

Description

Genetically modified microorganism and process for production of macrolide compound using the microorganism

Field of the invention

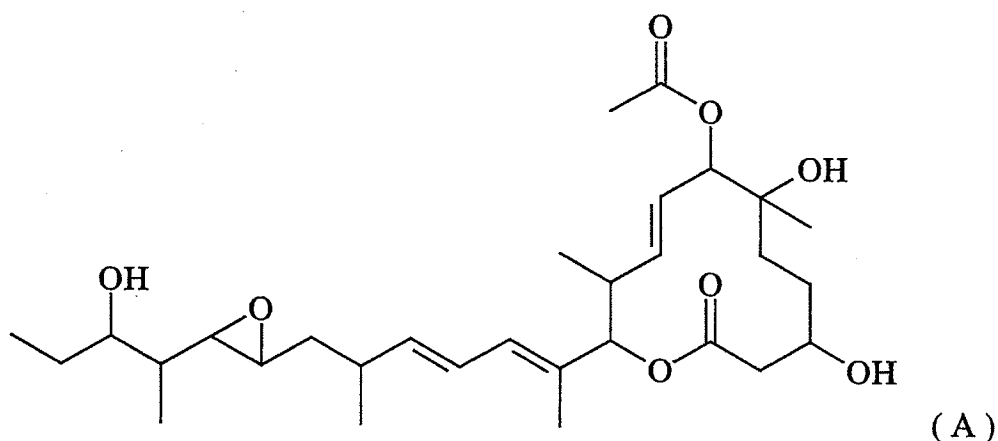
The present invention relates to recombinant microorganisms having the ability to produce 16-position hydroxyl macrolide compounds, and to methods for producing 16-position hydroxyl macrolide compounds using those microorganisms.

Related art

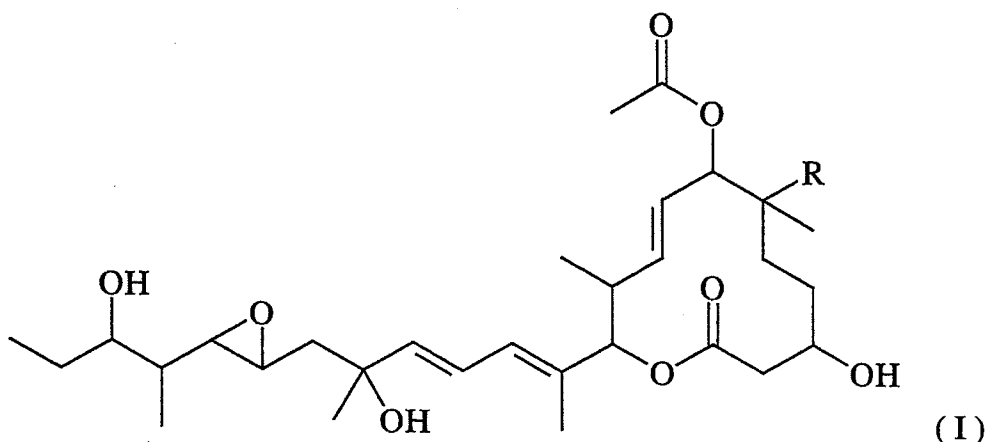
Important bioactive substances have been discovered among the various metabolites produced by the actinomycetes. In particular, many compounds have been found having a polyketide in the basic structure (hereunder called polyketide compounds). Compounds having various biological functions are known including for example the antibacterial agents erythromycin, josamycin, tylosin, midecamycin and mycinamicin, the antifungal agents nystatin and amphotericin, the insecticidal agents milbemycin and avermectin, the immune suppressors tacrolimus and rapamycin and the anti-tumor agents daunomycin, adriamycin, aclacinomycin and the like.

The polyketide compounds include a group of macrolide

compounds with excellent anti-tumor activity called pladienolides. "Pladienolide" is a general term for a group of compounds discovered in culture of the *Streptomyces* sp. Mer-11107 strain, and more than 50 relatives are known, beginning with 11107B (also called pladienolide B), which is represented by the formula (A) below (See WO 2002/060890).



Of these pladienolides, the 16-position hydroxyl substance represented by the following formula (I) has excellent properties including strong anti-tumor activity and the like, but it has been difficult to produce it efficiently because its productivity is inferior to that of 11107B.



(wherein R is a hydrogen atom or hydroxyl group).

The compound represented by the formula (I) may be called ME-282 when R is a hydrogen atom and 11107D or pladienolide D when R is a hydroxyl group.

A research group including the applicants in the present application has subsequently discovered actinomycetes which hydroxylates the 16-position of 11107B but which are not pladienolide-producing organisms (see WO 2003/099813 and WO 2004/050890), have obtained and identified DNA participating in a 16-position hydroxylation reaction of 11107B (DNA encoding for 16-position hydroxylating enzyme and those ferredoxins) from these actinomycetes, and have submitted a patent application therefor (see WO-A 2005/052152 published on June 9, 2005).

We have also obtained and identified from the pladienolide-producing strain *Streptomyces* sp. Mer-11107 polypeptides participating in biosynthesis of 11107B and DNA encoding those polypeptides, and have submitted a patent application therefor (see WO-A 2006/009276 published

on January 26, 2006).

It has thus become possible to produce 16-position hydroxyl substances more efficiently than before by producing 11107B or the like using the *Streptomyces* sp. Mer 11107 strain or variants thereof and then converting the resulting 11107B into 16-position hydroxyl substances using other actinomycetes capable of the 16-position hydroxylation reaction. However, because of the two-stage reaction, the operational difficulty is doubled, and it has been therefore desired to establish a more efficient production method.

Cases are known in which a gene coding for an enzyme which modifies a polyketide compound has been introduced so as to be expressed in a strain producing the polyketide compound, resulting in the direct production of a polyketide compound modified by an introduced enzyme. For example, it has been reported that direct production of 6-hydroxy-tetracenomycin C can be achieved by introducing the hydroxylation gene of *Streptomyces fradiae* Tu2717 into *Streptomyces glaucescens* GLA.0 that produces tetracenomycin C (see *J. Bacteriol.* 1995, Vol. 177, 6126-6136).

Cases are also known in which a polyketide compound biosynthesis gene has been introduced so as to be expressed in a strain having an enzyme which modifies the polyketide compound, resulting in direct production of a polyketide compound modified by an enzyme originally present in the

host strain. For example, it has been reported that direct production of 6-hydroxy-tetracenomycin C can be achieved by introducing the tetracenomycin C biosynthesis gene of *Streptomyces glaucescens* GLA.0 into *Streptomyces fradiae* Tu2717, which possesses an enzyme that hydroxylates the 6 position of tetracenomycin C (see *J. Bacteriol.* 1995, Vol. 177, 6126-6136).

In some cases, production of a modified polyketide compound can be achieved with a strain of a different species by introducing the polyketide compound biosynthesis gene and a gene encoding an enzyme which modifies the polyketide compound together into a strain of a different species. For example, it has been reported that direct production of 8,8a-dihydroxy-6-erythronolide B was achieved by introducing both a 6-erythronolide B biosynthesis gene of *Saccharopolyspora erythraea* and a P-450 gene of *Streptomyces antibioticus* ATCC11891 into *Streptomyces lividans* K4-114 as the host (see *J. Antibiot.* 2000, Vol. 53, 502-508).

Disclosure of the invention

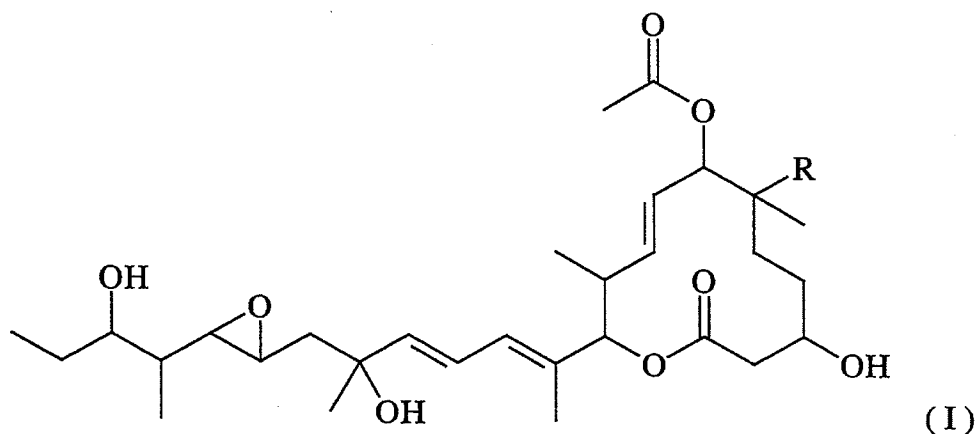
It is an object of the present invention to provide recombinant microorganisms having the ability to produce the 16-position hydroxyl macrolide compound represented by the above formula (I). Another object is to provide a method of producing a 16-position hydroxyl macrolide

compound using these recombinant microorganisms.

In an effort to resolve the aforementioned problems, the inventors used genetic engineering methods to prepare (breed) microorganisms having both the previously isolated and identified DNA participating in biosynthesis of 11107B and DNA participating in a 16-position hydroxylation reaction of 11107B, and they have found that a 16-position hydroxyl macrolide compound accumulated in large quantities in the resulting culture broth. Thus, they accomplished the present invention.

That is, the present invention relates to the following (1) to (21).

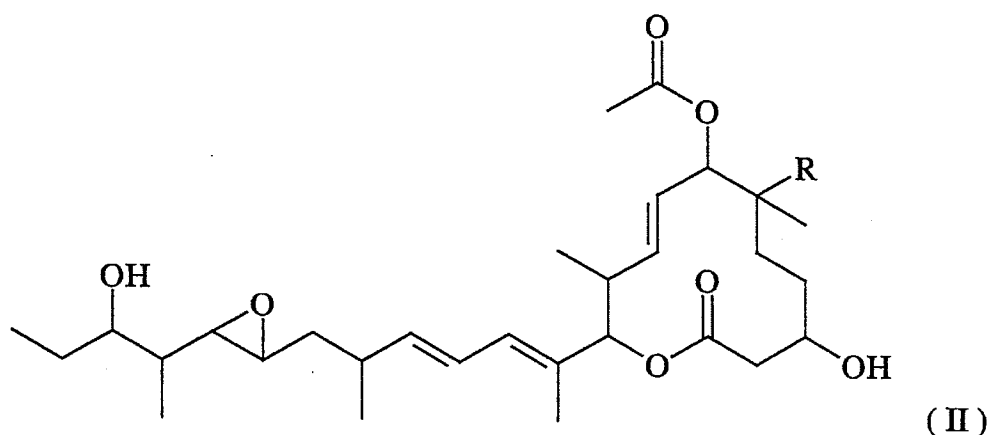
(1) A recombinant microorganism, which is a microorganism having ability to produce 16-position hydroxyl macrolide compounds represented by the formula (I):



(wherein R represents a hydrogen atom or hydroxyl group), and, which comprises:

(a) DNA encoding wholly or partially a polypeptide

participating in biosynthesis of the macrolide compounds represented by the formula (II):



(wherein R represents a hydrogen atom or hydroxyl group);
and

(b) DNA encoding wholly or partially a polypeptide having 16-position hydroxylating enzyme activity with respect to the macrolide compounds represented by the above formula (II).

(2) The recombinant microorganism described in (1), wherein the DNA described in the above (a) is DNA encoding wholly or partially a polypeptide having polyketide synthetase activity, a polypeptide having 7-position acetylase activity, a polypeptide having 18,19-position epoxidase activity and a polypeptide having transcriptional regulator factor activity.

(3) The recombinant microorganism described in (1) or (2), wherein the DNA described under (a) above is DNA comprising:

(a-1) the continuous nucleotide sequence of nucleotides

8340 to 27935 in the sequence No. 1;
(a-2) the continuous nucleotide sequence of nucleotides 28021 to 49098 in the sequence No. 1;
(a-3) the continuous nucleotide sequence of nucleotides 49134 to 60269 in the sequence No. 1;
(a-4) the continuous nucleotide sequence of nucleotides 60269 to 65692 in the sequence No. 1;
(a-5) the continuous nucleotide sequence of nucleotides 68160 to 66970 in the sequence No. 1;
(a-6) the continuous nucleotide sequence of nucleotides 69568 to 68270 in the sequence No. 1; and
(a-7) the continuous nucleotide sequence of nucleotides 72725 to 70020 in the sequence No. 1, or a variant thereof.

(4) The recombinant microorganism described in (2) or (3), wherein the DNA described under (a) above further includes DNA encoding wholly or partially a polypeptide having 6-position hydroxylating enzyme activity.

(5) The recombinant microorganism described in (4), wherein the DNA encoding wholly or partially a polypeptide having 6-position hydroxylating enzyme activity is DNA comprising the continuous nucleotide sequence of nucleotides 65707 to 66903 in the sequence No. 1, or a variant thereof.

(6) The recombinant microorganism described in any of (1) to (5), wherein the DNA described in the above (b) above is DNA selected from the group consisting of:

- (b-1) the continuous nucleotide sequence of nucleotides 1322 to 2548 in the sequence No. 2;
- (b-2) the continuous nucleotide sequence of nucleotides 420 to 1604 in the sequence No. 3; and
- (b-3) the continuous nucleotide sequence of nucleotides 172 to 1383 in the sequence No. 4, or a variant thereof.

(7) The recombinant microorganism described in any one of (1) to (6), wherein the DNA represented by (b) above is a variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the sequence No. 2 which has 1 or 2 or more modified sites selected from the group consisting of 1) to 6) below:

- 1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding an amino acid other than arginine;
- 2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding an amino acid other than methionine;
- 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding an amino acid other than arginine;
- 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding an amino acid other than arginine;
- 5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding an amino acid

other than isoleucine; and

6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding an amino acid other than leucine.

(8) The recombinant microorganism described in any one of (1) to (6), wherein the DNA represented by (b) above is a variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the sequence No. 2 which has 1 or 2 or more modified sites selected from the group consisting of 1) to 6) below:

1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding cysteine, methionine or proline;

2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding threonine or serine;

3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding leucine, isoleucine, proline, tyrosine or phenylalanine;

4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding leucine, isoleucine or proline;

5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding phenylalanine or valine; and

6) a modified site wherein the sequence ctg of nucleotides

2528 to 2530 is modified to a codon encoding methionine, cysteine or isoleucine.

(9) The recombinant microorganism described in any one of (1) to (6), wherein the DNA represented by (b) above is a variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the sequence No. 2 which has 1 or 2 or more modified sites selected from the group consisting of 1) to 6) below:

- 1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding cysteine;
 - 2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding threonine;
 - 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding leucine or tyrosine;
 - 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding leucine;
 - 5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding phenylalanine;
- and
- 6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding methionine.

(10) The recombinant microorganism described in (1), comprising a host having the DNA described under (a) above into which has been incorporated the DNA described under (b) above derived from a microorganism of a different

species.

(11) The recombinant microorganism described in (1), comprising a host having the DNA described under (b) above into which has been incorporated the DNA described under (a) above derived from a different species.

(12) The recombinant microorganism described in (1), comprising a host having none of the DNA described under (a) and (b) above into which has been incorporated the DNA described under both (a) and (b) above derived from different species.

