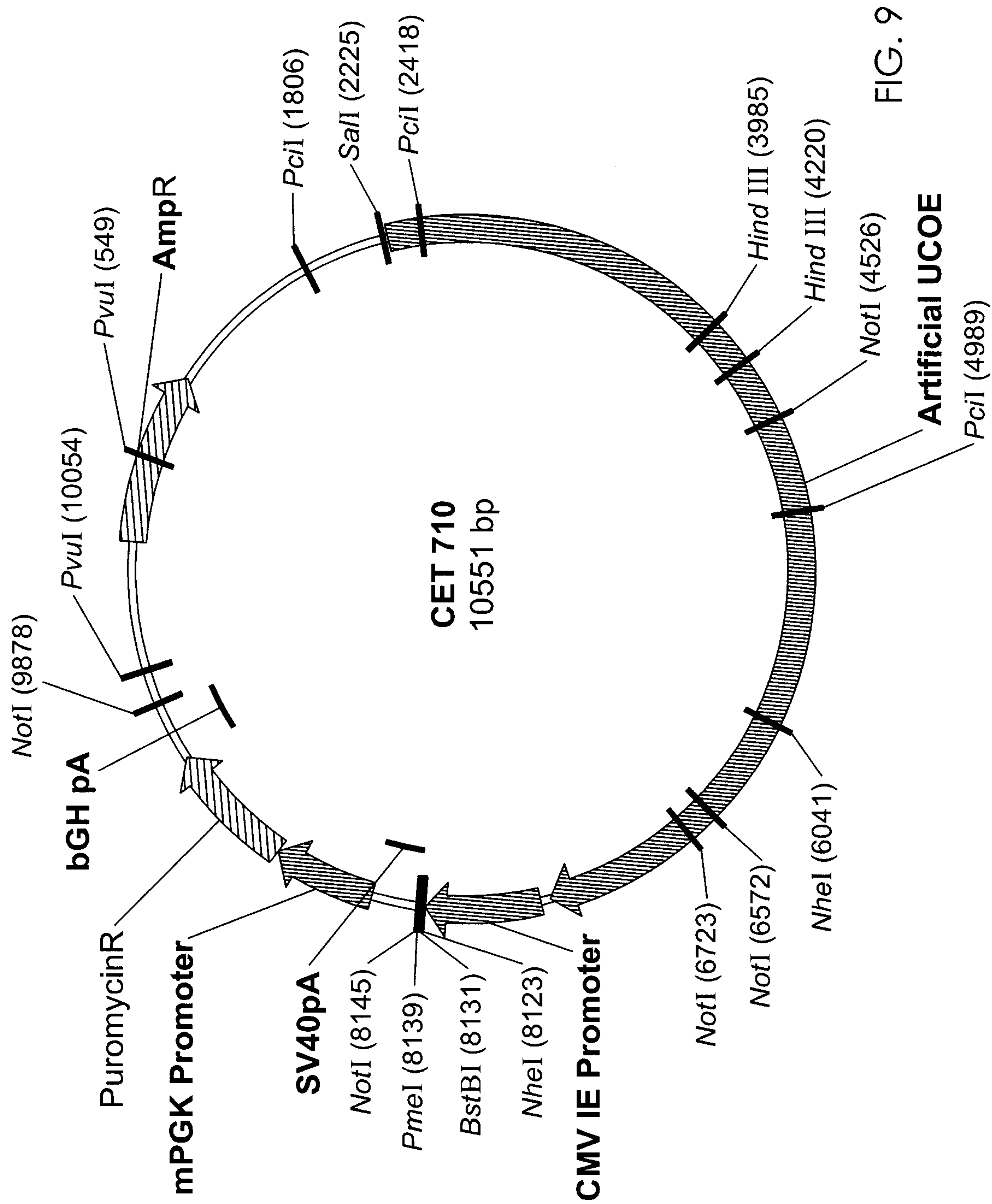


FIG. 9

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	TGTAAAGGCA
151	GTCGCCCTTA	TTCCCTTTTT	TGCGGCATTT	TGCCTTCCTG	TTTTTGCTCA
	CAGCGGGAAT	AAGGGAAAAA	ACGCCGTAAA	ACGGAAGGAC	AAAAACGAGT
201	CCCAGAAACG	CTGGTGAAAG	TAAAAGATGC	TGAAGATCAG	TTGGGTGCAC
	GGGTCTTTGC	GACCACTTTC	ATTTTCTACG	ACTTCTAGTC	AACCCACGTG
251	GAGTGGGTTA	CATCGAACTG	GATCTCAACA	GCGGTAAGAT	CCTTGAGAGT
	CTCACCCAAT	GTAGCTTGAC	CTAGAGTTGT	CGCCATTCTA	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	TGTTGCTAGC
551	GAGGACCGAA	GGAGCTAACC	GCTTTTTTGC	ACAACATGGG	GGATCATGTA
	CTCCTGGCTT	CCTCGATTGG	CGAAAAAACG	TGTTGTACCC	CCTAGTACAT
601	ACTCGCCTTG	ATCGTTGGGA	ACCGGAGCTG	AATGAAGCCA	TACCAAACGA
	TGAGCGGAAC	TAGCAACCCT	TGGCCTCGAC	TTACTTCGGT	ATGGTTTGCT
651	CGAGCGTGAC	ACCACGATGC	CTGTAGCAAT	GGCAACAACG	TTGCGCAAAC
	GCTCGCACTG	TGGTGCTACG	GACATCGTTA	CCGTTGTTGC	AACGCGTTTG
701	TATTAAGTGG	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	CGTAACCAT	GACAGTCTGG
1001	AAGTTTACTC	ATATATACTT	TAGATTGATT	TAAAACTTCA	TTTTTAATTT
	TTCAAATGAG	TATATATGAA	ATCTAACTAA	ATTTTGAAGT	AAAAATTAAA
1051	AAAAGGATCT	AGGTGAAGAT	CCTTTTTGAT	AATCTCATGA	CCAAAATCCC
	TTTTCCTAGA	TCCACTTCTA	GGAAAAACTA	TTAGAGTACT	GGTTTTAGGG
1101	TTAACGTGAG	TTTTCGTTCC	ACTGAGCGTC	AGACCCCGTA	GAAAAGATCA
	AATTGCACTC	AAAAGCAAGG	TGACTCGCAG	TCTGGGGCAT	CTTTTCTAGT
1151	AAGGATCTTC	TTGAGATCCT	TTTTTTCTGC	GCGTAATCTG	CTGCTTGCAA
	TTCCTAGAAG	AACTCTAGGA	AAAAAAGACG	CGCATTAGAC	GACGAACGTT
1201	ACAAAAAAC	CACCGCTACC	AGCGGTGGTT	TGTTTGCCGG	ATCAAGAGCT
	TGTTTTTTTG	GTGGCGATGG	TCGCCACCAA	ACAAACGGCC	TAGTTCTCGA
1251	ACCAACTCTT	TTTCCGAAGG	TAAGTGGCTT	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	CTATTCAGCA	CAGAATGGCC	CAACCTGAGT	TCTGCTATCA
1451	TACCGGATAA	GGCGCAGCGG	TCGGGCTGAA	CGGGGGGTTC	GTGCACACAG
	ATGGCCTATT	CCGCGTCGCC	AGCCCGACTT	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	TGTCCTCTCG	CGTGCTCCCT	CGAAGGTCCC
1651	GGAAACGCCT	GGTATCTTTA	TAGTCCTGTC	GGGTTTCGCC	ACCTCTGACT
	CCTTTGCGGA	CCATAGAAAT	ATCAGGACAG	CCCAAAGCGG	TGGAGACTGA
1701	TGAGCGTCGA	TTTTTGTGAT	GCTCGTCAGG	GGGGCGGAGC	CTATGGAAAA
	ACTCGCAGCT	AAAAACACTA	CGAGCAGTCC	CCCCGCCTCG	GATACCTTTT
1751	ACGCCAGCAA	CGCGGCCTTT	TTACGGTTCC	TGGCCTTTTG	CTGGCCTTTT
	TGCGGTTCGTT	GCGCCGGA	AATGCCAAGG	ACCGGAAAAC	GACCGGAAAA
1801	GCTCACATGT	TCTTTCCTGC	GTTATCCCCT	GATTCTGTGG	ATAACCGTAT
	CGAGTGTACA	AGAAAGGACG	CAATAGGGGA	CTAAGACACC	TATTGGCATA
1851	TACCGCCTTT	GAGTGAGCTG	ATACCGCTCG	CCGCAGCCGA	ACGACCGAGC
	ATGGCGGAAA	CTCACTCGAC	TATGGCGAGC	GGCGTCGGCT	TGCTGGCTCG
1901	GCAGCGAGTC	AGTGAGCGAG	GAAGCGGAAG	AGCGCCCAAT	ACGCAAACCG
	CGTCGCTCAG	TCACTCGCTC	CTTCGCCTTC	TCGCGGGTTA	TGCGTTTGGC
1951	CCTCTCCCCG	CGCGTTGGCC	GATTCATTAA	TGCAGCTGGC	ACGACAGGTT
	GGAGAGGGGC	GCGCAACCGG	CTAAGTAATT	ACGTCGACCG	TGCTGTCCAA
2001	TCCCGACTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC
	AGGGCTGACC	TTTCGCCCCG	CACTCGCGTT	GCGTTAATTA	CACTCAATCG
2051	TCACTCATTA	GGCACCCAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG
	AGTGAGTAAT	CCGTGGGGTC	CGAAATGTGA	AATACGAAGG	CCGAGCATAC
2101	TTGTGTGGAA	TTGTGAGCGG	ATAACAATTT	CACACAGGAA	ACAGCTATGA
	AACACACCTT	AACACTCGCC	TATTGTATAA	GTGTGTCCTT	TGTCGATACT
2151	CCATGATTAC	GCCAAGCGCG	CAATTAACCC	TACTAAAGG	GAACAAAAGC
	GGTACTAATG	CGGTTTCGCG	GTTAATTGGG	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	AGGTCAGGCA
	ACCTCACAAAC	GAAGGGTAAT	ATAGGGACAA	TAGTTCGTGT	TCCAGTCCGT
2401	CAGAGTAAGA	CTCAAAACAT	GTTTTGGAAT	GTATGACTGG	TATGAACTAC
	GTCTCATTCT	GAGTTTTGTA	CAAAACCTTA	CATACTGACC	ATACTTGATG
2451	AAACCAGTAA	GCTGATGTTT	TCATTTTGAG	TCTATAAATC	TAATTTTGTG
	TTTGGTCATT	CGACTACAAA	AGTAAACTC	AGATATTTAG	ATTAAAACAC
2501	GTGGTTTTGT	GTATGGCTCA	AGGCTCAAAT	TGTAAAATTT	AATATTATGT
	CACCAAAACA	CATACCGAGT	TCCGAGTTTA	ACATTTTAAA	TTATAATACA
2551	GACCAAAGAA	AGTTATACCC	AGAACCTCAA	TTTCCTCACC	TTCAAAATGG
	CTGGTTTCTT	TCAATATGGG	TCTTGAGATT	AAAGGAGTGG	AAGTTTTACC
2601	GGCAGTTTCT	CACTCATTGG	TCTGCTGTCA	CGATTTTAAT	GAGCTCATGC
	CCGTCAAAGA	GTGAGTAACC	AGACGACAGT	GCTAAAATTA	CTCGAGTACG
2651	ACAAACAGCC	CTTTATATAA	GGTAAGTGCT	GGATAAATGT	TGGCTACTAT
	TGTTTGTCGG	GAAATATATT	CCATTCACGA	CCTATTTACA	ACCGATGATA

11/93

FIG. 10 (contd.)

2701	AATAAAATAA	GCCTCTAAGA	TACTTGGTCA	GCACAAGTAC	TACCCAAGAG
	TTATTTTATT	CGGAGATTCT	ATGAACCAGT	CGTGTTTCATG	ATGGGTTCTC
2751	TATGCACTGT	AAGTAAACTG	ACAAAATTGT	GTATCTAAAA	CTGGCCAGAT
	ATACGTGACA	TTCATTTGAC	TGTTTTAACA	CATAGATTTT	GACCGGTCTA
2801	GAAAGAGAAA	CTTTTAAGGG	GCCCTTCTGC	GTGCCCCGACA	CTGTGCTAGG
	CTTTCTCTTT	GAAAATTCCC	CGGGAAGACG	CACGGGCTGT	GACACGATCC
2851	CACTCACACT	ATCCCGACCC	GAGAAACCGA	TCTGCGACCC	AGAGGAACTT
	GTGAGTGTGA	TAGGGCTGGG	CTCTTTGGCT	AGACGCTGGG	TCTCCTTGAA
2901	ACCAAGCCTC	CAGCATCTTG	TGCAGCCCTA	CTCATGGGAC	CATCTGGATA
	TGGTTCGGAG	GTCGTAGAAC	ACGTCGGGAT	GAGTACCCTG	GTAGACCTAT
2951	CCCACCCTTG	TCTTTACAGG	GAGCAGAACA	CACCTCTTAT	GTGTCAGAAA
	GGGTGGGAAC	AGAAATGTCC	CTCGTCTTGT	GTGGAGAATA	CACAGTCTTT
3001	ACAAAGTCCA	GGAAGTATAT	TTTTACCTGA	GGCAATATCT	GAAAATTGTA
	TGTTTTCAGGT	CCTTCATATA	AAAATGGACT	CCGTTATAGA	CTTTTAACAT
3051	TGCTACAGCC	TCCAAAGTGA	GTCTTCCTCT	CAGTACCTCT	CTTCTAGGCA
	ACGATGTCGG	AGGTTTCACT	CAGAAGGAGA	GTCATGGAGA	GAAGATCCGT
3101	CATGGAGCCC	TTTCTTCCAA	GTATTATGTT	TAACCACTTA	ATGAATGAAG
	GTACCTCGGG	AAAGAAGGTT	CATAATACAA	ATTGGTGAAT	TACTTACTTC
3151	TCCTGAAACT	GCTTACCCAT	GCTCCCTATA	ATCTCTGAGT	AATCTTCCTT
	AGGACTTTGA	CGAATGGGTA	CGAGGGATAT	TAGAGACTCA	TTAGAAGGAA
3201	TTCCACAACC	TCAGGCATAA	TCTCATCTTC	TGTTTCTATT	ACAATTTCAA
	AAGGTGTTGG	AGTCCGTATT	AGAGTAGAAG	ACAAAGATAA	TGTTAAAGTT
3251	ATTCTGGAAA	AAGGAAGTTG	TGGTCTGGAA	TTATATGGTC	CAGATGATCT
	TAAGACCTTT	TTCCTTCAAC	ACCAGACCTT	AATATACCAG	GTCTACTAGA
3301	GAAACAAAAA	GGACAGCACT	ATTAGTAATC	ATTTAGTTTT	GAAGACAGTC
	CTTTGTTTTT	CCTGTCGTGA	TAATCATTAG	TAAATCAAAA	CTTCTGTCAG
3351	TAATAATTTG	CTGTCTCTAA	AGTACTATAT	TCCCTATAGT	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
	TTTTTTAAAT	TTCATGATAT	GGAAAGGATT	AAAATGGAAA	TTTCAACAGG
3551	TGGAAATATC	TGGGTTGACA	AAGGCGATGA	AACTGAACTG	AGACTTAAAA
	ACCTTTATAG	ACCCAACGTG	TTCCGCTACT	TTGACTTGAC	TCTGAATTTT
3601	AAAAGATTAC	CCACCTGGTT	GTGCACAAGC	CTGCTTATGT	CCCAATCTCC
	TTTTCTAATG	GGTGGACCAA	CACGTGTTCTG	GACGAATACA	GGGTTAGAGG
3651	AGTCTAGGGT	CTGATGCTCC	TTGCTGCAGT	AATATGCTTT	GTGGCATCTG
	TCAGATCCCA	GACTACGAGG	AACGACGTCA	TTATACGAAA	CACCGTAGAC
3701	GAGCACGTTT	TGGGGCCTAA	ACAGCCACAA	ACCCTGCAGA	GATGAGCACC
	CTCGTGCAAA	ACCCCGGATT	TGTCGGGTGTT	TGGGACGTCT	CTACTCGTGG
3751	AGACTTAAGC	TGGAGACACA	CTGATTCTCC	TGTTTCTGGG	GGAGGATTCT
	TCTGAATTCTG	ACCTCTGTGT	GACTAAGAGG	ACAAAGACCC	CCTCCTAAGA
3801	CAGAAGGTGG	CTCATATGAG	TAAAAATCGT	TTTTCCTGGG	TAGTTGATTCT
	GTCTTCCACC	GAGTATACTC	ATTTTTAGCA	AAAAGGACCC	ATCAACTAAG
3851	CTAAAAACTA	AAAAAGAATA	CAGAGAAAAG	TTTTATCTTC	AAACAAAACA
	GATTTTTTGAT	TTTTTCTTAT	GTCTCTTTTC	AAAATAGAAG	TTTGTTTTGT
3901	GCAATTCAAC	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

12/93

FIG. 10 (contd.)

4051	TTGAGTTAAT	CAAATTAATA	AGACTGATCG	TTAAGAACGA	CTGCCAAAAA
	AACTCAATTA	GTTTAATTAT	TCTGACTAGC	AATTCTTGCT	GACGGTTTTT
4101	TACGAAAAAG	CTACTGGGAT	CCATCTTTCC	AAGACAATTT	CTATTATCTG
	ATGCTTTTTT	GATGACCCTA	GGTAGAAAGG	TTCTGTTAAA	GATAATAGAC
4151	AATTAACACC	ATACCTGGTA	CCCCTGATT	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	GCGCGCTCTC	CCTCGAGACT
4301	GGCTGGGGCG	GCAGCGGTGC	TTGCCGCCGT	CCCCCTGGTC	GCTCCCGGAA
	CCGACCCCGC	CGTCGCCACG	AACGGCGGCA	GGGGGACCAG	CGAGGGCCTT
4351	TTAACGCCGC	GCACGCGTCG	GAGGCATGGC	CCCGTCCCGA	CCCCGTTTGG
	AATTGCGGCG	CGTGCGCAGC	CTCCGTACCG	GGGCAGGGCT	GGGGCAAACC
4401	CGGCTCACCT	CGCAGGCCGG	CACAGCACGG	CTGCTCGCGG	CAGCAGAAGA
	GCCGAGTGGA	GCGTCCGGCC	GTGTCGTGCC	GACGAGCGCC	GTCGTCTTCT
4451	GGAAGATGCA	GCGGTGGAAG	GCGTCCGGGC	GGCCAGGCAG	CGGCGCATAC
	CCTTCTACGT	CGCCACCTTC	CGCAGGCCCG	CCGGTCCGTC	GCCGCGTATG
4501	ACCTGCAGCA	GGAAGGAGAG	CGGGCGGGCG	CACAGCTCGC	AGGCCAGGGC
	TGGACGTCGT	CCTTCCTCTC	GCCCGCCGGC	GTGTCGAGCG	TCCGGTCCCG
4551	CTGGGGCCCC	GGCAGCCCGG	CCGCGCCCAG	CCATGCCGGC	CGCCCGCCCA
	GACCCCGGGG	CCGTGCGGGC	GGCGCGGGTC	GGTACGGCCG	GCGGGCGGGT
4601	CCTTGCTGGG	GAACTGCTCG	CTGCGCAGTC	GCCACGCCGG	CGCCGACTCG
	GGAACGACCC	CTTGACGAGC	GACGCGTCAG	CGGTGCGGCC	GCGGCTGAGC
4651	GCGAAGCCCA	GCTCCACAGG	CCTGGCCCCG	GCGGCAGCCA	TGCGGGGCGC
	CGCTTCGGGT	CGAGGTGTCC	GGACCGGGGC	CGCCGTCCGT	ACGCCCCGCG
4701	GGGCTGGCGT	GGGGCGCAGC	CCACAGCTGG	GTCGGAAGGC	GGAAATCGGG
	CCCGACCGCA	CCCCGCGTCG	GGTGTGACCC	CAGCCTTCCG	CCTTTAGCCC
4751	CGCCGGGGCC	GAAGGCAAGA	GGCGGGCACC	TTTCCGGAGG	ACAGGAGGCG
	GCGGCCCCGG	CTTCCGTTCT	CCGCCCGTGG	AAAGGCCTCC	TGTCCTCCGC
4801	GAAACGCGTC	TGACGGGAGC	GGTTGCAGGA	CCAATGCGAG	GGAACGGGGC
	CTTTGCGCAG	ACTGCCCTCG	CCAACGTCTT	GGTTACGCTC	CCTTGCCCCG
4851	AGAGGAAACC	TCTCGGCATC	AGCCCCGCCC	CTGGCGCCTC	TGCCTCCGAG
	TCTCCTTTGG	AGAGCCGTAG	TCGGGGCGGG	GACCGCGGAG	ACGGAGGCTC
4901	CCGCTTTTCT	GGTGCCCTCC	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	TTCCAACCTG	TCCATTTTAC
5151	AGGACCCTAC	AGCGGAAACC	AAGAATCCTG	TGTGCCTGAG	AGTAATGAAG
	TCCTGGGATG	TCGCCTTTGG	TTCTTAGGAC	ACACGGACTC	TCATTACTTC
5201	AAGCCTCTGC	AGAAGAGTCT	TTTCTGTCAG	TCTTAAGGTC	TCTGTTTTAA
	TTCGGAGACG	TCTTCTCAGA	AAAGACAGTC	AGAATTCCAG	AGACAAAATT
5251	TGTTAGTGCT	GGCTTGCTGT	ACCTGAATTC	CAAGGGAGGA	GTGTATAATG
	ACAATCACGA	CCGAACGACA	TGGACTTAAG	GTTCCCTCCT	CACATATTAC
5301	AGGCATGGCC	AACCCCCACT	TCCCATCATT	GCCTGAACTA	GTTTTTCAGG
	TCCGTACCGG	TTGGGGGTGA	AGGGTAGTAA	CGGACTTGAT	CAAAAAGTCC
5351	TTAACTTCAG	AATGCCCTTG	GTACCGCGGG	CCCCCTCTGT	GGTCCCACGC
	AATTGAAGTC	TTACGGGAAC	CATGGCGCCC	GGGGGAGACA	CCAGGGTGCG

13/93

FIG. 10 (contd.)

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

14/93

FIG. 10 (contd.)

6751	GCCGTCTGGG	CCGCAGCGGG	GGGCGCATGG	GGCGGAACCG	GACCGCCGTG
	CGGCAGACCC	GGCGTCGCCC	CCCGCGTACC	CCGCCTTGGC	CTGGCGGCAC
6801	GGGGGCGCGG	GAGAAGCCCC	TGGGCCTCCG	GAGATGGGGG	ACACCCACAG
	CCCCCGCGCC	CTCTTCGGGG	ACCCGGAGGC	CTCTACCCCC	TGTGGGGTGC
6851	CCAGTTCGCA	GGCGCGAGGC	CGCGCTCGGG	CGGGCGCGCT	CCGGGGGTGC
	GGTCAAGCGT	CCGCGCTCCG	GCGCGAGCCC	GCCCGCGCGA	GGCCCCACAG
6901	CGCTCTCGGG	GCGGGGGCAA	CCGGCGGGGT	CTTTGTCTGA	GCCGGGCTCT
	GCGAGAGCCC	CGCCCCCGTT	GGCCGCCCCA	GAAACAGACT	CGGCCCGAGA
6951	TGCCAATGGG	GATCGCACGG	TGGGCGCGGC	GTAGCCCCCG	TCAGGCCCGG
	ACGGTTACCC	CTAGCGTGCC	ACCCGCGCCG	CATCGGGGGC	AGTCCGGGCC
7001	TGGGGGCTGG	GGCGCCATGC	GCGTGCGCGC	TGGTCCTTTG	GGCGCTAACT
	ACCCCCGACC	CCGCGGTACG	CGCACGCGCG	ACCAGGAAAC	CCGCGATTGA
7051	GCGTGCGCGC	TGGGAATTGG	CGCTAATTGC	GCGTGCGCGC	TGGGACTCAA
	CGCACGCGCG	ACCCTTAACC	GCGATTAAAC	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	CCCGGTGGG
	TGGGCCAGCG	ACAAACTTGG	CCCGCCTCCG	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	GCGCCCGGCT
	CGGCCTTCAC	CGGTCCCGCC	CCCGCTGAAG	CCGAGTGTCG	CGCGGGCCGA
7451	ATTCTCGCAG	CTCACCATGC	CGGTCGCCAC	CATGAGCTTG	ATATCGAATT
	TAAGAGCGTC	GAGTGGTACG	GCCAGCGGTG	GTAAGCTGAC	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	GCGGTTTTGG	CAGTACATCA	ATGGGCGTGG	ATAGCGGTTT	GAATCACGGG
	CGCCAAAACC	GTCATGTAGT	TACCCGCACC	TATCGCCAAA	CTGAGTGCCC
7951	GATTTCCAAG	TCTCCACCCC	ATTGACGTCA	ATGGGAGTTT	GTTTTGGCAC
	CTAAAGGTTC	AGAGGTGGGG	TAACTGCAGT	TACCCTCAAA	CAAAACCGTG
8001	CAAAATCAAC	GGGACTTTCC	AAAATGTCGT	AACAACCTCC	CCCCATTGAC
	GTTTTAGTTG	CCCTGAAAGG	TTTTACAGCA	TTGTTGAGGC	GGGGTAACTG
8051	GCAAATGGGC	GGTAGGCGTG	TACGGTGGGA	GGTCTATATA	AGCAGAGCTG
	CGTTTACCCG	CCATCCGCAC	ATGCCACCCT	CCAGATATAT	TCGTCTCGAC

15/93

FIG. 10 (contd.)

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

16/93

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

FIG. 10 (contd.)

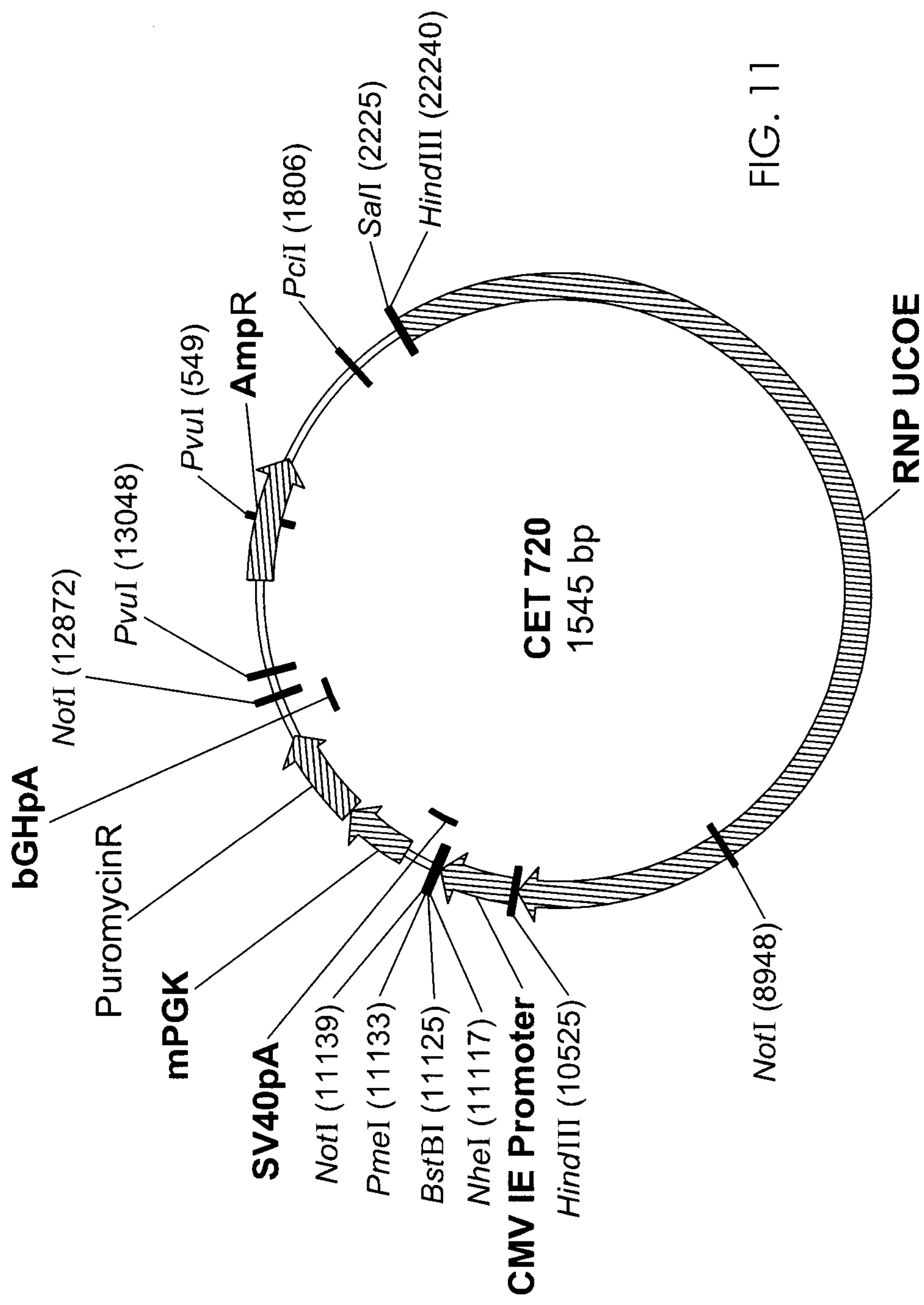


FIG. 11

18/93

FIG. 12 CET720 nucleotide sequence

1	GGTGGCACTT	TTCGGGGGAAA	TGTGCGCGGA	ACCCCTATTT	GTTTATTTTT
	CCACCGTGAA	AAGCCCCTTT	ACACGCGCCT	TGGGGATAAA	CAAATAAAAA
51	CTAAATACAT	TCAAATATGT	ATCCGCTCAT	GAGACAATAA	CCCTGATAAA
	GATTTATGTA	AGTTTATACA	TAGGCGAGTA	CTCTGTTATT	GGGACTATTT
101	TGCTTCAATA	ATATTGAAAA	AGGAAGAGTA	TGAGTATTCA	ACATTTCCGT
	ACGAAGTTAT	TATAACTTTT	TCCTTCTCAT	ACTCATAAGT	TGTAAAGGCA
151	GTCGCCCTTA	TTCCCTTTTT	TGCGGCATTT	TGCCTTCCTG	TTTTTGCTCA
	CAGCGGGAAT	AAGGGAAAAA	ACGCCGTAAA	ACGGAAGGAC	AAAAACGAGT
201	CCCAGAAACG	CTGGTGAAAG	TAAAAGATGC	TGAAGATCAG	TTGGGTGCAC
	GGGTCTTTGC	GACCACTTTC	ATTTTCTACG	ACTTCTAGTC	AACCCACGTG
251	GAGTGGGTTA	CATCGAACTG	GATCTCAACA	GCGGTAAGAT	CCTTGAGAGT
	CTCACCCAAT	GTAGCTTGAC	CTAGAGTTGT	CGCCATTCTA	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	TACTGAACC	AACTCATGAG	TGGTCAGTGT
451	GAAAAGCATC	TTACGGATGG	CATGACAGTA	AGAGAATTAT	GCAGTGCTGC
	CTTTTCGTAG	AATGCCTACC	GTACTGTCAT	TCTCTTAATA	CGTCACGACG
501	CATAACCATG	AGTGATAACA	CTGCGGCCAA	CTTACTTCTG	ACAACGATCG
	GTATTGGTAC	TCACTATTGT	GACGCCGGTT	GAATGAAGAC	TGTTGCTAGC
551	GAGGACCGAA	GGAGCTAACC	GCTTTTTTGC	ACAACATGGG	GGATCATGTA
	CTCCTGGCTT	CCTCGATTGG	CGAAAAAACG	TGTTGTACCC	CCTAGTACAT
601	ACTCGCCTTG	ATCGTTGGGA	ACCGGAGCTG	AATGAAGCCA	TACCAAACGA
	TGAGCGGAAC	TAGCAACCCT	TGGCCTCGAC	TTACTTCGGT	ATGGTTTGCT
651	CGAGCGTGAC	ACCACGATGC	CTGTAGCAAT	GGCAACAACG	TTGCGCAAAC
	GCTCGCACTG	TGGTGCTACG	GACATCGTTA	CCGTTGTTGC	AACGCGTTTG
701	TATTAAGTGG	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	CATTTCGGGAG	GGCATAGCAT
901	GTTATCTACA	CGACGGGGAG	TCAGGCAACT	ATGGATGAAC	GAAATAGACA
	CAATAGATGT	GCTGCCCTTC	AGTCCGTTGA	TACCTACTTG	CTTTATCTGT
951	GATCGCTGAG	ATAGGTGCCT	CACTGATTAA	GCATTGGTAA	CTGTCAGACC
	CTAGCGACTC	TATCCACGGA	GTGACTAATT	CGTAACCATT	GACAGTCTGG
1001	AAGTTTACTC	ATATATACTT	TAGATTGATT	TAAAACCTCA	TTTTTAATTT
	TTCAAATGAG	TATATATGAA	ATCTAACTAA	ATTTTGAAGT	AAAAATTAAT
1051	AAAAGGATCT	AGGTGAAGAT	CCTTTTTTGAT	AATCTCATGA	CCAAAATCCC
	TTTTTCCTAGA	TCCACTTCTA	GGAAAAACTA	TTAGAGTACT	GGTTTTAGGG
1101	TTAACGTGAG	TTTTTCGTTCC	ACTGAGCGTC	AGACCCCGTA	GAAAAGATCA
	AATTGCACTC	AAAAGCAAGG	TGACTCGCAG	TCTGGGGCAT	CTTTTCTAGT
1151	AAGGATCTTC	TTGAGATCCT	TTTTTTCTGC	GCGTAATCTG	CTGCTTGCAA
	TTCTTAGAAG	AACTCTAGGA	AAAAAAGACG	CGCATTAGAC	GACGAACGTT
1201	ACAAAAAAAC	CACCGCTACC	AGCGGTGGTT	TGTTTGCCGG	ATCAAGAGCT
	TGTTTTTTTG	GTGGCGATGG	TCGCCACCAA	ACAAACGGCC	TAGTTCTCGA
1251	ACCAACTCTT	TTTCCGAAGG	TAACTGGCTT	CAGCAGAGCG	CAGATACCAA
	TGGTTGAGAA	AAAGGCTTCC	ATTGACCGAA	GTCGTCTCGC	GTCTATGGTT
1301	ATACTGTCCT	TCTAGTGTAG	CCGTAGTTAG	GCCACCACTT	CAAGAACTCT
	TATGACAGGA	AGATCACATC	GGCATCAATC	CGGTGGTGAA	GTTCTTGAGA

FIG. 12(contd.) 19/93

1351	GTAGCACCGC	CTACATACCT	CGCTCTGCTA	ATCCTGTTAC	CAGTGGCTGC
	CATCGTGGCG	GATGTATGGA	GCGAGACGAT	TAGGACAATG	GTCACCGACG
1401	TGCCAGTGGC	GATAAGTCGT	GTCTTACCGG	GTTGGACTCA	AGACGATAGT
	ACGGTCACCG	CTATTCAGCA	CAGAATGGCC	CAACCTGAGT	TCTGCTATCA
1451	TACCGGATAA	GGCGCAGCGG	TCGGGGCTGAA	CGGGGGGTTT	GTGCACACAG
	ATGGCCTATT	CCGCGTCGCC	AGCCCGACTT	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
	GCCATTGCGC	GTCCCAGCCT	TGTCCCTCTCG	CGTGCTCCCT	CGAAGGTCCC
1651	GGAAACGCCT	GGTATCTTTA	TAGTCCTGTC	GGGTTTCGCC	ACCTCTGACT
	CCTTTGCGGA	CCATAGAAAT	ATCAGGACAG	CCCAAAGCGG	TGGAGACTGA
1701	TGAGCGTCGA	TTTTTGTGAT	GCTCGTCAGG	GGGGCGGAGC	CTATGGAAAA
	ACTCGCAGCT	AAAAACACTA	CGAGCAGTCC	CCCCGCCTCG	GATACCTTTT
1751	ACGCCAGCAA	CGCGGCCTTT	TTACGGTTCC	TGGCCTTTTG	CTGGCCTTTT
	TGCGGTGCGT	GCGCCGGAAA	AATGCCAAGG	ACCGGAAAAC	GACCGGAAAA
1801	GCTCACATGT	TCTTTCCTGC	GTTATCCCCT	GATTCTGTGG	ATAACCGTAT
	CGAGTGTACA	AGAAAGGACG	CAATAGGGGA	CTAAGACACC	TATTGGCATA
1851	TACCGCCTTT	GAGTGAGCTG	ATACCGCTCG	CCGCAGCCGA	ACGACCGAGC
	ATGGCGGAAA	CTCACTCGAC	TATGGCGAGC	GGCGTCGGCT	TGCTGGCTCG
1901	GCAGCGAGTC	AGTGAGCGAG	GAAGCGGAAG	AGCGCCCAAT	ACGCAAACCG
	CGTCGCTCAG	TCACTCGCTC	CTTCGCCTTC	TCGCGGGTTA	TGCGTTTGGC
1951	CCTCTCCCCG	CGCGTTGGCC	GATTCATTAA	TGCAGCTGGC	ACGACAGGTT
	GGAGAGGGGC	GCGCAACCGG	CTAAGTAATT	ACGTCGACCG	TGCTGTCCAA
2001	TCCCGACTGG	AAAGCGGGCA	GTGAGCGCAA	CGCAATTAAT	GTGAGTTAGC
	AGGGCTGACC	TTTCGCCCCG	CACTCGCGTT	GCGTTAATTA	CACTCAATCG
2051	TCACTCATTA	GGCACCCCAG	GCTTTACACT	TTATGCTTCC	GGCTCGTATG
	AGTGAGTAAT	CCGTGGGGTC	CGAAATGTGA	AATACGAAGG	CCGAGCATAC
2101	TTGTGTGGAA	TTGTGAGCGG	ATAACAATTT	CACACAGGAA	ACAGCTATGA
	AACACACCTT	AACACTCGCC	TATTGTTAAA	GTGTGTCCTT	TGTCGATACT
2151	CCATGATTAC	GCCAAGCGCG	CAATTAACCC	TCACTAAAGG	GAACAAAAGC
	GGTACTAATG	CGGTTGCGCG	GTTAATTGGG	AGTGATTTCC	CTTGTTTTTCG
2201	TGGGTACCGG	GCCCCCCCCTC	GAGGTCGACG	GTATCGATAA	GCTTCAATGT
	ACCCATGGCC	CGGGGGGGGAG	CTCCAGCTGC	CATAGCTATT	CGAAGTTACA
2251	TTTTAGCACC	CTCTGTGTGG	AGGAAAATAA	TGCAGATTAT	TCTAATTAGT
	AAAATCGTGG	GAGACACACC	TCCTTTTATT	ACGTCTAATA	AGATTAATCA
2301	GTAATATCTA	ACCACATTAA	AATATATTAC	ATAGTAAACT	ACACTCCATA
	CATTATAGAT	TGGTGTAATT	TTATATAATG	TATCATTTGA	TGTGAGGTAT
2351	ATTTTATAAA	TTTGACTCCC	CAGGGTAATA	AACTAGTCTC	TAGTCTGCTC
	TAAAATATTT	AAACTGAGGG	GTCCCATTTAT	TTGATCAGAG	ATCAGACGAG
2401	ACCTTCAACT	GTACAATAAA	GTCTTG GTTC	TTTTGAAATA	GACCTCAAAT
	TGGAAGTTGA	CATGTTATTT	CAGAACCAAG	AAAACCTTTAT	CTGGAGTTTA
2451	GAGACACCTA	AAATTCAAAG	TGTCTTTACA	TTTAAAGACA	CCTACAGGAA
	CTCTGTGGAT	TTTAAGTTTC	ACAGAAATGT	AAATTTCTGT	GGATGTCCTT
2501	AGCAGGTAAA	AGAGCCAGGT	TAAAAACAAA	TTCTAAAACC	ACTTAGCTGC
	TCGTCCATTT	TCTCGGTCCA	ATTTTTGTTT	AAGATTTTGG	TGAATCGACG
2551	AGTTAAACAT	ATAGTAAAGA	TGCACTAAAG	TTTCTTACTC	TGTAAATCCC
	TCAATTTGTA	TATCATTTCT	ACGTGATTTT	AAAGAATGAG	ACATTTAGGG
2601	TTCCACTTCA	GGAAATATTC	CACTTTCCCA	TTCCTACAC	GTGATCTAG
	AAGGTGAAGT	CCTTTATAAG	GTGAAAGGGT	AAGTGATGTG	CAGCTAGATC
2651	TACTTTTTTCC	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	ACTTAAACT	CGTGTAAG	TTCCGAGTTA

FIG. 12(contd.) 20/93

2801	TGCTTATATT	ATTATCACAT	TCGACATAAA	TTTTTACTTC	TATATCCCAG
	ACGAATATAA	TAATAGTGTA	AGCTGTATTT	AAAAATGAAG	ATATAGGGTC
2851	GGCAGACACC	TTCTGGAAAG	ATTAAAAGTC	AACAGACAAT	AAAATAAAAG
	CCGTCTGTGG	AAGACCTTTC	TAATTTTTCAG	TTGTCTGTTA	TTTTATTTTC
2901	AATGCTTTAT	CTTGTTTCATT	TAGTTCAAAC	TTACAACCCA	CCACCAAAT
	TTACGAAATA	GAACAAGTAA	ATCAAGTTTG	AATGTTGGGT	GGTGGTTTTA
2951	AATACAATAA	AAAAACACTA	TCTGGAAACA	GTTATTTTTT	TCCAGTCTTT
	TTATGTTATT	TTTTTGTGAT	AGACCTTTGT	CAATAAAAAA	AGGTCAGAAA
3001	TTTTTTGAGA	CAGGGTCTCA	CACTCTTGTC	GCCCAGGCTG	GAGTGCAGTG
	AAAAAACTCT	GTCCCAGAGT	GTGAGAACAG	CGGGTCCGAC	CTCACGTCAC
3051	GCGTGATCTC	AGCTCACTGC	AACCTCCGCC	TCCCCAGGTT	CAAGCAGTTC
	CGCACTAGAG	TCGAGTGACG	TTGGAGGCGG	AGGGGTCCAA	GTTCGTCAAG
3101	TCATGCCTCA	GCCTCCAGAG	TAGCTGGGAT	TATAGGCGGA	TGCCACCATG
	AGTACGGAGT	CGGAGGTCTC	ATCGACCCTA	ATATCCGCCT	ACGGTGGTAC
3151	CCGGGCTAAT	TTTTTTTGTG	TTTTTATTAG	AAACAGGGTT	TCACCATGTT
	GGCCCGATTA	AAAAAAACAC	AAAAATAATC	TTTGTCCCAA	AGTGGTACAA
3201	GACCAGGCTG	GTCTCAAAC	CCTGACCTGA	AGTGATTAC	CAGCCTGGGC
	CTGGTCCGAC	CAGAGTTTGA	GGACTGGACT	TCACTAAGTG	GTGCGACCCG
3251	CTCCCAAAGT	GCTGGCATT	CAGGCGTGAG	CCACTGCGCC	CGGCCCTGTA
	GAGGGTTTCA	CGACCGTAAT	GTCCGCACTC	GGTGACGCGG	GCCGGGACAT
3301	GTCTTAAAAG	ACCAAGTTTA	CTAATTTTCA	CTCATTTTAA	CAACACTGCA
	CAGAATTTTC	TGGTTCAAAT	GATTAAAAGT	GAGTAAAATT	GTTGTGACGT
3351	ACAAACAAC	ATGCAGGAAG	TACCTAAAGG	GTGATCCAGA	GAAGCAAGTA
	TGTTTGTTGA	TACGTCCTTC	ATGGATTTCC	CACTAGGTCT	CTTCGTTTCAT
3401	GTAGTGACAG	GTCTTAGGTG	AACCTATGAC	AGACCTTGTA	TCCACCCCCA
	CATCACTGTC	CAGAATCCAC	TTGGATACTG	TCTGGAACAT	AGGTGGGGGT
3451	GATGGTAAAA	GCCCCAGCCC	CCTTCTCAAT	TCAAATATTA	ATGTCAAAAG
	CTACCATTTT	CGGGGTCGGG	GGAAGAGTTA	AGTTTATAAT	TACAGTTTTC
3501	CATCAATGAT	ACAGAGAAAA	GATAAATGCA	GAATGAAAAC	ATGGTTCAAA
	GTAGTTACTA	TGTCTCTTTT	CTATTTACGT	CTTACTTTTG	TACCAAGTTT
3551	ATCCTGATAC	CAACTGCAGG	GTCAACTATA	GAGACCACTA	GGAGGTTCAA
	TAGGACTATG	GTTGACGTCC	CAGTTGATAT	CTCTGGTGAT	CCTCCAAGTT
3601	TTAAAGGACA	AGATTATTTT	TCCATAATCT	CTGTAGATAA	TATTTCCCTAC
	AATTTCCCTGT	TCTAATAAAA	AGGTATTAGA	GACATCTATT	ATAAAGGATG
3651	CACTTAGAAC	AAAACATATA	AGCTATCACT	TCAAGAGACC	AACATTACAA
	GTGAATCTTG	TTTTGATATT	TCGATAGTGA	AGTTCTCTGG	TTGTAATGTT
3701	ATTTATTTTA	ATTCCTAAG	GTGAAAAAAA	TCCTTCCTTC	CTGGTTTCTC
	TAAATAAAAT	TAAGGGATTC	CACTTTTTTT	AGGAAGGAAG	GACCAAAGAG
3751	AAGAGAAAGT	CTATACTGGT	AACCAAATTC	ACTTTAAACA	GGCATTTTCT
	TTCTCTTTCA	GATATGACCA	TTGGTTTAAAG	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	ATTAAAAAATG	GGGGAAAAAA	AACCCTGAAA	TGGTTAATCA
	AATACTCTCC	TAATTTTTTAC	CCCCTTTTTT	TTGGGACTTT	ACCAATTAGT
3951	GAAGATCCTA	TGGGCTGAGA	AGGAATCCAT	CTTAACATTT	CATCTTAAAG
	CTTCTAGGAT	ACCCGACTCT	TCCTTAGGTA	GAATTGTAAA	GTAGAATTTT
4001	CAAATGCTAT	TGCCGGGGGG	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	TTAATCGGTC
4151	GCATAGTGGT	GCATGCCTGT	AATCCCAGCT	ACTTGGGAGG	CTGAGGCAGG
	CGTATCACCA	CGTACGGACA	TTAGGGTCGA	TGAACCCTCC	GACTCCGTCC
4201	AGAACTGCTT	GAACCCAGGA	GGCTTAAGTT	GCGGTGAGCC	AAGATCACGC
	TCTTGACGAA	CTTGGGTCCT	CCGAATTCAA	CGCCACTCGG	TTCTAGTGCG

