

(12) STANDARD PATENT
(19) AUSTRALIAN PATENT OFFICE

(11) Application No. **AU 2002253226 B2**

(54) Title
Synthetic genes and bacterial plasmids devoid of CpG

(51) International Patent Classification(s)
C12N 15/09 (2006.01) **C12N 15/31** (2006.01)
C07K 14/245 (2006.01) **C12N 15/54** (2006.01)
C12N 1/21 (2006.01) **C12N 15/70** (2006.01)
C12N 9/10 (2006.01)

(21) Application No: **2002253226** (22) Date of Filing: **2002.03.11**

(87) WIPO No: **WO02/072846**

(30) Priority Data

(31) Number	(32) Date	(33) Country
01/03274	2001.03.09	FR

(43) Publication Date: **2002.09.24**

(43) Publication Journal Date: **2003.03.20**

(44) Accepted Journal Date: **2007.11.15**

(71) Applicant(s)
Cayla

(72) Inventor(s)
Drocourt, Daniel;Reynes, Jean-Paul;Tiraby, Gerard

(74) Agent / Attorney
Griffith Hack, Level 3 509 St Kilda Road, Melbourne, VIC, 3004

(56) Related Art
Skopek et al (1996) Mutation Research, Vol. 349, pages 163-172

(12) DEMANDE INTERNATIONALE PUBLIÉE EN VERTU DU TRAITÉ DE COOPÉRATION
EN MATIÈRE DE BREVETS (PCT)

VERSION CORRIGÉE

(19) Organisation Mondiale de la Propriété
Intellectuelle
Bureau international



(43) Date de la publication internationale
19 septembre 2002 (19.09.2002)

PCT

(10) Numéro de publication internationale
WO 02/072846 A3

(51) Classification internationale des brevets⁷ :

C12N 15/70, 15/31, 1/21, 9/10, 15/54

SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VN,
YU, ZA, ZM, ZW.

(21) Numéro de la demande internationale :

PCT/FR02/00862

(84) États désignés (*régional*) : brevet ARIPO (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), brevet eurasien (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), brevet européen (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), brevet OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

(22) Date de dépôt international : 11 mars 2002 (11.03.2002)

(25) Langue de dépôt :

français

(26) Langue de publication :

français

Déclarations en vertu de la règle 4.17 :

— relative à la qualité d'inventeur (règle 4.17.iv)) pour US seulement

— relative à des divulgations non opposables ou à des exceptions au défaut de nouveauté (règle 4.17.v)) pour toutes les désignations

(30) Données relatives à la priorité :

01/03274 9 mars 2001 (09.03.2001) FR

(71) Déposant (*pour tous les États désignés sauf US*) : CAYLA [FR/FR]; ZI Montaudran BP4437, 5, Rue Jean Rodier, F-31405 Toulouse Cedex 4 (FR).

Publiée :

avec rapport de recherche internationale
avec une déclaration relative à des divulgations non opposables ou à des exceptions au défaut de nouveauté

(72) Inventeurs; et

(75) Inventeurs/Déposants (*pour US seulement*) : DROCOURT, Daniel [FR/FR]; 10, rue de la Désirade, F-31650 Saint-Orens de Gameville (FR). REYNES, Jean-Paul [FR/FR]; 17, rue Canto-gril, F-31750 Escalquens (FR). TIRABY, Gérard [FR/FR]; 99, chemin de la Salade-Pon-san, F-31400 Toulouse (FR).

(88) Date de publication du rapport de recherche internationale:

27 février 2003

(74) Mandataires : PARIS, Fabienne etc.; Cabinet Plasseraud, 84, rue d'Amsterdam, F-75440 Paris Cedex 09 (FR).

(48) Date de publication de la présente version corrigée:

20 mars 2003

(81) États désignés (*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,

(15) Renseignements relatifs à la correction:

voir la Gazette du PCT n° 12/2003 du 20 mars 2003, Section II

En ce qui concerne les codes à deux lettres et autres abréviations, se référer aux "Notes explicatives relatives aux codes et abréviations" figurant au début de chaque numéro ordinaire de la Gazette du PCT.

(54) Title: SYNTHETIC GENES AND BACTERIAL PLASMIDS DEVOID OF CPG

(54) Titre : GENES SYNTHETIQUES ET PLASMIDES BACTÉRIENS DEPOURVUS DE CPG

(57) Abstract: The invention relates to a new series of bacterial plasmid vectors which are fully devoid of CpG and which can express synthetic genes which do not contain CpG in the bacteria Escherichia coli.

(57) Abrégé : La présente demande est relative à une série de vecteurs plasmidiques bactériens entièrement dépourvus de CpG qui peuvent exprimer chez la bactérie Escherichia coli des gènes synthétiques ne contenant pas de CpG.



WO 02/072846 A3

IN THE MATTER OF an Australian
Application corresponding to
PCT Application PCT/FR02/00862

I, Elaine Patricia PARRISH BSc, PhD,
translator to RWS Group plc, of Europa House, Marsham Way, Gerrards Cross,
Buckinghamshire, England, do solemnly and sincerely declare that I am conversant with the
English and French languages and am a competent translator thereof, and that to the best of
my knowledge and belief the following is a true and correct translation of the PCT
Application filed under No. PCT/FR02/00862.

Date: 14 August 2003



E. P. PARRISH

For and on behalf of RWS Group plc

Synthetic genes and bacterial plasmids devoid of CpG

FIELD OF THE INVENTION

- 5 The present application relates to synthetic genes and to plasmids entirely devoid of CpG.

TECHNOLOGICAL BACKGROUND

- 10 All references, including any patents or patent applications, cited in this specification are hereby incorporated by reference. No admission is made that any reference constitutes prior art. The discussion of the references states what their authors assert, and the
15 applicants reserve the right to challenge the accuracy and pertinency of the cited documents. It will be clearly understood that, although a number of prior art publications are referred to herein, this reference does not constitute an admission that any of these documents
20 forms part of the common general knowledge in the art, in Australia or in any other country.

- Plasmids are genetic elements essentially found in bacteria, made up of a molecule of deoxyribonucleic acid,
25 which is most commonly circular and the replication of which is autonomous and independent from that of genomic DNA. Natural plasmids isolated from a very broad variety of bacteria are capable of accomplishing several cellular functions. The first, which is vital for all plasmids, is
30 that responsible for their replication, generally carried out in a manner which is synchronized with replication of the genomic DNA and cell division. Besides the region required for replication of the plasmid, all natural plasmids carry genes which encode proteins, the function
35 of which most commonly remains unknown due to a lack of scientific investigation with regard to these genes. The number of genes present on a plasmid determines the size

- 1a -

of this plasmid, the smallest natural plasmids containing only two or three genes. The properties of plasmids attracted research scientists to them very early on, to make them vehicles for transporting and expressing genes
5 in prokaryotic cells, as well as in eukaryotic cells. The very rapid progress observed in the fields of the molecular biology of nucleic acids and proteins, over the past two decades, can in part be attributed to the exploitation of recombined plasmids constructed from
10 fragments of natural DNA of plasmid origin or other cellular DNAs, and even chemically synthesized.

The four bases adenine (A), guanine (G), cytosine (C) and thymine (T) which constitute deoxyribonucleic acid (DNA) are distributed in 16 dinucleotide configurations, namely CG, GC, TA, AT, CC, GG, TT, AA, TG, CA, AG, CT, AC, GT, GA and TC. Analysis of the qualitative distribution of the dinucleotides of the DNA of thousands of plasmids for which sequences are known reveals that the 16 dinucleotides are always present in all natural plasmids or plasmids constructed in the laboratory. However, analysis of the quantitative distribution of the dinucleotides of plasmids shows great disparities which depend, only partly, on the percentage of each one of the four bases of the DNA. Specifically, comparison of the frequencies observed for each one of the dinucleotides with those of the frequencies calculated on the basis of a random association between two bases, for a given plasmid, can demonstrate major differences for several dinucleotides in terms of an over representation, or, on the contrary, an under representation (Campbell A., Mrazek J. and Karlin S. (1999) *Proc Natl Acad Sci USA* 96, 9184-9). The differences observed in the distribution of certain dinucleotides, not always the same, of natural plasmids isolated from bacteria of phylogenically distant species have been explained by differences in specificity in the mechanisms of repair, recombination and replication acting on the cellular DNAs.

Gene transfers *in vitro* into cells in culture and *in vivo* into various animals are common practices undergoing great development, on the one hand, for the purpose of gaining a better understanding of cell function and, on the other hand, in order to apply these techniques to cell and gene therapies. None of the viral vectors and plasmid vectors among the panoply of vectors available for gene transfer in animals has taken a decisive advantage over the others, since each one has advantages, but also disadvantages. There is,

however, an application in which naked plasmid DNA or plasmid DNA complexed with various substances to facilitate DNA transport to the nucleus is the subject of intense research activity, namely that of immunizing DNA. The principle of immunizing DNA is based on the immune responses observed in laboratory animals treated, by intramuscular or intradermal injection or by inhalation, with plasmid DNA encoding an antigenic peptide. It is now well established that a first consequence of introducing a plasmid DNA derived from the bacteria *E. coli* into the body of an experimental animal intravenously and intramuscularly is the rapid production of various cytokines by the guard cells of the immune system (Krieg A. M. and Kline J. N. (2000) *Immunopharmacology* 48, 303-305). This response is extremely specific for bacterial DNA since DNA extracted from animal cells does not cause such an induction of cytokines under the same conditions. The cellular mechanisms involved in this immune response are far from being fully understood. However, it is known that the recognition which discriminates between bacterial DNA and DNA of animal origin takes place at the level of structural differences relating to the methylation of certain cytosines of the molecule. Specifically, mammalian DNA is naturally methylated at the cytosine of all CG dinucleotides (subsequently written CpG), with the exception of short regions of high CpG density, called CpG islands, present in functional regions in some promoters. DNA extracted from *E. coli* does not exhibit this type of methylation due to the absence of the enzyme activity capable of accomplishing this modification in this bacterium. It is, however, possible to methylate the CpGs of plasmid DNA extracted from *E. coli* in a test tube with an appropriate enzyme. Under these conditions, DNA methylated *in vitro* loses a great deal of its immunostimulant activity compared to control nonmethylated DNA. The *E. coli* strain K12, from which virtually all the mutant strains used to produce

plasmid DNAs are derived, contains an enzyme activity (DNA methylase dcm (Palmer B. R. and Marinus M. G. (1994) *Gene* 143, 1-12)) which leads to methylation of cytosine occurring in the nucleic acid context CC(A/T)GG. All plasmid vectors for gene transfer contain this sequence in varying number and, as a result, their DNA molecule contains methylated cytosines which are not found in mammalian DNA. This form of methylation specific to *E. coli* thus introduces another difference into the cytosine methylations between bacterial DNA and that of mammals, which might contribute to the immunostimulant capacity of plasmid DNA.

The CpG frequency in primate and rodent DNAs is, overall, much lower than that expected on the basis of the frequency of cytosines and guanines. The CpG deficiency is dependent, for a given DNA fragment, on the biological role of this fragment, intergenic regions containing only a fifth of the expected frequency, while exons have a less marked deficiency and, at the other extreme, some promoters containing a large CpG island exhibit a CpG percentage close to that expected. Analysis of the data from sequencing human cDNAs and chromosomes reveals, however, broad heterogeneities in the CpG frequency for promoter regions and cDNAs. This observation is illustrated by the cDNA of the human gene encoding interleukin 2, which has just one CpG. Similarly, a portion of the promoter of this gene containing the TATA box does not contain any CpG, but, on the other hand, the upstream portion rich in transcription factor recognition sites contains CpGs. The regions positioned 3' of the genes, formed by the 3' UTRs (untranslated regions) and the polyadenylation and end of transcription sequences are rather poor in CpG. In humans, it is not unusual to find regions immediately downstream of the genes, which are devoid of CpG. However, the human sequencing data available at the end of 2000 have not made it possible

to demonstrate a single transcriptional unit, made up of the transcription promoter regions, a gene with or without an intron and the polyadenylation region, which is completely devoid of CpG. The situation of the CpGs in *E. coli* is quite different from that of animal cells since the frequency of CpGs in the genomic DNA of this bacterium is slightly greater compared to the calculated frequency. The same is true for the CpGs of natural plasmids isolated from hospital strains of *E. coli*. The recombined plasmids resulting from genetic manipulations, used for gene transfer, exhibit variations in their CpG numbers which depend on the origin of the fragments inserted into the vector. Analysis of the sequences of several tens of recombined *E. coli* plasmids randomly taken from the GenBank databank shows that the plasmids most lacking in CpG have, at the very most, a 50% deficiency in the number of their CpGs.

As regards the present invention, it provides products and methods for synthesizing plasmid DNA in *E. coli* which is completely devoid of CpG and in which the cytosines placed in the context CC(A/T)GG are not methylated. To the applicant's knowledge, this is the first description of such products which exhibit such a structure while at the same time having conserved their function.

DESCRIPTION OF THE INVENTION

30

The present invention provides means for producing plasmids which are functional in a prokaryotic organism such as *Escherichia coli*, but which are nevertheless completely devoid of CpG. More particularly it provides means for producing plasmids which are completely devoid of CpG, and which are also free of cytosine methylation in the nucleic acid context CC(A/T)GG.

The present application thus relates to methods for producing such plasmids, and also to the elements constituting these plasmids, namely genes devoid of CpG which can be expressed in *E. coli*, promoters devoid of
5 CpG which are suitable for the expression of said genes, and origins of replication devoid of CpG which are suitable for the bacterial transformation of said plasmids. The present application is also directed towards the biotechnological and medical applications
10 of these products. Each one of these products has the particular characteristic of being completely devoid of CpG, while at the same time having conserved its functionality in a prokaryote such as *E. coli*.

The present application also provides an *E. coli* strain
15 specially suited to the production of the plasmids according to the invention, this strain having the particular characteristic of allowing stable replication of these plasmids and of the genetic material which they transport, without impairing their
20 function, and without, however, inducing methylation at CC(A/T)GG sites (strain comprising an inactivated *dcm* gene).

One of the common concepts linking the various aspects of the invention is therefore to make it possible to
25 produce plasmids which are completely devoid of CpG and which have, despite everything, conserved their functional properties in a prokaryote such as *E. coli*. To the applicant's knowledge, this is the first description of such means.

30 The present application is thus directed toward a method for producing a plasmid which is a vector of at least one gene, and which is completely devoid of CpG characterized in that a plasmid is constructed by
35 assembling, by enzyme ligation, DNA fragments, all devoid of CpG, corresponding to an origin of replication for the plasmid and to elements constituting a transcriptional unit for said at least one gene, and in that this plasmid is transferred into

an *Escherichia coli* strain expressing the *pi* protein for replication of the plasmid.

Plasmids isolated from wild-type bacterial strains
5 generally accomplish three functions in relation to replication, namely initiation of DNA replication, control of replication and stable maintenance of the plasmid during successive divisions. Plasmids constructed in the laboratory do not always exhibit all
10 of these functions. The number of copies of the plasmids is, for example, quite often increased compared to the parent plasmid, denoting that replication control elements have been modified. The plasmid R6K contains three origins of replication,
15 *alpha*, *gamma* and *beta*, linked on the same DNA fragment (Filutowicz M. and Rakowski S. A. (1998) *Gene* 223, 195-204). Each one of the origins is activated by the R6K specific *pi* initiation protein encoded by the R6K *pir* gene. In order to be functional, the three origins need
20 a 277 bp sequence, known by those skilled in the art as "core", located at the center of the fragment carrying the three origins, and also an additional single fragment positioned in *cis*, i.e. present on the same DNA molecule. When the sequences of the *alpha* and *beta*
25 origins are deleted, the remaining *gamma* origin allows autonomous duplication of the plasmid on the condition that the *pir* gene is present in *cis* on the plasmid or in *trans* on the chromosome of the bacterium. The inventors chose to focus more particularly on the
30 smallest of the three origins, namely the *gamma* origin, which has the advantage of containing all the elements required for controlled replication of the plasmid, namely the core and an adjacent activating sequence. The core is made up of a *pi* protein-binding sequence
35 repeated 7 times, and an AT-rich sequence. The activating region contains binding sites for several cellular proteins of the bacterium, required for stably maintaining the plasmid. The number of copies of the plasmids containing only the *gamma* origin depends on

the *pi* protein; mutant forms of *pi* leading to a large increase in the number of copies of the plasmid have been isolated and characterized. As shown in greater detail in the examples below, the inventors have
5 succeeded in constructing, from the gamma origin of replication of the plasmid R6K, origins of replication which no longer exhibit any CpG, while at the same time having conserved their functionality intact. It may be noted that the specific choice of R6K gamma as starting
10 material, namely the choice of a small replicon which exhibits only a small number of CpG, is a particularly relevant choice insofar as, when other plasmids such as those of the pUC series are used as starting material, CpG-free plasmids which remain functional are not
15 successfully obtained: all the attempts made by the inventors to chemically reconstitute the minimum pUC sequence by replacing the cytosines of CpGs with guanine or adenine resulted in DNA fragments which had lost all functional replication activity. As regards
20 the plasmids which comprise a CpG-free origin of replication according to the invention, they have conserved their ability to replicate stably within a prokaryote cell such as *E. coli*, and *E. coli* K12 in particular, provided, of course, that they are provided
25 with the *pi* protein required to activate the replication (wild-type *pir* or mutated *pir* such as *pir* 116 in *cis* or in *trans*). An origin of replication for plasmid according to the invention is characterized in that its sequence corresponds to that of the R6K gamma
30 origin of replication in which each G of the CpGs of the repeat-region of the core has been replaced with an A, a C or a T, or each C of the CpGs has been replaced with a G, an A or a T. Various CpG-free origins of replication have thus been obtained, which,
35 surprisingly, are still capable of performing the functions of origin of replication of plasmids in *E. coli* and, what is more, are capable of performing these functions for genes and transcriptional units which are themselves devoid of CpG. The examples below

give some illustrations thereof (cf. origins R6K gamma M2A, R6K gamma M2C, R6K gamma M2T in examples 7-10). The present application is directed more particularly toward any origin of replication whose sequence
5 comprises the sequence SEQ ID No. 12 or SEQ ID No. 13 (Figures 12 and 14). It has also been demonstrated that the pi protein-binding sequence can be not repeated 7 times, as is observed in the standard R6K gamma origin with CpG, but that the number thereof can be limited to
10 5 or 6, without however impairing the functions of the origin of replication. The present application is thus directed toward any origin of replication according to the invention as defined above, which would comprise only 5 or 6 repeats of the pi protein-binding sequence.
15 By virtue of these CpG-free functional origins of replication, the inventors have been able to construct various plasmids which, notably, have conserved their transfection vector functions.

20 The creation of *E. coli* plasmids devoid of CpG necessarily requires having functional genes (which can be expressed in prokaryotes such as *E. coli*) which do not contain any CpG. Thus, selection of the bacteria transformed with a recombined plasmid DNA involves a
25 gene whose protein confers a dominant advantage on the bacterium. Most commonly, the selective marker is introduced by a gene for resistance to an antibiotic which is active on the *E. coli* bacterium. Analysis of the wide variety of resistance genes used in *E. coli*
30 shows that, without exception, they all contain CpGs, very often in very high numbers for resistance genes originating from *Streptomyces* which produce the antibiotic for selection. Similarly, it is necessary to have reporter genes which are CpG-free while at the
35 same time remaining functional.

Analysis of several hundred chromosomal and plasmid genes of the *E. coli* bacterium, the well-characterized sequences of which are available in several databanks,

reveals that all genes greater than 250 pb in size, without exception, consist of 16 dinucleotides.

The present invention demonstrates that it is, despite
5 everything, possible to construct genes which are functional in *E. coli* and which are devoid of CpG. The inventors have in fact developed a method for obtaining genes devoid of CpG while at the same time being able to be expressed in *E. coli*. This method is based on the
10 synthesis of a polynucleotide chain by following the amino acid chain of a protein which can be expressed in *E. coli*, assigning to each amino acid a nucleotide codon chosen from those which, according to the genetic code, and taking into account the degeneracy of this
15 code, correspond to this amino acid, but eliminating from this choice:

i. all codons containing a CpG in their sequence: this concerns the codons ACG (Thr), CCG (Pro), GCG
20 (Ala), TCG (Ser), CGA (Arg), CGC (Arg), CGG (Arg) and CGT (Arg), and

ii. codons which finish with a C when the codon which follows it directly begins with a G. Examples of a gene
25 thus obtained comprise the Neo Δ CpG gene (SEQ ID No 316; cf. example 11).

According to one variant of implementation of the invention, the codons for which the frequencies are low
30 in proteins of human origin will also be eliminated from said choice: this concerns the codons ATA (Ile), CTA (Leu), GTA (Val) and TTA (Leu). The set of possible codons therefore, according to this variant, corresponds to the following set:

35 GCA (Ala), GCC (Ala), GCT (Ala), AGA (Arg), AGG (Arg), AAC (Asn), AAT (Asn), GAC (Asp), GAT (Asp), TGC (Cys), TGT (Cys), CAA (Gln), CAG (Gln), GAA (Glu), GAG (Glu), GGA (Gly), GGC (Gly), GGG (Gly), GGT (Gly), CAC (His), CAT (His), ATC (Ile), ATT (Ile), CTC (Leu), CTG (Leu),

CTT (Leu), TTG (Leu), AAA (Lys), AAG (Lys), TTC (Phe),
TTT (Phe), CCA (Pro), CCC (Pro), CCT (Pro), TCA (Ser),
TCC (Ser), TCT (Ser), AGC (Ser), AGT (Ser), ACA (Thr),
ACC (Thr), ACT (Thr), TAC (Tyr), TAT (Tyr), GTC (Val),
5 GTG (Val), GTT (Val),
to which rule ii. above should of course be applied.
Examples of a gene obtained in accordance with this
variant of implementation comprise in particular the
LacZΔCpG gene (positions 3 to 3056 of SEQ ID No 9; cf.
10 example 5).

Preferably, said choice of codon will also be made so
as to avoid structures which are unfavorable for the
messenger RNA, such as the presence of splice
15 sequences, of direct or inverted repeat sequences, of
stem-loop structures or of polyadenylation signals. The
number and the size variety of the genes synthesized by
this method are illustrated in the examples below,
which show that it is thus possible to envision the
20 synthesis of genes completely devoid of CpG which
nevertheless remain functional in *E. coli*. As a
reference protein, any protein which can be expressed
by *E. coli* can be chosen, for example a protein encoded
by a gene for resistance to an antibiotic, such as the
25 genes for resistance to zeocin® (phleomycin), to
hygromycin, to blasticidin or to puromycin, or a
protein encoded by reporter genes such as *lacZ*.

The present application is also directed toward such a
30 method for obtaining genes devoid of CpG, which can be
expressed in *E. coli*, and also any gene of at least
250 bp which can be obtained using this method. More
particularly, the present application is directed
toward:

35

- any gene the sequence of which comprises the
sequence SEQ ID No. 1 from position 3 to
position 374 (Figure 1), the sequence SEQ ID
No. 3 from position 3 to position 1025 (Figure

3), the sequence SEQ ID No. 5 from position 3 to position 422 (Figure 5), the sequence SEQ ID No. 7 from position 3 to position 599 (Figure 7), and also any use of these genes as selection markers, and

5

- any gene the sequence of which comprises the sequence SEQ ID No. 9 from position 3 to position 3056 (cf. Figure 9), and any gene the sequence of which comprises the sequence SEQ ID No. 316 (positions 3 to 797 of the DNA sequence presented in Figure 18 (encoding SEQ ID No. 317)), and also any use of such a gene as a reporter gene.

10

15

The expression of a plasmid gene also requires having promoters suitable for the cell harboring the plasmid. The fact that the *E. coli* genome has been entirely known for a few years has facilitated the study of various noncoding elements exhibiting specific functions. Results from studies relating to the nature of the *E. coli* promoters are continually being updated and made public on the PromEC site accessible via the Internet (<http://bioinfo.md.huji.ac.il/marg/promec>). An analysis of the 471 well-defined promoters from base -75 to +25 relative to the transcription initiation point +1 reveals that only 6 of them do not possess CpG. The addition of each one of the 6 promoters synthesized chemically and placed upstream of the lacZ gene encoding *E. coli* β -galactosidase has proved to be negative for detection of the activity of this reporter gene. The lack of strong homology with the consensus sequences of the -10 and -35 canonic boxes suggests that these promoters are low strength promoters and, alternatively, that said promoters might be regulated by induction conditions yet to be defined for each one of them. An analysis of the well-characterized promoters of *E. coli* has revealed that only about 10 do not contain any CpG in the specific boxes for

20

25

30

35

recognition by RNA polymerase. A bibliographical search has also revealed that these promoters are all inducible by various stimuli, a situation which is sometimes desired but most often relinquished for
5 expression which is constitutive in nature. As regards the inventors, they have succeeded, by random PCR assembly of fragments exhibiting short consensus sequences devoid of CpG, drawn from several strong promoters, in developing novel promoters which are
10 suitable for the expression of genes without CpG in *E. coli*, and which have the particular advantage of being very strong constitutive promoters and of being completely devoid of CpG. Example 6 below illustrates the construction of the novel promoter EM2K using this
15 technology. The present application is more particularly directed toward any promoter, the sequence of which comprises the sequence SEQ ID No. 11 (cf. Figure 11).

20 The characterized transcription terminators in *E. coli* are made of short sequences, several of which do not possess any CpG, and the inventors have been able to verify that such terminators effectively perform their function when they are associated, in *E. coli*, with a
25 CpG-free promoter and gene according to the invention.

The present application is thus directed toward any transcriptional unit which comprises at least one CpG-free gene according to the invention, and at least one
30 CpG-free promoter according to the invention. Such a transcriptional unit may also comprise at least one CpG-free terminator. The invention thus provides, for the first time, a nucleotide group which is completely devoid of CpG, and which can nevertheless be expressed
35 in *E. coli*, performing its normal functions therein.

The present application is thus directed toward any plasmid which comprises an origin of replication according to the invention. Such plasmids may also

comprise a CpG-free gene according to the invention and/or a CpG-free promoter according to the invention and/or a CpG-free transcription terminator, or a transcriptional unit according to the invention. The
5 plasmids according to the invention therefore have the advantage of being able to exhibit no CpG in their structure, while at the same time still being capable of performing expression vector functions. Examples of such plasmids are given in the examples below.

10

The present application is more particularly directed toward any plasmid of SEQ ID No. 14 (Figure 15).

Any cell transformed with at least one element selected
15 from the group consisting of the CpG-free genes according to the invention, the CpG-free promoters according to the invention, the CpG-free origins of replication according to the invention and the CpG-free plasmids according to the invention also falls within
20 the field of the present application. Such a cell according to the invention may also comprise a gene encoding a *pi* protein, such as wild-type *pir* or mutated *pir*, *pir 116*. Advantageously, a transformed cell according to the invention is an *E. coli* cell.

25

To replicate a sufficient number of functional copies of a plasmid according to the invention, those skilled in the art have at their disposal many bacteria, such as, for example, *E. coli* K12 bacteria which are
30 conventionally used for the purposes of plasmid replication. The original K12 strain of *E. coli* has a DNA methylase which introduces a methyl group onto all the cytosines placed in the context CC(A/T)GG of the genomic and plasmid DNAs of the bacterium. All of the
35 various strains of the K12 line have this activity due to a methylase encoded by the *dcm* gene (Palmer B. R. and Marinus M. G. (1994) *Gene* 143, 1-12). Since methylation of the plasmid DNAs prepared from *dcm*⁺ strains of *E. coli* leads to a modification of the DNA

molecule which is undesirable for the transfer of genes into eukaryote cells, the inventors have developed a strain which makes it possible both for plasmids with the CpG-free R6K *gamma* origin in accordance with the invention to function, and plasmid DNA devoid of methylation on the *dcm* sites to be obtained. For this, a new gene was constructed by the inventors, by deleting the *dcm* gene from position +3 after the ATG to position -14 before the TGA in the *dcm* gene of a *pir* 116 strain (cf. Example 10 below).

The *dcm* gene is located in a chromosomal region the sequence of which can be obtained via GenBank with the accession number D90835 (clone Kohara #344: 43.5-43.9 min).

A deletion in the gene (-) introduces the additional advantage of avoiding any reversion of the gene to the wild-type form. *dcm*⁻ mutant strains have thus been produced by the inventors; they exhibit no negative phenotype which might impair bacterial growth or modify the quality and quantity of the plasmid DNA. More particularly, an optimized strain of *E. coli* has been constructed by targeted inactivation of the *dcm* gene using a parent strain expressing a mutated *pi* protein at a site leading to an increase in the number of copies of CpG-free plasmids. This optimized strain allows quality and abundant production of the plasmid DNAs which are the subjects of this invention, and which are devoid of CpG and free of methylation on the cytosines of the *dcm* sites. The present application is therefore directed toward any cell comprising a gene encoding the *pi* protein, which is transformed with the deleted *dcm* gene according to the invention, and any method of replicating plasmids, which comprises transforming such a cell with a plasmid, and culturing the transformed cell under conditions suitable for replication of this plasmid.

The present application is thus directed toward a method for producing a plasmid completely devoid of CpG and free of methylation on cytosine in the nucleic acid context CC(A/T)GG, characterized in that a plasmid according to
5 the invention is produced by replication in an *Escherichia coli* strain expressing the *pi* protein, which is deficient for the *dcm* methylation system.

Any kit for producing plasmids, which comprises at least
10 one cell according to the invention, also falls within the field of the present application. These kits are particularly suitable for the replication of plasmids according to the invention, in order to avoid these plasmids, the structure of which is devoid of CpG, being,
15 moreover, methylated on CC(A/T)GG during their replication.

The invention thus provides a complete set of transformation means devoid of CpG; CpG-free genes, CpG-
20 free promoters, CpG-free transcriptional units, CpG-free origins of replication for plasmid, CpG-free plasmids, cells specially suited to replication of plasmids without methylation of cytosines. These novel means find direct applications for genetic transformation of cells for
25 biotechnological or medical purposes. Such products are in fact exceptionally well suited to the production of DNA vaccine compositions intended for humans or animals.

In the claims which follow and in the description of the
30 invention, except where the context requires otherwise due to express language or necessary implication, the word "comprise" or variations such as "comprises" or "comprising" is used in an inclusive sense, i.e. to specify the presence of the stated features but not to
35 preclude the presence or addition of further features in various embodiments of the invention.

The invention is illustrated by the following examples, in which reference is made to the figures:

- 5
- Figure 1: Sequence of the Sh ble Δ CpG gene (CpG-free),
 - Figure 2: List of the oligonucleotides used for assembling the Sh ble Δ CpG gene,
 - Figure 3: Sequence of the Hph Δ CpG gene,

- Figure 4: List of the oligonucleotides used to assemble the Hph Δ CpG gene,
- Figure 5: Sequence of the Bsr Δ CpG gene,
- 5 - Figure 6: List of the oligonucleotides used to assemble the Bsr Δ CpG gene,
- Figure 7: Sequence of the Pac Δ CpG gene,
- Figure 8: List of the oligonucleotides used to assemble the Pac Δ CpG gene,
- Figure 9: Sequence of the LacZ Δ CpG gene,
- 10 - Figure 10A: List of the oligonucleotides used to assemble the first third of the LacZ Δ CpG gene,
- Figure 10B: List of the oligonucleotides used to assemble the second third of the LacZ Δ CpG gene,
- 15 - Figure 10C: List of the oligonucleotides used to assemble the third third of the LacZ Δ CpG gene,
- Figure 11: Sequence of the EM7 promoter (1-), and of the degenerative oligonucleotides (2-) used to construct the EM2K promoter devoid of CpG (3-),
- 20 - Figure 12: Sequence of the R6K gamma M2A origin of replication,
- 25 - Figure 13: List of the oligonucleotides used to assemble the R6K gamma M2A origin of replication,
- Figure 14: Sequence of the R6K gamma origin of the plasmid pGTR6Kneoc9 delimited by the PacI sites,
- 30 - Figure 15: Sequence of the plasmid pSh-LacZ Δ CpG gene,
- Figure 16: Map of the plasmid pSh Δ CpG,
- Figure 17: Map of the plasmid pSh-LacZ Δ CpG,
- 35 - Figure 18: Sequence of the Neo Δ CpG gene (CpG-free) [position 3 to 797 of DNA sequence = SEQ ID No. 316; protein sequence = SEQ ID No. 317),

- Figure 19: Sequence of the oligonucleotides SEQ ID No. 318 to SEQ ID No. X used to assemble the Neo Δ CpG gene.