(13) The recombinant microorganism described in any of (1) to (12), wherein the recombinant microorganism having the ability to produce the 16-position hydroxyl macrolide compound represented by the formula (I) is a strain belonging to *Streptomyces* genus.

(14) The recombinant microorganism described in any of (1) to (13), wherein 50 mg or more of the 16-position hydroxyl macrolide compound represented by the formula (I) per 1 liter of culture broth can be produced when the recombinant microorganism is inoculated into a 250 mL Erlenmeyer flask containing 30 mL of the following medium, cultured at 25°C for 4 days with rotary shaking (220 rpm) and extracted by addition of 270 mL of acetonitrile, and the resulting extract is analyzed by HPLC under the following measurement conditions to assay the amount of the 16-position hydroxyl macrolide compound represented by the

formula (I).

(Medium)

Soluble starch 5%, glucose 1%, Pharmamedia 3%, CaCO_3 0.1%, pH 7.5

(HPLC measurement conditions)

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (ϕ 4.6 mm x 50 mm 3 μm)

Mobile phase: 45%-55% methanol (0-5 min), 55% methanol (5-13 min), 55% to 70% methanol (13-21 min), 70% methanol (21-25 min)

Flow rate: 1.2 mL/min

Detection: UV 240 nm

Injection volume: 5 μL

Column temperature: 40°C

(15) A method of producing 16-position hydroxyl macrolide compounds represented by the formula (I) or pharmacologically acceptable salts thereof, or hydrates of them, wherein the recombinant microorganism described in any of (1) to (14) is cultured in nutritious medium, and the 16-position hydroxyl macrolide compounds represented by the formula (I) are collected from the culture broth thereof.

(16) The method described in (15), wherein a cyclodextrin is present in the culture broth.

(17) The method described in (16), wherein the cyclodextrin is a cyclodextrin selected from the group

consisting of β -cyclodextrin, γ -cyclodextrin, partially methylated β -cyclodextrin, dimethyl- β -cyclodextrin, trimethyl- β -cyclodextrin, glycosyl- β -cyclodextrin and hydroxypropyl- β -cyclodextrin.

(18) A variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the sequence No. 2 which have 1 or 2 or more modified sites selected from the group consisting of 1) to 6) below:

- 1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding an amino acid other than arginine;
- 2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding an amino acid other than methionine;
- 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding an amino acid other than arginine;
- 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding an amino acid other than arginine;
- 5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding an amino acid other than isoleucine; and
- 6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding an amino acid other than leucine.

(19) A variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the sequence No. 2 which have 1 or 2 or more modified sites selected from the group consisting of 1) to 6) below:

- 1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding cysteine, methionine or proline;
- 2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding threonine or serine;
- 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding leucine, isoleucine, proline, tyrosine or phenylalanine;
- 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding leucine, isoleucine or proline;
- 5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding phenylalanine or valine; and
- 6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding methionine, cysteine or isoleucine.

(20) A variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the sequence No. 2 which have 1 or 2 or more modified sites selected from the group comprising 1) to 6) below:

- 1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding cysteine;
 - 2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding threonine;
 - 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding leucine or tyrosine;
 - 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding leucine;
 - 5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding phenylalanine;
- and
- 6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding methionine.

(21) A polypeptide encoded by the variant of DNA described in any of (18) to (20).

The terms in the present specifications are defined as follows.

A "recombinant microorganism" is a microorganism (bacteria, actinomycete, yeast, mould or the like) comprising a gene derived from another organism introduced into a specific microorganism by genetic recombination, and the genetic introduction techniques used therefor are not limited to genetic recombination using plasmids and other vectors but include techniques of homologous recombination and the like.

"Variant of DNA" means:

- (1) DNA that hybridizes under stringent conditions with the original DNA;
- (2) DNA with a nucleotide sequence having 70% or greater homology with the nucleotide sequence of the original DNA;
- (3) DNA having a nucleotide sequence complementary to the nucleotide sequence of the original DNA; or
- (4) DNA which does not hybridize under stringent conditions with the original DNA due to degeneracy of a gene codon, but which has a nucleotide sequence coding for the same amino acid sequence as the amino acid sequence coded for by the DNA defined in any of (1) to (3).

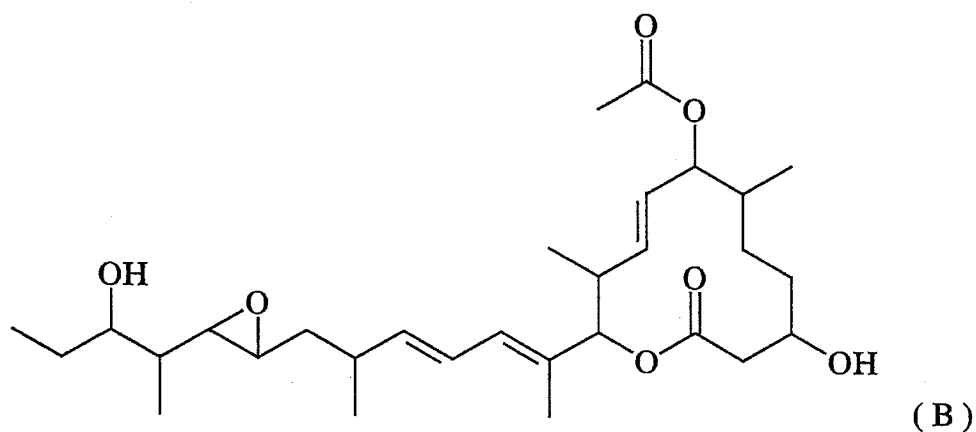
"DNA that hybridizes under stringent conditions" means for example DNA obtained by colony hybridization, plaque hybridization, Southern hybridization or the like using the original DNA as the probe, and specifically may be DNA which can be identified using a filter having DNA derived from a colony or plaque fixed thereon by first hybridizing at 65°C in the presence of 0.7 to 1.0 M sodium chloride, and then washing the filter at 65°C using an 0.1 to 2x concentration SSC solution (an SSC solution with a 1x concentration comprising 150 mM sodium chloride and 15 mM sodium citrate).

A "relative" is a compound that has the same framework characterizing the chemical structure but that differs in terms of the form of modification or the form of the side

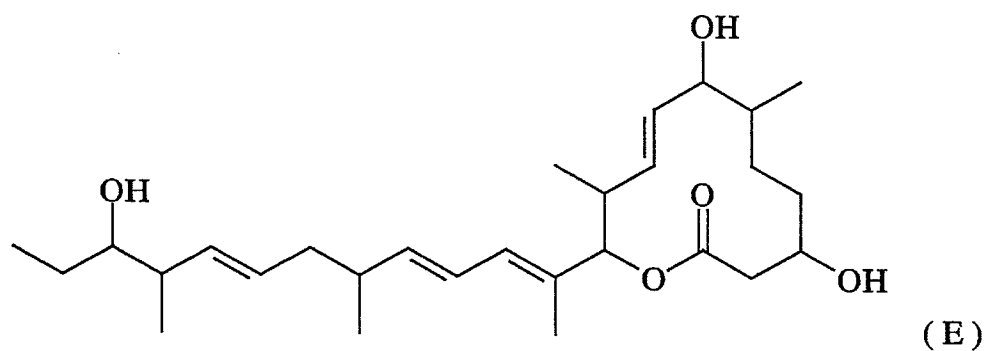
chain.

"16-position hydroxylating enzyme activity" means the activity of converting the 16-position hydrogen atom of macrolide compound (II) into a hydroxyl group.

"6-position hydroxylating enzyme activity" means the activity of converting the 6-position hydrogen atom of the macrolide compound (sometimes called ME-265) represented by the following formula (B) into a hydroxyl group.

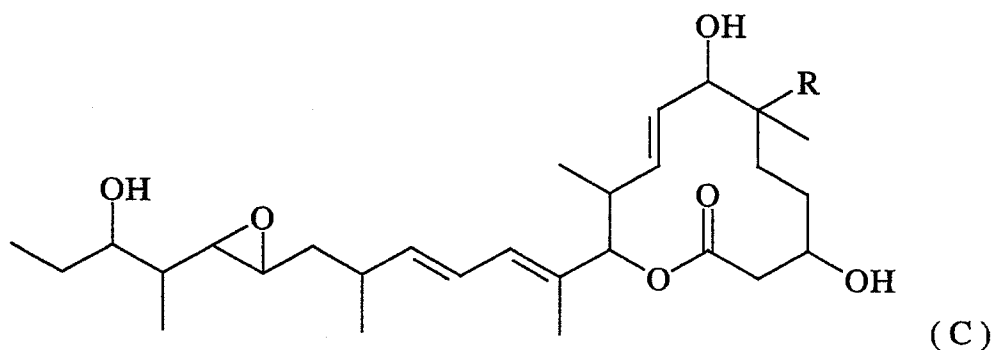


"Polyketide synthetase activity" means the activity of synthesizing the macrolide compound represented by the following formula (E):



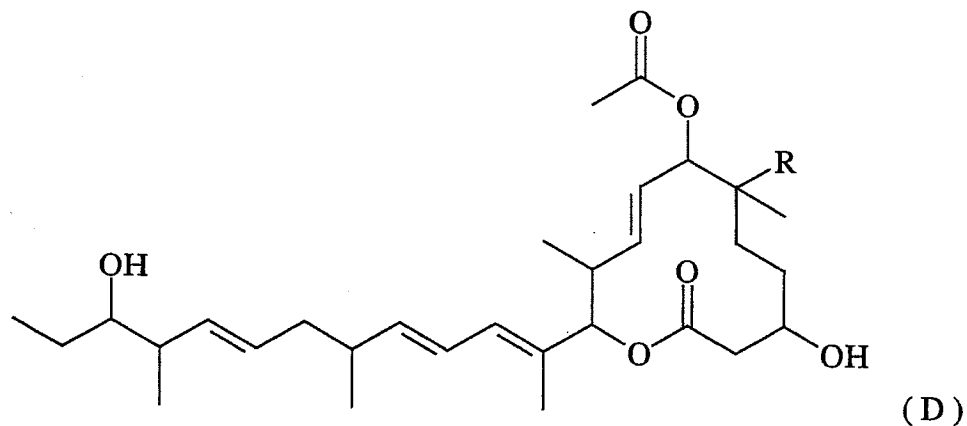
"7-position acetylase activity" means the activity of converting the 7-position hydroxyl group in the macrolide

compound represented by the following formula (C) into an acetyl group.



(wherein R represents a hydrogen atom or hydroxyl group).

"18,19-position epoxidase activity" means the activity of converting an epoxy group for the 18,19-position double bond in the macrolide compound represented by the following formula (D):



(wherein R represents a hydrogen atom or hydroxyl group).

"Transcription regulator activity" means the activity of regulating the transcription of DNA encoding a polypeptide participating in biosynthesis of macrolide compound (II).

According to the present invention, a recombinant

microorganism can be obtained having DNA encoding a polypeptide participating in biosynthesis of macrolide compound (II) and having DNA encoding a polypeptide having 16-position hydroxylating enzyme activity with respect to macrolide compound (II), and capable of direct fermentative production of 16-position hydroxyl macrolide compound (I). 16-position hydroxyl macrolide compound (I) can be produced efficiently by culturing that recombinant microorganism and collecting the compound from the culture broth.

Particularly, 16-position hydroxyl macrolide compound (I) can be produced more efficiently and selectively, by applying DNA encoding a polypeptide which has 16-position hydroxylating enzyme activity, in which polypeptide the amino acid sequence has been partially modified, to the above-mentioned method.

Brief description of the drawings

Fig. 1 shows the biosynthesis pathway of pladienolides in Mer-11107.

Fig. 2 shows the correspondence between cosmid and each of ORFs of DNA participating in biosynthesis of pladienolides in Mer-11107.

Fig. 3 shows the structure of plasmid pKU253.

Fig. 4 shows the structure of plasmid pUC19aph::oriT::intphiC31.

Detailed Description of the Invention

Embodiments of the present invention are explained in detail below.

DNA encoding a polypeptide participating in biosynthesis of macrolide compound (II) in the present invention can be obtained by a method known in the field (such as the colony hybridization method described in Molecular Cloning, 2nd Edition) from the cultured mycelia of a microorganism capable of producing that macrolide compound. This microorganism may be any capable of producing the macrolide compound, regardless of species or strain, but strains belonging to the *Streptomyces* genus of the actinomycetes are preferred. One example is the *Streptomyces* sp. Mer-11107 strain, which has been isolated from soil. Mer-11107 was deposited as FERM P-18144 at the National Institute of Bioscience and Human-Technology Agency of Industrial Science and Technology (1-3, Higashi 1-chome Tsukuba-shi, Ibaraki-ken 305-8566 Japan) as of December 19, 2000, and then transferred to International Deposit FERM BP-7812 at International Patent Organism Depositary (IPOD) National Institute of Advanced Industrial Science and Technology (Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken 305-8566 Japan) as of November 27, 2001.

The operations for obtaining the target DNA are outlined below using the Mer-11107 strain as an example.

Recombinant DNA obtained by connecting genomic DNA of Mer-11107 that has been partially digested with a suitable restriction enzyme together with a cosmid vector capable of replication in *E. coli* that has been digested with a suitable restriction enzyme is incorporated into *E. coli* to obtain transduced strains. The multiple transduced strains obtained above are then screened using a fragment of a known hydroxylating enzyme (cytochrome P450 enzyme) gene as the probe, and the transduced strains that hybridize to the probe are selected. DNA that hybridizes to the hydroxylating enzyme gene present in the selected cosmids is then obtained and sequenced. This is then introduced into *E. coli*, and once it has been confirmed that the transformed *E. coli* has 6-position hydroxylating enzyme activity, DNA coding for this 6-position hydroxylating enzyme can be used as the probe in Southern hybridization to select, cosmids comprising a macrolide compound biosynthesis gene cluster neighboring DNA coding for the 6-position hydroxylating enzyme from the multiple positive clones (cosmids) obtained above, and the target DNA is obtained by organizing these cosimids. Details of all these operations are described in Reference Example 1 and also in WO-A 2006/009276, specifications submitted by the applicants in this case, and these can be used for reference.

A nucleotide sequence obtained in this way (amino acid

sequence in the coding sequence) comprising the region surrounding of DNA coding for a polypeptide participating in biosynthesis of macrolide compound (II) is represented by the Sequence No. 1 in the sequence listing.

The DNA represented by the Sequence No. 1 includes 8 open reading frames (ORFs): pldA I (nucleotides 8340 to 27935), pldA II (nucleotides 28021 to 49098), pldA III (nucleotides 49134 to 60269), pldA IV (nucleotides 60269 to 65692), pldB (nucleotides 65707 to 66903), pldC (nucleotides 68160 to 66970), pldD (nucleotides 69568 to 68270) and pldR (nucleotides 72725 to 70020).

Of these DNAs, pldA I, pldA II, pldA III and pldA IV have several transcription reading frames each including one or more repeating units called modules in the same way as other known polyketide biosynthesis genes. As explained below, each module codes for all or some domains selected from the acyl carrier protein (ACP), β -ketoacyl ACP synthetase (KS), acyl transferase (AT) domains, which participate in condensation reactions in polyketide synthesis, and the ketoacyl reductase (KR), dehydrogenase (DH) and enoyl reductase (ER) domains, which participate in β -position carbonyl group modification reactions, and the final module includes the thioesterase (TE) domain, which removes the polyketide chain from the polyketide synthetase.