FIG. 12(contd.) 21/93

4251	CATTGCACTC	TAGCCTGGAC	AACAAGAGAA	AAACTCTGTC	TCAAAAAAAC
	GTAACGTGAG	ATCGGACCTG	TTGTTCTCTT	TTTGAGACAG	AGTTTTTTTG
4301	ACAAAAACAA	AAAACCCAAA	TACTATTTAA	AAAAGATAAA	CCTTAATTGC
	TGTTTTTTGTT	TTTTGGGTTT	ATGATAAATT	TTTTCTATTT	GGAATTAACG
4351	TCAATCATTA	AAGCCATCCC	ACAAGTAAAG	CAGCAAGCAG	AAAAAAGTTA
	AGTTAGTAAT	TTCCGGTAGGG	TGTTCAATTC	GTCGTTTCGTC	TTTTTTTCAAT
4401	AGAACACCTC	AAGGCTACAG	AAGGACATTT	CAAGCTATGC	AGGCATATGA
	TCTTGTGGAG	TTCCGATGTC	TTCCTGTAAA	GTTGATACG	TCCGTATACT
4451	AGTGTGCAGA	CAGATATGTA	AGAAAGGCCT	CAAGACTGCA	AAAGGGCATT
	TCACACGTCT	GTCTATACAT	TCTTTCCGGA	GTTCTGACGT	TTTCCCGTAA
4501	TCAAGCTATG	CAAGCATATA	GGTAACACAT	ACACACACAC	AAAATAAAAT
	AGTTCGATAC	GTTCGTATAT	CCATTGTGTA	TGTGTGTGTG	TTTTATTTTA
4551	CCCCTGAAAT	ACAAAAACAT	GCAGCAAACA	CCTGACGTTT	TTGGATACCA
	GGGGACTTTA	TGTTTTTTGTA	CGTCGTTTGT	GGACTGCAAA	AACCTATGGT
4601	TTTCTAAGTC	AGGTGTTATG	ATTCTCATTA	GTCAAGATAC	TTGAGTACTG
	AAAGATTCAG	TCCACAATAC	TAAGAGTAAT	CAGTTCTATG	AACTCATGAC
4651	GGCCCAAACA	GCTTTCTGCC	ACTGTACAGT	ACAAGAAGGT	AGGAATAATG
	CCGGGTTTGT	CGAAAGACGG	TGACATGTCA	TGTTCTTCCA	TCCTTATTAC
4701	GTGGGAGGAG	CAAAGACAAA	CTGTAATAGA	CAGAAGTGTA	TCAGATACCT
	CACCCTCCTC	GTTTCTGTTT	GACATTATCT	GTCTTCACAT	AGTCTATGGA
4751	ATACTACATG	AAAAACAAAA	CAGCTACTGC	CACAAAGGGA	GAAGGCTAAC
	TATGATGTAC	TTTTTGTTTT	GTGATGACG	GTGTTTCCCT	CTTCCGATTG
4801	AAAATAAAGT	CAACAATAAA	TACAGAAAAT	GAAAAGGATA	CACACTAAGG
	TTTTATTTCA	GTTGTTATTT	ATGTCTTTTA	CTTTTCCTAT	GTGTGATTCC
4851	TTTACAAAAA	AAAAAAGGCA	GACAAAATGC	CATACAGTAT	TCATTCACTA
	AAATGTTTTT	TTTTTTCCGT	CTGTTTTACG	GTATGTCATA	AGTAAGTGAT
4901	CTATGGCATT	CATAAGCTAG	TTTCAAATGC	TCACTATTTT	CTTTTATAGT
	GATACCGTAA	GTATTCGATC	AAAGTTTACG	AGTGATAAAA	GAAAATATCA
4951	ATATATTTGC	CTTAACCCAG	CACTTTTTTC	CAAAGTGGA	TGAGTCAAAA
	TATATAACG	GAATTGGGTC	GTGAAAAAAG	GTTTTTCACCT	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	TGTTTTTATTT	GATTTATGTT	ATTGTCTGGG	GTTTCCGATA
5151	CAATTGCGTG	CAAAAACAAC	TTCTGTTTTT	CAGGGTAAAC	AGAATCTAAT
	GTTAACGCAC	GTTTTTTGTTG	AAGACAAAAG	GTCCCATTTG	TCTTAGATTA
5201	GCAGAATCTA	ATGCAGGGTA	AACAGACTTA	ATGCAGAATC	TAATGATGGC
	CGTCTTAGAT	TACGTCCCAT	TTGTCTGAAT	TACGTCTTAG	ATTACTACCG
5251	ACAAATTAAT	AATCACTAAC	GTGCCCTTTT	TAGTGTGAAA	CCCAGAGAGA
	TGTTTAATTT	TTAGTGATTG	CACGGGAAAA	ATCACACTTT	GGGTCTCTCT
5301	GCACATACAA	GCCAAAAACA	AATGCTTTAT	TTTACCTAGG	AGACATTAAC
	CGTGTATGTT	CGGTTTTTTGT	TTACGAAATA	AAATGGATCC	TCTGTAATTG
5351	ATTCACCTTT	ACGTGTTTAA	GATTAATGCA	ATGTTAAATA	TTGTGAAAAC
	TAAGTGGAAA	TGCACAAATT	CTAATTACGT	TACAATTTAT	AACACTTTTG
5401	TGTAACCTTG	AATTTTCATGA	TTTTTATGTG	AATATTCCAG	GGTTTAAAAA
	ACATTGAAAC	TTAAAGTACT	AAAAATACAC	TTATAAGGTC	CCAAATTTTT
5451	AACTTGTAAC	ATGACATGGC	TGAATAAGAT	AAAAAATAAA	TCTAGCCTTT
	TTGAACATTG	TACTGTACCG	ACTTATTCTA	TTTTTTTTTT	AGATCGGAAA
5501	TCTCCCTTCT	GGCTCATATT	TGCGATTTCT	ATCATTTTGT	TTAAAAACA
	AGAGGGAAGA	CCGAGTATAA	ACGCTAAAGC	TAGTAAAACA	AATTTTTTTG
5551	AAACACTGCA	ATGAATTAAT	CTTAATATTC	TTCTATGTTT	TAGAGTAAGT
	TTTGTGACGT	TACTTAATTT	GAATTATAAG	AAGATACAAA	ATCTCATTCA
5601	TAAAACAAGA	TAAAGTGACC	AAAGTAATTT	GAAAGATTCA	ATGACTTTTG
	ATTTTGTCT	ATTTCACTGG	TTTCATTAAA	CTTTCTAAGT	TACTGAAAAC
5651	CTCCAACCTA	GGTGACACAAG	GTACCTTGTT	CTTTAAATTG	GGCTTTAATG
	GAGGTTGGAT	CCACGTGTTT	CATGGAACAA	GAAATTTAAC	CCGAAATTAC

FIG. 12(contd.) 22/93

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

FIG. 12(contd.) 23/93

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

FIG. 12(contd.) 24/93

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

FIG. 12(contd.) 25/93

10051	ACAAGTAACC	ATTTCAAAAG	ACTGCTTTGT	GTCTCATAGG	AGCTAGTTTG
	TGTTTCATTGG	TAAAGTTTTTC	TGACGAAACA	CAGAGTATCC	TCGATCAAAC
10101	ATCATTTCAG	TTAATTTTTT	CTTTAATTTT	TACGAGTCAT	GAAAACCTACA
	TAGTAAAGTC	AATTAAAAAA	GAAATTAATA	ATGCTCAGTA	CTTTTGATGT
10151	GGAAAAAAA	TCTGAACTGG	GTTTTACCAC	TACTTTTTTAG	GAGTTGGGAG
	CCTTTTTTTT	AGACTTGACC	CAAAATGGTG	ATGAAAAATC	CTCAACCCCTC
10201	CATGCGAATG	GAGGGAGAGC	TCCGTAGAAC	TGGGATGAGA	GCAGCAATTA
	GTACGCTTAC	CTCCCTCTCG	AGGCATCTTG	ACCCTACTCT	CGTCGTTAAT
10251	ATGCTGCTTG	CTAGGAACAA	AAAATAATTG	ATTGAAAATT	ACGTGTGACT
	TACGACGAAC	GATCCTTGTT	TTTTATTAA	TAACTTTTAA	TGCACACTGA
10301	TTTGTAGTTG	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	ATCATGTCGT	ATATTTCTAA
10451	TCAGCTCTGA	GGTTTTTTCCT	ATTGACTTGG	AAAATTGTTT	TGTGCCTGTC
	AGTCGAGACT	CCAAAAAGGA	TAAGTGAACC	TTTTAACAAC	ACACGGACAG
10501	GCTTGCCACA	TGGCCAATCA	AGTAAGCTTG	ATTAATAGTA	ATCAATTACG
	CGAACGGTGT	ACCGGTTAGT	TCATTTCGAAC	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	AAAGGTAAC
10701	CGTCAATGGG	TGGAGTATTT	ACGGTAAACT	GCCCCACTTG	CAGTACATCA
	GCAGTTACCC	ACCTCATAAA	TGCCATTTGA	CGGGTGAACC	GTCATGTAGT
10751	AGTGTATCAT	ATGCCAAGTA	CGCCCCCTAT	TGACGTCAAT	GACGGTAAAT
	TCACATAGTA	TACGGTTCAT	GCGGGGGATA	ACTGCAGTTA	CTGCCATTTA
10801	GGCCCGCCTG	GCATTATGCC	CAGTACATGA	CCTTATGGGA	CTTTCCTACT
	CCGGGCGGAC	CGTAATACGG	GTCATGTACT	GGAATACCCT	GAAAGGATGA
10851	TGGCAGTACA	TCTACGTATT	AGTCATCGCT	ATTACCATGG	TGATGCGGTT
	ACCGTCATGT	AGATGCATAA	TCAGTAGCGA	TAATGGTACC	ACTACGCCAA
10901	TTGGCAGTAC	ATCAATGGGC	GTGGATAGCG	GTTTGACTCA	CGGGGATTTT
	AACCGTCATG	TAGTTACCCG	CACCTATCGC	CAAACCTGAGT	GCCCCCTAAAG
10951	CAAGTCTCCA	CCCCATTGAC	GTCAATGGGA	GTTTGTTTTT	GCACCAAAAT
	GTTTCAGAGGT	GGGGTAACTG	CAGTTACCCT	CAAACAAAAC	CGTGGTTTTT
11001	CAACGGGACT	TTCCAAAATG	TCGTAACAAC	TCCGCCCCAT	TGACGCAAAT
	GTTGCCCTGA	AAGGTTTTTAC	AGCATTGTTG	AGGCGGGGTA	ACTGCGTTTA
11051	GGGCGGTAGG	CGTGTACGGT	GGGAGGTCTA	TATAAGCAGA	GCTGGTTTTAG
	CCCGCCATCC	GCACATGCCA	CCCTCCAGAT	ATATTCGTCT	CGACCAAATC
11101	TGAACCGTCA	GATCCGCTAG	CGTTCGAAGT	TTAAACGCGG	CCGCGACTCT
	ACTTGGCAGT	CTAGGCGATC	GCAAGCTTCA	AATTTGCGCC	GGCGCTGAGA
11151	AGATCATAAT	CAGCCATACC	ACATTTGTAG	AGGTTTTTACT	TGCTTTAAAA
	TCTAGTATTA	GTCGGTATGG	TGTAAACATC	TCCAAAATGA	ACGAAATTTT
11201	AACCTCCCAC	ACCTCCCCCT	GAACCTGAAA	CATAAAATGA	ATGCAATTGT
	TTGGAGGGTG	TGGAGGGGGA	CTTGGACTTT	GTATTTTACT	TACGTTAACA
11251	TGTTGTTAAC	TTGTTTATTG	CAGCTTATAA	TGGTTACAAA	TAAAGCAATA
	ACAACAATTG	AACAAATAAC	GTCGAATATT	ACCAATGTTT	ATTTCGTTAT
11301	GCATCACAAA	TTTCACAAAT	AAAGCATTTT	TTTCACTGCA	TTCTAGTTGT
	CGTAGTGTTT	AAAGTGTTTA	TTTCGTAAAA	AAAGTGACGT	AAGATCAACA
11351	GGTTTGTCCA	AACTCATCAA	TGTATCTTAA	ATCGAATTCT	ACCGGGTAGG
	CCAAACAGGT	TTGAGTAGTT	ACATAGAATT	TAGCTTAAGA	TGGCCCATCC
11401	GGAGGCGCTT	TTCCCAAGGC	AGTCTGGAGC	ATGCGCTTTA	GCAGCCCCGC
	CCTCCGCGAA	AAGGGTTCCG	TCAGACCTCG	TACGCGAAAT	CGTCGGGGCG
11451	TGGGCACTTG	GCGCTACACA	AGTGGCCTCT	GGCCTCGCAC	ACATTCCACA
	ACCCGTGAAC	CGCGATGTGT	TCACCGGAGA	CCGGAGCGTG	TGTAAGGTGT

FIG. 12(contd.) 26/93

11501	TCCACCGGTA	GGCGCCAACC	GGCTCCGTTC	TTTGGTGGCC	CCTTCGCGCC
	AGGTGGCCAT	CCGCGGTTGG	CCGAGGCAAG	AAACCACCGG	GGAAGCGCGG
11551	ACCTTCTACT	CCTCCCCTAG	TCAGGAAGTT	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	GTGAAACGA	GGAAGCGAAA	GACCCGAGTC	TCCGACCCTT
11751	GGGGTGGGTC	CGGGGGCGGG	CTCAGGGGCG	GGCTCAGGGG	CGGGGCGGGC
	CCCCACCCAG	GCCCCCGCCC	GAGTCCCCGC	CCGAGTCCCC	GCCCCGCCCC
11801	GCCCGAAGGT	CCTCCGGAGG	CCCGGCATTC	TGCACGCTTC	AAAAGCGCAC
	CGGGCTTCCA	GGAGGCCTCC	GGGCCGTAAG	ACGTGCGAAG	TTTTCGCGTG
11851	GTCTGCCGCG	CTGTTCTCCT	CTTCCTCATC	TCCGGGCCTT	TCGACCAGCT
	CAGACGGCGC	GACAAGAGGA	GAAGGAGTAG	AGGCCCGGAA	AGCTGGTCGA
11901	TACCATGACC	GAGTACAAGC	CCACGGTGCG	CCTCGCCACC	CGCGACGACG
	ATGGTACTGG	CTCATGTTTC	GGTGCCACGC	GGAGCGGTGG	GCGCTGCTGC
11951	TCCCCAGGGC	CGTACGCACC	CTCGCCGCCG	CGTTCGCCGA	CTACCCCGCC
	AGGGGTCCCG	GCATGCGTGG	GAGCGGCGGC	GCAAGCGGCT	GATGGGGCGG
12001	ACGCGCCACA	CCGTCGATCC	GGACCGCCAC	ATCGAGCGGG	TCACCGAGCT
	TGCGCGGTGT	GGCAGCTAGG	CCTGGCGGTG	TAGCTCGCCC	AGTGGCTCGA
12051	GCAAGAACTC	TTCCTCACGC	GCGTCGGGCT	CGACATCGGC	AAGGTGTGGG
	CGTTCCTTGAG	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	CGGTTCCCGG	CTGGCCGCGC	AGCAACAGAT	GGAAGGCCTC	CTGGCGCCGC
	GCCAAGGGCC	GACCGGCGCG	TCGTTGTCTA	CCTTCCGGAG	GACCGCGGCG
12251	ACCGGCCCAA	GGAGCCCGCG	TGGTTCCTGG	CCACCGTCGG	CGTCTCGCCC
	TGGCCGGGTT	CCTCGGGCGC	ACCAAGGACC	GGTGGCAGCC	GCAGAGCGGG
12301	GACCACCAGG	GCAAGGGTCT	GGGCAGCGCC	GTCGTGCTCC	CCGGAGTGGA
	CTGGTGGTCC	CGTTCCCAGA	CCCGTCGCGG	CAGCACGAGG	GGCCTCACCT
12351	GGCGGCCGAG	CGCGCCGGGG	TGCCCCGCTT	CCTGGAGACC	TCCGCGCCCC
	CCGCCGGCTC	GCGCGGCCCC	ACGGGCGGAA	GGACCTCTGG	AGGCGCGGGG
12401	GCAACCTCCC	CTTCTACGAG	CGGCTCGGCT	TCACCGTCAC	CGCCGACGTC
	CGTTGGAGGG	GAAGATGCTC	GCCGAGCCGA	AGTGGCAGTG	GCGGCTGCAG
12451	GAGGTGCCCC	AAGGACCGCG	CACCTGGTGC	ATGACCCGCA	AGCCCGGTGC
	CTCCACGGGC	TTCCTGGCGC	GTGGACCACG	TACTGGGCGT	TCGGGGCCACG
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	GGTGTCATTC	TATTCTGGGG	GGTGGGGTGG	GGCAGGACAG
	ACAGACTCAT	CCACAGTAAG	ATAAGACCCC	CCACCCACC	CCGTCTGTGC
12751	CAAGGGGGGG	GATTGGGRAG	ACAATAGCAG	GCATGCTGGG	GGGGCGGTGG
	GTTCCCCCCC	CTAACCCYTC	TGTTATCGTC	CGTACGACCC	CCCCGCCACC
12801	GGGCTATGGC	TTCTGAGGCG	GAAAGAACCA	GCTGGGGGCTC	GAGATCCACT
	CCCGATACCG	AAGACTCCGC	CTTTCTTGGT	CGACCCCGAG	CTCTAGGTGA
12851	AGTTCTAGCC	TCGAGGCTAG	AGCGGCCGCC	ACCGCGGTGG	AGCTCCAATT
	TCAAGATCGG	AGCTCCGATC	TCGCCGGCGG	TGGCGCCACC	TCGAGGTTAA
12901	CGCCCTATAG	TGAGTCGTAT	TACGCGCGCT	CACTGGCCGT	CGTTTTACAA
	GCGGGATATC	ACTCAGCATA	ATGCGCGCGA	GTGACCGGCA	GCAAAATGTT

27/93

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

FIG. 12(contd.)

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 GCGGGCGCAA 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 ATCGGCCCGC GCATGGCCGA GTTGAGCGGT
   GCCCCCGCCA CAAGCGGCTC TAGCCGGGCG CGTACCGGCT CAACTCGCCA
301  TCCCGGCTGG CCGCGCAGCA ACAGATGGAA GGCCTCCTGG CGCCGCACCG
   AGGGCCGACC GGC GCGTCGT TGTCTACCTT CCGGAGGACC GCGGCGTGGC
351  GCCCAAGGAG CCCGCGTGGT TCCTGGCCAC CGTCGGCGTC TCGCCCGACC
   CGGGTTCCTC GGGCGCACCA AGGACCGGTG GCAGCCGCAG AGCGGGCTGG
401  ACCAGGGCAA GGGTCTGGGC AGCGCCGTCG TGCTCCCCGG AGTGGAGGCG
   TGGTCCCGTT CCCAGACCCG TCGCGGCAGC ACGAGGGGCC TCACCTCCGC
451  GCCGAGCGCG CCGGGGTGCC CGCCTTCCTG GAGACCTCCG CGCCCCGCAA
   CGGCTCGCGC GGCCCCACGG GCGGAAGGAC CTCTGGAGGC GCGGGGCGTT
501  CCTCCCCTTC TACGAGCGGC TCGGCTTCAC CGTCACCGCC GACGTCGAGG
   GGAGGGGAAG ATGCTCGCCG AGCCGAAGTG GCAGTGGCGG CTGCAGCTCC
551  TGCCCGAAGG ACCGCGCACC TGGTGCATGA CCCGCAAGCC CGGTGCCTGA
   ACGGGCTTCC TGGCGCGTGG ACCACGTACT GGGCGTTCGG GCCACGGACT
```

29/93

FIG. 14

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

```
1  ATGACTGAAT ACAAACCAAC TGTTGCGCTG GCAACTCGTG ATGATGTTCC
   TACTGACTTA TGTTTGGTTG 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 AGACAAC TTC
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 AACTGACGA CTACACCTTC
551 TGCCAGAAGG CCCACGCACT TGGTGCATGA CTCGCAAACC AGGTGCTTAA
   ACGGTCTTCC GGGTGCGTGA ACCACGTACT GAGCGTTTGG TCCACGAATT
```

30/93

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

```

1  ATGGACGACA GCACGTTGCG CCGGAAGTAC CCGCACCCACG AGTGGCACGC
   TACCTGCTGT CGTGCAACGC GGCCTTCATG GGCGTGGTGC TCACCGTGCG
51 AGTGAACGAA GGAGACTCGG GCGCCTTCGT CTACCAGCTC ACCGGCGGCC
   TCACTTGCTT CCTCTGAGCC CGCGGAAGCA GATGGTCGAG TGGCCGCCGG
101 CCGAGCCCCA GCCCCGAGCTC TACGCGAAGA TCGCCCCCCG CCCCCCGAG
   GGCTCGGGGT CGGGCTCGAG ATGCGCTTCT AGCGGGGGGC GCGGGGGCTC
151 AACTCCGCCT TCGACCTGTC CGGCGAGGCC GACCGGCTGG AGTGGCTCCA
   TTGAGGCGGA AGCTGGACAG GCCGCTCCGG CTGGCCGACC TCACCGAGGT
201 CCGCCACGGG ATCCCCGTCC CCCGCGTCGT CGAGCGCGGT GCCGACGACA
   GGCGGTGCCC TAGGGGCAGG GGGCGCAGCA GCTCGCGCCA CGGCTGCTGT
251 CCGCCGCGTG GCTCGTCACG GAGGCCGTCC CCGGCGTCGC GGCGGCCGAG
   GGCGGCGCAC CGAGCAGTGC CTCCGGCAGG GGCCGCAGCG CCGCCGGCTC
301 GAGTGGCCCG AGCACCAGCG GTTCGCCGTG GTCGAGGCGA TGGCGGAGCT
   CTCACCGGGC TCGTGGTCGC CAAGCGGCAC CAGCTCCGCT ACCGCCTCGA
351 GGCCCGCGCC CTCCACGAGC TGCCCGTGGA GGACTGCCCC TCCGACCGGC
   CCGGGCGCGG GAGGTGCTCG ACGGGCACCT CCTGACGGGG AGGCTGGCCG
401 GCCTCGACGC GGCGGTGCGC GAGGCCCGGC GGAACGTCGC CGAGGGCTTG
   CGGAGCTGCG CCGCCAGCGG CTCCGGGCCG CCTTGCAGCG 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 CCCCGGGACC
   AGACGGTACC GCTGGACACG GGGTTGTTGC AGGACGAGCT GGGGCCCTGG
601 TGCCGGGTCA CCGGCGTGAT CGACGTCGGC CGCCTCGGGG TCGCCGACCG
   ACGGCCAGT GGCCGCACTA GCTGCAGCCG GCGGAGCCCC AGCGGCTGGC
651 CCACGCCGAC ATCGCCTTGG CCGCCCGCGA GCTGGAGATC GACGAGGACC
   GGTGCGGCTG TAGCGGAACC GGCGGGCGCT CGACCTCTAG CTGCTCCTGG
701 CCTGGTTCGG CCCC GCCTAC GCCGAGCGGT TCCTGGAGCG GTACGGCGCC
   GGACCAAGCC GGGGCGGATG CGGCTCGCCA AGGACCTCGC CATGCCGCGG
751 CACCGCGTCG ACAAGGAGAA GCTGGCCTTC TACCAGCTTC TCGACGAGTT
   GTGGCGCAGC TGTTCTCTT CGACCGGAAG ATGGTCGAAG AGCTGCTCAA
801 CTTCTAG
   GAAGATC

```

31/93

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

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

32/93

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

```

1  ATGAAAAAGC CTGAACTCAC CGCGACGTCT GTCGCGAAGT TTCTGATCGA
   TACTTTTTCG GACTTGAGTG GCGCTGCAGA CAGCGCTTCA AAGACTAGCT
51  AAAGTTCGAC AGCGTCTCCG ACCTGATGCA GCTCTCGGAG GGCGAAGAAT
   TTTCAAGCTG TCGCAGAGGC TGGACTACGT CGAGAGCCTC CCGCTTCTTA
101 CTCGTGCTTT CAGCTTCGAT GTAGGAGGGC GTGGATATGT CCTGCGGGTA
   GAGCACGAAA GTCGAAGCTA CATCCTCCCG CACCTATACA GGACGCCCAT
151 AATAGCTGCG CCGATGGTTT CTACAAAGAT CGTTATGTTT ATCGGCACTT
   TTATCGACGC GGCTACCAAA GATGTTTCTA GCAATACAAA TAGCCGTGAA
201 TGCATCGGCC GCGCTCCCGA TTCCGGAAGT GCTTGACATT GGGGAATTCA
   ACGTAGCCGG CGCGAGGGCT AAGGCCTTCA CGAACTGTAA CCCCTTAAGT
251 GCGAGAGCCT GACCTATTGC ATCTCCCGCC GTGCACAGGG TGTACCGTTG
   CGCTCTCGGA CTGGATAACG TAGAGGGCGG CACGTGTCCC ACAGTGCAAC
301 CAAGACCTGC CTGAAACCGA ACTGCCCGCT GTTCTGCAAC CCGTCGCGGA
   GTTCTGGACG GACTTTGGCT TGACGGGCGA CAAGACGTTG GGCAGCGCCT
351 GCTCATGGAT GCGATCGCTG CGGCCGATCT TAGCCAGACG AGCGGGTTCG
   CGAGTACCTA CGCTAGCGAC GCCGGCTAGA ATCGGTCTGC TCGCCCAAGC
401 GCCCATTCGG ACCGCAAGGA ATCGGTCAAT AACTACATG GCGTGATTTC
   CGGGTAAGCC TGGCGTTCCT TAGCCAGTTA TGTGATGTAC CGCACTAAAG
451 ATATGCGCGA TTGCTGATCC CCATGTGTAT CACTGGCAA CTGTGATGGA
   TATACGCGCT AACGACTAGG GGTACACATA GTGACCGTTT GACACTACCT
501 CGACACCGTC AGTGCGTCCG TCGCGCAGGC TCTCGATGAG CTGATGCTTT
   GCTGTGGCAG TCACGCAGGC AGCGCGTCCG AGAGCTACTC GACTACGAAA
551 GGGCCGAGGA CTGCCCCGAA GTCCGGCACC TCGTGCACGC GGATTTCGGC
   CCCGGCTCCT GACGGGGCTT CAGGCCGTGG AGCACGTGCG CCTAAAGCCG
601 TCCAACAATG TCCTGACGGA CAATGGCCGC ATAACAGCGG TCATTGACTG
   AGGTTGTTAC AGGACTGCCT GTTACCGGCG TATTGTCGCC AGTAACTGAC
651 GAGCGAGGCG 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 GCCGAGGCCG GCATATACGA
801 CCGCATTTGG CTTGACCAAC TCTATCAGAG CTTGGTTGAC GGCAATTTTCG
   GCGGTAACCA GAACTGGTTG AGATAGTCTC GAACCAACTG CCGTTAAAGC
851 ATGATGCAGC TTGGGCGCAG GGTGATGCG ACGCAATCGT CCGATCCGGA
   TACTACGTCT 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 CGTGCGTCCA AGAGGCCGGC GAACCCACCT
51  GAGGCTATTC GGCTATGACT GGGCACAACA GACAATCGGC TGCTCTGATG
   CTCCGATAAG CCGATACTGA CCCGTGTTGT CTGTTAGCCG ACGAGACTAC
101 CCGCCGTGTT CCGGCTGTCA GCGCAGGGGC GCCCGGTTCT TTTTGTCAAG
   GGCGGCACAA GGCCGACAGT CGCGTCCCCG CGGGCCAAGA AAAACAGTTC
151 ACCGACCTGT CCGGTGCCCT GAATGAACTG CAAGACGAGG CAGCGCGGCT
   TGGCTGGACA GGCCACGGGA CTTACTTGAC GTTCTGCTCC GTCGCGCCGA
201 ATCGTGGCTG GCCACGACGG GCGTTCCTTG CGCAGCTGTG CTCGACGTTG
   TAGCACCGAC CCGTGCTGCC CGCAAGGAAC GCGTCGACAC GAGCTGCAAC
251 TCACTGAAGC GGAAGGGAC TGGCTGCTAT TGGGCGAAGT GCCGGGGCAG
   AGTGACTTCG CCCTTCCCTG ACCGACGATA ACCCGCTTCA CGGCCCCGTC
301 GATCTCCTGT CATCTCACCT TGCTCCTGCC GAGAAAGTAT CCATCATGGC
   CTAGAGGACA GTAGAGTGGA ACGAGGACGG CTCTTTCATA GGTAGTACCG
351 TGATGCAATG CGGCGGCTGC ATACGCTTGA TCCGGCTACC TGCCCATTCG
   ACTACGTTAC GCCGCCGACG TATGCGAACT AGGCCGATGG ACGGGTAAGC
401 ACCACCAAGC GAAACATCGC ATCGAGCGAG CACGTACTCG GATGGAAGCC
   TGGTGGTTCG CTTTGTAGCG TAGCTCGCTC GTGCATGAGC CTACCTTCGG
451 GGTCTTGTCG ATCAGGATGA TCTGGACGAA GAGCATCAGG GGCTCGCGCC
   CCAGAACAGC TAGTCCTACT AGACCTGCTT CTCGTAGTCC CCGAGCGCGG
501 AGCCGAACTG TTCGCCAGGC TCAAGGCGAG CATGCCCGAC GGCGAGGATC
   TCGGCTTGAC AAGCGGTCCG AGTTCCGCTC GTACGGGCTG CCGCTCCTAG
551 TCGTCGTGAC CCATGGCGAT GCCTGCTTGC CGAATATCAT GGTGGAAAAT
   AGCAGCACTG GGTACCGCTA CGGACGAACG GCTTATAGTA CCACCTTTTA
601 GGCCGCTTTT CTGGATT CAT 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	GGAAC TAGGT	CACATTCGTG
51	CTCCTAGAAA	GTCAGTAGAC	AATAAAACAA	AAGTTCTGCT	TCACCGATTT
	GAGGATCTTT	CAGTCATCTG	TTATTTTGT	TTCAAGACGA	AGTGGCTAAA
101	ACATTTATAA	CCAAATACCC	TTACCAATA	CAATAAAAAA	ACAAAACAAC
	TGTAAATATT	GGTTTATGGG	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	TCTTTTGTAC	TCTGTTGTTA	CCTCTCGACT
301	GTAAC TATTA	GTAGTAGTAC	TTTACCAGAG	AATGGCCTCT	ATAGGCTCAC
	CATTGATAAT	CATCATCATG	AAATGGTCTC	TTACCGGAGA	TATCCGAGTG
351	ATGTAGGAAT	GGTTGGTCCC	CAGGTGGTAG	GTAGAGCTGT	TTGAGGATTA
	TACATCCTTA	CCAACCAGGG	GTCCACCATC	CATCTCGACA	AACTCCTAAT
401	CGTGGCCTTC	TTGGATGGGG	GGTGGGGGTG	GGGTGGGAGG	GTTGGGTGGT
	GCACCGGAAG	AACCTACCCC	CCACCCCCAC	CCCACCTCC	CAACCCACCA
451	GGGTACTTAA	GAGGTTTCAA	AAGTCAATAT	TGTTTGCATT	TAGCTCTTCC
	CCCATGAATT	CTCCAAAGTT	TTCAGTTATA	ACAAACGTAA	ATCGAGAAGG
501	TTGTACTTGT	GGATCAAACA	CAACCTGTCA	GCTACTGCTT	CAAATGTCAT
	AACATGAACA	CCTAGTTTGT	GTTGGACAGT	CGATGACGAA	GTTTACAGTA
551	GCCTGCTGCC	ATCTTCTCAG	CAGGATGGTC	ATGGCCTCAC	CCTCTTCAAC
	CGGACGACGG	TAGAAGAGTC	GTCCTACCAG	TACCGGAGTG	GGAGAAGTTG
601	TGTAAATCTT	TCTTTCTTTT	CTTCTTTTTC	TTTTGGTTTC	GAGACAGGGT
	ACATTTAGAA	AGAAAGAAAA	GAAGAAAAAG	AAAACCAAAG	CTCTGTCCCA
651	TTCTCTGTAT	AGTCTTGGCT	GTCCTGGAAC	TCACTTTGTA	GACCAGGCTG
	AAGAGACATA	TCAGGACCGA	CAGGACCTTG	AGTGAAACAT	CTGGTCCGAC
701	GCCTTGAACT	CAGAAATCCG	CCTGCCTCTG	CCTCCCTAGC	ACTGGGATTA
	CGGAAC TTGA	GTCCTTAGGC	GGACGGAGAC	GGAGGGATCG	TGACCCTAAT
751	AAGGCGTGCG	CCACCACGCC	CAGCTTTCAA	CTGGAAATCT	TAATAAACTT
	TTCCGCACGC	GGTGGTGCGG	GTCGAAAGTT	GACCTTTAGA	ATTATTTGAA
801	TCCTAGAAGT	GGCCTTGGTT	ATGGGAGCTT	ATCACAGCAA	TAGAACAGCA
	AGGATCTTCA	CCGGAACCAA	TACCCTCGAA	TAGTGTGCTT	ATCTTGTGCT
851	ATTATGACTG	GAGTATGATA	GTAAAAACA	AGCAAGCAAG	CAAGCAAACA
	TAATACTGAC	CTCATACTAT	CAATTTTGT	TCGTTCGTTC	GTTGCTTTGT
901	CACACACCAA	AACAACAAAA	CCCCAAGACA	GAGTCACATG	TAGCCCAGGC
	GTGTGTGGTT	TTGTTGTTTT	GGGGTTCTGT	CTCAGTGTAC	ATCGGGTCCG
951	TAGCCTCCAA	ATTCACTATA	TAAC TGAAGA	AGACCCCTAA	TTCCCATTCC
	ATCGGAGGTT	TAAGTGATAT	ATTGACTTCT	TCTGGGGATT	AAGGGTAAGG
1001	TCTAGAATCT	ATACCTCAAG	TACTGAATGG	CTTGGTTCAC	AATACCCAC
	AGATCTTAGA	TATGGAGTTC	ATGACTTACC	GAACCAAGTG	TTATGGGGTG
1051	TAAATGATTG	GTCTTACTAA	GTGCAACAAG	GTAACCTAA	AACTTCAGCC
	ATTTACTAAC	CAGAATGATT	CACGTTGTTC	CATTTGGATT	TTGAAGTCGG
1101	CTCAGACATC	CCTTTTCCAG	TATCAATTTA	TAAAATTAGA	TCCCAAGGAT
	GAGTCTGTAG	GGAAAAGGTC	ATAGTTAAAT	ATTTTAATCT	AGGGTTCCTA
1151	AAAAATTAAT	TGTAAAGTAA	AATCAGAGTT	CTAGCATCAA	CTACAGGCTC
	TTTTTAATTA	ACATTTCAAT	TTAGTCTCAA	GATCGTAGTT	GATGTCCGAG
1201	AACCATGGGG	ACCACAAATA	AACTAAAAGG	GATAAGACTG	GCTTCCCCAT
	TTGGTACCCC	TGGTGTTTAT	TTGATTTTCC	CTATTCTGAC	CGAAGGGGTA
1251	AATTATTACA	TTTAGATAAT	TTTCCTGACT	ACTCAACAAA	GCTAAAATAT
	TTAATAATGT	AAATCTATTA	AAAGGACTGA	TGAGTTGTTT	CGATTTTATA
1301	CACCACTGGT	TTATTTTCTC	CTTCTAGGGT	TTAAGCTCAC	TCTGAGGAGG
	GTGGTGAQCA	AATAAAAGAG	GAAGATCCCA	AATTCGAGTG	AGACTCCTCC

FIG. 19(contd.)

35/93

1351	GGCATGCGGC	ACACACTCAT	AGCATCCAGG	AAATAGAAAT	ATGGTGA
	CCGTACGCCG	TGTGTGAGTA	TCGTAGGTCC	TTTATCTTTA	TACCACTGAT
1401	TCATGGGTTC	AGGGCCAACC	TAGGCTTTAG	AGAAAAACCT	TGTCCCACAA
	AGTACCCAAG	TCCCGGTTGG	ATCCGAAATC	TCTTTTTTGA	ACAGGGTGTT
1451	ACCAAAAATG	TCTCTTTTTT	ATTCTATCAG	GGGTGGATGG	ATTTGTTAAA
	TGGTTTTTAC	AGAGAAAAAA	TAAGATAGTC	CCCACCTACC	TAAACAATTT
1501	GAAGTGCTTT	TAAAAACCTT	GAGATGGTTA	TTTAGAAGTC	CCCATGGGAT
	CTTCACGAAA	ATTTTTTGGA	CTCTACCAAT	AAATCTTCAG	GGGTACCCTA
1551	ACCAAAATAA	CCCCTATTTT	ATATGCCCAA	GCATTTTCACC	TCCACAACAG
	TGGTTTTTATT	GGGTGATAAA	TATACGGGTT	CGTAAAGTGG	AGGTGTTGTC
1601	TGCTATGCAC	CCTTTAACAT	TTTTGAGACA	GTAGCCCAGT	CTAGTCTTTA
	ACGATACGTG	GGAAATTGTA	AAAACCTCTG	CATCGGGTCA	GATCAGAAAT
1651	ACTTGCAGTG	ATTTTTTCCTG	ATTCAGCTTC	TCCCAGTGCT	GGAATTATAG
	TGAACGTCAC	TAAAAAGGAC	TAAGTCGAAG	AGGGTCACGA	CCTTAATATC
1701	GTATGCACCA	CCATGTGTAA	CTACAGATGC	TACTTAAAAA	TTTTTTTAAAG
	CATACGTGGT	GGTACACATT	GATGTCTACG	ATGAATTTTT	AAAAAATTTT
1751	GAATCACAAA	AATAACCCCC	TATCAAATGC	CTAGTCCCTC	TAACCATCAC
	CTTAGTGTTT	TTATTGGGGG	ATAGTTTACG	GATCAGGGAG	ATTGGTAGTG
1801	CAAGTGAAGG	ATCACGCAGG	AAAAAAAAAA	TCACCAGCAG	CACCTCAGAA
	GTTCACTTCC	TAGTGCCTCC	TTTTTTTTTT	AGTGGTCGTC	GTGGAGTCTT
1851	CCAGGATACT	CAGTCCATCA	GCATCCAGGG	CCATACCCAC	ACTCACAGCA
	GGTCCTATGA	GTCAGGTAGT	CGTAGGTCCC	GGTATGGGTG	TGAGTGTCGT
1901	TCTCCACAGT	TTACCAGATG	ATTCATGCTT	ATCACTGTAT	TGGGTCATCT
	AGAGGTGTCA	AATGGTCTAC	TAAGTACGAA	TAGTGACATA	ACCCAGTAGA
1951	AAGAGTGACC	ATCAGGGCTT	CTGATCACAG	AATCTAGTCC	ACTTTGCAGA
	TTCTCACTGG	TAGTCCCGAA	GACTAGTGTC	TTAGATCAGG	TGAAACGTCT
2001	CCAGTTGAAG	TCATGCACTA	TATGAGATAG	AAATACCCTC	TTGCTCATTT
	GGTCAACTTC	AGTACGTGAT	ATACTCTATC	TTTATGGGAG	AACGAGTAAA
2051	TGGTCAGAAA	TTCAAGGATA	AAAACCCATG	TTTTGTTAAT	GCACACCTCC
	ACCAGTCTTT	AAGTTCCTAT	TTTTGGGTAC	AAAACAATTA	CGTGTGGAGG
2101	ATATGATTGA	GATCAATGTG	TCCTAATTAA	TGTAGAAACC	ACAACGTGTA
	TATACTAACT	CTAGTTACAC	AGGATTAATT	ACATCTTTGG	TGTTGACATT
2151	ATTTCACTCT	TTTGACATGA	ATCTTTTTCT	AGACAGGGTC	TTGGATGCAG
	TAAAGTGAGA	AAACTGTACT	TAGAAAAAGA	TCTGTCCCAG	AACCTACGTC
2201	CCCCGACTAC	CCAGAATTTT	GGAATCCAGG	CTAGCCTCAA	ACTCAAGGCA
	GGGGCTGATG	GGTCTTAAAA	CCTTAGGTCC	GATCGGAGTT	TGAGTTCCGT
2251	ATCTGCTTGC	TTCAGCTTCT	CACAGGCTGG	ATCACAAACA	TACACCTTCA
	TAGACGAACG	AAGTCGAAGA	GTGTCCGACC	TAGTGTTTGT	ATGTGGAAGT
2301	GACCCATTTT	TTTTTCCTCC	CTCCGTTTTT	GGTTTCTCTG	TGTAGCCCTG
	CTGGGTAAAA	AAAAAGGAGG	GAGGCAAAAA	CCAAAGAGAC	ACATCGGGAC
2351	GGTGTCCGTG	GACTCGCTGT	GTAGATCTAT	CTACCAGCCT	CTGTCTTGGA
	CCACAGGCAC	CTGAGCGACA	CATCTAGATA	GATGGTCGGA	GACAGAACCT
2401	GTA	TAAAGTTGTG	GGCTACCACT	GCCTGGCTGA	CCCAGTTTAA
	CATGACCCTA	ATTTCAACAC	CCGATGGTGA	CGGACCGACT	GGGTCAAAAT
2451	TTTATTTTAA	ATATAACTTG	ACAAAAATAA	ATTTGTCTAA	CTTACTAGAA
	AAATAAAATT	TATATTGAAC	TGTTTTTTATT	TAAACAGATT	GAATGATCTT
2501	ATCCCAAGAA	AACTAACACT	GGATTTAGCA	ACAGTCAGAA	ATCGCTGAAA
	TAGGGTTCTT	TTGATTGTGA	CCTAAATCGT	TGTCAGTCTT	TAGCGACTTT
2551	AGAAACAGAA	TTGATCTAAC	AGTCTTAGAT	CACTCCTAGA	CAGTTTGTA
	TCTTTGTCTT	AACTAGATTG	TCAGAATCTA	GTGAGGATCT	GTCAAACATT
2601	TTCTTGCTCA	TGGCAACGTG	AGCTCTATCT	AACTCACTCT	CTGTGCACTA
	AAGAACGAGT	ACCGTTGCAC	TCGAGATAGA	TTGAGTGAGA	GACACGTGAT
2651	ATGAATGCTC	AGTGTCTCCA	GAACAGCACA	GCTTCCAGGG	TAATCATGCC
	TACTTACGAG	TCACAGAGGT	CTTGTCGTGT	CGAAGGTCCC	ATTAGTACGG
2701	AACCCACAAG	ACTTTTATAG	AGCTGTCCAC	GACTCTTCCC	CCATTTCAGCT
	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	GCATTTTTTA	CCTCAAAAGC	ATCCAGCAAA
	CAAAATATTG	GGTGTTTGTA	CGTAAAAAAT	GGAGTTTTTCG	TAGGTCGTTT
2901	AACTGTACAA	CGCTTTTTCA	AAAAAATGTA	TTGTGATCCT	CCTTAAGAAA
	TTGACATGTT	GCGAAAAAGT	TTTTTTTACAT	AACACTAGGA	GGAATTCTTT
2951	AGCCTTACTT	AGTGTTAATT	CCTTTTTTCTT	TAGAATGCTG	GTAAATACAA
	TCGGAATGAA	TCACAATTAA	GGAAAAAGAA	ATCTTACGAC	CATTTATGTT
3001	GGACTTAGGT	AGGCTGGCTT	CTAACAGCAA	TTCACCCACT	TATGATGGGA
	CCTGAATCCA	TCCGACCGAA	GATTGTCGTT	AAGTGGGTGA	ATACTACCCT
3051	TTAAAGGAAG	GCACAACCAT	GTCCACCACA	GGTTCTAGCT	CCCCCACCCA
	AATTTCTTTC	CGTGTTGGTA	CAGGTGGTGT	CCAAGATCGA	GGGGGTGGGT
3101	CACGCCCAGA	GAGGGTTTTT	CTGTGTAGCT	CTGACTATTC	TGGAATTCAC
	GTGCGGGTCT	CTCCCAAAAA	GACACATCGA	GACTGATAAG	ACCTTAAGTG
3151	ACTGCAGACC	AGGCTGGTCT	CGAACTCAGA	GATCCACCAC	CACATGGTTT
	TGACGTCTGG	TCCGACCAGA	GCTTGAGTCT	CTAGGTGGTG	GTGTACCAAA
3201	CTTAATTGTA	ATTTTAAAGA	AAAAAATAAA	TCCTTCAGTT	AAGATTCTTA
	GAATTAACAT	TAAAATTTCT	TTTTTTTTTTT	AGGAAGTCAA	TTCTAAGAAT
3251	TGTTCTAGGT	TTTCACAAAC	TTACCAATGT	AGTTTTATTG	GAGGCCATTT
	ACAAGATCCA	AAAGTGTTTG	AATGGTTACA	TCAAATAAAC	CTCCGGTAAA
3301	TTTAAATTTA	ATCGGAGACT	TGAAGAGCTA	TTGCAAGAAA	AAAAATGTAG
	AAATTTAAAT	TAGCCTCTGA	ACTTCTCGAT	AACGTTCTTT	TTTTTACATC
3351	GACAGTTAAA	ATTTCATGAC	ACACAAAAGG	CAGCTACAAG	TTTTGTGTGG
	CTGTCAATTT	TAAAGTACTG	TGTGTTTTCC	GTCGATGTTC	AAAACACACC
3401	ATTTCAACAT	GTAAATTTTCG	GGTAAAAATG	CAGGAAAACA	GTTGAGTTCC
	TAAAGTTGTA	CATTTAAAGC	CCATTTTTTAC	GTCCTTTTGT	CAACTCAAGG
3451	CGTGTTATTA	GTATGTTACT	AATAATTTCA	GTATGTTAGT	GAAAATAATC
	GCACAATAAT	CATACAATGA	TTATTAAAGT	CATACAATCA	CTTTTATTAG
3501	TTACTAAAAC	ACTGGTACCT	CAGACAACCT	TACATGGTGA	GGATTGTTAC
	AATGATTTTG	TGACCATGGA	GTCTGTTGAA	ATGTACCACT	CCTAACAATG
3551	TTTCCCAATC	CATATAGAAT	TTTAAACAATT	TTAGTGTTTA	TTTTGGATGA
	AAAGGGTTAG	GTATATCTTA	AAATTGTAA	AATCACAAT	AAAACCTACT
3601	AAGGAAATGA	CTATCTTTTG	TTAGCAAATT	ACCATAAGAT	CTTTTTCTTT
	TTCCTTTACT	GATAGAAAAC	AATCGTTTAA	TGGTATTCTA	GAAAAAGAAA
3651	AGATTTCTGA	ATACTCCAAG	GAGCTCATAT	AATTCCATCC	TTATTTTTTTC
	TCTAAAGACT	TATGAGGTTC	CTCGAGTATA	TTAAGGTAGG	AATAAAAAAG
3701	AGAGGCCCTC	CCTGTTCAAT	CACGGTATAA	AAAAAGGAAC	ACATTAAGAT
	TCTCCGGGAG	GGACAAGTTA	GTGCCATATT	TTTTTCCTTG	TGTAATTCTA
3751	GTCCCAGTCC	TATTTTCTGG	CTTTTTTTTTT	CCGGGGGTGG	TGGTGCGGTA
	CAGGGTCAGG	ATAAAAGACC	GAAAAAATAA	GGCCCCCACC	ACCACGCCAT
3801	ATCACTCTCT	ATAGTCCAGT	CTGGGCTTCA	ACGCCTGGCA	ATCCCCAGCC
	TAGTGAGAGA	TATCAGGTCA	GACCCGAAGT	TGCGGACCGT	TAGGGGTTCG
3851	TCAAGCTCCC	AAGTACTGTC	CTGATAAGGA	TAGAAGGAGT	CGACCTCCTT
	AGTTCGAGGG	TTTATGACAG	GACTATTCCT	ATCTTCCTCA	GCTGGAGGAA
3901	CACGCTCCCC	TCCGAGGAGG	GCTCCTTCCC	AGCTCCATTC	CCCGGTCGGG
	GTGCGAGGGG	AGGCTCCTCC	CGAGGAAGGG	TCGAGGTAAG	GGGCCAGCCC
3951	AGCCCGTCCT	CCACCCGAGA	GCGCGGGCCT	CGTGGTCAGC	GCCTCCGCGG
	TCGGGCAGGG	GGTGGGCTCT	CGCGCCCGGA	GCACCAGTCG	CGGAGGCGCC
4001	GGAGAAACAA	AGGCGGCGGC	GGGGGCTCAA	GGGCACTGCG	CCACGGGCCC
	CCTCTTTGTT	TCCGCCGCGG	CCCCCGAGTT	CCCGTGACGC	GGTGCCCGGG
4051	GCGCCTCCCC	CATCCGGCGG	CGGCCACGTA	GCCGGGAGCG	CGCCGCAGCC
	CGCGGAGGGG	GTAGGCCGCC	GCCGGTG CAT	CGGCCCTCGC	GCGGCGTCGG
4101	CGGAGCCTCG	GGCCTCGCAG	CTGCAGAGCC	TGAACCGCTC	TCTCCCTGCG
	GCCTCGGAGC	CCGGAGCGTC	GACGTCTCGG	ACTTGGCGAG	AGAGGGACGC
4151	GGCCTGCGAC	GAGGCTGGGG	GAGGGGAGGC	CCGCGCTTTG	TCTGGAGTCT
	CCGGACGCTG	CTCCGACCCC	CTCCCCCTCCG	GGCGCGAAAC	AGACCTCAGA
4201	CGGTAGCTGT	CATCCGGCTC	CCACCCTCAT	GCACAATTGT	CCCATCTCCC
	GCCATCGACA	GTAGGCCGAG	GGTGGGAGTA	CGTGTTAACA	GGGTAGAGGG

FIG. 19(contd.)