5 EXAMPLE 1: Construction of the Sh ble gene for resistance to zeocin devoid of CpG

The Sh ble Δ CpG gene, the sequence of which is given in Figure 1 (positions 3 to 377 of SEQ ID No. 1) was
10 synthesized from an assembly of overlapping oligonucleotides (20-40 pb in size), the sequences of which are given in Figure 2. The assembly method is carried out in three steps; the first step consists of phosphorylation of the oligonucleotides of the coding
15 strand, in a second step, all the oligonucleotides of both strands are combined by hybridization and ligation and, in the final step, the gene is amplified by PCR. This method was successfully used to synthesize all the synthetic genes mentioned in Examples 1, 2, 3, 4 and 5.
20 Details of the method are given for the Sh ble Δ CpG gene:

The 10 oligonucleotides, from OL26199 to OL27099 (Figure 2), corresponding to the coding strand, are phosphorylated according to the following procedure:
25 1 μ l of each one of the oligonucleotides taken up in water at 250 μ M are mixed in a microtube containing 15 μ l of water so as to bring the final solution to a concentration of 100 picomol per microliter. 5 μ l of this solution are then mixed with 10 μ l of 10-times
30 concentrated polynucleotide kinase buffer, 0.4 μ l of a 50 mM ATP solution, 85 μ l of water and 1 μ l of the enzyme (at 10 μ /l), and the entire mixture is incubated for 4 hours at 37°C (solution A).

35 A solution of the oligonucleotides of the noncoding strand is made up by mixing 1 μ l of each oligonucleotide (OL27199 to OL28199; cf. Figure 2) and 1 μ l of the oligonucleotide OL26099 (Figure 2), to which solution 43 μ l of water are added in order to

obtain a final solution at 54 picomol per μ l (solution B).

Assembly of the gene is carried out first by mixing
5 10 μ l of solution A, 1 μ l of solution B, 6 μ l of a
100 mM KCl solution, 3 μ l of a 0.5% NP-40 solution,
4 μ l of a 50 mM $MgCl_2$ solution, 3 μ l of a 10 mM ATP
solution and 7.5 μ l of Pfu ligase (30 units), and the
mixture is then heated in a programmable thermocycler
10 for 3 minutes at 95°C and then 3 minutes at 80°C, before
undergoing 3 cycles of one minute at 95°C, followed by
a change from 95°C to 70°C in 1 minute, and then a
change from 70°C to 55°C in 1 hour and, finally, 2 hours
at 55°C. The mixture of the assembled oligonucleotides
15 is then amplified with the primers OL26099 and OL27199.
The amplification product is purified on a promega
column, digested with the NcoI and NheI restriction
enzymes, and cloned into the plasmid pMOD1LacZ(wt)
linearized with NcoI and NheI. The sequences of the
20 plasmid DNA of 2 zeocin-resistant clones which appeared
after transformation of the *E. coli* strain GT100
(available from Invivogen) with the mixture of the
ligation between the vector fragment and the PCR
fragment, were found to be in accordance with the
25 desired sequence given in Figure 1. This synthetic
gene, placed under the control of the bacterial EM7
promoter (vector pMOD1Sh Δ CpG) confers a zeocin
resistance identical to that provided by the same
vector containing the native *Sh ble* gene with the
30 recipient *E. coli* strain GT100.

EXAMPLE 2: Construction of the Hph gene for resistance
to hygromycin devoid of CpG

35 The synthetic Hph Δ CpG (sequence SEQ ID No. 3 given in
Figure 3) was constructed according to the method
described in Example 1. The two strands were
synthesized using oligonucleotides of 60 bases plus two
oligonucleotides of 30 bases with an overlapping region

of 30 bases. The assembly of the various oligonucleotides given in Figure 4 was carried out through a final PCR with the sense oligonucleotide TTCAGCTGAGGAGGCACATC (SEQ ID No. 299) and the reverse
5 oligonucleotide CTCAGGATCCGCTAGCTAAT (SEQ ID No. 300) according to the experimental conditions mentioned in the preceding example.

The amplified and purified fragment (1068 pb) was then
10 digested with the BspHI and NheI restriction enzymes and cloned into the vector pMOD2 LacZ(wt), in which the NcoI site of pMOD1 is replaced with the BspHI site. The *E. coli* clones containing this recombinant vector were selected on FastMedia™ Hydro Agar medium (Cayla). The
15 sequence SEQ ID No. 3 in Figure 3 was confirmed by sequencing on both strands of the plasmid DNA of two hygromycin-resistant clones. This synthetic gene, placed under the control of bacterial EM7 promoter (vector pMOD2 Hph ΔCpG), confers a hygromycin B
20 resistance which is at least equal to that provided by the same vector containing the native Hph gene with the recipient *E. coli* strain GT100.

EXAMPLE 3: Construction of the Bsr gene for resistance
25 to blasticidin devoid of CpG

The Bsr ΔCpG gene, the sequence of which is given in Figure 5 (SEQ ID No. 5), was synthesized using the oligonucleotides indicated in 6, by following the
30 method described in Example 1. The mixture of the assembled oligonucleotides was amplified with the primers OL64 and OL76 (cf. Figure 6). The amplified and purified fragment was then digested with the BspHI and NheI restriction enzymes and cloned into the vector
35 pMOD2 LacZ(wt). The *E. coli* clones containing this recombinant vector were selected on FastMedia™ Blast Agar medium (Cayla). The sequence SEQ ID No. 5 in Figure 5 was confirmed by sequencing on both strands of the plasmid DNA of two blasticidin-resistant clones.

This synthetic gene, placed under the control of the bacterial EM7 promoter (vector pMOD2 Bsr Δ CpG), confers a resistance to blasticidin which is identical to that provided by the same vector containing the native Bsr gene with the recipient *E. coli* strain GT100.

EXAMPLE 4: Construction of the Pac gene for resistance to puromycin devoid of CpG

- 10 The BspHI-NheI fragment (Pac Δ CpG gene; SEQ ID No. 7), the sequence of which is given in Figure 7, was synthesized by assembling the oligonucleotides indicated in Figure 8.
- 15 The mixture of the assembled oligonucleotides was amplified with the sense primer pur24 (AGGACCATCATGACTGAG; SEQ ID No. 301) and the reverse primer pur25 (ATCATGTCGAGCTAGCTC; SEQ ID No. 302). The purified BspHI-NheI fragment was cloned into the
- 20 plasmid pMOD2LacZ (wt) between the BspHI and NheI sites. The sequences of the plasmid DNA of 2 puromycin-resistant clones of the GT100 strain, which appeared on the FastMedia™ puro Agar medium (Cayla) after transformation by the product of the ligation between
- 25 the vector fragment and the PCR fragment, were found to be in accordance with the desired sequence given in Figure 7. The synthetic Pac Δ CpG gene, placed under the control of the bacterial EM7 promoter (vector pMOD2 Pac Δ CpG), confers a puromycin resistance which is slightly
- 30 greater than that provided by the same vector containing the native *pac* gene with the recipient *E. coli* strain GT100.

EXAMPLE 5: Construction of the LacZ gene devoid of CpG encoding β -galactosidase of *E. coli*

The synthetic LacZ Δ CpG gene (SEQ ID No. 9 given in Figure 9) was constructed according to the method described in the preceding examples. Given the size of

the gene to be produced (more than 3000 pb), the construction was carried out in 3 distinct parts, conserving the EcoRV and SacI restriction sites at the same sites as on the native sequence of the *lacZ* gene.

5 For each part, the two strands were synthesized using oligonucleotides of 40 bases plus two oligonucleotides of 20 bases with an overlapping region of 20 bases.

The first region corresponds to the NcoI-EcoRV fragment
10 (Part I), the second region corresponds to the EcoRV-SacI fragment (Part II) and the third region corresponds to the SacI-NheI fragment. The assembly of the various oligonucleotides given in Figures 10A (oligonucleotides used to assemble part I), 10B
15 (oligonucleotides used to assemble part II), and 10C (oligonucleotides used to assemble part III) was carried out by PCR according to the same experimental conditions stated in the preceding examples. The gradual cloning of the three parts of the synthetic
20 gene was carried out in the vector pMOD1 *LacZ*(wt). The functionality of each cloned part and also that of the complete synthetic gene present on the vector pMOD1 *LacZ* was demonstrated by revealing the β -galactosidase activity on FastMedia™ Amp Xgal Agar medium (Cayla), of
25 the recombinant clones obtained in the MC1061 Δ Lac strain. The complete synthetic *LacZ* Δ CpG gene, placed under the control of the EM7 promoter, gives 30% less β -galactosidase activity (luminometric assay of protein extracts from culture) compared to the expression of
30 the native *LacZ* gene in the same plasmid environment.

EXAMPLE 6: Construction of a strong constitutive promoter for *E. coli*, devoid of CpG

35 The bacterial EM7 promoter present on vectors of the pMOD1 type is a synthetic promoter which is constitutive and strong in *E. coli*. Its sequence, which contains 3 CpG (SEQ ID No. 297 in Figure 11), was used as a reference to produce a bacterial promoter devoid

of CpG. We produced "linker" oligonucleotides which were degenerative at 4 places (indicated W, D, W and H on the sequence SEQ ID No. 298 in Figure 11) and compatible with the AseI and NcoI restriction sites.

5 These various oligos were hybridized and cloned into pMOD1 ShΔCpG between the AseI and NcoI restriction sites of the EM7 promoter. After selection of the recombinant clones on FastMedia™ Zeo Agar medium and determination of the promoter sequence of the most

10 zeocin-resistant clone, we selected the EM2K promoter (sequence SEQ ID No. 11 in Figure 11) as the bacterial promoter devoid of CpG.

EXAMPLE 7: Synthesis of the R6K *gamma* origins devoid of

15 CpG

The PacI DNA fragment containing the R6K gamma M2A origin (SED ID No. 12 in Figure 12) was synthesized by PCR from the assembly of the oligonucleotides indicated

20 in Figure 13. The R6K gamma M2A fragment assembly was amplified with the primers RK15 (GCAGGACTGAGGCTTAATTAAACCTTAAAC; SEQ ID No. 303) and RK16 (AAGTCTCCAGGTAAATTAAGATCAGCAGTTC; SEQ ID No. 304), and the fragments, after digestion with the PacI

25 enzyme, were cloned into a plasmid (pGTCMVneo) containing the kanamycin resistance gene and the pUC origin of replication bordered by 2 PacI sites. Many transformants of the GT97 strain (which expresses the pi protein) were analyzed and only clones containing a

30 high-copy plasmid conserved after several rounds of subculturing in the absence of kanamycin were selected. After sequencing, it was found that the ori fragment of most of these plasmids could have a lower number (5-6) of repeat sequences, instead of the 7, of the natural

35 origin of the R6K plasmid. One of these novel sequences of the synthetic R6K gamma origin devoid of CpG is given in SEQ ID No. 13 in Figure 14).

Two other versions of the R6K gamma origin, in which the G of each CpG present in the repeat sequences (22 bp element repeated several times in the pi protein-binding region) has been replaced with a C, to give the origin (R6K gamma M2C), or a T, to give the origin (R6K gamma M2T), were synthesized in a similar manner. The functionality of these novel R6K gamma origins in which the G of the CpGs of the repeat sequences is replaced with a C or with a T, added to the example of the origin of Figure 13, in which the G is replaced with an A, demonstrates that the CpGs of these repeat sequences do not play a role in the functionality of the origin.

EXAMPLE 8: Assembly of plasmid vectors completely devoid of CpG, expressing a gene for resistance in *E. coli*

Firstly, a PacI-PacI cassette containing the bacterial EM2K promoter and the Sh Δ CpG zeocin resistance gene followed by a CpG-free bacterial terminator was prepared. For this, "linker" oligonucleotides containing the sequence of the t1 terminator of the intergenic region *rpsO-pnp* of *E. coli* were hybridized and cloned between the NheI and PacI sites of the vector pMOD1 EM2K Sh Δ CpG:

"linker" oligonucleotides:

rpsO-1(5'→3'):

CTAGCTGAGTTTCAGAAAAGGGGGCCTGAGTGGCCCCTTTTTTCAACTTAAT

SEQ ID No. 305

rpsO-2(5'→3'):

TAAGTTGAAAAAAGGGGCCACTCAGGCCCTTTTCTGAAACTCAG

SEQ ID No. 306.

The recombinant vector obtained (pMOD1 EM2K sH Δ CpG Term) was verified by sequencing in the region of the terminator sequence which does not naturally contain any CpG. The EM2K-Sh Δ CpG-Term cassette contained in this vector was then amplified by PCR so as to flank

the two sides with PacI sites using the following primers:

PACI-UP (5'→3'):

ATCGTTAATTAAAACAGTAGTTGACAATTAAACATTGGC

5 SEQ ID No. 307

PACI-DOWN (5'→3'):

ATCGTTAATTAAAGTTGAAAAAAGGGGCC

SEQ ID No. 308.

10 This amplified fragment was then purified and cleaved with PacI, and then assembled with PacI fragment containing the R6K gamma ΔCpG origin described in Example 7. After transformation of this ligation mixture into the GT97 strain (which expresses the pi
15 protein) and selection on FastMedia™ Zeo medium, analysis of the recombinant clones obtained revealed two possible orientations of the PacI-PacI fragment containing the R6K gamma ΔCpG origin. The orientation selected in pSh ΔCpG is represented in Figure 16.

20

EXAMPLE 9: Assembly of a plasmid vector completely devoid of CpG, expressing the zeocin resistance gene and the β-galactosidase gene in *E. coli*.

25 The vector pSh ΔCpG described in Example 8 (Figure 16) was used to insert the synthetic LacZ gene devoid of CpG between the EcoRI and NheI sites. For this, EcoRI- and NcoI-compatible "linker" oligonucleotides containing a ribosome-binding site consensus sequence
30 from *E. coli* were hybridized and cloned with the NcoI-NheI LacZΔCpG fragment of pMOD1 LacZΔCpG, between the EcoRI and NheI sites of the vector pMOD1 EM2K ShΔCpG.

"linker" oligonucleotides used:

35 rbs-1 (5'→3'): AATTCTGAGGAGAAGCT SEQ ID No. 309

rbs-2 (5'→3'): CATGAGCTTCTCCTCAG SEQ ID No. 310

Transformation of this ligation mixture into the GT97 strain (which expresses the pi protein) and selection

on FastMedia™ Zeo Xgal medium made it possible to obtain the recombinant clones containing the vector pSh-LacZΔCpG (Figures 15 and 17). This vector co-expresses, under the control of the bacterial EM2K promoter, in an artificial operon system, the ShΔCpG and LacZΔCpG genes.

EXAMPLE 10: Production of an *E. coli* strain expressing the mutant protein *pir116* and carrying a deletion in the *dcm* gene

The *pir* gene encoding the *pi* protein which is essential for initiating replication of the R6K *gamma* origin, and also the mutated gene *pir116* which leads to an increase in the number of copies of R6K *gamma* plasmids, have been introduced, in a functional form, into various *E. coli* K12 strains by various groups. Strains of this type can be obtained from the *E. coli* Genetic Stock Center (<http://cgsc.biology.yale.edu>), and are also commercially available from companies specializing in supplying biological material for research. This is the case, for example, of the *pir1* (*pir116*) and *pir2* (wild-type *pir*) strains provided by the company Invitrogen, whose products can be purchased in all European countries. The GT97 strain of the K12 line, which has the genotype $\Delta lac169$ *hsdR514* *endA1* *recA1* *codBa* *uidA* (Δ *lul1*)::*pir 116* (available from InvivoGen), was chosen for its simplicity, the consistency of the R6K *gamma* plasmid DNA preparations and its high levels of competence, from several K12 *pir* strains of distinctive genotype for some genes. The introduction of a deletion into the *dcm* gene of the GT97 strain was carried out in the following way:

Two DNA regions, of 1.8 kb and 1.5 kb, flanking respectively the ATG initiation codon (fragment A) and the TGA stop codon (fragment B) of the *dcm* gene were amplified by PCR. The fragment A was then amplified with the pair of primers OLDcmAF (TTTTCGCGCCGCTTGCTGCGCCAGCAACTAATAACG; SEQ ID No. 311)

and OLdcmAR (CCTTGGATCCTGGTAAACACGCACTGTCCGCCAATCGATTC;
SEQ ID No. 312) and fragment B was amplified with the
pair of primers OLdcmBF
(TTTGGATCCTCAGCAAGAGGCACAACATG; SEQ ID No.313) and
5 OLdcmBR (TTTCTCGAGAAACGGCAGCTCTGATACTTGCTTC; SEQ ID
No. 314). The restriction sites for the NotI
(GCGGCCGC), BamHI (GGATCC) and XhoI (CTCGAG) enzymes
were introduced into the primers in order to combine
10 fragment A and fragment B with one another, forming a
genetic element flanked by the NotI and XhoI sites. The
region of the *dcm* gene is thus reconstituted, creating
a deletion which stems from position +3 after the ATG
to position -14 before the TGA. This genetic element
was cloned into pKO3 (Link A. J., Phillips D. and
15 Church G. M. (1997) *J Bacteriol* 179, 6228-37), a vector
developed for allele replacement in *Escherichia coli*,
with thermosensitive replication, between the NotI and
SalI sites, to give the plasmid named pKO3 Δ dcm. The GT97
strain was co-transformed with this plasmid and with a
20 plasmid which expresses the RecA protein (pFL352). A
transformant containing the two plasmids was cultured at a
nonpermissible temperature (42°C) in the presence of
chloramphenicol in order to select clones which have
integrated pKO3 Δ dcm into the bacterial chromosome by
25 homologous recombination. A subclone resistant to
chloramphenicol at 42°C was then cultured at 30°C on a
medium containing a high concentration of sucrose (5%)
in order to counter-select the strains which, after a
second homologous recombination event, have exchanged
30 the chromosomal region of the *dcm* gene with the
homologous fragment cloned into the plasmid. The
deletion introduced into the selected clone (GT106) was
verified by PCR with the pair of primers OldcmAF and
OldcmBR, generating a fragment smaller in size than
35 that obtained with the parental strain, and by a PCR
with the primer OldcmBR and a primer positioned outside
the exchanged region (OldcmCF
TTTGGCGCCGCGTTGCGGTATTACCCTTGTC; SEQ ID No.315).

The *dcm*⁻ genotype of the GT106 strain was confirmed by introducing into said strain and into GT106, a plasmid containing a restriction site for the SexAI enzyme, which is subject to *dcm* methylation. The plasmid
5 purified from GT106 is cleaved by SexAI, whereas it is resistant to the enzyme when it is purified from GT97.

The latter strain named GT106 exhibits the same growth characteristics as the parental strain GT97 and, as
10 expected, no negative modification of the amount of R6K *gamma* plasmid DNAs was observed, only the quality of the DNAs, assessed by the absence of methylation of the cytosines of the *dcm* sites, was improved. The GT106 strain will be available from the company Invivogen
15 from the day on which this patent application is filed.

EXAMPLE 11: Production of the Neo gene for resistance to neomycin, devoid of CpG

20 The Neo Δ CpG gene, the sequence of which is given in Figure 18 (position 3 to 797 of the DNA sequence given in Figure 18 = SEQ ID No. 316; protein sequence = SEQ ID No. 317), was synthesized from an assembly of overlapping oligonucleotides (20-40 pb in size), the
25 sequences of which are given in Figure 19. The assembly method is carried out in three steps; the first step consists of phosphorylation of the oligonucleotides of the coding strand, in a second step, all the oligonucleotides of both strands are combined by
30 hybridization and ligation and, in the final step, the gene is amplified by PCR.

The 20 oligonucleotides of SEQ ID No. 319 to SEQ ID No. 338 (Figure 19) corresponding to the coding strand are
35 phosphorylated according to the following procedure:
1 μ l of each one of the oligonucleotides, taken up in water at 250 μ M, are mixed in a microtube containing 50 μ l of water, so as to bring the final solution to a concentration of 100 picomol per microliter. 5 μ l of

this solution are then mixed with 10 μ l of 10-times concentrated polynucleotide kinase buffer, 0.4 μ l of a 50 mM ATP solution, 85 μ l of water and 1 μ l of the enzyme (at 10 μ /l), and the entire mixture is
5 incubated for 4 hours at 37°C and then 5 minutes at 95°C (solution A).

A solution of the oligonucleotides of the noncoding strand is made up by mixing 1 μ l of each
10 oligonucleotide (SEQ ID No. 339 to SEQ ID No. 360; Figure 19) and 1 μ l of the oligonucleotide SEQ ID No. 318 (Figure 19), to which solution 160 μ l of water are added in order to obtain a final solution at 54 picomol per μ l (solution B).

15 The assembly of the gene is carried out first by mixing 10 μ l of solution A, 1 μ l of solution B, 6 μ l of a 100 mM KCl solution, 3 μ l of a 0.5% solution of the surfactant NP-40, 4 μ l of a 50 mM MgCl₂ solution, 3 μ l
20 of a 10 mM ATP solution and 7.5 μ l of Pfu ligase (30 units), and the mixture is then heated in a programmable thermocycler for 3 minutes at 95°C and then 3 minutes at 80°C, before undergoing 3 cycles of one minute at 95°C, followed by a change from 95°C to
25 70°C in 1 minute, then a change from 70°C to 55°C in 1 hour and, finally, 2 hours at 55°C. The mixture of the assembled oligonucleotides is then amplified with the primers NO1 and NO22. The amplification product is purified on a Promega column, digested with the BspHI
30 and NheI restriction enzymes and cloned into the plasmid pMOD2LacZ(wt) linearized with BspHI and NheI. The sequences of the plasmid DNA of 2 kanamycin-resistant clones, which appeared after transformation of the *E. coli* strain GT100 (available from Invivogen)
35 with the mixture of the ligation between the vector fragment and the PCR fragment, were found to be in accordance with the sequence given in Figure 18. This synthetic gene, placed under the control of the bacterial EM7 promoter (vector pMOD2Neo Δ CpG), confers

a resistance to kanamycin identical to that provided by the same vector containing the native *neo* gene with the recipient *E. coli* strain GT100. The BspHI-NheI *neo* fragment of the plasmid pMOD2Neo Δ CpG was then
5 introduced into the plasmid pSh Δ CpG of Figure 16, linearized with NcoI-NheI, to give, after ligation and transformation in *E. coli*, the plasmid pNeo Δ CpG.

THE CLAIMS DEFINING THE INVENTION ARE AS FOLLOWS:

1. An origin of replication for a plasmid, wherein its sequence corresponds to that of the R6K gamma origin of replication in which each G of the CpGs of the repeat region of the core has been replaced with an A, a C or a T, or each C of the CpGs has been replaced with a G, an A or a T.
2. An origin of replication as claimed in claim 1, wherein its sequence comprises the sequence SEQ ID NO: 12 or the sequence SEQ ID NO: 13.
3. An origin of replication as claimed in either claim 1 or 2, wherein the *pi* protein-binding sequence is repeated 5 or 6 times.
4. A promoter, whose sequence comprises the sequence SEQ ID NO: 11.
5. A plasmid comprising an origin of replication as claimed in any one of claims 1 to 3.
6. The plasmid as claimed in claim 5, being completely devoid of CpG.
7. A plasmid of SEQ ID NO: 14.
8. A method for producing a plasmid completely devoid of CpG and free of methylation on cytosine in the nucleic acid context CC(A/T)GG, wherein a plasmid as claimed in claim 5 is produced by replication in an *Escherichia coli* strain expressing the *pi* protein, which is deficient for the *dcm* methylation system.
9. An *Escherichia coli* cell transformed with a plasmid as claimed in claim 5.

10. A transformed *Escherichia coli* cell as claimed in claim 9, which expresses a gene encoding a *pi* protein.

5 11. A transformed *Escherichia coli* cell as claimed in either claim 9 or 10, which further includes an inactivated *dcm* gene.

10 12. A kit for producing plasmids, comprising at least one cell as claimed in any one of claims 9 to 11.

NcoI
 1 CCATGGCCAAGTTGACCAAGTGCTGTCCCAGTGCTCACAGCCAGGGATGTGGCTGGAGCT
 1 MetAlaLysLeuThrSerAlaValProValLeuThrAlaArgAspValAlaGlyAla
 60 GTTGAGTTCTGGACTGACAGGTTGGGGTTCTCCAGAGATTTTGTGGAGGATGACTTTGCA
 20 ValGluPheTrpThrAspArgLeuGlyPheSerArgAspPheValGluAspAspPheAla
 120 GGTGTGGTCAGAGATGATGTCACCCTGTTTCATCTCAGCAGTCCAGGACCAGGTGGTGCCT
 40 GlyValValArgAspAspValThrLeuPheIleSerAlaValGlnAspGlnValValPro
 180 GACAACACCCTGGCTTGGGTGTGGGTGAGAGGACTGGATGAGCTGTATGCTGAGTGGAGT
 60 AspAsnThrLeuAlaTrpValTrpValArgGlyLeuAspGluLeuTyrAlaGluTrpSer
 240 GAGGTGGTCTCCACCAACTTCAGGGATGCCAGTGGCCCTGCCATGACAGAGATTGGAGAG
 80 GluValValSerThrAsnPheArgAspAlaSerGlyProAlaMetThrGluIleGlyGlu
 300 CAGCCCTGGGGGAGAGAGTTTGCCCTGAGAGACCCAGCAGGCAACTGTGTGCACTTTGTG
 100 GlnProTrpGlyArgGluPheAlaLeuArgAspProAlaGlyAsnCysValHisPheVal
 NheI
 360 GCAGAGGAGCAGGACTGAGGATAAGAATTCAGCTAGC SEQ ID No.1
 120 AlaGluGluGlnAsp*** SEQ ID No.2

Figure 1

OL26099:AACTCAGCTGAGGAGGCAGACCATGGCCAAGTTGACCACT	SEQ ID No.16
OL26199:GCTGTCCCACTGCTCACAGCCAGGGATGTGGCTGGAGCTG	SEQ ID No.17
OL26299:TTGAGTTCTGGACTGACAGGTTGGGGTTCTCCAGAGATTT	SEQ ID No.18
OL26399:TGTGGAGGATGACTTTGCAGGTGTGGTCAGAGATGATGTC	SEQ ID No.19
OL26499:ACCCGTTCATCTCAGCAGTCCAGGACCAGGTGGTGCCTG	SEQ ID No.20
OL26599:ACAACACCCTGGCTTGGGTGTGGGTGAGAGGACTGGATGA	SEQ ID No.21
OL26699:GCTGTATGCTGAGTGGAGTGAGGTGGTCTCCACCACTTC	SEQ ID No.22
OL26799:AGGGATGCCAGTGGCCCTGCCATGACAGAGATTGGAGAGC	SEQ ID No.23
OL26899:AGCCCTGGGGGAGAGAGTTTGGCCCTGAGAGACCCAGCAGG	SEQ ID No.24
OL26999:CAACTGTGTGCACTTTGTGGCAGAGGAGCAGGACTGAGGA	SEQ ID No.25
OL27099:TAAGAATTCAGCTAGCTCGAC	SEQ ID No.26
OL27199:GTCGAGCTAGCTGAATTCTTATCCTCAGTCCTGCTCCTCTG	SEQ ID No.27
OL27299:CCACAAAGTGCACACAGTTGCCTGCTGGGTCTCTCAGGGC	SEQ ID No.28
OL27399:AAACTCTCTCCCCCAGGGCTGCTCTCCAATCTCTGTCATG	SEQ ID No.29
OL27499:GCAGGGCCACTGGCATCCCTGAAGTTGGTGGAGACCACCT	SEQ ID No.30
OL27599:CACTCCACTCAGCATACAGCTCATCCAGTCCTCTCACCCA	SEQ ID No.31
OL27699:CACCCAAGCCAGGGTGTGTCAGGCACCACCTGGTCCTGG	SEQ ID No.32
OL27799:ACTGCTGAGATGAACAGGGTGACATCATCTCTGACCACAC	SEQ ID No.33
OL27899:CTGCAAAGTCATCCTCCACAAATCTCTGGAGAACCCCAA	SEQ ID No.34
OL27999:CCTGTCACTCCAGAACTCAACAGCTCCAGCCACATCCCTG	SEQ ID No.35
OL28099:GCTGTGAGCACTGGGACAGCACTGGTCAACTTGGCCATGG	SEQ ID No.36
OL28199:TCTGCCTCCTCAGCTGAGTT	SEQ ID No.37

Figure 2

BspHI

1 TCATGAAGAAACCTGAACTGACAGCAACTTCTGTTGAGAAGTTTCTCATTGAAAAATTT
 1 MetLysLysProGluLeuThrAlaThrSerValGluLysPheLeuIleGluLysPhe

60 GATTCTGTTTCTGATCTCATGCAGCTGTCTGAAGGTGAAGAAAGCAGAGCCTTTTCTTTT
 20 AspSerValSerAspLeuMetGlnLeuSerGluGlyGluGluSerArgAlaPheSerPhe

120 GATGTTGGAGGAAGAGGTTATGTTCTGAGGGTCAATTCTTGTGCTGATGGTTTTTACAAA
 40 AspValGlyGlyArgGlyTyrValLeuArgValAsnSerCysAlaAspGlyPheTyrLys

180 GACAGATATGTTTACAGACACTTTGCCTCTGCTGCTCTGCCAATTCAGAAGTTCTGGAC
 60 AspArgTyrValTyrArgHisPheAlaSerAlaAlaLeuProIleProGluValLeuAsp

240 ATTGGAGAAATTTTCTGAATCTCTCACCTACTGCATCAGCAGAAGAGCACAAGGAGTCACT
 80 IleGlyGluPheSerGluSerLeuThrTyrCysIleSerArgArgAlaGlnGlyValThr

300 CTCCAGGATCTCCCTGAAACTGAGCTGCCAGCTGTTCTGCAACCTGTTGCTGAAGCAATG
 100 LeuGlnAspLeuProGluThrGluLeuProAlaValLeuGlnProValAlaGluAlaMet

360 GATGCCATTGCAGCAGCTGATCTGAGCCAAACCTCTGGATTGGTCTTTTGGTCCCAA
 120 AspAlaIleAlaAlaAlaAspLeuSerGlnThrSerGlyPheGlyProPheGlyProGln

420 GGCATTGGTCAGTACACCCTTGGAGGGATTTCATTGTGCCATTGCTGATCCTCATGTC
 140 GlyIleGlyGlnTyrThrTrpArgAspPheIleCysAlaIleAlaAspProHisVal

480 TATCACTGGCAGACTGTGATGGATGACACAGTTTCTGCTTCTGTTGCTCAGGCCTGGAT
 160 TyrHisTrpGlnThrValMetAspAspThrValSerAlaSerValAlaGlnAlaLeuAsp

540 GAACTCATGCTGTGGGCAGAAGATTGTCTGAAGTCAGACACCTGGTCCATGCTGATTTT
 180 GluLeuMetLeuTrpAlaGluAspCysProGluValArgHisLeuValHisAlaAspPhe

600 GGAAGCAACAATGTTCTGACAGACAATGGCAGAATCACTGCAGTCATTGACTGGTCTGAA
 200 GlySerAsnAsnValLeuThrAspAsnGlyArgIleThrAlaValIleAspTrpSerGlu

660 GCCATGTTTGGAGATTCTCAATATGAGGTTGCCAACATTTTTTTTGGAGACCTTGGCTG
 220 AlaMetPheGlyAspSerGlnTyrGluValAlaAsnIlePhePheTrpArgProTrpLeu

720 GCTTGCAATGGAACAACAAGATATTTTGAAGAAGACACCCAGAACTGGCTGGTTCC
 240 AlaCysMetGluGlnGlnThrArgTyrPheGluArgArgHisProGluLeuAlaGlySer

780 CCCAGACTGAGAGCCTACATGCTCAGAATTGGCCTGGACCAACTGTATCAATCTCTGGTT
 260 ProArgLeuArgAlaTyrMetLeuArgIleGlyLeuAspGlnLeuTyrGlnSerLeuVal

840 GATGGAAACTTTGATGATGCTGCTTGGGCACAAGGAAGATGTGATGCCATTGTGAGGTCT
 280 AspGlyAsnPheAspAspAlaAlaTrpAlaGlnGlyArgCysAspAlaIleValArgSer

900 GGTGCTGGAACTGTTGGAAGAACTCAAATTGCAAGAAGGTCTGCTGCTGTTTGGACTGAT
 300 GlyAlaGlyThrValGlyArgThrGlnIleAlaArgArgSerAlaAlaValTrpThrAsp

960 GGATGTGTTGAAGTTCTGGCTGACTCTGGAAACAGGAGACCTCCACAAGACCCAGAGCC
 320 GlyCysValGluValLeuAlaAspSerGlyAsnArgArgProSerThrArgProArgAla