Fig. 1 shows the biosynthesis pathway of a macrolide compound in the Mer-11107 strain. Unlike the other modules

the loading module has the central cysteine replaced by glutamine, indicating that pldA I participates in the initial reaction. Module 10 includes a thioesterase (TE) domain, indicating that pldA IV participates in the final reaction of synthesizing the basic polyketide skeleton. Once the basic skeleton of the polyketide has been formed in this way, it is modified by the 18,19-position epoxidase coded for by pldD, the 7-position acetylase encoded by pldC and the 6-position hydroxylase encoded by pldB to biosynthesize macrolide compound (II). Macrolide compound (II) in which R is a hydrogen atom can be produced by either deleting pldB or modifying it so that it does not express its function. PldR, which has high homology with the aveR gene encoding a transcription regulator factor in avermectin biosynthesis, encodes a DNA transcription regulator factor participating in biosynthesis of macrolide compounds.

In the present invention, DNA encoding a polypeptide having 16-position hydroxylating enzyme activity with respect to macrolide compound (II) can be obtained by methods well known in the field (such as the colony hybridization method described in Molecular Cloning, 2nd Edition) from cultured mycelia of a microorganism having 16-position hydroxylation ability with respect to the macrolide compound. Any microorganism having the ability to convert the 16-position hydrogen atom of the macrolide

compound to a hydroxyl group can be used regardless of species or strain, but desirable examples include strains belonging to the actinomycetes. Examples of these include the Streptomyces sp. A-1544 strain and the Streptomyces sp. Mer-11107 strain and the Streptomyces sp. A-1560 strain. The A-1544 strain was deposited as FERM P-18943 at International Patent Organism Depositary National Institute of Advanced Industrial Science and Technology (Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken 305-8566 Japan) as of July 23, 2002, and then transferred to International Deposit FERM BP-8446 as of July 30, 2003, at International Patent Organism Depositary (IPOD) National Institute of Advanced Industrial Science and Technology (Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken 305-8566 Japan). The A-1560 strain was deposited as FERM P-19585 at International Patent Organism Depositary National Institute of Advanced Industrial Science and Technology (Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken 305-8566 Japan) as of November 13, 2003 and then transferred to International Deposit FERM BP-10102 as of August 19, 2004, at International Patent Organism Depositary (IPOD) National Institute of Advanced Industrial Science and Technology (Tsukuba Central 6, 1-1, Higashi 1-Chome, Tsukuba-shi, Ibaraki-ken 305-8566 Japan).

The operations for obtaining the target DNA are

outlined below using the A-1544 strain as an example. Genomic DNA of A-1544 is first prepared and subjected to PCR using primers prepared based on the sequence information for a known hydroxylating enzyme gene (cytochrome P450 enzyme) to obtain a specifically amplified DNA fragment which is then sequenced. The nucleotide sequence of the neighboring region upstream and downstream from the cloned DNA fragment is then amplified by an inverse PCR (Cell Technology Vol. 14, pp. 591-593, 1995) based on the resulting sequence and cloned, and the sequence is analyzed. Details of all these operations are described in Reference Example 2 (A-1544 strain), Reference Example 3 (Mer-11107 strain) and Reference Example 4 (A-1560 strain), and also in WO-A 2005/052152 published on June 9, 2005, and these can be used for reference.

The DNA (including neighboring region) thus obtained from A-1544 strain which encodes a polypeptide having 16-position hydroxylating enzyme activity with respect to macrolide compound (II) is represented by the Sequence No. 2 in the sequence listing. The DNA represented by the Sequence No. 2 includes two open reading frames (ORFs), psmA (nucleotides 1322 to 2548) and psmB (nucleotides 2564 to 2761), which encode 16-position hydroxylating enzyme and ferredoxin, respectively.

Similarly, DNA (including neighboring region) encoding a polypeptide having 16-position hydroxylating enzyme

activity with respect to macrolide compound (II) obtained from Mer-11107 is represented by the Sequence No. 3 in the sequence listing. This DNA represented by the Sequence No. 3 includes two open reading frames (ORFs), bpmA (nucleotides 420 to 1604) and bpmB (nucleotides 1643 to 1834), which encode 16-position hydroxylating enzyme and ferredoxin, respectively.

DNA (including neighboring region) encoding a polypeptide having 16-position hydroxylating enzyme activity with respect to macrolide compound (II) obtained from A-1560 is represented by the Sequence No. 4 in the sequence listing. This DNA represented by the Sequence No. 4 includes two open reading frames (ORFs), tpmA (nucleotides 172 to 1383) and tpmB (nucleotides 1399 to 1593), which encode 16-position hydroxylating enzyme and ferredoxin, respectively.

Further, variants of the above-mentioned various DNA which encodes a polypeptide having 16-position hydroxylating enzyme activity can be obtained by the following procedures. First, the site-directed modification is conducted, based on the original DNA sequence information, and then the resulting site-directed variant of DNA is introduced into the host. Thus, the objective variants can be selected and obtained by using the highly selective and active 16-position hydroxylating enzyme of the variant of DNA of the resulting recombinant

microorganism as the index. The site-directed modification, for example, can be conducted by using, as a template, a plasmid cloning the original DNA and using the primer constructed based on the sequence information thereof, using a commercially available kit such as QuickChange™ Site-Directed Mutagenesis Kit (Stratagene Co.).

Among these variants, as those having high 16-position hydroxylating enzyme activity and further high selectivity, for example, in the case of psmA encoding 16-position hydroxylating enzyme and being obtained from A-1544 strain, the variant of DNA having the continuous nucleotide sequence of nucleotides 1322 to 2548 in the Sequence No. 2 and 1 or 2 or more modified sites selected from the group consisting of the following 1) to 6) may be included:

- 1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding an amino acid, preferably cysteine, other than arginine;
- 2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding an amino acid, preferably threonine, other than methionine;
- 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding an amino acid, preferably leucine or tyrosine, other than arginine;
- 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding an amino acid, preferably leucine, other than arginine;

- 5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding an amino acid, preferably phenylalanine, other than isoleucine; and
- 6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding an amino acid, preferably methionine, other than leucine.

The recombinant of the present invention is a recombinant microorganism having the aforementioned "DNA encoding a polypeptide participating in biosynthesis of macrolide compound (II)" and "DNA encoding a polypeptide having 16-position hydroxylating enzyme activity with respect to macrolide compound (II)," and can be prepared by the methods shown under A) to C) below.

A) "DNA encoding a polypeptide having 16-position hydroxylating enzyme activity with respect to macrolide compound (II)" is incorporated into a host that originally has "DNA encoding a polypeptide participating in biosynthesis of macrolide compound (II)".

B) "DNA encoding a polypeptide participating in biosynthesis of macrolide compound (II)" is incorporated into a host that originally has "DNA encoding a polypeptide having 16-position hydroxylase activity with respect to macrolide compound (II)".

C) "DNA encoding a polypeptide participating in biosynthesis of macrolide compound (II)" and "DNA encoding a polypeptide having 16-position hydroxylating enzyme

activity with respect to macrolide compound (II)" are both incorporated into a host that originally possesses neither DNA.

The host for the recombinant of the present invention may be any microorganism capable of incorporating the target DNA and producing the target 16-position hydroxyl macrolide compound, but an example of a preferred microorganism is a strain belonging to the *Streptomyces* genus of actinomycetes. For example, *Streptomyces* sp. Mer-11107 or the like can be used in method A) above, while *Streptomyces* sp. A-1544, *Streptomyces* sp. Mer-11107 or *Streptomyces* sp. A-1560 can be used in method B) above. In method C), a strain of *Escherichia coli* or the like can be used.

There are no particular limits on the method of incorporating the target DNA and causing it to be expressed in a host, but for example the methods described in *Molecular Cloning*, 2nd Edition or *Current Protocols in Molecular Biology* can be used. There are no particular limits on the host and plasmid vector system as long as the target DNA can be stably maintained and expressed in the host. The plasmid may include self-replicating sequences, promoter sequences, terminator sequences, drug-resistance genes and the like in addition to the target DNA, and the plasmid may be either a self-replicating plasmid or an integrative plasmid having a sequence homologous to a

specific region in the genome of the anticipated host. The target DNA may be incorporated into a site either on the plasmid or on the genome of the host microorganism.

If the host is an actinomycete or related strain, the self-replicating vector pIJ6021 (*Gene*, 166, 133-137 (1995)) or the genome integration vector KC515 (*The Bacteria* Vol. 9, Antibiotic-producing *Streptomyces* (ed. Queener, S.E. and Day, L.E.), pp. 119-158, Academic Press, Orlando, Fla) can be used.

Examples of self-replicating sequences include rep pIJ101 (1988, *J. Bacteriol.*, 170, 4634-4651), rep SCP2 (1981, *J. Gen. Microbiol.*, 126, 427-442), pUC19 (*Gene*, 33 (1), 103-119 (1985)) and the like, while examples of promoter sequences include ermEp (*Mol. Microbiol.*, 1994, Nov; 14(3): 533-45), SnpAp (*J. Bacteriol.* 1992, May; 174(9): 2797-2808) and the like, examples of terminator sequences include ter(mmr) (*Gene*, 1987; 58(2-3): 229-41), fd terminator (*Nucleic Acids Res.*, 5(12), 4495-4503 (1978) and the like, and examples of drug-resistance genes include hyg (*Nucleic Acids Res.*, 4(14), 1565-1581 (1986)), tsr (*Mol. Gen. Genet.*, 199(1), 26-36 (1985)) and the like. The microorganism of the present invention can be easily prepared by a person skilled in the art following the descriptions of the present specification.

A recombinant microorganism prepared in this way can be cultured, and productivity of 16-position hydroxyl

macrolide compound (I) can be evaluated and selected for to obtain a strain with a high level of productivity of the macrolide compound. That is, a recombinant organism having the ability to produce 50 mg or more of the macrolide compound per 1 liter of culture broth can be obtained at a high rate of about 40%, while one having the ability to produce 100 mg or more per 1 liter of culture broth can be obtained at a high rate of about 30%.

(Method of evaluating productivity)

The recombinant microorganism is inoculated into a 250 mL Erlenmyer flask containing 30 mL of medium, (soluble starch 5%, glucose 1%, Pharmamedia 3%, CaCO_3 0.1%, pH 7.5), cultured with rotary shaking (220 rpm) for 4 days at 25°C and extracted by addition of 270 mL acetonitrile, and the resulting extract is analyzed by HPLC under the following measurement conditions to assay the concentration of accumulated 16-position hydroxyl macrolide compound (I).

(HPLC measurement conditions)

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (ϕ 4.6 mm $\times$ 50 mm 3 μm)

Mobile phase: 45%-55% methanol (0 to 5 min), 55% methanol (5 to 13 min), 55%-70% methanol (13 to 21 min), 70% methanol (21 to 25 min)

Flow rate: 1.2 mL/min

Detection: UV 240 nm

Injection volume: 5 μL

Column temperature: 40°C

The recombinant microorganism of the present invention can be inoculated into a nutrient medium and culture aerobically, and 16-position hydroxyl macrolide compound (I) or a pharmacologically acceptable salt thereof, or a hydrate of them can be obtained from the resulting culture broth.

In principle the recombinant microorganism can be cultured by ordinary culture methods for microorganisms, but culture under aerobic conditions such as culture with aeration and agitation or shaking culture in the case of a liquid culture is normally preferred. As the medium used for cultivation, any medium may be used insofar as it contains nutrients which these microorganisms can utilize, and various synthetic medium, semi-synthetic medium and organic medium may be all utilized. As the medium composition, various carbon sources such as glucose, sucrose, fructose, glycerin, dextrin, starch, molasses and soybean oil may be used either singly or in combinations. As the nitrogen source, there can be used a single or a combination of organic nitrogen sources such as Pharmamedia, peptone, meat extract, soybean meal, casein, amino acid, yeast extract or urea, and inorganic nitrogen sources such as sodium nitrate or ammonium sulfate. Additionally, for example, there can be added and used salts such as sodium chloride, potassium chloride, calcium carbonate, magnesium

sulfate, sodium phosphate, potassium phosphate or cobalt chloride; heavy metal salts; vitamins such as vitamin B or biotin, if necessary. Further, when foaming is remarkable during cultivation, various antifoaming agents can be appropriately added in the medium. Chemicals necessary for maintaining drug-resistant genes may also be added as necessary. When the antifoaming agent is added, it is required to set at a concentration for not adversely affecting the production of an objective substance.

The concentration of the accumulated 16-position hydroxyl macrolide compound can be increased by addition of cyclodextrins in the culture broth when producing 16-position hydroxyl macrolide compound (I) by cultivation of the recombinant microorganism of the present invention. Examples of cyclodextrins to be used in the present invention include β -cyclodextrin, γ -cyclodextrin, partially methylated β -cyclodextrin, dimethyl- β -cyclodextrin, trimethyl- β -cyclodextrin, glycosyl- β -cyclodextrin and hydroxypropyl- β -cyclodextrin. Each of them may be used singly or in a combination with another.

The amount of cyclodextrin added to the culture broth is preferably 0.5% to 5%. Addition is preferably at the beginning of cultivation or during cultivation.

The culture conditions can be selected appropriately within the range at which the microorganism grows well and can produce 16-position hydroxyl macrolide compound (I).

For example, the pH of the medium is kept at 5 to 9 or near normally the neutral range. The culture temperature is normally 20 to 40°C or preferably 23 to 35°C. Cultivation is 2 to 12 days or normally 3 to 10 days. Of course, the various culture conditions described above can be changed and optimal conditions selected appropriately depending on the type of microorganism used, the external conditions and the like. Common methods used for isolation and purification of microbial metabolites from culture broth can be used to isolate and purify the 16-position hydroxyl macrolide compound (I) which accumulates in the culture broth. Possible methods include all known methods such as organic solvent extraction with methanol, ethanol, butanol, ethyl acetate, chloroform, acetone, toluene or the like, various forms of ion-exchange chromatography, gel filtration chromatography using Sephadex LH-20 or the like, adsorption chromatography using activated charcoal, silica gel, adsorbent resin (Diaion HP20) or the like, adsorption-desorption by thin layer chromatography and high-performance liquid chromatography using a reverse-phase column and the like, and are not limited to the methods indicated herein. These methods may be used singly or in combinations of two or more in an optional order or repeatedly, which makes it possible to isolate and purify the target 16-position hydroxyl macrolide compound (I).

Examples

The present invention is explained in detail below using examples, but the present invention is not limited by these examples. In the explanations below, concentrations of medium components are expressed as weight percentages unless otherwise specified.

Reference Example 1: Obtaining DNA encoding polypeptide involved in biosynthesis of macrolide compound (II)

(1) Cultivation of Mer-11107 and isolation of genomic DNA

Hyphae of *Streptomyces* sp. Mer-11107 were inoculated into 25 mL of Tryptic Soy Broth, and cultured with shaking at 28°C for 3 days. Genomic DNA was prepared from the resulting culture broth according to the methods described under "Isolation genomic DNA" (pp. 162-170) in D.A. Hopwood et al.'s Practical Streptomyces Genetics (The John Innes Foundation, Norwich, England, 2000).