37/93

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

FIG. 19(contd.)

38/93

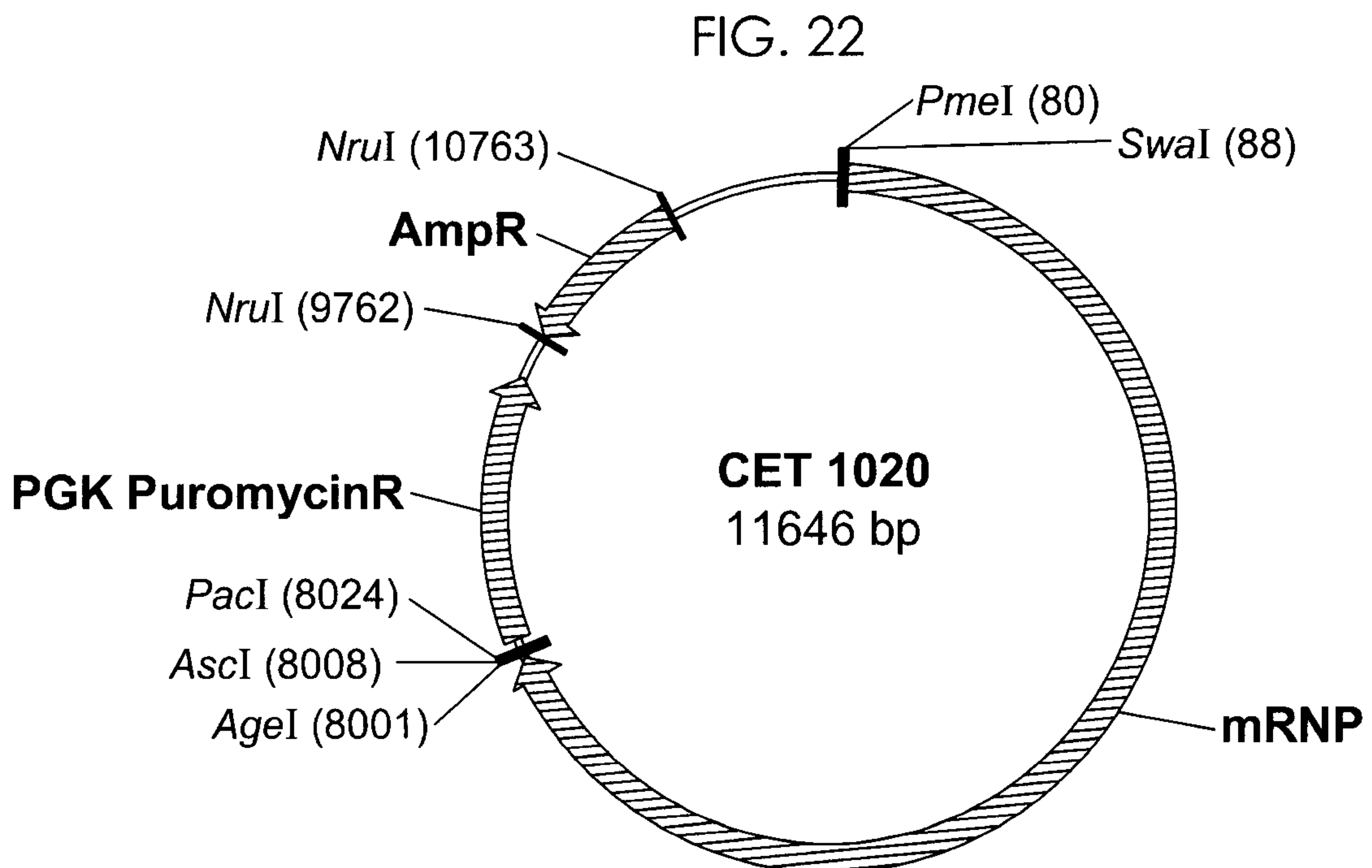
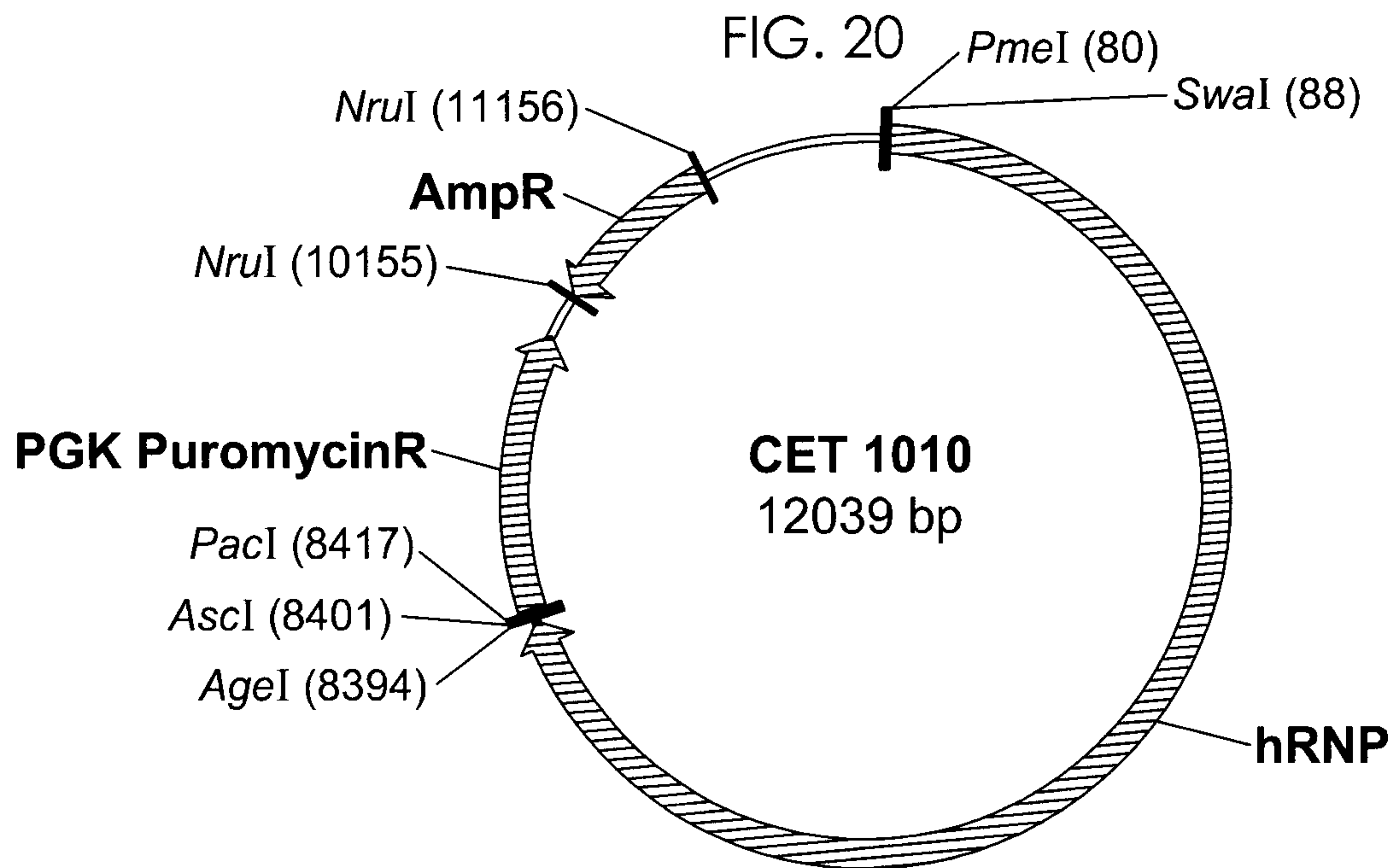
5701	ATGGGAAGTG	CCCGCCGCGA	CCGGGCTGGA	CCGCTAAAGT	AGCGCGTGAG
	TACCCTTCAC	GGGCGGCGCT	GGCCCGACCT	GGCGATTTC	TCGCGCACTC
5751	CGGGCCATCG	CTGGCCTTTC	GATGTGCGCG	GGCCTAGGGG	CTCGGTTGTG
	GCCCGGTAGC	GACCGGAAAG	CTACACGCGC	CCGGATCCCC	GAGCCAACAC
5801	TTCGCGGCGG	AACGTTTCTG	GGGCCCCCCC	GGCTTCCCGG	AGCGAGTCTG
	AAGCGCCGCC	TTGCAAAGAC	CCCGGGGGGG	CCGAAGGGCC	TCGCTCAGAC
5851	CGAAGCTAGC	TTCCCCTCCC	CCCTCTCCCG	GGAACCGGAT	TTGGCGGCCG
	GCTTCGATCG	AAGGGGAGGG	GGGAGAGGGC	CCTTGGCCTA	AACCGCCGGC
5901	CCATTTTCCC	GTCTCCTTCC	TCGCCACGAT	TTTGCTTTCA	ACGCTTTAGG
	GGTAAAAGGG	CAGAGGAAGG	AGCGGTGCTA	AAACGAAAGT	TGCGAAATCC
5951	TTTACTAGTT	TGGTTTTCTT	TTTTCAACCAC	TGCGTAGACG	TGTTTAGCGA
	AAATGATCAA	ACCAAAAGAA	AAAAGTGGTG	ACGCATCTGC	ACAAATCGCT
6001	TTTTCCTTTC	TTTTGGAAGT	CTTCATACCG	TTTCGAGGTG	GATTTAGCGT
	AAAAGGAAAG	AAAACCTTCA	GAAGTATGGC	AAAGCTCCAC	CTAAATCGCA
6051	TTTGAGCTTG	GGTCTTCAGC	GTCCTGCGCA	CCTCGCTAAA	GGCTCTCTGC
	AAACTCGAAC	CCAGAAGTCG	CAGGACGCGT	GGAGCGATTT	CCGAGAGACG
6101	CTTCCCCTCG	ACGAAATGGC	GCCATTGCTT	TCTGAAGCCA	CCGAGGCGCG
	GAAGGGGAGC	TGCTTTACCG	CGGTAACGAA	AGACTTCGGT	GGCTCCGCGC
6151	GGGTGGGGGC	GGGGTGGCGG	CGCTCCACGA	GCTTTACTGG	AACAGGCAGA
	CCCACCCCCG	CCCCACCGCC	GCGAGGTGCT	CGAAATGACC	TTGTCCGTCT
6201	GAGAACGTAG	TACAACCGAG	GCCTGGGCGG	GTGGCTGAAG	GCAGCGTCGC
	CTCTTGCAATC	ATGTTGGCTC	CGGACCCGCC	CACCGACTTC	CGTCGCAGCG
6251	TGCAAAGAGA	CCGTTTTTATT	TTTCATAATA	CGTAAGATTA	CGGGTGCTGT
	ACGTTTCTCT	GGCAAAATAA	AAAGTATTAT	GCATTCTAAT	GCCCACGACA
6301	AGTAAAGCAC	TTGAGCATTA	GTATAGTAGG	AGGAAGTCAA	AGTGGAAGAA
	TCATTTTCGTG	AACTCGTAAT	CATATCATCC	TCCTTCAGTT	TCACCTTTTT
6351	ATGGGAGCGC	TCATCAGGAA	GCTAGGGAGG	CTATGTTGAG	TGCAGGGTTA
	TACCCTCGCG	AGTAGTCCTT	CGATCCCTCC	GATACAACTC	ACGTCCCAAT
6401	CTTTCCTTTT	ATTGCAGAAC	TTTTATCTGC	TTAAAGGATC	CTCGGATCGA
	GAAAGGAAAA	TAACGTCTTG	AAAATAGACG	AATTTCTTAG	GAGCCTAGCT
6451	AATAATTCAA	ATTATAAGCA	TTTTTAAGGG	AATCTTCGAA	TTTGTGTTGA
	TTATTAAGTT	TAATATTCGT	AAAAATTCCC	TTAGAAGCTT	AAACAACCAT
6501	AAGTCAACGG	ATCCTTAGCA	CGTGGTGTTT	ACTTTAAGGA	AGTGAAATAG
	TTCAAGTTGCC	TAGGAATCGT	GCACCACAAG	TGAAATTCCT	TCACCTTATC
6551	CTGACTTTTC	ATAGTTAGCC	TTGCTTAAA	GCCTGGTTCA	GTGGACGAAA
	GACTGAAAAG	TATCAATCGG	AAGCGAATTT	CGGACCAAGT	CACCTGCTTT
6601	ATCCACGTCC	TGGCTATATA	AAAACCTAGT	TTGGGGTCAC	AGTGTGTTGAG
	TAGGTGCAGG	ACCGATATAT	TTTTGAATCA	AACCCAGTG	TCACAACTC
6651	CGTGGTCATT	CGGTTTTTTT	ATTTTTTTATT	TGTTTGAAAT	TATGATGCAT
	GCACCAGTAA	GCCAAAAAAA	TAAAAAATAA	ACAAACTTTA	ATACTACGTA
6701	CATTACACTG	ATAAGCATT	GCTTTCGAAT	TGAAAGGGGT	CTCCTTGGTT
	GTAATGTGAC	TATTCGTAAT	CGAAAGCTTA	ACTTTCCCCA	GAGGAACCAA
6751	ATTTTCTTTG	ACTCTAAGCA	CACTTATAAA	TAAAATAACC	TTGTTTATAA
	TAAAAGAAAC	TGAGATTCGT	GTGAATATTT	ATTTTATTGG	AACAAATATT
6801	TCGATAGTGG	ACGTCTGGTA	AGTTTGGAAG	AAACCCGAGG	TAAGTAAAGA
	AGCTATCACC	TGCAGACCAT	TCAAACCTTT	TTTGGGCTCC	ATTCATTTCT
6851	GCTTTTGCTT	TCGTTAGTGA	TATGAAAAAA	CAAGGTGTAT	TTAATACTTG
	CGAAAACGAA	AGCAATCACT	ATACTTTTTT	GTTCCACATA	AATTATGAAC
6901	CAACTTAGTT	TAAGGAAAGC	CAATTTACTG	ACATTTTAGT	AGAGCTACCA
	GTTGAATCAA	ATTCTTTTCG	GTTAAATGAC	TGTAAAATCA	TCTCGATGGT
6951	GAAACACTAT	TTGGAGTCCT	GATTAAGGCT	TTTGTAACCT	TTTTGACTAT
	CTTTGTGATA	AACCTCAGGA	CTAATTCCGA	AAACATTGAT	AAAACCTGATA
7001	TTAAAACAAT	TTTGGTCGTT	TTTATTAAAC	ATTTCAAAC	CTAAAAATTG
	AATTTTGTTA	AAACCAGCAA	AAATAATTTG	TAAAGTTTTG	GATTTTTAAC
7051	TAAACATTGG	CTTTTTGAGC	ACATTTTGGA	GAAACTTACA	AATTTAGGCT
	ATTTGTAACC	GAAAACTCG	TGTAAAACCT	CTTTGAATGT	TTAAATCCGA
7101	ATACAGTAAA	ATAACGGATT	TGTTTTATAA	TTTTGCTTTT	TCATTTTCGT
	TATGTCATTT	TATTGCCTAA	ACAAAATATT	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	AAAAAAATTT
7451	TCTAAAAAGT	AGTAAATTTT	GGACTTGGGT	TGTAGAATTT	AATGAATTAC
	AGATTTTTCA	TCATTTAAAA	CCTGAACCCA	ACATCTTAAA	TTACTTAATG
7501	AAAAGAATTC	TTTAATACCC	TTTAAATGAC	CTAAGAGCTG	GGTATGGTTT
	TTTTCTTAAG	AAATTATGGG	AAATTTACTG	GATTCTCGAC	CCATACCAAA
7551	TTCTGAATTT	TTTTGAAGAA	AATCTAAGAA	AGTTTACGTG	AATTAGAAGT
	AAGACTTAAA	AAAACCTTCT	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	AAAACTTTAG	AAACTGTTCC	TTTGGAGAGG
	CAAAAAAAG	AGGGGGGGTC	TTTTGAAATC	TTTGACAAGG	AAACCTCTCC
7751	AAAAAGGTAC	TCTACCAGCA	GGTCACCTCA	TATTTAAGAA	TTTAATTTCC
	TTTTTCCATG	AGATGGTCGT	CCAGTGGAGT	ATAAATTCTT	AAATTAAAGG
7801	TGCATACAAA	GAAAGTGTA	ATAAAAATTG	AAATGGTATT	TCCCTTTGCA
	ACGTATGTTT	CTTTCACATT	TATTTTTAAC	TTTACCATAA	AGGGAAACGT
7851	GAGAGAAAAG	GAACAGTTCC	GAAAGCTCTT	TATTGGTGGC	TTAAGCTT
	CTCTCTTTTC	CTTGTCAGG	CTTTCGAGAA	ATAACCACCG	AATTCGAA

FIG. 19(contd.)

40/93



41/93

FIG. 21 CET1010 nucleotide sequence

1	CGTTGTAAAA	CGACGGCCAG	TGAATTGTAA	TACGACTCAC	TATAGGGCGA
	GCAACATTTT	GCTGCCGGTC	ACTTAACATT	ATGCTGAGTG	ATATCCCGCT
51	ATTGGGTACC	GGGCCCCCCC	TCGAAGTTTA	AACATTTAAA	TCTAGAAGCT
	TAACCCATGG	CCCGGGGGGG	AGCTTCAAAT	TTGTAAATTT	AGATCTTCGA
101	TCAATGTTTT	TAGCACCCCTC	TGTGTGGAGG	AAAATAATGC	AGATTATTCT
	AGTTACAAAA	ATCGTGGGAG	ACACACCTCC	TTTTATTACG	TCTAATAAGA
151	AATTAGTGTA	ATATCTAACC	ACATTAAAAT	ATATTACATA	GTAAACTACA
	TTAATCACAT	TATAGATTGG	TGTAATTTTA	TATAATGTAT	CATTTGATGT
201	CTCCATAATT	TTATAAATTT	GACTCCCCAG	GGTAATAAAC	TAGTCTCTAG
	GAGGTATTAA	AATATTTAAA	CTGAGGGGTC	CCATTATTTG	ATCAGAGATC
251	TCTGCTCACC	TTCAACTGTA	CAATAAAGTC	TTGGTTCTTT	TGAAATAGAC
	AGACGAGTGG	AAGTTGACAT	GTTATTTTCAG	AACCAAGAAA	ACTTTATCTG
301	CTCAAATGAG	ACACCTAAAA	TTCAAAGTGT	CTTTACATTT	AAAGACACCT
	GAGTTTACTC	TGTGGATTTT	AAGTTTCACA	GAAATGTAAA	TTTCTGTGGA
351	ACAGGAAAGC	AGGTAAAAGA	GCCAGGTAA	AAACAAATTC	TAAAACCACT
	TGTCCTTTTCG	TCCATTTTCT	CGGTCCAATT	TTTGTTTAAG	ATTTTGGTGA
401	TAGCTGCAGT	TAAACATATA	GTAAAGATGC	ACTAAAGTTT	CTTACTCTGT
	ATCGACGTCA	ATTTGTATAT	CATTTCTACG	TGATTTCAAA	GAATGAGACA
451	AAATCCCTTC	CACTTCAGGA	AATATTCCAC	TTTCCCATT	ACTACACGTC
	TTTAGGGAAG	GTGAAGTCCT	TTATAAGGTG	AAAGGGTAAG	TGATGTGCAG
501	GATCTAGTAC	TTTTTCCACG	ACAAATTCTT	CAGGCTCTGC	CTCTTCAACT
	CTAGATCATG	AAAAAGGTGC	TGTTTAAGAA	GTCCGAGACG	GAGAAGTTGA
551	TTTTTACTCT	TTCCATTCTG	TTTTTTTCCC	ATTTTTTGCT	AAAATAAAAC
	AAAAATGAGA	AAGGTAAGAC	AAAAAAAGGG	TAAAAAACGA	TTTTATTTTG
601	AAAAGAGAAA	TTAAGAAATA	TTCCTCTTGA	ATTTTGAGCA	CATTTTCAAG
	TTTTCTCTTT	AATTCTTTAT	AAGGAGAACT	TAAAACTCGT	GTAAAAGTTC
651	GCTCAATTGC	TTATATTATT	ATCACATTCT	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	TGTTATTTTT	TTGTGATAGA	CCTTTGTCAA	TAAAAAAAGG
851	AGTCTTTTTT	TTTGAGACAG	GGTCTCACAC	TCTTGTCCGC	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	TCAAACCTCT	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	AACAACCTATG	CAGGAAGTAC	CTAAAGGGTG	ATCCAGAGAA
	GTGACGTTGT	TTGTTGATAC	GTCCTTCATG	GATTTCCAC	TAGGTCTCTT
1251	GCAAGTAGTA	GTGACAGGTC	TTAGGTGAAC	CTATGACAGA	CCTTGTATCC
	CGTTCATCAT	CACTGTCCAG	AATCCACTTG	GATACTGTCT	GGAACATAGG
1301	ACCCCCAGAT	GGTAAAAGCC	CCAGCCCCCT	TCTCAATTCA	AATATTAATG
	TGGGGGTCTA	CCATTTTTCGG	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
	CAAGTTTTAG	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	AAAAAAATCC	TTCCTTCCTG
	TAATGTTTAA	ATAAAATTAA	GGGATTCCAC	TTTTTTTAGG	AAGGAAGGAC
1601	GTTTCTCAAG	AGAAAGTCTA	TACTGGTAAC	CAAATTCACT	TTAAACAGGC
	CAAAGAGTTC	TCTTTCAGAT	ATGACCATTG	GTTTAAGTGA	AATTTGTCCG
1651	ATTTTCTTTG	GTATGACACT	ATTTAAGAGA	AGCAGGAAAC	CAACGTGAAC
	TAAAAGAAAC	CATACTGTGA	TAAATTCTCT	TCGTCCTTTG	GTTGCACTTG
1701	CAGCTCTTTC	CAATGGCTCA	AGATTTCCTA	TGAGAGGACT	AAAAATGGGG
	GTCGAGAAAG	GTTACCGAGT	TCTAAAGGAT	ACTCTCCTGA	TTTTTACCCC
1751	AAAATTTTTA	TGAGAGGATT	AAAAATGGGG	GAAAAAAAC	CCTGAAATGG
	TTTTTAAAAAT	ACTCTCCTAA	TTTTTACCCC	CTTTTTTTTG	GGACTTTACC
1801	TTAATCAGAA	GATCCTATGG	GCTGAGAAGG	AATCCATCTT	AACATTTTCAT
	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	GGGTTCGATGA	ACCCTCCGAC
2051	AGGCAGGAGA	ACTGCTTGAA	CCCAGGAGGC	TTAAGTTGCG	GTGAGCCAAG
	TCCGTCCTCT	TGACGAACTT	GGGTCCCTCCG	AATTCAACGC	CACTCGGTTC
2101	ATCACGCCAT	TGCACTCTAG	CCTGGACAAC	AAGAGAAAAA	CTCTGTCTCA
	TAGTGCGGTA	ACGTGAGATC	GGACCTGTTG	TTCTCTTTTTT	GAGACAGAGT
2151	AAAAAACACA	AAAACAAAAA	ACCCAAATAC	TATTTAAAAA	AGATAAACCT
	TTTTTTGTGT	TTTTGTTTTTT	TGGGTTTATG	ATAAATTTTTT	TCTATTTGGA
2201	TAATTGCTCA	ATCATTAAG	CCATCCCACA	AGTAAAGCAG	CAAGCAGAAA
	ATTAACGAGT	TAGTAATTTT	GGTAGGGTGT	TCATTTTCGTC	GTTTCGTCTTT
2251	AAAGTTAAGA	ACACCTCAAG	GCTACAGAAG	GACATTTCAA	GCTATGCAGG
	TTTCAATTCT	TGTGGAGTTC	CGATGTCTTC	CTGTAAAGTT	CGATACGTCC
2301	CATATGAAGT	GTGCAGACAG	ATATGTAAGA	AAGGCCTCAA	GACTGCAAAA
	GTATACTTCA	CACGTCTGTC	TATACATTCT	TTCCGGAGTT	CTGACGTTTTT
2351	GGGCATTTCA	AGCTATGCAA	GCATATAGGT	AACACATACA	CACACACAAA
	CCCGTAAAGT	TCGATACGTT	CGTATATCCA	TTGTGTATGT	GTGTGTGTTT
2401	ATAAAATCCC	CTGAAATACA	AAAACATGCA	GCAAACACCT	GACGTTTTTTG
	TATTTTAGGG	GACTTTATGT	TTTTGTACGT	CGTTTGTGGA	CTGCAAAAAC
2451	GATACCATTT	CTAAGTCAGG	TGTTATGATT	CTCATTAGTC	AAGATACTTG
	CTATGGTAAA	GATTCAGTCC	ACAATACTAA	GAGTAATCAG	TTCTATGAAC
2501	AGTACTGGGC	CCAAACAGCT	TTCTGCCACT	GTACAGTACA	AGAAGGTAGG
	TCATGACCCG	GGTTTGTGCA	AAGACGGTGA	CATGTCATGT	TCTTCCATCC
2551	AATAATGGTG	GGAGGAGCAA	AGACAAACTG	TAATAGACAG	AAGTGTATCA
	TTATTACCAC	CCTCCTCGTT	TCTGTTTGAC	ATTATCTGTC	TTCACATAGT
2601	GATACCTATA	CTACATGAAA	AACAAAACAG	CTACTGCCAC	AAAGGGAGAA
	CTATGGATAT	GATGTACTTT	TTGTTTTGTC	GATGACGGTG	TTTCCCTCTT
2651	GGCTAACAAA	ATAAAGTCAA	CAATAAATAC	AGAAAATGAA	AAGGATACAC
	CCGATTGTTT	TATTTTCAGTT	GTTATTTATG	TCTTTTACTT	TTCTTATGTG
2701	ACTAAGGTTT	ACAAAAAAA	AAAGGCAGAC	AAAATGCCAT	ACAGTATTCA
	TGATTCCAAA	TGTTTTTTTTT	TTTCCGTCTG	TTTTACGGTA	TGTCATAAGT
2751	TTCACTACTA	TGGCATTTCAT	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	ATTTCCCAT	ATTTAAGTGA	AATTAACAGC	ACACATATCT
	CAGTTTTATT	TAAAGGGTAA	TAAATTCAC	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	GTCCCATTTG	TCTGAATTAC	GTCTTAGATT
3101	TGATGGCACA	AATTAAAAAT	CACTAACGTG	CCCTTTTTAG	TGTGAAACCC
	ACTACCGTGT	TTAATTTTTA	GTGATTGCAC	GGGAAAATC	ACACTTTGGG
3151	AGAGAGAGCA	CATACAAGCC	AAAAACAAAT	GCTTTATTTT	ACCTAGGAGA
	TCTCTCTCGT	GTATGTTTCG	TTTTTGTTTA	CGAAATAAAA	TGGATCCTCT
3201	CATTAACATT	CACCTTTACG	TGTTTAAGAT	TAATGCAATG	TTAAATATTG
	GTAATTGTAA	GTGGAAATGC	ACAAATTCTA	ATTACGTTAC	AATTTATAAC
3251	TGAAAACCTG	AACTTTGAAT	TCATGATTTT	TTATGTGAAT	ATTCCAGGGT
	ACTTTTGACA	TTGAAACTTA	AAGTACTAAA	AATACACTTA	TAAGGTCCCA
3301	TTAAAAAAAC	TTGTAACATG	ACATGGCTGA	ATAAGATAAA	AAAAAAATCT
	AATTTTTTTG	AACATTGTAC	TGTACCGACT	TATTCATTTT	TTTTTTTTAGA
3351	AGCCTTTTCT	CCCTTCTGGC	TCATATTTGC	GATTTTCGATC	ATTTTGTTTA
	TCGGAAAAGA	GGGAAGACCG	AGTATAAACG	CTAAAGCTAG	TAAAACAAAT
3401	AAAAACAAAA	CACTGCAATG	AATTAAACTT	AATATTCTTC	TATGTTTTAG
	TTTTTGTTTT	GTGACGTTAC	TTAATTTGAA	TTATAAGAAG	ATACAAAATC
3451	AGTAAGTTAA	AACAAGATAA	AGTGACCAAA	GTAATTTGAA	AGATTCAATG
	TCATTCAATT	TTGTTCTATT	TCACTGGTTT	CATTAAACTT	TCTAAGTTAC
3501	ACTTTTGCTC	CAACCTAGGT	GCACAAGGTA	CCTTGTTCTT	TAAATTGGGC
	TGAAAACGAG	GTTGGATCCA	CGTGTTCCAT	GGAACAAGAA	ATTTAACCCG
3551	TTTAATGAAA	ATACTTCTCC	AGAATTCTGG	GGATTTAAGA	AAAATTATGC
	AAATTACTTT	TATGAAGAGG	TCTTAAGACC	CCTAAATTCT	TTTTAATACG
3601	CAACCAACAA	GGGCTTTACC	ATTTTATGTA	ACATTTTTCA	ACGCTGCAAA
	GTTGGTTGTT	CCCGAAATGG	TAAAATACAT	TGTAAAAAGT	TGCGACGTTT
3651	AATGTGTGTA	TTTCTATTTG	AAGATAAAAA	TCCTCAGCAA	AATCCACATT
	TTACACACAT	AAAGATAAAC	TTCTATTTTT	AGGAGTCGTT	TTAGGTGTAA
3701	GCACTGTCCT	TCAAAGATTA	GCCTTCTTTG	AACTAGTTAA	GACACTATTA
	CGTGACAGGA	AGTTTCTAAT	CGGAAGAAAC	TTGATCAATT	CTGTGATAAT
3751	AGCCAAGCCA	GTATCTCCCT	GTAATGAATT	CGTTTTTCTC	TTAATTTTCC
	TCGGTTCGGT	CATAGAGGGA	CATTACTTAA	GCAAAAAGAG	AATTAAAAGG
3801	CCTGTAATTT	ACACTGGGAG	AGCTGGGAAA	TATGTGGATG	TAAATTTCTC
	GGACATTAAA	TGTGACCCTC	TCGACCCTTT	ATACACCTAC	ATTTAAAGAG
3851	AGCCACAGAG	ATGCAAAGTT	ATACTGTGGG	GAAAAAAAC	TTGAGTTAAA
	TCGGTGTCTC	TACGTTTCAA	TATGACACCC	CTTTTTTTTG	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	AAAGGTAACT	AAATACAAAT
	CTGTCAATTTC	CATTCAAACG	TACTGTGTTT	TTTCCATTGA	TTTATGTTTA
4051	TCTGTTTGGA	TTCCAACCCC	CAAGTAGAGA	GCGCACACTT	TCAAACGTGA
	AGACAAACCT	AAGGTTGGGG	GTTTCATCTCT	CGCGTGTGAA	AGTTTGCACT
4101	ATACAAATCC	AGAGTAGATC	TGCGCTCCTA	CCTACATTGC	TTATGATGTA
	TATGTTTAGG	TCTCATCTAG	ACGCGAGGAT	GGATGTAACG	AATACTACAT
4151	CTTAAGTACG	TGTCCTAACC	ATGTGAGTCT	AGAAAGACTT	TACTGGGGAT
	GAATTCATGC	ACAGGATTGG	TACACTCAGA	TCTTTCTGAA	ATGACCCCTA
4201	CCTGGTACCT	AAAACAGCTT	CACATGGCTT	AAAATAGGGG	ACCAATGTCT
	GGACCATGGA	TTTTGTCGAA	GTGTACCGAA	TTTTATCCCC	TGGTTACAGA

FIG. 21(contd.) 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	TTAATTTATT	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
	CTCCGGATTC	GCGCCACGCG	GCGGGGGTGG	GCCCTCGCGC	CCGGAGCACC
4601	TCAGCGCATC	CGCGGGGAGA	AACAAAGGCC	GCGGCACGGG	GGCTCAAGGG
	AGTCGCGTAG	GCGCCCCCTCT	TTGTTTCCGG	CGCCGTGCCC	CCGAGTTCCC
4651	CACTGCGCCA	CACCGCACGC	GCCTACCCCC	GCGCGGCCAC	GTTAACTGGC
	GTGACGCGGT	GTGGCGTGCG	CGGATGGGGG	GCGCGCGGTG	CAATTGACCG
4701	GGTCGCCGCA	GCCTCGGGAC	AGCCGGCCGC	GCGCCGCCAG	GCTCGCGGAC
	CCAGCGGCGT	CGGAGCCCTG	TCGGCCGGCG	CGCGGCGGTC	CGAGCGCCTG
4751	GCGGGACCAC	GCGCCGCCCT	CCGGGAGGCC	CAAGTCTCGA	CCCAGCCCCG
	CGCCCTGGTG	CGCGGCGGGA	GGCCCTCCGG	GTTCAAGAGCT	GGGTCGGGGC
4801	CGTGGCGCTG	GGGGAGGGGG	CGCCTCCGCC	GGAACGCGGG	TGGGGGAGGG
	GCACCGCGAC	CCCCTCCCCC	GCGGAGGCGG	CCTTGCGCCC	ACCCCTCCC
4851	GAGGGGGAAA	TGCGCTTTGT	CTCGAAATGG	GGCAACCGTC	GCCACAGCTC
	CTCCCCCTTT	ACGCGAAACA	GAGCTTTACC	CCGTTGGCAG	CGGTGTCGAG
4901	CCTACCCCCT	CGAGGGCAGA	GCAGTCCCCC	CACTAACTAC	CGGGCTGGCC
	GGATGGGGGA	GCTCCCGTCT	CGTCAGGGGG	GTGATTGATG	GCCCGACCGG
4951	GCGCGCCAGG	CCAGCCGCGA	GGCCACCGCC	CGACCCTCCA	CTCCTTCCCG
	CGCGCGGTCC	GGTCGGCGCT	CCGGTGGCGG	GCTGGGAGGT	GAGGAAGGGC
5001	CAGCTCCCCG	CGCGGGGTCC	GGCGAGAAGG	GGAGGGGAGG	GGAGCGGAGA
	GTCGAGGGCC	GCGCCCCAGG	CCGCTCTTCC	CCTCCCCCTC	CCTCGCCTCT
5051	ACCGGGCCCC	CGGGACGCGT	GTGGCATCTG	AAGCACCACC	AGCGAGCGAG
	TGGCCCCGGG	GCCCTGCGCA	CACCGTAGAC	TTCGTGGTGG	TCGCTCGCTC
5101	AGCTAGAGAG	AAGGAAAGCC	ACCGACTTCA	CCGCCTCCGA	GCTGCTCCGG
	TCGATCTCTC	TTCTTTTCGG	TGGCTGAAGT	GGCGGAGGCT	CGACGAGGCC
5151	GTCGCGGGTC	TGCAGCGTCT	CCGGCCCTCC	GCGCCTACAG	CTCAAGCCAC
	CAGCGCCCAAG	ACGTCGCAGA	GGCCGGGAGG	CGCGGATGTC	GAGTTCGGTG
5201	ATCCGAAGGG	GGAGGGAGCC	GGGAGCTGCG	CGCGGGGCCG	CCGGGGGGAG
	TAGGCTTCCC	CCTCCCTCGG	CCCTCGACGC	GCGCCCCGGC	GGCCCCCTC
5251	GGGTGGCACC	GCCCACGCCG	GGCGGCCACG	AAGGGCGGGG	CAGCGGGCGC
	CCCACCGTGG	CGGGTGCGGC	CCGCCGGTGC	TTCCCGCCCC	GTCGCCCCGC
5301	GCGCGCGGCG	GGGGGAGGGG	CCGGCGCCGC	GCCCGCTGGG	AATTGGGGCC
	CGCGCGCCGC	CCCCCTCCCC	GGCCGCGGCG	CGGGCGACCC	TTAACCCCGG
5351	CTAGGGGGAG	GGCGGAGGCG	CCGACGACCG	CGGCACTTAC	CGTTCGCGGC
	GATCCCCCTC	CCGCCTCCGC	GGCTGCTGGC	GCCGTGAATG	GCAAGCGCCG
5401	GTGGCGCCCG	GTGGTCCCCA	AGGGGAGGGA	AGGGGGAGGC	GGGGCGAGGA
	CACCGCGGGC	CACCAGGGGT	TCCCCTCCCT	TCCCCCTCCG	CCCCGCTCCT
5451	CAGTGACCGG	AGTCTCCTCA	GCGGTGGCTT	TTCTGCTTGG	CAGCCTCAGC
	GTCACTGGCC	TCAGAGGAGT	CGCCACCGAA	AAGACGAACC	GTCGGAGTCG
5501	GGCTGGCGCC	AAAACCGGAC	TCCGCCCACT	TCCTCGCCCC	CCGGTGCAGG
	CCGACCGCGG	TTTTGGCCTG	AGGCGGGTGA	AGGAGCGGGC	GGCCACGCTC
5551	GGTGTGGAAT	CCTCCAGACG	CTGGGGGAGG	GGGAGTTGGG	AGCTTAAAAA
	CCACACCTTA	GGAGGTCTGC	GACCCCTCC	CCCTCAACCC	TCGAATTTTT
5601	CTAGTACCQC	TTTGGGACCA	CTTTCAGCAG	CGAACTCTCC	TGTACACCAG
	GATCATGGGG	AAACCCTGGT	GAAAGTCGTC	GCTTGAGAGG	ACATGTGGTC
5651	GGGTCAGTTC	CACAGACGCG	GGCCAGGGGT	GGGTCATTGC	GGCGTGAACA
	CCCAGTCAAG	GTGTCTGCGC	CCGGTCCCCA	CCCAGTAACG	CCGCACTTGT