NheI

1020 AAGGAATGAATATTAGCTAGC
 340 LysGlu***

SEQ. ID. No. 3
 SEQ ID No 4

Figure 3

HS01: TGAGATCACC GGTT CAGCTGAGGAGGCACATCATGAAGAAACCTGAACTGACAGCAACTT SEQ ID No. 38
 HS02: CTGTTGAGAGTTTCTCATTGAAAAATTTGATTCTGTTTCTGATCTCATGCAGCTGTCTG SEQ ID No. 39
 HS03: AAGGTGAAGAAAGCAGAGCCTTTTCTTTTGATGTTGGAGGAAGAGGTTATGTTCTGAGGG SEQ ID No. 40
 HS04: TCAATTCTTGCTGATGGTTTTTACAAAGACAGATATGTTTACAGACACTTTGCCCTCTG SEQ ID No. 41
 HS05: CTGCTCTGCCAATTCAGAACTTCTGGACATGGAGAAATTTCTGAATCTCTCACCTACT SEQ ID No. 42
 HS06: GCATCAGCAGAAGAGCACAAGGAGTCACTCTCCAGSATCTCCCTGAACTGAGCTGCCAG SEQ ID No. 43
 HS07: CTGTTCTGCAACCTGTGCTGAAGCAATGGATGCCATTGCAGCAGCTGATCTGAGCCAA SEQ ID No. 44
 HS08: CCTCTGGATTGGTCTTTTGGTCCCCAAGGCATTGGTCAGTACACCACTTGGAGGGATT SEQ ID No. 45
 HS09: TCATTGTGCCATTGCTGATCCTCATGTCTATCACTGGCAGACTGTGATGGATGACACAG SEQ ID No. 46
 HS10: TTTCTGCTTCTGTTGCTCAGGCACTGGATGAACTCATGCTGTGGGCAGAAGATTGTCTCTG SEQ ID No. 47
 HS11: AAGTCAGACACCTGGTCCATGCTGATTTTGGAAAGCAACAATGTTCTGACAGACAATGGCA SEQ ID No. 48
 HS12: GAATCACTGCAGTCAATTGACTGGTCTGAAGCCATGTTTGGAGATTCTCAATATGAGGTTG SEQ ID No. 49
 HS13: CCAACATTTTTTTTGGAGACCTTGGCTGGCTTGCATGGAAACAACAAGATATTTTG SEQ ID No. 50
 HS14: AAAGAAGACACCCAGAACTGGCTGGTTCCTCCAGACTGAGAGCCTACATGCTCAGAAATG SEQ ID No. 51
 HS15: GCCTGGACCAACTGTATCAATCTCTGTTGATGGAAACTTTGATGATGCTGCTTGGGCAC SEQ ID No. 52
 HS16: AAGGAAGATGTGATGCCATTGTGAGGTCTGGTCTGGAACTGTTGGAAGAACTCAAAATG SEQ ID No. 53
 HS17: CAAGAAGGTCTGCTGCTGTTTGGACTGATGGATGTTGAAGTTCTGGCTGACTCTGGAA SEQ ID No. 54
 HS18: ACAGGAGACCCCTCCACAAGACCCAGAGCCAGGAATGAATATTAGCTAGCGGATCCTGAG SEQ ID No. 55
 HS19: CTCAGGATCCGCTAGCTAATATTCAATTCCT SEQ ID No. 56
 HS20: TGGCTCTGGGTCTTGTGGAGGGTCTCTGTTTCCAGAGTCAGCCAGAAGTTCAACACATC SEQ ID No. 57
 HS21: CATCAGTCCAAACAGCAGCAGACCTTCTTGCAATTTGAGTTCTTCCAACAGTTCAGCAC SEQ ID No. 58
 HS22: CAGACCTCACAAATGGCATCACATCTTCTTGTGCCCCAAGCAGCATCATCAAAGTTTCCAT SEQ ID No. 59
 HS23: CAACCAGAGATTGATACAGTTGGTCCAGGCCAATTTCTGAGCATGTAGGCTCTCAGTCTGG SEQ ID No. 60
 HS24: GGGAAACCAAGCTTCTGGGTGTCTTCTTTCAAATATCTTGTGTTGTTGTTCCATGCAAG SEQ ID No. 61
 HS25: CCAGCCAAGGTCTCCAAAAAATGTTGGCAACCTCATATTGAGAATCTCCAAACATGG SEQ ID No. 62
 HS26: CTTCAAGACAGTCAATGACTGCAGTGATTCTGCCATTGTCTGTCAAGAACATTGTTGCTTC SEQ ID No. 63
 HS27: CAAAATCAGCATGGACCAAGTGTCTGACTTCAGGACAATCTTCTGCCACAGCATGAGTT SEQ ID No. 64
 HS28: CATCCAGTGCCTGAGCAACAGAAAGCAGAACTGTGTATCCATCACAGTCTGCCAGTGAT SEQ ID No. 65
 HS29: AGACATGAGGATCAGCAATGGCACAATGAAATCCCTCCAAGTGGTACTGACCAATGC SEQ ID No. 66
 HS30: CTTGGGGACCAAAAGGACCAAAATCCAGAGGTTTGGCTCAGATCAGCTGCTGCAATGGCAT SEQ ID No. 67
 HS31: CCATTGCTTCAGCAACAGGTTGCAGAACAGCTGGCAGCTCAGTTTCAGGGAGATCCTGGA SEQ ID No. 68
 HS32: GAGTGACTCTTGTGCTCTTCTGCTGATGCAGTAGGTGAGAGATTCAAGAAATCTCCAA SEQ ID No. 69
 HS33: TGTCCAGAACTTCTGGAATTGGCAGAGCAGCAGAGGCAAGTGTCTGTAAACATATCTGT SEQ ID No. 70
 HS34: CTTTGTAAAAACATCAGCAACAAGATTGACCCTCAGAACATAACCTCTTCTCCAAACAT SEQ ID No. 71
 HS35: CAAAAGAAAAGGCTCTGCTTTCTTCACTTCAGACAGCTGCATGAGATCAGAAACAGAA SEQ ID No. 72
 HS36: CAAATTTTCAATGAGAACTTCTCAACAGAAAGTTGCTGTCAAGTTTCAAGTTTCTTATGA SEQ ID No. 73
 HS37: TGTGCTCTCTCAGCTGAACCGGTGATCTCA SEQ ID No. 74

Figure 4

```

      BspHI
1  TCATGAAGACCTTCAACATCTCTCAGCAGGATCTGGAGCTGGTGGAGGTCGCCACTGAG
1  MetLysThrPheAsnIleSerGlnGlnAspLeuGluLeuValGluValAlaThrGlu

60 AAGATCACCATGCTCTATGAGGACAACAAGCACCATGTCGGGGCGGCCATCAGGACCAAG
20 LysIleThrMetLeuTyrGluAspAsnLysHisHisValGlyAlaAlaIleArgThrLys

120 ACTGGGGAGATCATCTCTGCTGTCCACATTGAGGCCTACATTGGCAGGGTCACTGTCTGT
40 ThrGlyGluIleIleSerAlaValHisIleGluAlaTyrIleGlyArgValThrValCys

180 GCTGAAGCCATTGCCATTGGGTCTGCTGTGAGCAACGGGCAGAAGGACTTTGACACCATT
60 AlaGluAlaIleAlaIleGlySerAlaValSerAsnGlyGlnLysAspPheAspThrIle

240 GTGGCTGTCAGGCACCCCTACTCTGATGAGGTGGACAGATCCATCAGGGTGGTCAGCCCC
80 ValAlaValArgHisProTyrSerAspGluValAspArgSerIleArgValValSerPro

300 TGTGGCATGTGCAGAGAGCTCATCTCTGACTATGCTCCTGACTGCTTTGTGCTCATTGAG
100 CysGlyMetCysArgGluLeuIleSerAspTyrAlaProAspCysPheValLeuIleGlu

360 ATGAATGGCAAGCTGGTCAAAACCACCATTGAGGAACTCATCCCCCTCAAGTACACCAGG
120 MetAsnGlyLysLeuValLysThrThrIleGluGluLeuIleProLeuLysTyrThrArg
      NheI
420 AACTAAACCTGAATTCAGCTAGC          SEQ ID No 5
140 Asn***                          SEQ ID No 6

```

Figure 5

OL64: AGGAGGCACATCATGAAGACCTTCAACATCTCTCAGCAGG	SEQ ID No.75
OL65: ATCTGGAGCTGGTGGAGGTCGCCACTGAGAAGATCACCAT	SEQ ID No.76
OL66: GCTCTATGAGGACAACAAGCACCATGTCGGGGCGGCCATC	SEQ ID No.77
OL67: AGGACCAAGACTGGGGAGATCATCTCTGCTGTCCACATTG	SEQ ID No.78
OL68: AGGCCCTACATTGGCAGGGTCACTGTCTGTGCTGAAGCCAT	SEQ ID No.79
OL69: TGCCATTGGGTCTGCTGTGAGCAACGGGCAGAAAGGACTTT	SEQ ID No.80
OL70: GACACCATTGTGGCTGTGAGGCACCCCTACTCTGATGAGG	SEQ ID No.81
OL71: TGGACAGATCCATCAGGGTGGTCAGCCCCTGTGGCATGTG	SEQ ID No.82
OL72: CAGAGAGCTCATCTCTGACTATGCTCCTGACTGCTTTGTG	SEQ ID No.83
OL73: CTCATTGAGATGAATGGCAAGCTGGTCAAACCACCATTG	SEQ ID No.84
OL74: AGGAACTCATCCCCCTCAAGTACACCAGGAATAAACCTG	SEQ ID No.85
OL75: AATTGAGCTAGCTCGACATGA	SEQ ID No.86
OL76: TCATGTCGAGCTAGCTGAATTGAGTTTAGTTCTTGGTGTA	SEQ ID No.87
OL77: CTTGAGGGGGATGAGTTCTCAATGGTGGTTTGACCAGC	SEQ ID No.88
OL78: TTGCCATTCTCAATGAGCACAAAGCAGTCAGGAGCAT	SEQ ID No.89
OL79: AGTCAGAGATGAGCTCTCTGCACATGCCACAGGGGCTGAC	SEQ ID No.90
OL80: CACCCTGATGGATCTGTCCACCTCATCAGAGTAGGGGTGC	SEQ ID No.91
OL81: CTGACAGCCACAATGGTGTCAAAGTCCTTCTGCCCGTTGC	SEQ ID No.92
OL82: TCACAGCAGACCCAATGGCAATGGCTTCAGCACAGACAGT	SEQ ID No.93
OL83: GACCCTGCCAATGTAGGCCTCAATGTGGACAGCAGAGATG	SEQ ID No.94
OL84: ATCTCCCCAGTCTTGGTCCTGATGGCCGCCCGACATGGT	SEQ ID No.95
OL85: GCTTGTGTCTCTCATAGAGCATGGTGATCTTCTCAGTGCC	SEQ ID No.96
OL86: GACCTCCACCAGCTCCAGATCCTGCTGAGAGATGTTGAAG	SEQ ID No.97
OL87: GTCTTCATGATGTGCCTCT	SEQ ID No.98

Figure 6

BspHI
 1 TCATGACTGAGTACAAACCCACAGTGAGGCTGGCAACCAGAGATGATGTTCCAAGAGCT
 1 MetThrGluTyrLysProThrValArgLeuAlaThrArgAspAspValProArgAla
 60 GTGAGAACACTGGCTGCTGCTTTTGCAGACTACCCTGCAACAAGGCACACAGTTGACCCT
 20 ValArgThrLeuAlaAlaAlaPheAlaAspTyrProAlaThrArgHisThrValAspPro
 120 GACAGGCACATTGAGAGGGTGACAGAACTGCAAGAACTCTCCTCACCAGAGTGGGACTG
 40 AspArgHisIleGluArgValThrGluLeuGlnGluLeuPheLeuThrArgValGlyLeu
 180 GACATTGGAAAAGTTTGGGTTGCAGATGATGGAGCTGCTGTTGCAGTTTGGACAACACCT
 60 AspileGlyLysValTrpValAlaAspAspGlyAlaAlaValAlaValTrpThrThrPro
 240 GAGTCTGTTGAAGCTGGTGCTGTTTTTGCTGAAATTGGACCAAGAAATGGCTGAGCTCTCT
 80 GluSerValGluAlaGlyAlaValPheAlaGluIleGlyProArgMetAlaGluLeuSer
 300 GGAAGCAGGCTGGCAGCACAACAACAAATGGAAGGTCTGCTGGCACCACACAGGCCAAAA
 100 GlySerArgLeuAlaAlaGlnGlnGlnMetGluGlyLeuLeuAlaProHisArgProLys
 360 GAGCCAGCTTGGTTTCTGGCAACTGTTGGAGTGAGCCCTGACCACCAGGAAAGGGTCTG
 120 GluProAlaTrpPheLeuAlaThrValGlyValSerProAspHisGlnGlyLysGlyLeu
 420 GGATCTGCTGTTGTTCTGCTGGAGTTGAAGCTGCTGAAAGGGCTGGAGTTCCTGCCTTT
 140 GlySerAlaValValLeuProGlyValGluAlaAlaGluArgAlaGlyValProAlaPhe
 480 CTGGAAACTTCTGCTCCAGAACTGCCTTTTATGAAAGACTGGGATTCCTGTGACA
 160 LeuGluThrSerAlaProArgAsnLeuProPheTyrGluArgLeuGlyPheThrValThr
 540 GCTGATGTTGAGGTTCCAGAAGGCCCAAGAACTTGGTGCAAGGAAGCCTGGAGCT
 180 AlaAspValGluValProGluGlyProArgThrTrpCysMetThrArgLysProGlyAla
 NheI
 600 TAAACCTGAGCTAGC SEQ ID No 7
 200 *** SEQ ID No 8

Figure 7

```

pur1:CTCACTATAGGAGGACCATCATGACTGAGTACAAACCCACAGTGAGGCTGGCAACCAGAG SEQ ID No.99
pur2:ATGATGTTCCAAGAGCTGTGAGAACACTGGCTGCTGCTTTGTCAGACTACCTTGCAACAA SEQ ID No.100
pur3:GGCACACAGTTGACCCCTGACAGGCACATTGAGAGGGTGACAGAACTGCAAGAACTCTTCC SEQ ID No.101
pur4:TCACCAGAGTGGGACTGGACATTGGAAAAGTTTGGGTTGCAGATGATGGAGCTGCTGTTG SEQ ID No.102
pur5:CAGTTTGGACAACACCTGAGTCTGTTGAAGCTGCTGCTGTTTTGCTGAAATTGGACCAA SEQ ID No.103
pur6:GAATGGCTGAGCTCTCTGGAAGCAGGCTGGCAGCACAAACAATGGAAGGTCTGCTGG SEQ ID No.104
pur7:CACCACACAGGCCAAAGAGCCAGCTTGGTTTCTGGCAACTGTTGGAGTGAGCCCTGACC SEQ ID No.105
pur8:ACCAGGAAAGGGTCTGGGATCTGCTGTTGTTCTGCTGGAGTTGAAGCTGCTGAAAGGG SEQ ID No.106
pur9:CTGGAGTTCTGCTCTTCTGGAAACTTCTGCTCCAGAAACCTGCTTTTATGAAAGAC SEQ ID No.107
pur10:TGGGATTCACTGTGACAGCTGATGTTGAGGTTCCAGAGGCCCAAGAACTTGGTGCAATGA SEQ ID No.108
pur11:CAAGGAAGCCTGGAGCTTAAACCTGAGCTAGCTCGACATGATAAGATACATTGATGAGTT SEQ ID No.109
pur12:AACTCATCAATGTATCTTATCATGTCGAGC SEQ ID No.110
pur13:TAGCTCAGGTTTAAGCTCCAGGCTTCTTGTCTATGCACCAAGTTCTTGGGCTTCTGGAA SEQ ID No.111
pur14:CCTCAACATCAGCTGTCAAGTGAATCCAGTCTTTCATAAAAAGGCAGGTTTCTGGGAG SEQ ID No.112
pur15:CAGAAGTTTCCAGAAAGGCAGGAATCCAGCCCTTTCAGCAGCTTCAACTCCAGGCAGAA SEQ ID No.113
pur16:CAACAGCAGATCCCAGACCTTTCCTGCTGCTCAGGCTCACTCCAAAGCTTCCAGAA SEQ ID No.114
pur17:ACCAAGCTGGCTCTTTTGGCTGTGTGGTGCCAGCAGACCTTCCATTGTTGTGTGCTG SEQ ID No.115
pur18:CCAGCTGCTTCCAGAGAGCTCAGCCATTCTTGGTCCAAATTCAGCAAAAACAGCACCAG SEQ ID No.116
pur19:CTTCAACAGACTCAGGTGTTGTCCAACTGCAACAGCAGCTCCATCATCTGCAACCCAAA SEQ ID No.117
pur20:CTTTTCCAATGTCCAGTCCCACTCTGCTGAGGAAGAGTTCTTGCAAGTTCTGTCAACCTCT SEQ ID No.118
pur21:CAATGTGCTGTGACGGTCAACTGTGTGCTTGTGTCAGGGTAGTCTGCAAAAGCAGCAG SEQ ID No.119
pur22:CCAGTGTCTCAAGCTCTTGGAAACATCATCTCTGCTTGGCAGCCTCACTGTGGGTTTGT SEQ ID No.120
pur23:ACTCAGTCATGATGGTCTCTCTATAGTGAG SEQ ID No.121

```

Figure 8

9/27

NcoI

1 CCATGGACCCCTGTTGTGCTGCAAAGGAGAGACTGGGAGAACCCTGGAGTGACCCAGCTC SEQ ID No 9
1 MetAspProValValLeuGlnArgArgAspTrpGluAsnProGlyValThrGlnLeu SEQ ID No 10

60 AACAGACTGGCTGCCCCACCCTCCCTTTGCCTCTTGGAGGAACCTCTGAGGAAGCCAGGACA
20 AsnArgLeuAlaAlaHisProProPheAlaSerTrpArgAsnSerGluGluAlaArgThr

120 GACAGGCCCCAGCCAGCAGCTCAGGTCTCTCAATGGAGAGTGGAGGTTTGCCTGGTTCCCT
40 AspArgProSerGlnGlnLeuArgSerLeuAsnGlyGluTrpArgPheAlaTrpPhePro

180 GCCCCTGAAGCTGTGCCTGAGTCTTGGCTGGAGTGTGACCTCCAGAGGCTGACACTGTT
60 AlaProGluAlaValProGluSerTrpLeuGluCysAspLeuProGluAlaAspThrVal

240 GTGGTGGCCAGCAACTGGCAGATGCATGGCTATGATGCCCCCATCTACACCAATGTCACC
80 ValValProSerAsnTrpGlnMetHisGlyTyrAspAlaProIleTyrThrAsnValThr

300 TACCCCATCACTGTGAACCCCTTTTGTGCCCCACTGAGAACCCCACTGGCTGCTACAGC
100 TyrProIleThrValAsnProProPheValProThrGluAsnProThrGlyCysTyrSer

360 CTGACCTTCAATGTTGATGAGAGCTGGCTGCAAGAAGGCCAGACCAGGATCATCTTTGAT
120 LeuThrPheAsnValAspGluSerTrpLeuGlnGluGlyGlnThrArgIleIlePheAsp

420 GGAGTCAACTCTGCCTTCCACCTCTGGTGCATGGCAGGTGGGTTGGCTATGGCCAAGAC
140 GlyValAsnSerAlaPheHisLeuTrpCysAsnGlyArgTrpValGlyTyrGlyGlnAsp

480 AGCAGGCTGCCCTCTGAGTTTGACCTCTCTGCCTTCTCAGAGCTGGAGAGAACAGGCTG
160 SerArgLeuProSerGluPheAspLeuSerAlaPheLeuArgAlaGlyGluAsnArgLeu

540 GCTGTCTAGGTGCTCAGGTGGTCTGATGGCAGCTACCTGGAAGACCAAGACATGTGGAGG
180 AlaValMetValLeuArgTrpSerAspGlySerTyrLeuGluAspGlnAspMetTrpArg

600 ATGTCTGGCATCTTCAGGGATGTGAGCCTGTGTCACAAGCCACCACCCAGATTCTTGAC
200 MetSerGlyIlePheArgAspValSerLeuLeuHisLysProThrThrGlnIleSerAsp

660 TTCCATGTTGCCACCAGGTTCAATGATGACTTCAGCAGAGCTGTGCTGGAGGCTGAGGTG
220 PheHisValAlaThrArgPheAsnAspAspPheSerArgAlaValLeuGluAlaGluVal

720 CAGATGTGTGGAGAACTCAGAGACTACCTGAGAGTCACAGTGAGCCTCTGGCAAGGTGAG
240 GlnMetCysGlyGluLeuArgAspTyrLeuArgValThrValSerLeuTrpGlnGlyGlu

780 ACCCAGGTGGCCTCTGGCACAGCCCCCTTTGGAGGAGAGATCATTGATGAGAGAGGAGGC
260 ThrGlnValAlaSerGlyThrAlaProPheGlyGlyGluIleIleAspGluArgGlyGly

840 TATGCTGACAGAGTCACCTGAGGCTCAATGTGGAGAACCCCAAGCTGTGGTCTGCTGAG
280 TyrAlaAspArgValThrLeuArgLeuAsnValGluAsnProLysLeuTrpSerAlaGlu

900 ATCCCCAACCTCTACAGGGCTGTTGTGGAGCTGCACACTGCTGATGGCACCTGATTGAA
300 IleProAsnLeuTyrArgAlaValValGluLeuHisThrAlaAspGlyThrLeuIleGlu

960 GCTGAAGCCTGTGATGTTGGATTGAGAGAAGTCAGGATTGAGAAATGGCCTGCTGCTGCTC
320 AlaGluAlaCysAspValGlyPheArgGluValArgIleGluAsnGlyLeuLeuLeuLeu

1020 AATGGCAAGCCTCTGCTCATCAGGGGAGTCAACAGGCATGAGCACCCCTCTGCTGGA
340 AsnGlyLysProLeuLeuIleArgGlyValAsnArgHisGluHisHisProLeuHisGly

EcoRV

1080 CAAGTGATGGATGAACAGACAATGGTGCAGATATCTGCTAATGAAGCAGAACAACTTC

10/27

360 GlnValMetAspGluGlnThrMetValGlnAspIleLeuLeuMetLysGlnAsnAsnPhe
1140 AATGCTGTCAGGTGCTCTCACTACCCCAACCACCCTCTCTGGTACACCCTGTGTGACAGG
380 AsnAlaValArgCysSerHisTyrProAsnHisProLeuTrpTyrThrLeuCysAspArg
1200 TATGGCCTGTATGTTGTTGATGAAGCCAACATTGAGACACATGGCATGGTGGCCCATGAAC
400 TyrGlyLeuTyrValValAspGluAlaAsnIleGluThrHisGlyMetValProMetAsn
1260 AGGCTCACAGATGACCCAGGTGGCTGCCATGTCTGAGAGAGTGACCAGGATGGTG
420 ArgLeuThrAspAspProArgTrpLeuProAlaMetSerGluArgValThrArgMetVal

1320 CAGAGAGACAGGAACCACCCCTCTGTGATCATCTGGTCTCTGGGCAATGAGTCTGGACAT
440 GlnArgAspArgAsnHisProSerValIleIleTrpSerLeuGlyAsnGluSerGlyHis
1380 GGAGCCAACCATGATGCTCTCTACAGGTGGATCAAGTCTGTTGACCCAGCAGACCTGTG
460 GlyAlaAsnHisAspAlaLeuTyrArgTrpIleLysSerValAspProSerArgProVal
1440 CAGTATGAAGGAGGTGGAGCAGACACCACAGCCACAGACATCATCTGCCCCATGTATGCC
480 GlnTyrGluGlyGlyAlaAspThrThrAlaThrAspIleIleCysProMetTyrAla
1500 AGGGTTGATGAGGACCAGCCCTTCCCTGCTGTGCCCAAGTGGAGCATCAAGAAGTGGCTC
500 ArgValAspGluAspGlnProPheProAlaValProLysTrpSerIleLysLysTrpLeu
1560 TCTCTGCCTGGAGAGACCAGACCTCTGATCCTGTGTGAATATGCATGCAATGGGCAAC
520 SerLeuProGlyGluThrArgProLeuIleLeuCysGluTyrAlaHisAlaMetGlyAsn
1620 TCTCTGGGAGGCTTTGCCAAGTACTGGCAAGCCTTCAGACAGTACCCAGGCTGCAAGGA
540 SerLeuGlyGlyPheAlaLysTyrTrpGlnAlaPheArgGlnTyrProArgLeuGlnGly
1680 GGATTTGTGTGGGACTGGGTGGACCAATCTCTCATCAAGTATGATGAGAATGGCAACCCC
560 GlyPheValTrpAspTrpValAspGlnSerLeuIleLysTyrAspGluAsnGlyAsnPro
1740 TGGTCTGCCTATGGAGGAGACTTTGGTGACACCCCCAATGACAGGCAGTTCTGCATGAAT
580 TrpSerAlaTyrGlyGlyAspPheGlyAspThrProAsnAspArgGlnPheCysMetAsn
1800 GGCCTGGTCTTTGCAGACAGGACCCCTCACCCCTGCCCTCACAGAGGCCAAGCACCCAGAA
600 GlyLeuValPheAlaAspArgThrProHisProAlaLeuThrGluAlaLysHisGlnGln
1860 CAGTTCTTCCAGTTCAGGCTGTCTGGACAGACCATTGAGGTGACATCTGAGTACCTCTTC
620 GlnPhePheGlnPheArgLeuSerGlyGlnThrIleGluValThrSerGluTyrLeuPhe

SacI
1920 AGGCACTCTGACAATGAGCTCCTGCACTGGATGGTGGCCCTGGATGGCAAGCCTCTGGCT
640 ArgHisSerAspAsnGluLeuLeuHisTrpMetValAlaLeuAspGlyLysProLeuAla
1980 TCTGGTGAGGTGCCTCTGGATGTGGCCCTCAAGGAAAGCAGCTGATTGAACTGCCTGAG
660 SerGlyGluValProLeuAspValAlaProGlnGlyLysGlnLeuIleGluLeuProGlu
2040 CTGCCTCAGCCAGAGTCTGCTGGACAACCTGTGGCTAACAGTGAGGGTGGTTTCAGCCCAAT
680 LeuProGlnProGluSerAlaGlyGlnLeuTrpLeuThrValArgValValGlnProAsn
2100 GCAACAGCTTGGTCTGAGGCAGGCCACATCTCTGCATGGCAGCAGTGGAGGCTGGCTGAG
700 AlaThrAlaTrpSerGluAlaGlyHisIleSerAlaTrpGlnGlnTrpArgLeuAlaGlu
2160 AACCTCTCTGTGACCTGCCTGCTGCCTCTCATGCCATCCCTCACCTGACAACATCTGAA
720 AsnLeuSerValThrLeuProAlaAlaSerHisAlaIleProHisLeuThrThrSerGlu
2220 ATGGACTTCTGCATTGAGCTGGGCAACAAGAGATGGCAGTTCAACAGGCAGTCTGGCTTC
740 MetAspPheCysIleGluLeuGlyAsnLysArgTrpGlnPheAsnArgGlnSerGlyPhe

2280 CTGTCTCAGATGTGGATTGGAGACAAGAAGCAGCTCCTCACCCTCTCAGGGACCAATTC
 760 LeuSerGlnMetTrpIleGlyAspLysLysGlnLeuLeuThrProLeuArgAspGlnPhe
 2340 ACCAGGGCTCCTCTGGACAATGACATTGGAGTGTCTGAGGCCACCAGGATTGACCCAAAT
 780 ThrArgAlaProLeuAspAsnAspIleGlyValSerGluAlaThrArgIleAspProAsn
 2400 GCTTGGGTGGAGAGGTGGAAGGCTGCTGGACACTACCAGGCTGAGGCTGCCCTGCTCCAG
 800 AlaTrpValGluArgTrpLysAlaAlaGlyHisTyrGlnAlaGluAlaAlaLeuLeuGln
 2460 TGCACAGCAGACCCCTGGCTGATGCTGTTCTGATCACCACAGCCCATGCTTGGCAGCAC
 820 CysThrAlaAspThrLeuAlaAspAlaValLeuIleThrThrAlaHisAlaTrpGlnHis
 2520 CAAGGCAAGACCTGTTTCATCAGCAGAAAGACCTACAGGATTGATGGCTCTGGACAGATG
 840 GlnGlyLysThrLeuPheIleSerArgLysThrTyrArgIleAspGlySerGlyGlnMet
 2580 GCAATCAGTGGATGTGGAGGTGCCTCTGACACACCTCACCTGCAAGGATTGGCCTG
 860 AlaIleThrValAspValGluValAlaSerAspThrProHisProAlaArgIleGlyLeu
 2640 AACTGTCAACTGGCACAGGTGGCTGAGAGGCTGAAGTGGCTGGGCTTAGGCCCTCAGGAG
 880 AsnCysGlnLeuAlaGlnValAlaGluArgValAsnTrpLeuGlyLeuGlyProGlnGlu
 2700 AACTACCTGACAGGCTGACAGCTGCCTGCTTTCAGAGGTGGGACCTGCCTCTGTCTGAC
 900 AsnTyrProAspArgLeuThrAlaAlaCysPheAspArgTrpAspLeuProLeuSerAsp
 2760 ATGTACACCCCTTATGTGTTCCCTTCTGAGAATGGCCTGAGGTGTGGCACCAGGGAGCTG
 920 MetTyrThrProTyrValPheProSerGluAsnGlyLeuArgCysGlyThrArgGluLeu
 2820 AACTATGGTCCTCACCAGTGGAGGGGAGACTTCCAGTTCAACATCTCCAGGTACTCTCAG
 940 AsnTyrGlyProHisGlnTrpArgGlyAspPheGlnPheAsnIleSerArgTyrSerGln
 2880 CAACAGCTCATGGAAACCTCTCACAGGCACCTGCTCCATGCAGAGGAGGGAACCTGGCTG
 960 GlnGlnLeuMetGluThrSerHisArgHisLeuLeuHisAlaGluGluGlyThrTrpLeu
 2940 AACATTGATGGCTTCCACATGGGCATTGGAGGAGATGACTCTTGGTCTCCTTCTGTGTCT
 980 AsnIleAspGlyPheHisMetGlyIleGlyGlyAspAspSerTrpSerProSerValSer
 3000 GCTGAGTTCCAGTTATCTGCTGGCAGGTACCACTATCAGCTGGTGTGGTGCCAGAAGTAA
 1000 AlaGluPheGlnLeuSerAlaGlyArgTyrHisTyrGlnLeuValTrpCysGlnLys***
 NheI
 3060 ACCTGAGCTAGC