(2) Preparation of Mer-11107 genomic library

160 µL of sterile purified water, 200 µL of Mer-11107 genome DNA solution (1 mg/mL), 40 µL of 10x concentration M buffer solution (100 mM Tris-HCl (pH 7.5), 100 mM MgCl₂, 10 mM dithiothreitol, 500 mM NaCl) and 1 µL of restriction enzyme Sau3AI (1 unit/µL) were mixed and incubated at 37°C for 3 minutes. 50 µL was then taken out and extracted with 50 µL of phenol-chloroform mixture (phenol:chloroform:isoamyl alcohol=25:24:1, volume ratio), the aqueous layer was collected and extracted again with 50

μL of chloroform , and the aqueous layer was again collected. 5 μL of 3 M sodium acetate (pH 6.0) and 150 μL of ethanol were added to the liquid, which was then left at -80°C for 30 minutes and centrifuged to collect the precipitated DNA. After being washed in 70% ethanol, this DNA was dissolved in 90 μL of sterile purified water, and incubated at 37°C for 3 hours after addition of 10 μL of 10 times concentration BAP buffer solution (500 mM Tris-HCl (pH 9.0), 10 mM MgCl₂) and 5 μL of bacterial alkaline phosphatase (0.5 unit/μL, Takara Bio Inc.). The reaction liquid was extracted with 100 μL of phenol-chloroform mixture (phenol:chloroform:isoamyl alcohol=25:24:1, volume ratio), the aqueous layer was collected and extracted again with 100 μL of chloroform, and the aqueous layer was again collected. 10 μL of 3 M sodium acetate (pH 6.0) and 300 μL of ethanol were added to this liquid, which was then left at -80°C for 30 minutes and centrifuged to collect the precipitated DNA. After being washed in 70% ethanol, this DNA was dissolved in 20 μL of TE buffer solution (10 mM Tris-HCl (pH 8.0), 1 mM EDTA).

Meanwhile, 10 μg of SuperCos cosmid vector (Stratagene Co.) was digested with restriction enzyme XbaI in accordance with the Stratagene manual, the DNA terminals were de-phosphorylated with calf intestinal alkaline phosphatase (Takara Bio Inc.), and after being digested with restriction enzyme BamHI and purified, this was

dissolved in 10 μ L of TE buffer solution. 2.5 μ L of the Sau3A1 partial digest solution of Mer-11107 DNA described above was added to 1 μ L of this cosmid DNA solution, and 1.5 μ L of sterile purified water, 5 μ L of DNA Ligation Kit (Takara Bio Inc.) Solution II and 10 μ L of Solution I were added in that order and incubated at 23°C for 10 minutes. 4 μ L of the reaction liquid was packaged into a lambda-phage using Gigapack III XL Kit (Stratagene Co.). When the resulting packaged liquid (total 500 μ L) was subjected to a transduction test, colony formation ability was tested at 380 cfu (colony forming units)/ μ L.

(3) Preparation of various probes

(3-1) Preparation of probes comprising keto synthetase coding regions

The following two primers, KS-3F and KS-4R, were synthesized based on sequences that are normally conserved in the keto synthesis domains of polyketide synthetases (see the Sequence Nos. 5 and 6).

KS-3F: 5'-GACCGCGGCTGGGACGTGGAGGG-3'

KS-4R: 5'-GTGCCCCGATGTTGGA CTCAACGA-3'

These primers were used to carry out PCR under the following conditions.

(PCR reaction liquid composition)

Sterile purified water	31 μ L
2x concentration GC buffer	50 μ L
dNTP mixed solution (2.5 mM each dATP, dGTP, dTTP and	

dCTP)	16 μ L
KS-3F (100 pmol/ μ L)	0.5 μ L
KS-4R (100 pmol/ μ L)	0.5 μ L
Mer-11107 total DNA (100 ng/ μ L)	1 μ L
LA Taq polymerase (5u/ μ L, Takara Bio Inc.)	1 μ L

(Reaction temperature conditions)

95°C 3 minutes

(98°C 20 sec, 63°C 30 sec, 68°C 2 min) 30 cycles

72°C 5 minutes

The 930 bp DNA fragments amplified as a result of this reaction were electrophoresed on 0.8% agarose gel, and the isolated 930 bp DNA fragments were excised and collected and purified using SUPREC-01 (Takara Bio Inc.). Using 10 ng of the resulting DNA fragments as the template, 930 bp DNA fragments comprising the keto synthetase coding region were amplified again under the same PCR conditions as above except that the number of reaction cycles was changed to 20. These DNA fragments were concentrated and purified using SUPREC-02 (Takara Bio Inc.), and 50 μ L of the resulting TE solution was taken as the probe solution.

(3-2) Preparation of probe comprising cytochrome P450 gene region

Two known cytochrome P450 genes were amplified from actinomycetes for purposes of preparing a cytochrome P450 gene probe. That is, the two primers CB-1F and CB-2R comprising the sequences showed below (see the Sequence

Nos.7 and 8) were synthesized for purposes of amplifying the ORF-A gene derived from *Streptomyces thermotolerans* ATCC11416 (*Biosci. Biotechnol. Biochem.* 59: 582-588, 1995).

CB-1F: 5'-ATGACAGCTTTGAATCTGATGGATCCC-3'

CB-2R: 5'-TCAGAGACGGACCGGCAGACTCTTCAGACG-3'

Meanwhile, the two primers PKC-1F and PKC-2R comprising the sequences shown below (see the Sequence Nos. 9 and 10) were synthesized for purposes of amplifying the pik-C gene derived from *Streptomyces venezuelae* ATCC15439 (*Chem. Biol.* 5: 661-667, 1998).

PKC-1F: 5'-GTGCGCCGTACCCAGCAGGGAACGACC-3'

PKC-2R: 5'-TCACGCGCTCTCCGCCCCGCCCCCTGCC-3'

These primers were used to carry out PCR under the following conditions.

(PCR reaction liquid composition)

Sterile purified water	31 μ L
2x concentration GC buffer	50 μ L
dNTP mixed solution (2.5 mM each dATP, dGTP, dTTP and dCTP)	16 μ L
primer-F (100 pmol/ μ L)	0.5 μ L
primer-R (100 pmol/ μ L)	0.5 μ L
ATCC11416 or ATCC15439 genome DNA (100 ng/ μ L)	1 μ L
LA Taq polymerase (5u/ μ L), Takara Bio Inc.)	1 μ L

(Reaction temperature conditions)

95°C, 3 minutes

(98°C 20 sec, 63°C 30 sec, 68°C 2 min) 30 cycles

72°C, 5 minutes

The two 1.2 kb DNA fragments amplified as a result of this reaction were purified by QIAGEN PCR Purification Kit (QIAGEN Co.), and a mixed solution comprising 10 ng/μL of each DNA fragment was prepared and used as the probe.

(4) Screening using probe comprising keto synthetase coding region

An *E. coli* XL-1Blue MR host (Stratagene Co.) was transduced with the Mer-11107 genome DNA library packaged solution prepared in the above (2) in accordance with the Stratagene manual. After transduction the bacterial suspension was dispensed and spread onto ten LB-50 μg/mL ampicillin-1.5% agar medium plates (inner diameter 90 mm, height 15 mm), and cultured for 18 hours at 37°C. The colonies growing on each plate were transferred to Hybondon+ filters (Amersham Biosciences), alkali and neutral treated under the conditions described in the manual for the Hybondon+ filters, and dried for 2 hours at 80°C to fix DNA derived from the colonies onto the filters.

The genome DNA library was screened by colony hybridization using an AlkPhos Direct System (Amersham Biosciences) with 100 ng of the 930 bp DNA fragment comprising the keto synthetase region prepared under (3-1) above as the probe. Hybridization was performed for 2 hours at 65°C at a salt concentration of 0.5 M NaCl. The conditions for hybridization and detection were those

described in the manual attached to the AlkPhos Direct System. Of the roughly 7,600 colonies tested, 59 colonies which hybridized strongly with the alkali phosphatase-labeled probe were isolated. Cosmids were extracted and purified from *E. coli* clones derived from these colonies.

(5) Selection and verification of cosmid clones having pladienolide biosynthesis gene region using probe comprising cytochrome P450 gene region

Two μ L of each of the cosmid DNA solutions obtained under the above (4) was spotted onto a HybondN+ filter, alkali and neutral treated under the conditions described in the attached manual, and dried for 2 hours at 80°C to fix the DNA on the filter. Hybridization was performed with these filters under the same conditions as in the above (4) using the cytochrome P450 gene fragment described under the above (3) as the probe. One cosmid that hybridized strongly with the probe was selected as a result and named pKS58.

The pKS58 DNA was partially digested with Sau3AI restriction enzyme, ligated with the BamHI-CIAP treated phage vector Zap Express (Stratagene Co.), and packaged into a lambda phage using a Gigapack III XL Kit (Stratagene Co.). *E. coli* XL1-Blue MRF' was infected with this phage solution, and made to form a plaque. Plaque hybridization was performed using the cytochrome P450 gene probe prepared in the above (3-2) to subclone an approximately 2 kb length

DNA fragment containing cytochrome P450 gene.

This cytochrome P450 gene DNA fragment was sequenced, and two primers PDL58-1F and PDL58-2R (see the Sequence Nos. 11 and 12) having the sequences shown below were synthesized from the N- and C-terminals, which are considered to be the cytochrome P450 coding regions.

PDL58-1F: 5'-GCCCCGCATATGGATCTGGAAACCCAACTTCTC-3'

PDL58-2R: 5'-GCACTAGTCAGCCGCGCTCGACGAGGAGGTG-3'

These primers were used to carry out PCR under the following conditions.

(PCR reaction liquid composition)

Sterile purified water	31 μ L
2 $\times$ concentration GC buffer	50 μ L
dNTP mixed solution (2.5 mM each dATP, dGTP, dTTP and dCTP)	16 μ L
PDL58-1F (100 pmol/ μ L)	0.5 μ L
PDL58-2R (100 pmol/ μ L)	0.5 μ L
pKS58DNA (10 ng/ μ L)	1 μ L
LA Taq polymerase (5u/ μ L, Takara Bio Inc.)	1 μ L

(Reaction temperature conditions)

95°C, 3 minutes

(98°C 20 sec, 63°C 30 sec, 68°C 2 min) 20 cycles

72°C, 5 minutes

The 1.2 kb DNA fragment amplified as a result of this reaction was purified with QIAGEN PCR Purification Kit (QIAGEN Co.), and digested with NdeI and SpeI restriction

enzymes. After the reaction the DNA was electrophoresed on 0.8% agarose gel, the isolated 1.2 kb DNA fragment was excised and DNA was collected and purified using QIAGEN GelExtraction Kit (QIAGEN Co.). This DNA fragment was inserted into the NdeI and SpeI sites of the cytochrome P450 gene expression plasmid pT7NS-camAB (see Reference Example 5 and WO 03/087381) to construct pPDL96.

E. coli BL21 (DE3) was transformed using this plasmid and cultured in M9CG medium (1.28% $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 0.3% KH_2PO_4 , 0.05% NaCl, 0.1% NH_4Cl , 1% casamino acid, 0.4% glucose, 1 mM MgCl_2 , 100 μM CaCl_2 , 50 $\mu\text{g/mL}$ ampicillin) to a density of 0.8 OD_{600} (optical density at 600 nm). 5-Aminolevulinic acid was added to 80 $\mu\text{g/mL}$ and IPTG to 0.1 mM, and cultivation was continued at 22°C for 25 hours to induce the cytochrome P450 protein. After induction, the mycelia were collected and suspended in 5 mL of CV buffer solution (50 mM NaPO_4 (pH 7.3), 1 mM EDTA, 10% glycerol, 1 mM glucose). 1 mL of this suspension was taken in a test tube and 5 μL of a DMSO solution (50 mg/mL) of ME-265 (the 6-position deoxide of pladienolide B) was added and incubated at 28°C for 15 hours. 1 mL of acetonitrile was added and mixed with this reaction solution, which was then centrifuged and supernatant analyzed by HPLC under the following conditions to confirm conversion to pladienolide B. These results lead the pladienolide biosynthesis gene region is involved in pKS58.

(HPLC analysis conditions)

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (ϕ 4.6 mm $\times$ 50 mm 3 μ m)

Mobile phase: 45% to 55% methanol (0 to 5 min)

55% methanol (5 to 13 min)

55% to 70% methanol (13 to 21 min)

45% methanol (21 to 25 min)

Flow rate: 1.2 mL/min

Detection: UV 240 nm

Injection volume: 5 μ L

Column temperature: 40°C

Analysis time: 25 minutes

Retention time: ME-265: 20 min, Pladienolide B: 13 min

(6) Selection of cosmid comprising biosynthesis gene cluster neighboring cytochrome P450 gene

A cosmid comprising the biosynthesis gene cluster neighboring the cytochrome P450 gene obtained in the above (5) was selected from the 59 cosmid DNA samples obtained in the above (4).

The 59 cosmid DNA samples were digested with restriction enzymes EcoRI and BamHI, and the DNA obtained in each case was electrophoresed on agarose gel and subjected to Southern hybridization using as probes the KS domain DNA (aveA2 KS6 domain) and the AT domain DNA (aveA1 AT2 domain) of the avermectin aglycone biosynthesis gene (see *Proc. Natl. Acad. Sci. USA* 96 (1999) 9509-9514; JP-A

2000-245457; or WO 00/50605) and the cytochrome P450 gene obtained in the above (5).

Those cosmids having DNA fragments that hybridized at the same length were grouped on the basis of the electrophoresis patterns of DNA digested with restriction enzymes EcoRI and BamHI, and the hybridization band patterns using the various probes. Of these, all but one of the cosmids exhibiting similar patterns were deleted, and the remaining cosmids were organized according to partially matching band patterns. Beginning with the pKS58 cosmid comprising the cytochrome P450 gene obtained in (5), pKS54 was selected as a cosmid neighboring the side comprising the polyketide synthetase gene from the cytochrome P450 gene side, and pKS35 was selected as a cosmid neighboring pKS54. pKS23 was also selected as a cosmid neighboring the cosmid pKS58 from cytochrome P450 gene side to the side not comprising the polyketide synthetase gene. As a result, as shown in Figure 2, pKS23, pKS58, pKS54 and pKS35 were selected as cosmid clones encompassing the pladienolide biosynthesis gene cluster.

(7) Determining nucleotide sequence of pladienolide biosynthesis gene cluster

The nucleotide sequence of a DNA group coding for the pladienolide biosynthesis gene was determined. Each cosmid, selected in the above (6), was first isolated by the cesium chloride method, and then sheared to about 1 kb using

HydroShear (Genomic Solutions K.K.) and subcloned using a BKL Kit (Takara Bio Inc.).

The resulting subclones were subjected to a cycle sequence reaction (Amersham Biosciences Co.) using fluorescent-labeled primers, and the nucleotide sequences of the respective fragments were determined (MegaBACE 1000: Amersham Biosciences Co.). Thus, a roughly 75 kb nucleotide sequence comprising DNA associated with pladienolide biosynthesis was determined (see the Sequence No. 1).

(8) Preparation of pladienolide 6-position hydroxylase gene (pldB)-deficient strain.

It has been shown that a pladienolide is biosynthesized by the biosynthesis pathway shown in Fig. 1 from the roughly 75 kb nucleotide sequence comprising DNA participating in pladienolide biosynthesis which was sequenced in the above (7) (see the Sequence No. 1). A pldB-deficient strain was thus prepared by the following methods with the idea that a strain producing only ME-265 (the 6-position deoxide of pladienolide B) could be obtained by destroying only the cytochrome P450 gene pldB.