FIG. 21(contd.) 45/93

5701	ATAATTTGAC	TAGAAGTTGA	TTCGGGTGTT	TCCGGAAGGG	GCCGAGTCAA
	TATTAAACTG	ATCTTCAACT	AAGCCCACAA	AGGCCTTCCC	CGGCTCAGTT
5751	TCCGCCGAGT	TGGGGCACGG	AAAACAAAAA	GGGAAGGCTA	CTAAGATTTT
	AGGCGGCTCA	ACCCCGTGCC	TTTTGTTTTT	CCCTTCCGAT	GATTCTAAAA
5801	TCTGGCGGGG	GTTATCATTG	GCGTAACTGC	AGGGACCACC	TCCCGGGTTG
	AGACCGCCCC	CAATAGTAAC	CGCATTGACG	TCCCTGGTGG	AGGGCCCAAC
5851	AGGGGGCTGG	ATCTCCAGGC	TGCGGATTAA	GCCCCTCCCG	TCGGCGTTAA
	TCCCCCGACC	TAGAGGTCCG	ACGCCTAATT	CGGGGAGGGC	AGCCGCAATT
5901	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
	GTAAC TCGC	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	GCCGCCGTCTG	CCGAGATCGC
6201	GCAGTAGCAG	CAGCGCCGGG	TCCCGTGCGG	AGGTGCTCCT	CGCAGAGTTG
	CGTCATCGTC	GTGCGGGCCC	AGGGCACGCC	TCCACGAGGA	GCGTCTCAAC
6251	TTTCTCGAGC	AGCGGCAGTT	CTCACTACAG	CGCCAGGACG	AGTCCGGTTC
	AAAGAGCTCG	TCGCCGTCAA	GAGTGATGTC	GCGGTCTTGC	TCAGGCCAAG
6301	GTGTTCGTCC	GCGGAGATCT	CTCTCATCTC	GCTCGGCTGC	GGGAAATCGG
	CACAAGCAGG	CGCCTCTAGA	GAGAGTAGAG	CGAGCCGACG	CCCTTTAGCC
6351	GCTGAAGCGA	CTGAGTCCGC	GATGGAGGTA	ACGGGTTTGA	AATCAATGAG
	CGACTTCGCT	GACTCAGGCG	CTACCTCCAT	TGCCCAAACCT	TTAGTTACTC
6401	TTATTGAAAA	GGGCATGGCG	AGGCCGTTGG	CGCCTCAGTG	GAAGTCGGCC
	AATAACTTTT	CCCGTACCGC	TCCGGCAACC	GCGGAGTCAC	CTTCAGCCGG
6451	AGCCGCCTCC	GTGGGAGAGA	GGCAGGAAAT	CGGACCAATT	CAGTAGCAGT
	TCGGCGGAGG	CACCCTCTCT	CCGTCCTTTA	GCCTGGTTAA	GTCTATCGTCA
6501	GGGGCTTAAG	GTTTATGAAC	GGGGTCTTGA	GCGGAGGCCT	GAGCGTACAA
	CCCCGAATTC	CAAATACTTG	CCCCAGAACT	CGCCTCCGGA	CTCGCATGTT
6551	ACAGCTTCCC	CACCCTCAGC	CTCCCGGCGC	CATTTCCCTT	CACTGGGGGT
	TGTCGAAGGG	GTGGGAGTCG	GAGGGCCGCG	GTAAAGGGAA	GTGACCCCCA
6601	GGGGGATGGG	GAGCTTTCAC	ATGGCGGACG	CTGCCCCGCT	GGGGTGAAAG
	CCCCCTACCC	CTCGAAAGTG	TACCGCCTGC	GACGGGGCGA	CCCCACTTTC
6651	TGGGGCGCGG	AGGCGGGAAT	TCTTATTCCC	TTTCTAAAGC	ACGCTGCTTC
	ACCCCGCGCC	TCCGCCCTTA	AGAATAAGGG	AAAGATTTTCG	TGCGACGAAG
6701	GGGGGCCACG	GCGTCTCCTC	GGCGAGCGTT	TCGGCGGGCA	GCAGGTCCTC
	CCCCCGGTGC	CGCAGAGGAG	CCGCTCGCAA	AGCCGCCCGT	CGTCCAGGAG
6751	GTGAGCGAGG	CTGCGGAGCT	TCCCCCTCCC	CTCTCTCCC	GGAACCGATT
	CACTCGCTCC	GACGCCTCGA	AGGGGAGGGG	GAGAGAGGGC	CCTTGGCTAA
6801	TGGCGGCCGC	CATTTTTCATG	GCTCGCCTTC	CTCTCAGCGT	TTTCCTTATA
	ACCGCCGGCG	GTAAAAGTAC	CGAGCGGAAG	GAGAGTCGCA	AAAGGAATAT
6851	ACTCTTTTAT	TTTCTTAGTG	TGCTTTCTCT	ATCAAGAAGT	AGAAGTGGTT
	TGAGAAAATA	AAAGAATCAC	ACGAAAGAGA	TAGTTCTTCA	TCTTCACCAA
6901	AACTATTTTT	TTTTTCTTCT	CGGGCTGTTT	TCATATCGTT	TCGAGGTGGA
	TTGATAAAAA	AAAAAGAAGA	GCCCGACAAA	AGTATAGCAA	AGCTCCACCT
6951	TTTGGAGTGT	TTTGTGAGCT	TGGATCTTTA	GAGTCCTGCG	CACCTCATTA
	AAACCTCACA	AAACACTCGA	ACCTAGAAAT	CTCAGGACGC	GTGGAGTAAT
7001	AAGGCGCTCA	GCCTTCCCCCT	CGATGAAATG	GCGCCATTGC	GTTCCGAAGC
	TTCCGCGAGT	CGGAAGGGGA	GCTACTTTAC	CGCGGTAACG	CAAGCCTTCG
7051	CACACCGAAG	AGCGGGGAGG	GGGGGTGCTC	CGGGTTTGCG	GGCCCGGTTT
	GTGTGGCTTC	TCGCCCCCTCC	CCCCCACGAG	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	GGGATATTCT	CGTCAGAATC	CTTTACTGTT
	CACCAAAAGA	CCCGCCCCCT	CCCTATAAGC	GCAGTCTTAG	GAAATGACAA
7351	CTTAAGGATT	CCGTTTAAGT	TGTAGAGCTG	ACTCATTTTA	AGTAATGTTG
	GAATTCCTAA	GGCAAATTCA	ACATCTCGAC	TGAGTAAAAT	TCATTACAAC
7401	TTACTGAGAA	GTTTAACCCT	TACGGGACAG	ATCCATGGAC	CTTTATAGAT
	AATGACTCTT	CAAATTGGGA	ATGCCCTGTC	TAGGTACCTG	GAAATATCTA
7451	GATTACGAGG	AAAGTGAAAT	AACGATTTTG	TCCTTAGTTA	TACTTCGATT
	CTAATGCTCC	TTTCACTTTA	TTGCTAAAAC	AGGAATCAAT	ATGAAGCTAA
7501	AAAACATGGC	TTCAGAGGCT	CCTTCCTGTA	ATGCGTATGG	ATTGATGTGC
	TTTTGTACCG	AAGTCTCCGA	GGAAGGACAT	TACGCATACC	TAACCTACACG
7551	AAAACGTGTT	TGGGCCTGGG	CCGCTCTGTA	TTTGAACCTT	GTTACTTTTC
	TTTTGACAAA	ACCCGGACCC	GGCGAGACAT	AAACTTGAAA	CAATGAAAAG
7601	TCATTTTGTT	TGCAATCTTG	GTTGAACATT	ACATTGATAA	GCATAAGGTC
	AGTAAAACAA	ACGTTAGAAC	CAACTTGTA	TGTAACCTATT	CGTATTCCAG
7651	TCAAGCGAAG	GGGGTCTACC	TGGTTATTTT	TCTTTGACCC	TAAGCACGTT
	AGTTCGCTTC	CCCCAGATGG	ACCAATAAAA	AGAAACTGGG	ATTCGTGCAA
7701	TATAAAATAA	CATTGTTTAA	AATCGATAGT	GGACATCGGG	TAAGTTTGGA
	ATATTTTATT	GTAACAAATT	TTAGCTATCA	CCTGTAGCCC	ATTCAAACCT
7751	TAAATTGTGA	GGTAAGTAAT	GAGTTTTTGC	TTTTTGTTAG	TGATTTGTAA
	ATTTAACACT	CCATTCATTA	CTCAAAAACG	AAAAACAATC	ACTAAACATT
7801	AACTTGTTAT	AAATGTACAT	TATCCGTAAT	TTCAAGTTAG	AGATAACCTA
	TTGAACAATA	TTTACATGTA	ATAGGCATTA	AAGTCAAATC	TCTATTGGAT
7851	TGTGCTGACG	ACAATTAAGA	ATAAAAACCTA	GCTGAAAAAA	TGAAAATAAC
	ACACGACTGC	TGTTAATTCT	TATTTTTGAT	CGACTTTTTT	ACTTTTATTG
7901	TATCGTGACA	AGTAACCATT	TCAAAAGACT	GCTTTGTGTC	TCATAGGAGC
	ATAGCACTGT	TCATTGGTAA	AGTTTTCTGA	CGAAACACAG	AGTATCCTCG
7951	TAGTTTGATC	ATTTCAAGTTA	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
	AACTGAAAAA	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	AAAGATTTCA	GCTCTGAGGT	TTTTCTTATT	GACTTGGA	ATTGTTTTGT
	TTTCTAAAGT	CGAGACTCCA	AAAAGGATAA	CTGAACCTTT	TAACAAAACA
8351	GCCTGTCGCT	TGCCACATGG	CCAATCAAGT	AAGCTTATCG	ATACCGGTGG
	CGGACAGCGA	ACGGTGTACC	GGTTAGTTCA	TTCAATAGC	TATGGCCACC
8401	CGCGCCAATT	GTTAATTAAG	ATCTGGCCCA	ATGGGCCGTA	CGAATTCCTT
	GCGCGGTTAA	CAATTAATTC	TAGACCGGGT	TACCCGGCAT	GCTTAAGGAA
8451	AGGCTACCGG	GTAGGGGAGG	CGCTTTTCCC	AAGGCAGTCT	GGAGCATGCG
	TCCGATGGCC	CATCCCCCTC	GCGAAAAGGG	TTCCGTCAGA	CCTCGTACGC
8501	CTTTAGCAGC	CCCGCTGGGC	ACTTGGCGCT	ACACAAGTGG	CCTCTGGCCT
	GAAATCGTGG	GGGCGACCCG	TGAACCGCGA	TGTGTTTACC	GGAGACCGGA
8551	CGCACACATT	CCACATCCAC	CGGCCGGTAG	GCGCCAACCG	GCTCCGTTCT
	GCGTGTGTAA	GGTGTAGGTG	GCCGGCCATC	CGCGGTGGC	CGAGGCAAGA

FIG. 21(contd.) 47/93

8601	TTGGTGGCCC	CTTCGCGCCA	CCTTCTACTC	CTCCCCTAGT	CAGGAAGTTC
	AACCACCGGG	GAAGCGCGGT	GGAAGATGAG	GAGGGGATCA	GTCCTTCAAG
8651	CCCCCGCCC	CGCAGCTCGC	GTCGTGCAGG	ACGTGACAAA	TGGAAGTAGC
	GGGGGGCGGG	GCGTCGAGCG	CAGCACGTCC	TGCACTGTTT	ACCTTCATCG
8701	ACGTCTCACT	AGTCTCGTGC	AGATGGACAG	CACCGCTGAG	CAATGGAAGC
	TGCAGAGTGA	TCAGAGCACG	TCTACCTGTC	GTGGCGACTC	GTTACCTTCG
8751	GGGTAGGCCT	TTGGGGCAGC	GGCCAATAGC	AGCTTTGCTC	CTTCGCTTTC
	CCCATCCGGA	AACCCCGTCG	CCGGTTATCG	TCGAAACGAG	GAAGCGAAAG
8801	TGGGCTCAGA	GGCTGGGAAG	GGGTGGGTCC	GGGGGCGGGC	TCAGGGGCGG
	ACCCGAGTCT	CCGACCCTTC	CCCACCCAGG	CCCCCGCCCG	AGTCCCCGCC
8851	GCTCAGGGGC	GGGGCGGGCG	CCCGAAGGTC	CTCCGGAGGC	CCGGCATTTCT
	CGAGTCCCCG	CCCCGCCCCG	GGGCTTCCAG	GAGGCCTCCG	GGCCGTAAGA
8901	GCACGCTTCA	AAAGCGCACG	TCTGCCGCGC	TGTTCTCCTC	TTCTTCATCT
	CGTGCGAAGT	TTTCGCGTGC	AGACGGCGCG	ACAAGAGGAG	AAGGAGTAGA
8951	CCGGGCCTTT	CGACCAGCTT	ACCATGACCG	AGTACAAGCC	CACGGTGCGC
	GGCCCCGAAA	GCTGGTCGAA	TGGTACTGGC	TCATGTTTCG	GTGCCACGCG
9001	CTCGCCACCC	GCGACGACGT	CCCCAGGGCC	GTACGCACCC	TCGCCGCCGC
	GAGCGGTGGG	CGCTGCTGCA	GGGGTCCCGG	CATGCGTGGG	AGCGGCGGGC
9051	GTTCGCCGAC	TACCCCGCCA	CGCGCCACAC	CGTCGATCCG	GACCGCCACA
	CAAGCGGCTG	ATGGGGCGGT	GCGCGGTGTG	GCAGCTAGGC	CTGGCGGTGT
9101	TCGAGCGGGT	CACCGAGCTG	CAAGAACTCT	TCCTCACGCG	CGTCGGGCTC
	AGCTCGCCCA	GTGGCTCGAC	GTTCTTGAGA	AGGAGTGCGC	GCAGCCCGAG
9151	GACATCGGCA	AGGTGTGGGT	CGCGGACGAC	GGCGCCGCGG	TGGCGGTCTG
	CTGTAGCCGT	TCCACACCCA	GCGCCTGCTG	CCGCGGCGCC	ACCGCCAGAC
9201	GACCACGCCG	GAGAGCGTCG	AAGCGGGGGC	GGTGTTTCGC	GAGATCGGCC
	CTGGTGCGGC	CTCTCGCAGC	TTCGCCCCCG	CCACAAGCGG	CTCTAGCCGG
9251	CGCGCATGGC	CGAGTTGAGC	GGTTCCCGGC	TGGCCGCGCA	GCAACAGATG
	GCGCGTACCG	GCTCAACTCG	CCAAGGGCCG	ACCGGCGCGT	CGTTGTCTAC
9301	GAAGGCCTCC	TGGCGCCGCA	CCGGCCCAAG	GAGCCCGCGT	GGTTCCTGGC
	CTTCCGGAGG	ACCGCGGCGT	GGCCGGGTTC	CTCGGGCGCA	CCAAGGACCG
9351	CACCGTCGGC	GTCTCGCCCC	ACCACCAGGG	CAAGGGTCTG	GGCAGCGCCG
	GTGGCAGCCG	CAGAGCGGGC	TGGTGGTCCC	GTTCCTCAGC	CCGTGCGGGC
9401	TCGTGCTCCC	CGGAGTGGAG	GCGGCCGAGC	GCGCCGGGGT	GCCCCCCTTC
	AGCACGAGGG	GCCTCACCTC	CGCCGGCTCG	CGCGGCCCCA	CGGGCGGAAG
9451	CTGGAGACCT	CCGCGCCCCG	CAACCTCCCC	TTCTACGAGC	GGCTCGGCTT
	GACCTCTGGA	GGCGCGGGGC	GTTGGAGGGG	AAGATGCTCG	CCGAGCCGAA
9501	CACCGTCACC	GCCGACGTCG	AGGTGCCCGA	AGGACCGCGC	ACCTGGTGCA
	GTGGCAGTGG	CGGCTGCAGC	TCCACGGGCT	TCCTGGCGCG	TGGACCACGT
9551	TGACCCGCAA	GCCCCGTGCC	TGACGCCCGC	CCCACGACCC	GCAGCGCCCC
	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	CGTCCTGTCT	TTCCCCCCCC	TAACCCYTCT	GTTATCGTCC
9851	CATGCTGGGG	GGGCGGTGGG	GGCTATGGCT	TCTGAGGCGG	AAAGAACCAG
	GTACGACCCC	CCCGCCACCC	CCGATACCGA	AGACTCCGCC	TTTCTTGGTC
9901	CTGGGGCTCG	AGGGCCGCCA	CCGCGGTGGA	GCTCCAGCTT	TTGTTCCCTT
	GACCCCGAGC	TCCCGGCGGT	GGCGCCACCT	CGAGGTGCAA	AACAAGGGAA
9951	TAGTGAGGGT	TAATTTTCGAG	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	GGTGCCCTAAT	GAGTGAGCTA	ACTCACATTA
	CGTATTTTCAC	ATTTCCGGACC	CCACGGATTA	CTCACTCGAT	TGAGTGTAAT
10101	ATTGCGTTGC	GCTCACTGCC	CGCTTTCCAG	TCGGGAAACC	TGTCGTGCCA
	TAACGCAACG	CGAGTGACGG	GCGAAAGGTC	AGCCCTTTGG	ACAGCACGGT
10151	GCATCGCGAG	CACTTTTTCGG	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	CTTTTTCCCT	CTCATACTCA	TAAGTTGTAA
10301	TCCGTGTCGC	CCTTATTCCC	TTTTTTGCGG	CATTTTGCCCT	TCCTGTTTTT
	AGGCACAGCG	GGAATAAGGG	AAAAAACGCC	GTAAAACGGA	AGGACAAAAA
10351	GCTCACCCAG	AAACGCTGGT	GAAAGTAAAA	GATGCTGAAG	ATCAGTTGGG
	CGAGTGGGTC	TTTGCGACCA	CTTTCATTTT	CTACGACTTC	TAGTCAACCC
10401	TGCACGAGTG	GGTTACATCG	AACTGGATCT	CAACAGCGGT	AAGATCCTTG
	ACGTGCTCAC	CCAATGTAGC	TTGACCTAGA	GTTGTCGCCA	TTCTAGGAAC
10451	AGAGTTTTTCG	CCCCGAAGAA	CGTTTTCCAA	TGATGAGCAC	TTTTAAAGTT
	TCTCAAAGC	GGGGCTTCTT	GCAAAAGGTT	ACTACTCGTG	AAAATTTCAA
10501	CTGCTATGTG	GCGCGGTATT	ATCCCCTATT	GACGCCGGGC	AAGAGCAACT
	GACGATACAC	CGCGCCATAA	TAGGGCATAA	CTGCGGCCCG	TTCTCGTTGA
10551	CGGTCGCCGC	ATACACTATT	CTCAGAATGA	CTTGGTTGAG	TACTCACCAG
	GCCAGCGGCG	TATGTGATAA	GAGTCTTACT	GAACCAACTC	ATGAGTGGTC
10601	TCACAGAAAA	GCATCTTACG	GATGGCATGA	CAGTAAGAGA	ATTATGCAGT
	AGTGTCTTTT	CGTAGAATGC	CTACCGTACT	GTCATTCTCT	TAATACGTCA
10651	GCTGCCATAA	CCATGAGTGA	TAACACTGCG	GCCAACTTAC	TTCTGACAAC
	CGACGGTATT	GGTACTCACT	ATTGTGACGC	CGGTGAATG	AAGACTGTTG
10701	GATCGGAGGA	CCGAAGGAGC	TAACCGCTTT	TTTGCACAAC	ATGGGGGATC
	CTAGCCTCCT	GGCTTCCTCG	ATTGGCGAAA	AAACGTGTTG	TACCCCTAG
10751	ATGTAACCTCG	CCTTGATCGT	TGGGAACCGG	AGCTGAATGA	AGCCATACCA
	TACATTGAGC	GGAAC TAGCA	ACCCTTGGCC	TCGACTTACT	TCGGTATGGT
10801	AACGACGAGC	GTGACACCAC	GATGCCTGTA	GCAATGGCAA	CAACGTTGCG
	TTGCTGCTCG	CACTGTGGTG	CTACGGACAT	CGTTACCGTT	GTTGCAACGC
10851	CAAAC TATTA	ACTGGCGAAC	TACTTACTCT	AGCTTCCCGG	CAACAATTAA
	GTTTGATAAT	TGACCGCTTG	ATGAATGAGA	TCGAAGGGCC	GTTGTTAATT
10901	TAGACTGGAT	GGAGGCGGAT	AAAGTTGCAG	GACCACTTCT	GCGCTCGGCC
	ATCTGACCTA	CCTCCGCCTA	TTTCAACGTC	CTGGTGAAGA	CGCGAGCCGG
10951	CTTCCGGCTG	GCTGGTTTAT	TGCTGATAAA	TCTGGAGCCG	GTGAGCGTGG
	GAAGGCCGAC	CGACCAAATA	ACGACTATTT	AGACCTCGGC	CACTCGCACC
11001	GTCTCGCGGT	ATCATTGCAG	CACTGGGGCC	AGATGGTAAG	CCCTCCCGTA
	CAGAGCGCCA	TAGTAACGTC	GTGACCCCGG	TCTACCATTC	GGGAGGGCAT
11051	TCGTAGTTAT	CTACACGACG	GGGAGTCAGG	CAACTATGGA	TGAACGAAAT
	AGCATCAATA	GATGTGCTGC	CCCTCAGTCC	GTTGATACCT	ACTTGCTTTA
11101	AGACAGATCG	CTGAGATAGG	TGCCTCACTG	ATTAAGCATT	GGTAACTGTC
	TCTGTCTAGC	GACTCTATCC	ACGGAGTGAC	TAATTCGTAA	CCATTGACAG
11151	AGACTCGCGA	CACTGCATTA	ATGAATCGGC	CAACGCGCGG	GGAGAGGCGG
	TCTGAGCGCT	GTGACGTAAT	TACTTAGCCG	GTTGCGCGCC	CCTCTCCGCC
11201	TTTGCGTATT	GGGCGCTCTT	CCGCTTCCTC	GCTCACTGAC	TCGCTGCGCT
	AAACGCATAA	CCCGCGAGAA	GGCGAAGGAG	CGAGTGACTG	AGCGACGCGA
11251	CGGTGCTTCG	GCTGCGGCGA	GCGGTATCAG	CTCACTCAAA	GGCGGTAATA
	GCCAGCAAGC	CGACGCCGCT	CGCCATAGTC	GAGTGAGTTT	CCGCCATTAT
11301	CGGTTATCCA	CAGAATCAGG	GGATAACGCA	GGAAAGAACA	TGTGAGCAAA
	GCCAATAGGT	GTCTTAGTCC	CCTATTGCGT	CCTTTCTTGT	AACTCGTTT
11351	AGGCCAGCAA	AAGGCCAGGA	ACCGTAAAAA	GGCCGCGTTG	CTGGCGTTTT
	TCCGGTCGTT	TTCCGGTCCT	TGGCATT TTT	CCGGCGCAAC	GACCGCAAAA
11401	TCCATAGGCT	CCGCCCCCT	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	GTTGCTCCA	AGCTGGGCTG
	GCGACATCCA	TAGAGTCAAG	CCACATCCAG	CAAGCGAGGT	TCGACCCGAC
11651	TGTGCACGAA	CCCCCGTTC	AGCCCGACCG	CTGCGCCTTA	TCCGGTAACT
	ACACGTGCTT	GGGGGGCAAG	TCGGGCTGGC	GACGCGGAAT	AGGCCATTGA
11701	ATCGTCTTGA	GTCCAACCCG	GTAAGACACG	ACTTATCGCC	ACTGGCAGCA
	TAGCAGAACT	CAGGTTGGGC	CATTCTGTGC	TGAATAGCGG	TGACCGTCGT
11751	GCCACTGGTA	ACAGGATTAG	CAGAGCGAGG	TATGTAGGCG	GTGCTACAGA
	CGGTGACCAT	TGTCCTAATC	GTCTCGCTCC	ATACATCCGC	CACGATGTCT
11801	GTTCTTGAAG	TGGTGGCCTA	ACTACGGCTA	CACTAGAAGG	ACAGTATTTG
	CAAGAACTTC	ACCACCGGAT	TGATGCCGAT	GTGATCTTCC	TGTCATAAAC
11851	GTATCTGCGC	TCTGCTGAAG	CCAGTTACCT	TCGGAAAAAG	AGTTGGTAGC
	CATAGACGCG	AGACGACTTC	GGTCAATGGA	AGCCTTTTTT	TCAACCATCG
11901	TCTTGATCCG	GCAAACAAAC	CACCGCTGGT	AGCGGTGGTT	TTTTTGTTTG
	AGAACTAGGC	CGTTTGTTTG	GTGGCGACCA	TCGCCACCAA	AAAAACAAAC
11951	CAAGCAGCAG	ATTACGCGCA	GAAAAAAAGG	ATCTCAAGAA	GATCCTTTGA
	GTTGGTCGTC	TAATGCGCGT	CTTTTTTTTCC	TAGAGTTCTT	CTAGGAAACT
12001	TCTTTTCTAC	GGGGTCTGAC	GCTCAGTGGA	ACGAAAACCT	A
	AGAAAAGATG	CCCCAGACTG	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	ATAGGGGAAAT	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	ACCCAACCA	TCTGAGAAAT	AATCTTCTCC	TTTCCCAGCT	TTATTCCCAG
	TGGGGTTGGT	AGACTCTTTA	TTAGAAGAGG	AAAGGGTCGA	AATAAGGGTC
301	GATTCTACAT	GACCAAATTA	CCAGAGTCAC	CACTCATTTT	AATCACAACA
	CTAAGATGTA	CTGGTTTAAT	GGTCTCAGTG	GTGAGTAAAA	TTAGTGTTGT
351	TAGTGTCAAA	TAAC TAGAAA	ACATGAGACA	ACAATGGAGA	GCTGAGTAAC
	ATCACAGTTT	ATTGATCTTT	TGTACTCTGT	TGTTACCTCT	CGACTCATTG
401	TATTAGTAGT	AGTACTTTAC	CAGAGAATGG	CCTCTATAGG	CTCACATGTA
	ATAATCATCA	TCATGAAATG	GTCTCTTACC	GGAGATATCC	GAGTGTACAT
451	GGAATGGTTG	GTCCCCAGGT	GGTAGGTAGA	GCTGTTTGAG	GATTACGTGG
	CCTTACCAAC	CAGGGGTCCA	CCATCCATCT	CGACAAACTC	CTAATGCACC
501	CCTTCTTGGA	TGGGGGGTGG	GGGTGGGGTG	GGAGGGTTGG	GTGGTGGGTA
	GGAAGAACCT	ACCCCCCACC	CCCACCCAC	CCTCCCAACC	CACCACCCAT
551	CTTAAGAGGT	TTCAAAAGTC	AATATTGTTT	GCATTTAGCT	CTTCCTTGTA
	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	ATCTTTCTTT	CTTTTCTTCT	TTTTCTTTTG	GTTTCGAGAC	AGGGTTTCTC
	TAGAAAGAAA	GAAAAGAAGA	AAAAGAAAAC	CAAAGCTCTG	TCCCAAAGAG
751	TGTATAGTCC	TGGCTGTCCT	GGAAC TCACT	TTGTAGACCA	GGCTGGCCTT
	ACATATCAGG	ACCGACAGGA	CCTTGAGTGA	AACATCTGGT	CCGACCGGAA
801	GAACTCAGAA	ATCCGCCTGC	CTCTGCCTCC	CTAGCACTGG	GATTAAAGGC
	CTTGAGTCTT	TAGGCGGACG	GAGACGGAGG	GATCGTGACC	CTAATTTCCG
851	GTGCGCCACC	ACGCCCAGCT	TTCAACTGGA	AATCTTAATA	AACTTTCCTA
	CACGCGGTGG	TGCGGGTCGA	AAGTTGACCT	TTAGAATTAT	TTGAAAGGAT
901	GAAGTGGCCT	TGGTTATGGG	AGCTTATCAC	AGCAATAGAA	CAGCAATTAT
	CTTCACCGGA	ACCAATACCC	TCGAATAGTG	TCGTATCTTT	GTCGTTAATA
951	GACTGGAGTA	TGATAGTTAA	AAACAAGCAA	GCAAGCAAGC	AAACACACAC
	CTGACCTCAT	ACTATCAATT	TTTGTTCGTT	CGTTCGTTTCG	TTTGTGTGTG
1001	ACCAAAACAA	CAAAACCCCA	AGACAGAGTC	ACATGTAGCC	CAGGCTAGCC
	TGGTTTTGTT	GTTTTGGGGT	TCTGTCTCAG	TGTACATCGG	GTCCGATCGG
1051	TCCAAATTCA	CTATATAACT	GAAGAAGACC	CCTAATTCCC	ATTCCTCTAG
	AGGTTTAAGT	GATATATTGA	CTTCTTCTGG	GGATTAAGGG	TAAGGAGATC
1101	AATCTATACC	TCAAGTACTG	AATGGCTTGG	TTCAACAATAC	CCCCTAAAT
	TTAGATATGG	AGTTCATGAC	TTACCGAACC	AAGTGTTATG	GGGTGATTTA
1151	GATTGGTCTT	ACTAAGTGCA	ACAAGGTAAA	CCTAAAACCTT	CAGCCCTCAG
	CTAACCAGAA	TGATTCACGT	TGTTCCATTT	GGATTTTGAA	GTCGGGAGTC
1201	ACATCCCTTT	TCCAGTATCA	ATTTATAAAA	TTAGATCCCA	AGGATAAAAA
	TGTAGGGAAA	AGGTCATAGT	TAAATATTTT	AATCTAGGGT	TCCTATTTTT
1251	TTAATTGTAA	AGTAAAATCA	GAGTTCTAGC	ATCAACTACA	GGCTCAACCA
	AATTAACATT	TCATTTT TAGT	CTCAAGATCG	TAGTTGATGT	CCGAGTTGGT
1301	TGGGGACCAC	AAATAA ACTA	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	TGGA ACTCTA	CCAATAAATC	TTCAAGGGTA	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	TGTA ACTACA	GATGCTACTT	AAAAATTTTT	TAAAGGAATC
	GTGGTGGTAC	ACATTGATGT	CTACGATGAA	TTTTTAAAAA	ATTTCTTAG
1851	ACAAAAATAA	CCCCCTATCA	AATGCCTAGT	CCCTCTAACC	ATCACCAAGT
	TGTTTTTTATT	GGGGGATAGT	TTACGGATCA	GGGAGATTGG	TAGTGGTTCA
1901	GAAGGATCAC	GCAGGAAAAA	AAAAATCACC	AGCAGCACCT	CAGAACCAGG
	CTTCCTAGTG	CGTCCTTTTT	TTTTTAGTGG	TCGTCGTGGA	GTCTTGGTCC
1951	ATACTCAGTC	CATCAGCATC	CAGGGCCATA	CCCACACTCA	CAGCATCTCC
	TATGAGTCAG	GTAGTCGTAG	GTCCCGGTAT	GGGTGTGAGT	GTCGTAGAGG
2001	ACAGTTTACC	AGATGATTCA	TGCTTATCAC	TGTATTGGGT	CATCTAAGAG
	TGTCAAATGG	TCTACTAAGT	ACGAATAGTG	ACATAACCCA	GTAGATTCTC
2051	TGACCATCAG	GGCTTCTGAT	CACAGAATCT	AGTCCACTTT	GCAGACCAGT
	ACTGGTAGTC	CCGAAGACTA	GTGTCTTAGA	TCAGGTGAAA	CGTCTGGTCA
2101	TGAAGTCATG	CACTATATGA	GATAGAAATA	CCCTCTTGCT	CATTTTGGTC
	ACTTCAGTAC	GTGATATACT	CTATCTTTAT	GGGAGAACGA	GTAAAACCAG
2151	AGAAATTCAA	GGATAAAAAC	CCATGTTTTG	TTAATGCACA	CCTCCATATG
	TCTTTAAGTT	CCTATTTTTG	GGTACAAAAC	AATTACGTGT	GGAGGTATAC
2201	ATTGAGATCA	ATGTGTCCTA	ATTAATGTAG	AAACCACAAC	TGTAAATTTT
	TAACTCTAGT	TACACAGGAT	TAATTACATC	TTTGGTGTG	ACATTTAAAG
2251	ACTCTTTTGA	CATGAATCTT	TTTCTAGACA	GGGTCTTGGA	TGCAGCCCCG
	TGAGAAAACT	GTA CTTAGAA	AAAGATCTGT	CCCAGAACCT	ACGTGCGGGC
2301	ACTACCCAGA	ATTTTGGAAT	CCAGGCTAGC	CTCAA ACTCA	AGGCAATCTG
	TGATGGGTCT	TAAAACCTTA	GGTCCGATCG	GAGTTTGAGT	TCCGTTAGAC
2351	CTTGCTTCAG	CTTCTCACAG	GCTGGATCAC	AAACATACAC	CTTCAGACCC
	GAACGAAGTC	GAAGAGTGTC	CGACCTAGTG	TTTGTATGTG	GAAGTCTGGG
2401	ATTTTTTTTT	CCTCCCTCCG	TTTTTGGTTT	CTCTGTGTAG	CCCTGGGTGT
	TAAAAA AAAA	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	AAGAAA ACTA	ACACTGGATT	TAGCAACAGT	CAGAAATCGC	TGAAAAGAAA
	TTCTTTTGAT	TGTGACCTAA	ATCGTTGTCA	GTCTTTAGCG	ACTTTTCTTT
2651	CAGAATTGAT	CTAACAGTCT	TAGATCACTC	CTAGACAGTT	TGTAATTCTT
	GTCTTAACTA	GATTGTCAGA	ATCTAGTGAG	GATCTGTCAA	ACATTAAGAA
2701	GCTCATGGCA	ACGTGAGCTC	TATCTAACTC	ACTCTCTGTG	CACTAATGAA
	CGAGTACCGT	TGCACTCGAG	ATAGATTGAG	TGAGAGACAC	GTGATTACTT
2751	TGCTCAGTGT	CTCCAGAACAA	GCACAGCTTC	CAGGGTAATC	ATGCCAACCC
	ACGAGTCACA	GAGGTCTTGT	CGTGTCGAAG	GTCCCATTAG	TACGGTTGGG

FIG. 23(contd.) 52/93

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	GCAAAAACCTG
	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	ATGTTCCCTGA
3101	TAGGTAGGCT	GGCTTCTAAC	AGCAATTCAC	CCACTTATGA	TGGGATTAAA
	ATCCATCCGA	CCGAAGATTG	TCGTTAAGTG	GGTGAATACT	ACCCTAATTT
3151	GGAAGGCACA	ACCATGTCCA	CCACAGGTTC	TAGCTCCCCC	ACCCACACGC
	CCTTCCGTGT	TGGTACAGGT	GGTGTCCAAG	ATCGAGGGGG	TGGGTGTGCG
3201	CCAGAGAGGG	TTTTTCTGTG	TAGCTCTGAC	TATTCTGGAA	TTCACACTGC
	GGTCTCTCCC	AAAAAGACAC	ATCGAGACTG	ATAAGACCTT	AAGTGTGACG
3251	AGACCAGGCT	GGTCTCGAAC	TCAGAGATCC	ACCACCACAT	GGTTTCTTAA
	TCTGGTCCGA	CCAGAGCTTG	AGTCTCTAGG	TGGTGGTGTA	CCAAAGAATT
3301	TTGTAATTTT	AAAGAAAAAA	AAAAATCCTT	CAGTTAAGAT	TCTTATGTTC
	AACATTAAAA	TTTCTTTTTT	TTTTTAGGAA	GTCAATTCTA	AGAATACAAG
3351	TAGGTTTTCA	CAAACCTACC	AATGTAGTTT	TATTGGAGGC	CATTTTTTTAA
	ATCCAAAAGT	GTTTGAATGG	TTACATCAAA	ATAACCTCCG	GTAAAAAATT
3401	ATTTAATCGG	AGACTTGAAG	AGCTATTGCA	AGAAAAAAA	TGTAGGACAG
	TAAATTAGCC	TCTGAACCTC	TCGATAACGT	TCTTTTTTTT	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	TCTTAAATTT	GTTAAATCA	CAAATAAAAC	CTACTTTTCT
3701	AATGACTATC	TTTTGTTAGC	AAATTACCAT	AAGATCTTTT	TCTTTAGATT
	TTACTGATAG	AAAACAATCG	TTTAATGGTA	TTCTAGAAAA	AGAAATCTAA
3751	TCTGAATACT	CCAAGGAGCT	CATATAATTC	CATCCTTATT	TTTTTCAGAGG
	AGACTTATGA	GGTTCCTCGA	GTATATTAAG	GTAGGAATAA	AAAAGTCTCC
3801	CCCTCCCTGT	TCAATCACGG	TATAAAAAAA	GGAACACATT	AAGATGTCCC
	GGGAGGGA	AGTTAGTGCC	ATATTTTTTT	CCTTGTGTAA	TTCTACAGGG
3851	AGTCCTATTT	TCTGGCTTTT	TTTTTCCGGG	GGTGGTGGTG	CGGTAATCAC
	TCAGGATAAA	AGACCGAAAA	AAAAGGCCC	CCACCACCAC	GCCATTAGTG
3901	TCTCTATAGT	CCAGTCTGGG	CTTCAACGCC	TGGCAATCCC	CAGCCTCAAG
	AGAGATATCA	GGTCAGACCC	GAAGTTGCGG	ACCGTTAGGG	GTCGGAGTTC
3951	CTCCCAAGTA	CTGTCCTGAT	AAGGATAGAA	GGAGTCGACC	TCCTTCACGC
	GAGGGTTCAT	GACAGGACTA	TTCTATCTT	CCTCAGCTGG	AGGAAGTGCG
4001	TCCCCCTCCGA	GGAGGGCTCC	TTCCCAGCTC	CATTCCCCGG	TCGGGAGCCC
	AGGGGAGGCT	CCTCCCGAGG	AAGGGTCGAG	GTAAGGGGCC	AGCCCTCGGG
4051	GTCCCCCA	CGAGAGCGCG	GGCCTCGTGG	TCAGCGCCTC	CGCGGGGAGA
	CAGGGGGTGG	GCTCTCGCGC	CCGGAGCACC	AGTCGCGGAG	GCGCCCCCTCT
4101	AACAAAGGCG	GCGGCGGGGG	CTCAAGGGCA	CTGCGCCACG	GGCCCGCGCC
	TTGTTTCCGC	CGCCGCCCCC	GAGTTCCCGT	GACGCGGTGC	CCGGGCGCGG
4151	TCCCCCATCC	GGCGGCGGCC	ACGTAGCCGG	GAGCGCGCCG	CAGCCCGGAG
	AGGGGGTAGG	CCGCCGCCGG	TGCATCGGCC	CTCGCGCGGC	GTCGGGCCTC
4201	CCTCGGGCCT	CGCAGCTGCA	GAGCCTGAAC	CGCTCTCTCC	CTGCGGGCCT
	GGAGCCCGGA	GCGTCGACGT	CTCGGACTTG	GCGAGAGAGG	GACGCCCGGA

FIG. 23(contd.) 53/93

4251	GCGACGAGGC	TGGGGGAGGG	GAGGCCCGCG	CTTTGTCTGG	AGTCTCGGTA
	CGCTGCTCCG	ACCCCTCCC	CTCCGGGCGC	GAAACAGACC	TCAGAGCCAT
4301	GCTGTCATCC	GGCTCCCACC	CTCATGCACA	ATTGTCCCAT	CTCCCCACG
	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	TTCGAGCCTG	CGTGCCCTCGT
4451	TCCTAAACCC	CACCACACGC	AAGATCGAAA	AAAAGCAAAG	GCACGAACTT
	AGGATTTGGG	GTGGTGTGCG	TTCTAGCTTT	TTTTTCGTTC	CGTGCTTGAA
4501	CACCGCTCCG	ATGCTCAGGG	CCGCGGATCC	TGCAGAGTCT	CCCGCCTGCG
	GTGGCGAGGC	TACGAGTCCC	GGCGCCTAGG	ACGTCTCAGA	GGGCGGACGC
4551	CGCTTCGGTT	CAGCCACATC	CGAGGGGAGG	GGGCGCGGGC	AGCTCCGCCG
	GCGAAGCCAA	GTCGGTGTAG	GCTCCCTCC	CCGCGCCCCG	TCGAGGCGGC
4601	GGGGGGAGGG	GGAGCACCGC	CCACGCCCTG	GCCGCGCGGG	GCCCGCCGGG
	CCCCCTCCC	CCTCGTGGCG	GGTGCGGGAC	CGGCGCGCCC	CGGGCGGCC
4651	AACGCGTCCT	GCGGGGGGCG	GCGCGCGCAA	TGCTCACCGT	CCGCGGCGTG
	TTGCGCAGGA	CGCCCCCGC	CGCGCGCGTT	ACGAGTGGCA	GGCGCCGCAC
4701	GCGCCAGGG	GGTCTCCTGG	CTGGGGGGAG	GGGGGGGAAG	GCGGGCAGGA
	CGCGGGTCCC	CCAGAGGACC	GACCCCTC	CCCCCTTC	CGCCCGTCCT
4751	AGGACCGCGG	AGGCCTCTCT	GCGTCTCGGA	GCGCGCCAAA	GCGGGGCTCC
	TCCTGGCGCC	TCCGGAGAGA	CGCAGAGCCT	CGCGCGGTTT	CGCCCCGAGG
4801	ACCCACCTCC	TTGCCCGGAT	CTTGAAGGCC	GGGGAGATAA	ACAGCGGGGT
	TGGGTGGAGG	AACGGGCCTA	GAACTTCCGG	CCCCTCTATT	TGTCGCCCA
4851	TCTTTAAGCA	CCACCTCTCA	CTAGGCGCGG	GATCCCAAGG	CTTGTGGCAT
	AGAAATTCGT	GGTGGAGAGT	GATCCGCGCC	CTAGGGTTCC	GAACACCGTA
4901	CCGGGGTGGT	ACTTGACTA	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
	CTGCCCCGAAC	CGTTTCCGCA	TCGACGTCTC	TCGTGCGAAC	CCACCTCTCC
5051	GCCGCACGCC	CCAGCGCCGG	CCTAAGCCCC	TCCCGACGGC	GTTATTTCAA
	CGGCGTGCGG	GGTCGCGGCC	GGATTCGGGG	AGGGCTGCCG	CAATAAAGTT
5101	ACTGCGCGAC	CGTTTCTCCG	CTCCCTACGC	GGAGGTGGGG	GCCGGACCTA
	TGACGCGCTG	GCAAAGAGGC	GAGGGATGCG	CCTCCACCCC	CGGCCTGGAT
5151	GTTCCGGACG	TAGTAACACG	CCGAGCGCGA	GCCTTCCGCA	ATTCACGGAA
	CAAGGCCTGC	ATCATTGTGC	GGCTCGCGCT	CGGAAGGCGT	TAAGTGCCTT
5201	CACAGTTGCG	CAAGTGATGT	AAAGCAGTCC	CGCTGTACCT	AAAGGGGGAG
	GTGTCAACGC	GTTCACTACA	TTTCGTCAGG	GCGACATGGA	TTTCCCCCTC
5251	TGTCACGTAC	TTGGCGTAAG	GAGAGTGTAG	GCCCTTCCCG	CCATTGGCGG
	ACAGTGCATG	AACCGCATTC	CTCTCACATC	CGGGAAGGGC	GGTAACCGCC
5301	CGGTTAGGGC	GTTTACGTAA	CGGCGTGACG	TAAGCGGAGA	CGCGTTAGTG
	GCCAATCCCG	CAAATGCATT	GCCGCACTGC	ATTCGCCTCT	GCGCAATCAC
5351	GGGGGAAGGT	TCTAGAAAAG	CGGCGGTCTC	GGCTCCAGCG	GCAGTAGCAG
	CCCCCTTCCA	AGATCTTTTC	GCCGCCAGAG	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
	GACTCAGGCG	CTACCTCCAT	TGCCCAAAC	TTAGTTACTC	AATAATTTT
5601	TGGCATGGCG	AGGCCGTAGG	CACCGCAATG	GAAACCGGCC	ACCCGCCTCC
	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	TCCTCCGGAG	ACTAACAGGC	TGGCGGAAGG
5751	CCGCCCTCAG	CCGCCCGGCG	CCATTTCCCT	CAGTTGGGGT	GGGGGATGGG
	GGCGGGAGTC	GGCGGGCCGC	GGTAAAGGGA	GTCAACCCCA	CCCCCTACCC
5801	AAGTGCCCGC	CGCGACCGGG	CTGGACCGCT	AAAGTAGCGC	GTGAGCGGGC
	TTACAGGGCG	GCGCTGGCCC	GACCTGGCGA	TTTCATCGCG	CACTCGCCCCG
5851	CATCGCTGGC	CTTTCGATGT	GCGCGGGCCT	AGGGGCTCGG	TTGTGTTTCG
	GTAGCGACCG	GAAAGCTACA	CGCGCCCGGA	TCCCCGAGCC	AACACAAGCG
5901	GGCGGAACGT	TTCTGGGGCC	CCCCCGGCTT	CCCGGAGCGA	GTCTGCGAAG
	CCGCCTTGCA	AAGACCCCGG	GGGGGCCGAA	GGGCCTCGCT	CAGACGCTTC
5951	CTAGCTTCCC	CTCCCCCCTC	TCCCGGGAAC	CGGATTTGGC	GGCCGCCATT
	GATCGAAGGG	GAGGGGGGAG	AGGGCCCTTG	GCCTAAACCG	CCGGCGGTAA
6001	TTCCCGTCTC	CTTCCTCGCC	ACGATTTTGC	TTTCAACGCT	TTAGGTTTAC
	AAGGGCAGAG	GAAGGAGCGG	TGCTAAAACG	AAAGTTGCGA	AATCCAAATG
6051	TAGTTTGGTT	TTCTTTTTTC	ACCACTGCGT	AGACGTGTTT	AGCGATTTTC
	ATCAAACCAA	AAGAAAAAAG	TGGTGACGCA	TCTGCACAAA	TCGCTAAAAG
6101	CTTTCTTTTG	GAAGTCTTCA	TACCGTTTCG	AGGTGGATTT	AGCGTTTTGA
	GAAAGAAAAC	CTTCAGAAGT	ATGGCAAAGC	TCCACCTAAA	TCGCAAAACT
6151	GCTTGGGTCT	TCAGCGTCCT	GCGCACCTCG	CTAAAGGCTC	TCTGCCTTCC
	CGAACCCAGA	AGTCGCAGGA	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	GGCGGGTGCG	TGAAGGCAGC	GTGCTGCAA
	GCATCATGTT	GGCTCCGGAC	CCGCCCACCG	ACTTCCGTCG	CAGCGACGTT
6351	AGAGACCGTT	TTATTTTTCA	TAATACGTAA	GATTACGGGT	GCTGTAGTAA
	TCTCTGGCAA	AATAAAAAGT	ATTATGCATT	CTAATGCCCA	CGACATCATT
6401	AGCACTTGAG	CATTAGTATA	GTAGGAGGAA	GTCAAAGTGG	AAAAAATGGG
	TCGTGAACTC	GTAATCATAT	CATCCTCCTT	CAGTTTCACC	TTTTTTACCC
6451	AGCGCTCATC	AGGAAGCTAG	GGAGGCTATG	TTGAGTGCAG	GGTACTTTTC
	TCGCGAGTAG	TCCTTCGATC	CCTCCGATAC	AACTCACGTC	CCAATGAAAG
6501	CTTTTATTGC	AGAACTTTTA	TCTGCTTAAA	GGATCCTCGG	ATCGAAATAA
	GAAAATAACG	TCTTGAAAAT	AGACGAATTT	CCTAGGAGCC	TAGCTTTATT
6551	TTCAAATTAT	AAGCATTTTT	AAGGGAATCT	TCGAATTTGT	TGGTAAAGTC
	AAGTTTAATA	TTCGTAAAAA	TTCCCTTAGA	AGCTTAAACA	ACCATTTTCAG
6601	AACGGATCCT	TAGCACGTGG	TGTTCACTTT	AAGGAAGTGA	AATAGCTGAC
	TTGCCTAGGA	ATCGTGCACC	ACAAGTGAAA	TTCTTCACT	TTATCGACTG
6651	TTTTCATAGT	TAGCCTTCGC	TTAAAGCCTG	GTTCAGTGGA	CGAAAATCCA
	AAAAGTATCA	ATCGGAAGCG	AATTTCCGGAC	CAAGTCACCT	GCTTTTAGGT
6701	CGTCCTGGCT	ATATAAAAAC	TTAGTTTGGG	GTCACAGTGT	TTGAGCGTGG
	GCAGGACCGA	TATATTTTTG	AATCAAACCC	CAGTGTCA	AACTCGCACC
6751	TCATTCGGTT	TTTTTATTTT	TTATTTGTTT	GAAATTATGA	TGCATCATTA
	AGTAAGCCAA	AAAAATAAAA	AATAAACAAA	CTTTAATACT	ACGTAGTAAT
6801	CACTGATAAG	CATTAGCTTT	CGAATTGAAA	GGGGTCTCCT	TGGTTATTTT
	GTGACTATTC	GTAATCGAAA	GCTTAACTTT	CCCCAGAGGA	ACCAATAAAA
6851	CTTTGACTCT	AAGCACACTT	ATAAATAAAA	TAACCTTGTT	TATAATCGAT
	GAAACTGAGA	TTCGTGTGAA	TATTTATTTT	ATTGGAACAA	ATATTAGCTA
6901	AGTGGACGTC	TGGTAAGTTT	GGAAAAAACC	CGAGGTAAGT	AAAGAGCTTT
	TCACCTGCAG	ACCATTCAAA	CCTTTTTTTG	GCTCCATTCA	TTTCTCGAAA
6951	TGCTTTCGTT	AGTGATATGA	AAAAACAAGG	TGTATTTAAT	ACTTGCAACT
	ACGAAAGCAA	TCACTATACT	TTTTTGTTCC	ACATAAATTA	TGAACGTTGA
7001	TAGTTTAAGG	AAAGCCAATT	TACTGACATT	TTAGTAGAGC	TACCAGAAAC
	ATCAAATTCC	TTTCGGTTAA	ATGACTGTAA	AATCATCTCG	ATGGTCTTTG
7051	ACTATTTGGA	GTCCTGATTA	AGGCTTTTGT	AACTATTTTG	ACTATTTAAA
	TGATAAACCT	CAGGACTAAT	TCCGAAAACA	TTGATAAAAC	TGATAAATTT
7101	ACAATTTTGG	TCGTTTTTAT	TAAACATTTT	AAAACCTAAA	AATTGTAAAC
	TGTTAAAACC	AGCAAAAATA	ATTTGTAAAG	TTTTGGATTT	TTAACATTTG