Figure 9

BGS001:	ATCACTATAGGAGGGCCACCATGGACCTGTTGTGCTGCA	SEQ ID No.122
BGS002:	GGTGGCCCTCCTATAGTGAT	SEQ ID No.123
BGS003:	AAGGAGAGACTGGGAGAACCCTGGAGTGACCCAGCTCAAC	SEQ ID No.124
BGS004:	GGTTCTCCCACTCTCTCCTTTGCAGCACACAGGGTCCAT	SEQ ID No.125
BGS005:	AGACTGGCTGCCCCACCCTCCCTTTGCCCTCTTGAGGAACT	SEQ ID No.126
BGS006:	GGAGGGTGGGCAGCCAGTCTGTTGAGCTGGGTCACTCCAG	SEQ ID No.127
BGS007:	CTGAGGAAGCCAGGACAGACAGGCCAGCCAGCAGCTCAG	SEQ ID No.128
BGS008:	GTCTGTCTGCTGCTTCTCTCAGAGTTCCCTCAAGAGGCAAAAG	SEQ ID No.129
BGS009:	GTCTCTCAATGGAGAGTGGAGGTTTGCCCTGGTTCCCTGCC	SEQ ID No.130
BGS010:	TCCACTCTCCATTGAGAGACCTGAGCTGCTGGCTGGGCCT	SEQ ID No.131
BGS011:	CCTGAAGCTGTGCCCTGAGTCTTGCTGGAGTGTGACCTCC	SEQ ID No.132
BGS012:	GACTCAGGCACAGCTTCAGGGGCAGGGAACAGGCAAAACC	SEQ ID No.133
BGS013:	CAGAGGCTGACACTGTTGTGGTGGCCAGCAACTGGCAGAT	SEQ ID No.134
BGS014:	CACAACAGTGTGAGCTCTGGGAGGTCACTCCAGCCAA	SEQ ID No.135
BGS015:	GCATGGCTATGATGCCCCCATCTACCAATGTCACTTAC	SEQ ID No.136
BGS016:	TGGGGGCATCATAGCCATGCATCTGCCAGTGTGCTGGGCAC	SEQ ID No.137
BGS017:	CCCATCACTGTGAACCCCTTTTGTGCCCACTGAGAAACC	SEQ ID No.138
BGS018:	GGGGGGTTCACAGTGTGAGGTAGGTGACATTGGTGTAGA	SEQ ID No.139
BGS019:	CCACTGGCTGTACAGCCTGACCTTCAATGTTGATGAGAG	SEQ ID No.140
BGS020:	CAGGCTGTAGCAGCCAGTGGGTTCTCAGTGGGCACAAA	SEQ ID No.141
BGS021:	CTGGCTGCAAGAAGGCCAGACCAGGATCATCTTTGATGGA	SEQ ID No.142
BGS022:	TCTGGCCTTCTGCGCCAGCTCTCATCAACATGAAGGT	SEQ ID No.143
BGS023:	GTCAACTCTGCTTCCACCTCTGGTGCAATGGCAGGTGGG	SEQ ID No.144
BGS024:	AGGTGGAAGGCAGAGTTGACTCCATCAAGAGATGATCCTGG	SEQ ID No.145
BGS025:	TTGGCTATGGCCAAGACAGCAGGCTGCCCTCTGAGTTTGA	SEQ ID No.146
BGS026:	GCTGTCTTGCCCATAGCCAAACCCACCTGCCATTGCACCAG	SEQ ID No.147
BGS027:	CCTCTCTGCTTCCCTCAGAGCTGGAGAGAACAGGCTGGCT	SEQ ID No.148
BGS028:	CTCTGAGGAAGGCAGAGAGGTCAAACCTCAGAGGGCAGCCT	SEQ ID No.149
BGS029:	GTGATGGTGCTCAGGTGGTCTGATGGCAGCTACCTGGAAG	SEQ ID No.150
BGS030:	GACCACCTGAGCACCATGACAGCCAGCCTGTTCTCTCCAG	SEQ ID No.151
BGS031:	ACCAAGACATGTGGAGGATGTCTGGCATCTTCAGGGATGT	SEQ ID No.152
BGS032:	CATCCTCCACATGTCTTGGTCTTCCAGGTAGCTGCCATCA	SEQ ID No.153
BGS033:	GAGCCTGCTGCACAAGCCCCACCCAGATTTCTGACTTC	SEQ ID No.154
BGS034:	TGGGCTTGTGCAAGGCTCACATCCCTGAAGATGCCAGA	SEQ ID No.155
BGS035:	CATGTTGCCACCAGGTTCAATGATGACTTCAGCAGAGCTG	SEQ ID No.156
BGS036:	TTGAACCTGGTGGCAACATGGAAGTCAGAAATCTGGGTGG	SEQ ID No.157
BGS037:	TGCTGGAGGCTGAGGTGCAGATGTGTGGAGAACTCAGAGA	SEQ ID No.158
BGS038:	CTGCACCTCAGCCTCCAGCACAGCTCTGCTGAAGTCATCA	SEQ ID No.159
BGS039:	CTACCTGAGAGTCACAGTGAGCCTCTGGCAAGGTGAGACC	SEQ ID No.160
BGS040:	TCACTGTGACTCTCAGGTAGTCTCTGAGTTCTCCACACAT	SEQ ID No.161
BGS041:	CAGGTGGCCTCTGGCACAGCCCCCTTTGGAGGAGAGATCA	SEQ ID No.162
BGS042:	GCTGTGCCAGAGGCCACCTGGGTCTCACCTTGCCAGAGGC	SEQ ID No.163
BGS043:	TTGATGAGAGAGGAGCTATGCTGACAGAGTCAACCTGAG	SEQ ID No.164
BGS044:	ATAGCCTCCTCTCTCATCAATGATCTCTCTCCAAAGGGG	SEQ ID No.165
BGS045:	GCTCAATGTGGAGAACCCCAAGCTGTGGTCTGCTGAGATC	SEQ ID No.166
BGS046:	TGGGGTTCTCCACATTGAGCCTCAGGGTGACTCTGTGAGC	SEQ ID No.167
BGS047:	CCCAACCTCTACAGGGCTGTTGTGGAGCTGCACACTGCTG	SEQ ID No.168
BGS048:	ACAGCCCTGTAGAGGTTGGGGATCTCAGCAGACCACAGCT	SEQ ID No.169
BGS049:	ATGGCACCTGATTGAAGCTGAAGCCTGTGATGTTGGATT	SEQ ID No.170
BGS050:	AGCTTCAATCAGGGTGCCATCAGCAGTGTGAGCTCCACA	SEQ ID No.171
BGS051:	CAGAGAAATCAGGATTGAGAATGGCCTGCTGCTGCTCAAT	SEQ ID No.172
BGS052:	TCTCAATCCTGACTTCTCTGAATCCAACATCACAGGCTTC	SEQ ID No.173
BGS053:	GGCAAGCCTCTGCTCATCAGGGGAGTCAACAGGCATGAGC	SEQ ID No.174
BGS054:	CTGATGAGCAGAGGCTTGCCATTGAGCAGCAGGCCAT	SEQ ID No.175
BGS055:	ACCACCTCTGCATGGACAAGTGATGGATGAACAGACAAT	SEQ ID No.176

13/27

BGS056: TTGTCCATGCAGAGGGTGGTGCTCATGCCTGTTGACTCCC
BGS057: GGTGCAAGATATCCTGCTGATGAAGCAGAACTCCGCCTAC
BGS058: TCAGCAGGATATCTTGACCATTTGTCTGTTTCATCCATCAC
BGS059: GTAGGCGGAGTTCTGCTTCA

SEQ ID No.177
SEQ ID No.178
SEQ ID No.179
SEQ ID No.180

Figure 10a

BGS060: TCATTAGCAGGATATCTTGC	SEQ ID No.181
BGS061: GCAAGATATCTGCTAATGAAGCAGAACAACCTTCAATGCT	SEQ ID No.182
BGS062: GGGTAGTGAGAGCACCTGACAGCATTGAAGTTGTTCTGCT	SEQ ID No.183
BGS063: GTCAGGTGCTCTCACTACCCCAACCACCTCTCTGGTACA	SEQ ID No.184
BGS064: GCCATACCTGTCACACAGGGGTACCAGAGAGGGTGGTTG	SEQ ID No.185
BGS065: CCCTGTGTGACAGGTATGGCCTGTATGTTGTTGATGAAGC	SEQ ID No.186
BGS066: TGCCATGTGTCTCAATGTTGGCTTCATCAACAACATACAG	SEQ ID No.187
BGS067: CAACATTGAGACACATGGCATGGTGCCCATGAACAGGCTC	SEQ ID No.188
BGS068: AGCCACCTGGGGTCACTGTGAGCCTGTTTCATGGGCACCA	SEQ ID No.189
BGS069: ACAGATGACCCAGGTGGCTGCCTGCCATGCTGAGAGAG	SEQ ID No.190
BGS070: TCTCTGCACCATCCTGGTCACTCTCTCAGACATGGCAGGC	SEQ ID No.191
BGS071: TGACCAGGATGGTGCAGAGAGACAGGAACCACCCCTCTGT	SEQ ID No.192
BGS072: TGCCCAGAGACCAGATGATCAGAGGGGTGGTTCTGTCT	SEQ ID No.193
BGS073: GATCATCTGGTCTCTGGGCAATGAGTCTGGACATGGAGCC	SEQ ID No.194
BGS074: CTGTAGAGAGCATCATGGTTGGCTCCATGTCCAGACTCAT	SEQ ID No.195
BGS075: AACCATGATGCTCTCTACAGGTGGATCAAGTCTGTTGACC	SEQ ID No.196
BGS076: ATACTGCACAGGTCTGCTGGGGTCAACAGACTTGATCCAC	SEQ ID No.197
BGS077: CCAGCAGACCTGTGCAGTATGAAGGAGGTGGAGCAGACAC	SEQ ID No.198
BGS078: AGATGATGTCTGTGGCTGTGGTGTCTGCTCCACCTCCTTC	SEQ ID No.199
BGS079: CACAGCCACAGACATCATCTGCCCATGTATGCCAGGGTT	SEQ ID No.200
BGS080: GGGAAAGGGCTGGTCTCTCATCAACCTGGCATAACATGGGGC	SEQ ID No.201
BGS081: GATGAGGACCAGCCCTTCCCTGCTGTGCCCAAGTGAGCA	SEQ ID No.202
BGS082: CAGAGAGAGCCACTTCTTGATGCTCCACTTGGGCACAGCA	SEQ ID No.203
BGS083: TCAAGAAAGTGGCTCTCTCTGCTGGAGAGACCAGACCTCT	SEQ ID No.204
BGS084: GTGCATATTACACAGGATCAGAGGTCTGGTCTCTCCAGG	SEQ ID No.205
BGS085: GATCCTGTGTGAATATGCACATGCAATGGGCAACTCTCTG	SEQ ID No.206
BGS086: CAGTACTTGGCAAAGCCTCCAGAGAGTTGCCCATGTCAT	SEQ ID No.207
BGS087: GGAGGCTTTGCCAAGTACTGGCAAGCCTTCAGACAGTACC	SEQ ID No.208
BGS088: AAATCCTCCTTGCAAGCCTGGGGTACTGTCTGAAGGCTTGC	SEQ ID No.209
BGS089: CCAGGCTGCAAGGAGGATTTGTGTGGGACTGGGTGGACCA	SEQ ID No.210
BGS090: CATCATACTTGATGAGAGATTGGTCCACCCAGTCCACAC	SEQ ID No.211
BGS091: ATCTCTCATCAAGTATGATGAGAATGGCAACCCCTGGTCT	SEQ ID No.212
BGS092: CCAAAGTCTCCTCCATAGGCAGACCAGGGGTGCCATTCT	SEQ ID No.213
BGS093: GCCTATGGAGGAGACTTTGGTGACACCCCAATGACAGGC	SEQ ID No.214
BGS094: CAGGCCATTTCATGCAGAACTGCCTGTCTTGGGGTGTCA	SEQ ID No.215
BGS095: AGTTCTGCATGAATGGCCTGGTCTTTGCAGACAGGACCCC	SEQ ID No.216
BGS096: CCTCTGTGAGGGCAGGGGTGAGGGTCTGTCTGCAAAAGAC	SEQ ID No.217
BGS097: TCACCCCTGCCCTCACAGAGCCAAGCACCAGCAACAGTTC	SEQ ID No.218
BGS098: CCAGACAGCCTGAACTGGAAGAACTGTTGCTGGTGTCTGG	SEQ ID No.219
BGS099: TTCCAGTTGAGGCTGTCTGGACAGACCATTGAGGTGACAT	SEQ ID No.220
BGS100: GTGCCTGAAGAGGTACTCAGATGTCACCTCAATGGTCTGT	SEQ ID No.221
BGS101: CTGAGTACCTCTTCAGGCACTCTGACAATGAGTCTCCTGCA	SEQ ID No.222
BGS102: TGCAGGAGCTCATTTGTGAGA	SEQ ID No.223

Figure 10b

BGS103: GTAATTTAACAATGAGCTCCTGCACTGGATGGTGGCCCTG	SEQ ID No.224
BGS104: GGAGCTCATTGTAAATTAC	SEQ ID No.225
BGS105: GATGGCAAGCCTCTGGCTTCTGGTGAGGTGCCTCTGGATG	SEQ ID No.226
BGS106: GAAGCCAGAGGCTTGCCATCCAGGGCCACCATCCAGTGCA	SEQ ID No.227
BGS107: TGGCCCTCAAGGAAAGCAGCTGATTGAACCTGCCCTGAGCT	SEQ ID No.228
BGS108: CTGCTTTCTTGAGGGGCCACATCCAGAGGCACCTCACCA	SEQ ID No.229
BGS109: GECTCAGCCAGAGTCTGCTGGACAACCTGTGGCTAACAGTG	SEQ ID No.230
BGS110: CAGCAGACTCTGGCTGAGGCAGCTCAGGCAGTTCAATCAG	SEQ ID No.231
BGS111: AGGGTGGTTTCAAGCCCAATGCAACAGCTTGGTCTGAGGCAG	SEQ ID No.232
BGS112: GCATTGGGCTGAACCACCTCACTGTTAGCCACAGTTGTC	SEQ ID No.233
BGS113: GCCACATCTCTGCATGGCAGCAGTGGAGGCTGGCTGAGAA	SEQ ID No.234
BGS114: CTGCCATGCAGAGATGTGGCCTGCCTCAGACCAAGCTGTT	SEQ ID No.235
BGS115: CCTCTCTGTGACCTGCCTGCTGCCTCTCATGCCATCCCT	SEQ ID No.236
BGS116: CAGGCAGGGTCACAGAGAGGTTCTCAGCCAGCCTCCACTG	SEQ ID No.237
BGS117: CACCTGACAACATCTGAAATGGACTTCTGCATTGAGCTGG	SEQ ID No.238
BGS118: ATTTTCAGATGTTGTGAGGTGAGGGATGGCATGAGAGGCAG	SEQ ID No.239
BGS119: GCAACAAGAGATGGCAGTTCAACAGGCAGTCTGGCTTCCT	SEQ ID No.240
BGS120: GAACTGCCATCTCTTGTGCCCAGCTCAATGCAGAAGTCC	SEQ ID No.241
BGS121: GTCTCAGATGTGGATTGGAGACAAGAAGCAGCTCCTCACC	SEQ ID No.242
BGS122: CTCCAATCCACATCTGAGACAGGAAGCCAGACTGCCTGTT	SEQ ID No.243
BGS123: CCTCTCAGGGACCAATTACACAGGGCTCCTCTGGACAATG	SEQ ID No.244
BGS124: GTGAATTGGTCCCTGAGAGGGGTGAGGAGCTGCTTCTTGT	SEQ ID No.245
BGS125: ACATTGGAGTGTCTGAGGCCACCAGGATTGACCCAAATGC	SEQ ID No.246
BGS126: GGCCTCAGACACTCCAATGTCTTGTCCAGAGGAGCCCTG	SEQ ID No.247
BGS127: TTGGGTGGAGAGGTGGAAGGCTGCTGGCACTACCAGGCT	SEQ ID No.248
BGS128: CCTTCCACCTCTCCACCCAAGCATTGGGGTCAATCCTGGT	SEQ ID No.249
BGS129: GAGGCTGCCCTGCTCCAGTGACAGCAGACACCTGGCTG	SEQ ID No.250
BGS130: CACTGGAGCAGGGCAGCCTCAGCCTGGTAGTGTCCAGCAG	SEQ ID No.251
BGS131: ATGCTGTTCTGATCACACAGCCCATGCTTGGCAGCACCA	SEQ ID No.252
BGS132: TGTGGTGATCAGAACAGCATCAGCCAGGGGTGCTGCTGTG	SEQ ID No.253
BGS133: AGGCAAGACCCTGTTTCATCAGCAGAAAGACCTACAGGATT	SEQ ID No.254
BGS134: TGATGAACAGGGTCTTGCCTTGGTCTGCCAAGCATGGGC	SEQ ID No.255
BGS135: GATGGCTCTGGACAGATGGCAATCACAGTGGATGTGGAGG	SEQ ID No.256
BGS136: GCCATCTGTCCAGAGCCATCAATCCTGTAGGTCTTTCTGC	SEQ ID No.257
BGS137: TTGCCTCTGACACACCTCACCTGCAAGGATTGGCCTGAA	SEQ ID No.258
BGS138: GTGAGGTGTGTGACAGGGCAACCTCCACATCCACTGTGATT	SEQ ID No.259
BGS139: CTGTCAACTGGCACAGGTGGCTGAGAGGGTGAACCTGGCTG	SEQ ID No.260
BGS140: CCACCTGTGCCAGTTGACAGTTTCAAGCCAACTCTTGCAGG	SEQ ID No.261
BGS141: GGCTTAGGGCCTCAGGAGAACTACCTGACAGGCTGACAG	SEQ ID No.262
BGS142: TTCTCCTGAGGGCCTAAGCCAGCCAGTTTCAACCTCTCAG	SEQ ID No.263
BGS143: CTGCCCTGCTTTGACAGGTGGGACCTGCCTCTGTCTGACAT	SEQ ID No.264
BGS144: CCACCTGTCAAAGCAGGCAGCTGTGAGCCTGTGAGGGTAG	SEQ ID No.265
BGS145: GTACACCCCTTATGTGTTCCCTTCTGAGAATGGCTGAGG	SEQ ID No.266
BGS146: GGAACACATAAGGGGTGTACATGTGACAGAGAGGCAGGTC	SEQ ID No.267
BGS147: TGTGGCACCAAGGAGCTGAACTATGGTCCCTCACCAGTGGA	SEQ ID No.268
BGS148: TTCAGCTCCCTGGTGCCACACCTCAGGCCATTCTCAGAAG	SEQ ID No.269
BGS149: GGGGAGACTTCCAGTTCAACATCTCCAGGTAATCTCTCAGCA	SEQ ID No.270
BGS150: GTTGAACCTGGAAGTCTCCCTCCACTGGTGAGGACCATAG	SEQ ID No.271
BGS151: ACAGCTCATGGAAACCTCTCAGGCACCTGCTCCATGCA	SEQ ID No.272
BGS152: GAGAGGTTTCCATGAGCTGTTGCTGAGAGTACCTGGAGAT	SEQ ID No.273
BGS153: GAGGAGGGAACCTGGCTGAACATTGATGGCTTCCACATGG	SEQ ID No.274
BGS154: TTCAGCCAGGTTCCCTCCTCTGATGGAGCAGGTGCCTGT	SEQ ID No.275
BGS155: GCATTGGAGGAGATGACTCTTGGTCTCCTTCTGTGTCTGC	SEQ ID No.276
BGS156: AGAGTCATCTCTCCAATGCCCATGTGGAAGCCATCAATG	SEQ ID No.277

16/27

BGS157: TGAGTTCCAGTTATCTGCTGGCAGGTACCACTATCAGCTG
BGS158: CAGCAGATAACTGGAACTCAGCAGACACAGAAGGAGACCA
BGS159: GTGTGGTGCCAGAAGTAAACCTGAGCTAGCAGTCCATGAT
BGS160: GTTTACTTCTGGCACCACACCAGCTGATAGTGGTACCTGC
BGS161: ATCATGGACTGCTAGCTCAG

SEQ ID No.278
SEQ ID No.279
SEQ ID No.280
SEQ ID No.281
SEQ ID No.282

Figure 10c

AseI	NcoI
1-CAATTAATCATcggGCATAGTATATcgGCATAGTATAATAcgACTCACTATAGGAGGGGCCCATGG	SEQ ID No.297
2-CAATTAATCATDGGCATAGTATATCwGCATAGTATAATACHACTCACTATAGGAGGGGCCCATGG	SEQ ID No.298
3-CAATTAATCATtGGCATAGTATATCtGCATAGTATAATAcACTCACTATAGGAGGGGCCCATGG	SEQ ID No.11

Figure 11

1 GCAGGACTGAGGCTTAATTAAACCTTAAACCTTTAAAAGCCTTATATATTCTTTTTT
60 TCTTATAAACTTAAACCTTAGAGGCTATTTAAGTTGCTGATTATATTAATTTIATTG
120 TTCAAACATGAGAGCTTAGTACATGAAACATGAGAGCTTAGTACATTAGCCATGAGAGCT
180 TAGTACATTAGCCATGAGGCTTTAGTTCATTAAACATGAGAGCTTAGTACATTAAACATG
240 AGAGCTTAGTACATGAAACATGAGAGCTTAGTACATACTATCAACAGGTTGAACTGCTGA
300 TCTTAATTAACTGGAGACTT

PacI
PacI

SEQ ID No 12

Figure 12

RK1 : GCAGGACTGAGGCTTAATTAAACCTTAAACCTTTAAAAG	SEQ ID No. 283
RK2 : CCTTATATATTCTTTTTTCTTATAAACTTAAACCTT	SEQ ID No. 284
RK3 : AGAGGCTATTTAAGTTGCTGATTTATATTAAATTTATTGT	SEQ ID No. 285
RK4 : TCAAAACATGAGAGCTTAGTACATGAAACATGAGAGCTTAGTACATTAGCC	SEQ ID No. 286
RK5 : ATGAGAGCTTAGTACATTAGCCATGAGGTTTAGTTTCATTAACATGAGAGCTTAGTACA	SEQ ID No. 287
RK6 : TTAAACATGAGAGCTTAGTACATGAAACATGAGAGCTTAGTACA_TACTATCAACAGGTTG	SEQ ID No. 288
RK7 : AACTGCTGATCTTAATTAACCTGGAGACTT	SEQ ID No. 289
RK8 : AAGTCTCCAGGTTAATTAAGATCAGCAGTTCAACCTGTTGATAGTATGTAAGCTCTC	SEQ ID No. 290
RK9 : ATGTTTCATGTACTAAGCTCTCATGTTTAATGTACTAAGCTCTCATGTTTAATGAACTAA	SEQ ID No. 291
RK10 : ACCCTCATGGCTAA_TGTACTAAGCTCTCATGGCTAATGTACTAAGCTCTCATGTTTCATG	SEQ ID No. 292
RK11 : TACTAAGCTCTCATGTTTGAACAATAAAATTAATATAAAT	SEQ ID No. 293
RK12 : CAGCAACTTAAATAGCCTCTAAGGTTTAAAGTTTATAAG	SEQ ID No. 294
RK13 : AAAAAAAGAAATATATAAGGCTTTTAAAGGTTTAAAGTT	SEQ ID No. 295
RK14 : TAATTAAGCCTCAGTCCTGC	SEQ ID No. 296

Figure 13

20/27

PacI
1 TTAATTACCTTAAAACCTTTAAAAGCCTTATATATTCTTTTTTTCTTATAAACTTA
60 AAACCTTAGAGGCTATTTAAGTTGCTGATTTATATTAATTTATTGTTCAAACATGAGAG
120 CTTAGTACATGAAACATGAGAGCTTAGTACATTAGCCATGAGAGCTTAGTACATTAGCCA
180 TGAGGGTTTAGTTCATTAAACATGAGAGCTTAGTACATTAAACATGAGAGCTTAGTACAT
PacI
240 ACTATCAACAGGTTGAACTGCTGATCTTAATTAA SEQ ID No 13

Figure 14

PacI

1 TTAATTAACAGTAGTTGACAATTAAACATTGGCATAGTATATCTGCATAGTATAATAC
NcoI

61 AACTCACTATAGGAGGGCCACCATGGCCAAGTTGACCAGTGCTGTCCCAGTGCTCACAGC
MetAlaLysLeuThrSerAlaValProValLeuThrAl
121 CAGGGATGTGGCTGGAGCTGTTGAGTTCTGGACTGACAGGTGGGGTTCTCCAGAGATTT
aArgAspValAlaGlyAlaValGluPheTrpThrAspArgLeuGlyPheSerArgAspPh
181 TGTGGAGGATGACTTTGCAGGTGTGGTCAGAGATGATGTACCCTGTTCATCTCAGCAGT
eValGluAspAspPheAlaGlyValValArgAspAspValThrLeuPheIleSerAlaVa
241 CCAGGACCAGGTGGTGCCTGACCAACCCCTGGCTGGGTGTGGGTGAGAGGACTGGATGA
lGlnAspGlnValValProAspAsnThrLeuAlaTrpValTrpValArgGlyLeuAspGl
301 GCTGTATGCTGAGTGGAGTGAGGTGGTCTCCACCACTTCAGGGATGCCAGTGGCCCTGC
uLeuTyrAlaGluTrpSerGluValValSerThrAsnPheArgAspAlaSerGlyProAl
361 CATGACAGAGATTGGAGAGCAGCCCTGGGGAGAGAGTTTGCCTGAGAGACCCAGCAGG
aMetThrGluIleGlyGluGlnProTrpGlyArgGluPheAlaLeuArgAspProAlaGl

EcoRI

421 CAACTGTGTGCACTTGTGGCAGAGGAGCAGGACTGAGGATAAGAATTCTGAGGAGAAGC
yAsnCysValHisPheValAlaGluGluGlnAsp***
481 TCATGGACCCTGTTGTGCTGCAAAGGAGAGACTGGGAGAACCCCTGGAGTGACCCAGCTCA
MetAspProValValLeuGlnArgArgAspTrpGluAsnProGlyValThrGlnLeuA
541 ACAGACTGGCTGCCACCCCTCCCTTGCCTCTTGGAGGAACCTCTGAGGAAGCCAGGACAG
snArgLeuAlaAlaHisProProPheAlaSerTrpArgAsnSerGluGluAlaArgThrA
601 ACAGGCCAGCCAGCAGCTCAGGTCTCTCAATGGAGAGTGGAGTTTGCCTGCTTCCCTG
spArgProSerGlnGlnLeuArgSerLeuAsnGlyGluTrpArgPheAlaTrpPheProA
661 CCCCTGAAGCTGTGCCTGAGTCTTGGCTGGAGTGTGACCTCCCAGAGGCTGACACTGTTG
laProGluAlaValProGluSerTrpLeuGluCysAspLeuProGluAlaAspThrValV
721 TGGTGGCCAGCAACTGGCAGATGCATGGCTATGATGCCCCATCTACACCAATGTCACCT
alValProSerAsnTrpGlnMetHisGlyTyrAspAlaProIleTyrThrAsnValThrT
781 ACCCCATCACTGTGAACCCCTTTTGTGCCCACTGAGAACCCCACTGGCTGCTACAGCC
yrProIleThrValAsnProProPheValProThrGluAsnProThrGlyCysTyrSerL
841 TGACCTTCAATGTTGATGAGAGCTGGCTGCAAGAAGGCCAGACCAGGATCATCTTGATG
euThrPheAsnValAspGluSerTrpLeuGlnGluGlyGlnThrArgIleIlePheAspG
901 GACTCAACTCTGCCTTCCACCTCTGCTGCAATGGCAGGTGGGTGGCTATGGCCAGACA
lyValAsnSerAlaPheHisLeuTrpCysAsnGlyArgTrpValGlyTyrGlyGlnAspS
961 GCAGGCTGCCCTCTGAGTTGACCTCTCTGCCTTCTCAGAGCTGGAGAGAACAGGCTGG
ergLeuProSerGluPheAspLeuSerAlaPheLeuArgAlaGlyGluAsnArgLeuA
1021 CTGTGCATGGTGCTCAGGTGGTCTGATGGCAGCTACCTGGAGACCAAGACATGTGGAGGA
laValMetValLeuArgTrpSerAspGlySerTyrLeuGluAspGlnAspMetTrpArgM
1081 TGCTTGGCATCTTCAGGGATGTGAGCCTGCTGCACAAGCCCAACCCAGATTCTGACT
etSerGlyIlePheArgAspValSerLeuLeuHisLysProThrThrGlnIleSerAspP
1141 TCCATGTTGCCACCAGGTTCAATGATGACTTCAGCAGAGCTGTGCTGGAGGCTGAGGTGC
heHisValAlaThrArgPheAsnAspAspPheSerArgAlaValLeuGluAlaGluValG
1201 AGATGTGTGGAGAACTCAGAGACTACCTGAGAGTCACAGTGAGCCTCTGGCAAGGTGAGA
lnMetCysGlyGluLeuArgAspTyrLeuArgValThrValSerLeuTrpGlnGlyGluT
1261 CCCAGGTGGCCTCTGGCACAGCCCCCTTTGGAGGAGAGATCATTGATGAGAGAGGAGGCT
hrGlnValAlaSerGlyThrAlaProPheGlyGlyGluIleIleAspGluArgGlyGlyT
1321 ATGCTGACAGAGTCACCCCTGAGGCTCAATGTGGAGAACCCCAAGCTGTGGTCTGCTGAGA
yrAlaAspArgValThrLeuArgLeuAsnValGluAsnProLysLeuTrpSerAlaGluI
1381 TCCCCAACCTCTACAGGGCTGTTGTGGAGCTGCACACTGCTGATGGCACCTGATTGAAG
leProAsnLeuTyrArgAlaValValGluLeuHisThrAlaAspGlyThrLeuIleGluA
1441 CTGAAGCCTGTGATGTTGGATTTCAGAGAAGTCAGGATTGAGAATGGCCTGCTGCTGCTCA
laGluAlaCysAspValGlyPheArgGluValArgIleGluAsnGlyLeuLeuLeuA
1501 ATGGCAAGCCTCTGCTCATCAGGGGAGTCAACAGGCATGAGCACCACCTCTGCATGGAC
snGlyLysProLeuLeuIleArgGlyValAsnArgHisGluHisHisProLeuHisGlyG
1561 AAGTGATGGATGAACAGACAAATGGTGCAAGATATCTGCTAATGAAGCAGAACAACCTCA
lnValMetAspGluGlnThrMetValGlnAspIleLeuLeuMetLysGlnAsnAsnPheA
1621 ATGCTGTGAGGTGCTCTCACTACCCCAACACCTCTCTGGTACACCTGTGTGACAGGT
snAlaValArgCysSerHisTyrProAsnHisProLeuTrpTyrThrLeuCysAspArgT
1681 ATGGCCTGTATGTTGTTGATGAAGCCAACATTGAGACACATGGCATGGTGGCCATGAACA

yrGlyLeuTyrValValAspGluAlaAsnIleGluThrHisGlyMetValProMetAsnA
 1741 GGCTCACAGATGACCCAGGTGGCTGCCTGCCATGTCTGAGAGAGTGACCAGGATGGTGC
 rgLeuThrAspAspProArgTrpLeuProAlaMetSerGluArgValThrArgMetValG
 1801 AGAGAGACAGGAACCAACCCCTCTGTGATCATCTGGTCTCTGGGCAATGAGTCTGGACATG
 lnArgAspArgAsnHisProSerValIleIleTrpSerLeuGlyAsnGluSerGlyHisG
 1861 GAGCCAACCATGATGCTCTCTACAGGTGGATCAAGTCTGTTGACCCAGCAGACCTGTGC
 lyAlaAsnHisAspAlaLeuTyrArgTrpIleLysSerValAspProSerArgProValG
 1921 AGTATGAAGGAGGTGGAGCAGACACCACAGCCACAGACATCATCTGCCCATGTATGCCA
 lnTyrGluGlyGlyGlyAlaAspThrThrAlaThrAspIleIleCysProMetTyrAlaA
 1981 GGGTTGATGAGGACCACCCCTTCCCTGCTGTGCCCAAGTGGAGCATCAAGAAGTGGCTCT
 rgValAspGluAspGlnProPheProAlaValProLysTrpSerIleLysLysTrpLeuS
 2041 CTCTGCTGGAGAGACCAGACCTCTGATCCTGTGTAATATGCACATGCAATGGGCAACT
 erLeuProGlyGluThrArgProLeuIleLeuCysGluTyrAlaHisAlaMetGlyAsnS
 2101 CTCTGGGAGGCTTTGCCAAGTACTGGCAAGCCTTCAGACAGTACCCAGGCTGCAAGGAG
 erLeuGlyGlyPheAlaLysTyrTrpGlnAlaPheArgGlnTyrProArgLeuGlnGlyG
 2161 GATTGTGTGGGACTGGGTGGACCAATCTCTCATCAAGTATGATGAGAATGGCAACCCCT
 lyPheValTrpAspTrpValAspGlnSerLeuIleLysTyrAspGluAsnGlyAsnProT
 2221 GGTCTGCCTATGAGGAGACTTTGGTGACACCCCAATGCAGGCAGTTCCTGCATGAATG
 rpSerAlaTyrGlyGlyAspPheGlyAspThrProAsnAspArgGlnPheCysMetAsnG
 2281 GCCTGGTCTTTGCAGACAGGACCCCTCACCTGCCCTCACAGAGGCCAAGCACCAGCAAC
 lyLeuValPheAlaAspArgThrProHisProAlaLeuThrGluAlaLysHisGlnGlnG
 2341 AGTTCTTCCAGTTCAGGCTGTCTGGACAGACCATGAGGTGACATCTGAGTACCTCTTCA
 lnPhePheGlnPheArgLeuSerGlyGlnThrIleGluValThrSerGluTyrLeuPheA
 2401 GGCACCTCTGACAATGAGCTCCTGCCTGGATGGTGGCCCTGGATGGCAAGCCTCTGGCTT
 rgHisSerAspAsnGluLeuLeuHisTrpMetValAlaLeuAspGlyLysProLeuAlaS
 2461 CTGGTGAGGTGCCTCTGGATGTGGCCCTCAAGGAAAGCAGCTGATGAACTGCTGAGC
 erGlyGluValProLeuAspValAlaProGlnGlyLysGlnLeuIleGluLeuProGluL
 2521 TGCTCAGCCAGAGTCTGCTGGACAACCTGTGGCTAACAGTGAGGGTGGTTCAGCCCAATG
 euProGlnProGluSerAlaGlyGlnLeuTrpLeuThrValArgValValGlnProAsnA
 2581 CAACAGCTTGGTCTGAGGCAGGCCACATCTCTGCATGGCAGCAGTGGAGGCTGGCTGAGA
 laThrAlaTrpSerGluAlaGlyHisIleSerAlaTrpGlnGlnTrpArgLeuAlaGluA
 2641 ACCTCTCTGTGACCTGCCTGCTGCTCTCATGCCATCCCTCACCTGACAACATCTGAAA
 snLeuSerValThrLeuProAlaAlaSerHisAlaIleProHisLeuThrThrSerGluM
 2701 TGGACTTCTGCATTGAGCTGGGCAACAAGAGATGGCAGTTCAACAGGCAGTCTGGCTTCC
 etAspPheCysIleGluLeuGlyAsnLysArgTrpGlnPheAsnArgGlnSerGlyPheL
 2761 TGTCTCAGATGTGGATTGGAGACAAGAAGCAGCTCCTCACCCCTCTCAGGGACCAATCA
 euSerGlnMetTrpIleGlyAspLysLysGlnLeuLeuThrProLeuArgAspGlnPheT
 2821 CCAGGGCTCCTCTGGACAATGACATTGGAGTGTCTGAGGCCACCAGGATTGACCCAAATG
 hrArgAlaProLeuAspAsnAspIleGlyValSerGluAlaThrArgIleAspProAsnA
 2881 CTTGGGTGGAGAGGTGGAAGGCTGCTGGACACTACCAGGCTGAGGCTGCCCTGCTCCAGT
 laTrpValGluArgTrpLysAlaAlaGlyHisTyrGlnAlaGluAlaAlaLeuLeuGlnC
 2941 GCACAGCAGACACCTGGCTGATGCTGTTCTGATCACCACAGCCCATGTCTGGCAGCACC
 ysThrAlaAspThrLeuAlaAspAlaValLeuIleThrThrAlaHisAlaTrpGlnHisG
 3001 AAGGCAAGACCCCTGTTTCATCAGCAGAAAGACCTACAGGATTGATGGCTCTGGACAGATGG
 lnGlyLysThrLeuPheIleSerArgLysThrTyrArgIleAspGlySerGlyGlnMetA
 3061 CAATCAGTGGATGTGGAGGTTCCTCTGACACACCTCACCTGCAAGGATTGGCCTGA
 laIleThrValAspValGluValAlaSerAspThrProHisProAlaArgIleGlyLeuA
 3121 ACTGTCAACTGGCACAGGTGGCTGAGAGGGTGAAGTGGCTGGGCTTAGGCCCTCAGGAGA
 snCysGlnLeuAlaGlnValAlaGluArgValAsnTrpLeuGlyLeuGlyProGlnGluA
 3181 ACTACCTGACAGGCTGACAGCTGCCTTTGACAGGTGGGACCTGCCTCTGTCTGACA
 snTyrProAspArgLeuThrAlaAlaCysPheAspArgTrpAspLeuProLeuSerAspM
 3241 TGTACACCCCTTATGTGTTCCCTTCTGAGAATGGCCTGAGGTGTGGCACCAGGGAGCTGA
 etTyrThrProTyrValPheProSerGluAsnGlyLeuArgCysGlyThrArgGluLeuA
 3301 ACTATGGTCTCACCAGTGGAGGGAGACTTCCAGTTCAACATCTCCAGTACTCTCAGC
 snTyrGlyProHisGlnTrpArgGlyAspPheGlnPheAsnIleSerArgTyrSerGlnG
 3361 AACAGCTCATGAAACCTCTCACAGGCACCTGCTCCATGCAGAGGAGGGAACTGGCTGA
 lnGlnLeuMetGluThrSerHisArgHisLeuLeuHisAlaGluGluGlyThrTrpLeuA
 3421 ACATTATGGCTTCCCATGGGATGGAGGAGATGACTCTGGTCTCCTTCTGTGTCTG
 snIleAspGlyPheHisMetGlyIleGlyGlyAspAspSerTrpSerProSerValSerA