The four primers pldB-L-Bgl2F, pldB-L-Hind3R, pldB-R-Hind3F and pldB-R-Bgl2R (see the sequence nos. 13, 14, 15 and 16) having the sequences shown below were synthesized based on the nucleotide sequence of the sequence no. 1.

pldB-L-Bgl2F: 5'-GGGAGATCTAGAGGCCGGTTACCTCTACGAGTA-3'

pldB-L-Hind3R: 5'-GGGAAGCTTGCGATGAGCTGTGCCAGATAG-3'

pldB-R-Hind3F: 5'-GGGAAGCTTGAAGTGGCGCGACAGTGTCTT-3'

pldB-R-Bgl2R: 5'-GGGAGATCTGCAGCGGATCGTCTTCGAGACCCTT-3'

PCR was performed under the following conditions using these primers.

(PCR reaction liquid composition)

Sterile purified water	30 μ L
2 x concentration GC buffer	50 μ L
dNTP mixed solution (2.5 mM each dATP, dGTP, dTTP and dCTP)	16 μ L
pldB-L-Bgl2F or pldB-R-Hind3F (50 pmol/ μ L)	1 μ L
pldB-L-Hind3R or pldB-R-Bgl2R (50 pmol/ μ L)	1 μ L
Mer-11107 total DNA (100 ng/ μ L)	1 μ L
LA Taq polymerase (5 u/ μ L, Takara Bio Inc.)	1 μ L

(Reaction temperature conditions)

95°C, 3 minutes

(98°C 20 sec, 63°C 30 sec, 68°C 2 min) 30 cycles

72°C, 5 minutes

As a result, a 1.57 kb DNA fragment (DNA fragment L1) comprising nucleotides 64756 to 66302 in the Sequence No. 1 was amplified by the reaction using pldB-L-Bgl2F and pldB-L-Hind3R, while a 1.54 kb DNA fragment (DNA fragment R1) comprising nucleotides 66849 to 68368 in the Sequence No. 1 was amplified from the reaction using pldB-R-Hind3F and pldB-R-Bgl2R. DNA fragments L1 and R1 were purified with a QIAGEN PCR purification Kit (QIAGEN Co.), and digested with

restriction enzymes BglIII and HindIII.

The DNA fragments L1 and R1 which had been digested with restriction enzymes BglIII and HindIII, a 2.3 kb hygromycin B resistance gene (derived from pHP45omegahyg: Gene 190, 315-317, 1997, sometimes abbreviated hereunder as "hyg") which had been digested with restriction enzyme HindIII and the shuttle vector pKU253 (see Fig. 3) which had been digested with restriction enzyme BamHI were all four connected to DNA ligation kit Ver. 2.1 (Takara Bio Inc.). A roughly 5.4 kb DNA fragment having the hygromycin B resistance gene inserted between DNA fragments L1 and R1 was thus inserted into pKU253 to construct a roughly 21.4 kb plasmid called pKU253-L1-hyg-R1.

The resulting pKU253-L1-hyg-R1 was transformed into conjugated *E. coli* S17-1 by electroporation to obtain S17-1/pKU253-L1-hyg-R1. The resulting S17-1/pKU253-L1-hyg-R1 was inoculated into 10 mL of LB medium (1% bacto tryptone, 0.5% yeast extract, 0.5% NaCl) comprising 25 µg/mL of kanamycin and 100 µg/mL of hygromycin B and shaking cultured at 30°C for 2 hours, and the mycelia were collected, washed twice with 10 mL of LB medium and suspended in 5 mL of LB medium. This was the donor suspension.

While the donor suspension was being prepared, Mer-11107 was inoculated into 10 mL of TSB medium (Trypto-Soya broth: Nissui Seiyaku) and shaking cultured at 30°C for 5

hours, and the mycelia were collected, washed twice with 10 mL of sterile water and suspended in 1 mL of sterile water. This was the recipient suspension. 500 μ L of the S17-1/pKU253-L1-hyg-R1 donor suspension was mixed with 10 μ L of the Mer-11107 recipient suspension, and applied to Actino Medium No. 4 agar medium. After incubated at 30°C for 18 hours, this was covered with 2.5 mL of SNA (0.8% nutrient medium: Difco, 0.4% agar) comprising 2 mg/mL ribostamycin, and incubated at 30°C for 7 days to obtain a ribostamycin-resistant pKU253-L1-hyg-R1 transformant strain.

The resulting pKU253-L1-hyg-R1 transformant strain was inoculated into 10 mL of TSB medium containing no ribostamycin, and shaking cultured at 30°C for 24 hours. Mycelia were collected from the pKU253-L1-hyg-R1 transformant culture broth, washed twice with 10 mL of sterilized water and suspended in 10 mL of sterilized water. After being diluted appropriately, the suspension was applied to YMS agar medium (0.4% yeast extract, 1% wheat germ extract, 0.4% soluble starch, 2% agar, 10 mM calcium chloride) comprising 200 μ g/mL hygromycin B, and incubated at 30°C for 4 days. Single colonies growing on the YMS agar medium comprising hygromycin B were transplanted to YMS agar medium comprising 200 μ g/mL of hygromycin B and YMS agar medium comprising 200 μ g/mL of ribostamycin, and incubated at 30°C for 2 days. After incubated, a hygromycin B -resistant, ribostamycin-sensitive strain was

selected. The resulting strain, called Mer-11107 *pldB::hyg*, was a *pldB*-deficient strain lacking 546 bp (nucleotides 66303 to 66848 in the Sequence No. 1) of the *pldB* gene from the genome, with the hygromycin B resistance gene inserted in its place.

(9) Pladienolide production test of pladienolide 6-position hydroxylase gene (*pldB*)-deficient strain

200 μ L of frozen seed of the Mer-11107 *pldB::hyg* strain obtained in the above (8) was inoculated into 20 mL of seed medium (soluble starch 2%, soy bean powder (ESUSAN-MEAT manufactured by Ajinomoto Co. Ltd.) 2%, yeast extract 0.5%, K_2HPO_4 0.1%, $MgSO_4 \cdot 7H_2O$ 0.25%, $CaCO_3$ 0.3%, pH not adjusted) and incubated at 25°C for 2 days. 300 μ L of the resulting seed culture broth was inoculated into 30 mL of seed culture medium (5% soluble starch, 1% glucose, 3% Pharmamedia, 2% β -cyclodextrin, 0.1% $CaCO_3$, pH 7.5) and cultured at 25°C for 4 and 5 days. After the completion of the cultivation, 20 mL of the resulting culture liquid was extracted by adding an equal amount of acetonitrile thereto. Part of this extract was taken and diluted with 5 times the amount of acetonitrile, and levels of pladienolide B and ME-265 were measured by HPLC under the following conditions. The measurement results are shown in Table 1.

(HPLC analysis conditions)

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (ϕ 4.6 mm $\times$ 50 mm 3 μ m)

Mobile phase: 45% through 55% methanol (0 to 5 min)

55% methanol (5 to 13 min)

55% to 70% methanol (13 to 17 min)

70% methanol (17 to 35 min)

45% methanol (35 to 40 min)

Flow rate: 1.2 mL/minute

Detection: UV 240 nm

Injection volume: 10 µL

Column temperature: 40°C

Analysis time: 35 minutes

Retention time: ME-265: 22 min, Pladienolide B: 16 min

Table 1

Mer-11107 pldB::hyg strain	ME-265 (mg/L)	Pladienolide B (mg/L)
cultured for 4 days (96 hours)	1247.7	0.0
cultured for 5 days (120 hours)	1316.6	0.0

Reference Example 2: DNA-1 encoding polypeptide having the 16-position hydroxylating enzyme activity on the macrolide compound (II)

(1) Preparation of genomic DNA of *Streptomyces* sp. A-1544 strain

The A-1544 strain was inoculated into a medium containing 1% of glucose, 0.4% of malt extract and 1% of yeast extract and incubated at 28°C for 3 days. The obtained culture broth was centrifuged at 3000 rpm for 10 minutes to collect the mycelia. A genomic DNA was prepared

using Blood & Cell Culture Kit (QIAGEN Co.) from the mycelia.

(2) Cloning of a partial sequence of a DNA encoding a protein having the activity in hydroxylating the 16-position of the macrolide compound 11107B

The following mix primers (5Dm-3F and 5Dm-3R) were designed and produced on reference to the amino acid sequence assumed to be that of the cytochrome P450 (CYP105D5) of *Streptomyces coelicolor* A3 (2) (see the Sequence Nos. 17 and 18).

5Dm-3F: 5'-TTCGCCTSCCSGTCCCSTCSATGGTSAT-3'

5Dm-3R: 5'-GTTGATSAYSGASGTSGAGAA-3'

In order to promote reactivity taking the fluctuation of a codon into account, mixed bases S(=C+G) and Y(=C+T) were used.

Next, these two types of primers (5Dm-3F and 5Dm-3R) and the A-1544 strain genomic DNA as a template obtained in the above (1), were used to run a PCR reaction. The PCR reaction was accomplished by repeating a three-stage reaction including denaturing carried out at 98°C for 20 seconds, annealing carried out at 50°C for 2 minutes and elongation carried out at 68°C for 30 seconds 35 times by using Takara LA Taq (Takara Bio Inc.) and a PCR amplifier (T Gradient, Biometra Co.). As a result, a DNA fragment (hereinafter referred to as a DNA fragment-A1) having a size of about 500 bp was amplified. It is highly possible

that this DNA fragment-A1 is a part of the DNA encoding a protein having hydroxylating activity. The DNA fragment-A1 amplified by a PCR reaction was recovered from the reaction solution by SUPREC PCR (Takara Bio Inc.).

In order to obtain the DNA fragment-A1 in an amount enough to analyze the nucleotide sequence of the obtained DNA fragment-A1, the DNA fragment was combined with a plasmid vector pT7Blue T (Novagen Co.) by using DNA Ligation kit ver.2 (Takara Bio Inc.) to transform E.coli JM109 strain. Thereafter, the transformed E. coli was selected using a L-broth agar media (1.0% bacto tryptone, 0.5% yeast extract, 0.5% NaCl, 1.5% agar) containing ampicillin (50 µg/mL), X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside; 40 µg/mL) and IPTG (isopropyl-β-D-thiogalactopyranoside; 100 µM). The colony of the transformed E. coli thus isolated was cultured in a L-broth liquid medium (1% bacto tryptone, 0.5% yeast extract, 0.5% NaCl) containing ampicillin (50 µg/mL). A plasmid DNA was separated from the mycelia of the proliferated transformed E. coli and purified by using a plasmid purifying kit (QIA filter Plasmid Midi Kit, QIAGEN Co.), to obtain enough amount of the DNA fragment-A1.

(3) Analysis of the nucleotide sequence of the cloned DNA fragment-A1

The nucleotide sequence of the DNA fragment-A1 obtained in the above (2) was analyzed using a DNA

nucleotide sequence analyzer (PE Biosystems 377XL) according to a dye terminator cycle sequence method. As the result of the nucleotide sequence analysis, it was clarified that the DNA fragment-A1 amplified by a PCR reaction had an exact size of 528 bp though it had been found to have a size of about 500 bp by the above measurement using electrophoresis (see the nucleotide sequence 1775 to nucleotide sequence 2302 of the Sequence No. 2). Since DNA sequences corresponding to the two types of primers used in the above PCR reaction were found at both ends of the above cloned 528 bp DNA sequence, it was clarified that the DNA fragment-A1 was singularly amplified by these two types of primers (5Dm-3F and 5Dm-3R) in the above PCR reaction.

(4) Analysis of the neighboring region of the DNA fragment-A1

As mentioned above, the partial sequence of the DNA encoding a protein having the hydroxylating activity derived from the A-1544 strain was determined. Therefore, the amplification, cloning and sequence analysis of the nucleotide sequence in the neighboring region extending to the upstream side and downstream side of the cloned fragment were accomplished by an inverse PCR method (Cell Technology vol. 14, p.591-593, 1995). Specifically, the A-1544 strain genomic DNA (see the above (1)) was digested by respective restriction enzymes PstI and SalI in a H buffer

solution (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 10 mM dithiothreitol and 100 mM NaCl). The obtained each DNA fragment cleaved by the restriction enzymes was self-circularized by using DNA Ligation Kit ver.2 (Takara Bio Inc.).

Meanwhile, the following primers (6PIN-2F and 6PIN-2R) were designed and produced based on the nucleotide sequence of DNA fragment-A1 (see the Sequence Nos. 19 and 20).

6PIN-2F: 5'-GCTGCGCCTGGCCCTGGAGGACATCGAGAT-3'

6PIN-2R: 5'-CTGTTTCCTCGAAGAACTCGTGGTCGGCGTA-3'

Next, these two primers (6PIN-2F and 6PIN-2R) and the above self-circularized A-1544 strain genomic DNA as a template, were used to run a PCR reaction. In the PCR reaction, the cycle of a two-stage reaction including denaturing carried out at 98°C for 20 seconds and annealing and elongation carried out at 68°C for 5 minutes was repeated 35 times by using Takara LA Taq (Takara Bio Inc.) and a PCR amplifier (T Gradient, Biometra Co.).

As a result, a DNA fragment (DNA fragment-B1) about 3.5 kbp in size and a DNA fragment (DNA fragment-C1) about 2.8 kbp in size were amplified. It was highly possible that these DNA fragments were a DNA encoding a protein having hydroxylating activity and a DNA having a DNA sequence including the upstream and downstream regions of the former DNA.

The DNA fragment-B1 and the DNA fragment-C1 were

recovered from the PCR amplified reaction solution by SUPREC PCR (Takara Bio Inc.). Next, as to the obtained DNA fragment-B1 and DNA fragment-C1, in order to obtain each DNA fragment in an amount enough to analyze the nucleotide sequence of the obtained DNA fragment, a plasmid vector pT7Blue T (Novagen Co.), DNA Ligation kit ver.2 (Takara Bio Inc.), E. coli JM109 strain and a plasmid purifying kit (QIA filter Plasmid Midi kit, QIAGEN Co.) were used in the same manner as the above (2), to obtain enough amount of each DNA fragment.

(5) Analysis of each nucleotide sequence of the DNA fragment-B1 (about 3.5 kbp in size) and the DNA fragment-C1 (about 2.8 kbp in size)

Each nucleotide sequence of the DNA fragment-B1 and DNA fragment-C1 obtained in the above (4) was analyzed using a DNA nucleotide sequence analyzer (PE Biosystems 377XL) according to a dye terminator cycle sequence method. The nucleotide sequence was thus analyzed to obtain the information of the nucleotide sequence of 3793 bp shown in the Sequence No. 2 from each sequence of the DNA fragment-B1 and DNA fragment-C1.

An open reading frame (ORF) in this 3793 bp was retrieved, to find that the two proteins were coded. Each amino acid sequence of these proteins was retrieved by the BLAST search, and as a result, an ORF (hereinafter referred to as PsmA) coding a protein consisting of 409 amino acids

having high homology to cytochrome P450 existed in the base 1322 to base 2548 of the Sequence No. 2. The psmA had the highest homology (homology: 72.6%) to the amino acid sequence assumed to be that of cytochrome P450 (CYP105D5) of the *Streptomyces coelicolor* A3 (2) and to the amino acid sequence assumed to be that of cytochrome P450 (CYP105D4) of the *Streptomyces lividans*, and also had a relatively high homology (homology: 69.4%) to cytochrome P450 soy (Soy C) of *Streptomyces griseus*. It was considered from this fact that the psmA was highly possibly a gene coding hydroxylating enzyme of the cytochrome P-450 type.