FIG. 23(contd.) 55/93

7151	ATTGGCTTTT	TGAGCACATT	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	AAGTCGTAA	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	TGGGTGTAG	AATTTAATGA	ATTACAAAAG
	TTTCATCATT	TAAAACCTGA	ACCCAACATC	TTAAATTACT	TAATGTTTTC
7601	AATTCTTTAA	TACCCTTTAA	ATGACCTAAG	AGCTGGGTAT	GGTTTTTCTG
	TTAAGAAATT	ATGGGAAATT	TACTGGATT	TCGACCCATA	CCAAAAAGAC
7651	AATTTTTTTG	AAGAAAATCT	AAGAAAGTTT	ACGTGAATTA	GAAGTTAGAT
	TTAAAAAAAC	TTCTTTTAGA	TTCTTTCAAA	TGCACTTAAT	CTTCAATCTA
7701	CGAATATTAG	TGACTTTGAA	ACTTGTATAG	CTCAGGCAAT	TTTTGGTGTA
	GCTTATAATC	ACTGAAACTT	TGAACATATC	GAGTCCGTTA	AAAACCACAT
7751	ACACAACATA	TATGCAGTTT	AACATATGGT	TTAAATTTGA	TGTAAGTTTTT
	TGTGTTGATT	ATACGTCAAA	TTGTATACCA	AATTTAAACT	ACATTCAAAA
7801	TTTTCTCCCC	CCCAGAAAAC	TTTAGAAACT	GTTCTTTTGG	AGAGGAAAAA
	AAAAGAGGGG	GGGTCTTTTG	AAATCTTTGA	CAAGGAAACC	TCTCCTTTTTT
7851	GGTACTCTAC	CAGCAGGTCA	CCTCATATTT	AAGAATTTAA	TTTCTGTCAT
	CCATGAGATG	GTCGTCCAGT	GGAGTATAAA	TTCTTAAATT	AAAGGACGTA
7901	ACAAAGAAAG	TGTAAATAAA	AATTGAAATG	GTATTTCCCT	TTGCAGAGAG
	TGTTTCTTTC	ACATTTATTT	TTAACTTTAC	CATAAAGGGA	AACGTCTCTC
7951	AAAAGGAACA	GTTCCGAAAG	CTCTTTATTG	GTGGCTTAAG	CTTATCGATA
	TTTTCTTTGT	CAAGGCTTTC	GAGAAATAAC	CACCGAATTC	GAATAGCTAT
8001	CCGGTGGCGC	GCCAATTGTT	AATTAAGATC	TGGCCCAATG	GGCCGTACGA
	GGCCACCGCG	CGGTTAACAA	TTAATTCTAG	ACCGGGTTAC	CCGGCATGCT
8051	ATTCCTTAGG	CTACCGGGTA	GGGGAGGCGC	TTTTCCCAAG	GCAGTCTGGA
	TAAGGAATCC	GATGGCCCAT	CCCCTCCGCG	AAAAGGGTTC	CGTCAGACCT
8101	GCATGCGCTT	TAGCAGCCCC	GCTGGGCACT	TGGCGCTACA	CAAGTGGCCT
	CGTACGCGAA	ATCGTCGGGG	CGACCCGTGA	ACCGCGATGT	GTTACCGGGA
8151	CTGGCCTCGC	ACACATTCCA	CATCCACCGG	CCGGTAGGCG	CCAACCGGCT
	GACCGGAGCG	TGTGTAAGGT	GTAGGTGGCC	GGCCATCCGC	GGTTGGCCGA
8201	CCGTTCCTTG	GTGGCCCCTT	CGCGCCACCT	TCTACTCCTC	CCCTAGTCAG
	GGCAAGAAAC	CACCGGGGAA	GCGCGGTGGA	AGATGAGGAG	GGGATCAGTC
8251	GAAGTTCCCC	CCCGCCCCGC	AGCTCGCGTC	GTGCAGGACG	TGACAAATGG
	CTTCAAGGGG	GGGCGGGGCG	TCGAGCGCAG	CACGTCCTGC	ACTGTTTACC
8301	AAGTAGCACG	TCTCACTAGT	CTCGTGCAGA	TGGACAGCAC	CGCTGAGCAA
	TTCATCGTGC	AGAGTGATCA	GAGCACGTCT	ACCTGTCGTG	GCGACTCGTT
8351	TGGAAGCGGG	TAGGCCTTTG	GGGCAGCGGC	CAATAGCAGC	TTTGCTCCTT
	ACCTTCGCCC	ATCCGGAAAC	CCCGTCGCCG	GTTATCGTCG	AAACGAGGAA
8401	CGCTTTCTGG	GCTCAGAGGC	TGGGAAGGGG	TGGGTCCGGG	GGCGGGCTCA
	GCGAAAGACC	CGAGTCTCCG	ACCCTTCCCC	ACCCAGGCCC	CCGCCCCGAGT
8451	GGGGCGGGCT	CAGGGGCGGG	GCGGGCGCCC	GAAGGTCCTC	CGGAGGCCCCG
	CCCCGCCCCG	GTCCCCGCCC	CGCCCCGCGG	CTTCCAGGAG	GCCTCCGGGC
8501	GCATTCTGCA	CGCTTCAAAA	GCGCACGTCT	GCCGCGCTGT	TCTCCTCTTC
	CGTAAGACGT	GCGAAGTTTT	CGCGTGCAGA	CGGCGCGACA	AGAGGAGAAG
8551	CTCATCTCCG	GGCCTTTCGA	CCAGCTTACC	ATGACCGAGT	ACAAGCCCAC
	GAGTAGAGGC	CCGGAAAGCT	GGTCGAATGG	TACTGGCTCA	TGTTCCGGGTG

FIG. 23(contd.) 56/93

8601	GGTGCGCCTC	GCCACCCGCG	ACGACGTCCC	CAGGGCCGTA	CGCACCCCTCG
	CCACGCGGAG	CGGTGGGCGC	TGCTGCAGGG	GTCCCGGCAT	GCGTGGGAGC
8651	CCGCCGCGTT	CGCCGACTAC	CCCGCCACGC	GCCACACCGT	CGATCCGGAC
	GGCGGCGCAA	GCGGCTGATG	GGGCGGTGCG	CGGTGTGGCA	GCTAGGCCTG
8701	CGCCACATCG	AGCGGGTCAC	CGAGCTGCAA	GAACTCTTCC	TCACGCGCGT
	GCGGTGTAGC	TCGCCCAGTG	GCTCGACGTT	CTTGAGAAGG	AGTGCGCGCA
8751	CGGGCTCGAC	ATCGGCAAGG	TGTGGGTGCG	GGACGACGGC	GCCGCGGTGG
	GCCCGAGCTG	TAGCCGTTCC	ACACCCAGCG	CCTGCTGCCG	CGGCGCCACC
8801	CGGTCTGGAC	CACGCCGGAG	AGCGTCGAAG	CGGGGGCGGT	GTTCGCCGAG
	GCCAGACCTG	GTGCGGCCCTC	TCGCAGCTTC	GCCCCCGCCA	CAAGCGGCTC
8851	ATCGGCCCGC	GCATGGCCGA	GTTGAGCGGT	TCCCGGCTGG	CCGCGCAGAA
	TAGCCGGGCG	CGTACCGGCT	CAACTCGCCA	AGGGCCGACC	GGCGCGTCTT
8901	CAGATGGAAG	GCCTCCTGGC	GCCGCACCGG	CCCAAGGAGC	CCGCGTGGTT
	GTCTACCTTC	CGGAGGACCG	CGGCGTGGCC	GGGTTCTTCG	GGCGCACCAA
8951	CCTGGCCACC	GTCGCGTCTC	GCCCGACCAC	CAGGGCAAGG	GTCTGGGCAG
	GGACCGGTGG	CAGCGCAGAG	CGGGCTGGTG	GTCCCGTTCC	CAGACCCGTC
9001	CGCCGTCGTG	CTCCCCGAG	TGGAGGCGGC	CGAGCGCGCC	GGGGTGCCCG
	GCGGCAGCAC	GAGGGGCCTC	ACCTCCGCCG	GCTCGCGCGG	CCCCACGGGC
9051	CCTTCCTGGA	GACCTCCGCG	CCCCGCAACC	TCCCCTTCTA	CGAGCGGCTC
	GGAAGGACCT	CTGGAGGCGC	GGGGCGTTGG	AGGGGAAGAT	GCTCGCCGAG
9101	GGCTTCACCG	TCACCGCCGA	CGTCGAGGTG	CCCGAAGGAC	CGCGCACCTG
	CCGAAGTGGC	AGTGGCGGCT	GCAGCTCCAC	GGGCTTCCTG	GCGCGTGGAC
9151	GTGCATGACC	CGCAAGCCCG	GTGCCTGACG	CCCGCCCCAC	GACCCGCAGC
	CACGTACTGG	GCGTTCGGGC	CACGGACTGC	GGGCGGGGTG	CTGGGCGTCG
9201	GCCCGACCGA	AAGGAGCGCA	CGACCCCATG	CATCGTAGAG	CTCGCTGATC
	CGGGCTGGCT	TTCCTCGCGT	GCTGGGGTAC	GTAGCATCTC	GAGCGACTAG
9251	AGCCTCGACT	GTGCCTTCTA	GTTGCCAGCC	ATCTGTTGTT	TGCCCCCTCC
	TCGGAGCTGA	CACGGAAGAT	CAACGGTCGG	TAGACAACAA	ACGGGGAGGG
9301	CCGTGCCTTC	CTTGACCCTG	GAAGGTGCCA	CTCCCCTGT	CCTTTCCTAA
	GGCACGGAAG	GAAGTGGGAC	CTTCCACGGT	GAGGGTGACA	GGAAAGGATT
9351	TAAAATGAGG	AAATTGCATC	GCATTGTCTG	AGTAGGTGTC	ATTCTATTCT
	ATTTTACTCC	TTTAACGTAG	CGTAACAGAC	TCATCCACAG	TAAGATAAGA
9401	GGGGGGTGGG	GTGGGGCAGG	ACAGCAAGGG	GGGGGATTGG	GRAGACAATA
	CCCCCACCC	CACCCCGTCC	TGTCGTTCCC	CCCCCTAACC	CYTCTGTTAT
9451	GCAGGCATGC	TGGGGGGGCG	GTGGGGGCTA	TGGCTTCTGA	GGCGGAAAGA
	CGTCCGTACG	ACCCCCCGCG	CACCCCCGAT	ACCGAAGACT	CCGCCTTTCT
9501	ACCAGCTGGG	GCTCGAGGGC	CGCCACCGCG	GTGGAGCTCC	AGCTTTTGT
	TGGTCGACCC	CGAGCTCCCC	GCGGTGGCGC	CACCTCGAGG	TCGAAAACAA
9551	CCCTTTAGTG	AGGGTTAATT	TCGAGCTTGG	CGTAATCATG	GTCATAGCTG
	GGGAAATCAC	TCCCAATTAA	AGCTCGAACC	GCATTAGTAC	CAGTATCGAC
9601	TTTCCTGTGT	GAAATTGTTA	TCCGCTCACA	ATTCCACACA	ACATACGAGC
	AAAGGACACA	CTTTAACAAT	AGGCGAGTGT	TAAGGTGTGT	TGTATGCTCG
9651	CGGAAGCATA	AAGTGTAAG	CCTGGGGTGC	CTAATGAGTG	AGCTAACTCA
	GCCTTCGTAT	TTACATTTT	GGACCCACG	GATTACTCAC	TCGATTGAGT
9701	CATTAATTGC	GTTGCGCTCA	CTGCCCGCTT	TCCAGTCGGG	AAACCTGTG
	GTAATTAACG	CAACGCGAGT	GACGGGCGAA	AGGTCAGCCC	TTTGGACAGC
9751	TGCCAGCATC	GCGAGCACTT	TTCGGGGAAA	TGTGCGCGGA	ACCCCTATTT
	ACGGTCGTAG	CGCTCGTGAA	AAGCCCCCTT	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	TGCCTTCTG
	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

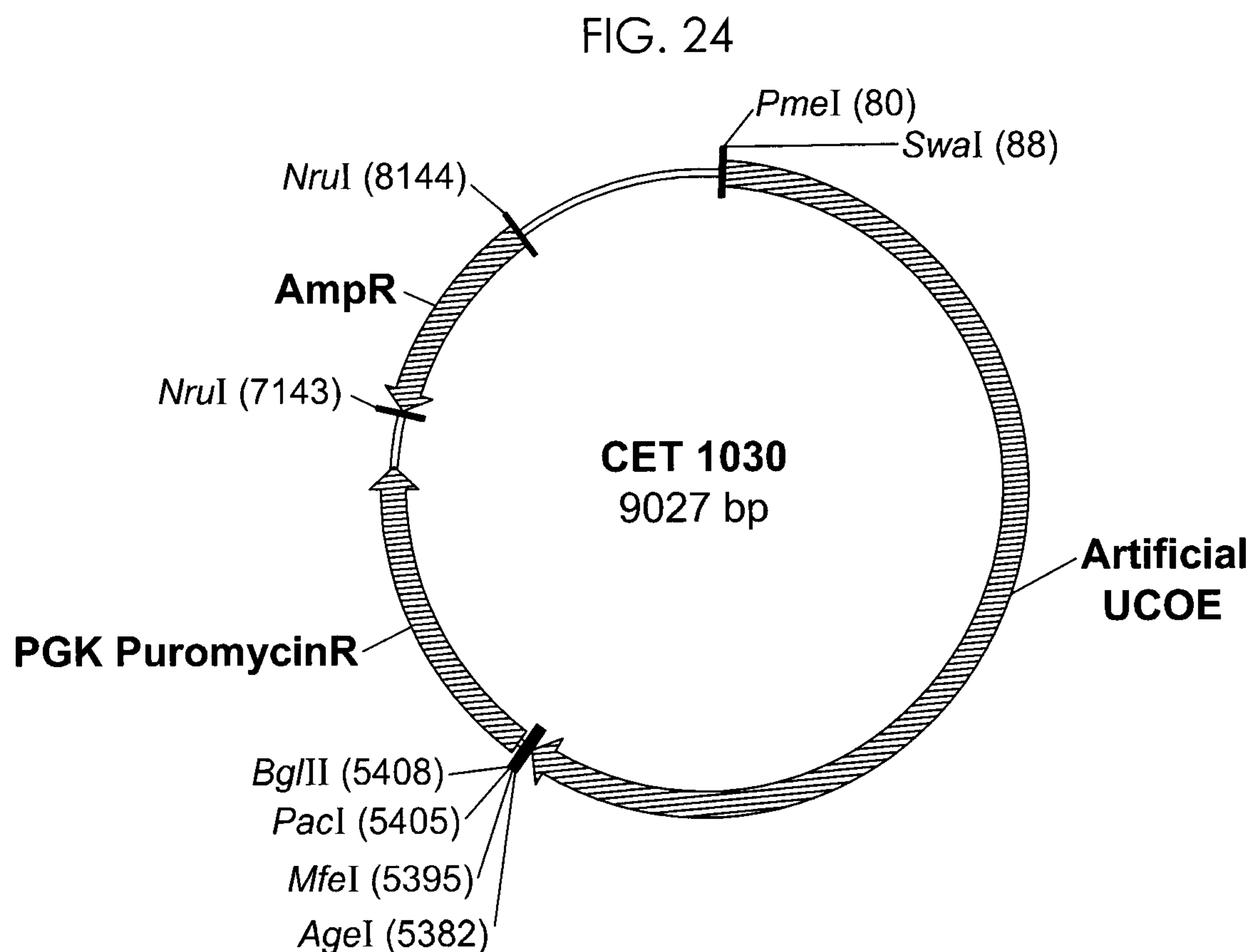
FIG. 23(contd.) 57/93

10051	CCTTGAGAGT	TTTCGCCCCG	AAGAACGTTT	TCCAATGATG	AGCACTTTTA
	GGAAGTCTCA	AAAGCGGGGC	TTCTTGCAAA	AGGTTACTAC	TCGTGAAAAT
10101	AAGTTCTGCT	ATGTGGCGCG	GTATTATCCC	GTATTGACGC	CGGGCAAGAG
	TTCAAGACGA	TACACCGCGC	CATAATAGGG	CATAACTGCG	GCCCGTTCTC
10151	CAACTCGGTC	GCCGCATACA	CTATTCTCAG	AATGACTTGG	TTGAGTACTC
	GTTGAGCCAG	CGGCGTATGT	GATAAGAGTC	TTACTGAACC	AACTCATGAG
10201	ACCAGTCACA	GAAAAGCATC	TTACGGATGG	CATGACAGTA	AGAGAATTAT
	TGGTCAGTGT	CTTTTCGTAG	AATGCCTACC	GTACTGTCAT	TCTCTTAATA
10251	GCAGTGCTGC	CATAACCATG	AGTGATAACA	CTGCGGCCAA	CTTACTTCTG
	CGTCACGACG	GTATTGGTAC	TCACTATTGT	GACGCCGGTT	GAATGAAGAC
10301	ACAACGATCG	GAGGACCGAA	GGAGCTAACC	GCTTTTTTGC	ACAACATGGG
	TGTTGCTAGC	CTCCTGGCTT	CCTCGATTGG	CGAAAAAACG	TGTTGTACCC
10351	GGATCATGTA	ACTCGCCTTG	ATCGTTGGGA	ACCGGAGCTG	AATGAAGCCA
	CCTAGTACAT	TGAGCGGAAC	TAGCAACCCT	TGGCCTCGAC	TTACTTCGGT
10401	TACCAAACGA	CGAGCGTGAC	ACCACGATGC	CTGTAGCAAT	GGCAACAACG
	ATGGTTTGCT	GCTCGCACTG	TGGTGCTACG	GACATCGTTA	CCGTTGTTGC
10451	TTGCGCAAAC	TATTAACCTG	CGAACTACTT	ACTCTAGCTT	CCCGGCAACA
	AACGCGTTTG	ATAATTGACC	GCTTGATGAA	TGAGATCGAA	GGGCCGTTGT
10501	ATTAATAGAC	TGGATGGAGG	CGGATAAAGT	TGCAGGACCA	CTTCTGCGCT
	TAATTATCTG	ACCTACCTCC	GCCTATTTC	ACGTCCTGGT	GAAGACGCGA
10551	CGGCCCTTCC	GGCTGGCTGG	TTTATTGCTG	ATAAATCTGG	AGCCGGTGAG
	GCCGGGAAGG	CCGACCGACC	AAATAACGAC	TATTTAGACC	TCGGCCACTC
10601	CGTGGGTCTC	GCGGTATCAT	TGCAGCACTG	GGGCCAGATG	GTAAGCCCTC
	GCACCCAGAG	CGCCATAGTA	ACGTCGTGAC	CCCGGTCTAC	CATTGCGGAG
10651	CCGTATCGTA	GTTATCTACA	CGACGGGGAG	TCAGGCAACT	ATGGATGAAC
	GGCATAGCAT	CAATAGATGT	GCTGCCCCCT	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	GTTTCGGCTG	GGCGAGCGGT	ATCAGCTCAC	TCAAAGGCGG
	CGCGAGCCAG	CAAGCCGACG	CCGCTCGCCA	TAGTCGAGTG	AGTTTCCGCC
10901	TAATACGGTT	ATCCACAGAA	TCAGGGGATA	ACGCAGGAAA	GAACATGTGA
	ATTATGCCAA	TAGGTGTCTT	AGTCCCCTAT	TGCGTCCTTT	CTTGTACACT
10951	GCAAAAGGCC	AGCAAAAGGC	CAGGAACCGT	AAAAAGGCCG	CGTTGCTGGC
	CGTTTTCCGG	TCGTTTTCCG	GTCCTTGGCA	TTTTTCCGGC	GCAACGACCG
11001	GTTTTTCCAT	AGGCTCCGCC	CCCCTGACGA	GCATCACAAA	AATCGACGCT
	CAAAAAGGTA	TCCGAGGCGG	GGGGACTGCT	CGTAGTGTTT	TTAGCTGCGA
11051	CAAGTCAGAG	GTGGCGAAAC	CCGACAGGAC	TATAAAGATA	CCAGGCGTTT
	GTTCACTCTC	CACCGCTTTG	GGCTGTCCTG	ATATTTCTAT	GGTCCGCAAA
11101	CCCCCTGGAA	GCTCCCTCGT	GCGCTCTCCT	GTTCCGACCC	TGCCGCTTAC
	GGGGGACCTT	CGAGGGAGCA	CGCGAGAGGA	CAAGGCTGGG	ACGGCGAATG
11151	CGGATACCTG	TCCGCCTTTC	TCCCTTCGGG	AAGCGTGGCG	CTTTCTCATA
	GCCTATGGAC	AGGCGGAAAG	AGGGAAGCCC	TTGCGACCGC	GAAAGAGTAT
11201	GCTCACGCTG	TAGGTATCTC	AGTTCGGTGT	AGGTCGTTCG	CTCCAAGCTG
	CGAGTGCGAC	ATCCATAGAG	TCAAGCCACA	TCCAGCAAGC	GAGGTTGCGC
11251	GGCTGTGTGC	ACGAACCCCC	CGTTCAGCCC	GACCGCTGCG	CCTTATCCGG
	CCGACACACG	TGCTTGGGGG	GCAAGTCGGG	CTGGCGACGC	GGAATAGGCC
11301	TAACATATCGT	CTTGAGTCCA	ACCCGGTAAG	ACACGACTTA	TCGCCACTGG
	ATTGATAGCA	GAAGTCAGGT	TGGGCCATTG	TGTGCTGAAT	AGCGGTGACC
11351	CAGCAGCCAC	TGGTAACAGG	ATTAGCAGAG	CGAGGTATGT	AGGCGGTGCT
	GTCGTGCGTG	ACCATTTGTC	TAATCGTCTC	GCTCCATACA	TCCGCCACGA
11401	ACAGAGTTCT	TGAAGTGGTG	GCCTAACTAC	GGCTACACTA	GAAGGACAGT
	TGTCTCAAGA	ACTTCACCAC	CGGATTGATG	CCGATGTGAT	CTTCTGTCA
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
CAAACGTTTC TCGTCTAATG CGCGTCTTTT TTCCTAGAG TTCTTCTAGG
11601 TTTGATCTTT TCTACGGGGT CTGACGCTCA GTGGAACGAA AACTCA
AAACTAGAAA AGATGCCCCA GACTGCGAGT CACCTTGCTT TTGAGT

FIG. 23(contd.)



59/93

FIG. 25 **cET1030 nucleotide sequence**

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

FIG. 25(contd.) 60/93

1351	CTTTTTTTTT	TTTTTACAAA	GGTGTTCAG	GCATGAAAAA	TTTTAAAGTA
	GAAAAA	AAAAATGTTT	CCACAAGGTC	CGTACTTTTT	AAAATTTTCAT
1401	CTATACCTTT	CCTAATTTTA	CCTTTAAAGT	TGTCCTGGAA	ATATCTGGGT
	GATATGGAAA	GGATTAAAT	GGAAATTTCA	ACAGGACCTT	TATAGACCCA
1451	TGACAAAGGC	GATGAAACTG	AACTGAGACT	TAAAAAAAAG	ATTACCCACC
	ACTGTTTCCG	CTACTTTGAC	TTGACTCTGA	ATTTTTTTTC	TAATGGGTGG
1501	TGGTTGTGCA	CAAGCCTGCT	TATGTCCCAA	TCTCCAGTCT	AGGGTCTGAT
	ACCAACACGT	GTTCGGACGA	ATACAGGGTT	AGAGGTCAGA	TCCCAGACTA
1551	GCTCCTTGCT	GCAGTAATAT	GCTTTGTGGC	ATCTGGAGCA	CGTTTTGGGG
	CGAGGAACGA	CGTCATTATA	CGAAACACCG	TAGACCTCGT	GCAAAACCCC
1601	CCTAAACAGC	CACAAACCCT	GCAGAGATGA	GCACCAGACT	TAAGCTGGAG
	GGATTTGTCTG	GTGTTTGGGA	CGTCTCTACT	CGTGGTCTGA	ATTGACCTC
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	TTTTCAAAT	AGAAGTTTGT	TTTGTCTGTA	AGTGTATAAA
1801	TATCCTCTGC	ACGTAAAAC	GAAAATAACA	ACAACAAAAA	AGAAATGAAA
	ATAGGAGACG	TGCATTTTGA	CTTTTATTGT	TGTTGTTTTT	TCTTTACTTT
1851	GTTTTTGCTT	TCAGGAATAA	GCTTTTAAAA	TCCAGAAACT	AGATTTCTGC
	CAAAAACGAA	AGTCCTTATT	CGAAAATTTT	AGGTCTTTGA	TCTAAAGCAG
1901	CGGTACACGC	AACTGAGTTG	CCTCCTAGAG	GTGGTTTGAG	TTAATCAAAT
	GCCATGTGCG	TTGACTCAAC	GGAGGATCTC	CACCAAAC	AATTAGTTTA
1951	TAATAAGACT	GATCGTTAAG	AACGACTGCC	AAAAATACGA	AAAAGCTACT
	ATTATTCTGA	CTAGCAATTC	TTGCTGACGG	TTTTTATGCT	TTTTTCGATGA
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	CCCGCCGTCTG
2201	GGTGCTTGCC	GCCGTCCCCC	TGGTCGCTCC	CGGAATTAAC	GCCGCGCACG
	CCACGAACGG	CGGCAGGGGG	ACCAGCGAGG	GCCTTAATTG	CGGCGCGTGC
2251	CGTCGGAGGC	ATGGCCCCGT	CCCGACCCCG	TTTGGCGGCT	CACCTCGCAG
	GCAGCCTCCG	TACCGGGGCA	GGGCTGGGGC	AAACCGCCGA	GTGGAGCGTC
2301	GCCGGCACAG	CACGGCTGCT	CGCGGCAGCA	GAAGAGGAAG	ATGCAGCGGT
	CGGCCGTGTC	GTGCCGACGA	GCGCCGTCGT	CTTCTCCTTC	TACGTCGCCA
2351	GGAAGGCGTC	CGGGCGGCCA	GGCAGCGGCG	CATACACCTG	CAGCAGGAAG
	CCTTCCGCAG	GCCCGCCGGT	CCGTGCGCCG	GTATGTGGAC	GTCGTCCTTC
2401	GAGAGCGGGC	GGCCGCACAG	CTCGCAGGCC	AGGGCCTGGG	GCCCCGGCAG
	CTCTCGCCCC	CCGGCGTGTC	GAGCGTCCGG	TCCCGGACCC	CGGGGCCGTC
2451	CCCGGCCGCG	CCCAGCCATG	CCGGCCGCCC	GCCCACCTTG	CTGGGGAAC
	GGGCCGGCGC	GGGTCGGTAC	GGCCGGCGGG	CGGGTGGAAC	GACCCCTTGA
2501	GCTCGCTGCG	CAGTCGCCAC	GCCGGCGCCG	ACTCGGCGAA	GCCCAGCTCC
	CGAGCGACGC	GTCAGCGGTG	CGGCCGCGGC	TGAGCCGCTT	CGGGTCGAGG
2551	ACAGGCCTGG	CCCCGGCGGC	AGCCATGCGG	GGCGCGGGCT	GGCGTGGGGC
	TGTCCGGACC	GGGGCCGCCG	TCGGTACGCC	CCGCGCCCGA	CCGCACCCCG
2601	GCAGCCCACA	GCTGGGTGCG	AAGGCGGAAA	TCGGGCGCCG	GGCCGGAAGG
	CGTCGGGTGT	CGACCCAGCC	TTCCGCCTTT	AGCCCGCGGC	CCGGCCTTCC
2651	CAAGAGGCGG	GCACCTTTCC	GGAGGACAGG	AGGCGGAAAC	GCGTCTGACG
	GTTCTCCGCC	CGTGGAAGG	CCTCCTGTCC	TCCGCCTTTG	CGCAGACTGC
2701	GGAGCGGTTG	CAGGACCAAT	GCGAGGGAAC	GGGGCAGAGG	AAACCTCTCG
	CCTCGCCAAC	GTCCTGGTTA	CGCTCCCTTG	CCCCGTCTCC	TTTGGAGAGC
2751	GCATCAGCCC	CGCCCCCTGGC	GCCTCTGCCT	CCGAGCCGCT	TTCCTGGTGC
	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	GACGCCCACC
2901	CCGCCAGGGA	GGCAGAGCCC	TCCCGCAAAC	CTTCCCTGCT	GGTGTCCACC
	GGCGGTCCCT	CCGTCTCGGG	AGGGCGTTTG	GAAGGGACGA	CCACAGGTGG
2951	TCAGGGTGTG	GGAAACCTGT	GCGCTGGCCG	AGTGCTAACC	AAGAGTAGGC
	AGTCCCACAC	CCTTTGGACA	CGCGACCGGC	TCACGATTGG	TTCTCATCCG
3001	AGTGAAAGAC	AAATGAAGGT	TGAACAGGTA	AAGTGAGGAC	CCTACAGCGG
	TCACTTTCTG	TTTACTTCCA	ACTTGTC CAT	TTCACTCCTG	GGATGTCGCC
3051	AAACCAAGAA	TCCTGTGTGC	CTGAGAGTAA	TGAAGAAGCC	TCTGCAGAAG
	TTTGGTTCTT	AGGACACACG	GACTCTCATT	ACTTCTTCGG	AGACGTCTTC
3101	AGTCTTTTCT	GTCAGTCTTA	AGGTCTCTGT	TTTAATGTTA	GTGCTGGCTT
	TCAGAAAAGA	CAGTCAGAAT	TCCAGAGACA	AAATTACAAT	CACGACCGAA
3151	GCTGTACCTG	AATTCCAAGG	GAGGAGTGTA	TAATGAGGCA	TGGCCAACCC
	CGACATGGAC	TTAAGGTTCC	CTCCTCACAT	ATTACTCCGT	ACCGGTTGGG
3201	CCACTTCCCA	TCATTGCCTG	AACTAGTTTT	TCAGGTTAAC	TTCAGAATGC
	GGTGAAGGGT	AGTAACGGAC	TTGATCAAAA	AGTCCAATTG	AAGTCTTACG
3251	CCTTGGTACC	GCGGGCCCCC	TCTGTGGTCC	CACGCCACTG	ATCGCTGCAT
	GGAACCATGG	CGCCCGGGGG	AGACACCAGG	GTGCGGTGAC	TAGCGACGTA
3301	GCCCACCACC	TGGGTACACA	CAGTCTGTGA	TTCCCGGAGC	AGAACGGACC
	CGGGTGGTGG	ACCCATGTGT	GTCAGACACT	AAGGGCCTCG	TCTTGCCTGG
3351	CTGCCCACCC	GGTCTTGTGT	GCTACTCAGT	GGACAGACCC	AAGGCAAGAA
	GACGGGTGGG	CCAGAACACA	CGATGAGTCA	CCTGTCTGGG	TTCCGTTCTT
3401	AGGGTGACAA	GGACAGGGTC	TTCCCAGGCT	GGCTTTGAGT	TCCTAGCACC
	TCCCAC TGTT	CCTGTCC CAG	AAGGGTCCGA	CCGAAACTCA	AGGATCGTGG
3451	GCCCCGCCCC	CAATCCTCTG	TGGCACATGG	AGTCTTGGTC	CCCAGAGTCC
	CGGGGCGGGG	GTTAGGAGAC	ACCGTG TACC	TCAGAACCAG	GGGTCTCAGG
3501	CCCAGCGGCC	TCCAGATGGT	CTGGGAGGGC	AGTTCAGCTG	TGGCTGCGCA
	GGGTGCGCCG	AGGTCTACCA	GACCCTCCCG	TCAAGTCGAC	ACCGACGCGT
3551	TAGCAGACAT	ACAACGGACG	GTGGGCCCAG	ACCCAGGCTG	TGTAGACCCA
	ATCGTCTGTA	TGTTGCCTGC	CACCCGGGTC	TGGGTCCGAC	ACATCTGGGT
3601	GCCCCCCCCG	CCCGCAGTGC	CTAGGT CACC	CACTAACGCC	CCAGGCCTGG
	CGGGGGGGCG	GGGCGTCACG	GATCCAGTGG	GTGATTGCGG	GGTCCGGACC
3651	TCTTGGCTGG	GCGTGACTGT	TACCCTCAA	AGCAGGCAGC	TCCAGGGTAA
	AGAACCGACC	CGCACTGACA	ATGGGAGTTT	TCGTCCGTCG	AGGTCCCATT
3701	AAGGTGCCCT	GCCCTGTAGA	GCCCAC TTCC	TTCCCAGGGC	TGCGGCTGGG
	TTCCACGGGA	CGGGACATCT	CGGGTGAAGG	AAGGGTCCCG	ACGCCGACCC
3751	TAGGTTTGTA	GCCTTCATCA	CGGGCCACCT	CCAGCCACTG	GACCGCTGGC
	ATCCAAACAT	CGGAAGTAGT	GCCCGGTGGA	GGTCGGTGAC	CTGGCGACCG
3801	CCCTGCCCTG	TCCTGGGGAG	TGTGGT CCTG	C GACTCTAAT	GGCCGCAAGC
	GGGACGGGAC	AGGACCCCTC	ACACCAGGAC	GCTGAGATTA	CCGGCGTTCC
3851	CACCTGACTC	CCCCAACACC	ACACTCTACC	TCTCAAGCCC	AGGTCTCTCC
	GTGGACTGAG	GGGGTTGTGG	TGTGAGATGG	AGAGTTCGGG	TCCAGAGAGG
3901	CTAGTGACCC	ACCCAGCACA	TTTAGCTAGC	TGAGCCCCAC	AGCCAGAGGT
	GATCACTGGG	TGGGTCGTGT	AAATCGATCG	ACTCGGGGTG	TCCGTCTCCA
3951	CCTCAGGCCC	TGCTTTCAGG	GCAGTTGCTC	TGAAGTCGGC	AAGGGGGAGT
	GGAGTCEGGG	ACGAAAGTCC	CGTCAACGAG	ACTTCAGCCG	TTCCCCCTCA
4001	GACTGCCTGG	CCACTCCATG	CCCTCCAAGA	GCTCCTTCTG	CAGGAGCGTA
	CTGACGGACC	GGTGAGGTAC	GGGAGGTTCT	CGAGGAAGAC	GTCTCTGCAT
4051	CAGAACCCAG	GGCCCTGGCA	CCCGTGCAGA	CCCTGGCCCA	CCCCACCTGG
	GTCTTGGGTC	CCGGGACCGT	GGGCACGTCT	GGGACCGGGT	GGGGTGGACC
4101	GCGCTCAGTG	CCCAAGAGAT	GTCCACACCT	AGGATGTCCC	GCGGTGGGTG
	CGCGAGTCAC	GGGTCTCTA	CAGGTGTGGA	TCCTACAGGG	CGCCACCCAC
4151	GGGGGCCCCG	GAGACGGGCA	GGCCGGGGGC	AGGCCTGGCC	ATGCGGGGCC
	CCCCCGGGCT	CTCTGCCCGT	CCGGCCCCCG	TCCGGACCGG	TACGCCCCGG
4201	GAACCGGGCA	CTGCCCAGCG	TGGGGCGCGG	GGGCCACGGC	GCGCGCCCCC
	CTTGGCCCGT	GACGGGT CGC	ACCCCGCGCC	CCCGGTGCCG	CGCGCGGGGG

FIG. 25(contd.) 62/93

4251	AGCCCCCGGG	CCCAGCACCC	CAAGGCGGCC	AACGCCAAAA	CTCTCCCTCC
	TCGGGGGGCCC	GGGTCGTGGG	GTTCCGCCGG	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	GCTTTC AACG	GAAAATACCG
4451	TCGAGCGGCC	GCGGCGGCGC	CCTATAAAAC	CCAGCGGCGC	GACGCGCCAC
	AGCTCGCCGG	CGCCGCGCGG	GGATATTTTG	GGTCGCGCGC	CTGCGCGGTG
4501	CACCGCCGAG	ACCGCGTCCG	CCCGCGAGCA	CAGAGCCTCG	CCTTTGCCGA
	GTGGCGGCTC	TGGCGCAGGC	GGGCGCTCGT	GTCTCGGAGC	GGAAACGGCT
4551	TCCGCCGCCC	GTCCACACCC	GCCGCCAGGT	AAGCCCGGCC	AGCCGACCGG
	AGGCGGCGGG	CAGGTGTGGG	CGGCGGTCCA	TTGCGGCGCG	TCGGCTGGCC
4601	GGCATGCGGC	CGCGGCCCTT	CGCCCGTGCA	GAGCCGCGGT	CTGGGCGCGA
	CCGTACGCCG	GCGCCGGGAA	GCGGGCACGT	CTCGGCGGCA	GACCCGGCGT
4651	GCGGGGGGCG	CATGGGGGCG	AACCGGACCG	CCGTGGGGGG	CGCGGGAGAA
	CGCCCCCGCG	GTACCCCGCC	TTGGCCTGGC	GGCACCCCCC	GCGCCCTCTT
4701	GCCCCCTGGG	CTCCGGAGAT	GGGGGACACC	CCACGCCAGT	TCGCAGGCGC
	CGGGGACCCG	GAGGCCTCTA	CCCCCTGTGG	GGTGCGGTCA	AGCGTCCGCG
4751	GAGGCCGCGC	TCGGGCGGGC	GCGCTCCGGG	GGTGCCGCTC	TCGGGGCGGG
	CTCCGGCGCG	AGCCCGCCCC	CGCGAGGCCC	CCACGGCGAG	AGCCCGCCCC
4801	GGCAACCGGC	GGGGTCTTTG	TCTGAGCCGG	GCTCTTGCCA	ATGGGGATCG
	CCGTGCGCCG	CCCCAGAAAC	AGACTCGGCC	CGAGAACGGT	TACCCCTAGC
4851	CACGGTGGGC	GCGGCGTAGC	CCCCGTGAGG	CCCGGTGGGG	GCTGGGGCGC
	GTGCCACCCG	CGCCGCATCG	GGGGCAGTCC	GGGCCACCCC	CGACCCCGCG
4901	CATGCGCGTG	GCGGCTGGTC	CTTTGGGCGC	TAACTGCGTG	GCGGCTGGGA
	GTACGCGCAC	GCGCGACCAG	GAAACCCGCG	ATTGACGCAC	GCGCGACCCT
4951	ATTGGCGCTA	ATTGCGCGTG	GCGGCTGGGA	CTCAATGGCG	CTAATCGCGC
	TAACCGCGAT	TAACGCGCAC	GCGCGACCCT	GAGTTACCGC	GATTAGCGCG
5001	GTGCGTTCTG	GGGCCCGGGC	GCTTGCGCCA	CTTCCTGCCC	GAGCCGCTGG
	CACGCAAGAC	CCCGGGCCCC	CGAACGCGGT	GAAGGACGGG	CTCGGCGACC
5051	CGCCCGAGGG	TGTGGCCGCT	GCGTGCGCGC	GCGCGACCCG	GTCGCTGTTT
	GCGGGCTCCC	ACACCGGCGA	CGCACGCGCG	CGCGCTGGGC	CAGCGACAAA
5101	GAACCGGGCG	GAGGCGGGGC	TGGCGCCCCG	TTGGGAGGGG	GTTGGGGCCT
	CTTGGCCCCG	CTCCGCCCCG	ACCGCGGGCC	AACCCTCCCC	CAACCCCGGA
5151	GGCTTCCTGC	GCGCGCGCCG	GGGGACGCCT	CCGACCAGTG	TTTGCCTTTT
	CCGAAGGACG	GCGCGCGGCG	CCCCTGCGGA	GGCTGGTCAC	AAACGGAAAA
5201	ATGGTAATAA	GCGGCGCGGC	CCGGCTTCCT	TTGTCCCCAA	TCTGGGCGCG
	TACCATTATT	GCGCCGGCCG	GGCCGAAGGA	AACAGGGGTT	AGACCCGCGC
5251	CGCCGGCGCC	CCCTGGCGGC	CTAAGGACTC	GGCGCGCCGG	AAGTGGCCAG
	GCGGCCGCGG	GGGACCGCCG	GATTCCTGAG	CCGCGCGGCC	TTACCCGGTC
5301	GGCGGGGGCG	ACTTCGGCTC	ACAGCGCGCC	CGGCTATTCT	GCGAGCTCAC
	CCGCCCCCGC	TGAAGCCGAG	TGTCGCGCGG	GCCGATAAGA	GCGTCGAGTG
5351	CATGCCGGTC	GCCACCATGA	GCTTATCGAT	ACCGGTGGCG	CGCCAATTGT
	GTACGGCCAG	CGGTGGTACT	CGAATAGCTA	TGGCCACCGC	GCGGTAAACA
5401	TAATTAAGAT	CTGGCCCAAT	GGGCCGTACG	AATTCCTTAG	GCTACCGGGT
	ATTAATTCTA	GACCGGGTTA	CCCGGCATGC	TTAAGGAATC	CGATGGCCCA
5451	AGGGGAGGCG	CTTTTCCCAA	GGCAGTCTGG	AGCATGCGCT	TTAGCAGCCC
	TCCCCTCCGC	GAAAAGGGTT	CCGTCAGACC	TCGTACGCGA	AATCGTCGGG
5501	CGCTGGGCAC	TTGGCGCTAC	ACAAGTGGCC	TCTGGCCTCG	CACACATTCC
	GCGACCCGTG	AACCGCGATG	TGTTCAACCG	AGACCGGAGC	GTGTGTAAGG
5551	ACATCCACCG	GCCGGTAGGC	GCCAACCGGC	TCCGTTCTTT	GGTGGCCCCCT
	TGTAGGTGGC	CGGCCATCCG	CGGTTGGCCG	AGGCAAGAAA	CCACCGGGGA
5601	TCGCGCCACC	TTCTACTCCT	CCCCTAGTCA	GGAAGTTCCC	CCCCGCCCCG
	AGCGCGGTGG	AAGATGAGGA	GGGGATCAGT	CCTTCAAGGG	GGGGCGGGGC
5651	CAGCTCGCGT	CGTGCAAGAC	GTGACAAATG	GAAGTAGCAC	GTCTCACTAG
	GTCGAGCGCA	GCACGTCCTG	CACTGTTTAC	CTTCATCGTG	CAGAGTGATC