3481 CTGAGTTCAGTTATCTGCTGGCAGGTACCACTATCAGCTGGTGTGGTGCCAGAAGTAAG
laGluPheGlnLeuSerAlaGlyArgTyrHisTyrGlnLeuValTrpCysGlnLys*** SEQ ID No.15
NheI PacI
3541 CTAGCTGAGTTTCAGAAAAGGGGGCCTGAGTGGCCCCTTTTTTCAACTTAATTAACCTTA
3601 AAACCTTTAAAAGCCTTATATATTCTTTTTTTCTTATAAACTTAAACCTTAGAGGCT
3661 ATTTAAGTTGCTGATTTATATTAATTTTATTGTTCAAACATGAGAGCTTAGTACATGAAA
3721 CATGAGAGCTTAGTACATTAGCCATGAGAGCTTAGTACATTAGCCATGAGGGTTTAGTTC
3781 ATTAACATGAGAGCTTAGTACATTAAACATGAGAGCTTAGTACATACTATCAACAGGTT
3841 GAACTGCTGATC SEQ ID No.14

Figure 15

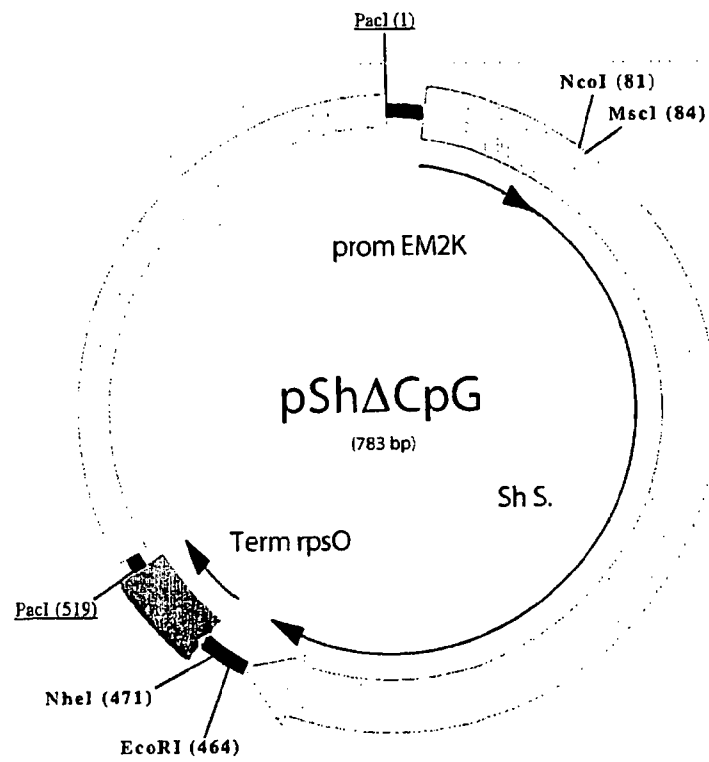


FIG. 16

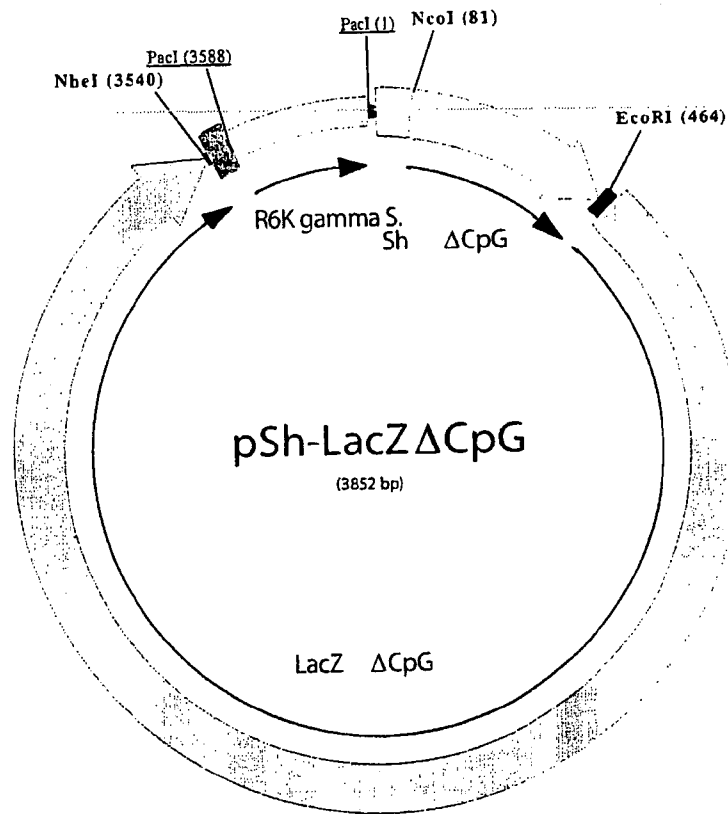


FIG. 17

BspHI
 1 TCATGATTGAACAAGATGGCCTACATGCAGGTTCTCCAGCTGCCTGGGTTGAGAGACTG
 1 MetIleGluGlnAspGlyLeuHisAlaGlySerProAlaAlaTrpValGluArgLeu
 60 TTTGGCTATGACTGGGCACAGCAGACCATTGGTTGCTCTGATGCAGCAGTTTTCAGACTT
 20 PheGlyTyrAspTrpAlaGlnGlnThrIleGlyCysSerAspAlaAlaValPheArgLeu
 120 TCAGCCCAAGGCAGGCCAGTCCTTTTGTAAAGACAGACCTCAGTGGGGCTCTCAATGAG
 40 SerAlaGlnGlyArgProValLeuPheValLysThrAspLeuSerGlyAlaLeuAsnGlu
 180 CTCCAGGATGAGGCTGCCAGACTCTCCTGGTTGGCAACAACCTGGGGTCCCCTGTGCAGCT
 60 LeuGlnAspGluAlaAlaArgLeuSerTrpLeuAlaThrThrGlyValProCysAlaAla
 240 GTCCTTGATGTGGTCACAGAAGCTGGAAGGGACTGGCTCCTACTAGGTGAGGTGCCTGGG
 80 ValLeuAspValValThrGluAlaGlyArgAspTrpLeuLeuLeuGlyGluValProGly
 300 CAGGACCTCCTTTCTCTCACCTAGCTCCAGCTGAGAAAGTGTCAATCATGGCTGATGCC
 100 GlnAspLeuLeuSerSerHisLeuAlaProAlaGluLysValSerIleMetAlaAspAla
 360 ATGAGAAGACTCCACACCCTTGACCCAGCCACCTGCCCTTTGACCACCAGGCCAAGCAC
 120 MetArgArgLeuHisThrLeuAspProAlaThrCysProPheAspHisGlnAlaLysHis
 420 AGGATAGAGAGGGCCAGAACCAGGATGGAGGCTGGCCTGGTGGACCAAGATGACTTGGAT
 140 ArgIleGluArgAlaArgThrArgMetGluAlaGlyLeuValAspGlnAspAspLeuAsp
 480 GAAGAACACCAGGGCCTGGCCCCTGCTGAACTATTTGCCAGGCTCAAGGCATCCATGCCA
 160 GluGluHisGlnGlyLeuAlaProAlaGluLeuPheAlaArgLeuLysAlaSerMetPro
 540 GATGGTGAGGACCTAGTGGTGACTCATGGGGATGCCTGCCTTCCCAACATCATGGTTGAA
 180 AspGlyGluAspLeuValValThrHisGlyAspAlaCysLeuProAsnIleMetValGlu
 600 AATGGAAGGTTCTCTGGCTTCATAGACTGTGGCAGGCTGGGAGTGGCTGACAGGTACCAG
 200 AsnGlyArgPheSerGlyPheIleAspCysGlyArgLeuGlyValAlaAspArgTyrGln
 660 GACATTGCCCTAGCAACCAGGGACATAGCAGAAGAGCTAGGGGGAGAGTGGGCAGACAGG
 220 AspIleAlaLeuAlaThrArgAspIleAlaGluGluLeuGlyGlyGluTrpAlaAspArg
 720 TTCCTAGTGCTCTATGGCATTGCAGCCCCTGACTCCCAGAGAATTGCCTTCTACAGACTT
 240 PheLeuValLeuTyrGlyIleAlaAlaProAspSerGlnArgIleAlaPheTyrArgLeu
 NheI
 780 CTTGATGAGTTCTTCTAAAGCTAGC SEQ ID NO: 316 (positions 3 to 797)
 260 LeuAspGluPhePhe*** SEQ ID NO: 317

Figure 18

SEQ ID NO: 318 CATTACCGGTAGGCACATCATGATTGAACAAGATGGCCTA
 SEQ ID NO: 319 CATGCAGGTTCTCCAGCTGCCTGGGTTGAGAGACTGTTTG
 SEQ ID NO: 320 GCTATGACTGGGCACAGCAGACCATTGGTTGCTCTGATGC
 SEQ ID NO: 321 AGCAGTTTTTCAGACTTTCAGCCCAAGGCAGGCCAGTCCTT
 SEQ ID NO: 322 TTTGTAAAGACAGACCTCAGTGGGGCTCTCAATGAGCTEC
 SEQ ID NO: 323 AGGATGAGGCTGCCAGACTCTCCTGGTTGGCAACAACTGG
 SEQ ID NO: 324 GGTCCCCTGTGCAGCTGTCCTTGATGTGGTCAAGAAGCT
 SEQ ID NO: 325 GGAAGGGACTGGCTCCTACTAGGTGAGGTGCCTGGGCAGG
 SEQ ID NO: 326 ACCTCCTTTCCTCTCACCTAGCTCCAGCTGAGAAAGTGTC
 SEQ ID NO: 327 AATCATGGCTGATGCCATGAGAAGACTCCACACCCCTTGAC
 SEQ ID NO: 328 CCAGCCACCTGCCCCCTTGACCACCAGGCCAAGCACAGGA
 SEQ ID NO: 329 TAGAGAGGGCCAGAACCAGGATGGAGGCTGGCCTGGTGGA
 SEQ ID NO: 330 CCAAGATGACTTGGATGAAGAACACCAGGGCCTGGCCCCCT
 SEQ ID NO: 331 GCTGAACATTTTGCCAGGCTCAAGGCATCCATGCCAGATG
 SEQ ID NO: 332 GTGAGGACCTAGTGGTGACTCATGGGGATGCCTGCCTTCC
 SEQ ID NO: 333 CAACATCATGGTTGAAAAATGGAAGGTTCTCTGGCTTCATA
 SEQ ID NO: 334 GACTGTGGCAGGCTGGGAGTGGCTGACAGGTACCAGGACA
 SEQ ID NO: 335 TTGCCCTAGCAACCAGGGACATAGCAGAAGAGCTAGGGGG
 SEQ ID NO: 336 AGAGTGGGCAGACAGGTTCTAGTGCTCTATGGCATTGCA
 SEQ ID NO: 337 GCCCCTGACTCCCAGAGAATTGCCTTCTACAGACTTCTTG
 SEQ ID NO: 338 ATGAGTTCTTCTAAAGCTAGCTGATCCTGATAGCTGTTCC
 SEQ ID NO: 339 CGAACAGCTATCAGGATCAG
 SEQ ID NO: 340 CTAGCTTTAGAAGAACTCATCAAGAAGTCTGTAGAAGGCA
 SEQ ID NO: 341 ATTCTCTGGGAGTCAGGGGCTGCAATGCCATAGAGCACTA
 SEQ ID NO: 342 GGAACCTGTCTGCCCCACTCTCCCCCTAGCTCTTCTGCTAT
 SEQ ID NO: 343 GTCCCTGGTTGCTAGGGCAATGTCCTGGTACCTGTGAGCC
 SEQ ID NO: 344 ACTCCCAGCCTGCCACAGTCTATGAAGCCAGAGAACCCTTC
 SEQ ID NO: 345 CATTTTCAACCATGATGTTGGGAAGGCAGGCATCCCCATG
 SEQ ID NO: 346 AGTCACCACTAGGTCCTCACCATCTGGCATGGATGCCTTG
 SEQ ID NO: 347 AGCCTGGCAAATAGTTTACAGCAGGGGCCAGGCCCTGGTGTT
 SEQ ID NO: 348 CTTTCATCCAAGTCATCTTGGTCCACCAGGCCAGCCTCCAT
 SEQ ID NO: 349 CCTGGTTCTGGCCCTCTCTATCCTGTGCTTGGCCTGGTGG
 SEQ ID NO: 350 TCAAAGGGGCGAGGTGGCTGGGTCAAGGGTGTGGAGTCTTC
 SEQ ID NO: 351 TCATGGCATCAGCCATGATTGACACTTTCTCAGCTGGAGC
 SEQ ID NO: 352 TAGGTGAGAGGAAAGGAGGTCCTGCCCAGGCACCTCACCT
 SEQ ID NO: 353 AGTAGGAGCCAGTCCCTTCCAGCTTCTGTGACCACATCAA
 SEQ ID NO: 354 GGACAGCTGCACAGGGGACCCAGTTGTTGCCAACCAGGA
 SEQ ID NO: 355 GAGTCTGGCAGCCTCATCCTGGAGCTCATTGAGAGCCCCA
 SEQ ID NO: 356 CTGAGGTCTGTCTTTACAAAAAGGACTGGCCTGCCTTGGG
 SEQ ID NO: 357 CTGAAAGTCTGAAAAGTCTGTCATCAGAGCAACCAATGGT
 SEQ ID NO: 358 CTGCTGTGCCAGTCATAGCCAAACAGTCTCTCAACCCAG
 SEQ ID NO: 359 GCAGCTGGAGAACCTGCATGTAGGCCATCTTGTTCATCA
 SEQ ID NO: 360 TGATGTGCCTACCGGTAATG

FIGURE 19

SEQUENCE LISTING

<110> CAYLA

<120> SYNTHETIC GENES AND BACTERIAL PLASMIDS DEVOID OF CpG

<130> FP - D. 0200862

<140> PCT/FR02/00862 <141> 2002-03-11

<150> FR01/03274 <151> 2001-03-09

<160> 360

<170> PatentIn version 3.1

<210> 1 <211> 396 <212> DNA <213> Artificial Sequence

<220> <223> CpG-Free Sh ble

<220> <221> CDS <222> (3)..(374) <223>

<400> 1 cc atg gcc aag ttg acc agt gct gtc cca gtg ctc aca gcc agg gat 47

Met Ala Lys Leu Thr Ser Ala Val Pro Val Leu Thr Ala Arg Asp

1 5 10 15

gtg gct gga gct gtt gag ttc tgg act gac agg ttg ggg ttc tcc aga 95 Val Ala Gly Ala Val Glu Phe Trp Thr Asp Arg
Leu Gly Phe Ser Arg

20 25 30

gat ttt gtg gag gat gac ttt gca ggt gtg gtc aga gat gat gtc acc 143 Asp Phe Val Glu Asp Asp Phe Ala Gly Val Val
Arg Asp Asp Val Thr

35 40 45

ctg ttc atc tca gca gtc cag gac cag gtg gtg cct gac aac acc ctg 191 Leu Phe Ile Ser Ala Val Gln Asp Gln Val Val
Pro Asp Asn Thr Leu

50 55 60

gct tgg gtg tgg gtg aga gga ctg gat gag ctg tat gct gag tgg agt 239 Ala Trp Val Trp Val Arg Gly Leu Asp Glu Leu
Tyr Ala Glu Trp Ser

65 70 75

gag gtg gtc tcc acc aac ttc agg gat gcc agt ggc cct gcc atg aca 287 Glu Val Val Ser Thr Asn Phe Arg Asp Ala Ser
Gly Pro Ala Met Thr 80 85 90 95

gag att gga gag cag ccc tgg ggg aga gag ttt gcc ctg aga gac cca 335 Glu Ile Gly Glu Gln Pro Trp Gly Arg Glu Phe
Ala Leu Arg Asp Pro

100 105 110

gca ggc aac tgt gtg cac ttt gtg gca gag gag cag gac tgaggataag 384 Ala Gly Asn Cys Val His Phe Val Ala Glu Glu
Gln Asp

115 120

aattcagcta gc 396

<210> 2 <211> 124 <212> PRT <213> Artificial Sequence

<220> <223> CpG-Free Sh ble

<400> 2

Met Ala Lys Leu Thr Ser Ala Val Pro Val Leu Thr Ala Arg Asp Val 1

5

10

15

Ala Gly Ala Val Glu Phe Trp Thr Asp Arg Leu Gly Phe Ser Arg Asp
20 25 30

Phe Val Glu Asp Asp Phe Ala Gly Val Val Arg Asp Asp Val Thr Leu
35 40 45

Phe Ile Ser Ala Val Gln Asp Gln Val Val Pro Asp Asn Thr Leu Ala
50 55 60

Trp Val Trp Val Arg Gly Leu Asp Glu Leu Tyr Ala Glu Trp Ser Glu 65 70 75 80

Val Val Ser Thr Asn Phe Arg Asp Ala Ser Gly Pro Ala Met Thr Glu
85 90 95

Ile Gly Glu Gln Pro Trp Gly Arg Glu Phe Ala Leu Arg Asp Pro Ala
100 105 110

Gly Asn Cys Val His Phe Val Ala Glu Glu Gln Asp
115 120

<210> 3 <211> 1040 <212> DNA <213> Artificial Sequence

<220> <223> CpG-Free Hph

<220> <221> CDS <222> (3)..(1025) <223>

<400> 3 tc atg aag aaa cct gaa ctg aca gca act tct gtt gag aag ttt etc 47
Met Lys Lys Pro Glu Leu Thr Ala Thr Ser Val Glu Lys Phe Leu
1 5 10 15

att gaa aaa ttt gat tct gtt tct gat ctc atg cag ctg tct gaa ggt 95 Ile Glu Lys Phe Asp Ser Val Ser Asp Leu Met Gln
Leu Ser Glu Gly
20 25 30

gaa gaa agc aga gcc ttt tct ttt gat gtt gga gga aga ggt tat gtt 143 Glu Glu Ser Arg Ala Phe Ser Phe Asp Val Gly
Gly Arg Gly Tyr Val
35 40 45

ctg agg gtc aat tct tgt gct gat ggt ttt tac aaa gac aga tat gtt 191 Leu Arg Val Asn Ser Cys Ala Asp Gly Phe Tyr Lys
Asp Arg Tyr Val
50 55 60

tac aga cac ttt gcc tct gct gct ctg cca att cca gaa gtt ctg gac 239 Tyr Arg His Phe Ala Ser Ala Ala Leu Pro Ile Pro
Glu Val Leu Asp
65 70 75

att gga gaa ttt tct gaa tct ctc acc tac tgc atc agc aga aga gca 287 Ile Gly Glu Phe Ser Glu Ser Leu Thr Tyr Cys Ile
Ser Arg Arg Ala 80 85 90 95

caa gga gtc act ctc cag gat ctc cct gaa act gag ctg cca gct gtt 335 Gln Gly Val Thr Leu Gln Asp Leu Pro Glu Thr
Glu Leu Pro Ala Val
100 105 110

ctg caa cct gtt gct gaa gca atg gat gcc att gca gca gct gat ctg 383 Leu Gln Pro Val Ala Glu Ala Met Asp Ala Ile
Ala Ala Ala Asp Leu
115 120 125

agc caa acc tct gga ttt ggt cct ttt ggt ccc caa ggc att ggt cag 431 Ser Gln Thr Ser Glv Phe Glv Pro Phe Glv Pro Gln

Gly Ile Gly Gln
130 135 140

tac acc act tgg agg gat ttc att tgt gcc att gct gat cct cat gtc 479 Tyr Thr Thr Trp Arg Asp Phe Ile Cys Ala Ile Ala
Asp Pro His Val
145 150 155

tat cac tgg cag act gtg atg gat gac aca gtt tct gct tct gtt gct 527 Tyr His Trp Gln Thr Val Met Asp Asp Thr Val Ser
Ala Ser Val Ala 160 165 170 175

cag gca ctg gat gaa ctc atg ctg tgg gca gaa gat tgt cct gaa gtc 575 Gln Ala Leu Asp Glu Leu Met Leu Trp Ala Glu
Asp Cys Pro Glu Val
180 185 190

aga cac ctg gtc cat gct gat ttt gga agc aac aat gtt ctg aca gac 623 Arg His Leu Val His Ala Asp Phe Gly Ser Asn
Asn Val Leu Thr Asp
195 200 205

aat ggc aga atc act gca gtc att gac tgg tct gaa gcc atg ttt gga 671 Asn Gly Arg Ile Thr Ala Val Ile Asp Trp Ser Glu
Ala Met Phe Gly
210 215 220

gat tct caa tat gag gtt gcc aac att ttt ttt tgg aga cct tgg ctg 719 Asp Ser Gln Tyr Glu Val Ala Asn Ile Phe Phe Trp
Arg Pro Trp Leu
225 230 235

gct tgc atg gaa caa caa aca aga tat ttt gaa aga aga cac cca gaa 767 Ala Cys Met Glu Gln Gln Thr Arg Tyr Phe Glu
Arg Arg His Pro Glu 240 245 250 255

ctg gct ggt tcc ccc aga ctg aga gcc tac atg ctc aga att ggc ctg 815 Leu Ala Gly Ser Pro Arg Leu Arg Ala Tyr Met
Leu Arg Ile Gly Leu
260 265 270

gac caa ctg tat caa tct ctg gtt gat gga aac ttt gat gat gct gct 863 Asp Gln Leu Tyr Gln Ser Leu Val Asp Gly Asn
Phe Asp Asp Ala Ala
275 280 285

tgg gca caa gga aga tgt gat gcc att gtg agg tct ggt gct gga act 911 Trp Ala Gln Gly Arg Cys Asp Ala Ile Val Arg
Ser Gly Ala Gly Thr
290 295 300

gtt gga aga act caa att gca aga agg tct gct gct gtt tgg act gat 959 Val Gly Arg Thr Gln Ile Ala Arg Arg Ser Ala Ala
Val Trp Thr Asp
305 310 315

gga tgt gtt gaa gtt ctg gct gac tct gga aac agg aga ccc tcc aca 1007 Gly Cys Val Glu Val Leu Ala Asp Ser Gly Asn
Arg Arg Pro Ser Thr 320 325 330 335

aga ccc aga gcc aag gaa tgaatattag ctage 1040 Arg Pro Arg Ala Lys Glu
340

<210> 4 <211> 341 <212> PRT <213> Artificial Sequence

<220> <223> CpG-Free Hph

<400> 4

Met Lys Lys Pro Glu Leu Thr Ala Thr Ser Val Glu Lys Phe Leu Ile 1 5 10 15

Glu Lys Phe Asp Ser Val Ser Asp Leu Met Gln Leu Ser Glu Gly Glu
20 25 30

Glu Ser Arg Ala Phe Ser Phe Asp Val Glv Glv Arg Glv Tyr Val Leu

35	40	45			
Arg Val Asn Ser Cys Ala Asp Gly Phe Tyr Lys Asp Arg Tyr Val Tyr					
50	55	60			
Arg His Phe Ala Ser Ala Ala Leu Pro Ile Pro Glu Val Leu Asp Ile 65					
70	75	80			
Gly Glu Phe Ser Glu Ser Leu Thr Tyr Cys Ile Ser Arg Arg Ala Gln					
85	90	95			
Gly Val Thr Leu Gln Asp Leu Pro Glu Thr Glu Leu Pro Ala Val Leu					
100	105	110			
Gln Pro Val Ala Glu Ala Met Asp Ala Ile Ala Ala Ala Asp Leu Ser					
115	120	125			
Gln Thr Ser Gly Phe Gly Pro Phe Gly Pro Gln Gly Ile Gly Gln Tyr					
130	135	140			
Thr Thr Trp Arg Asp Phe Ile Cys Ala Ile Ala Asp Pro His Val Tyr 145					
150	155	160			
His Trp Gln Thr Val Met Asp Asp Thr Val Ser Ala Ser Val Ala Gln					
165	170	175			
Ala Leu Asp Glu Leu Met Leu Trp Ala Glu Asp Cys Pro Glu Val Arg					
180	185	190			
His Leu Val His Ala Asp Phe Gly Ser Asn Asn Val Leu Thr Asp Asn					
195	200	205			
Gly Arg Ile Thr Ala Val Ile Asp Trp Ser Glu Ala Met Phe Gly Asp					
210	215	220			
Ser Gln Tyr Glu Val Ala Asn Ile Phe Phe Trp Arg Pro Trp Leu Ala 225					
230	235	240			
Cys Met Glu Gln Gln Thr Arg Tyr Phe Glu Arg Arg His Pro Glu Leu					
245	250	255			
Ala Gly Ser Pro Arg Leu Arg Ala Tyr Met Leu Arg Ile Gly Leu Asp					
260	265	270			
Gln Leu Tyr Gln Ser Leu Val Asp Gly Asn Phe Asp Asp Ala Ala Trp					
275	280	285			
Ala Gln Gly Arg Cys Asp Ala Ile Val Arg Ser Gly Ala Gly Thr Val					
290	295	300			
Gly Arg Thr Gln Ile Ala Arg Arg Ser Ala Ala Val Trp Thr Asp Gly 305					
310	315	320			

Cys Val Glu Val Leu Ala Asp Ser Gly Asn Arg Arg Pro Ser Thr Arg
325 330 335

Pro Arg Ala Lys Glu
340

<210> 5 <211> 442 <212> DNA <213> Artificial Sequence

<220> <223> CpG-Free Bsr

<220> <221> CDS <222> (3)..(422) <223>

<400> 5 tc atg aag acc ttc aac atc tct cag cag gat ctg gag ctg gtg gag 47
Met Lys Thr Phe Asn Ile Ser Gln Gln Asp Leu Glu Leu Val Glu
1 5 10 15

gtc gcc act gag aag atc acc atg ctc tat gag gac aac aag cac cat 95 Val Ala Thr Glu Lys Ile Thr Met Leu Tyr Glu
Asp Asn Lys His His
20 25 30

gtc ggg gcg gcc atc agg acc aag act ggg gag atc atc tct gct gtc 143 Val Gly Ala Ala Ile Arg Thr Lys Thr Gly Glu
Ile Ile Ser Ala Val
35 40 45

cac att gag gcc tac att ggc agg gtc act gtc tgt gct gaa gcc att 191 His Ile Glu Ala Tyr Ile Gly Arg Val Thr Val Cys
Ala Glu Ala Ile
50 55 60

gcc att ggg tct gct gtg agc aac ggg cag aag gac ttt gac acc att 239 Ala Ile Gly Ser Ala Val Ser Asn Gly Gln Lys
Asp Phe Asp Thr Ile
65 70 75

gtg gct gtc agg cac ccc tac tct gat gag gtg gac aga tcc atc agg 287 Val Ala Val Arg His Pro Tyr Ser Asp Glu Val
Asp Arg Ser Ile Arg 80 85 90 95

gtg gtc agc ccc tgt ggc atg tgc aga gag ctc atc tct gac tat gct 335 Val Val Ser Pro Cys Gly Met Cys Arg Glu Leu
Ile Ser Asp Tyr Ala
100 105 110

cct gac tgc ttt gtg ctc att gag atg aat ggc aag ctg gtc aaa acc 383 Pro Asp Cys Phe Val Leu Ile Glu Met Asn Gly
Lys Leu Val Lys Thr
115 120 125

acc att gag gaa ctc atc ccc ctc aag tac acc agg aac taacctgaa 432 Thr Ile Glu Glu Leu Ile Pro Leu Lys Tyr Thr
Arg Asn
130 135 140

ttcagctagc 442

<210> 6 <211> 140 <212> PRT <213> Artificial Sequence

<220> <223> CpG-Free Bsr

<400> 6

Met Lys Thr Phe Asn Ile Ser Gln Gln Asp Leu Glu Leu Val Glu Val 1 5 10 15

Ala Thr Glu Lys Ile Thr Met Leu Tyr Glu Asp Asn Lys His His Val
20 25 30

Gly Ala Ala Ile Arg Thr Lys Thr Gly Glu Ile Ile Ser Ala Val His
35 40 45

Ile Glu Ala Tyr Ile Gly Arg Val Thr Val Cys Ala Glu Ala Ile Ala
50 55 60

Ile Gly Ser Ala Val Ser Asn Gly Gln Lys Asp Phe Asp Thr Ile Val 65 70 75 80

Ala Val Arg His Pro Tyr Ser Asp Glu Val Asp Arg Ser Ile Arg Val
85 90 95

Val Ser Pro Cys Gly Met Cys Arg Glu Leu Ile Ser Asp Tyr Ala Pro
100 105 110

Asp Cys Phe Val Leu Ile Glu Met Asn Gly Lys Leu Val Lys Thr Thr
115 120 125

Ile Glu Glu Leu Ile Pro Leu Lys Tyr Thr Arg Asn
130 135 140

<210> 7 <211> 614 <212> DNA <213> Artificial Sequence

<220> <223> CpG-Free Pac

<220> <221> CDS <222> (3)..(599) <223>

<400> 7 tc atg act gag tac aaa ccc aca gtg agg ctg gca acc aga gat gat 47

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp

1 5 10 15

gtt cca aga gct gtg aga aca ctg gct gct gct ttt gca gac tac cct 95 Val Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe
Ala Asp Tyr Pro

20 25 30

gca aca agg cac aca gtt gac cct gac agg cac att gag agg gtg aca 143 Ala Thr Arg His Thr Val Asp Pro Asp Arg His
Ile Glu Arg Val Thr

35 40 45

gaa ctg caa gaa ctc ttc ctc acc aga gtg gga ctg gac att gga aaa 191 Glu Leu Gln Glu Leu Phe Leu Thr Arg Val Gly
Leu Asp Ile Gly Lys

50 55 60

gtt tgg gtt gca gat gat gga gct gct gtt gca gtt tgg aca aca cct 239 Val Trp Val Ala Asp Asp Gly Ala Ala Val Ala Val
Trp Thr Thr Pro

65 70 75

gag tct gtt gaa gct ggt gct gtt ttt gct gaa att gga cca aga atg 287 Glu Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile
Gly Pro Arg Met 80 85 90 95

gct gag ctc tct gga agc agg ctg gca gca caa caa caa atg gaa ggt 335 Ala Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln
Gln Gln Met Glu Gly

100 105 110

ctg ctg gca cca cac agg cca aaa gag cca gct tgg ttt ctg gca act 383 Leu Leu Ala Pro His Arg Pro Lys Glu Pro Ala
Trp Phe Leu Ala Thr