Also, an ORF (hereinafter referred to as psmB) encoding a protein having a high homology to ferredoxin of a 2Fe-2S type existed just downstream (the base 2564 to base 2761 of the Sequence No. 2) of the psmA. The protein encoded by the psmB consists of 66 amino acids, and had the highest homology (83.3%) to the amino acid sequence assumed to be that of ferredoxin just downstream of the amino acid sequence assumed to be that of cytochrome P450 (CYP105D5) of the *Streptomyces coelicolor* A3(2) and a relatively higher homology (homology: 57.6%) to ferredoxin soy (soyB) of *Streptomyces griseus*. Therefore, it was considered that the psmB serves to transfer electrons and codes ferredoxin participating in hydroxylation together with the psmA.

(6) Construction of a plasmid pTC-DM

pT7NS-CamAB (see Reference Example 5 and WO 03/087381)

was digested by respective restriction enzymes NdeI and SpeI in a H buffer solution (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 10 mM dithiothreitol and 100 mM NaCl) to obtain a plasmid digested products. Meanwhile, a primer DM-NdeF (5'-GCCCCCATATGACGGAAGTACGGACATCA-3': see the Sequence No. 33) obtained by adding a NdeI site to the 5'-terminal and a primer DM-SpeR (5'-GGGCCACTAGTCAGCCGGCCGGTTCGGTCA-3': see the Sequence No. 34) obtained by adding a SpeI site to the 5'-terminal were designed and produced on reference to the nucleotide sequence of the Sequence No. 2. Next, these two types of primers (DM-NdeF and DM-SpeR) and the A-1544 strain genomic DNA obtained in the above (1) as a template, were used to run a PCR reaction. The PCR reaction was accomplished by repeating a two-stage reaction including denaturing carried out at 98°C for 20 seconds and annealing and elongation carried out at 68°C for 2 minutes 30 times by using Takara LA Taq (TAKARA HOLDINGS INC.) and a PCR amplifier (T Gradient, Biometra Co.). From the PCR amplification reaction solution, a DNA fragment having a size of about 1.5 kbp and containing the psmA and the psmB was recovered by SUPREC PCR (TAKARA HOLDINGS INC.). The resulting DNA fragment was digested with restriction enzymes NdeI and SpeI, and the digest and the aforementioned pT7NS-camAB plasmid digest were connected using DNA Ligation Kit Ver. 2 (Takara Bio Inc.). Thus, a roughly 9.5 kbp plasmid (sometimes called plasmid pTC-DM)

was constructed comprising the pT7NS-camAB plasmid connected to a DNA fragment containing both the psmA and the psmB.

Reference Example 3: DNA-2 encoding polypeptide having the 16-position hydroxylating enzyme activity on the macrolide compound (II)

(1) Preparation of genomic DNA of *Streptomyces* sp. Mer-11107 strain

The Mer-11107 strain was inoculated into a medium containing 1% of glucose, 0.4% of malt extract and 1% of yeast extract and incubated at 28°C for 3 days. The obtained culture broth was centrifuged at 3000 rpm for 10 minutes to collect the mycelia. A genomic DNA was prepared using Blood & Cell Culture Kit (QIAGEN Co.) from the mycelia.

(2) Cloning of a partial sequence of a DNA encoding a protein having the activity in hydroxylating the 16-position of the macrolide compound 11107B

The following mix primers (5Dm-3F and 5D-1R) were designed and produced on reference to the amino acid sequence assumed to be that of the cytochrome P450 (CYP105D5) of *Streptomyces coelicolor* A3(2) (see the Sequence Nos. 17 and 21).

5Dm-3F: 5'-TTCGCSC TSCCSGTCCCSTCSATGGTSAT-3'

5D-1R: 5'-AGGTGCCCAGCGAGATCATGTT-3'

In order to promote reactivity taking the fluctuation

of a codon into account, mixed bases S(=C+G) and Y(=C+T) were used.

Next, these two primers (5Dm-3F and 5D-1R) and the Mer-11107 genomic DNA obtained in the above (1) as the template were used to run a PCR reaction. The PCR reaction was accomplished by repeating a three-stage reaction including denaturing carried out at 98°C for 20 seconds, annealing carried out at 50°C for 2 minutes and elongation carried out at 68°C for 30 seconds 35 times by using Takara LA Taq (Takara Bio Inc.) and a PCR amplifier (T Gradient, Biometra Co.). A roughly 300 bp DNA fragment (hereunder called DNA fragment A2) was amplified as a result. It is highly possible that this DNA fragment A2 is a part of the DNA encoding a protein having hydroxylation activity. The DNA fragment-A2 amplified by the PCR reaction was recovered from the reaction solution by SUPREC PCR (Takara Bio Inc.).

Next, in order to obtain the DNA fragment-A2 in an amount enough to analyze the nucleotide sequence of the obtained DNA fragment-A2, the DNA fragment-A2 was combined with a plasmid vector pT7Blue T (Novagen Co.) using DNA Ligation Kit Ver. 2 (Takara Bio Inc.), to transform E. coli JM109 strain. The transformed E. coli was subsequently selected using a L-broth agar media(1.0% bacto tryptone, 0.5% yeast extract, 0.5% NaCl, 1.5% agar) containing ampicillin (50 µg/mL), X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside; 40 µg/mL) and IPTG (isopropyl-β-D-

thiogalactopyranoside; 100 μ M). The transformed *E. coli* colonies isolated in this way were cultured in a L-broth liquid medium (1% bacto tryptone, 0.5% yeast extract, 0.5% NaCl) containing ampicillin (50 μ g/mL). A plasmid DNA was isolated and purified from the mycelia of the proliferated transformed *E. coli* using a plasmid purifying kit (QIA filter Plasmid Midi Kit, QIAGEN Co.) to obtain enough amount of DNA fragment-A2.

(3) Analysis of the nucleotide sequence of the cloned DNA fragment-A2

The nucleotide sequence of the DNA fragment-A2 obtained in the above (2) was analyzed using a DNA nucleotide sequence analyzer (PE Biosystems 377XL) according to a dye terminator cycle sequence method. While the DNA fragment-A2 amplified by a PCR reaction was measured at about 300 bp by electrophoresis, it was clarified by the nucleotide sequence analysis to be exactly 325 bp in length (see the nucleotide sequence 837 to nucleotide sequence 1161 of the Sequence No. 3). Since DNA sequences corresponding to the two types of primers used in the aforementioned PCR reaction were found at both ends of the above cloned 325 bp DNA sequence, it was clarified that the DNA fragment-A2 had been specifically amplified by these two types of primers (5Dm-3F and 5D-1R) in the above-mentioned PCR reaction.

(4) Analysis of the neighboring region of the DNA fragment-

A2

Once the partial sequence of the DNA encoding a protein having the hydroxylating activity derived from the Mer-11107 strain had been determined as described above, the nucleotide sequence of the neighboring region extending the upstream side and downstream side of the cloned fragment was amplified, cloned and sequenced by an inverse PCR method (Cell Technology Vol. 14, pp. 591-593, 1995). Specifically, the Mer-11107 strain genomic DNA (see the above (1)) was digested with a restriction enzyme BamHI in a K buffer solution (50 mM Tris-HCl, pH 8.5, 10 mM MgCl₂, 1 mM dithiothreitol, 100 mM KCl) and with a restriction enzyme SalI in a H buffer solution (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 1 mM dithiothreitol, 100 mM NaCl). The obtained each DNA fragment cleaved by the restriction enzymes was self-circularized by using DNA Ligation Kit Ver. 2 (Takara Bio Inc.).

Meanwhile, the following primers (7PIN-2F and 6PIN-2R) were designed and prepared from the nucleotide sequence of the DNA fragment-A2. (see the Sequence Nos. 22 and 20)

7PIN-2F: 5'-CCATGATCCTGCTGGTGGCCGGCCATGAGA-3'

6PIN-2R: 5'-CTGTTTCCTCGAAGAACTCGTGGTCGGCGTA-3'

Next, these two types of primers (7PIN-2F and 6PIN-2R) and the above self-circularized Mer-11107 strain genomic DNA as a template, were used to run a PCR reaction. In the PCR reaction, the cycle of a two-stage reaction including

denaturing carried out at 98°C for 20 seconds and annealing and elongation carried out at 68°C for 5 minutes was repeated 35 times by using Takara LA Taq (Takara Bio Inc.) and a PCR amplifier (T Gradient, Biometra Co.).

As a result, a DNA fragment (DNA fragment-B2) about 1.3kbp in size and a DNA fragment (DNA fragment-C2) about 1.4kbp in size were amplified. It is highly possible that these DNA fragments were respectively a DNA encoding a protein having hydroxylating activity and a DNA having a DNA sequence including those in the upstream and downstream regions of the former DNA.

The DNA fragment-B2 and DNA fragment-C2 were recovered from the PCR amplified reaction solution by SUPREC PCR (Takara Bio Inc.). Next, as to the obtained DNA fragment-B2 and DNA fragment-C2, in order to obtain each DNA fragment in an amount enough to analyze the nucleotide sequence of the obtained DNA fragment, a plasmid vector pT7Blue T (Novagen Co.), DNA Ligation kit ver.2 (TAKARA HOLDINGS INC.), E. coli JM109 strain and a plasmid purifying kit (QIA filter Plasmid Midi Kit, QIAGEN Co.) were used in the same manner as the above (2), to obtain enough amount of each DNA fragment.

(5) Analysis of each nucleotide sequence of the DNA fragment-B2 (about 1.3 kbp in size) and the DNA fragment-C2 (about 1.4 kbp in size)

Each nucleotide sequence of the DNA fragment-B2 and

DNA fragment-C2 obtained in the above (4) was analyzed using a DNA nucleotide sequence analyzer (PE Biosystems 377XL) according to a dye terminator cycle sequence method. The nucleotide sequences were thus analyzed to obtain the information of the nucleotide sequence of 2329 bp shown in the Sequence No. 3 from each sequence of the DNA fragment-B2 and DNA fragment-C2.

An open reading frame (ORF) in this 2329 bp was retrieved, to find that the two kinds of protein were coded. Each amino acid sequence of these proteins was retrieved by the BLAST search, and as a result, an ORF (hereinafter referred to as bpmA) encoding a protein consisting of 395 amino acids having high homology to cytochrome P450 existed in the base 420 to base 1604 of the Sequence No. 3. The bpmA had the highest homology (homology: 67.4%) to the amino acid sequence of the psmA isolated from the A-1544 strain and also had a relatively high homology (homology: 64.8%) to cytochrome P450 soy (Soy C) of *Streptomyces griseus*. It was considered from this fact that the bpmA highly possibly encoded hydroxylating enzyme of the cytochrome P-450 type.

Also, an ORF (hereinafter referred to as bpmB) encoding a protein having a high homology to ferredoxin of a 2Fe-2S type that existed just downstream (the base 1643 to base 1834 of the Sequence No. 3) of the bpmA. The protein encoded by the bpmB consisted of 64 amino acids,

and had the highest homology (81.0%) to the amino acid sequence of the psmB isolated from the A-1544 strain and had a relatively higher homology (homology: 76.2%) to the amino acid sequence assumed to be that of ferredoxin just downstream of the amino acid sequence assumed to be cytochrome P450 (CYP105D5) of *Streptomyces coelicolor* A3(2). Therefore, it was considered that the bpmB served to transfer electrons and participated in hydroxylation together with the bpmA.

Reference Example 4: DNA-3 encoding polypeptide having the 16-position hydroxylating enzyme activity on the macrolide compound (II)

(1) Preparation of genomic DNA of A-1560 strain

The A-1560 strain was inoculated into medium containing 1% glucose, 0.4% wheat germ extract and 1% yeast extract, and incubated at 28°C for 3 days. The obtained culture broth was centrifuged at 3000 rpm for 10 minutes to collect the mycelia. A genomic DNA of A-1560 strain was prepared from the mycelia using Blood & Cell Culture Kit (QIAGEN Co.).

(2) Cloning of a partial sequence of a DNA encoding a protein having the activity in hydroxylating the 16-position of the macrolide compound 11107B

The following mixed primers (5Dm-3F and 5Dm-2R) were designed and produced on reference to the amino acid sequence assumed to be that of the cytochrome P450

(CYP105D5) of *Streptomyces coelicolor* A3(2) (see the Sequence Nos. 17 and 23).

5Dm-3F: 5'-TTCGCCTSCCSTCCSTCSATGGTSAT-3'

5Dm-2R: 5'-CTGGATSGTGTCCSCGGYTT-3'

In order to promote reactivity taking the fluctuation of a codon into account, mixed bases S (=C+G) and Y (=C+T) were used.

Next, these two types of primers (5Dm-3F and 5Dm-2R) and the A-1560 genomic DNA obtained in the above (1) as the template, were used to run a PCR reaction. The PCR reaction was accomplished by repeating a three-stage reaction including denaturing carried out at 98°C for 20 seconds, annealing carried out at 50°C for 2 minutes and elongation carried out at 68°C for 30 seconds 35 times by using Takara LA Taq (Takara Bio Inc.) and a PCR amplifier (T Gradient, Biometra Co.). A roughly 750 bp DNA fragment (hereunder called DNA fragment-A3) was amplified as a result. It is highly possible that this DNA fragment-A3 is a part of DNA encoding a protein having hydroxylating activity. The DNA fragment-A3 amplified by a PCR reaction was recovered from the reaction solution by SUPREC PCR (Takara Bio Inc.).

Next, in order to obtain the DNA fragment-A3 in an amount enough to analyze the nucleotide sequence of the obtained DNA fragment-A3, DNA fragment-A3 was combined with a plasmid vector pT7Blue T (Novagen Co.) by using DNA

Ligation Kit Ver. 2 (Takara Bio Inc.) to transform a colibacillus JM109 strain. (Stratagene Co.). Thereafter, the transformed colibacillus was selected using a L-broth agar media (1.0% bacto tryptone, 0.5% yeast extract, 0.5% NaCl, 1.5% agar) containing ampicillin (50 µg/mL), X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactoside; 40 µg/mL) and IPTG (isopropyl-β-D-thiogalactopyranoside; 100 µM). The colony of the transformed colibacillus isolated in this way were cultured in a L-broth liquid medium (1% bacto tryptone, 0.5% yeast extract, 0.5% NaCl) containing ampicillin (50 µg/mL). A plasmid DNA was isolated and purified from the mycelia of the proliferated transformed colibacillus by using a plasmid purifying kit (QIA filter Plasmid Midi Kit, QIAGEN Co.), to obtain enough amount of DNA fragment-A3.

(3) Analysis of the nucleotide sequence of cloned DNA fragment-A3

The nucleotide sequence of the DNA fragment-A3 obtained in the above (2) was analyzed using a DNA nucleotide sequence analyzer (PE Biosystems 377XL) according to a dye terminator cycle sequence method. While the DNA fragment-A3 amplified by a PCR reaction was measured at about 750 bp by electrophoresis, it was clarified by the nucleotide sequence analysis to be exactly 741 bp in length (see the nucleotide sequence 616 to nucleotide sequence 1356 of the Sequence No. 4). Since DNA sequences corresponding to the two types of primers used in

the aforementioned PCR reaction were found at both ends of the above cloned 741 bp DNA sequence, it was clarified that the DNA fragment-A3 had been specifically amplified by these two types of primers (5Dm-3F and 5Dm-2R) in the aforementioned PCR reaction.