FIG. 25(contd.) 63/93

5701	TCTCGTGCAG	ATGGACAGCA	CCGCTGAGCA	ATGGAAGCGG	GTAGGCCTTT
	AGAGCACGTC	TACCTGTTCGT	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	CCGGAGGGCC	GGCATTCCTG	ACGCTTCAAA
	CCGCCCCGCG	GCTTCCAGGA	GGCCTCCGGG	CCGTAAGACG	TGCGAAGTTT
5901	AGCGCACGTC	TGCCGCGCTG	TTCTCCTCTT	CCTCATCTCC	GGGCCTTTTC
	TCGCGTGCAG	ACGGCGCGAC	AAGAGGAGAA	GGAGTAGAGG	CCCGGAAAGC
5951	ACCAGCTTAC	CATGACCGAG	TACAAGCCCA	CGGTGCGCCT	CGCCACCCGC
	TGGTCGAATG	GTAAGGCTC	ATGTTTCGGG	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	AGCCCGAGCT	GTAGCCGTTC
6151	GTGTGGGTCG	CGGACGACGG	CGCCGCGGTG	GCGGTCTGGA	CCACGCCGGA
	CACACCCAGC	GCCTGCTGCC	GCGGCGCCAC	CGCCAGACCT	GGTGCGGCCT
6201	GAGCGTCGAA	GCGGGGGCGG	TGTTTCGCCA	GATCGGGCCG	CGCATGGCCG
	CTCGCAGCTT	CGCCCCCGCC	ACAAGCGGCT	CTAGCCGGGC	GCGTACCGGC
6251	AGTTGAGCGG	TTCCCGGCTG	GCCGCGCAGA	ACAGATGGAA	GGCCTCCTGG
	TCAACTCGCC	AAGGGCCGAC	CGGCGCGTCT	TGTCTACCTT	CCGGAGGACC
6301	CGCCGCACCG	GCCCAAGGAG	CCGCGGTGGT	TCCTGGCCAC	CGTCGCGTCT
	GCGGCGTGGC	CGGGTTCCTC	GGGCGCACCA	AGGACCGGTG	GCAGCGCAGA
6351	CGCCCGACCA	CCAGGGCAAG	GGTCTGGGCA	GCGCCGTCGT	GCTCCCCGGA
	GCGGGCTGGT	GGTCCCCTTC	CCAGACCCGT	CGCGGCAGCA	CGAGGGGCCT
6401	GTGGAGGCGG	CCGAGCGCGC	CGGGGTGCC	GCCTTCCTGG	AGACCTCCGC
	CACCTCCGCC	GGCTCGCGCG	GCCCCACGGG	CGGAAGGACC	TCTGGAGGCG
6451	GCCCCGCAAC	CTCCCCCTTCT	ACGAGCGGCT	CGGCTTCACC	GTCACCGCCG
	CGGGGCGTTG	GAGGGGAAGA	TGCTCGCCGA	GCCGAAGTGG	CAGTGGCGGC
6501	ACGTCGAGGT	GCCCGAAGGA	CCGCGCACCT	GGTGCATGAC	CCGCAAGCCC
	TGCAGCTCCA	CGGGCTTCCT	GGCGCGTGGA	CCACGTACTG	GGCGTTCGGG
6551	GGTGCCTGAC	GCCCCGCCCA	CGACCCGCAG	CGCCCGACCG	AAAGGAGCGC
	CCACGGACTG	CGGGCGGGGT	GCTGGGCGTC	GCGGGCTGGC	TTTCCTCGCG
6601	ACGACCCCAT	GCATCGTAGA	GCTCGCTGAT	CAGCCTCGAC	TGTGCCTTCT
	TGCTGGGGTA	CGTAGCATCT	CGAGCGACTA	GTCGGAGCTG	ACACGGAAGA
6651	AGTTGCCAGC	CATCTGTTGT	TTGCCCCCTC	CCCGTGCCTT	CCTTGACCCT
	TCAACGGTCG	GTAACAACA	AACGGGGAGG	GGGCACGGAA	GGAATGGGA
6701	GGAAGGTGCC	ACTCCCACTG	TCCTTTCTTA	ATAAAATGAG	GAAATTGCAT
	CCTTCCACGG	TGAGGGTGAC	AGGAAAGGAT	TATTTTACTC	CTTTAACGTA
6751	CGCATTGTCT	GAGTAGGTGT	CATTCTATTC	TGGGGGGTGG	GGTGGGGCAG
	GCGTAACAGA	CTCATCCACA	GTAAGATAAG	ACCCCCACC	CCACCCCGTC
6801	GACAGCAAGG	GGGGGGATTG	GGRAGACAAT	AGCAGGCATG	CTGGGGGGGC
	CTGTCGTTCC	CCCCCCTAAC	CCYTCTGTTA	TCGTCCGTAC	GACCCCCCG
6851	GGTGGGGGCT	ATGGCTTCTG	AGGCGGAAAG	AACCAGCTGG	GGCTCGAGGG
	CCACCCCCGA	TACCGAAGAC	TCCGCCTTTC	TTGGTTCGAC	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	AATTCACAC	AACATACGAG	CCGGAAGCAT	AAAGTGTAAG
	TAGGCGAGTG	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

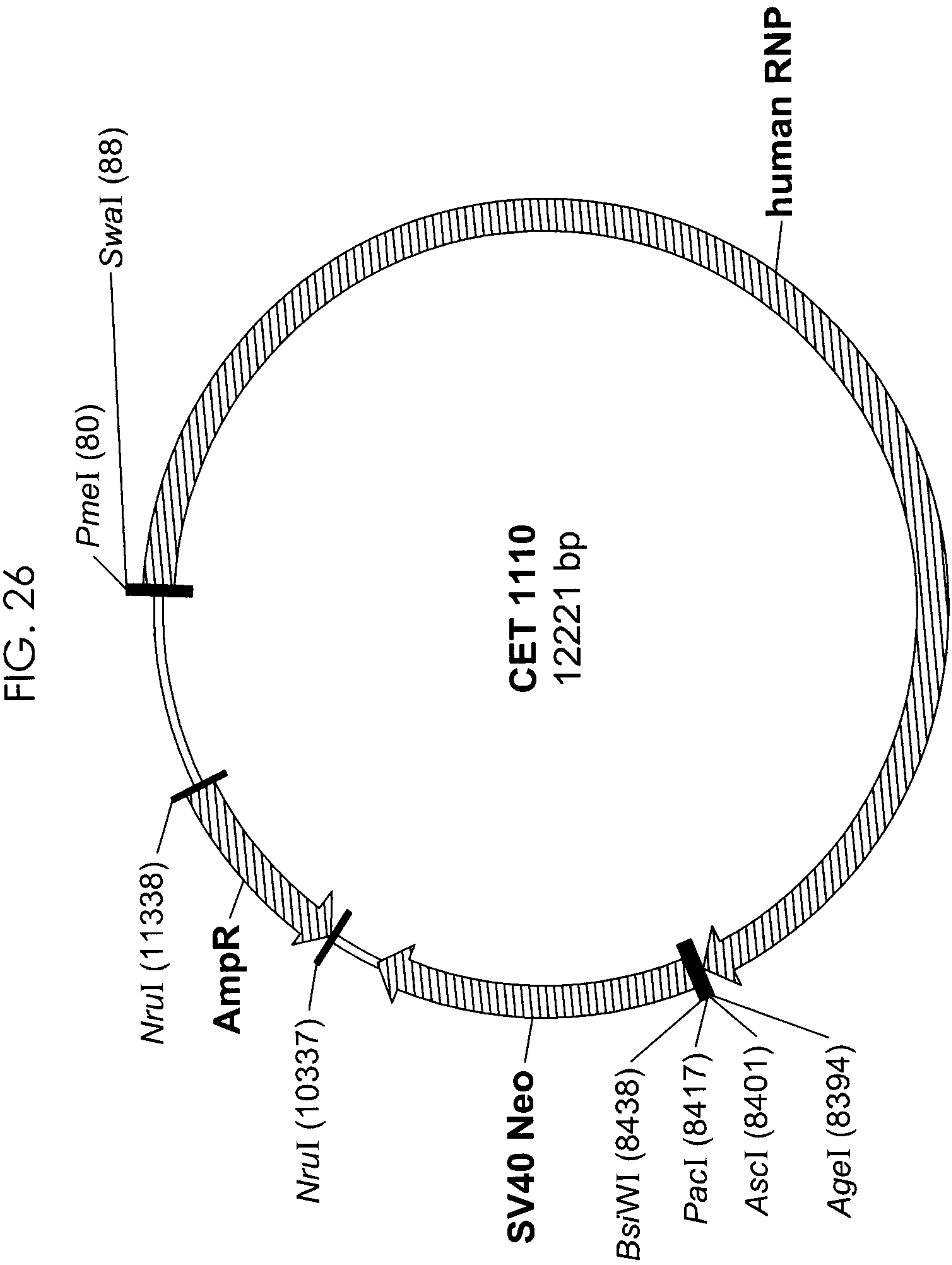
FIG. 25(contd.) 64/93

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	TTCTTCTCA	TACTCATAAG	TTGTAAAGGC	ACAGCGGGAA
7301	ATTCCCTTTT	TTGCGGCATT	TTGCCTTCCT	GTTTTTGTCT	ACCCAGAAAC
	TAAGGGAAAA	AACGCCGTAA	AACGGAAGGA	CAAAAACGAG	TGGGTCTTTG
7351	GCTGGTGAAA	GTAAAAGATG	CTGAAGATCA	GTTGGGTGCA	CGAGTGGGT
	CGACCACTTT	CATTTTCTAC	GACTTCTAGT	CAACCCACGT	GCTCACCCAA
7401	ACATCGAACT	GGATCTCAAC	AGCGGTAAGA	TCCTTGAGAG	TTTTCGCCCC
	TGTAGCTTGA	CCTAGAGTTG	TCGCCATTCT	AGGAACTCTC	AAAAGCGGGG
7451	GAAGAACGTT	TTCCAATGAT	GAGCACTTTT	AAAGTTCTGC	TATGTGGCGC
	CTTCTTGCAA	AAGGTTACTA	CTCGTGAAAA	TTTCAAGACG	ATACACCGCG
7501	GGTATTATCC	CGTATTGACG	CCGGGCAAGA	GCAACTCGGT	CGCCGCATAC
	CCATAATAGG	GCATAACTGC	GGCCCGTTCT	CGTTGAGCCA	GCGGCGTATG
7551	ACTATTCTCA	GAATGACTTG	GTTGAGTACT	CACCAGTCAC	AGAAAAGCAT
	TGATAAGAGT	CTTACTGAAC	CAACTCATGA	GTGGTCAGTG	TCTTTTCGTA
7601	CTTACGGATG	GCATGACAGT	AAGAGAATTA	TGCAGTGCTG	CCATAACCAT
	GAATGCCTAC	CGTACTGTCA	TTCTCTTAAT	ACGTCACGAC	GGTATTGGTA
7651	GAGTGATAAC	ACTGCGGCCA	ACTTACTTCT	GACAACGATC	GGAGGACCGA
	CTCACTATTG	TGACGCCGGT	TGAATGAAGA	CTGTTGCTAG	CCTCCTGGCT
7701	AGGAGCTAAC	CGCTTTTTTG	CACAACATGG	GGGATCATGT	AACTCGCCTT
	TCCTCGATTG	GCGAAAAAAC	GTGTTGTACC	CCCTAGTACA	TTGAGCGGAA
7751	GATCGTTGGG	AACCGGAGCT	GAATGAAGCC	ATACCAAACG	ACGAGCGTGA
	CTAGCAACCC	TTGGCCTCGA	CTTACTTCGG	TATGGTTTGC	TGCTCGCACT
7801	CACCACGATG	CCTGTAGCAA	TGGCAACAAC	GTTGCGCAA	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	CCATTCGGGA	GGGCATAGCA	TCAATAGATG
8051	ACGACGGGGA	GTCAGGCAAC	TATGGATGAA	CGAAATAGAC	AGATCGCTGA
	TGCTGCCCTT	CAGTCCGTTG	ATACCTACTT	GCTTTATCTG	TCTAGCGACT
8101	GATAGGTGCC	TCACTGATTA	AGCATTTGGTA	ACTGTCAGAC	TCGCGACACT
	CTATCCACGG	AGTGACTAAT	TCGTAACCAT	TGACAGTCTG	AGCGCTGTGA
8151	GCATTAATGA	ATCGGCCAAC	GCGCGGGGAG	AGGCGGTTTG	CGTATTGGGC
	CGTAATTACT	TAGCCGGTTG	CGCGCCCCTC	TCCGCCAAAC	GCATAACCCG
8201	GCTCTTCCGC	TTCTTCGCTC	ACTGACTCGC	TGCGCTCGGT	CGTTCGGCTG
	CGAGAAGGCG	AAGGAGCGAG	TGACTGAGCG	ACGCGAGCCA	GCAAGCCGAC
8251	CGGCGAGCGG	TATCAGCTCA	CTCAAAGGCG	GTAATACGGT	TATCCACAGA
	GCCGCTCGCC	ATAGTCGAGT	GAGTTTCCGC	CATTATGCCA	ATAGGTGTCT
8301	ATCAGGGGAT	AACGCAGGAA	AGAACATGTG	AGCAAAAGGC	CAGCAAAAGG
	TAGTCCCCTA	TTGCGTCCTT	TCTTGTACAC	TCGTTTTCCG	GTCGTTTTCC
8351	CCAGGAACCG	TAAAAAGGCC	GCGTTGCTGG	CGTTTTTCCA	TAGGCTCCGC
	GGTCCTTGGC	ATTTTTCCGG	CGCAACGACC	GCAAAAAGGT	ATCCGAGGCG
8401	CCCCCTGACG	AGCATCACAA	AAATCGACGC	TCAAGTCAGA	GGTGGCGAAA
	GGGGGACTGC	TCGTAGTGTT	TTTAGCTGCG	AGTTCAGTCT	CCACCGCTTT
8451	CCCGACAGGA	CTATAAAGAT	ACCAGGCGTT	TCCCCCTGGA	AGCTCCCTCG
	GGGCTGTCCT	GATATTTCTA	TGGTCCGCAA	AGGGGGACCT	TCGAGGGAGC
8501	TGCGCTCTCC	TGTTCCGACC	CTGCCGCTTA	CCGGATACCT	GTCCGCCTTT
	ACGCGAGAGG	ACAAGGCTGG	GACGGCGAAT	GGCCTATGGA	CAGGCGGAAA
8551	CTCCCTTCGG	GAAGCGTGGC	GCTTTCTCAT	AGCTCACGCT	GTAGGTATCT
	GAGGGAAGCC	CTTCGCACCG	CGAAAGAGTA	TCGAGTGCGA	CATCCATAGA

65/93

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

FIG. 25(contd.)



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	TAGCACCCCTC	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
	TGTCCTTTTCG	TCCATTTTCT	CGGTCCAATT	TTTGTTTAAG	ATTTTGGTGA
401	TAGCTGCAGT	TAAACATATA	GTAAAGATGC	ACTAAAGTTT	CTTACTCTGT
	ATCGACGTCA	ATTTGTATAT	CATTTCTACG	TGATTTCAAA	GAATGAGACA
451	AAATCCCTTC	CACTTCAGGA	AATATTCCAC	TTTCCCATTC	ACTACACGTC
	TTTAGGGAAG	GTGAAGTCCT	TTATAAGGTG	AAAGGGTAAG	TGATGTGCAG
501	GATCTAGTAC	TTTTTCCACG	ACAAATTCTT	CAGGCTCTGC	CTCTTCAACT
	CTAGATCATG	AAAAAGGTGC	TGTTTAAGAA	GTCCGAGACG	GAGAAGTTGA
551	TTTTTACTCT	TTCCATTCTG	TTTTTTTCCC	ATTTTTTGCT	AAAATAAAAC
	AAAAATGAGA	AAGGTAAGAC	AAAAAAAGGG	TAAAAAACGA	TTTTATTTTG
601	AAAAGAGAAA	TTAAGAAATA	TTCTCTTGA	ATTTTGAGCA	CATTTTCAAG
	TTTTCTCTTT	AATTCTTTAT	AAGGAGAACT	TAAAACCTCGT	GTAAAAGTTC
651	GCTCAATTGC	TTATATTATT	ATCACATTCG	ACATAAATTT	TTACTTCTAT
	CGAGTTAACG	AATATAATAA	TAGTGTAAGC	TGTATTTAAA	AATGAAGATA
701	ATCCCAGGGC	AGACACCTTC	TGGAAAGATT	AAAAGTCAAC	AGACAATAAA
	TAGGGTCCCG	TCTGTGGAAG	ACCTTTCTAA	TTTTTCAGTTG	TCTGTTATTT
751	ATAAAAGAAT	GCTTTATCTT	GTTCATTTAG	TTCAAACCTA	CAACCCACCA
	TATTTTCTTA	CGAAATAGAA	CAAGTAAATC	AAGTTTGAAT	GTTGGGTGGT
801	CCAAAATAAT	ACAATAAAAA	AACACTATCT	GGAAACAGTT	ATTTTTTTTCC
	GGTTTTATTA	TGTTATTTTT	TTGTGATAGA	CCTTTGTCAA	TAAAAAAAGG
851	AGTCTTTTTT	TTTGAGACAG	GGTCTCACAC	TCTTGTCGCC	CAGGCTGGAG
	TCAGAAAAAA	AAACTCTGTC	CCAGAGTGTG	AGAACAGCGG	GTCCGACCTC
901	TGCAGTGGCG	TGATCTCAGC	TCACTGCAAC	CTCCGCCTCC	CCAGGTTCAA
	ACGTCACCGC	ACTAGAGTCG	AGTGACGTTG	GAGGCGGAGG	GGTCCAAGTT
951	GCAGTTCTCA	TGCCTCAGCC	TCCAGAGTAG	CTGGGATTAT	AGGCGGATGC
	CGTCAAGAGT	ACGGAGTCGG	AGGTCTCATC	GACCCTAATA	TCCGCCTACG
1001	CACCATGCCG	GGCTAATTTT	TTTTGTGTTT	TTATTAGAAA	CAGGGTTTCA
	GTGGTACGGC	CCGATTAAAA	AAAACACAAA	AATAATCTTT	GTCCCAAAGT
1051	CCATGTTGAC	CAGGCTGGTC	TCAAACCTCT	GACCTGAAGT	GATTCAACCAG
	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	GTCCTTCATG	GATTTCCAC	TAGGTCTCTT
1251	GCAAGTAGTA	GTGACAGGTC	TTAGGTGAAC	CTATGACAGA	CCTTGTATCC
	CGTTCATCAT	CACTGTCCAG	AATCCACTTG	GATACTGTCT	GGAACATAGG
1301	ACCCCAGAT	GGTAAAAGCC	CCAGCCCCCT	TCTCAATTCA	AATATTAATG
	TGGGGGTCTA	CCATTTTTCGG	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	GGTTCAATTA	AAGGACAAGA	TTATTTTTC	ATAATCTCTG	TAGATAATAT
	CCAAGTTAAT	TTCCTGTTCT	AATAAAAAGG	TATTAGAGAC	ATCTATTATA
1501	TTCCTACCAC	TTAGAACAAA	ACTATAAAGC	TATCACTTCA	AGAGACCAAC
	AAGGATGGTG	AATCTTGTTT	TGATATTTTCG	ATAGTGAAGT	TCTCTGGTTG
1551	ATTACAAATT	TATTTTAATT	CCCTAAGGTG	AAAAAAATCC	TTCCTTCCTG
	TAATGTTTAA	ATAAAATTAA	GGGATTCCAC	TTTTTTTAGG	AAGGAAGGAC
1601	GTTTCTCAAG	AGAAAGTCTA	TACTGGTAAC	CAAATTCACT	TTAAACAGGC
	CAAAGAGTTC	TCTTTCAGAT	ATGACCATTG	GTTTAAGTGA	AATTTGTCCG
1651	ATTTTCTTTG	GTATGACACT	ATTTAAGAGA	AGCAGGAAAC	CAACGTGAAC
	TAAAAGAAAC	CATACTGTGA	TAAATTCTCT	TCGTCCTTTG	GTTGCACTTG
1701	CAGCTCTTTC	CAATGGCTCA	AGATTTTCCTA	TGAGAGGACT	AAAAATGGGG
	GTCGAGAAAG	GTTACCGAGT	TCTAAAGGAT	ACTCTCCTGA	TTTTTACCCC
1751	AAAATTTTTA	TGAGAGGATT	AAAAATGGGG	GAAAAAAAAC	CCTGAAATGG
	TTTTAAAAAT	ACTCTCCTAA	TTTTTACCCC	CTTTTTTTTG	GGACTTTACC
1801	TTAATCAGAA	GATCCTATGG	GCTGAGAAGG	AATCCATCTT	AACATTTTCAT
	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	TTATGTTTTA
2001	TAGCCAGGCA	TAGTGGTGCA	TGCTGTAAAT	CCCAGCTACT	TGGGAGGCTG
	ATCGGTCCGT	ATCACCACGT	ACGGACATTA	GGGTCGATGA	ACCCTCCGAC
2051	AGGCAGGAGA	ACTGCTTGAA	CCCAGGAGGC	TTAAGTTGCG	GTGAGCCAAG
	TCCGTCCTCT	TGACGAACCT	GGGTCCTCCG	AATTCAACGC	CACTCGGTTC
2101	ATCACGCCAT	TGCACTCTAG	CCTGGACAAC	AAGAGAAAAA	CTCTGTCTCA
	TAGTGCGGTA	ACGTGAGATC	GGACCTGTTG	TTCTCTTTTT	GAGACAGAGT
2151	AAAAAACACA	AAAACAAAAA	ACCCAAATAC	TATTTAAAAA	AGATAAACCT
	TTTTTTGTGT	TTTTGTTTTT	TGGGTTTATG	ATAAATTTTT	TCTATTTGGA
2201	TAATTGCTCA	ATCATTAAAG	CCATCCCACA	AGTAAAGCAG	CAAGCAGAAA
	ATTAACGAGT	TAGTAATTTT	GGTAGGGTGT	TCATTTTCGTC	GTTTCGTCTT
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	GACGTTTTTTG
	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	TTACACATAGT
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	TTTCCGTCTG	TTTTACGGTA	TGTCATAAGT
2751	TTCACTACTA	TGGCATTCAT	AAGCTAGTTT	CAAATGCTCA	CTATTTTCTT
	AAGTGATGAT	ACCGTAAGTA	TTCGATCAAA	GTTTACGAGT	GATAAAAGAA

FIG. 27(contd.)

69/93

2801	TTATAGTATA	TATTTGCCTT	AACCCAGCAC	TTTTTTCCAA	AAGTGGATGA
	AATATCATAT	ATAAACGGAA	TTGGGTCGTG	AAAAAAGGTT	TTCACCTACT
2851	GTCAAAATAA	ATTTCCCAT	ATTTAAGTGA	AATTAACAGC	ACACATATCT
	CAGTTTTATT	TAAAGGGTAA	TAAATTCAC	TTAATTGTCT	TGTGTATAGA
2901	CACAACACTA	ATGAATTTTT	AAAATGGAAA	GTTAAGAACT	TTTAAAGTGG
	GTGTTGTGAT	TACTTAAAAA	TTTTACCTTT	CAATTCTTGA	AAATTTTACC
2951	CCAACCTGTG	ATCCTTCACA	AAATAAACTA	AATACAATAA	CAGACCCCAA
	GGTTGGACAC	TAGGAAGTGT	TTTATTTGAT	TTATGTTATT	GTCTGGGGTT
3001	AGGCTATCAA	TTGCGTGCAA	AAACAACCTC	TGTTTTCCAG	GGTAAACAGA
	TCCGATAGTT	AACGCACGTT	TTTGTGTAAG	ACAAAAGGTC	CCATTTGTCT
3051	ATCTAATGCA	GAATCTAATG	CAGGGTAAAC	AGACTTAATG	CAGAATCTAA
	TAGATTACGT	CTTAGATTAC	GTCCCATTTG	TCTGAATTAC	GTCTTAGATT
3101	TGATGGCACA	AATTA AAAAT	CAC TAACGTG	CCCTTTT TAG	TGTGAAACCC
	ACTACCGTGT	TTAATTTTTA	GTGATTGCAC	GGGAAAATC	ACACTTTGGG
3151	AGAGAGAGCA	CATACAAGCC	AAAAACAAAT	GCTTTATTTT	ACCTAGGAGA
	TCTCTCTCGT	GTATGTTCTG	TTTTTGT TTA	CGAAATAAAA	TGGATCCTCT
3201	CATTAACATT	CACCTTTACG	TGTTTAAGAT	TAATGCAATG	TTAAATATTG
	GTAATTGTAA	GTGGAAATGC	ACAAATTCTA	ATTACGTTAC	AATTTATAAC
3251	TGAAAAC TGT	AAC TTTGAAT	TTCATGATTT	TTATGTGAAT	ATTCCAGGGT
	ACTTTTGACA	TTGAAACTTA	AAGTACTAAA	AATACACTTA	TAAGGTCCCA
3301	TTAAAAAAAC	TTGTAAACATG	ACATGGCTGA	ATAAGATAAA	AAAAAAATCT
	AATTTTTTTTG	AACATTGTAC	TGTACCGACT	TATTCTATTT	TTTTTTT TAGA
3351	AGCCTTTTCT	CCCTTCTGGC	TCATATTTGC	GATTTCTGATC	ATTTTGT TTA
	TCGGAAAAGA	GGGAAGACCG	AGTATAAACG	CTAAAGCTAG	TAAACAAAT
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	TTTAAATACG
3601	CAACCAACAA	GGGCTTTACC	ATTTTATGTA	ACATTTT TCA	ACGCTGCAAA
	GTTGGTTGTT	CCCGAAATGG	TAAAATACAT	TGTAAAAAGT	TGCGACGTTT
3651	AATGTGTGTA	TTTCTATTTG	AAGATAAAAA	TCCTCAGCAA	AATCCACATT
	TTACACACAT	AAAGATAAAC	TTCTATTTTT	AGGAGTCGTT	TTAGGTGTAA
3701	GCACTGTCCT	TCAAAGATTA	GCCTTCTTTG	AACTAGTTAA	GACACTATTA
	CGTGACAGGA	AGTTTCTAAT	CGGAAGAAAC	TTGATCAATT	CTGTGATAAT
3751	AGCCAAGCCA	GTATCTCCCT	GTAATGAATT	CGTTTTTCTC	TTAATTTTCC
	TCGGTTCGGT	CATAGAGGGA	CATTACTTAA	GCAAAAAGAG	AATTAAAAGG
3801	CCTGTAATTT	ACACTGGGAG	AGCTGGGAAA	TATGTGGATG	TAAATTTCTC
	GGACATTAAA	TGTGACCCTC	TCGACCCTTT	ATACACCTAC	ATTTAAAGAG
3851	AGCCACAGAG	ATGCAAAGTT	ATACTGTGGG	GAAAAAAAAC	TTGAGTTAAA
	TCGGTGTCTC	TACGTTTCAA	TATGACACCC	CTTTTTTTTG	AACTCAATTT
3901	TCCTTACATA	TTTTAGGTTT	TCATTAACTT	ACCAATGTAG	TTTTGTGGA
	AGGAATGTAT	AAAATCCAAA	AGTAATTGAA	TGGTTACATC	AAAACAACCT
3951	GGCCATTTTT	TTTATTGCAG	ACTTGAAGAG	CTATTACTAG	AAAAATGCAT
	CCGGTAAAAA	AAATAACGTC	TGAACCTCTC	GATAATGATC	TTTTTACGTA
4001	GACAGTTAAG	GTAAGTTTGC	ATGACACAAA	AAAGGTAACT	AAATACAAAT
	CTGTCAATTC	CATTCAAACG	TACTGTGTTT	TTTCCATTGA	TTTATGTTTA
4051	TCTGTTTGGA	TTCCAACCCC	CAAGTAGAGA	GCGCACACTT	TCAAACGTGA
	AGACAAACCT	AAGGTTGGGG	GTTTCATCTCT	CGCGTGTGAA	AGTTTGC ACT
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	TTTTGTCTGAA	GTGTACCGAA	TTTTATCCCC	TGGTTACAGA

FIG. 27(contd.) 70/93

4251	TTTCCAATCT	AAGTCCCATT	TATAATAAAG	TCCATGTTCC	ATTTTTTAAAG
	AAAGGTTAGA	TTCAGGGTAA	ATATTATTTT	AGGTACAAGG	TAAAAATTTC
4301	GACAATCCTT	TCGGTTTAAA	ACCAGGCACG	ATTACCCAAA	CAACTCACAA
	CTGTTAGGAA	AGCCAAATTT	TGGTCCGTGC	TAATGGGTTT	GTTGAGTGTT
4351	CGGTAAAGCA	CTGTGAATCT	TCTCTGTTCT	GCAATCCCAA	CTTGGTTTCT
	GCCATTTCGT	GACACTTAGA	AGAGACAAGA	CGTTAGGGTT	GAACCAAAGA
4401	GCTCAGAAAC	CCTCCCTCTT	TCCAATCGGT	AATTAAATAA	CAAAAGGAAA
	CGAGTCTTTG	GGAGGGAGAA	AGGTTAGCCA	TTAATTTATT	GTTTTCTCTT
4451	AAACTTAAGA	TGCTTCAACC	CCGTTTTCGTG	ACACTTTGAA	AAAAGAATCA
	TTTGAATTCT	ACGAAGTTGG	GGCAAAGCAC	TGTGAAACTT	TTTTCTTAGT
4501	CCTCTTGCAA	ACACCCGCTC	CCGACCCCCG	CCGCTGAAGC	CCGGCGTCCA
	GGAGAACGTT	TGTGGGCGAG	GGCTGGGGGC	GGCGACTTCG	GGCCGCAGGT
4551	GAGGCCTAAG	CGCGGGTGCC	CGCCCCCACC	CGGGAGCGCG	GGCCTCGTGG
	CTCCGGATTC	GCGCCCACGG	GCGGGGGTGG	GCCCTCGCGC	CCGGAGCACC
4601	TCAGCGCATC	CGCGGGGAGA	AACAAAGGCC	GCGGCACGGG	GGCTCAAGGG
	AGTCGCGTAG	GCGCCCCTCT	TTGTTTCCGG	CGCCGTGCCC	CCGAGTTCCT
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	GGGTGCGGGC
4801	CGTGGCGCTG	GGGGAGGGGG	CGCCTCCGCC	GGAACGCGGG	TGGGGGAGGG
	GCACCGCGAC	CCCCTCCCCC	GCGGAGGCGG	CCTTGCGCCC	ACCCCTCTCC
4851	GAGGGGGAAA	TGCGCTTTGT	CTCGAAATGG	GGCAACCGTC	GCCACAGCTC
	CTCCCCCTTT	ACGCGAAACA	GAGCTTTACC	CCGTTGGCAG	CGGTGTCGAG
4901	CCTACCCCTT	CGAGGGCAGA	GCAGTCCCCC	CACTAACTAC	CGGGCTGGCC
	GGATGGGGGA	GCTCCCGTCT	CGTCAGGGGG	GTGATTGATG	GCCCGACCGG
4951	GCGCGCCAGG	CCAGCCGCGA	GGCCACCGCC	CGACCCTCCA	CTCCTTCCCG
	CGCGCGGTCC	GGTCGGCGCT	CCGGTGGCGG	GCTGGGAGGT	GAGGAAGGGC
5001	CAGCTCCCGG	CGCGGGGTCC	GGCGAGAAGG	GGAGGGGAGG	GGAGCGGAGA
	GTCGAGGGCC	GCGCCCCAGG	CCGCTCTTCC	CCTCCCCCTC	CCTCGCCTCT
5051	ACCGGGCCCC	CGGGACGCGT	GTGGCATCTG	AAGCACCACC	AGCGAGCGAG
	TGGCCCGGGG	GCCCTGCGCA	CACCGTAGAC	TTCGTGGTGG	TCGCTCGCTC
5101	AGCTAGAGAG	AAGGAAAGCC	ACCGACTTCA	CCGCCTCCGA	GCTGCTCCGG
	TCGATCTCTC	TTCCTTTCGG	TGGCTGAAGT	GGCGGAGGCT	CGACGAGGCC
5151	GTCGCGGGTC	TGCAGCGTCT	CCGGCCCTCC	GCGCCTACAG	CTCAAGCCAC
	CAGCGCCAG	ACGTCGCAGA	GGCCGGGAGG	CGCGGATGTC	GAGTTCGGTG
5201	ATCCGAAGGG	GGAGGGAGCC	GGGAGCTGCG	CGCGGGGCCG	CCGGGGGGAG
	TAGGCTTCCC	CCTCCCTCGG	CCCTCGACGC	GCGCCCCGGC	GGCCCCCTC
5251	GGGTGGCACC	GCCACGCGCG	GGCGGCCACG	AAGGGCGGGG	CAGCGGGCGC
	CCCACCGTGG	CGGGTGCGGC	CCGCCGGTGC	TTCCCGCCCC	GTCGCCCCGC
5301	GCGCGCGGCG	GGGGGAGGGG	CCGGCGCCCG	GCCCGCTGGG	AATTGGGGCC
	CGCGCGCCGC	CCCCCTCCCC	GGCCGCGGCG	CGGGCGACCC	TTAACCCCGG
5351	CTAGGGGGAG	GGCGGAGGCG	CCGACGACCG	CGGCACTTAC	CGTTCGCGGC
	GATCCCCCTC	CCGCCTCCGC	GGCTGCTGGC	GCCGTGAATG	GCAAGCGCCG
5401	GTGGCGCCCC	GTGGTCCCCA	AGGGGAGGGA	AGGGGGAGGC	GGGGCGAGGA
	CACCGCGGGC	CACCAGGGGT	TCCCCTCCCT	TCCCCCTCCG	CCCCGCTCCT
5451	CAGTGACCGG	AGTCTCCTCA	GCGGTGGCTT	TTCTGCTTGG	CAGCCTCAGC
	GTCACTGGCC	TCAGAGGAGT	CGCCACCGAA	AAGACGAACC	GTCGGAGTCG
5501	GGCTGGCGCC	AAAACCGGAC	TCCGCCCACT	TCCTCGCCCC	CCGGTGCAG
	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	GGCGTGAACA
	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	GCCCCTCCCG	TCGGCGTTAA
	TCCCCCGACC	TAGAGGTCCG	ACGCCTAATT	CGGGGAGGGC	AGCCGCAATT
5901	TTTCAAAC TG	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
	GTAAC TCGC	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	GCGGTCTTGC	TCAGGCCAAG
6301	GTGTTCTGTC	GCGGAGATCT	CTCTCATCTC	GCTCGGCTGC	GGGAAATCGG
	CACAAGCAGG	CGCCTCTAGA	GAGAGTAGAG	CGAGCCGACG	CCCTTTAGCC
6351	GCTGAAGCGA	CTGAGTCCGC	GATGGAGGTA	ACGGGTTTGA	AATCAATGAG
	CGACTTCGCT	GACTCAGGCG	CTACCTCCAT	TGCCCAAAC T	TTAGTTACTC
6401	TTATTGAAAA	GGGCATGGCG	AGGCCGTTGG	CGCCTCAGTG	GAAGTCGGCC
	AATAACTTTT	CCCGTACCGC	TCCGGCAACC	GCGGAGTCAC	CTTCAGCCGG
6451	AGCCGCCTCC	GTGGGAGAGA	GGCAGGAAAT	CGGACCAATT	CAGTAGCAGT
	TCGGCGGAGG	CACCCTCTCT	CCGTCCTTTA	GCCTGGTTAA	GTCATCGTCA
6501	GGGGCTTAAG	GTTTATGAAC	GGGGTCTTGA	GCGGAGGCCT	GAGCGTACAA
	CCCCGAATTC	CAAATACTTG	CCCCAGAACT	CGCCTCCGGA	CTCGCATGTT
6551	ACAGCTTCCC	CACCCTCAGC	CTCCCGGCGC	CATTTCCCTT	CACTGGGGGT
	TGT CGAAGG	GTGGGAGTCG	GAGGGCCGCG	GTAAAGGGAA	GTGACCCCCA
6601	GGGGGATGGG	GAGCTTTCAC	ATGGCGGACG	CTGCCCCGCT	GGGGTGAAAG
	CCCCCTACCC	CTCGAAAGTG	TACCGCCTGC	GACGGGGCGA	CCCCACTTTC
6651	TGGGGCGCGG	AGGCGGGAAT	TCTTATTCCC	TTTCTAAAGC	ACGCTGCTTC
	ACCCCGCGCC	TCCGCCCTTA	AGAATAAGGG	AAAGATTTCG	TGCGACGAAG
6701	GGGGGCCACG	GCGTCTCCTC	GGCGAGCGTT	TCGGCGGGCA	GCAGGTCTTC
	CCCCCGGTGC	CGCAGAGGAG	CCGCTCGCAA	AGCCGCCCGT	CGTCCAGGAG
6751	GTGAGCGAGG	CTGCGGAGCT	TCCCCTCCCC	CTCTCTCCCG	GGAACCGATT
	CACTCGCTCC	GACGCCTCGA	AGGGGAGGGG	GAGAGAGGGC	CCTTGGCTAA
6801	TGGCGGCCGC	CATTTTCATG	GCTCGCCTTC	CTCTCAGCGT	TTTCCTTATA
	ACCGCCGGCG	GTAAAAGTAC	CGAGCGGAAG	GAGAGTCGCA	AAAGGAATAT
6851	ACTCTTTTAT	TTTCTTAGTG	TGCTTTCTCT	ATCAAGAAGT	AGAAGTGGTT
	TGAGAAAATA	AAAGAATCAC	ACGAAAGAGA	TAGTTCTTCA	TCTTCACCAA
6901	AACTATTTTT	TTTTTCTTCT	CGGGCTGTTT	TCATATCGTT	TCGAGGTGGA
	TTGATAAAAA	AAAAAGAAGA	GCCCGACAAA	AGTATAGCAA	AGCTCCACCT
6951	TTTGGAGTGT	TTTGTGAGCT	TGGATCTTTA	GAGTCCTGCG	CACCTCATTA
	AAACCTCACA	AAACACTCGA	ACCTAGAAAT	CTCAGGACGC	GTGGAGTAAT
7001	AAGGCGCTCA	GCCTTCCCCCT	CGATGAAATG	GCGCCATTGC	GTTCCGGAAGC
	TTCCGCGAGT	CGGAAGGGGA	GCTACTTTAC	CGCGGTAACG	CAAGCCTTCG
7051	CACACCGAAG	AGCGGGGAGG	GGGGGTGCTC	CGGGTTTGCG	GGCCCGGTTT
	GTGTGGCTTC	TCGCCCTCC	CCCCCACGAG	GCCCAAACGC	CCGGGCCAAA
7101	CAGAGAAGAT	ATCACCACCC	AGGGCGTCGG	GCCGGGTTCA	ATGCGAGCCG
	GTCTCTTCTA	TAGTGGTGGG	TCCCGCAGCC	CGGCCCAAGT	TACGCTCGGC

FIG. 27(contd.)

72/93

7151	TAGGACAAAG	AAACCATTTT	ATGTTTTTCC	TGTCTTTTTT	TTCCTTTGAG
	ATCCTGTTTC	TTTGGTAAAA	TACAAAAAAG	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	GGGATATTCT	CGTCAGAATC	CTTTACTGTT
	CACCAAAAGA	CCCGCCCCCT	CCCTATAAGC	GCAGTCTTAG	GAAATGACAA
7351	CTTAAGGATT	CCGTTTAAGT	TGTAGAGCTG	ACTCATTTTA	AGTAATGTTG
	GAATTCCTAA	GGCAAATTCA	ACATCTCGAC	TGAGTAAAAT	TCATTACAAC
7401	TTACTGAGAA	GTTTAACCCCT	TACGGGACAG	ATCCATGGAC	CTTTATAGAT
	AATGACTCTT	CAAATTGGGA	ATGCCCTGTC	TAGGTACCTG	GAAATATCTA
7451	GATTACGAGG	AAAGTGAAAT	AACGATTTTG	TCCTTAGTTA	TACTTCGATT
	CTAATGCTCC	TTTCACTTTA	TTGCTAAAAC	AGGAATCAAT	ATGAAGCTAA
7501	AAAACATGGC	TTCAGAGGCT	CCTTCCTGTA	ATGCGTATGG	ATTGATGTGC
	TTTTGTACCG	AAGTCTCCGA	GGAAGGACAT	TACGCATACC	TAACTACACG
7551	AAAAC TGTTT	TGGGCCTGGG	CCGCTCTGTA	TTTGAAC TTT	GTTACTTTTC
	TTTTGACAAA	ACCCGGACCC	GGCGAGACAT	AAACTTGAAA	CAATGAAAAG
7601	TCATTTTGTT	TGCAATCTTG	GTTGAACATT	ACATTGATAA	GCATAAGGTC
	AGTAAAACAA	ACGTTAGAAC	CAACTTGTA	TGTAAC TATT	CGTATTCCAG
7651	TCAAGCGAAG	GGGGTCTACC	TGGTTATTTT	TCTTTGACCC	TAAGCACGTT
	AGTTCGCTTC	CCCCAGATGG	ACCAATAAAA	AGAAACTGGG	ATTCGTGCAA
7701	TATAAAATAA	CATTGTTTAA	AATCGATAGT	GGACATCGGG	TAAGTTTGGA
	ATATTTTATT	GTAACAAATT	TTAGCTATCA	CCTGTAGCCC	ATTCAAACCT
7751	TAAATTGTGA	GGTAAGTAAT	GAGTTTTTGC	TTTTTGTTAG	TGATTTGTAA
	ATTTAACACT	CCATTCATTA	CTCAAAAACG	AAAAACAATC	ACTAAACATT
7801	AACTTGTTAT	AAATGTACAT	TATCCGTAAT	TTCAGTTTAG	AGATAACCTA
	TTGAACAATA	TTTACATGTA	ATAGGCATTA	AAGTCAAATC	TCTATTGGAT
7851	TGTGCTGACG	ACAATTAAGA	ATAAAA ACTA	GCTGAAAAAA	TGAAAATAAC
	ACACGACTGC	TGTTAATTCT	TATTTTTTGAT	CGACTTTTTT	ACTTTTATTG
7901	TATCGTGACA	AGTAACCATT	TCAAAAGACT	GCTTTGTGTC	TCATAGGAGC
	ATAGCACTGT	TCATTGGTAA	AGTTTTCTGA	CGAAACACAG	AGTATCCTCG
7951	TAGTTTGATC	ATTT CAGTTA	ATTTTTTTCTT	TAATTTTTTAC	GAGTCATGAA
	ATCAAAC TAG	TAAAGTCAAT	TAAAAAAGAA	ATTA AAAATG	CTCAGTACTT
8001	AACTACAGGA	AAAAAAATCT	GAAGTGGGTT	TTACCACTAC	TTTTTAGGAG
	TTGATGTCCT	TTTTTTT TAGA	CTTGACCCAA	AATGGTGATG	AAAAATCCTC
8051	TTGGGAGCAT	GCGAATGGAG	GGAGAGCTCC	GTAGAACTGG	GATGAGAGCA
	AACCCTCGTA	CGCTTACCTC	CCTCTCGAGG	CATCTTGACC	CTACTCTCGT
8101	GCAATTAATG	CTGCTTGCTA	GGAACAAAAA	ATAATTGATT	GAAAATTACG
	CGTTAATTAC	GACGAACGAT	CCTTGTTTTT	TATTA ACTAA	CTTTTAATGC
8151	TGTGACTTTT	TAGTTTG CAT	TATGCGTTTG	TAGCAGTTGG	TCCTGGATAT
	ACACTGAAAA	ATCAAACGTA	ATACGCAAAC	ATCGTCAACC	AGGACCTATA
8201	CACTTTCTCT	CGTTTGAGGT	TTTTTAACTT	AGTTAACTTT	TAAGACAGGT
	GTGAAAGAGA	GCAA ACTCCA	AAAAATTGGA	TCAATTGAAA	ATTCTGTCCA
8251	TTCCTTAACA	TTCATAAGTG	CCCAGAATAC	AGCTGTGTAG	TACAGCATAT
	AAGGAATTGT	AAGTATT CAC	GGGTCTTATG	TCGACACATC	ATGTCGTATA
8301	AAAGATTTCA	GCTCTGAGGT	TTTTCTTATT	GA CTTGGA	ATTGTTTTGT
	TTTCTAAAGT	CGAGACTCCA	AAAAGGATAA	CTGAACCTTT	TAACAAAACA
8351	GCCTGTGCT	TGCCACATGG	CCAATCAAGT	AAGCTTATCG	ATACCGGTGG
	CGGACAGCGA	ACGGTGTACC	GGTTAGTTCA	TTCGAATAGC	TATGGCCACC
8401	CGCGCCAATT	GTTAATTAAG	ATCTGGCCCA	ATGGGCCGTA	CGAATTTGAG
	GCGCGGTTAA	CAATTAATTC	TAGACCGGGT	TACCCGGCAT	GCTTAAACTC
8451	GCGGAAAGAA	CCAGCTGTGG	AATGTGTGTC	AGTTAGGGTG	TGGAAAGTCC
	CGCCTTTCTT	GGTCGACACC	TTACACACAG	TCAATCCCAC	ACCTTTCAGG
8501	CCAGGCTCCC	CAGCAGGCAG	AAGTATGCAA	AGCATGCATC	TCAATTAGTC
	GGTCCGAGGG	GTCGTCCGTC	TTCATACGTT	TCGTACGTAG	AGTTAATCAG
8551	AGCAACCAGG	TGTGGAAAGT	CCCAGGCTC	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	TTTTTTATTT	ATGCAGAGGC	CGAGGCCGCC	TCGGCCTCTG
	GACTGATTAA	AAAAAATAAA	TACGTCTCCG	GCTCCGGCGG	AGCCGGAGAC
8751	AGCTATTCCA	GAAGTAGTGA	GGAGGCTTTT	TTGGAGGCCT	AGGCTTTTGC
	TCGATAAGGT	CTTCATCACT	CCTCCGAAAA	AACCTCCGGA	TCCGAAAACG
8801	AAAGATCGAT	CAAGAGACAG	GATGAGGATC	GTTTCGCATG	ATTGAACAAG
	TTTCTAGCTA	GTTCTCTGTC	CTACTCCTAG	CAAAGCGTAC	TAAC TTGTTC
8851	ATGGATTGCA	CGCAGGTTCT	CCGGCCGCTT	GGGTGGAGAG	GCTATTCGGC
	TACCTAACGT	GCGTCCAAGA	GGCCGGCGAA	CCCACCTCTC	CGATAAGCCG
8901	TATGACTGGG	CACAACAGAC	AATCGGCTGC	TCTGATGCCG	CCGTGTTCCG
	ATACTGACCC	GTGTTGTCTG	TTAGCCGACG	AGACTACGGC	GGCACAAGGC
8951	GCTGTCAGCG	CAGGGGCGCC	CGGTTCTTTT	TGTCAAGACC	GACCTGTCCG
	CGACAGTCGC	GTCCCCGCGG	GCCAAGAAAA	ACAGTTCTGG	CTGGACAGGC
9001	GTGCCCTGAA	TGAACTGCAA	GACGAGGCAG	CGCGGCTATC	GTGGCTGGCC
	CACGGGACTT	ACTTGACGTT	CTGCTCCGTC	GCGCCGATAG	CACCGACCGG
9051	ACGACGGGCG	TTCCTTGCGC	AGCTGTGCTC	GACGTTGTCA	CTGAAGCGGG
	TGCTGCCCGC	AAGGAACGCG	TCGACACGAG	CTGCAACAGT	GACTTCGCCC
9101	AAGGGACTGG	CTGCTATTGG	GCGAAGTGCC	GGGGCAGGAT	CTCCTGTCAT
	TTCCCTGACC	GACGATAACC	CGCTTCACGG	CCCCGTCTTA	GAGGACAGTA
9151	CTCACCTTGC	TCCTGCCGAG	AAAGTATCCA	TCATGGCTGA	TGCAATGCGG
	GAGTGGAACG	AGGACGGCTC	TTTCATAGGT	AGTACCGACT	ACGTTACGCC
9201	CGGCTGCATA	CGCTTGATCC	GGCTACCTGC	CCATTTCGACC	ACCAAGCGAA
	GCCGACGTAT	GCGAACTAGG	CCGATGGACG	GGTAAGCTGG	TGGTTCGCTT
9251	ACATCGCATC	GAGCGAGCAC	GTA CTGGAT	GGAAGCCGGT	CTTGTCGATC
	TGTAGCGTAG	CTCGCTCGTG	CATGAGCCTA	CCTTCGGCCA	GAACAGCTAG
9301	AGGATGATCA	AGAGCATCAG	GGGCTCGCGC	CAGCCGAACT	GTTTCGCCAGG
	TCCTACTAGT	TCTCGTAGTC	CCCGAGCGCG	GTCGGCTTGA	CAAGCGGTCC
9351	CTCAAGGCGA	GCATGCCCGA	CGGCGAGGAT	CTCGTCGTGA	CCCATGGCGA
	GAGTTCCGCT	CGTACGGGCT	GCCGCTCCTA	GAGCAGCACT	GGGTACCGCT
9401	TGCCTGCTTG	CCGAATATCA	TGGTGGAAAA	TGGCCGCTTT	TCTGGATTCA
	ACGGACGAAC	GGCTTATAGT	ACCACCTTTT	ACCGGCGAAA	AGACCTAAGT
9451	TCGACTGTGG	CCGGCTGGGT	GTGGCGGACC	GCTATCAGGA	CATAGCGTTG
	AGCTGACACC	GGCCGACCCA	CACCGCCTGG	CGATAGTCCT	GTATCGCAAC
9501	GCTACCCGTG	ATATTGCTGA	AGAGCTTGGC	GGCGAATGGG	CTGACCGCTT
	CGATGGGCAC	TATAACGACT	TCTCGAACCG	CCGCTTACCC	GA CTGGCGAA
9551	CCTCGTGCTT	TACGGTATCG	CCGCTCCCGA	TTCGCAGCGC	ATCGCCTTCT
	GGAGCACGAA	ATGCCATAGC	GGCGAGGGCT	AAGCGTCGCG	TAGCGGAAGA
9601	ATCGCCTTCT	TGACGAGTTC	TTCTGAGCGG	GACTCTGGGG	TTCGAAATGA
	TAGCGGAAGA	ACTGCTCAAG	AAGACTCGCC	CTGAGACCCC	AAGCTTTACT
9651	CCGACCAAGC	GACGCCCAAC	CTGCCATCAC	GAGATTTCTGA	TTCCACCGCC
	GGCTGGTTCTG	CTGCGGGTTG	GACGGTAGTG	CTCTAAAGCT	AAGGTGGCGG
9701	GCCTTCTATG	AAAGGTTGGG	CTTCGGAATC	GTTTTCCGGG	ACGCCGGCTG
	CGGAAGATAC	TTTCCAACCC	GAAGCCTTAG	CAAAAGGCC	TGCGGCCGAC
9751	GATGATCCTC	CAGCGCGGGG	ATCTCATGCT	GGAGTTCTTC	GCCCACCTTA
	CTACTAGGAG	GTCGCGCCCC	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	TGTTTCATAAA	CGCGGGGTTT	GGTCCCAGGG	CTGGCACTCT	GTCGATACCC
	ACAAGTATTT	GCGCCCCAAG	CCAGGGTCCC	GACCGTGAGA	CAGCTATGGG
9951	CACCGAGACC	CCATTGGGGC	CAATACGCCC	GCGTTTCTTC	CTTTTCCCCA
	GTGGCTCTGG	GGTAACCCCG	GTTATGCGGG	CGCAAAGAAG	GAAAAGGGGT
10001	CCCCACCCCC	CAAGTTCGGG	TGAAGGCCCA	GGGCTCGCAG	CCAACGTCGG
	GGGGTGGGGG	GTTCAAGCCC	ACTTCCGGGT	CCCGAGCGTC	GGTTGCAGCC

FIG. 27(contd.)