115 120 125

gtt gga gtg agc cct gac cac cag gga aag ggt ctg gga tct gct gtt 431 Val Gly Val Ser Pro Asp His Gln Gly Lys Gly
Leu Gly Ser Ala Val
130 135 140

gtt ctg cct gga gtt gaa gct gct gaa agg gct gga gtt cct gcc ttt 479 Val Leu Pro Gly Val Glu Ala Ala Glu Arg Ala
Gly Val Pro Ala Phe
145 150 155

ctg gaa act tct gct ccc aga aac ctg cct ttt tat gaa aga ctg gga 527 Leu Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe
Tyr Glu Arg Leu Gly 160 165 170 175

ttc act gtg aca gct gat gtt gag gtt cca gaa ggc cca aga act tgg 575 Phe Thr Val Thr Ala Asp Val Glu Val Pro Glu
Gly Pro Arg Thr Trp
180 185 190

tgc atg aca agg aag cct gga gct taaacctgag ctage 614 Cys Met Thr Arg Lys Pro Gly Ala
195

<210> 8 <211> 199 <212> PRT <213> Artificial Sequence

<220> <223> CpG-Free Pac

<400> 8

Met Thr Glu Tyr Lys Pro Thr Val Arg Leu Ala Thr Arg Asp Asp Val 1 5 10 15

Pro Arg Ala Val Arg Thr Leu Ala Ala Ala Phe Ala Asp Tyr Pro Ala
20 25 30

Thr Arg His Thr Val Asp Pro Asp Arg His Ile Glu Arg Val Thr Glu
35 40 45

Leu Gln Glu Leu Phe Leu Thr Arg Val Gly Leu Asp Ile Gly Lys Val
50 55 60

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

Ser Val Glu Ala Gly Ala Val Phe Ala Glu Ile Gly Pro Arg Met Ala
85 90 95

Glu Leu Ser Gly Ser Arg Leu Ala Ala Gln Gln Gln Met Glu Gly Leu
100 105 110

Leu Ala Pro His Arg Pro Lys Glu Pro Ala Trp Phe Leu Ala Thr Val
115 120 125

Gly Val Ser Pro Asp His Gln Gly Lys Gly Leu Gly Ser Ala Val Val
130 135 140

Leu Pro Gly Val Glu Ala Ala Glu Arg Ala Gly Val Pro Ala Phe Leu 145 150 155 160

Glu Thr Ser Ala Pro Arg Asn Leu Pro Phe Tyr Glu Arg Leu Gly Phe
165 170 175

Thr Val Thr Ala Asp Val Glu Val Pro Glu Gly Pro Arg Thr Trp Cys
180 185 190

Met Thr Arg Lys Pro Gly Ala
195

<210> 9 <211> 3071 <212> DNA <213> Artificial Sequence

<220> <223> CpG-Free LacZ

<220> <221> CDS <222> (3)..(3056) <223>

<400> 9 cc atg gac cct gtt gtg ctg caa agg aga gac tgg gag aac cct gga 47
Met Asp Pro Val Val Leu Gln Arg Arg Asp Trp Glu Asn Pro Gly
1 5 10 15

gtg acc cag ctc aac aga ctg gct gcc cac cct ccc ttt gcc tct tgg 95 Val Thr Gln Leu Asn Arg Leu Ala Ala His Pro
Pro Phe Ala Ser Trp
20 25 30

agg aac tct gag gaa gcc agg aca gac agg ccc agc cag cag ctc agg 143 Arg Asn Ser Glu Glu Ala Arg Thr Asp Arg
Pro Ser Gln Gln Leu Arg
35 40 45

tct ctc aat gga gag tgg agg ttt gcc tgg ttc cct gcc cct gaa gct 191 Ser Leu Asn Gly Glu Trp Arg Phe Ala Trp Phe
Pro Ala Pro Glu Ala
50 55 60

gtg cct gag tct tgg ctg gag tgt gac ctc cca gag gct gac act gtt 239 Val Pro Glu Ser Trp Leu Glu Cys Asp Leu Pro
Glu Ala Asp Thr Val
65 70 75

gtg gtg ccc agc aac tgg cag atg cat ggc tat gat gcc ccc atc tac 287 Val Val Pro Ser Asn Trp Gln Met His Gly Tyr
Asp Ala Pro Ile Tyr 80 85 90 95

acc aat gtc acc tac ccc atc act gtg aac ccc cct ttt gtg ccc act 335 Thr Asn Val Thr Tyr Pro Ile Thr Val Asn Pro Pro
Phe Val Pro Thr
100 105 110

gag aac ccc act ggc tgc tac agc ctg acc ttc aat gtt gat gag agc 383 Glu Asn Pro Thr Gly Cys Tyr Ser Leu Thr Phe
Asn Val Asp Glu Ser
115 120 125

tgg ctg caa gaa ggc cag acc agg atc atc ttt gat gga gtc aac tct 431 Trp Leu Gln Glu Gly Gln Thr Arg Ile Ile Phe
Asp Gly Val Asn Ser
130 135 140

gcc ttc cac ctc tgg tgc aat ggc agg tgg gtt ggc tat ggc caa gac 479 Ala Phe His Leu Trp Cys Asn Gly Arg Trp Val
Gly Tyr Gly Gln Asp
145 150 155

agc agg ctg ccc tct gag ttt gac ctc tct gcc ttc ctc aga gct gga 527 Ser Arg Leu Pro Ser Glu Phe Asp Leu Ser Ala
Phe Leu Arg Ala Gly 160 165 170 175

gag aac agg ctg gct gtc atg gtg ctc agg tgg tct gat ggc agc tac 575 Glu Asn Arg Leu Ala Val Met Val Leu Arg Trp
Ser Asp Gly Ser Tyr
180 185 190

ctg gaa gac caa gac atg tgg agg atg tct ggc atc ttc agg gat gtg 623 Leu Glu Asp Gln Asp Met Trp Arg Met Ser Gly
Ile Phe Arg Asp Val
195 200 205

agc ctg ctg cac aag ccc acc acc cag att tct gac ttc cat gtt gcc 671 Ser Leu Leu His Lys Pro Thr Thr Gln Ile Ser Asp

Phe His Val Ala				
210	215	220		
acc agg ttc aat gat gac ttc agc aga gct gtg ctg gag gct gag gtg			719 Thr Arg Phe Asn Asp Asp Phe Ser Arg Ala Val	
Leu Glu Ala Glu Val				
225	230	235		
cag atg tgt gga gaa ctc aga gac tac ctg aga gtc aca gtg agc ctc			767 Gln Met Cys Gly Glu Leu Arg Asp Tyr Leu	
Arg Val Thr Val Ser Leu 240	245	250	255	
tgg caa ggt gag acc cag gtg gcc tct ggc aca gcc ccc ttt gga gga			815 Trp Gln Gly Glu Thr Gln Val Ala Ser Gly Thr	
Ala Pro Phe Gly Gly				
260	265	270		
gag atc att gat gag aga gga ggc tat gct gac aga gtc acc ctg agg			863 Glu Ile Ile Asp Glu Arg Gly Gly Tyr Ala Asp	
Arg Val Thr Leu Arg				
275	280	285		
ctc aat gtg gag aac ccc aag ctg tgg tct gct gag atc ccc aac ctc			911 Leu Asn Val Glu Asn Pro Lys Leu Trp Ser Ala	
Glu Ile Pro Asn Leu				
290	295	300		
tac agg gct gtt gtg gag ctg cac act gct gat ggc acc ctg att gaa			959 Tyr Arg Ala Val Val Glu Leu His Thr Ala Asp	
Gly Thr Leu Ile Glu				
305	310	315		
gct gaa gcc tgt gat gtt gga ttc aga gaa gtc agg att gag aat ggc			1007 Ala Glu Ala Cys Asp Val Gly Phe Arg Glu Val	
Arg Ile Glu Asn Gly 320	325	330	335	
ctg ctg ctg ctc aat ggc aag cct ctg ctc atc agg gga gtc aac agg			1055 Leu Leu Leu Leu Asn Gly Lys Pro Leu Leu Ile	
Arg Gly Val Asn Arg				
340	345	350		
cat gag cac cac cct ctg cat gga caa gtg atg gat gaa cag aca atg			1103 His Glu His His Pro Leu His Gly Gln Val Met	
Asp Glu Gln Thr Met				
355	360	365		
gtg caa gat atc ctg cta atg aag cag aac aac ttc aat gct gtc agg			1151 Val Gln Asp Ile Leu Leu Met Lys Gln Asn Asn	
Phe Asn Ala Val Arg				
370	375	380		
tgc tct cac tac ccc aac cac cct ctc tgg tac acc ctg tgt gac agg			1199 Cys Ser His Tyr Pro Asn His Pro Leu Trp Tyr	
Thr Leu Cys Asp Arg				
385	390	395		
tat ggc ctg tat gtt gtt gat gaa gcc aac att gag aca cat ggc atg			1247 Tyr Gly Leu Tyr Val Val Asp Glu Ala Asn Ile	
Glu Thr His Gly Met 400	405	410	415	
gtg ccc atg aac agg ctc aca gat gac ccc agg tgg ctg cct gcc atg			1295 Val Pro Met Asn Arg Leu Thr Asp Asp Pro	
Arg Trp Leu Pro Ala Met				
420	425	430		
tct gag aga gtg acc agg atg gtg cag aga gac agg aac cac ccc tct			1343 Ser Glu Arg Val Thr Arg Met Val Gln Arg	
Asp Arg Asn His Pro Ser				
435	440	445		
gtg atc atc tgg tct ctg ggc aat gag tct gga cat gga gcc aac cat			1391 Val Ile Ile Trp Ser Leu Gly Asn Glu Ser Gly His	
Gly Ala Asn His				
450	455	460		
gat gct ctc tac agg tgg atc aag tct gtt gac ccc agc aga cct gtg			1439 Asp Ala Leu Tyr Arg Trp Ile Lys Ser Val Asp	
Pro Ser Arg Pro Val				
465	470	475		
cag tat gaa gga ggt gga gca gac acc aca gcc aca gac atc atc tgc			1487 Gln Tyr Glu Gly Gly Gly Ala Asp Thr Thr	
Ala Thr Asp Ile Ile Cys 480	485	490	495	

ccc atg tat gcc agg gtt gat gag gac cag ccc ttc cct gct gtg ccc Phe Pro Ala Val Pro 500 505 510	1535 Pro Met Tyr Ala Arg Val Asp Glu Asp Gln Pro
aag tgg agc atc aag aag tgg ctc tct ctg cct gga gag acc aga cct Gly Glu Thr Arg Pro 515 520 525	1583 Lys Trp Ser Ile Lys Lys Trp Leu Ser Leu Pro
ctg atc ctg tgt gaa tat gca cat gca atg ggc aac tct ctg gga ggc Asn Ser Leu Gly Gly 530 535 540	1631 Leu Ile Leu Cys Glu Tyr Ala His Ala Met Gly
ttt gcc aag tac tgg caa gcc ttc aga cag tac ccc agg ctg caa gga Pro Arg Leu Gln Gly 545 550 555	1679 Phe Ala Lys Tyr Trp Gln Ala Phe Arg Gln Tyr
gga ttt gtg tgg gac tgg gtg gac caa tct ctc atc aag tat gat gag Ile Lys Tyr Asp Glu 560 565 570	1727 Gly Phe Val Trp Asp Trp Val Asp Gln Ser Leu 575
aat ggc aac ccc tgg tct gcc tat gga gga gac ttt ggt gac acc ccc Phe Gly Asp Thr Pro 580 585 590	1775 Asn Gly Asn Pro Trp Ser Ala Tyr Gly Gly Asp
aat gac agg cag ttc tgc atg aat ggc ctg gtc ttt gca gac agg acc Val Phe Ala Asp Arg Thr 595 600 605	1823 Asn Asp Arg Gln Phe Cys Met Asn Gly Leu
cct cac cct gcc ctc aca gag gcc aag cac cag caa cag ttc ttc cag Gln Gln Phe Phe Gln 610 615 620	1871 Pro His Pro Ala Leu Thr Glu Ala Lys His Gln
ttc agg ctg tct gga cag acc att gag gtg aca tct gag tac ctc ttc Glu Tyr Leu Phe 625 630 635	1919 Phe Arg Leu Ser Gly Gln Thr Ile Glu Val Thr Ser
agg cac tct gac aat gag ctc ctg cac tgg atg gtg gcc ctg gat ggc Val Ala Leu Asp Gly 640 645 650	1967 Arg His Ser Asp Asn Glu Leu Leu His Trp Met 655
aag cct ctg gct tct ggt gag gtg cct ctg gat gtg gcc cct caa gga Val Ala Pro Gln Gly 660 665 670	2015 Lys Pro Leu Ala Ser Gly Glu Val Pro Leu Asp
aag cag ctg att gaa ctg cct gag ctg cct cag cca gag tct gct gga Pro Glu Ser Ala Gly 675 680 685	2063 Lys Gln Leu Ile Glu Leu Pro Glu Leu Pro Gln
caa ctg tgg cta aca gtg agg gtg gtt cag ccc aat gca aca gct tgg Asn Ala Thr Ala Trp 690 695 700	2111 Gln Leu Trp Leu Thr Val Arg Val Val Gln Pro
tct gag gca ggc cac atc tct gca tgg cag cag tgg agg ctg gct gag Trp Arg Leu Ala Glu 705 710 715	2159 Ser Glu Ala Gly His Ile Ser Ala Trp Gln Gln
aac ctc tct gtg acc ctg cct gct gcc tct cat gcc atc cct cac ctg Ala Ile Pro His Leu 720 725 730	2207 Asn Leu Ser Val Thr Leu Pro Ala Ala Ser His 735
aca aca tct gaa atg gac ttc tgc att gag ctg ggc aac aag aga tgg Gly Asn Lys Arg Trp 740 745 750	2255 Thr Thr Ser Glu Met Asp Phe Cys Ile Glu Leu
cag ttc aac agg cag tct ggc ttc ctg tct cag atg tgg att gga gac Met Trp Ile Gly Asp 755 760 765	2303 Gln Phe Asn Arg Gln Ser Gly Phe Leu Ser Gln

aag aag cag ctc ctc acc cct ctc agg gac caa ttc acc agg gct cct Phe Thr Arg Ala Pro 770 775 780	2351 Lys Lys Gln Leu Leu Thr Pro Leu Arg Asp Gln
ctg gac aat gac att gga gtg tct gag gcc acc agg att gac cca aat Arg Ile Asp Pro Asn 785 790 795	2399 Leu Asp Asn Asp Ile Gly Val Ser Glu Ala Thr
gct tgg gtg gag agg tgg aag gct gct gga cac tac cag gct gag gct Tyr Gln Ala Glu Ala 800 805 810 815	2447 Ala Trp Val Glu Arg Trp Lys Ala Ala Gly His
gcc ctg ctc cag tgc aca gca gac acc ctg gct gat gct gtt ctg atc Asp Ala Val Leu Ile 820 825 830	2495 Ala Leu Leu Gln Cys Thr Ala Asp Thr Leu Ala
acc aca gcc cat gct tgg cag cac caa ggc aag acc ctg ttc atc agc Thr Leu Phe Ile Ser 835 840 845	2543 Thr Thr Ala His Ala Trp Gln His Gln Gly Lys
aga aag acc tac agg att gat ggc tct gga cag atg gca atc aca gtg Met Ala Ile Thr Val 850 855 860	2591 Arg Lys Thr Tyr Arg Ile Asp Gly Ser Gly Gln
gat gtg gag gtt gcc tct gac aca cct cac cct gca agg att ggc ctg Ala Arg Ile Gly Leu 865 870 875	2639 Asp Val Glu Val Ala Ser Asp Thr Pro His Pro
aac tgt caa ctg gca cag gtg gct gag agg gtg aac tgg ctg ggc tta Asn Trp Leu Gly Leu 880 885 890 895	2687 Asn Cys Gln Leu Ala Gln Val Ala Glu Arg Val
ggc cct cag gag aac tac cct gac agg ctg aca gct gcc tgc ttt gac Ala Ala Cys Phe Asp 900 905 910	2735 Gly Pro Gln Glu Asn Tyr Pro Asp Arg Leu Thr
agg tgg gac ctg cct ctg tct gac atg tac acc cct tat gtg ttc cct Pro Tyr Val Phe Pro 915 920 925	2783 Arg Trp Asp Leu Pro Leu Ser Asp Met Tyr Thr
tct gag aat ggc ctg agg tgt ggc acc agg gag ctg aac tat ggt cct Leu Asn Tyr Gly Pro 930 935 940	2831 Ser Glu Asn Gly Leu Arg Cys Gly Thr Arg Glu
cac cag tgg agg gga gac ttc cag ttc aac atc tcc agg tac tet cag Ser Arg Tyr Ser Gln 945 950 955	2879 His Gln Trp Arg Gly Asp Phe Gln Phe Asn Ile
caa cag ctc atg gaa acc tet cac agg cac ctg ctc cat gca gag gag Leu Leu His Ala Glu Glu 960 965 970 975	2927 Gln Gln Leu Met Glu Thr Ser His Arg His
gga acc tgg ctg aac att gat ggc ttc cac atg ggc att gga gga gat Gly Ile Gly Gly Asp 980 985 990	2975 Gly Thr Trp Leu Asn Ile Asp Gly Phe His Met
gac tct tgg tct cct tct gtg tct gct gag ttc cag tta tct gct ggc Leu Ser Ala Gly 995 1000 1005	3023 Asp Ser Trp Ser Pro Ser Val Ser Ala Glu Phe Gln
agg tac cac tat cag ctg gtg tgg tgc cag aag taaacctgag ctage 1010 1015	3071 Arg Tyr His Tyr Gln Leu Val Trp Cys Gln Lys

<210> 10 <211> 1018 <212> PRT <213> Artificial Sequence

<220> <223> CpG-Free LacZ

<400> 10

Met Asp Pro Val Val Leu Gln Arg Arg Asp Trp Glu Asn Pro Gly Val 1	5	10	15
Thr Gln Leu Asn Arg Leu Ala Ala His Pro Pro Phe Ala Ser Trp Arg			
20 25 30			
Asn Ser Glu Glu Ala Arg Thr Asp Arg Pro Ser Gln Gln Leu Arg Ser			
35 40 45			
Leu Asn Gly Glu Trp Arg Phe Ala Trp Phe Pro Ala Pro Glu Ala Val			
50 55 60			
Pro Glu Ser Trp Leu Glu Cys Asp Leu Pro Glu Ala Asp Thr Val Val 65	70	75	80
Val Pro Ser Asn Trp Gln Met His Gly Tyr Asp Ala Pro Ile Tyr Thr			
85 90 95			
Asn Val Thr Tyr Pro Ile Thr Val Asn Pro Pro Phe Val Pro Thr Glu			
100 105 110			
Asn Pro Thr Gly Cys Tyr Ser Leu Thr Phe Asn Val Asp Glu Ser Trp			
115 120 125			
Leu Gln Glu Gly Gln Thr Arg Ile Ile Phe Asp Gly Val Asn Ser Ala			
130 135 140			
Phe His Leu Trp Cys Asn Gly Arg Trp Val Gly Tyr Gly Gln Asp Ser 145	150	155	160
Arg Leu Pro Ser Glu Phe Asp Leu Ser Ala Phe Leu Arg Ala Gly Glu			
165 170 175			
Asn Arg Leu Ala Val Met Val Leu Arg Trp Ser Asp Gly Ser Tyr Leu			
180 185 190			
Glu Asp Gln Asp Met Trp Arg Met Ser Gly Ile Phe Arg Asp Val Ser			
195 200 205			
Leu Leu His Lys Pro Thr Thr Gln Ile Ser Asp Phe His Val Ala Thr			
210 215 220			
Arg Phe Asn Asp Asp Phe Ser Arg Ala Val Leu Glu Ala Glu Val Gln 225	230	235	240
Met Cys Gly Glu Leu Arg Asp Tyr Leu Arg Val Thr Val Ser Leu Trp			
245 250 255			
Gln Gly Glu Thr Gln Val Ala Ser Gly Thr Ala Pro Phe Gly Gly Glu			
260 265 270			
Ile Ile Asp Glu Arg Gly Gly Tyr Ala Asp Arg Val Thr Leu Arg Leu			

275	280	285			
Asn Val Glu Asn Pro Lys Leu Trp Ser Ala Glu Ile Pro Asn Leu Tyr					
290	295	300			
Arg Ala Val Val Glu Leu His Thr Ala Asp Gly Thr Leu Ile Glu Ala 305					
			310	315	320
Glu Ala Cys Asp Val Gly Phe Arg Glu Val Arg Ile Glu Asn Gly Leu					
325	330	335			
Leu Leu Leu Asn Gly Lys Pro Leu Leu Ile Arg Gly Val Asn Arg His					
340	345	350			
Glu His His Pro Leu His Gly Gln Val Met Asp Glu Gln Thr Met Val					
355	360	365			
Gln Asp Ile Leu Leu Met Lys Gln Asn Asn Phe Asn Ala Val Arg Cys					
370	375	380			
Ser His Tyr Pro Asn His Pro Leu Trp Tyr Thr Leu Cys Asp Arg Tyr 385					
			390	395	400
Gly Leu Tyr Val Val Asp Glu Ala Asn Ile Glu Thr His Gly Met Val					
405	410	415			
Pro Met Asn Arg Leu Thr Asp Asp Pro Arg Trp Leu Pro Ala Met Ser					
420	425	430			
Glu Arg Val Thr Arg Met Val Gln Arg Asp Arg Asn His Pro Ser Val					
435	440	445			
Ile Ile Trp Ser Leu Gly Asn Glu Ser Gly His Gly Ala Asn His Asp					
450	455	460			
Ala Leu Tyr Arg Trp Ile Lys Ser Val Asp Pro Ser Arg Pro Val Gln 465					
			470	475	480
Tyr Glu Gly Gly Gly Ala Asp Thr Thr Ala Thr Asp Ile Ile Cys Pro					
485	490	495			
Met Tyr Ala Arg Val Asp Glu Asp Gln Pro Phe Pro Ala Val Pro Lys					
500	505	510			
Trp Ser Ile Lys Lys Trp Leu Ser Leu Pro Gly Glu Thr Arg Pro Leu					
515	520	525			
Ile Leu Cys Glu Tyr Ala His Ala Met Gly Asn Ser Leu Gly Gly Phe					
530	535	540			
Ala Lys Tyr Trp Gln Ala Phe Arg Gln Tyr Pro Arg Leu Gln Gly Gly 545					
			550	555	560

Phe Val Trp Asp Trp Val Asp Gln Ser Leu Ile Lys Tyr Asp Glu Asn
565 570 575

Gly Asn Pro Trp Ser Ala Tyr Gly Gly Asp Phe Gly Asp Thr Pro Asn
580 585 590

Asp Arg Gln Phe Cys Met Asn Gly Leu Val Phe Ala Asp Arg Thr Pro
595 600 605

His Pro Ala Leu Thr Glu Ala Lys His Gln Gln Gln Phe Phe Gln Phe
610 615 620

Arg Leu Ser Gly Gln Thr Ile Glu Val Thr Ser Glu Tyr Leu Phe Arg 625 630 635 640

His Ser Asp Asn Glu Leu Leu His Trp Met Val Ala Leu Asp Gly Lys
645 650 655

Pro Leu Ala Ser Gly Glu Val Pro Leu Asp Val Ala Pro Gln Gly Lys
660 665 670

Gln Leu Ile Glu Leu Pro Glu Leu Pro Gln Pro Glu Ser Ala Gly Gln
675 680 685

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

Glu Ala Gly His Ile Ser Ala Trp Gln Gln Trp Arg Leu Ala Glu Asn 705 710 715 720

Leu Ser Val Thr Leu Pro Ala Ala Ser His Ala Ile Pro His Leu Thr
725 730 735

Thr Ser Glu Met Asp Phe Cys Ile Glu Leu Gly Asn Lys Arg Trp Gln
740 745 750

Phe Asn Arg Gln Ser Gly Phe Leu Ser Gln Met Trp Ile Gly Asp Lys
755 760 765

Lys Gln Leu Leu Thr Pro Leu Arg Asp Gln Phe Thr Arg Ala Pro Leu
770 775 780

Asp Asn Asp Ile Gly Val Ser Glu Ala Thr Arg Ile Asp Pro Asn Ala 785 790 795 800

Trp Val Glu Arg Trp Lys Ala Ala Gly His Tyr Gln Ala Glu Ala Ala
805 810 815

Leu Leu Gln Cys Thr Ala Asp Thr Leu Ala Asp Ala Val Leu Ile Thr
820 825 830

Thr Ala His Ala Trp Gln His Gln Gly Lys Thr Leu Phe Ile Ser Arg
835 840 845

Lys Thr Tyr Arg Ile Asp Gly Ser Gly Gln Met Ala Ile Thr Val Asp
850 855 860

Val Glu Val Ala Ser Asp Thr Pro His Pro Ala Arg Ile Gly Leu Asn 865 870 875 880

Cys Gln Leu Ala Gln Val Ala Glu Arg Val Asn Trp Leu Gly Leu Gly
885 890 895

Pro Gln Glu Asn Tyr Pro Asp Arg Leu Thr Ala Ala Cys Phe Asp Arg
900 905 910

Trp Asp Leu Pro Leu Ser Asp Met Tyr Thr Pro Tyr Val Phe Pro Ser
915 920 925

Glu Asn Gly Leu Arg Cys Gly Thr Arg Glu Leu Asn Tyr Gly Pro His
930 935 940

Gln Trp Arg Gly Asp Phe Gln Phe Asn Ile Ser Arg Tyr Ser Gln Gln 945 950 955 960

Gln Leu Met Glu Thr Ser His Arg His Leu Leu His Ala Glu Glu Gly
965 970 975

Thr Trp Leu Asn Ile Asp Gly Phe His Met Gly Ile Gly Gly Asp Asp
980 985 990

Ser Trp Ser Pro Ser Val Ser Ala Glu Phe Gln Leu Ser Ala Gly Arg
995 1000 1005

Tyr His Tyr Gln Leu Val Trp Cys Gln Lys
1010 1015

<210> 11 <211> 66 <212> DNA <213> Artificial Sequence

<220> <223> CpG-Free constitutive promotor

<400> 11 caattaaaca ttgcatagt atatctgc atgataatac aactactat aggaggcca 60
ccatgg 66

<210> 12 <211> 320 <212> DNA <213> Artificial Sequence

<220> <223> CpG-Free replication origin

<400> 12 gcaggactga ggcttaatta aacctaaaa cctttaaag ccttatatat tctttttt 60
cttataaac ttaaacctt agaggctatt taagtgtctg atttatatta attttattgt 120
tcaaactaga gagcttagta catgaacat gagagcttag tacattagcc atgagagctt 180
agtacattag ccatgaggtt ttattcatt aaacatgaga gcttagtaca ttaaactaga 240

gagcttagta catgaaacat gagagcttag tacatactat caacagggtg aactgctgat 300

cttaattaac ctggagactt 320

<210> 13 <211> 273 <212> DNA <213> Artificial Sequence

<220> <223> CpG-Free replication origin

<400> 13 ttaattaacc ttaaaacctt taaaagcctt atatattctt tttttctta taaaacttaa 60

aaccttagag gctatttaag ttgctgattt atattaattt tattgttcaa acatgagagc 120

ttagtacatg aaacatgaga gcttagtaca ttagccatga gagcttagta cattagccat 180

gaggggttag ttcattaac atgagagctt agtacattaa acatgagagc ttagtacata 240

ctatcaacag gttgaactgc tgatcttaat taa 273

<210> 14 <211> 3852 <212> DNA <213> Artificial Sequence

<220> <223> SEQ ID NO:14 = CpG-Free pSh-LacZ plasmid ; SEQ ID NO:15= SEQ ID N
O:14 CDS from 83 to 454 (SEQ ID NO:2) + SEQ ID NO:14 CDS from 483
to 3536 (SEQ ID NO:10)

<220> <221> CDS <222> (83)..(454) <223>

<220> <221> CDS <222> (483)..(3536) <223>

<400> 14 ttaattaaaa cagtagttga caattaaaca ttggcatagt atatctgcat agtataatac 60

aactcactat aggagggccca cc atg gcc aag ttg acc agt gct gtc cca gtc 112

Met Ala Lys Leu Thr Ser Ala Val Pro Val

1 5 10

ctc aca gcc agg gat gtg gct gga gct gtt gag ttc tgg act gac agg 160 Leu Thr Ala Arg Asp Val Ala Gly Ala Val Glu
Phe Trp Thr Asp Arg

15 20 25

ttg ggg ttc tcc aga gat ttt gtg gag gat gac ttt gca ggt gtg gtc 208 Leu Gly Phe Ser Arg Asp Phe Val Glu Asp Asp
Phe Ala Gly Val Val

30 35 40

aga gat gat gtc acc ctg ttc atc tca gca gtc cag gac cag gtg gtg 256 Arg Asp Asp Val Thr Leu Phe Ile Ser Ala Val
Gln Asp Gln Val Val

45 50 55

cct gac aac acc ctg gct tgg gtg tgg gtg aga gga ctg gat gag ctg 304 Pro Asp Asn Thr Leu Ala Trp Val Trp Val Arg
Gly Leu Asp Glu Leu

60 65 70

tat gct gag tgg agt gag gtg gtc tcc acc aac ttc agg gat gcc agt 352 Tyr Ala Glu Trp Ser Glu Val Val Ser Thr Asn
Phe Arg Asp Ala Ser 75 80 85 90

ggc cct gcc atg aca gag att gga gag cag ccc tgg ggg aga gag ttt 400 Gly Pro Ala Met Thr Glu Ile Gly Glu Gln Pro
Trp Gly Arg Glu Phe

95 100 105

gcc ctg aga gac cca gca ggc aac tgt gtg cac ttt gtg gca gag gag 448 Ala Leu Arg Asp Pro Ala Gly Asn Cys Val His
Phe Val Ala Glu Glu

110 115 120

cag gac tgaggataag aattctgagg agaagctc atg gac cct gtt gtg ctg Val Leu	500 Gln Asp	Met Asp Pro Val
125 130		
caa agg aga gac tgg gag aac cct gga gtg acc cag ctc aac aga ctg Thr Gln Leu Asn Arg Leu	548 Gln Arg Arg Asp Trp Glu Asn Pro Gly Val	
135 140 145		
gct gcc cac cct ccc ttt gcc tct tgg agg aac tct gag gaa gcc agg Ser Glu Glu Ala Arg	596 Ala Ala His Pro Pro Phe Ala Ser Trp Arg Asn	
150 155 160		
aca gac agg ccc agc cag cag ctc agg tct ctc aat gga gag tgg agg Leu Asn Gly Glu Trp Arg	644 Thr Asp Arg Pro Ser Gln Gln Leu Arg Ser	
165 170 175		
ttt gcc tgg ttc cct gcc cct gaa gct gtg cct gag tct tgg ctg gag Ser Trp Leu Glu	692 Phe Ala Trp Phe Pro Ala Pro Glu Ala Val Pro Glu	
180 185 190		
tgt gac ctc cca gag gct gac act gtt gtg gtg ccc agc aac tgg cag Pro Ser Asn Trp Gln 195	740 Cys Asp Leu Pro Glu Ala Asp Thr Val Val Val 210	
200 205		
atg cat ggc tat gat gcc ccc atc tac acc aat gtc acc tac ccc atc Thr Tyr Pro Ile	788 Met His Gly Tyr Asp Ala Pro Ile Tyr Thr Asn Val	
215 220 225		
act gtg aac ccc cct ttt gtg ccc act gag aac ccc act ggc tgc tac Pro Thr Gly Cys Tyr	836 Thr Val Asn Pro Pro Phe Val Pro Thr Glu Asn	
230 235 240		
agc ctg acc ttc aat gtt gat gag agc tgg ctg caa gaa ggc cag acc Gln Glu Gly Gln Thr	884 Ser Leu Thr Phe Asn Val Asp Glu Ser Trp Leu	
245 250 255		
agg atc atc ttt gat gga gtc aac tct gcc ttc cac ctc tgg tgc aat Leu Trp Cys Asn	932 Arg Ile Ile Phe Asp Gly Val Asn Ser Ala Phe His	
260 265 270		
ggc agg tgg gtt ggc tat ggc caa gac agc agg ctg ccc tct gag ttt Leu Pro Ser Glu Phe 275	980 Gly Arg Trp Val Gly Tyr Gly Gln Asp Ser Arg 290	
280 285		
gac ctc tct gcc ttc ctc aga gct gga gag aac agg ctg gct gtc atg Arg Leu Ala Val Met	1028 Asp Leu Ser Ala Phe Leu Arg Ala Gly Glu Asn	
295 300 305		
gtg ctc agg tgg tct gat ggc agc tac ctg gaa gac caa gac atg tgg Asp Gln Asp Met Trp	1076 Val Leu Arg Trp Ser Asp Gly Ser Tyr Leu Glu	
310 315 320		
agg atg tct ggc atc ttc agg gat gtg agc ctg ctg cac aag ccc acc Leu His Lys Pro Thr	1124 Arg Met Ser Gly Ile Phe Arg Asp Val Ser Leu	
325 330 335		
acc cag att tct gac ttc cat gtt gcc acc agg ttc aat gat gac ttc Asn Asp Asp Phe	1172 Thr Gln Ile Ser Asp Phe His Val Ala Thr Arg Phe	
340 345 350		
agc aga gct gtg ctg gag gct gag gtg cag atg tgt gga gaa ctc aga Cys Gly Glu Leu Arg 355	1220 Ser Arg Ala Val Leu Glu Ala Glu Val Gln Met 370	
360 365		
gac tac ctg aga gtc aca gtg agc ctc tgg caa ggt gag acc cag gtg Gly Glu Thr Gln Val	1268 Asp Tyr Leu Arg Val Thr Val Ser Leu Trp Gln	
375 380 385		
ggc tct ggc aca gcc ccc ttt gga gga gag atc att gat gag aga gga	1316 Ala Ser Gly Thr Ala Pro Phe Gly Gly Glu Ile	