(4) Analysis of the neighboring region of the DNA fragment-A3

Once the partial sequence of the DNA encoding a protein having hydroxylating activity derived from the A-1560 strain had been determined as described above, the nucleotide sequence of the neighboring region extending the upstream side and downstream side of the cloned fragment was amplified, cloned and sequenced by an inverse PCR method (Cell Technology Vol. 14, pp. 591-593, 1995). Specifically, the A-1560 genomic DNA (see the above (1)) was digested with a restriction enzyme BamHI in K buffer solution (50 mM Tris-HCl, pH 8.5, 10 mM MgCl₂, 1 mM dithiothreitol and 100 mM KCl), with a restriction enzyme KpnI in a L buffer solution (10 mM Tris-HCl, pH 7.5, 10 mM MgCl₂ and 1 mM dithiothreitol) and with a restriction enzyme SalI in a H buffer solution (50 mM Tris-HCl, pH 7.5, 10 mM MgCl₂, 1 mM dithiothreitol and 100 mM NaCl). The obtained each DNA fragment cleaved by the restriction enzymes was self-circularized by using DNA Ligation Kit Ver. 2 (Takara Bio Inc.).

Meanwhile, the following primers (5PIN-2F and 6PIN-2R)

were designed and prepared from the nucleotide sequence of the DNA fragment-A3 (see the Sequence Nos. 24 and 20).

5PIN-2F: 5'-CGGAATCCACCAGTGCCTCGGCCAGAACCT-3'

6PIN-2R: 5'-CTGTTCTCGAAGAACTCGTGGTCGGCGTA-3'

Next, these two types of primers (5PIN-2F and 6PIN-2R) and the above self-circularized A-1560 strain genomic DNA as a template, were used to run a PCR reaction. In the PCR reaction, the cycle of a two-stage reaction including denaturing carried out at 98°C for 20 seconds and annealing and elongation carried out at 68°C for 5 minutes was repeated 35 times by using Takara LA Taq (Takara Bio Inc.) and a PCR amplifier (T Gradient, Biometra Co.).

As a result, a DNA fragment (DNA fragment-B3) about 4.5 kbp in size, a DNA fragment (DNA fragment-C3) about 3.0 kbp in size and a DNA fragment (DNA fragment-D3) about 1.7 kbp in size were amplified. It was highly possible that these DNA fragments were a DNA encoding a protein having hydroxylating activity and a DNA having a DNA sequence including those in the upstream and downstream regions of the former DNA.

The DNA fragment-B3, the DNA fragment-C3 and the DNA fragment-D3 were recovered from the PCR amplified reaction solution by SUPREC PCR (Takara Bio Inc.). Next, as to the obtained DNA fragment-B3, DNA fragment-C3 and DNA fragment-D3, in order to obtain each DNA fragment in an amount enough to analyze the nucleotide sequence of the obtained

DNA fragment, a plasmid vector pT7Blue T (Novagen Co.), DNA Ligation kit ver.2 (Takara Bio Inc.), E. coli JM109 strain and a plasmid purifying kit (QIA filter Plasmid Midi Kit, QIAGEN Co.) were used in the same manner as the above (2), to obtain enough amount of each DNA fragment.

(5) Analysis of each nucleotide sequence of the DNA fragment-B3 (about 4.5 kbp in size), the DNA fragment-C3 (about 3.0 kbp in size) and the DNA fragment-D3 (about 1.7 kbp in size)

Each nucleotide sequence of the DNA fragment-B3, DNA fragment-C3 and DNA fragment-D3 obtained in the above (4) was analyzed using a DNA nucleotide sequence analyzer (PE Biosystems 377XL) according to a dye terminator cycle sequence method. The nucleotide sequence was thus analyzed to obtain the information of the nucleotide sequence of 1860 bp shown in the Sequence No. 4 from each sequence of the DNA fragment-B3, DNA fragment-C3 and DNA fragment-D3.

An open reading frame (ORF) in this 1860 bp was retrieved, to find that the two kinds of protein were encoded. Each amino acid sequence of these proteins was retrieved by the BLAST search, and as a result, an ORF (hereinafter referred to as tpmA) encoding a protein consisting of 404 amino acids having high homology to cytochrome P450 existed in the base 172 to base 1383 of the Sequence No. 4. The tpmA had the highest homology (homology: 77.4%) to the amino acid sequence assumed to be

that of cytochrome P450 (CYP105D5) of *Streptomyces coelicolor* A3(2) and also a high homology (homology: 76.6%) to the amino acid sequence of the psmA isolated from the A-1544 strain. It was considered from this fact that the tpmA was highly possibly a gene encoding hydroxylating enzyme of the cytochrome P-450 type.

Also, an ORF (hereinafter referred to as tpmB) encoding a protein having a high homology to ferredoxin of a 2Fe-2S type existed just downstream (the base 1399 to base 1593 of the Sequence No. 4) of the tpmA. The protein encoded by the tpmB consisted of 65 amino acids, and had the highest homology (87.3%) to the amino acid sequence of the psmB isolated from the A-1544 strain and also a high homology (homology: 82.5%) to the amino acid sequence assumed to be that of ferredoxin just downstream of the amino acid sequence assumed to be cytochrome P450 (CYP105D5) of *Streptomyces coelicolor* A3 (2). Therefore, it was considered that the tpmB served to transfer electrons and coded ferredoxin participating in hydroxylation together with the tpmA.

Reference Example 5: Construction of plasmid pT7NS-camAB

The primers PRR-1F (see the Sequence No. 29) and PRR-2R (see the Sequence No. 30) with the sequences shown below were used to carry out PCR under the following conditions. with *Pseudomonas putida* ATCC 17453 genome DNA as the template.

PRR-1F: 5'-GCCCCCATATGAACGCAAACGACAACGTGGTCATC-3'

PRR-2R: 5'-GCGGATCCTCAGGCACTACTCAGTTCAGCTTTGGC-3'

(Reaction liquid composition)

Sterile purified water:	15 μ L
2x concentration GC buffer 1 (TAKARA HOLDINGS INC.)	25 μ L
dNTP mixed solution (2.5 mM each of dATP, dGTP, dTTP and dCTP)	8 μ L
PRR-1F primer (100 pmol/ μ L)	0.5 μ L
PRR-2R primer (100 pmol/ μ L)	0.5 μ L
<i>Pseudomonas putida</i> ATCC 17453 genome DNA (10 ng/ μ L)	0.5 μ L
LA Taq (5 units/ μ L, TAKARA HOLDINGS INC.)	0.5 μ L

(Temperature conditions)

95°C, 3 minutes

(98°C 20 sec, 63°C 30 sec, 68°C 2 min) 30 cycles

72°C, 5 minutes

The resulting 1.5 kb amplified fragment (camAB fragment) comprising the putidaredoxin reductase gene (camA) and the putidaredoxin gene just downstream from that gene (camB) was digested with restriction enzymes Nde I and Bam HI, and electrophoresed on 0.8% agarose gel. After electrophoresis, the camAB fragment was collected and purified using SUPREC-01 (TAKARA HOLDINGS INC.) from a gel section comprising that fragment which had been excised from the gel. The fragment was connected by T4 DNA ligase to the Nde I site and Bam HI site of *E. coli* plasmid vector

pET11a (Stratagene Co.), and then transformed into *E. coli* DH5 α to construct plasmid pT7-camAB. Next, a linker 1 molecule obtained by annealing the two synthetic oligo DNAs SP-1 (see the Sequence No. 31) and SP-2 (see the Sequence No. 32) with the sequences shown below was connected to the Nde I site of the plasmid by T4 DNA ligase and transformed into *E. coli* DH5 α to construct plasmid pT7NS-camAB.

SP-1: 5'-TATGCGTCACTAGTCGGGAGTGCGTTA-3'

SP-2: 5'-TATAACGCACTCCCGACTAGTGACGCA-3'

Example 1: Preparation of plasmid pPapsmAB by connection of pldA promoter and psmAB

The pldA promoter region of the pladienolide B biosynthesis gene was connected to psmAB in order to express psmAB in conjunction with the pladienolide B biosynthesis gene. The two primers pldApro-SpeNdeF (SpeI and NdeI sites added to 5'-terminal: see the Sequence No. 25) and pldApro-NdeR (NdeI site added to 5'-terminal: see the Sequence No. 26) comprising the sequences shown below were synthesized based on the nucleotide sequence information of the Sequence No. 1 for purposes of amplifying the pldA promoter region.

pldApro-SpeNdeF: 5'-GGGCATATGACTAGTAGCCGTGTCCTGCCCCGCC-3'

pldApro-NdeR: 5'-GGGCATATGTTCGGACGTGAATTCATTCG-3'

PCR was carried out under the following conditions using these primers.

(PCR reaction liquid composition: using Toyo Boseki KOD-

plus-)

Sterile purified water:	71 μ L
10 x concentration PCR buffer:	10 μ L
dNTP mixed solution (2 mM each dATP, dGTP, dCTP and dTTP):	10 μ L
MgSO ₄ (25 mM)	4 μ L
KOD+	2 μ L
pIdApro-SpeNdeF (50 pmol/ μ L)	1 μ L
pIdApro-NdeR (50 pmol/ μ L)	1 μ L
pKS35 (Referential Example 1 (6), 10 ng/ μ L)	1 μ L

(Reaction conditions: using Biometra T GRADIENT)

95°C, 5 minutes

(98°C 20 sec, 63°C 1 min, 68°C 1 min) 25 cycles

72°C, 5 minutes

The 360 bp DNA fragment amplified as a result of this reaction was collected with a QIAquick PCR Purification Kit (QIAGEN Co.). The resulting 360 bp DNA fragment was digested with restriction enzyme NdeI. The plasmid pTC-DM (Reference Example 2(6)) was similarly digested with restriction enzyme NdeI. The resulting 360 bp DNA fragment digest and plasmid digest were connected using DNA Ligation Kit Ver. 2.1 (Takara Bio Inc.). A roughly 9.9 kbp plasmid pPapsmaAB comprising the pIdA promoter connected to psmAB was constructed in this way.

Example 2: Preparation of plasmid for recombination

The pPapsmaAB plasmid prepared in Example 1 was

digested with restriction enzyme SpeI, and blunt-ended using BKL Kit (Takara Bio Inc.). The resulting DNA was electrophoresed (Mupid-ex: Advance Co.) on 0.8% Molecular Biology Agarose™ (BIO-RAD), and the isolated 1.8 kbp DNA fragment comprising the *pIdA* promoter and *psmAB* (sometimes called *PapsmAB* hereunder) was excised and collected and purified using a QIAquick Gel Extraction Kit (QIAGEN Co.). The recombination vector *pUC19aph::oriT::intphiC31* shown in Fig. 4 was also digested with restriction enzyme BamHI, and blunt-ended with BKL Kit (Takara Bio Inc.). The resulting vector and *PapsmAB* were connected using DNA Ligation Kit Ver. 2.1 (Takara Bio Inc.). Thus, a recombination plasmid *pAOC::PapsmAB* was constructed having *PapsmAB* inserted into the recombination vector *pUC19aph::oriT::intphiC31*.

Example 3: Introduction of recombination plasmid *pAOC::PapsmAB* into Mer-11107

The resulting *pAOC::PapsmAB* was transformed into conjugated *E. coli* S17-1 (ATCC47055) by electroporation to obtain S17-1/*pAOC::PapsmAB*. The resulting S17-1/*pAOC::PapsmAB* was inoculated into 10 mL of LB medium (1% bacto tryptone, 0.5% yeast extract, 0.5% NaCl) containing 25 µg/ml kanamycin and shaking cultured at 30°C for 2 hours, and the mycelia were collected, washed twice with 10 mL of LB medium and suspended in 5 mL of LB medium. This was used as the donor suspension. While the donor suspension was being prepared, Mer-11107 was inoculated into 10 mL of

TSB medium (Trypto-Soya broth: Nissui Seiyaku) and shaking incubated at 30°C for 5 hours, and the mycelia were collected, washed twice with 10 mL of sterilized water and suspended in 1 mL of sterilized water. This was used as the recipient suspension. 500 µL of the S17-1/pAOC::PapsmAB donor suspension was mixed with 10 µL of the resulting Mer-11107 recipient suspension, and applied to Actino Medium No. 4 agar medium (Nissui Pharmaceutical). After incubated at 30°C for 18 hours, this was covered with 2.5 mL of SNA (0.8% nutrient medium: Difco, 0.4% agar) containing 2 mg/mL ribostamycin, and incubated at 30°C for 7 days to obtain a the ribostamycin-resistant pAOC::PapsmAB-transformed strain Mer-11107::pAOC::PapsmAB.

Example 4: Pladienolide production test of Mer-11107::pAOC::PapsmAB strain

The Mer-11107::pAOC::PapsmAB obtained in Example 3 and the parent Mer-11107 strain as a control were tested for productivity of pladienolide B and its 16-position hydroxide pladienolide D.

300 µL each of frozen seed of Mer-11107::pAOC::PapsmAB and Mer-11107 was inoculated into 60 mL of seed medium (soluble starch 2%, soy bean meal (ESUSAN-MEAT manufactured by Ajinomoto Co. Ltd.) 2%, yeast extract 0.3%, K₂HPO₄ 0.1%, MgSO₄·7H₂O 0.25%, CaCO₃ 0.3%, pH not adjusted) and incubated at 25°C for 3 days. 600 µL of the resulting seed culture broth was inoculated into 60 mL of main culture medium

(soluble starch 5%, glucose 1%, Pharmamedia 3%, adekanol LG-126 0.05%, CaCO₃ 0.1%, pH 7.5), and cultured at 25°C for 3 days, 4 days, 5 days and 6 days. After the completion of the cultivation, the resulting culture broth was extracted by addition of 4 times the amount of acetonitrile. The amounts of pladienolide B and its 16-position hydroxide pladienolide D in the resulting extract were measured by HPLC. The measurement results are shown in Table 2. The HPLC measurement conditions were as follows.

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (φ 4.6 mm × 50 mm 3 μm)

Mobile phase: 45% to 55% methanol (0 to 5 min)

55% methanol (5 to 13 min)

55% to 70% methanol (13 to 21 min)

70% methanol (21 to 25 min)

Flow rate: 1.2 mL/min

Detection: UV 240 nm

Injection volume: 5 μL

Column temperature: 40°C

Analysis time: 25 min

Retention time: Pladienolide B 12.4 min

Pladienolide D 4.3 min

Table 2

	Mer-11107::pAOC::PapsmAB strain		Mer-11107 strain	
	Pladienolide B (mg/L)	Pladienolide D (mg/L)	Pladienolide B (mg/L)	Pladienolide D (mg/L)
cultured for 3 days (72 hours)	36.2	148.4	180.5	0.0
cultured for 4 days (96 hours)	93.2	289.3	338.1	0.0
cultured for 5 days (120 hours)	122.3	420.6	402.2	0.0
cultured for 6 days (144 hours)	118.3	535.8	372.2	0.0

While almost no production of pladienolide D was seen with the parent strain Mer-11107, pladienolide D was produced by the Mer-11107::pAOC::PapsmAB strain having the introduced pAOC::PapsmAB. This shows that direct production of pladienolide D can be achieved by introducing the pladienolide B 16-position hydroxylase gene psmAB into the pladienolide B-producing strain Mer-11107.