74/93

10051	GGCGGCAGGC	CCTGCCATAG	CCTCAAATTC	CTTAGGCTCG	AGGGCCGCCA
	CCGCCGTCCG	GGACGGTATC	GGAGTTTAAG	GAATCCGAGC	TCCCGGCGGT
10101	CCGCGGTGGA	GCTCCAGCTT	TTGTTCCCTT	TAGTGAGGGT	TAATTTTCGAG
	GGCGCCACCT	CGAGGTCGAA	AACAAGGGAA	ATCACTCCCA	ATTAAAGCTC
10151	CTTGGCGTAA	TCATGGTCAT	AGCTGTTTCC	TGTGTGAAAT	TGTTATCCGC
	GAACCGCATT	AGTACCAGTA	TCGACAAAGG	ACACACTTTA	ACAATAGGCG
10201	TCACAATTCC	ACACAACATA	CGAGCCGGAA	GCATAAAGTG	TAAAGCCTGG
	AGTGTTAAGG	TGTGTTGTAT	GCTCGGCCTT	CGTATTTTAC	ATTTTCGGACC
10251	GGTGCCTAAT	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	CCTTATTCCC
	CTTTTTTCCTT	CTCATACTCA	TAAGTTGTAA	AGGCACAGCG	GGAATAAGGG
10501	TTTTTTTGCGG	CATTTTGCCCT	TCCTGTTTTT	GCTCACCCAG	AAACGCTGGT
	AAAAAACGCC	GTAAACCGGA	AGGACAAAAA	CGAGTGGGTC	TTTGCGACCA
10551	GAAAGTAAAA	GATGCTGAAG	ATCAGTTGGG	TGCACGAGTG	GGTTACATCG
	CTTTCATTTT	CTACGACTTC	TAGTCAACCC	ACGTGCTCAC	CCAATGTAGC
10601	AACTGGATCT	CAACAGCGGT	AAGATCCTTG	AGAGTTTTTCG	CCCCGAAGAA
	TTGACCTAGA	GTTGTCGCCA	TTCTAGGAAC	TCTCAAAAGC	GGGGCTTCTT
10651	CGTTTTTCCAA	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	ATGTAACCTCG	CCTTGATCGT
	ATTGGCGAAA	AAACGTGTTG	TACCCCCTAG	TACATTGAGC	GGAAC TAGCA
10951	TGGGAACCGG	AGCTGAATGA	AGCCATACCA	AACGACGAGC	GTGACACCAC
	ACCCTTGGCC	TCGACTTACT	TCGGTATGGT	TTGCTGCTCG	CACTGTGGTG
11001	GATGCCTGTA	GCAATGGCAA	CAACGTTGCG	CAAAC TATTA	ACTGGCGAAC
	CTACGGACAT	CGTTACCGTT	GTTGCAACGC	GTTTGATAAT	TGACCGCTTG
11051	TACTTACTCT	AGCTTCCCGG	CAACAATTAA	TAGACTGGAT	GGAGGCGGAT
	ATGAATGAGA	TCGAAGGGCC	GTTGTTAATT	ATCTGACCTA	CCTCCGCCTA
11101	AAAGTTGCAG	GACCACTTCT	GCGCTCGGCC	CTTCCGGCTG	GCTGGTTTAT
	TTTCAACGTC	CTGGTGAAGA	CGCGAGCCGG	GAAGGCCGAC	CGACCAAATA
11151	TGCTGATAAA	TCTGGAGCCG	GTGAGCGTGG	GTCTCGCGGT	ATCATTGCAG
	ACGACTATTT	AGACCTCGGC	CACTCGCACC	CAGAGCGCCA	TAGTAACGTC
11201	CACTGGGGCC	AGATGGTAAG	CCCTCCCGTA	TCGTAGTTAT	CTACACGACG
	GTGACCCCGG	TCTACCATTC	GGGAGGGCAT	AGCATCAATA	GATGTGCTGC
11251	GGGAGTCAGG	CAACTATGGA	TGAACGAAAT	AGACAGATCG	CTGAGATAGG
	CCCTCAGTCC	GTTGATACCT	ACTTGCTTTA	TCTGTCTAGC	GACTCTATCC
11301	TGCCTCACTG	ATTAAGCATT	GGTAACTGTC	AGACTCGCGA	CACTGCATTA
	ACGGAGTGAC	TAATTCGTAA	CCATTGACAG	TCTGAGCGCT	GTGACGTAAT
11351	ATGAATCGGC	CAACGCGCGG	GGAGAGGCGG	TTTGCGTATT	GGGCGCTCTT
	TACTTAGCCG	GTTGCGCGCC	CCTCTCCGCC	AAACGCATAA	CCCGCGAGAA
11401	CCGCTTCCTC	GCTCACTGAC	TCGCTGCGCT	CGGTCGTTCG	GCTGCGGCGA
	GCGGAAGGAG	CGAGTGACTG	AGCGACGCGA	GCCAGCAAGC	CGACGCCGCT
11451	GCGGTATCAG	CTCACTCAAA	GGCGGTAATA	CGGTTATCCA	CAGAATCAGG
	CGCCATAGTC	GAGTGAGTTT	CCGCCATTAT	GCCAATAGGT	GTCTTAGTCC