Ile Asp Glu Arg Gly 390 395 400	
ggc tat gct gac aga gtc acc ctg agg ctc aat gtg gag aac ccc aag Asn Val Glu Asn Pro Lys 405 410 415	1364 Gly Tyr Ala Asp Arg Val Thr Leu Arg Leu
ctg tgg tct gct gag atc ccc aac ctc tac agg gct gtt gtg gag ctg Ala Val Val Glu Leu 420 425 430	1412 Leu Trp Ser Ala Glu Ile Pro Asn Leu Tyr Arg
cac act gct gat ggc acc ctg att gaa gct gaa gcc tgt gat gtt gga Ala Cys Asp Val Gly 435 440 445	1460 His Thr Ala Asp Gly Thr Leu Ile Glu Ala Glu 450
ttc aga gaa gtc agg att gag aat ggc ctg ctg ctg ctc aat ggc aag Leu Leu Asn Gly Lys 455 460 465	1508 Phe Arg Glu Val Arg Ile Glu Asn Gly Leu Leu
cct ctg ctc atc agg gga gtc aac agg cat gag cac cac cct ctg cat His His Pro Leu His 470 475 480	1556 Pro Leu Leu Ile Arg Gly Val Asn Arg His Glu
gga caa gtg atg gat gaa cag aca atg gtg caa gat atc ctg cta atg Asp Ile Leu Leu Met 485 490 495	1604 Gly Gln Val Met Asp Glu Gln Thr Met Val Gln
aag cag aac aac ttc aat gct gtc agg tgc tct cac tac ccc aac cac His Tyr Pro Asn His 500 505 510	1652 Lys Gln Asn Asn Phe Asn Ala Val Arg Cys Ser
cct ctc tgg tac acc ctg tgt gac agg tat ggc ctg tat gtt gtt gat Leu Tyr Val Val Asp 515 520 525	1700 Pro Leu Trp Tyr Thr Leu Cys Asp Arg Tyr Gly 530
gaa gcc aac att gag aca cat ggc atg gtg ccc atg aac agg ctc aca Met Asn Arg Leu Thr 535 540 545	1748 Glu Ala Asn Ile Glu Thr His Gly Met Val Pro
gat gac ccc agg tgg ctg cct gcc atg tct gag aga gtg acc agg atg Glu Arg Val Thr Arg Met 550 555 560	1796 Asp Asp Pro Arg Trp Leu Pro Ala Met Ser
gtg cag aga gac agg aac cac ccc tct gtg atc atc tgg tct ctg ggc Ile Trp Ser Leu Gly 565 570 575	1844 Val Gln Arg Asp Arg Asn His Pro Ser Val Ile
aat gag tct gga cat gga gcc aac cat gat gct ctc tac agg tgg atc Leu Tyr Arg Trp Ile 580 585 590	1892 Asn Glu Ser Gly His Gly Ala Asn His Asp Ala
aag tct gtt gac ccc agc aga cct gtg cag tat gaa gga ggt gga gca Glu Gly Gly Gly Ala 595 600 605	1940 Lys Ser Val Asp Pro Ser Arg Pro Val Gln Tyr 610
gac acc aca gcc aca gac atc atc tgc ccc atg tat gcc agg gtt gat Tyr Ala Arg Val Asp 615 620 625	1988 Asp Thr Thr Ala Thr Asp Ile Ile Cys Pro Met
gag gac cag ccc ttc cct gct gtg ccc aag tgg agc atc aag aag tgg Ser Ile Lys Lys Trp 630 635 640	2036 Glu Asp Gln Pro Phe Pro Ala Val Pro Lys Trp
ctc tct ctg cct gga gag acc aga cct ctg atc ctg tgt gaa tat gca Leu Cys Glu Tyr Ala 645 650 655	2084 Leu Ser Leu Pro Gly Glu Thr Arg Pro Leu Ile
cat gca atg ggc aac tct ctg gga gcc ttt gcc aag tac tgg caa gcc Lvs Trp Trp Gln Ala	2132 His Ala Met Gly Asn Ser Leu Gly Gly Phe Ala

660	665	670	
ttc aga cag tac ccc agg ctg caa gga gga ttt gtg tgg gac tgg gtg	2180 Phe Arg Gln Tyr Pro Arg Leu Gln Gly Gly Phe		
Val Trp Asp Trp Val 675	680	685	690
gac caa tct ctc atc aag tat gat gag aat ggc aac ccc tgg tct gcc	2228 Asp Gln Ser Leu Ile Lys Tyr Asp Glu Asn Gly		
Asn Pro Trp Ser Ala	695	700	705
tat gga gga gac ttt ggt gac acc ccc aat gac agg cag ttc tgc atg	2276 Tyr Gly Gly Asp Phe Gly Asp Thr Pro Asn Asp		
Arg Gln Phe Cys Met	710	715	720
aat ggc ctg gtc ttt gca gac agg acc cct cac cct gcc ctc aca gag	2324 Asn Gly Leu Val Phe Ala Asp Arg Thr Pro His		
Pro Ala Leu Thr Glu	725	730	735
gcc aag cac cag caa cag ttc ttc cag ttc agg ctg tct gga cag acc	2372 Ala Lys His Gln Gln Gln Phe Phe Gln Phe Arg		
Leu Ser Gly Gln Thr	740	745	750
att gag gtg aca tct gag tac ctc ttc agg cac tct gac aat gag ctc	2420 Ile Glu Val Thr Ser Glu Tyr Leu Phe Arg His Ser		
Asp Asn Glu Leu 755	760	765	770
ctg cac tgg atg gtg gcc ctg gat ggc aag cct ctg gct tct ggt gag	2468 Leu His Trp Met Val Ala Leu Asp Gly Lys Pro		
Leu Ala Ser Gly Glu	775	780	785
gtg cct ctg gat gtg gcc cct caa gga aag cag ctg att gaa ctg cct	2516 Val Pro Leu Asp Val Ala Pro Gln Gly Lys Gln		
Leu Ile Glu Leu Pro	790	795	800
gag ctg cct cag cca gag tct gct gga caa ctg tgg cta aca gtg agg	2564 Glu Leu Pro Gln Pro Glu Ser Ala Gly Gln Leu		
Trp Leu Thr Val Arg	805	810	815
gtg gtt cag ccc aat gca aca gct tgg tct gag gca ggc cac atc tct	2612 Val Val Gln Pro Asn Ala Thr Ala Trp Ser Glu		
Ala Gly His Ile Ser	820	825	830
gca tgg cag cag tgg agg ctg gct gag aac ctc tct gtg acc ctg cct	2660 Ala Trp Gln Gln Trp Arg Leu Ala Glu Asn Leu		
Ser Val Thr Leu Pro 835	840	845	850
gct gcc tct cat gcc atc cct cac ctg aca aca tct gaa atg gac ttc	2708 Ala Ala Ser His Ala Ile Pro His Leu Thr Thr Ser		
Glu Met Asp Phe	855	860	865
tgc att gag ctg ggc aac aag aga tgg cag ttc aac agg cag tct ggc	2756 Cys Ile Glu Leu Gly Asn Lys Arg Trp Gln Phe		
Asn Arg Gln Ser Gly	870	875	880
ttc ctg tct cag atg tgg att gga gac aag aag cag ctc ctc acc cct	2804 Phe Leu Ser Gln Met Trp Ile Gly Asp Lys Lys		
Gln Leu Leu Thr Pro	885	890	895
ctc agg gac caa ttc acc agg gct cct ctg gac aat gac att gga gtg	2852 Leu Arg Asp Gln Phe Thr Arg Ala Pro Leu Asp		
Asn Asp Ile Gly Val	900	905	910
tct gag gcc acc agg att gac cca aat gct tgg gtg gag agg tgg aag	2900 Ser Glu Ala Thr Arg Ile Asp Pro Asn Ala Trp		
Val Glu Arg Trp Lys 915	920	925	930
gct gct gga cac tac cag gct gag gct gcc ctg ctc cag tgc aca gca	2948 Ala Ala Gly His Tyr Gln Ala Glu Ala Ala Leu		
Leu Gln Cys Thr Ala	935	940	945

gac acc ctg gct gat gct gtt ctg atc acc aca gcc cat gct tgg cag 2996 Asp Thr Leu Ala Asp Ala Val Leu Ile Thr Thr
Ala His Ala Trp Gln
950 955 960

cac caa ggc aag acc ctg ttc atc agc aga aag acc tac agg att gat 3044 His Gln Gly Lys Thr Leu Phe Ile Ser Arg Lys
Thr Tyr Arg Ile Asp
965 970 975

ggc tct gga cag atg gca atc aca gtg gat gtg gag gtt gcc tct gac 3092 Gly Ser Gly Gln Met Ala Ile Thr Val Asp Val
Glu Val Ala Ser Asp
980 985 990

aca cct cac cct gca agg att ggc ctg aac tgt caa ctg gca cag 3137 Thr Pro His Pro Ala Arg Ile Gly Leu Asn Cys
Gln Leu Ala Gln 995 1000 1005

gtg gct gag agg gtg aac tgg ctg ggc tta ggc cct cag gag aac 3182 Val Ala Glu Arg Val Asn Trp Leu Gly Leu
Gly Pro Gln Glu Asn 1010 1015 1020

tac cct gac agg ctg aca gct gcc tgc ttt gac agg tgg gac ctg 3227 Tyr Pro Asp Arg Leu Thr Ala Ala Cys Phe Asp
Arg Trp Asp Leu 1025 1030 1035

cct ctg tct gac atg tac acc cct tat gtg ttc cct tct gag aat 3272 Pro Leu Ser Asp Met Tyr Thr Pro Tyr Val Phe Pro
Ser Glu Asn 1040 1045 1050

ggc ctg agg tgt ggc acc agg gag ctg aac tat ggt cct cac cag 3317 Gly Leu Arg Cys Gly Thr Arg Glu Leu Asn
Tyr Gly Pro His Gln 1055 1060 1065

tgg agg gga gac ttc cag ttc aac atc tcc agg tac tct cag caa 3362 Trp Arg Gly Asp Phe Gln Phe Asn Ile Ser Arg
Tyr Ser Gln Gln 1070 1075 1080

cag ctc atg gaa acc tct cac agg cac ctg ctc cat gca gag gag 3407 Gln Leu Met Glu Thr Ser His Arg His Leu Leu
His Ala Glu Glu 1085 1090 1095

gga acc tgg ctg aac att gat ggc ttc cac atg ggc att gga gga 3452 Gly Thr Trp Leu Asn Ile Asp Gly Phe His Met
Gly Ile Gly Gly 1100 1105 1110

gat gac tct tgg tct cct tct gtg tct gct gag ttc cag tta tct 3497 Asp Asp Ser Trp Ser Pro Ser Val Ser Ala Glu Phe
Gln Leu Ser 1115 1120 1125

gct ggc agg tac cac tat cag ctg gtg tgg tgc cag aag taagctagct 3546 Ala Gly Arg Tyr His Tyr Gln Leu Val Trp Cys
Gln Lys 1130 1135 1140

gagtttcaga aaagggggccc tgagtggccc ctttttcaa ctttaattaac cttaaaacct 3606

ttaaaagcct tataatattct ttttttctt ataaaactta aaaccttaga ggctatttaa 3666

gttgctgatt tataattaatt ttattgttca aacatgagag cttagtacct gaaacatgag 3726

agcttagtac attagccatg agagcttagt acattagcca tgagggttta gttcattaaa 3786

catgagagct tagtacatta aacatgagag cttagtacct actatcaaca ggttgaactg 3846

ctgate 3852

<210> 15 <211> 1142 <212> PRT <213> Artificial Sequence

<220> <223> SEQ ID NO:14 = CpG-Free pSh-LacZ plasmid ; SEQ ID NO:15= SEQ ID N
O:14 CDS from 83 to 454 (SEQ ID NO:2) + SEQ ID NO:14 CDS from 483
to 3536 (SEQ ID NO:10)

<400> 15

Met Ala Lys Leu Thr Ser Ala Val Pro Val Leu Thr Ala Arg Asp Val 1

5

10

15

Ala Gly Ala Val Glu Phe Trp Thr Asp Arg Leu Gly Phe Ser Arg Asp
20 25 30

Phe Val Glu Asp Asp Phe Ala Gly Val Val Arg Asp Asp Val Thr Leu
35 40 45

Phe Ile Ser Ala Val Gln Asp Gln Val Val Pro Asp Asn Thr Leu Ala
50 55 60

Trp Val Trp Val Arg Gly Leu Asp Glu Leu Tyr Ala Glu Trp Ser Glu 65 70 75 80

Val Val Ser Thr Asn Phe Arg Asp Ala Ser Gly Pro Ala Met Thr Glu
85 90 95

Ile Gly Glu Gln Pro Trp Gly Arg Glu Phe Ala Leu Arg Asp Pro Ala
100 105 110

Gly Asn Cys Val His Phe Val Ala Glu Glu Gln Asp Met Asp Pro Val
115 120 125

Val Leu Gln Arg Arg Asp Trp Glu Asn Pro Gly Val Thr Gln Leu Asn
130 135 140

Arg Leu Ala Ala His Pro Pro Phe Ala Ser Trp Arg Asn Ser Glu Glu 145 150 155 160

Ala Arg Thr Asp Arg Pro Ser Gln Gln Leu Arg Ser Leu Asn Gly Glu
165 170 175

Trp Arg Phe Ala Trp Phe Pro Ala Pro Glu Ala Val Pro Glu Ser Trp
180 185 190

Leu Glu Cys Asp Leu Pro Glu Ala Asp Thr Val Val Val Pro Ser Asn
195 200 205

Trp Gln Met His Gly Tyr Asp Ala Pro Ile Tyr Thr Asn Val Thr Tyr
210 215 220

Pro Ile Thr Val Asn Pro Pro Phe Val Pro Thr Glu Asn Pro Thr Gly 225 230 235 240

Cys Tyr Ser Leu Thr Phe Asn Val Asp Glu Ser Trp Leu Gln Glu Gly
245 250 255

Gln Thr Arg Ile Ile Phe Asp Gly Val Asn Ser Ala Phe His Leu Trp
260 265 270

Cys Asn Gly Arg Trp Val Gly Tyr Gly Gln Asp Ser Arg Leu Pro Ser
275 280 285

Glu Phe Asp Leu Ser Ala Phe Leu Arg Ala Gly Glu Asn Arg Leu Ala

290	295	300			
Val Met Val Leu Arg Trp Ser Asp Gly Ser Tyr Leu Glu Asp Gln Asp 305			310	315	320
Met Trp Arg Met Ser Gly Ile Phe Arg Asp Val Ser Leu Leu His Lys					
325	330	335			
Pro Thr Thr Gln Ile Ser Asp Phe His Val Ala Thr Arg Phe Asn Asp					
340	345	350			
Asp Phe Ser Arg Ala Val Leu Glu Ala Glu Val Gln Met Cys Gly Glu					
355	360	365			
Leu Arg Asp Tyr Leu Arg Val Thr Val Ser Leu Trp Gln Gly Glu Thr					
370	375	380			
Gln Val Ala Ser Gly Thr Ala Pro Phe Gly Gly Glu Ile Ile Asp Glu 385			390	395	400
Arg Gly Gly Tyr Ala Asp Arg Val Thr Leu Arg Leu Asn Val Glu Asn					
405	410	415			
Pro Lys Leu Trp Ser Ala Glu Ile Pro Asn Leu Tyr Arg Ala Val Val					
420	425	430			
Glu Leu His Thr Ala Asp Gly Thr Leu Ile Glu Ala Glu Ala Cys Asp					
435	440	445			
Val Gly Phe Arg Glu Val Arg Ile Glu Asn Gly Leu Leu Leu Leu Asn					
450	455	460			
Gly Lys Pro Leu Leu Ile Arg Gly Val Asn Arg His Glu His His Pro 465			470	475	480
Leu His Gly Gln Val Met Asp Glu Gln Thr Met Val Gln Asp Ile Leu					
485	490	495			
Leu Met Lys Gln Asn Asn Phe Asn Ala Val Arg Cys Ser His Tyr Pro					
500	505	510			
Asn His Pro Leu Trp Tyr Thr Leu Cys Asp Arg Tyr Gly Leu Tyr Val					
515	520	525			
Val Asp Glu Ala Asn Ile Glu Thr His Gly Met Val Pro Met Asn Arg					
530	535	540			
Leu Thr Asp Asp Pro Arg Trp Leu Pro Ala Met Ser Glu Arg Val Thr 545			550	555	560
Arg Met Val Gln Arg Asp Arg Asn His Pro Ser Val Ile Ile Trp Ser					
565	570	575			

Leu Gly Asn Glu Ser Gly His Gly Ala Asn His Asp Ala Leu Tyr Arg
580 585 590

Trp Ile Lys Ser Val Asp Pro Ser Arg Pro Val Gln Tyr Glu Gly Gly
595 600 605

Gly Ala Asp Thr Thr Ala Thr Asp Ile Ile Cys Pro Met Tyr Ala Arg
610 615 620

Val Asp Glu Asp Gln Pro Phe Pro Ala Val Pro Lys Trp Ser Ile Lys 625 630 635 640

Lys Trp Leu Ser Leu Pro Gly Glu Thr Arg Pro Leu Ile Leu Cys Glu
645 650 655

Tyr Ala His Ala Met Gly Asn Ser Leu Gly Gly Phe Ala Lys Tyr Trp
660 665 670

Gln Ala Phe Arg Gln Tyr Pro Arg Leu Gln Gly Gly Phe Val Trp Asp
675 680 685

Trp Val Asp Gln Ser Leu Ile Lys Tyr Asp Glu Asn Gly Asn Pro Trp
690 695 700

Ser Ala Tyr Gly Gly Asp Phe Gly Asp Thr Pro Asn Asp Arg Gln Phe 705 710 715 720

Cys Met Asn Gly Leu Val Phe Ala Asp Arg Thr Pro His Pro Ala Leu
725 730 735

Thr Glu Ala Lys His Gln Gln Gln Phe Phe Gln Phe Arg Leu Ser Gly
740 745 750

Gln Thr Ile Glu Val Thr Ser Glu Tyr Leu Phe Arg His Ser Asp Asn
755 760 765

Glu Leu Leu His Trp Met Val Ala Leu Asp Gly Lys Pro Leu Ala Ser
770 775 780

Gly Glu Val Pro Leu Asp Val Ala Pro Gln Gly Lys Gln Leu Ile Glu 785 790 795 800

Leu Pro Glu Leu Pro Gln Pro Glu Ser Ala Gly Gln Leu Trp Leu Thr
805 810 815

Val Arg Val Val Gln Pro Asn Ala Thr Ala Trp Ser Glu Ala Gly His
820 825 830

Ile Ser Ala Trp Gln Gln Trp Arg Leu Ala Glu Asn Leu Ser Val Thr
835 840 845

Leu Pro Ala Ala Ser His Ala Ile Pro His Leu Thr Thr Ser Glu Met
850 855 860

Asp Phe Cys Ile Glu Leu Gly Asn Lys Arg Trp Gln Phe Asn Arg Gln 865 870 875 880

Ser Gly Phe Leu Ser Gln Met Trp Ile Gly Asp Lys Lys Gln Leu Leu
885 890 895

Thr Pro Leu Arg Asp Gln Phe Thr Arg Ala Pro Leu Asp Asn Asp Ile
900 905 910

Gly Val Ser Glu Ala Thr Arg Ile Asp Pro Asn Ala Trp Val Glu Arg
915 920 925

Trp Lys Ala Ala Gly His Tyr Gln Ala Glu Ala Ala Leu Leu Gln Cys
930 935 940

Thr Ala Asp Thr Leu Ala Asp Ala Val Leu Ile Thr Thr Ala His Ala 945 950 955 960

Trp Gln His Gln Gly Lys Thr Leu Phe Ile Ser Arg Lys Thr Tyr Arg
965 970 975

Ile Asp Gly Ser Gly Gln Met Ala Ile Thr Val Asp Val Glu Val Ala
980 985 990

Ser Asp Thr Pro His Pro Ala Arg Ile Gly Leu Asn Cys Gln Leu Ala
995 1000 1005

Gln Val Ala Glu Arg Val Asn Trp Leu Gly Leu Gly Pro Gln Glu
1010 1015 1020

Asn Tyr Pro Asp Arg Leu Thr Ala Ala Cys Phe Asp Arg Trp Asp
1025 1030 1035

Leu Pro Leu Ser Asp Met Tyr Thr Pro Tyr Val Phe Pro Ser Glu
1040 1045 1050

Asn Gly Leu Arg Cys Gly Thr Arg Glu Leu Asn Tyr Gly Pro His
1055 1060 1065

Gln Trp Arg Gly Asp Phe Gln Phe Asn Ile Ser Arg Tyr Ser Gln
1070 1075 1080

Gln Gln Leu Met Glu Thr Ser His Arg His Leu Leu His Ala Glu
1085 1090 1095

Glu Gly Thr Trp Leu Asn Ile Asp Gly Phe His Met Gly Ile Gly
1100 1105 1110

Gly Asp Asp Ser Trp Ser Pro Ser Val Ser Ala Glu Phe Gln Leu
 1115 1120 1125

Ser Ala Gly Arg Tyr His Tyr Gln Leu Val Trp Cys Gln Lys
 1130 1135 1140

<210> 16 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 16 aactcagctg aggaggcaga ccatggccaa gttgaccagt 40

<210> 17 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 17 gctgtcccag tgctcacagc cagggatgtg gctggagctg 40

<210> 18 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 18 ttgagttctg gactgacagg ttggggttct ccagagattt 40

<210> 19 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 19 tgtggaggat gactttgcag gtgtggtcag agatgatgtc 40

<210> 20 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 20 accctgttca tctcagcagt ccaggaccag gtggtgcctg 40

<210> 21 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 21 acaacaccct ggcttgggtg tgggtgagag gactggatga 40

<210> 22 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 22 gctgtatgct gattggagtg aggtggtctc caccaacttc 40

<210> 23 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 23 agggatgcca gtggccctgc catgacagag attggagagc 40

<210> 24 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 24 agccctgggg gagagagttt gccctgagag acccagcagg 40

<210> 25 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 25 caactgtgtg cactttgtgg cagaggagca ggactgagga 40

<210> 26 <211> 21 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 26 taagaattca gctagctcga c 21

<210> 27 <211> 41 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 27 gtcgagctag ctgaattctt atctctcagtc ctgctctct g 41

<210> 28 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 28 ccacaaagtg cacacagttg cctgctgggt ctctcagggc 40

<210> 29 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 29 aaactctctc ccccagggt gctctccaat ctctgtcatg 40

<210> 30 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 30 gcagggccac tggcatccct gaagttggtg gagaccacct 40

<210> 31 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 31 cactccactc agcatacagc tcattcagtc ctctcaccca 40

<210> 32 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 32 caccaagcc aggtgtgtgt caggcaccac ctggtcctgg 40

<210> 33 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 33 actctctgaga tgaacaggtt gacatcatct ctgaccacac 40

<210> 34 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 34 ctgcaaagtc atcctccaca aaatctctgg agaaccceaa 40

<210> 35 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 35 cctgtcagtc cagaactcaa cagctccagc cacatccctg 40

<210> 36 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 36 gctgtgagca ctgggacagc actggtaac ttggccatgg 40

<210> 37 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Sh ble

<400> 37 tctgcctcct cagctgagtt 20

<210> 38 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 38 tgagatcacc gggtcagctg aggaggcaca tcatgaagaa acctgaactg acagcaactt 60

<210> 39 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 39 ctgttgagaa gtttctcatt gaaaaatttg attctgttgc tgatctcatg cagctgtctg 60

<210> 40 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 40 aaggtgaaga aagcagagcc ttttcttttg atgttgagg aagagggtat gttctgaggg 60

<210> 41 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 41 tcaattcttg tgctgatgtg ttttacaag acagatatgt ttacagacac ttgcctctg 60

<210> 42 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 42 ctgctctgcc aattccagaa gttctggaca ttggagaatt ttctgaatct ctcactact 60

<210> 43 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 43 gcatcagcag aagagcaciaa ggagtcactc tccaggatct ccctgaaact gagtgccag 60

<210> 44 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 44 ctgttctgca acctgttgct gaagcaatgg atgccattgc agcagctgat ctgagccaaa 60

<210> 45 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 45 cctctggatt tggctctttt ggccccaaag gcattgtgca gtacaccact tggagggatt 60

<210> 46 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 46 tcatttgtgc cattgctgat cctcatgtct atcactggca gactgtgatg gatgacacag 60

<210> 47 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 47 ttctgcttc tggctgcag gcactggatg aactcatgct gtgggcagaa gattgtcctg 60

<210> 48 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 48 aagtcagaca cctgggccat gctgattttg gaagcaaciaa tggctgaca gacaatggca 60

<210> 49 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 49 gaatcactgc agtcattgac tggctgaag ccatgtttgg agattctcaa tatgaggttg 60

<210> 50 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 50 ccaacatttt tttttgaga ccttggtcgg ctgcatgga acaacaaca agatattttg 60

<210> 51 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 51 aaagaagaca cccagaactg gctgggtccc ccagactgag agcctacatg ctcagaattg 60

<210> 52 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 52 gcctggacca actgtatcaa tctctggttg atggaaactt tgatgatgct gcttgggcac 60

<210> 53 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 53 aaggaagatg tgatgccatt gtgaggtctg gtgctggaac tgttgaaga actcaaattg 60

<210> 54 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 54 caagaaggtc tgctgctgtt tggactgatg gatgtgttga agttctggct gactctggaa 60

<210> 55 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 55 acaggagacc ctccacaaga cccagagcca aggaatgaat attagctagc ggatcctgag 60

<210> 56 <211> 30 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 56 ctcaggatcc gctagctaat attcattcct 30

<210> 57 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 57 tggctctggg tcttgtggag ggtctctgtg ttccagagtc agccagaact tcaacacatc 60

<210> 58 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 58 catcagtcga aacagcagca gaccttcttg caatttgagt tcttccaaca gttccagcac 60

<210> 59 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 59 cagacctcac aatggcatca catcttctct gtgccaagc agcatcatca aagttccat 60

<210> 60 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 60 caaccagaga ttgatacagt tggccaggc caattctgag catgtagget ctacgtctgg 60

<210> 61 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 61 gggaaccagc cagttctggg tgtctctttt caaaatatct tgtttgtgt tccatgcaag 60

<210> 62 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 62 ccagccaagg tctccaaaa aaaatgttg caacctcata ttgagaatct ccaacatgg 60

<210> 63 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 63 ctcagacca gtcaatgact gcagtgatc tgccattgtc tgcagaaca ttgttgcttc 60

<210> 64 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 64 caaatcagc atggaccagg tgtctgactt caggacaatc ttctgccac agcatgagtt 60

<210> 65 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 65 catcagtc ctgagcaaca gaagcagaaa ctgtgtcatc catcacagtc tgccagtgt 60

<210> 66 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 66 agacatgagg atcagcaatg gcacaaatga aatccctcca agtgggtgtac tgaccaatgc 60

<210> 67 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 67 ctgggggacc aaaaggacca aatccagagg ttgggtcag atcagctgct gcaatggcat 60

<210> 68 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 68 ccattgcttc agcaacaggt tgcagaacag ctggcagctc agtttcaggg agatcctgga 60

<210> 69 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 69 gagtactcc ttgtgctctt ctgctgatgc agtaggtgag agattcagaa aatttccaa 60

<210> 70 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 70 tgccagaac ttctggaatt ggcagagcag cagaggcaaa gtgtctgtaa acatatctgt 60

<210> 71 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 71 ctttgtaaaa accatcagca caagaattga ccctcagaac ataacctctt cctccaacat 60

<210> 72 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 72 caaaagaaaa ggctctgctt tcttcacctt cagacagctg catgagatca gaaacagaat 60

<210> 73 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 73 caaatTTTtTc aatgagaaac ttctcaacag aagttgctgt cagttcaggt ttctcatga 60

<210> 74 <211> 30 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Hph

<400> 74 tgtgcctcct cagctgaacc ggtgatctca 30

<210> 75 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 75 aggaggcaca tcatgaagac cttcaacatc tctcagcagg 40

<210> 76 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 76 atctggagct ggtggaggtc gccactgaga agatcaccat 40

<210> 77 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 77 gctctatgag gacaacaagc accatgtcgg ggcggccatc 40

<210> 78 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 78 aggaccaaga ctggggagat catctctgct gtccacattg 40

<210> 79 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 79 aggcctacat tggcagggtc actgtctgtg ctgaagccat 40

<210> 80 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 80 tgccattggg tctgctgtga gcaacgggca gaaggacttt 40

<210> 81 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-Free Bsr
 <400> 81 gacaccattg tggctgtcag gcaccctac tctgatgagg 40

<210> 82 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-Free Bsr
 <400> 82 tggacagatc catcagggtg gtcagccct gtggcatgtg 40

<210> 83 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-Free Bsr
 <400> 83 cagagagctc atctctgact atgctctga ctgctttgtg 40

<210> 84 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-Free Bsr
 <400> 84 ctcatgaga tgaatggcaa gctggtcaaa accaccattg 40

<210> 85 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-Free Bsr
 <400> 85 aggaactcat cccctcaag tacaccagga actaacctg 40

<210> 86 <211> 21 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-Free Bsr
 <400> 86 aattcagcta gctcgacatg a 21

<210> 87 <211> 41 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-Free Bsr
 <400> 87 tcatgtcgag ctactgaat tcaggttag ttctggtgt a 41

<210> 88 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-Free Bsr
 <400> 88 cttgaggggg atgagttcct caatggtggt ttgaccagc 40

<210> 89 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-Free Bsr
 <400> 89 ttgccattca tctcaatgag cacaagcag tcaggagcat 40

<210> 90 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 90 agtcagagat gagctctctg cacatgccac aggggctgac 40

<210> 91 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 91 caccctgatg gatctgtcca cctcatcaga gtaggggtgc 40

<210> 92 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 92 ctgacagcca caatggtgtc aaagtccttc tgcccgttgc 40

<210> 93 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 93 tcacagcaga cccaatggca atggcttcag cacagacagt 40

<210> 94 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 94 gaccctgcca atgtaggcct caatgtggac agcagagatg 40

<210> 95 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 95 atctcccag tcttggtcct gatggccgcc ccgacatggt 40

<210> 96 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 96 gcttggtgtc ctcatagagc atggtgatct ttcagtggc 40

<210> 97 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 97 gacctccacc agctccagat cctgctgaga gatgttgaag 40

<210> 98 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Bsr

<400> 98 gtcttcatga tgtgcctcct 20

<210> 99 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 99 ctcaatatag gaggaccatc atgactgaat acaaaccac agtgaagctg gcaaccagag 60

<210> 100 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 100 atgatgttcc aagagctgtg agaacactgg ctgctgcttt tgcagactac cctgcaacaa 60

<210> 101 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 101 ggcacacagt tgacctgac aggcacattg agagggtgac agaactgcaa gaactcttcc 60

<210> 102 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 102 tcaccagagt gggactggac attggaaaag ttgggttgc agatgatgga gctgctgttg 60

<210> 103 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 103 cagtttggac aacacctgag tctgtgaag ctggtgctgt tttgctgaa attggacaa 60

<210> 104 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 104 gaatggctga gctctctgga agcaggctgg cagcacaaca acaaatggaa ggtctgctgg 60

<210> 105 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 105 caccacacag gccaaaagag ccagcttggc ttctggcaac tgttgagtg agccctgacc 60

<210> 106 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 106 accagggaaa gggcttgga tctgctgtt ttctgcctgg agtgaagct gctgaaaggg 60

<210> 107 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 107 ctggagtcc tgccttctg gaaactctg ctcccagaaa cctgcctttt tatgaaagac 60

<210> 108 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 108 tgggattcac tgtgacagt gatgtgagg ttccagaagg cccaagaact tggatcatga 60

<210> 109 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 109 caaggaagcc tggagcttaa acctgagcta gctcgacatg ataagataca ttgatgagtt 60

<210> 110 <211> 30 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 110 aactcatcaa tgtatcttat catgtcgagc 30

<210> 111 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 111 tagctcaggt ttaagctcca ggcttccttg tcatgcacca agttcttggg ccttctggaa 60