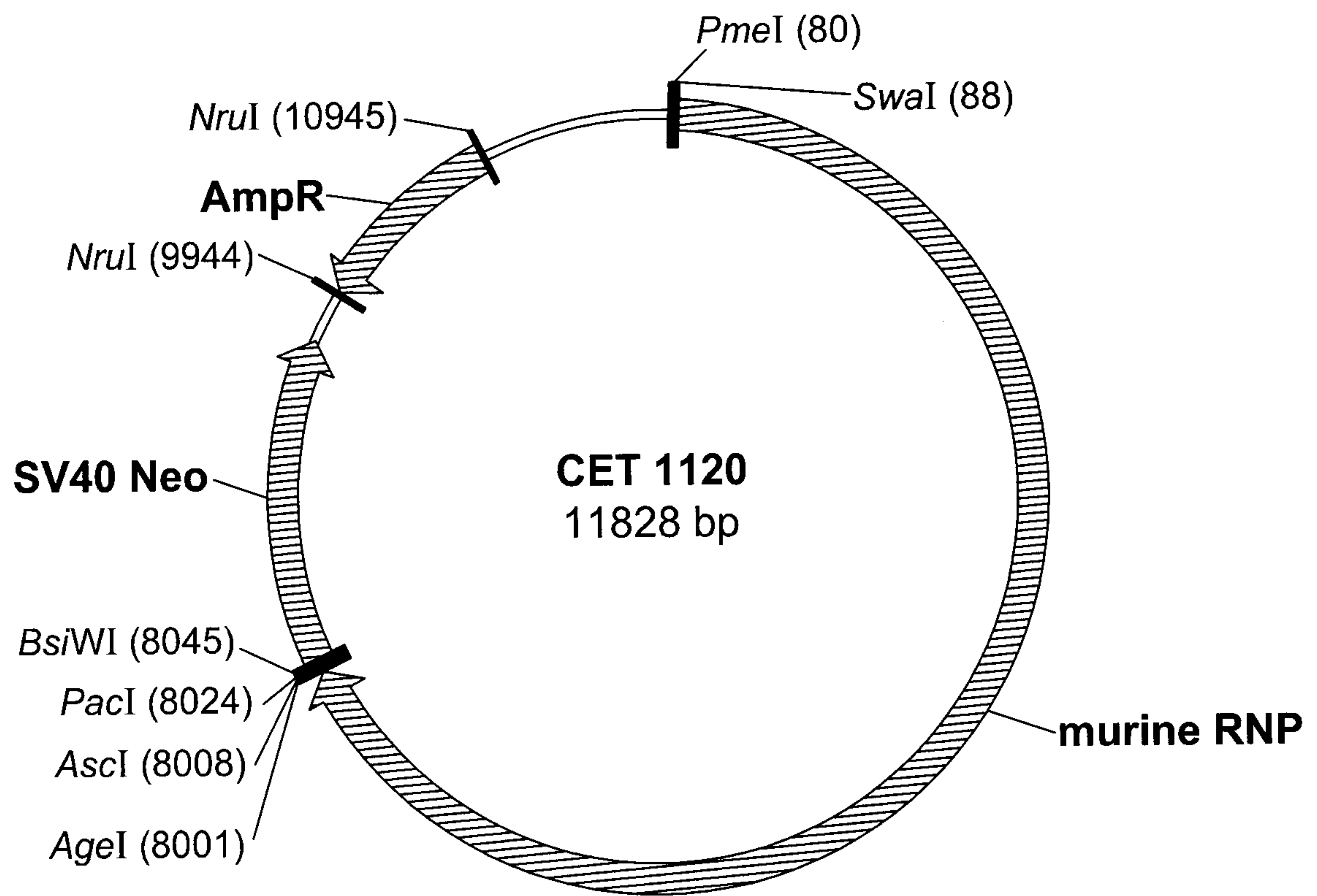
75/93

11501	GGATAACGCA	GGAAAGAACA	TGTGAGCAAA	AGGCCAGCAA	AAGGCCAGGA
	CCTATTGCGT	CCTTTCTTGT	ACACTCGTTT	TCCGGTCGTT	TTCCGGTCCT
11551	ACCGTAAAAA	GGCCGCGTTG	CTGGCGTTTT	TCCATAGGCT	CCGCCCCCCT
	TGGCATTTTT	CCGGCGCAAC	GACCGCAAAA	AGGTATCCGA	GGCGGGGGGA
11601	GACGAGCATC	ACAAAAATCG	ACGCTCAAGT	CAGAGGTGGC	GAAACCCGAC
	CTGCTCGTAG	TGTTTTTAGC	TGCGAGTTCA	GTCTCCACCG	CTTTGGGCTG
11651	AGGACTATAA	AGATACCAGG	CGTTTCCCCC	TGGAAGCTCC	CTCGTGCGCT
	TCCTGATATT	TCTATGGTCC	GCAAAGGGGG	ACCTTCGAGG	GAGCACGCGA
11701	CTCCTGTTCC	GACCCTGCCG	CTTACCGGAT	ACCTGTCCGC	CTTTCTCCCT
	GAGGACAAGG	CTGGGACGGC	GAATGGCCTA	TGGACAGGCG	GAAAGAGGGA
11751	TCGGGAAGCG	TGGCGCTTTC	TCATAGCTCA	CGCTGTAGGT	ATCTCAGTTC
	AGCCCTTCGC	ACCGCGAAAG	AGTATCGAGT	GCGACATCCA	TAGAGTCAAG
11801	GGTGTAGGTC	GTTCGCTCCA	AGCTGGGCTG	TGTGCACGAA	CCCCCCGTTC
	CCACATCCAG	CAAGCGAGGT	TCGACCCGAC	ACACGTGCTT	GGGGGGCAAG
11851	AGCCCGACCG	CTGCGCCTTA	TCCGGTAACT	ATCGTCTTGA	GTCCAACCCG
	TCGGGCTGGC	GACGCGGAAT	AGGCCATTGA	TAGCAGAACT	CAGGTTGGGC
11901	GTAAGACACG	ACTTATCGCC	ACTGGCAGCA	GCCACTGGTA	ACAGGATTAG
	CATTCTGTGC	TGAATAGCGG	TGACCGTCGT	CGGTGACCAT	TGTCCTAATC
11951	CAGAGCGAGG	TATGTAGGCG	GTGCTACAGA	GTTCTTGAAG	TGGTGGCCTA
	GTCTCGCTCC	ATACATCCGC	CACGATGTCT	CAAGAACTTC	ACCACCGGAT
12001	ACTACGGCTA	CACTAGAAGG	ACAGTATTTG	GTATCTGCGC	TCTGCTGAAG
	TGATGCCGAT	GTGATCTTCC	TGTCATAAAC	CATAGACGCG	AGACGACTTC
12051	CCAGTTACCT	TCGGAAAAAG	AGTTGGTAGC	TCTTGATCCG	GCAAACAAAC
	GGTCAATGGA	AGCCTTTTTT	TCAACCATCG	AGAACTAGGC	CGTTTGTTTG
12101	CACCGCTGGT	AGCGGTGGTT	TTTTTGTTTG	CAAGCAGCAG	ATTACGCGCA
	GTGGCGACCA	TCGCCACCAA	AAAAACAAAC	GTTTCGTCGTC	TAATGCGCGT
12151	GAAAAAAAGG	ATCTCAAGAA	GATCCTTTGA	TCTTTTCTAC	GGGGTCTGAC
	CTTTTTTTTCC	TAGAGTTCTT	CTAGGAAACT	AGAAAAGATG	CCCCAGACTG
12201	GCTCAGTGGA	ACGAAAATC	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 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 TTGTTGTTTTT
251 ACCCCAACCA TCTGAGAAAT AATCTTCTCC TTTCCCAGCT TTATTCCCAG
   TGGGGTTGGT AGACTCTTTA TTAGAAGAGG AAAGGGTCGA AATAAGGGTC
301 GATTCTACAT GACCAAATTA CCAGAGTCAC CACTCATTTT AATCACAACA
   CTAAGATGTA CTGGTTTAAT GGTCTCAGTG GTGAGTAAAA TTAGTGTTGT
351 TAGTGTCAAA TAACTAGAAA ACATGAGACA ACAATGGAGA GCTGAGTAAC
   ATCACAGTTT ATTGATCTTT TGTACTCTGT TGTTACCTCT CGACTCATTG
401 TATTAGTAGT AGTACTTTAC CAGAGAATGG CCTCTATAGG CTCACATGTA
   ATAATCATCA TCATGAAATG GTCTCTTACC GGAGATATCC GAGTGTACAT
451 GGAATGGTTG GTCCCCAGGT GGTAGGTAGA GCTGTTTGAG GATTACGTGG
   CCTTACCAAC CAGGGGTCCA CCATCCATCT CGACAAACTC CTAATGCACC
501 CCTTCTTGGA TGGGGGGGTG GGTGGGGTG GGAGGGTTGG GTGGTGGGTA
   GGAAGAACCT ACCCCCCACC CCCACCCAC CCTCCCAACC CACCACCCAT
551 CTTAAGAGGT TTCAAAGTC AATATTGTTT GCATTTAGCT CTTCTTGTA
   GAATTCTCCA AAGTTTTTCA 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 GGAAC TACT TGTAGACCA 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
   CTTACCCGGA ACCAATACCC TCGAATAGTG TCGTTATCTT GTCGTTAATA
951 GACTGGAGTA TGATAGTTAA AAACAAGCAA GCAAGCAAGC AAACACACAC
   CTGACCTCAT ACTATCAATT TTTGTTCGTT CGTTCGTTTCG TTTGTGTGTG
1001 ACCAAAACAA CAAAACCCCA AGACAGAGTC ACATGTAGCC CAGGCTAGCC
   TGGTTTTGTT GTTTTGGGGT TCTGTCTCAG TGTACATCGG GTCCGATCGG
1051 TCCAAATTCA CTATATAACT GAAGAAGACC CCTAATTCCC ATTCCTCTAG
   AGGTTTAAGT GATATATTGA CTTCTTCTGG GGATTAAGGG TAAGGAGATC
1101 AATCTATACC TCAAGTACTG AATGGCTTGG TTCACAATAC CCCACTAAAT
   TTAGATATGG AGTTCATGAC TTACCGAACC AAGTGTTATG GGGTGATTTA
1151 GATTGGTCTT ACTAAGTGCA ACAAGGTAAA CCTAAACTT CAGCCCTCAG
   CTAACCAGAA TGATTCACGT TGTTCCATTT GGATTTTGAA GTCGGGAGTC
1201 ACATCCCTTT TCCAGTATCA ATTTATAAAA TTAGATCCCA AGGATAAAAA
   TGTAGGGAAA AGGTCATAGT TAAATATTTT AATCTAGGGT TCCTATTTTT
1251 TTAATTGTAA AGTAAAATCA GAGTTCTAGC ATCAACTACA GGCTCAACCA
   AATTAACATT TCATTTTAGT CTCAAGATCG TAGTTGATGT CCGAGTTGGT
1301 TGGGGACCAC AAATAAATAA AAAGGGATAA GACTGGCTTC CCCATAATTA
   ACCCCTGGTG TTTATTTGAT TTTCCCTATT CTGACCGAAG GGGTATTAAT

```

FIG. 29(contd.) 78/93

1351	TTACATTTAG	ATAATTTTCC	TGACTACTCA	ACAAAGCTAA	AATATCACCA
	AATGTAAATC	TATTTAAAGG	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	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	ATTTCCCTAG
1851	ACAAAAATAA	CCCCCTATCA	AATGCCTAGT	CCCTCTAACC	ATCACCAAGT
	TGTTTTTTATT	GGGGGATAGT	TTACGGATCA	GGGAGATTGG	TAGTGGTTCA
1901	GAAGGATCAC	GCAGGAAAAA	AAAAATCACC	AGCAGCACCT	CAGAACCAGG
	CTTCCTAGTG	CGTCCTTTTT	TTTTTAGTGG	TCGTCGTGGA	GTCTTGGTCC
1951	ATACTCAGTC	CATCAGCATC	CAGGGCCATA	CCCACACTCA	CAGCATCTCC
	TATGAGTCAG	GTAGTCGTAG	GTCCCGGTAT	GGGTGTGAGT	GTCGTAGAGG
2001	ACAGTTTACC	AGATGATTCA	TGCTTATCAC	TGTATTGGGT	CATCTAAGAG
	TGTCAAATGG	TCTACTAAGT	ACGAATAGTG	ACATAACCCA	GTAGATTCTC
2051	TGACCATCAG	GGCTTCTGAT	CACAGAATCT	AGTCCACTTT	GCAGACCAGT
	ACTGGTAGTC	CCGAAGACTA	GTGTCTTAGA	TCAGGTGAAA	CGTCTGGTCA
2101	TGAAGTCATG	CACTATATGA	GATAGAAATA	CCCTCTTGCT	CATTTTGGTC
	ACTTCAGTAC	GTGATATACT	CTATCTTTAT	GGGAGAACGA	GTAACACCAG
2151	AGAAATTCAA	GGATAAAAAC	CCATGTTTTG	TTAATGCACA	CCTCCATATG
	TCTTTAAGTT	CCTATTTTTG	GGTACAAAAC	AATTACGTGT	GGAGGTATAC
2201	ATTGAGATCA	ATGTGTCCTA	ATTAATGTAG	AAACCACAAC	TGTAAATTTT
	TAACCTTAGT	TACACAGGAT	TAATTACATC	TTTGGTGTG	ACATTTAAAG
2251	ACTCTTTTGA	CATGAATCTT	TTTCTAGACA	GGGTCTTGGA	TGCAGCCCCG
	TGAGAAAAC	GTAAGTAGAA	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	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	TGTAATTCTT
	GTCTTAACTA	GATTGTCAGA	ATCTAGTGAG	GATCTGTCAA	ACATTAAGAA
2701	GCTCATGGCA	ACGTGAGCTC	TATCTAACTC	ACTCTCTGTG	CACTAATGAA
	CGAGTACCGT	TGCACTCGAG	ATAGATTGAG	TGAGAGACAC	GTGATTACTT
2751	TGCTCAGTGT	CTCCAGAAC	GCACAGCTTC	CAGGGTAATC	ATGCCAACCC
	ACGAGTCACA	GAGGTCTTGT	CGTGTCGAAG	GTCCCATTAG	TACGGTTGGG

FIG. 29(contd.) 79/93

2801	ACAAGACTTT	TATAGAGCTG	TCCACGACTC	TTCCCCCATT	CAGCTCATTA
	TGTTCTGAAA	ATATCTCGAC	AGGTGCTGAG	AAGGGGGTAA	GTCGAGTAAT
2851	ACAATATGAT	GGAGCTCCTG	TGTGGAAATC	AAGGCACACT	CTGGTAGAAA
	TGTTATACTA	CCTCGAGGAC	ACACCTTTAG	TTCCGTGTGA	GACCATCTTT
2901	CTTGTTTTTT	CTTTCCACTT	TTCCTTGGGC	TCTGAAGATT	GAGCTGTTTT
	GAACAAAAAA	GAAAGGTGAA	AAGGAACCCG	AGACTTCTAA	CTCGACAAAA
2951	ATAACCCACA	AACATGCATT	TTTTACCTCA	AAAGCATCCA	GCAAAAACCTG
	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	ACGACCATT	ATGTTCTCTGA
3101	TAGGTAGGCT	GGCTTCTAAC	AGCAATTCAC	CCACTTATGA	TGGGATTAAA
	ATCCATCCGA	CCGAAGATTG	TCGTTAAGTG	GGTGAATACT	ACCCTAATTT
3151	GGAAGGCACA	ACCATGTCCA	CCACAGGTTC	TAGCTCCCCC	ACCCACACGC
	CCTTCCGTGT	TGGTACAGGT	GGTGTCCAAG	ATCGAGGGGG	TGGGTGTGCG
3201	CCAGAGAGGG	TTTTTCTGTG	TAGCTCTGAC	TATTCTGGAA	TTCACACTGC
	GGTCTCTCCC	AAAAAGACAC	ATCGAGACTG	ATAAGACCTT	AAGTGTGACG
3251	AGACCAGGCT	GGTCTCGAAC	TCAGAGATCC	ACCACCACAT	GGTTTCTTAA
	TCTGGTCCGA	CCAGAGCTTG	AGTCTCTAGG	TGGTGGTGTA	CCAAAGAATT
3301	TTGTAATTTT	AAAGAAAAAA	AAAAATCCTT	CAGTTAAGAT	TCTTATGTTT
	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	TCTTTTTTTT	ACATCCTGTC
3451	TTAAAATTTT	ATGACACACA	AAAGGCAGCT	ACAAGTTTTG	TGTGGATTTC
	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	CTACTTTCCT
3701	AATGACTATC	TTTTGTTAGC	AAATTACCAT	AAGATCTTTT	TCTTTAGATT
	TTACTGATAG	AAAACAATCG	TTTAATGGTA	TTCTAGAAAA	AGAAATCTAA
3751	TCTGAATACT	CCAAGGAGCT	CATATAATTC	CATCCTTATT	TTTTTCAGAGG
	AGACTTATGA	GGTTCCTCGA	GTATATTAAG	GTAGGAATAA	AAAAGTCTCC
3801	CCCTCCCTGT	TCAATCACGG	TATAAAAAAA	GGAACACATT	AAGATGTCCC
	GGGAGGGACA	AGTTAGTGCC	ATATTTTTTT	CCTTGTGTAA	TTCTACAGGG
3851	AGTCCTATTT	TCTGGCTTTT	TTTTTCCGGG	GGTGGTGGTG	CGGTAATCAC
	TCAGGATAAA	AGACCGAAAA	AAAAAGGCCC	CCACCACCAC	GCCATTAGTG
3901	TCTCTATAGT	CCAGTCTGGG	CTTCAACGCC	TGGCAATCCC	CAGCCTCAAG
	AGAGATATCA	GGTCAGACCC	GAAGTTGCGG	ACCGTTAGGG	GTCGGAGTTC
3951	CTCCCAAGTA	CTGTCCTGAT	AAGGATAGAA	GGAGTCGACC	TCCTTCACGC
	GAGGGTTCAT	GACAGGACTA	TTCTATCTT	CCTCAGCTGG	AGGAAGTGCG
4001	TCCCCTCCGA	GGAGGGCTCC	TTCCCAGCTC	CATTCCCCGG	TCGGGAGCCC
	AGGGGAGGCT	CCTCCCGAGG	AAGGGTCGAG	GTAAGGGGCC	AGCCCTCGGG
4051	GTCCCCCACC	CGAGAGCGCG	GGCCTCGTGG	TCAGCGCCTC	CGCGGGGAGA
	CAGGGGGTGG	GCTCTCGCGC	CCGGAGCACC	AGTCGCGGAG	GCGCCCCTCT
4101	AACAAAGGCG	GCGGCGGGGG	CTCAAGGGCA	CTGCGCCACG	GGCCCGCGCC
	TTGTTTCCGC	CGCCGCCCCC	GAGTTCCCGT	GACGCGGTGC	CCGGGCGCGG
4151	TCCCCCATCC	GGCGGCGGCC	ACGTAGCCGG	GAGCGCGCCG	CAGCCCGGAG
	AGGGGGTAGG	CCGCCGCCGG	TGCATCGGCC	CTCGCGCGGC	GTCGGGCCTC
4201	CCTCGGGCCT	CGCAGCTGCA	GAGCCTGAAC	CGCTCTCTCC	CTGCGGGCCT
	GGAGCCCCGA	GCGTCGACGT	CTCGGACTTG	GCGAGAGAGG	GACGCCCGGA

FIG. 29(contd.) 80/93

4251	GCGACGAGGC	TGGGGGAGGG	GAGGCCCGCG	CTTTGTCTGG	AGTCTCGGTA
	CGCTGCTCCG	ACCCCTCCC	CTCCGGGCGC	GAAACAGACC	TCAGAGCCAT
4301	GCTGTCATCC	GGCTCCCACC	CTCATGCACA	ATTGTCCCAT	CTCCCCACG
	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	TTCGAGCCTG	CGTGCCTCGT
4451	TCCTAAACCC	CACCACACGC	AAGATCGAAA	AAAAGCAAAG	GCACGAACTT
	AGGATTTGGG	GTGGTGTGCG	TTCTAGCTTT	TTTTCGTTTC	CGTGCTTGAA
4501	CACCGCTCCG	ATGCTCAGGG	CCGCGGATCC	TGCAGAGTCT	CCCGCCTGCG
	GTGGCGAGGC	TACGAGTCCC	GGCGCCTAGG	ACGTCTCAGA	GGGCGGACGC
4551	CGCTTCGGTT	CAGCCACATC	CGAGGGGAGG	GGGCGCGGGC	AGCTCCGCCG
	GCGAAGCCAA	GTCGGTGTAG	GCTCCCCTCC	CCCGCGCCCG	TCGAGGCGGC
4601	GGGGGGAGGG	GGAGCACCGC	CCACGCCCTG	GCCGCGCGGG	GCCCGCCGGG
	CCCCCTCCC	CCTCGTGGCG	GGTGCGGGAC	CGGCGCGCCC	CGGGCGGCCC
4651	AACGCGTCCT	GCGGGGGGCG	GCGCGCGCAA	TGCTCACCGT	CCGCGGCGTG
	TTGCGCAGGA	CGCCCCCGC	CGCGCGCGTT	ACGAGTGGCA	GGCGCCGCAC
4701	GCGCCCAAGG	GGTCTCCTGG	CTGGGGGGAG	GGGGGGGAAG	GCGGGCAGGA
	CGCGGGTCCC	CCAGAGGACC	GACCCCCCTC	CCCCCCTTC	CGCCCGTCCT
4751	AGGACCGCGG	AGGCCTCTCT	GCGTCTCGGA	GCGCGCCAAA	GCGGGGCTCC
	TCCTGGCGCC	TCCGGAGAGA	CGCAGAGCCT	CGCGCGGTTT	CGCCCCGAGG
4801	ACCCACCTCC	TTGCCCGGAT	CTTGAAGGCC	GGGGAGATAA	ACAGCGGGGT
	TGGGTGGAGG	AACGGGCCTA	GAACTTCCGG	CCCCTCTATT	TGTCGCCCCA
4851	TCTTTAAGCA	CCACCTCTCA	CTAGGCGCGG	GATCCCAAGG	CTTGTGGCAT
	AGAAATTCTG	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	GGATTCGGGG	AGGGCTGCCG	CAATAAAGTT
5101	ACTGCGCGAC	CGTTTCTCCG	CTCCCTACGC	GGAGGTGGGG	GCCGGACCTA
	TGACGCGCTG	GCAAAGAGGC	GAGGGATGCG	CCTCCACCCC	CGGCCTGGAT
5151	GTTCGGGACG	TAGTAACACG	CCGAGCGCGA	GCCTTCCGCA	ATTCACGGAA
	CAAGGCCTGC	ATCATTTGTG	GGCTCGCGCT	CGGAAGGCGT	TAAGTGCCTT
5201	CACAGTTGCG	CAAGTGATGT	AAAGCAGTCC	CGCTGTACCT	AAAGGGGGAG
	GTGTCAACGC	GTTCACTACA	TTTCGTCAGG	GCGACATGGA	TTTCCCCCTC
5251	TGTCACGTAC	TTGGCGTAAG	GAGAGTGTAG	GCCCTTCCCG	CCATTGGCGG
	ACAGTGCATG	AACCGCATTC	CTCTCACATC	CGGGAAGGGC	GGTAACCGCC
5301	CGGTTAGGGC	GTTTACGTAA	CGGCGTGACG	TAAGCGGAGA	CGCGTTAGTG
	GCCAATCCCG	CAAATGCATT	GCCGCACTGC	ATTCGCCTCT	GCGCAATCAC
5351	GGGGGAAGGT	TCTAGAAAAG	CGGCGGTCTC	GGCTCCAGCG	GCAGTAGCAG
	CCCCCTTCCA	AGATCTTTTC	GCCGCCAGAG	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
	GACTCAGGCG	CTACCTCCAT	TGCCCAAAC	TTAGTTACTC	AATAATTTTT
5601	TGGCATGGCG	AGGCCGTAGG	CACCGCAATG	GAAACCGGCC	ACCCGCCTCC
	ACCGTACCGC	TCCGGCATCC	GTGGCGTTAC	CTTTGGCCGG	TGGGCGGAGG
5651	GTGGTCCGGC	GGAGGGGATG	CGGCCACTCG	AGTGGCGGTT	GGCCTTGGCG
	CACCAGGCCG	CCTCCCCTAC	GCCGGTGAGC	TCACCGCCAA	CCGGAACCGC

FIG. 29(contd.) 81/93

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

FIG. 29(contd.) 82/93

7151	ATTGGCTTTT	TGAGCACATT	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	AATTCTTTAA	TACCCTTTAA	ATGACCTAAG	AGCTGGGTAT	GGTTTTTCTG
	TTAAGAAATT	ATGGGAAATT	TACTGGATTC	TCGACCCATA	CCAAAAGAC
7651	AATTTTTTTG	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	TGTAAGTTTT
	TGTGTTGATT	ATACGTCAAA	TTGTATACCA	AATTTAAACT	ACATTCAAAA
7801	TTTTCTCCCC	CCCAGAAAAC	TTTAGAAACT	GTTCTTTTGG	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
	TTTTCTTTGT	CAAGGCTTTC	GAGAAATAAC	CACCGAATTC	GAATAGCTAT
8001	CCGGTGGCGC	GCCAATTGTT	AATTAAGATC	TGGCCCAATG	GGCCGTACGA
	GGCCACCGCG	CGGTTAACAA	TTAATTCTAG	ACCGGGTTAC	CCGGCATGCT
8051	ATTTGAGGCG	GAAAGAACCA	GCTGTGGAAT	GTGTGTCAGT	TAGGGTGTGG
	TAAACTCCGC	CTTTCTTGGT	CGACACCTTA	CACACAGTCA	ATCCCACACC
8101	AAAGTCCCCA	GGCTCCCCAG	CAGGCAGAAG	TATGCAAAGC	ATGCATCTCA
	TTTCAGGGGT	CCGAGGGGTC	GTCCGTCTTC	ATACGTTTCG	TACGTAGAGT
8151	ATTAGTCAGC	AACCAGGTGT	GGAAAGTCCC	CAGGCTCCCC	AGCAGGCAGA
	TAATCAGTCG	TTGGTCCACA	CCTTTCAGGG	GTCCGAGGGG	TCGTCCGTCT
8201	AGTATGCAAA	GCATGCATCT	CAATTAGTCA	GCAACCATAG	TCCCGCCCTT
	TCATACGTTT	CGTACGTAGA	GTTAATCAGT	CGTTGGTATC	AGGGCGGGGA
8251	AACTCCGCCC	ATCCCGCCCC	TAACCTCCGC	CAGTTCCGCC	CATTCTCCGC
	TTGAGGCGGG	TAGGGCGGGG	ATTGAGGCGG	GTCAAGGCGG	GTAAGAGGCG
8301	CCCATGGCTG	ACTAATTTTT	TTTATTTATG	CAGAGGCCGA	GGCCGCCTCG
	GGGTACCGAC	TGATTAAAAA	AAATAAATAC	GTCTCCGGCT	CCGGCGGAGC
8351	GCCTCTGAGC	TATTCCAGAA	GTAGTGAGGA	GGCTTTTTTT	GAGGCCTAGG
	CGGAGACTCG	ATAAGGTCTT	CATCACTCCT	CCGAAAAAAC	CTCCGGATCC
8401	CTTTTGCAAA	GATCGATCAA	GAGACAGGAT	GAGGATCGTT	TCGCATGATT
	GAAAACGTTT	CTAGCTAGTT	CTCTGTCCTA	CTCCTAGCAA	AGCGTACTAA
8451	GAACAAGATG	GATTGCACGC	AGGTTCTCCG	GCCGCTTGGG	TGGAGAGGCT
	CTTGTTCTAC	CTAACGTGCG	TCCAAGAGGC	CGGCGAACCC	ACCTCTCCGA
8501	ATTCGGCTAT	GACTGGGCAC	AACAGACAAT	CGGCTGCTCT	GATGCCGCCG
	TAAGCCGATA	CTGACCCGTG	TTGTCTGTTA	GCCGACGAGA	CTACGGCGGC
8551	TGTTCCGGCT	GTCAGCGCAG	GGGCGCCCGG	TTCTTTTTGT	CAAGACCGAC
	ACAAGGCCGA	CAGTCGCGTC	CCCGCGGGCC	AAGAAAAACA	GTTCTGGCTG

FIG. 29(contd.) 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	TTCACGGCCC	CGTCCTAGAG
8751	CTGTCATCTC	ACCTTGCTCC	TGCCGAGAAA	GTATCCATCA	TGGCTGATGC
	GACAGTAGAG	TGGAACGAGG	ACGGCTCTTT	CATAGGTAGT	ACCGACTACG
8801	AATGCGGCGG	CTGCATACGC	TTGATCCGGC	TACCTGCCCA	TTCGACCACC
	TTACGCCGCC	GACGTATGCG	AACTAGGCCG	ATGGACGGGT	AAGCTGGTGG
8851	AAGCGAAACA	TCGCATCGAG	CGAGCACGTA	CTCGGATGGA	AGCCGGTCTT
	TTCGCTTTGT	AGCGTAGCTC	GCTCGTGCAT	GAGCCTACCT	TCGGCCAGAA
8901	GTCGATCAGG	ATGATCAAGA	GCATCAGGGG	CTCGCGCCAG	CCGAAC TGTT
	CAGCTAGTCC	TACTAGTTCT	CGTAGTCCCC	GAGCGCGGTC	GGCTTGACAA
8951	CGCCAGGCTC	AAGGCGAGCA	TGCCCGACGG	CGAGGATCTC	GTCGTGACCC
	GCGGTCCGAG	TTCCGCTCGT	ACGGGCTGCC	GCTCCTAGAG	CAGCACTGGG
9001	ATGGCGATGC	CTGCTTGCCG	AATATCATGG	TGGAAAATGG	CCGCTTTTCT
	TACCGCTACG	GACGAACGGC	TTATAGTACC	ACCTTTTACC	GGCGAAAAGA
9051	GGATTTCATC	ACTGTGGCCG	GCTGGGTGTG	GCGGACCGCT	ATCAGGACAT
	CCTAAGTAGC	TGACACCGGC	CGACCCACAC	CGCCTGGCGA	TAGTCCTGTA
9101	AGCGTTGGCT	ACCCGTGATA	TTGCTGAAGA	GCTTGGCGGC	GAATGGGCTG
	TCGCAACCGA	TGGGCACTAT	AACGACTTCT	CGAACCGCCG	CTTACCCGAC
9151	ACCGCTTCCT	CGTGCTTTAC	GGTATCGCCG	CTCCCGATTG	GCAGCGCATC
	TGGCGAAGGA	GCACGAAATG	CCATAGCGGC	GAGGGCTAAG	CGTCGCGTAG
9201	GCCTTCTATC	GCCTTCTTGA	CGAGTTCTTC	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	CACGGTGTTG
	TTGGGCGCGA	TACTGCCGTT	ATTTTTCTGT	CTTATTTTGC	GTGCCACAAC
9501	GGTCGTTTGT	TCATAAACGC	GGGGTTCGGT	CCCAGGGCTG	GCACTCTGTC
	CCAGCAAACA	AGTATTTGCG	CCCCAAGCCA	GGGTCCCGAC	CGTGAGACAG
9551	GATACCCAC	CGAGACCCCA	TTGGGGGCCAA	TACGCCCCGCG	TTTCTTCCTT
	CTATGGGGTG	GCTCTGGGGT	AACCCCGGTT	ATGCGGGCGC	AAAGAAGGAA
9601	TTCCCCACCC	CACCCCCCAA	GTTCCGGGTGA	AGGCCAGGGG	CTCGCAGCCA
	AAGGGGTGGG	GTGGGGGGTT	CAAGCCCACT	TCCGGGTCCC	GAGCGTCGGT
9651	ACGTCGGGGC	GGCAGGCCCT	GCCATAGCCT	CAAATTCCTT	AGGCTCGAGG
	TGCAGCCCCG	CCGTCCGGGA	CGGTATCGGA	GTTTAAGGAA	TCCGAGCTCC
9701	GCCGCCACCG	CGGTGGAGCT	CCAGCTTTTG	TTCCCTTTAG	TGAGGGTTAA
	CGGCGGTGGC	GCCACCTCGA	GGTCGAAAAC	AAGGGAAATC	ACTCCCAATT
9751	TTTCGAGCTT	GGCGTAATCA	TGGTCATAGC	TGTTTCCTGT	GTGAAATTGT
	AAAGCTCGAA	CCGCATTAGT	ACCAGTATCG	ACAAAGGACA	CACTTTAACA
9801	TATCCGCTCA	CAATTCCACA	CAACATACGA	GCCGGAAGCA	TAAAGTGTA
	ATAGGCGAGT	GTTAAGGTGT	GTTGTATGCT	CGGCCTTCGT	ATTTACATT
9851	AGCCTGGGGT	GCCTAATGAG	TGAGCTAACT	CACATTAATT	GCGTTGCGCT
	TCGGACCCCA	CGGATTACTC	ACTCGATTGA	GTGTAATTAA	CGCAACGCGA
9901	CACTGCCCGC	TTTCCAGTCG	GGAAACCTGT	CGTGCCAGCA	TCGCGAGCAC
	GTGACGGGCG	AAAGGTCAGC	CCTTTGGACA	GCACGGTCGT	AGCGCTCGTG
9951	TTTTTCGGGGA	AATGTGCGCG	GAACCCCTAT	TTGTTTATTT	TTCTAAATAC
	AAAAGCCCCT	TTACACGCGC	CTTGGGGATA	AACAAATAAA	AAGATTTATG
10001	ATTCAAATAT	GTATCCGCTC	ATGAGACAAT	AACCCTGATA	AATGCTTCAA
	TAAGTTTATA	CATAGGCGAG	TACTCTGTTA	TTGGGACTAT	TTACGAAGTT

FIG. 29(contd.)

84/93

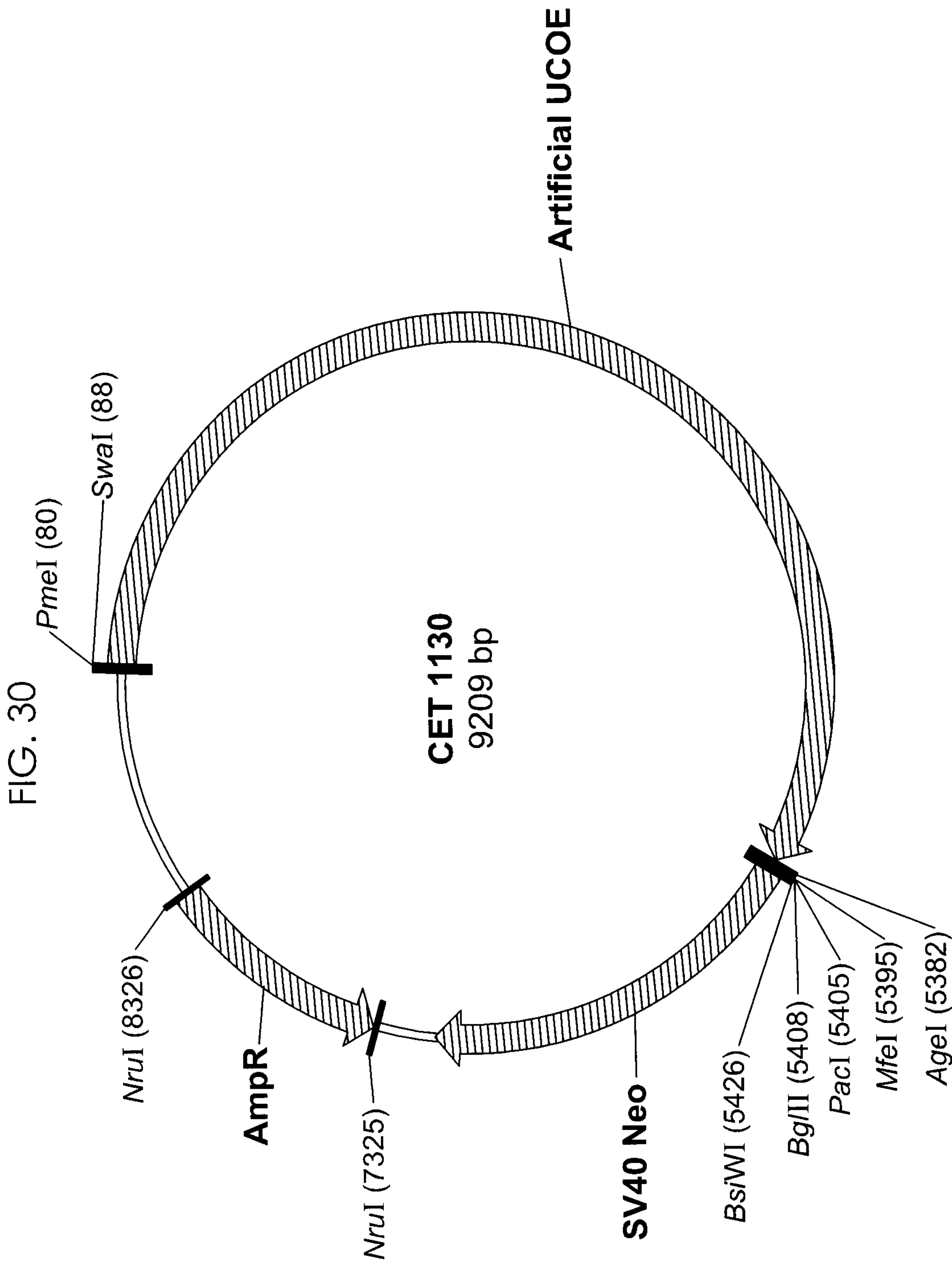
10051	TAATATTGAA	AAAGGAAGAG	TATGAGTATT	CAACATTTCC	GTGTCGCCCT
	ATTATAACTT	TTTCCTTCTC	ATACTCATAA	GTTGTAAAGG	CACAGCGGGA
10101	TATTCCCTTT	TTTGCGGCAT	TTTGCCCTTC	TGTTTTTGCT	CACCCAGAAA
	ATAAGGGAAA	AAACGCCGTA	AAACGGAAGG	ACAAAAACGA	GTGGGTCTTT
10151	CGCTGGTGAA	AGTAAAAGAT	GCTGAAGATC	AGTTGGGTGC	ACGAGTGGGT
	GCGACCACTT	TCATTTTCTA	CGACTTCTAG	TCAACCCACG	TGCTCACCCA
10201	TACATCGAAC	TGGATCTCAA	CAGCGGTAAG	ATCCTTGAGA	GTTTTTCGCCC
	ATGTAGCTTG	ACCTAGAGTT	GTCGCCATTC	TAGGAACTCT	CAAAGCGGG
10251	CGAAGAACGT	TTTCCAATGA	TGAGCACTTT	TAAAGTTCTG	CTATGTGGCG
	GCTTCTTGCA	AAAGGTTACT	ACTCGTGAAA	ATTTCAAGAC	GATACACCGC
10301	CGGTATTATC	CCGTATTGAC	GCCGGGCAAG	AGCAACTCGG	TCGCCGCATA
	GCCATAATAG	GGCATAACTG	CGGCCCGTTC	TCGTTGAGCC	AGCGGCGTAT
10351	CACTATTCTC	AGAATGACTT	GGTTGAGTAC	TCACCAGTCA	CAGAAAAGCA
	GTGATAAGAG	TCTTACTGAA	CCAACTCATG	AGTGGTCAGT	GTCTTTTCGT
10401	TCTTACGGAT	GGCATGACAG	TAAGAGAATT	ATGCAGTGCT	GCCATAACCA
	AGAATGCCTA	CCGTACTGTC	ATTCTCTTAA	TACGTCACGA	CGGTATTGGT
10451	TGAGTGATAA	CACTGCGGCC	AACTTACTTC	TGACAACGAT	CGGAGGACCG
	ACTCACTATT	GTGACGCCGG	TTGAATGAAG	ACTGTTGCTA	GCCTCCTGGC
10501	AAGGAGCTAA	CCGCTTTTTT	GCACAACATG	GGGGATCATG	TAACTCGCCT
	TTCCCTCGATT	GGCGAAAAAA	CGTGTTGTAC	CCCCTAGTAC	ATTGAGCGGA
10551	TGATCGTTGG	GAACCGGAGC	TGAATGAAGC	CATACCAAAC	GACGAGCGTG
	ACTAGCAACC	CTTGGCCTCG	ACTTACTTCG	GTATGGTTTG	CTGCTCGCAC
10601	ACACCACGAT	GCCTGTAGCA	ATGGCAACAA	CGTTGCGCAA	ACTATTAACT
	TGTGGTGCTA	CGGACATCGT	TACCGTTGTT	GCAACGCGTT	TGATAATTGA
10651	GGCGAACTAC	TTACTCTAGC	TTCCCGGCAA	CAATTAATAG	ACTGGATGGA
	CCGCTTGATG	AATGAGATCG	AAGGGCCGTT	GTAAATTATC	TGACCTACCT
10701	GGCGGATAAA	GTTGCAGGAC	CACTTCTGCG	CTCGGCCCTT	CCGGCTGGCT
	CCGCCTATTT	CAACGTCCTG	GTGAAGACGC	GAGCCGGGAA	GGCCGACCGA
10751	GGTTTATTGC	TGATAAATCT	GGAGCCGGTG	AGCGTGGGTC	TCGCGGTATC
	CCAAATAACG	ACTATTTAGA	CCTCGGCCAC	TCGCACCCAG	AGCGCCATAG
10801	ATTGCAGCAC	TGGGGCCAGA	TGGTAAGCCC	TCCCGTATCG	TAGTTATCTA
	TAACGTCGTG	ACCCCGGTCT	ACCATTCGGG	AGGGCATAGC	ATCAATAGAT
10851	CACGACGGGG	AGTCAGGCAA	CTATGGATGA	ACGAAATAGA	CAGATCGCTG
	GTGCTGCCCC	TCAGTCCGTT	GATACCTACT	TGCTTTATCT	GTCTAGCGAC
10901	AGATAGGTGC	CTCACTGATT	AAGCATTTGGT	AACTGTCAGA	CTCGCGACAC
	TCTATCCACG	GAGTGACTAA	TTCGTAACCA	TTGACAGTCT	GAGCGCTGTG
10951	TGCATTAATG	AATCGGCCAA	CGCGCGGGGA	GAGGCGGTTT	GCGTATTGGG
	ACGTAATTAC	TTAGCCGGTT	GCGCGCCCCT	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	AATCAGGGGA	TAACGCAGGA	AAGAACATGT	GAGCAAAAGG	CCAGCAAAAG
	TTAGTCCCCT	ATTGCGTCCT	TTCTTGTACA	CTCGTTTTTC	GGTCGTTTTT
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	TTCGAGGGAG
11301	GTGCGCTCTC	CTGTTCCGAC	CCTGCCGCTT	ACCGGATACC	TGTCCGCCTT
	CACGCGAGAG	GACAAGGCTG	GGACGGCGAA	TGGCCTATGG	ACAGGCGGAA
11351	TCTCCCTTCG	GGAAGCGTGG	CGCTTTCTCA	TAGCTCACGC	TGTAGGTATC
	AGAGGGAAGC	CCTTCGCACC	GCGAAAGAGT	ATCGAGTGCG	ACATCCATAG
11401	TCAGTTCGGT	GTAGGTCGTT	CGCTCCAAGC	TGGGCTGTGT	GCACGAACCC
	AGTCAAGCCA	CATCCAGCAA	GCGAGGTTCG	ACCCGACACA	CGTGCTTGGG
11451	CCCGTTCAGC	CCGACCGCTG	CGCCTTATCC	GGTAACTATC	GTCTTGAGTC
	GGGCAAGTCG	GGCTGGCGAC	GCGGAATAGG	CCATTGATAG	CAGAACTCAG

85/93

11501	CAACCCGGTA	AGACACGACT	TATCGCCACT	GGCAGCAGCC	ACTGGTAACA
	GTTGGGCCAT	TCTGTGCTGA	ATAGCGGTGA	CCGTCGTCGG	TGACCATTTGT
11551	GGATTAGCAG	AGCGAGGTAT	GTAGGCGGTG	CTACAGAGTT	CTTGAAGTGG
	CCTAATCGTC	TCGCTCCATA	CATCCGCCAC	GATGTCTCAA	GAAC TTCACC
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	AAA ACTCA		
	CAGACTGCGA	GTCACCTTGC	TTTTGAGT		

FIG. 29(contd.)

86/93



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	TGTTGCTTCC
	ACCCGAGGAG	GTATTCTGAC	ACTCGTAGAC	TGGTGACCTC	ACAACGAAGG
251	CATTATATCC	CTGTTATCAA	GCACAAGGTC	AGGCACAGAG	TAAGACTCAA
	GTAATATAGG	GACAATAGTT	CGTGTTCCAG	TCCGTGTCTC	ATTCTGAGTT
301	AACATGTTTT	GGAATGTATG	ACTGGTATGA	ACTACAAACC	AGTAAGCTGA
	TTGTACAAAA	CCTTACATAC	TGACCATACT	TGATGTTTGG	TCATTCGACT
351	TGTTTTTCATT	TTGAGTCTAT	AAATCTAATT	TTGTGGTGGT	TTTGTGTATG
	ACAAAAGTAA	AACTCAGATA	TTTAGATTAA	AACACCACCA	AAACACATAC
401	GCTCAAGGCT	CAAATTGTAA	AATTTAATAT	TATGTGACCA	AAGAAAGTTA
	CGAGTTCCGA	GTTTAAACATT	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	TAAAAC TGCC	CAGATGAAAG	AGAAACTTTT
	TTGACTGTTT	TAACACATAG	ATTTTGACCG	GTCTACTTTC	TCTTTGAAAA
701	AAGGGGCCCT	TCTGCGTGCC	CGACACTGTG	CTAGGCACTC	ACACTATCCC
	TTCCCCGGGA	AGACGCACGG	GCTGTGACAC	GATCCGTGAG	TGTGATAGGG
751	GACCCGAGAA	ACCGATCTGC	GACCCAGAGG	AACTTACCAA	GCCTCCAGCA
	CTGGGCTCTT	TGGCTAGACG	CTGGGTCTCC	TTGAATGGTT	CGGAGGTCGT
801	TCTTGTGCAG	CCCTACTCAT	GGGACCATCT	GGATACCCAC	CCTTGTCTTT
	AGAACACGTC	GGGATGAGTA	CCCTGGTAGA	CCTATGGGTG	GGAACAGAAA
851	ACAGGGAGCA	GAACACACCT	CTTATGTGTC	AGAAAACAAA	GTCCAGGAAG
	TGTCCCTCGT	CTTGTGTGGA	GAATACACAG	TCTTTTGT TT	CAGGTCCTTC
901	TATATTTT TA	CCTGAGGCAA	TATCTGAAAA	TTGTATGCTA	CAGCCTCCAA
	ATATAAAAAT	GGACTCCGTT	ATAGACTTTT	AACATACGAT	GTCGGAGGTT
951	AGTGAGTCTT	CCTCTCAGTA	CCTCTCTTCT	AGGCACATGG	AGCCCTTTCT
	TCACTCAGAA	GGAGAGTCAT	GGAGAGAAGA	TCCGTGTACC	TCGGGAAAGA
1001	TCCAAGTATT	ATGTTTAAAC	ACTTAATGAA	TGAAGTCCTG	AAACTGCTTA
	AGGTTCATAA	TACAAATTGG	TGAATTACTT	ACTTCAGGAC	TTTGACGAAT
1051	CCCATGCTCC	CTATAATCTC	TGAGTAATCT	TCCTTTTCCA	CAACCTCAGG
	GGGTACGAGG	GATATTAGAG	ACTCATTAGA	AGGAAAAGGT	GTTGGAGTCC
1101	CATAATCTCA	TCTTCTGTTT	CTATTACAAT	TTCAAATTCT	GGAAAAAGGA
	GTATTAGAGT	AGAAGACAAA	GATAATGTTA	AAGTTTAAGA	CCTTTTTCCT
1151	AGTTGTGGTC	TGGAATTATA	TGGTCCAGAT	GATCTGAAAC	AAAAAGGACA
	TCAACACCAG	ACCTTAATAT	ACCAGGTCTA	CTAGACTTTG	TTTTTCCTGT
1201	GCACTATTAG	TAATCATTTA	GTTTTGAAGA	CAGTCTAATA	ATTTGCTGTC
	CGTGATAATC	ATTAGTAAAT	CAAAACTTCT	GTCAGATTAT	TAAACGACAG
1251	TCTAAAGTAC	TATATTCCCT	ATAGTTCTGG	CATTTTAGAT	AAAGGGTCAT
	AGATTTTCATG	ATATAAGGGA	TATCAAGACC	GTAAAATCTA	TTTCCCAGTA
1301	AAATTAAATG	CCTATATGGT	GACATTATTC	AGTGATTTCAG	ACTTCACAGC
	TTTAATTTAC	GGATATACCA	CTGTAATAAG	TCACTAAGTC	TGAAGTGTCTG

FIG. 31(contd.) ^{88/93}

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	ATTTTTTTTC	TAATGGGTGG
1501	TGGTTGTGCA	CAAGCCTGCT	TATGTCCCAA	TCTCCAGTCT	AGGGTCTGAT
	ACCAACACGT	GTTTCGGACGA	ATACAGGGTT	AGAGGTCAGA	TCCCAGACTA
1551	GCTCCTTGCT	GCAGTAATAT	GCTTTGTGGC	ATCTGGAGCA	CGTTTTGGGG
	CGAGGAACGA	CGTCATTATA	CGAAACACCG	TAGACCTCGT	GCAAAACCCC
1601	CCTAAACAGC	CACAAACCCCT	GCAGAGATGA	GCACCAGACT	TAAGCTGGAG
	GGATTTGTCT	GTGTTTGGGA	CGTCTCTACT	CGTGGTCTGA	ATTCGACCTC
1651	ACACACTGAT	TCTCCTGTTT	CTGGGGGAGG	ATTCTCAGAA	GGTGGCTCAT
	TGTGTGACTA	AGAGGACAAA	GACCCCTCTC	TAAGAGTCTT	CCACCGAGTA
1701	ATGAGTAAAA	ATCGTTTTTC	CTGGGTAGTT	GATTCCTAAA	AACTAAAAAA
	TACTCATTTT	TAGCAAAAAG	GACCCATCAA	CTAAGGATTT	TTGATTTTTT
1751	GAATACAGAG	AAAAGTTTTA	TCTTCAAACA	AAACAGCAAT	TCACATATTT
	CTTATGTCTC	TTTTCAAAAT	AGAAGTTTGT	TTTGTCTGTA	AGTGTATAAA
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	TTTTTCGATGA
2001	GGGATCCATC	TTTCCAAGAC	AATTTCTATT	ATCTGAATTA	ACACCATACC
	CCCTAGGTAG	AAAGGTTCTG	TTAAAGATAA	TAGACTTAAT	TGTGGTATGG
2051	TGGTACCCAC	TGATTAAAAG	CTGGGGGTTA	CCAATGCGCG	TGGGCACAGT
	ACCATGGGTG	ACTAATTTTC	GACCCCAAT	GGTTACGCGC	ACCCGTGTCA
2101	TAGAAGCTTA	TGTAGCAAAA	ATGAGCACAT	CCTGGAAGGG	CCCGGGAGAA
	ATCTTCGAAT	ACATCGTTTT	TACTCGTGTA	GGACCTTCCC	GGGCCCTCTT
2151	GGTGCTCCTG	GGGCAGCGCG	GAGAGGGAGC	TCTGAGGCTG	GGGCGGCAGC
	CCACGAGGAC	CCCGTCGCGC	CTCTCCCTCG	AGACTCCGAC	CCCGCCGTCG
2201	GGTGCTTGCC	GCCGTCCCCC	TGGTCGCTCC	CGGAATTAAC	GCCGCGCACG
	CCACGAACGG	CGGCAGGGGG	ACCAGCGAGG	GCCTTAATTG	CGGCGCGTGC
2251	CGTCGGAGGC	ATGGCCCCGT	CCCGACCCCG	TTTGGCGGCT	CACCTCGCAG
	GCAGCCTCCG	TACCGGGGCA	GGGCTGGGGC	AAACCGCCGA	GTGGAGCGTC
2301	GCCGGCACAG	CACGGCTGCT	CGCGGCAGCA	GAAGAGGAAG	ATGCAGCGGT
	CGGCCGTGTC	GTGCCGACGA	GCGCCGTCGT	CTTCTCCTTC	TACGTCGCCA
2351	GGAAGGCGTC	CGGGCGGCCA	GGCAGCGGCG	CATACACCTG	CAGCAGGAAG
	CCTTCCGCAG	GCCCGCCGGT	CCGTGCGCCG	GTATGTGGAC	GTCGTCCTTC
2401	GAGAGCGGGC	GGCCGCACAG	CTCGCAGGCC	AGGGCCTGGG	GCCCCGGCAG
	CTCTCGCCCC	CCGGCGTGTC	GAGCGTCCGG	TCCCGGACCC	CGGGGCCGTC
2451	CCCGGCCGCG	CCCAGCCATG	CCGGCCGCCC	GCCCACCTTG	CTGGGGAACT
	GGGCCGGCGC	GGGTCGGTAC	GGCCGGCGGG	CGGGTGGAAC	GACCCCTTGA
2501	GCTCGCTGCG	CAGTCGCCAC	GCCGGCGCCG	ACTCGGCGAA	GCCCAGCTCC
	CGAGCGACGC	GTCAGCGGTG	CGGCCGCGGC	TGAGCCGCTT	CGGGTCGAGG
2551	ACAGGCCTGG	CCCCGGCGGC	AGCCATGCGG	GGCGCGGGCT	GGCGTGGGGC
	TGTCCGGACC	GGGGCCGCCG	TCGGTACGCC	CCGCGCCCGA	CCGCACCCCG
2601	GCAGCCACAA	GCTGGGTCGG	AAGGCGGAAA	TCGGGCGCCG	GGCCGGAAGG
	CGTCGGGTGT	CGACCCAGCC	TTCCGCCTTT	AGCCCGCGGC	CCGGCCTTCC
2651	CAAGAGGCGG	GCACCTTTCC	GGAGGACAGG	AGGCGGAAAC	GCGTCTGACG
	GTTCTCCGCC	CGTGGAAGAG	CCTCCTGTCC	TCCGCCTTTG	CGCAGACTGC
2701	GGAGCGGTTG	CAGGACCAAT	GCGAGGGAAC	GGGGCAGAGG	AAACCTCTCG
	CCTCGCCAAC	GTCCTGGTTA	CGCTCCCTTG	CCCCGTCTCC	TTTGGAGAGC
2751	GCATCAGCCC	CGCCCCTGGC	GCCTCTGCCT	CCGAGCCGCT	TTCTGGGTGC
	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	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	AACTAGTTTT	TCAGGTTAAC	TTCAGAATGC
	GGTGAAGGGT	AGTAACGGAC	TTGATCAAAA	AGTCCAATTG	AAGTCTTACG
3251	CCTTGGTACC	GCGGGCCCCC	TCTGTGGTCC	CACGCCACTG	ATCGCTGCAT
	GGAACCATGG	CGCCCCGGGG	AGACACCAGG	GTGCGGTGAC	TAGCGACGTA
3301	GCCCACCACC	TGGGTACACA	CAGTCTGTGA	TTCCCGGAGC	AGAACGGACC
	CGGGTGGTGG	ACCCATGTGT	GTCAGACACT	AAGGGCCTCG	TCTTGCCTGG
3351	CTGCCACCCC	GGTCTTGTGT	GCTACTCAGT	GGACAGACCC	AAGGCAAGAA
	GACGGGTGGG	CCAGAACACA	CGATGAGTCA	CCTGTCTGGG	TTCCGTTCTT
3401	AGGGTGACAA	GGACAGGGTC	TTCCCAGGCT	GGCTTTGAGT	TCCTAGCACC
	TCCCACGTGT	CCTGTCCCAG	AAGGGTCCGA	CCGAAACTCA	AGGATCGTGG
3451	GCCCCGCCCC	CAATCCTCTG	TGGCACATGG	AGTCTTGGTC	CCCAGAGTCC
	CGGGGCGGGG	GTTAGGAGAC	ACCGTGTACC	TCAGAACCAG	GGGTCTCAGG
3501	CCCAGCGGCC	TCCAGATGGT	CTGGGAGGGC	AGTTCAGCTG	TGGCTGCGCA
	GGGTGCGCCG	AGGTCTACCA	GACCCTCCCC	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	TCTTGGCTGG	GCGTGACTGT	TACCCTCAAA	AGCAGGCAGC	TCCAGGGTAA
	AGAACCGACC	CGCACTGACA	ATGGGAGTTT	TCGTCCGTCG	AGGTCCCATT
3701	AAGGTGCCCT	GCCCTGTAGA	GCCCACTTCC	TTCCCAGGGC	TGCGGCTGGG
	TTCCACGGGA	CGGGACATCT	CGGGTGAAGG	AAGGGTCCCC	ACGCCGACCC
3751	TAGGTTTGTA	GCCTTCATCA	CGGGCCACCT	CCAGCCACTG	GACCGCTGGC
	ATCCAAACAT	CGGAAGTAGT	GCCCGGTGGA	GGTCGGTGAC	CTGGCGACCG
3801	CCCTGCCCTG	TCCTGGGGAG	TGTGGTCCCTG	CGACTCTAAT	GGCCGCAAGC
	GGGACGGGAC	AGGACCCCTC	ACACCAGGAC	GCTGAGATTA	CCGGCGTTCC
3851	CACCTGACTC	CCCCAACACC	AACTCTACC	TCTCAAGCCC	AGGTCTCTCC
	GTGGACTGAG	GGGGTTGTGG	TGTGAGATGG	AGAGTTCGGG	TCCAGAGAGG
3901	CTAGTGACCC	ACCCAGCACA	TTTAGCTAGC	TGAGCCCCAC	AGCCAGAGGT
	GATCACTGGG	TGGGTCGTGT	AAATCGATCG	ACTCGGGGTG	TCGGTCTCCA
3951	CCTCAGGCC	TGCTTTCAGG	GCAGTTGCTC	TGAAGTCGGC	AAGGGGGAGT
	GGAGTCCGGG	ACGAAAGTCC	CGTCAACGAG	ACTTCAGCCG	TTCCCCCTCA
4001	GACTGCCTGG	CCACTCCATG	CCCTCCAAGA	GCTCCTTCTG	CAGGAGCGTA
	CTGACGGACC	GGTGAGGTAC	GGGAGGTTCT	CGAGGAAGAC	GTCTTCGCAT
4051	CAGAACCCAG	GGCCCTGGCA	CCCGTGCAGA	CCCTGGCCCA	CCCCACCTGG
	GTCTTGGGTC	CCGGGACCGT	GGGCACGTCT	GGGACCGGGT	GGGGTGGACC
4101	GCGCTCAGTG	CCCAAGAGAT	GTCCACACCT	AGGATGTCCC	GCGGTGGGTG
	CGCGAGTCAC	GGGTTCCTTA	CAGGTGTGGA	TCCTACAGGG	CGCCACCCAC
4151	GGGGGCCCCG	GAGACGGGCA	GGCCGGGGGC	AGGCCTGGCC	ATGCGGGGCC
	CCCCCGGGCT	CTCTGCCCGT	CCGGCCCCCG	TCCGGACCGG	TACGCCCCGG
4201	GAACCGGGCA	CTGCCCAGCG	TGGGGCGCGG	GGGCCACGGC	GCGCGCCCCC
	CTTGGCCCGT	GACGGGTCGC	ACCCCGCGCC	CCCGGTGCCG	CGCGCGGGGG

FIG. 31(contd.) ^{90/93}

4251	AGCCCCCGGG	CCCAGCACCC	CAAGGCGGCC	AACGCCAAAA	CTCTCCCTCC
	TCGGGGGGCCC	GGGTCGTGGG	GTTCCGCGCG	TTGCGGTTTT	GAGAGGGAGG
4301	TCCTCTTCCT	CAATCTCGCT	CTCGCTCTTT	TTTTTTTTTCG	CAAAAGGAGG
	AGGAGAAGGA	GTTAGAGCGA	GAGCGAGAAA	AAAAAAAAAGC	GTTTTCTCTCC
4351	GGAGAGGGGG	TAAAAAAATG	CTGCACTGTG	CGGCGAAGCC	GGTGAGTGAG
	CCTCTCCCCC	ATTTTTTTTAC	GACGTGACAC	GCCGCTTCGG	CCACTCACTC
4401	CGGCGCGGGG	CCAATCAGCG	TGCGCCGTTT	CGAAAGTTGC	CTTTTATGGC
	GCCGCGCCCC	GGTTAGTCGC	ACGCGGCAAG	GCTTTCAACG	GAAAATACCG
4451	TCGAGCGGCC	GCGGCGGCGC	CCTATAAAAC	CCAGCGGCGC	GACGCGCCAC
	AGCTCGCCGG	CGCCGCGCGC	GGATATTTTG	GGTCGCGCGC	CTGCGCGGTG
4501	CACCGCCGAG	ACCGCGTCCG	CCCGCGAGCA	CAGAGCCTCG	CCTTTGCCGA
	GTGGCGGCTC	TGGCGCAGGC	GGGCGCTCGT	GTCTCGGAGC	GGAAACGGCT
4551	TCCGCCGCCC	GTCCACACCC	GCCGCCAGGT	AAGCCCGGCC	AGCCGACCGG
	AGGCGGCGGG	CAGGTGTGGG	CGGCGGTCCA	TTCGGGGCCG	TCGGCTGGCC
4601	GGCATGCGGC	CGCGGCCCTT	CGCCCGTGCA	GAGCCGCCGT	CTGGGCGCGA
	CCGTACGCCG	GCGCCGGGAA	GCGGGCACGT	CTCGGCGGCA	GACCCGGCGT
4651	GCGGGGGGCG	CATGGGGCGG	AACCGGACCG	CCGTGGGGGG	CGCGGGAGAA
	CGCCCCCCGC	GTACCCCGCC	TTGGCCTGGC	GGCACCCCCC	GCGCCCTCTT
4701	GCCCCTGGGC	CTCCGGAGAT	GGGGGACACC	CCACGCCAGT	TCGCAGGCGC
	CGGGGACCCG	GAGGCCTCTA	CCCCCTGTGG	GGTGCGGTCA	AGCGTCCGCG
4751	GAGGCCGCGC	TCGGGCGGGC	GCGCTCCGGG	GGTGCCGCTC	TCGGGGCGGG
	CTCCGGCGCG	AGCCCGCCCC	CGCGAGGCC	CCACGGCGAG	AGCCCGCCCC
4801	GGCAACCGGC	GGGGTCTTTG	TCTGAGCCGG	GCTCTTGCCA	ATGGGGATCG
	CCGTGCGCCG	CCCCAGAAAC	AGACTCGGCC	CGAGAACGGT	TACCCCTAGC
4851	CACGGTGGGC	GCGGCGTAGC	CCCCGTCAGG	CCCGGTGGGG	GCTGGGGCGC
	GTGCCACCCG	CGCCGCATCG	GGGGCAGTCC	GGGCCACCCC	CGACCCCGCG
4901	CATGCGCGTG	CGCGCTGGTC	CTTTGGGCGC	TAAGTGCCTG	CGCGCTGGGA
	GTACGCGCAC	GCGCGACCAG	GAAACCCGCG	ATTGACGCAC	GCGCGACCCT
4951	ATTGGCGCTA	ATTGCGCGTG	CGCGCTGGGA	CTCAATGGCG	CTAATCGCGC
	TAACCGCGAT	TAACGCGCAC	GCGCGACCCT	GAGTTACCGC	GATTAGCGCG
5001	GTGCGTTCTG	GGGCCCCGGC	GCTTGCGCCA	CTTCCTGCCC	GAGCCGCTGG
	CACGCAAGAC	CCCGGGCCCC	CGAACGCGGT	GAAGGACGGG	CTCGGCGACC
5051	CGCCCGAGGG	TGTGGCCGCT	GCGTGCGCGC	GCGCGACCCG	GTGCTGTGTT
	GCGGGCTCCC	ACACCGGCGA	CGCACGCGCG	CGCGCTGGGC	CAGCGACAAA
5101	GAACCGGGCG	GAGGCGGGGC	TGGCGCCCCG	TTGGGAGGGG	GTTGGGGCCT
	CTTGGCCCCG	CTCCGCCCCC	ACCGCGGGCC	AACCCCTCCC	CAACCCCGGA
5151	GGCTTCCTGC	CGCGCGCCGC	GGGGACGCCT	CCGACCAGTG	TTTGCCTTTT
	CCGAAGGACG	GCGCGCGGCG	CCCCTGCGGA	GGCTGGTCAC	AAACGGAAAA
5201	ATGGTAATAA	CGCGGCCGGC	CCGGCTTCCT	TTGTCCCCAA	TCTGGGCGCG
	TACCATTATT	GCGCCGGCCG	GGCCGAAGGA	AACAGGGGTT	AGACCCGCGC
5251	CGCCGGCGCC	CCCTGGCGGC	CTAAGGACTC	GGCGCGCCGG	AAGTGGCCAG
	GCGGCCGCGG	GGGACCGCCG	GATTCCTGAG	CCGCGCGGCC	TTCACCGGTC
5301	GGCGGGGGCG	ACTTCGGCTC	ACAGCGCGCC	CGGCTATTCT	CGCAGCTCAC
	CCGCCCCCGC	TGAAGCCGAG	TGTCGCGCGG	GCCGATAAGA	GCGTCGAGTG
5351	CATGCCGGTC	GCCACCATGA	GCTTATCGAT	ACCGGTGGCG	CGCCAATTGT
	GTACGGCCAG	CGGTGGTACT	CGAATAGCTA	TGGCCACCGC	GCGGTTAACA
5401	TAATTAAGAT	CTGGCCCAAT	GGGCCGTACG	AATTTGAGGC	GGAAAGAACC
	ATTAATTCTA	GACCGGGTTA	CCCGGCATGC	TTAAACTCCG	CCTTTCTTGG
5451	AGCTGTGGAA	TGTGTGTCAG	TTAGGGTGTG	GAAAGTCCCC	AGGCTCCCCA
	TCGACACCTT	ACACACAGTC	AATCCCACAC	CTTTCAGGGG	TCCGAGGGGT
5501	GCAGGCAGAA	GTATGCAAAG	CATGCATCTC	AATTAGTCAG	CAACCAGGTG
	CGTCCGTCTT	CATACGTTTC	GTACGTAGAG	TTAATCAGTC	GTTGGTCCAC
5551	TGGAAAGTCC	CCAGGCTCCC	CAGCAGGCAG	AAGTATGCAA	AGCATGCATC
	ACCTTTCAGG	GGTCCGAGGG	GTCGTCCGTC	TTCATACGTT	TCGTACGTAG
5601	TCAATTAGTC	AGCAACCATA	GTCCCGCCCC	TAAGTCCGCC	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
	TCTCTGTCTT	ACTCCTAGCA	AAGCGTACTA	ACTTGTTCTA	CCTAACGTGC
5851	CAGGTTCTCC	GGCCGCTTGG	GTGGAGAGGC	TATTTCGGCTA	TGACTGGGCA
	GTCCAAGAGG	CCGGCGAACC	CACCTCTCCG	ATAAGCCGAT	ACTGACCCGT
5901	CAACAGACAA	TCGGCTGCTC	TGATGCCGCC	GTGTTCCGGC	TGTCAGCGCA
	GTTGTCTGTT	AGCCGACGAG	ACTACGGCGG	CACAAGGCCG	ACAGTCGCGT
5951	GGGGCGCCCG	GTTCTTTTTG	TCAAGACCGA	CCTGTCCGGT	GCCCTGAATG
	CCCCGCGGGC	CAAGAAAAAC	AGTTCTGGCT	GGACAGGCCA	CGGGACTTAC
6001	AACTGCAAGA	CGAGGCAGCG	CGGCTATCGT	GGCTGGCCAC	GACGGGCGTT
	TTGACGTTCT	GCTCCGTCGC	GCCGATAGCA	CCGACCGGTG	CTGCCCCGAA
6051	CCTTGCGCAG	CTGTGCTCGA	CGTTGTCACT	GAAGCGGGAA	GGGACTGGCT
	GGAACGCGTC	GACACGAGCT	GCAACAGTGA	CTTCGCCCTT	CCCTGACCGA
6101	GCTATTGGGC	GAAGTGCCGG	GGCAGGATCT	CCTGTCATCT	CACCTTGCTC
	CGATAACCCG	CTTCACGGCC	CCGTCCTAGA	GGACAGTAGA	GTGGAACGAG
6151	CTGCCGAGAA	AGTATCCATC	ATGGCTGATG	CAATGCGGCG	GCTGCATACG
	GACGGCTCTT	TCATAGGTAG	TACCGACTAC	GTTACGCCGC	CGACGTATGC
6201	CTTGATCCGG	CTACCTGCCC	ATTCGACCAC	CAAGCGAAAC	ATCGCATCGA
	GAAC TAGGCC	GATGGACGGG	TAAGCTGGTG	GTTTCGCTTTG	TAGCGTAGCT
6251	GCGAGCACGT	ACTCGGATGG	AAGCCGGTCT	TGTCGATCAG	GATGATCAAG
	CGCTCGTGCA	TGAGCCTACC	TTCGGCCAGA	ACAGCTAGTC	CTACTAGTTC
6301	AGCATCAGGG	GCTCGCGCCA	GCCGAACGTG	TCGCCAGGCT	CAAGGCGAGC
	TCGTAGTCCC	CGAGCGCGGT	CGGCTTGACA	AGCGGTCCGA	GTTCCGCTCG
6351	ATGCCCGACG	GCGAGGATCT	CGTCGTGACC	CATGGCGATG	CCTGCTTGCC
	TACGGGCTGC	CGCTCCTAGA	GCAGCACTGG	GTACCGCTAC	GGACGAACGG
6401	GAATATCATG	GTGGAAAATG	GCCGCTTTTC	TGGATTTCATC	GACTGTGGCC
	CTTATAGTAC	CACCTTTTAC	CGGCGAAAAG	ACCTAAGTAG	CTGACACCGG
6451	GGCTGGGTGT	GGCGGACCGC	TATCAGGACA	TAGCGTTGGC	TACCCGTGAT
	CCGACCCACA	CCGCCTGGCG	ATAGTCCTGT	ATCGCAACCG	ATGGGCACTA
6501	ATTGCTGAAG	AGCTTGCGCG	CGAATGGGCT	GACCGCTTCC	TCGTGCTTTA
	TAACGACTTC	TCGAACCGCC	GCTTACCCGA	CTGGCGAAGG	AGCACGAAAT
6551	CGGTATCGCC	GCTCCCGATT	CGCAGCGCAT	CGCCTTCTAT	CGCCTTCTTG
	GCCATAGCGG	CGAGGGCTAA	GCGTCGCGTA	GCGGAAGATA	GCGGAAGAAC
6601	ACGAGTTCTT	CTGAGCGGGA	CTCTGGGGTT	CGAAATGACC	GACCAAGCGA
	TGCTCAAGAA	GACTCGCCCT	GAGACCCCAA	GCTTTACTGG	CTGGTTCGCT
6651	CGCCCAACCT	GCCATCACGA	GATTTCGATT	CCACCGCCGC	CTTCTATGAA
	GCGGGTTGGA	CGGTAGTGCT	CTAAAGCTAA	GGTGGCGGCG	GAAGATACTT
6701	AGGTTGGGCT	TCGGAATCGT	TTTCCGGGAC	GCCGGCTGGA	TGATCCTCCA
	TCCAACCCGA	AGCCTTAGCA	AAAGGCCCTG	CGGCCGACCT	ACTAGGAGGT
6751	GCGCGGGGAT	CTCATGCTGG	AGTTCTTCGC	CCACCCTAGG	GGGAGGCTAA
	CGCGCCCCCTA	GAGTACGACC	TCAAGAAGCG	GGTGGGATCC	CCCTCCGATT
6801	CTGAAACACG	GAAGGAGACA	ATACCGGAAG	GAACCCGCGC	TATGACGGCA
	GACTTTGTGC	CTTCTCTGT	TATGGCCTTC	CTTGGGCGCG	ATACTGCCGT
6851	ATAAAAAGAC	AGAATAAAAC	GCACGGTGTT	GGGTCGTTTG	TTCATAAACG
	TATTTTTCTG	TCTTATTTTG	CGTGCCACAA	CCCAGCAAAC	AAGTATTTGC
6901	CGGGGTTCGG	TCCCAGGGCT	GGCACTCTGT	CGATACCCCA	CCGAGACCCC
	GCCCCAAGCC	AGGGTCCCGA	CCGTGAGACA	GCTATGGGGT	GGCTCTGGGG
6951	ATTGGGGCCA	ATACGCCCCG	GTTTCTTCCT	TTTCCCCACC	CCACCCCCCA
	TAACCCCGGT	TATGCGGGCG	CAAAGAAGGA	AAAGGGGTGG	GGTGGGGGGT
7001	AGTTCGGGTG	AAGGCCCAGG	GCTCGCAGCC	AACGTCGGGG	CGGCAGGCCC
	TCAAGCCAC	TTCCGGGTCC	CGAGCGTCGG	TTGCAGCCCC	GCCGTCCGGG
7051	TGCCATAGCC	TCAAATTCCT	TAGGCTCGAG	GGCCGCCACC	GCGGTGGAGC
	ACGGTATCGG	AGTTTAAGGA	ATCCGAGCTC	CCGGCGGTGG	CGCCACCTCG
7101	TCCAGCTTTT	GTTCCCTTTA	GTGAGGGTTA	ATTTTCGAGCT	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	TATTTACACAT	TTCGGACCCC	ACGGATTACT
7251	GTGAGCTAAC	TCACATTAAT	TGCGTTGCGC	TCACTGCCCC	CTTTCCAGTC
	CACTCGATTG	AGTGTAATTA	ACGCAACGCG	AGTGACGGGC	GAAAGGTCAG
7301	GGGAAACCTG	TCGTGCCAGC	ATCGCGAGCA	CTTTTCGGGG	AAATGTGCGC
	CCCTTTGGAC	AGCACGGTCG	TAGCGCTCGT	GAAAAGCCCC	TTTACACGCG
7351	GGAACCCCTA	TTTGTTTATT	TTTCTAAATA	CATTCAAATA	TGTATCCGCT
	CCTTGGGGAT	AAACAAATAA	AAAGATTTAT	GTAAGTTTAT	ACATAGGCGA
7401	CATGAGACAA	TAACCTTGAT	AAATGCTTCA	ATAATATTGA	AAAAGGAAGA
	GTACTCTGTT	ATTGGGACTA	TTTACGAAGT	TATTATAACT	TTTTCTTCT
7451	GTATGAGTAT	TCAACATTTC	CGTGTCGCCC	TTATTCCCTT	TTTTGCGGCA
	CATACTCATA	AGTTGTAAAG	GCACAGCGGG	AATAAGGGAA	AAAACGCCGT
7501	TTTTGCCTTC	CTGTTTTTGC	TCACCCAGAA	ACGCTGGTGA	AAGTAAAAGA
	AAAACGGAAG	GACAAAAACG	AGTGGGTCTT	TGCGACCACT	TTCATTTTCT
7551	TGCTGAAGAT	CAGTTGGGTG	CACGAGTGGG	TTACATCGAA	CTGGATCTCA
	ACGACTTCTA	GTCAACCCAC	GTGCTCACCC	AATGTAGCTT	GACCTAGAGT
7601	ACAGCGGTAA	GATCCTTGAG	AGTTTTTCGCC	CCGAAGAACG	TTTTCCAATG
	TGTCGCCATT	CTAGGAACTC	TCAAAAGCGG	GGCTTCTTGC	AAAAGGTTAC
7651	ATGAGCACTT	TTAAAGTTCT	GCTATGTGGC	GCGGTATTAT	CCCGTATTGA
	TACTCGTGAA	AATTTCAAGA	CGATACACCG	CGCCATAATA	GGGCATAACT
7701	CGCCGGGCAA	GAGCAACTCG	GTCGCCGCAT	ACACTATTCT	CAGAATGACT
	GCGGCCCGTT	CTCGTTGAGC	CAGCGGCGTA	TGTGATAAGA	GTCTTACTGA
7751	TGGTTGAGTA	CTCACCAGTC	ACAGAAAAGC	ATCTTACGGA	TGGCATGACA
	ACCAACTCAT	GAGTGGTCAG	TGTCPTTTTCG	TAGAATGCCT	ACCGTACTGT
7801	GTAAGAGAAT	TATGCAGTGC	TGCCATAAACC	ATGAGTGATA	ACACTGCGGC
	CATTCTCTTA	ATACGTCACG	ACGGTATTGG	TACTCACTAT	TGTGACGCCG
7851	CAACTTACTT	CTGACAACGA	TCGGAGGACC	GAAGGAGCTA	ACCGCTTTTT
	GTTGAATGAA	GACTGTTGCT	AGCCTCCTGG	CTTCCTCGAT	TGGCGAAAAA
7901	TGCACAACAT	GGGGGATCAT	GTAACCTCGCC	TTGATCGTTG	GGAACCGGAG
	ACGTGTTGTA	CCCCCTAGTA	CATTGAGCGG	AACTAGCAAC	CCTTGGCCTC
7951	CTGAATGAAG	CCATACCAA	CGACGAGCGT	GACACCACGA	TGCCTGTAGC
	GACTTACTTC	GGTATGGTTT	GCTGCTCGCA	CTGTGGTGCT	ACGGACATCG
8001	AATGGCAACA	ACGTTGCGCA	AACTATTAAC	TGGCGAACTA	CTTACTCTAG
	TTACCGTTGT	TGCAACGCGT	TTGATAATTG	ACCGCTTGAT	GAATGAGATC
8051	CTTCCCGGCA	ACAATTAATA	GACTGGATGG	AGGCGGATAA	AGTTGCAGGA
	GAAGGGCCGT	TGTTAATTAT	CTGACCTACC	TCCGCCTATT	TCAACGTCCT
8101	CCACTTCTGC	GCTCGGCCCT	TCCGGCTGGC	TGGTTTATTG	CTGATAAATC
	GGTGAAGACG	CGAGCCGGGA	AGGCCGACCG	ACCAAATAAC	GACTATTTAG
8151	TGGAGCCGGT	GAGCGTGGGT	CTCGCGGTAT	CATTGCAGCA	CTGGGGCCAG
	ACCTCGGCCA	CTCGCACCCA	GAGCGCCATA	GTAACGTCGT	GACCCCGGTC
8201	ATGGTAAGCC	CTCCCGTATC	GTAGTTATCT	ACACGACGGG	GAGTCAGGCA
	TACCATTCGG	GAGGGCATAG	CATCAATAGA	TGTGCTGCCC	CTCAGTCCGT
8251	ACTATGGATG	AACGAAATAG	ACAGATCGCT	GAGATAGGTG	CCTCACTGAT
	TGATACCTAC	TTGCTTTATC	TGTCTAGCGA	CTCTATCCAC	GGAGTGACTA
8301	TAAGCATTGG	TAAGTGTGAG	ACTCGCGACA	CTGCATTAAT	GAATCGGCCA
	ATTCGTAACC	ATTGACAGTC	TGAGCGCTGT	GACGTAATTA	CTTAGCCGGT
8351	ACGCGCGGGG	AGAGGCGGTT	TGCGTATTGG	GCGCTCTTCC	GCTTCCTCGC
	TGCGCGCCCC	TCTCCGCCAA	ACGCATAACC	CGCGAGAAGG	CGAAGGAGCG
8401	TCACTGACTC	GCTGCGCTCG	GTCGTTCTGGC	TGCGGCGAGC	GGTATCAGCT
	AGTGACTGAG	CGACGCGAGC	CAGCAAGCCG	ACGCCGCTCG	CCATAGTCGA
8451	CACTCAAAGG	CGGTAATACG	GTTATCCACA	GAATCAGGGG	ATAACGCAGG
	GTGAGTTTCC	GCCATTATGC	CAATAGGTGT	CTTAGTCCCC	TATTGCGTCC
8501	AAAGAACATG	TGAGCAAAAG	GCCAGCAAAA	GGCCAGGAAC	CGTAAAAAGG
	TTTCTTGTAC	ACTCGTTTTT	CGGTCTTTTT	CCGGTCCTTG	GCATTTTTTCC
8551	CCGCGTTGCT	GGCGTTTTTC	CATAGGCTCC	GCCCCCTGA	CGAGCATCAC
	GGCGCAACGA	CCGCAAAAAG	GTATCCGAGG	CGGGGGGACT	GCTCGTAGTG

93/93

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

FIG. 31(contd.)