<210> 112 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 112 cctcaacatc agctgtcaca gtgaatcca gtcttcata aaaaggcagg ttcttgggag 60

<210> 113 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 113 cagaagtttc cagaaaggca ggaactccag cctttcagc agttcaact ccaggcagaa 60

<210> 114 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 114 caacagcaga tccagaccc ttccctgggt ggtaggggt cactccaaca gttgccagaa 60

<210> 115 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 115 accaagctgg ctcttttggc ctgtgtggtg ccagcagacc ttccatttgt tgtgtgctg 60

<210> 116 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 116 ccagcctgct tccagagagc tcagccattc ttggtccaat ttcagcaaaa acagcaccag 60

<210> 117 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 117 cttcaacaga ctcaggtgtt gtccaaactg caacagcagc tccatcatct gcaacccaaa 60

<210> 118 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 118 ctttccaat gtccagtcct actctggtga ggaagagttc ttgcagttct gtcaccctct 60

<210> 119 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 119 caatgtgcct gtcagggtca actgtgtgcc ttgttcagg gtagtctgca aaagcagcag 60

<210> 120 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 120 ccagtgttct cacagctctt ggaacatcat ctctggttgc cagcctcact gtgggtttgt 60

<210> 121 <211> 30 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-Free Pac

<400> 121 actcagtcac gatggctctc ctatagttag 30

<210> 122 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 122 atcactatag gagggccacc atggaccctg ttgtgctgca 40

<210> 123 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 123 ggtggccctc ctatagttag 20

<210> 124 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 124 aaggagagac tgggagaacc ctggagtgc ccagctcaac 40

<210> 125 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 125 ggttctccca gtctctctt tgcagcaca cagggtccat 40

<210> 126 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 126 agactggctg cccaccctcc cttgcctct tggaggaact 40

<210> 127 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 127 ggagggtggg cagccagctt gttgagctgg gtcactccag 40

<210> 128 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 128 ctgaggaagc caggacagac aggcccagcc agcagctcag 40

<210> 129 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 129 gtctgtcctg gcttcctcag agttcctcca agaggcaaag 40

<210> 130 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 130 gtctctcaat ggagagtgga ggtttgcctg gttccctgcc 40

<210> 131 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 131 tcactctec attgagagac ctgagctgct ggctgggcct 40

<210> 132 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 132 cctgaagctg tgcctgagtc ttggctggag tgtgacctcc 40

<210> 133 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 133 gactcaggca cagcttcagg ggcagggaac caggcaaacc 40

<210> 134 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 134 cagaggtga cactgtgtg gtgccagca actggcagat 40

<210> 135 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 135 cacaacagtg tcagcctctg ggaggtcaca ctccagccaa 40

<210> 136 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 136 gcatggetat gatgccccca tctacacaa tgtcacctac 40

<210> 137 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ

<400> 137 tgggggcatc atagccatgc atctgccagt tgctgggcac	40
<210> 138 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 138 cccatcactg tgaaccccc tttgtgcc actgagaacc	40
<210> 139 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 139 ggggggttca cagtgatggg gtaggtgaca ttggtgtaga	40
<210> 140 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 140 ccactggctg ctacagcctg acctcaatg ttgatgagag	40
<210> 141 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 141 caggctgtag cagccagtgg ggttctcagt gggcacaaaa	40
<210> 142 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 142 ctggctgcaa gaaggccaga ccagatcat cttgatgga	40
<210> 143 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 143 tctggccttc ttgcagccag ctctcatcaa cattgaaggt	40
<210> 144 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 144 gtcaactctg cttccacct ctggtgcaat ggcaggtggg	40
<210> 145 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 145 agtggaagg cagagttgac tccatcaaag atgactctgg	40
<210> 146 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 146 ttgctatgg ccaagacagc aggetgccct ctgagttga	40

<210> 147 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 147 gctgtcttgg ccatagccaa cccacctgcc attgcaccag 40

<210> 148 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 148 cctctctgcc ttctcagag ctggagagaa caggctggct 40

<210> 149 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 149 ctctgaggaa ggcagagagg tcaaactcag agggcagcct 40

<210> 150 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 150 gtcattgtgc tcaggtggtc tgatggcagc tacctggaag 40

<210> 151 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 151 gaccacctga gcacatgac agccagcctg ttctctcag 40

<210> 152 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 152 accaagacat gtggaggatg tctggcatct tcagggatgt 40

<210> 153 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 153 catctccac atgtcttggc cttccagta gctgccatca 40

<210> 154 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 154 gagcctgctg cacaagccca ccaccagat ttctgacttc 40

<210> 155 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 155 tgggcttgctg cagcaggctc acatcctga agatgccaga 40

<210> 156 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 156 catgttgcca ccaggttcaa tgatgacttc agcagagctg 40

<210> 157 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 157 ttgaacctgg tggcaacatg gaagtcagaa atctgggtgg 40

<210> 158 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 158 tgctggaggc tgaggcag atgtgtggag aactcagaga 40

<210> 159 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 159 ctgcacctca gcctccagca cagctctgct gaagtcatca 40

<210> 160 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 160 ctacctgaga gtcacagtga gcctctggca aggtgagacc 40

<210> 161 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 161 tcactgtgac tctcaggtag tctctgagtt ctccacacat 40

<210> 162 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 162 caggtggcct ctggcacagc cccctttgga ggagagatca 40

<210> 163 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 163 gctgtgccag aggccacctg ggtctcacct tgccagaggc 40

<210> 164 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 164 ttgatgagag aggaggctat gctgacagag tcaccctgag 40

<210> 165 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 165 atagcctcct ctctcatcaa tcatctctcc tccaaagggg 40

<210> 166 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 166 gctcaatgtg gagaacccca agctgtggtc tgctgagatc 40

<210> 167 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 167 tggggttctc cacattgagc ctcagggtga ctctgtcagc 40

<210> 168 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 168 cccaacctct acagggtgtg tgtggagctg cacactgtg 40

<210> 169 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 169 acagccctgt agagggtggg gatctcagca gaccacagct 40

<210> 170 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 170 atggcacctt gattgaagct gaagcctgtg atgttgatt 40

<210> 171 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 171 agttcaatc aggttgccat cagcagtgtg cagctccaca 40

<210> 172 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 172 cagagaagtc aggattgaga atggcctgct gctgetcaat 40

<210> 173 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 173 tetcaatcct gacttctctg aatccaacat cacaggettc 40

<210> 174 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 174 ggcaagcctc tgctcatcag gggagtcaac aggcatgagc 40

<210> 175 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 175 ctgatgagca gaggcttgcc attgagcagc agcaggccat 40

<210> 176 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 176 accaccctct gcattggacaa gtgatggatg aacagacaat 40

<210> 177 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 177 ttgtccatgc agagggtggt gctcatgcct gttgactccc 40

<210> 178 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 178 ggtgcaagat atcctgctga tgaagcagaa ctccgcctac 40

<210> 179 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 179 tcagcaggat atcttgcacc attgtctgtt catccatcac 40

<210> 180 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 180 gtaggcggag ttctgcttca 20

<210> 181 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 181 tcattagcag gatattctgc 20

<210> 182 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 182 gcaagatatc ctgctaata agcagaacaa cttcaatgct 40

<210> 183 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 183 gggtagttag agcacctgac agcattgaag ttgttctgct 40

<210> 184 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 184 gtcaggtgct ctcactaccc caaccaccct ctctggtaca	40
<210> 185 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 185 gccatactg tcacacaggg tgtaccagag aggggtggtg	40
<210> 186 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 186 ccctgtgtga caggtatggc ctgtatgtg ttgatgaagc	40
<210> 187 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 187 tgccatgtgt ctcaatgttg gcttcatcaa caacatacag	40
<210> 188 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 188 caacattgag acacatggca tggtgcccat gaacaggctc	40
<210> 189 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 189 agccacctgg ggtcatctgt gacgctgttc atgggcacca	40
<210> 190 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 190 acagatgacc ccaggtggct gcctgccatg tctgagagag	40
<210> 191 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 191 ttctgcacc atcctgtgca ctctctcaga catggcaggc	40
<210> 192 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 192 tgaccaggat ggtgcagaga gacaggaacc acccctctgt	40
<210> 193 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 193 tgcccagaga ccagatgac acagaggggt ggttcctgtc	40

<210> 194 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 194 gatcatctgg tctctgggca atgagtctgg acatggagcc 40

<210> 195 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 195 ctgtagagag catcatgggt ggctccatgt ccagactcat 40

<210> 196 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 196 aaccatgatg ctctctacag gtggatcaag tctgttgacc 40

<210> 197 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 197 atactgcaca ggtctgtctgg gtcaacaga ctgatccac 40

<210> 198 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 198 ccagcagacc tgtgcagtat gaaggagggtg gagcagacac 40

<210> 199 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 199 agatgatgtc tgtggctgtg gtgtctgctc cacctccttc 40

<210> 200 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 200 cacagccaca gacatcatct gccccatgta tgccagggtt 40

<210> 201 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 201 gggaagggtt ggtctctatc aaccctggca tacatggggc 40

<210> 202 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 202 gatgaggacc agccctccc tgctgtgccc aagtggagca 40

<210> 203 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ

<400> 203 cagagagagc cacttcttga tgctccactt gggcacagca	40
<210> 204 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 204 tcaagaagtg gctctctctg cctggagaga ccagacctct	40
<210> 205 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 205 gtgcatattc acacaggatc agaggtctgg tctctccagg	40
<210> 206 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 206 gatcctgtgt gaatatgcac atgcaatggg caactctctg	40
<210> 207 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 207 cagtacttgg caaagcctcc cagagagttg cccattgcat	40
<210> 208 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 208 ggaggctttg ccaagtactg gcaagccttc agacagtacc	40
<210> 209 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 209 aaatcctcct tgcagcctgg ggtactgtct gaaggettgc	40
<210> 210 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 210 ccaggtgca aggaggattt gtgtgggact ggggtggacca	40
<210> 211 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 211 catcatactt gatgagagat tggtcacccc agtcccacac	40
<210> 212 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 212 atctctcatc aagtatgatg agaatggcaa cccctggtct	40

<210> 213 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 213 ccaaagtctc ctccataggc agaccagggg ttgccattct 40

<210> 214 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 214 gcctatggag gagactttgg tgacaccccc aatgacaggc 40

<210> 215 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 215 caggccattc atgcagaact gcctgtcatt gggggtgtca 40

<210> 216 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 216 agttctgcat gaatggcctg gtctttgcag acaggacccc 40

<210> 217 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 217 cctctgtgag ggcaggggtga ggggtcctgt ctgcaaagac 40

<210> 218 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 218 tcacctgcc ctcacagagg ccaagcacca gcaacagttc 40

<210> 219 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 219 ccagacagcc tgaactggaa gaactgttgc tgggtccttg 40

<210> 220 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 220 ttccagtca ggctgtctgg acagaccatt gaggtgacat 40

<210> 221 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 221 gtgcctgaag aggtactcag atgtcacctc aatggtctgt 40

<210> 222 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 222 ctgagtacct cttcaggcac tctgacaatg agctcctgca 40

<210> 223 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 223 tgcaggagct cattgtcaga 20

<210> 224 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 224 gtaatttaac aatgagctcc tgcactggat ggtggccctg 40

<210> 225 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 225 ggagctcatt gttaaattac 20

<210> 226 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 226 gatggcaagc ctctggcttc tggtgagggtg cctctggatg 40

<210> 227 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 227 gaagccagag gcttgccatc cagggccacc atccagtga 40

<210> 228 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 228 tggcccctca aggaaagcag ctgattgaac tgcctgagct 40

<210> 229 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 229 ctgctttctt tgaggggcca catccagagg cacctcacca 40

<210> 230 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 230 gcctcagcca gagtctgctg gacaactgtg gctaacagtg 40

<210> 231 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 231 cagcagactc tggctgaggc agctcaggca gttcaatcag 40

<210> 232 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 232 aggggtgggtc agcccaatgc aacagcttgg tctgaggcag 40

<210> 233 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 233 gcattgggct gaaccaccct cactgttagc cacagttgtc 40

<210> 234 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 234 gccacatctc tgcattggcag cagtggaggc tggctgagaa 40

<210> 235 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 235 ctgcatgca gagatgtggc ctgcctcaga ccaagctgtt 40

<210> 236 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 236 cctctctgtg accctgcctg ctgcctctca tgccatccct 40

<210> 237 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 237 caggcagggt cacagagagg ttctcagcca gcctccactg 40

<210> 238 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 238 cacctgacaa catctgaaat ggacttctgc attgagctgg 40

<210> 239 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 239 atttcagatg ttgtcagggt agggatggca ttagaggcag 40

<210> 240 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 240 gcaacaagag atggcagttc aacaggcagt ctggcttct 40

<210> 241 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 241 gaactgccat ctctgttgc ccagctcaat gcagaagtcc 40

<210> 242 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 242 gtctcagatg tggattggag acaagaagca gctcctcacc 40

<210> 243 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 243 ctccaatcca catctgagac aggaagccag actgcctgtt 40

<210> 244 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 244 cctctcaggg accaattcac cagggctcct ctggacaatg 40

<210> 245 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 245 gtgaattggt ccctgagagg ggtgaggagc tgcttctgt 40

<210> 246 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 246 acattggagt gtctgaggcc accaggattg acccaaatgc 40

<210> 247 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 247 ggcctcagac actccaatgt cattgtccag aggagccctg 40

<210> 248 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 248 ttgggtggag aggtggaagg ctgctggaca ctaccaggct 40

<210> 249 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 249 cttccacct ctccaccaa gcatttgggt caatcctggt 40

<210> 250 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 250 gaggtgccc tgctccagtg cacagcagac accctggctg	40
<210> 251 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 251 cactggagca gggcagcctc agcctggtag tgtccagcag	40
<210> 252 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 252 atgctgttct gatcaccaca gcccatgctt ggcagcacca	40
<210> 253 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 253 tgtgtgatc agaacagcat cagccagggt gtctgctgtg	40
<210> 254 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 254 aggcaagacc ctgttcatca gcagaaagac ctacaggatt	40
<210> 255 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 255 tgatgaacag ggtcttcct ttgtgctgcc aagcatgggc	40
<210> 256 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 256 gatggctctg gacagatggc aatcacagtg gatgtggagg	40
<210> 257 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 257 gccatctgtc cagagccatc aatcctgtag gtctttctgc	40
<210> 258 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 258 ttgcctetga cacacctcac cctgcaagga ttggcctgaa	40
<210> 259 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 259 gtgaggtgtg tcagaggcaa cctccacatc cactgtgatt	40

<210> 260 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 260 ctgtcaactg gcacaggtgg ctgagagggt gaactggctg 40

<210> 261 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 261 ccacctgtgc cagttgacag ttcaggccaa tccttcagg 40

<210> 262 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 262 ggcttaggcc ctcaggagaa ctaccctgac aggctgacag 40

<210> 263 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 263 ttctctgag ggctaagcc cagccagttc accctctcag 40

<210> 264 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 264 ctgcctgctt tgacaggtgg gacctgcctc tgtctgacat 40

<210> 265 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 265 ccacctgtca aagcaggcag ctgtcagcct gtcagggtag 40

<210> 266 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 266 gtacaccct tatgtgtcc cttctgagaa tggcctgagg 40

<210> 267 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 267 ggaacacata aggggtgtac atgtcagaca gaggcaggtc 40

<210> 268 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 268 tgtggcacca gggagctgaa ctatgtctct caccagtgga 40

<210> 269 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for CpG-free LacZ

<400> 269 ttcagctccc tgggtgccaca cctcaggcca ttctcagaag	40
<210> 270 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 270 ggggagactt ccagttcaac atctccaggt actctcagca	40
<210> 271 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 271 gttgaactgg aagtctcccc tccactggtg aggaccatag	40
<210> 272 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 272 acagctcatg gaaacctctc acaggcacct gctccatgca	40
<210> 273 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 273 gagaggttcc catgagctgt tgctgagagt acctggagat	40
<210> 274 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 274 gaggagggaa cctggctgaa cattgatggc ttccacatgg	40
<210> 275 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 275 ttcagccagg ttccctctc tgcattggagc aggtgcctgt	40
<210> 276 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 276 gcattggagg agatgactct tgggtctctt ctgtgtctgc	40
<210> 277 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 277 agagtcattc ctcacatgc ccatgtggaa gccatcaatg	40
<210> 278 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> assembling oligo for CpG-free LacZ	
<400> 278 tgagttccag ttatctgtgc gcaggtacca ctatcagctg	40

<210> 279 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 279 cagcagataa ctggaactca gcagacacag aaggagacca 40

<210> 280 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 280 gtgtggtgcc agaagtaaac ctgagctagc agtccatgat 40

<210> 281 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 281 gtttactct ggcaccacac cagctgatag tggtagctgc 40

<210> 282 <211> 20 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for CpG-free LacZ
 <400> 282 atcatggact gctagctcag 20

<210> 283 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for R6K gamma M2A replication origin
 <400> 283 gcaggactga ggcttaatta aaccttaaaa cctttaaag 40

<210> 284 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for R6K gamma M2A replication origin
 <400> 284 ccttatatat tctttttttt ctataaaac ttaaacctt 40

<210> 285 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for R6K gamma M2A replication origin
 <400> 285 agaggctatt taagtgctg atttatatta attttattgt 40

<210> 286 <211> 50 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for R6K gamma M2A replication origin
 <400> 286 tcaaacatga gagcttagta catgaacat gagagcttag tacattagcc 50

<210> 287 <211> 60 <212> DNA <213> Artificial Sequence
 <220> <223> assembling oligo for R6K gamma M2A replication origin
 <400> 287 atgagagctt agtacattag ccatgagggt ttagtcatt aaacatgaga gcttagtaca 60

<210> 288 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for R6K gamma M2A replication origin
<400> 288 ttaaacaatga gagcttagta catgaaacat gagagcttag tacatactat caacaggtg 60

<210> 289 <211> 30 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for R6K gamma M2A replication origin

<400> 289 aactgctgat ctaattaac ctggagactt 30

<210> 290 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for R6K gamma M2A replication origin

<400> 290 aagtctccag gtaattaag atcagcagtt caacctgttg atagtatga ctaagctctc 60

<210> 291 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for R6K gamma M2A replication origin

<400> 291 atgtttcatg tactaagctc tcattgttaa tgtactaagc tctcatgtt aatgaactaa 60

<210> 292 <211> 60 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for R6K gamma M2A replication origin

<400> 292 accctcatgg ctaatgtact aagctctcat ggctaataa ctaagctctc atgtttcatg 60

<210> 293 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for R6K gamma M2A replication origin

<400> 293 tactaagctc tcattgttga acaataaaat taatataaat 40

<210> 294 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for R6K gamma M2A replication origin

<400> 294 cagcaactta aatagctctt aaggttttaa gttttataag 40

<210> 295 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for R6K gamma M2A replication origin

<400> 295 aaaaaaaga atatataagg cttttaagg tttaaggtt 40

<210> 296 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> assembling oligo for R6K gamma M2A replication origin

<400> 296 taattaagcc tcagtctgc 20

<210> 297 <211> 66 <212> DNA <213> Artificial Sequence

<220> <223> EM7 promotor

<400> 297 caattaatca tcggcatagt atatcgcat agtataatac gactcactat aggaaggcca 60

<210> 298 <211> 66 <212> DNA <213> Artificial Sequence

<220> <223> degenerated oligo for assembling the GpG-Free EM2K promotor

<400> 298 caattaawca tdggcatagt atatcwgcat agtataatac hactcactat aggaggcca 60

<210> 299 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> sense oligo for the assembly of SEQ ID NO: 38-74 oligos

<400> 299 ttcagctgag gaggcacatc 20

<210> 300 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> reverse oligo for the assembly of SEQ ID NO: 38-74 oligos

<400> 300 ctcaggatcc gctagctaat 20

<210> 301 <211> 18 <212> DNA <213> Artificial Sequence

<220> <223> sense primer pur 24

<400> 301 aggaccatca tgactgag 18

<210> 302 <211> 18 <212> DNA <213> Artificial Sequence

<220> <223> reverse primer pur 25

<400> 302 atcatgtcga gctagctc 18

<210> 303 <211> 31 <212> DNA <213> Artificial Sequence

<220> <223> RK15 primer

<400> 303 gcaggactga ggcttaatta aaccttaaaa c 31

<210> 304 <211> 31 <212> DNA <213> Artificial Sequence

<220> <223> RK16 primer

<400> 304 aagtctccag gttaattaag atcagcagtt c 31

<210> 305 <211> 52 <212> DNA <213> Artificial Sequence

<220> <223> rps0-1 linker oligo

<400> 305 ctagctgagt tcagaaaag ggggcctgag tggcccttt ttcaactta at 52

<210> 306 <211> 46 <212> DNA <213> Artificial Sequence

<220> <223> rps0-2 linker oligo

<400> 306 taagttgaaa aaaggggcca ctcaggcccc cttttctgaa actcag	46
<210> 307 <211> 39 <212> DNA <213> Artificial Sequence	
<220> <223> PACI-UP primer	
<400> 307 atcgtaatt aaaacagtag ttgacaatta aacattggc	39
<210> 308 <211> 28 <212> DNA <213> Artificial Sequence	
<220> <223> PACI-DOWN primer	
<400> 308 atcgtaatt aagttgaaaa aaggggcc	28
<210> 309 <211> 17 <212> DNA <213> Artificial Sequence	
<220> <223> rbs-1 linker oligo	
<400> 309 aattctgagg agaagct	17
<210> 310 <211> 17 <212> DNA <213> Artificial Sequence	
<220> <223> rbs-2 linker oligo	
<400> 310 catgagcttc tectcag	17
<210> 311 <211> 36 <212> DNA <213> Artificial Sequence	
<220> <223> OL dcm AF primer	
<400> 311 tttgcggcc gcttgctgcg ccagcaacta ataacg	36
<210> 312 <211> 41 <212> DNA <213> Artificial Sequence	
<220> <223> OL dcm AR primer	
<400> 312 ccttgatcc tggtaaaca gcactgtccg ccaatcgatt c	41
<210> 313 <211> 30 <212> DNA <213> Artificial Sequence	
<220> <223> OL dcm BF primer	
<400> 313 ttttgatcc tcagcaagag gcacaacatg	30
<210> 314 <211> 35 <212> DNA <213> Artificial Sequence	
<220> <223> OL dcm BR primer	
<400> 314 ttttctcgag aaacggcagc tctgatactt gcttc	35
<210> 315 <211> 32 <212> DNA <213> Artificial Sequence	
<220> <223> OL dcm CF primer	
<400> 315 ttttgcggcc gcgttgcggt attacccttg tc	32

<210> 316 <211> 795 <212> DNA <213> Artificial Sequence

<220> <223> Neo DeltaCpG

<220> <221> CDS <222> (1)..(795) <223> neo gene without CpG

<400> 316 atg att gaa caa gat ggc cta cat gca ggt tct cca gct gcc tgg gtt 48 Met Ile Glu Gln Asp Gly Leu His Ala
Gly Ser Pro Ala Ala Trp Val 1 5 10 15

gag aga ctg ttt ggc tat gac tgg gca cag cag acc att ggt tgc tct 96 Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln
Thr Ile Gly Cys Ser
20 25 30

gat gca gca gtt ttc aga ctt tca gcc caa ggc agg cca gtc ctt ttt 144 Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly
Arg Pro Val Leu Phe
35 40 45

gta aag aca gac ctc agt ggg gct ctc aat gag ctc cag gat gag gct 192 Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu
Leu Gln Asp Glu Ala
50 55 60

gcc aga ctc tcc tgg ttg gca aca act ggg gtc ccc tgt gca gct gtc 240 Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val
Pro Cys Ala Ala Val 65 70 75 80

ctt gat gtg gtc aca gaa gct gga agg gac tgg ctc cta cta ggt gag 288 Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp
Leu Leu Leu Gly Glu
85 90 95

gtg cct ggg cag gac ctc ctt tcc tct cac cta gct cca gct gag aaa 336 Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu
Ala Pro Ala Glu Lys
100 105 110

gtg tca atc atg gct gat gcc atg aga aga ctc cac acc ctt gac cca 384 Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu
His Thr Leu Asp Pro
115 120 125

gcc acc tgc ccc ttt gac cac cag gcc aag cac agg ata gag agg gcc 432 Ala Thr Cys Pro Phe Asp His Gln Ala Lys
His Arg Ile Glu Arg Ala
130 135 140

aga acc agg atg gag gct ggc ctg gtg gac caa gat gac ttg gat gaa 480 Arg Thr Arg Met Glu Ala Gly Leu Val Asp
Gln Asp Asp Leu Asp Glu 145 150 155 160

gaa cac cag ggc ctg gcc cct gct gaa cta ttt gcc agg ctc aag gca 528 Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe
Ala Arg Leu Lys Ala
165 170 175

tcc atg cca gat ggt gag gac cta gtg gtg act cat ggg gat gcc tgc 576 Ser Met Pro Asp Gly Glu Asp Leu Val Val Thr
His Gly Asp Ala Cys
180 185 190

ctt ccc aac atc atg gtt gaa aat gga agg ttc tct ggc ttc ata gac 624 Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe
Ser Gly Phe Ile Asp
195 200 205

tgt ggc agg ctg gga gtg gct gac agg tac cag gac att gcc cta gca 672 Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln
Asp Ile Ala Leu Ala
210 215 220

acc agg gac ata gca gaa gag cta ggg gga gag tgg gca gac agg ttc 720 Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly
Glu Trp Ala Asp Arg Phe 225 230 235 240

cta gtg ctc tat ggc att gca gcc cct gac tcc cag aga att gcc ttc 768 Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln
Arg Ile Ala Phe
245 250 255

tac aga ctt ctt gat gag ttc ttc taa
260

795 Tyr Arg Leu Leu Asp Glu Phe Phe

<210> 317 <211> 264 <212> PRT <213> Artificial Sequence

<220> <223> Neo DeltaCpG

<400> 317

Met Ile Glu Gln Asp Gly Leu His Ala Gly Ser Pro Ala Ala Trp Val 1 5 10 15

Glu Arg Leu Phe Gly Tyr Asp Trp Ala Gln Gln Thr Ile Gly Cys Ser
20 25 30

Asp Ala Ala Val Phe Arg Leu Ser Ala Gln Gly Arg Pro Val Leu Phe
35 40 45

Val Lys Thr Asp Leu Ser Gly Ala Leu Asn Glu Leu Gln Asp Glu Ala
50 55 60

Ala Arg Leu Ser Trp Leu Ala Thr Thr Gly Val Pro Cys Ala Ala Val 65 70 75 80

Leu Asp Val Val Thr Glu Ala Gly Arg Asp Trp Leu Leu Leu Gly Glu
85 90 95

Val Pro Gly Gln Asp Leu Leu Ser Ser His Leu Ala Pro Ala Glu Lys
100 105 110

Val Ser Ile Met Ala Asp Ala Met Arg Arg Leu His Thr Leu Asp Pro
115 120 125

Ala Thr Cys Pro Phe Asp His Gln Ala Lys His Arg Ile Glu Arg Ala
130 135 140

Arg Thr Arg Met Glu Ala Gly Leu Val Asp Gln Asp Asp Leu Asp Glu 145 150 155
160

Glu His Gln Gly Leu Ala Pro Ala Glu Leu Phe Ala Arg Leu Lys Ala
165 170 175

Ser Met Pro Asp Gly Glu Asp Leu Val Val Thr His Gly Asp Ala Cys
180 185 190

Leu Pro Asn Ile Met Val Glu Asn Gly Arg Phe Ser Gly Phe Ile Asp
195 200 205

Cys Gly Arg Leu Gly Val Ala Asp Arg Tyr Gln Asp Ile Ala Leu Ala
210 215 220

Thr Arg Asp Ile Ala Glu Glu Leu Gly Gly Glu Trp Ala Asp Arg Phe 225 230 235 240

Leu Val Leu Tyr Gly Ile Ala Ala Pro Asp Ser Gln Arg Ile Ala Phe
245 250 255

Tyr Arg Leu Leu Asp Glu Phe Phe
260

<210> 318 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 318 cattaccggt aggcacatca tgattgaaca agatggccta 40

<210> 319 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 319 catgcaggtt ctccagctgc ctgggttgag agactgtttg 40

<210> 320 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 320 gctatgactg ggcacagcag accattgggt gctctgatgc 40

<210> 321 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 321 agcagttttc agactttcag cccaaggcag gccagtcctt 40

<210> 322 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 322 ttgttaaaga cagacctcag tggggctctc aatgagctcc 40

<210> 323 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 323 aggatgagge tgccagactc tctggttg caacaactgg 40

<210> 324 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 324 ggtcccctgt gcagctgtcc ttgatgtggt cacagaagct 40

<210> 325 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 325 ggaagggact ggctcctact aggtgaggtg cctgggcagg 40

<210> 326 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> oligo for construction of Neo DeltaCpG
 <400> 326 acctccttc ctctcaccta gctccagctg agaaagtgc 40

<210> 327 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> oligo for construction of Neo DeltaCpG
 <400> 327 aatcatggct gatgccatga gaagactcca cacccttgac 40

<210> 328 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> oligo for construction of Neo DeltaCpG
 <400> 328 ccagccacct gccctttga ccaccaggcc aagcacagga 40

<210> 329 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> oligo for construction of Neo DeltaCpG
 <400> 329 tagagagggc cagaaccagg atggaggctg gcctggtgga 40

<210> 330 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> oligo for construction of Neo DeltaCpG
 <400> 330 ccaagatgac ttggatgaag aacaccaggc cctggccct 40

<210> 331 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> oligo for construction of Neo DeltaCpG
 <400> 331 gctgaactat ttgccaggct caaggcatcc atgccagatg 40

<210> 332 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> oligo for construction of Neo DeltaCpG
 <400> 332 gtgaggacct agtggtgact catggggatg cctgccttc 40

<210> 333 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> oligo for construction of Neo DeltaCpG
 <400> 333 caacatcatg gttgaaaatg gaaggttctc tggcttcata 40

<210> 334 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> oligo for construction of Neo DeltaCpG
 <400> 334 gactgtggca ggctgggagt ggctgacagg taccaggaca 40

<210> 335 <211> 40 <212> DNA <213> Artificial Sequence
 <220> <223> oligo for construction of Neo DeltaCpG

<400> 335 ttgccctagc aaccaggac atagcagaag agctaggggg	40
<210> 336 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 336 agagtgggca gacaggttcc tagtgctcta tggcattgca	40
<210> 337 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 337 gccctgact cccagagaat tgccttctac agacttcttg	40
<210> 338 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 338 atgagttctt ctaaagctag ctgacctga tagctgttcg	40
<210> 339 <211> 20 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 339 cgaacagcta tcaggatcag	20
<210> 340 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 340 ctactttag aagaactcat caagaagtct gtagaaggca	40
<210> 341 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 341 attctctggg agtcaggggc tgcaatgccca tagagcacta	40
<210> 342 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 342 ggaacctgtc tgcccactct cccctagct cttctgtat	40
<210> 343 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 343 gtccttggtt gctagggcaa tgtcctggta cctgtcagcc	40
<210> 344 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 344 actcccagcc tgccacagtc tatgaagcca gagaaccttc	40

<210> 345 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 345 cattttcaac catgatgttg ggaaggcagg catcccatg	40
<210> 346 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 346 agtcaccact aggtcctcac catctggcat ggatgccttg	40
<210> 347 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 347 agcctggcaa atagttcagc aggggccagg ccctggtgtt	40
<210> 348 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 348 cttcatccaa gtcattcttg tccaccaggc cagcctccat	40
<210> 349 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 349 cctggttctg gcctctcta tctgtgctt ggccgtgtgg	40
<210> 350 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 350 tcaaaggggc aggtggctgg gtcaagggtg tggagtcttc	40
<210> 351 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 351 tcatggcatc agccatgatt gacactttet cagctggagc	40
<210> 352 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 352 tagtgagag gaaaggaggt cctgccagg cacctcacct	40
<210> 353 <211> 40 <212> DNA <213> Artificial Sequence	
<220> <223> oligo for construction of Neo DeltaCpG	
<400> 353 agtaggagcc agtccttcc agcttctgtg accacatcaa	40
<210> 354 <211> 40 <212> DNA <213> Artificial Sequence	

<220> <223> oligo for construction of Neo DeltaCpG

<400> 354 ggacagctgc acaggggacc ccagttgtg ccaaccagga 40

<210> 355 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 355 gagtctggca gcctcatcct ggagtcatt gagagcccca 40

<210> 356 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 356 ctgaggtctg tctttacaaa aaggactggc ctgccttggg 40

<210> 357 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 357 ctgaaagtct gaaaactgct gcatcagagc aaccaatggt 40

<210> 358 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 358 ctgctgtgcc cagtcatagc caaacagtct ctcaaccag 40

<210> 359 <211> 40 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 359 gcagctggag aacctgcatg taggcatct tgttcaatca 40

<210> 360 <211> 20 <212> DNA <213> Artificial Sequence

<220> <223> oligo for construction of Neo DeltaCpG

<400> 360 tgatgtgcct accgtaatg 20