Example 5: Effects of β -CD addition on pladienolide D production by Mer-11107::pAOC::PapsmAB strain

The effects of β -cyclodextrin (β -CD) addition on productivity of pladienolide B and its 16-position hydroxide pladienolide D by the Mer-11107::pAOC::PapsmAB strain obtained in Example 3 were investigated. 300 μ L of frozen Mer-11107::pAOC::PapsmAB seed was inoculated into 30 mL of seed culture medium (soluble starch 2%, soy bean meal (ESUSAN-MEAT manufactured by Ajinomoto Co. Ltd.) 2%, yeast extract 0.3%, K_2HPO_4 0.1%, $MgSO_4 \cdot 7H_2O$ 0.25%, $CaCO_3$ 0.3%, pH

not adjusted), and incubated at 25°C for 3 days. 300 µL of the resulting seed culture broth was inoculated into 30 mL of main culture medium to which 2% β-CD was added (soluble starch 5%, glucose 1%, Pharmamedia 3%, adekanol LG-126 0.05%, β-CD 2.0%, CaCO₃ 0.1%, pH 7.5) or 30 mL of main culture medium without β-CD added (soluble starch 5%, glucose 1%, Pharmamedia 3%, adekanol LG-126 0.05%, CaCO₃ 0.1%, pH 7.5), and incubated for 3 days, 4 days, 5 days and 6 days at 25°C. After the completion of the cultivation, the resulting culture liquid was extracted by addition of 4 times the amount of acetonitrile. The amounts of pladienolide B and its 16-position hydroxide pladienolide D in the resulting extract were measured by HPLC. The measurement results are shown in Table 3. The HPLC measurement conditions were as follows.

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (φ 4.6 mm × 50 mm 3 µm)

Mobile phase: 45% to 55% methanol (0 to 5 min)

55% methanol (5 to 13 min)

55% to 70% methanol (13 to 21 min)

70% methanol (21 to 25 min)

Flow rate: 1.2 mL/min

Detection: UV 240 nm

Injection volume: 5 µL

Column temperature: 40°C

Analysis time: 25 min

Retention time: Pladienolide B 12.4 min

Pladienolide D 4.3 min

Table 3

	2% β -CD added		No β -CD added	
	Pladienolide B (mg/L)	Pladienolide D (mg/L)	Pladienolide B (mg/L)	Pladienolide D (mg/L)
cultured for 3 days (72 hours)	199.4	202.0	36.2	148.4
cultured for 4 days (96 hours)	362.8	350.5	93.2	289.3
cultured for 5 days (120 hours)	484.3	580.1	122.3	420.6
cultured for 6 days (144 hours)	433.1	774.1	118.3	535.8

As a result, addition of 2% β -CD resulted in a roughly 1.4 times increase in pladienolide D accumulation, showing that addition of β -CD has an effect on direct production of pladienolide D.

Example 6: Pladienolide production by Mer-11107::pAOC::PapsmAB strain (2)

300 μ L of frozen seed of the Mer-11107::pAOC::PapsmAB strain obtained in Example 3 was inoculated into 30 mL of seed culture medium (soluble starch 2%, soy bean meal (ESUSAN-MEAT manufactured by Ajinomoto Co. Ltd.) 2%, yeast extract 0.3%, K_2HPO_4 0.1%, $MgSO_4 \cdot 7H_2O$ 0.25%, $CaCO_3$ 0.3%, pH not adjusted) and incubated at 25°C for 3 days. 300 μ L of the resulting seed culture broth was inoculated into 30 mL

of main culture medium (soluble starch 5%, glucose 1%, Pharmamedia 3%, soy bean powder (J-Oil Mills: Honen SoyPro) 1%, adekanol LG-126 0.05%, β -CD 2.0%, CaCO_3 0.1%, pH 7.5), and cultured for 3 days, 4 days, 5 days, 6 days and 7 days at 25°C.

After the completion of the cultivation, the resulting culture liquid was extracted by addition of 4 times the amount of acetonitrile. The amounts of pladienolide B and its 16-position hydroxide pladienolide D in the resulting extract were measured by HPLC. The measurement results are shown in Table 4. The HPLC measurement conditions were as follows.

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (ϕ 4.6 mm $\times$ 50 mm 3 μ m)

Mobile phase: 45% to 55% methanol (0 to 5 min)

55% methanol (5 to 13 min)

55% to 70% methanol (13 to 21 min)

70% methanol (21 to 25 min)

Flow rate: 1.2 mL/min

Detection: UV 240 nm

Injection volume: 5 μ L

Column temperature: 40°C

Analysis time: 25 minutes

Retention time: Pladienolide B 12.4 minutes

Pladienolide D 4.3 minutes

Table 4

	Pladienolide B (mg/L)	Pladienolide D (mg/L)
cultured for 3 days (72 hours)	378.3	230.6
cultured for 4 days (96 hours)	601.4	567.9
cultured for 5 days (120 hours)	480.4	874.2
cultured for 6 days (144 hours)	387.4	985.4
cultured for 7 days (168 hours)	439.1	1215.9

50 mL of extract obtained by adding an equal amount of acetonitrile to the aforementioned culture broth was filtered, and the mycelia were washed with 30 mL of water. 70 mL of water was added to the filtrate, which was then extracted once with 100 mL and twice with 50 mL of ethyl acetate. The ethyl acetate layers were combined, washed twice with 50 mL of brine and dried over anhydrous sodium sulfate. After removing the solvent, the resulting residue was purified by thin layer chromatography (MERCK Silicagel 60 F254 0.5 mm developing solution; toluene:acetone=1:1) to obtain 12.8 mg of pladienolide B and 23.1 mg of pladienolide D. The ^1H -NMR analysis results for pladienolide B and pladienolide D were as follows.

Pladienolide B;

^1H -NMR spectrum (CD_3OD , 500 MHz): δ ppm (integration, multiplicity, coupling constant J (Hz)):

0.93 (3H,d,J=7.0Hz), 0.94 (3H,d,J=6.8Hz), 0.98
 (3H,t,J=8.0Hz), 1.12 (3H,d,J=6.8Hz), 1.23 (3H,s), 1.25
 (1H,m), 1.42 (2H,m), 1.53-1.70 (6H,m), 1.79 (3H,d,J=1.0Hz),
 2.10 (3H,s), 2.52 (1H,m), 2.56 (2H,m), 2.60 (1H,m), 2.70
 (1H,dd,J=2.4,8.3Hz), 2.76 (1H,dt,J=2.4,5.7Hz), 3.56
 (1H,dt,J=8.3,4.4Hz), 3.82 (1H,m), 5.08 (2H,d,J=9.8Hz), 5.60
 (1H,dd,J=9.8,15.2Hz), 5.70 (1H,dd,J=8.3,15.2Hz), 5.74
 (1H,dd,J=9.8,15.2Hz), 6.13 (1H,d,J=9.8Hz), 6.36
 (1H,dd,J=9.8,15.2Hz)

Pladienolide D;

¹H-NMR spectrum (CD₃OD, 500 MHz): δ ppm (integration,
 multiplicity, coupling constant J (Hz)):

0.93 (3H,d,J=7.0Hz), 0.95 (3H,d,J=6.8Hz), 0.98
 (3H,t,J=8.0Hz), 1.23 (3H,s), 1.30 (1H,m), 1.36-1.66 (9H,m),
 1.70 (1H,dd,J=6.4,14.2Hz), 1.82 (3H,d,J=1.0Hz), 1.90
 (1H,dd,J=6.4,14.2Hz), 2.10 (3H,s), 2.52 (2H,m), 2.62 (1H,m),
 2.72 (1H,dd,J=2.4,8.3Hz), 2.94 (1H,dt,J=2.4,5.7Hz), 3.55
 (1H,dt,J=8.3,4.4Hz), 3.82 (1H,m), 5.10 (1H,d,J=9.8Hz), 5.11
 (1H,d,J=10.8Hz), 5.60 (1H,dd,J=9.8,15.2Hz), 5.74
 (1H,dd,J=8.3,15.2Hz), 5.92 (1H,d,J=15.2Hz), 6.18
 (1H,d,J=10.8Hz), 6.57 (1H,dd,J=10.8,15.2Hz)

Example 7: Introduction of recombination plasmid
 pAOC::PapsmAB into Mer-11107 pldB::hyg strain

The S17-1/pAOC::PapsmAB strain obtained in Example 3
 was inoculated into 10 mL of LB medium (1% bacto tryptone,
 0.5% yeast extract, 0.5% NaCl) comprising 25 µg/mL

kanamycin and shaking cultured for 2 hours at 30°C, and the mycelia were collected, washed twice with 10 mL of LB medium and suspended in 5 mL of LB medium. This was the donor suspension.

While the donor suspension was being prepared, the Mer-11107 pldB::hyg strain prepared in Reference Example 1(8) was inoculated into 10 mL of TSB medium (Trypto-Soya broth: Nissui Pharmaceutical) and shaking cultured at 30°C for 5 hours, and the mycelia were collected, washed twice with 10 mL of sterilized water and suspended in 1 mL of sterilized water. This was the recipient suspension. 500 µL of the S17-1/pAOC::PapsmAB donor suspension was mixed with 10 µL of the resulting Mer-11107 pldB::hyg recipient suspension, and applied to Actino Medium No. 4 agar medium (Nippon Seiyaku). After incubated at 30°C for 18 hours, this was covered with 2.5 mL of SNA (0.8% nutrient medium: Difco, 0.4% agar) containing 2 mg/mL ribostamycin and incubated at 30°C for 7 days to obtain a ribostamycin-resistant pAOC::PapsmAB-transformed Mer-11107 pldB::hyg::pAOC::PapsmAB strain.

Example 8: Pladienolide production test of Mer-11107 pldB::hyg::pAOC::PapsmAB strain

The Mer-11107 pldB::hyg::pAOC::PapsmAB strain obtained in Example 7 and the parent Mer-11107 pldB::hyg strain as a control were tested for productivity of ME-265 and its 16-position hydroxide ME-282. 300 µL each of frozen seed of

Mer-11107 pldB::hyg::pAOC::PapsmAB and Mer-11107 pldB::hyg was inoculated into 30 mL of seed medium (soluble starch 2%, soy bean meal (ESUSAN-MEAT manufactured by Ajinomoto Co. Ltd.) 2%, yeast extract 0.3%, K_2HPO_4 0.1%, $MgSO_4 \cdot 7H_2O$ 0.25%, $CaCO_3$ 0.3%, pH not adjusted) and incubated at 25°C for 3 days. 300 μ L of the resulting seed culture broth was inoculated into 30 mL of main culture medium with 2% β -CD added thereto (soluble starch 5%, glucose 1%, Pharmamedia 3%, adekanol LG-126 0.05%, β -CD 2.0%, $CaCO_3$ 0.1%, pH 7.5), and incubated at 25°C for 3 days, 4 days, 5 days and 6 days. After the completion of the cultivation, the culture broth was extracted by addition of 4 times the amount of acetonitrile. The amounts of ME-265 and its 16-position hydroxide ME-282 in the resulting extract were measured by HPLC. The measurements results are shown in Table 5. The HPLC measurement conditions were as follows.

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (ϕ 4.6 mm $\times$ 50 mm 3 μ m)

Mobile phase: 45% to 55% methanol (0 to 5 min)

55% methanol (5 to 13 min)

55% to 70% methanol (13 to 17 min)

70% methanol (1 to 35 min)

Flow rate: 1.2 mL/min

Detection: UV 240 nm

Injection volume: 5 μ L

Column temperature: 40°C

Analysis time: 35 minutes

Retention time: ME-265 21.0 min

ME-282 15.6 min

Table 5

	Mer-11107 pldB::hyg::pAOC::PapsmAB strain		Mer-11107 pldB::hyg strain	
	ME-265 (mg/L)	ME-282 (mg/L)	ME-265 (mg/L)	ME-282 (mg/L)
cultured for 3 days (72 hours)	47.8	153.8	485.7	0.0
cultured for 4 days (96 hours)	63.5	222.1	907.0	0.0
cultured for 5 days (120 hours)	70.5	280.5	1534.4	0.0
cultured for 6 days (144 hours)	71.8	309.6	2031.3	0.0

While almost no production of ME-282 was seen with the parent strain Mer-11107 pldB::hyg, ME-282 was produced by the Mer-11107 pldB::hyg::pAOC::PapsmAB strain into which pAOC::PapsmAB was introduced. This shows that direct production of ME-282 can be achieved by introducing the ME-265 16-position hydroxylating enzyme gene psmAB into the ME-265 producing strain Mer-11107 pldB::hyg.

A filtration aid was added to about 90 mL of extract obtained by addition of an equal amount of acetonitrile to the aforementioned Mer-11107 pldB::hyg::pAOC::PapsmAB culture broth, which was then mixed and filtered through 60 ϕ Kiriya funnel (Kiriya filter paper No. 4) to isolate the mycelia. The mycelia were washed with 50 mL of water to obtain a mycelia isolation filtrate. The mycelia

isolation filtrate was extracted twice with 100 mL of ethyl acetate. After drying over anhydrous sodium sulfate, the solvent was removed to obtain the residue. This was purified by thin layer chromatography (MERCK Silicagel 60 F254 0.5 mm developing solution; toluene:acetone=1:1) to obtain 0.5 mg of ME-265 and 12.5 mg of ME-282. The ^1H -NMR analysis results for the ME-265 and ME-282 were as follows. ME-265;

^1H -NMR spectrum (CD_3OD , 500 MHz): δ ppm (integration, multiplicity, coupling constant J (Hz)):

0.87 (3H, d, J=7.0Hz), 0.90 (3H, d, J=7.0Hz), 0.94 (3H, d, J=7.3Hz), 0.97 (3H, d, J=7.0Hz), 1.08 (3H, d, J=7.0Hz), 1.17-1.21 (1H, m), 1.24-1.36 (2H, m), 1.42-1.52 (3H, m), 1.61-1.66 (3H, m), 1.74 (3H, d, J=1.1Hz), 1.89-1.96 (1H, m), 2.00 (3H, s), 2.41-2.47 (1H, m), 2.43 (1H, dd, J=5.5, 13.9Hz), 2.51-2.58 (1H, m), 2.56 (1H, dd, J=3.7, 13.9Hz), 2.65 (1H, dd, J=2.2, 8.1Hz), 2.72 (1H, dt, J=2.2, 5.9Hz), 3.51 (1H, dt, J=4.4, 8.4Hz), 3.75-3.80 (1H, m), 4.91 (1H, dd, J=8.8, 10.6Hz), 5.00 (1H, d, J=10.6Hz), 5.42 (1H, dd, J=9.2, 15.0Hz), 5.49 (1H, dd, J=9.2, 15.0Hz), 5.65 (1H, dd, J=8.4, 15.0Hz), 6.08 (1H, d, J=10.6Hz), 6.32 (1H, dd, J=10.6, 15.0Hz)

ME-282;

^1H -NMR spectrum (CD_3OD , 500 MHz): δ ppm (integration, multiplicity, coupling constant J (Hz)):

0.87 (3H, d, J=7.0Hz), 0.90 (3H, d, J=7.0Hz), 0.94 (3H, t, J=7.3Hz), 0.97 (3H, d, J=6.6Hz), 1.21-1.26 (1H, m), 1.29-1.37 (3H, m),

1.34 (3H,s), 1.44-1.52 (2H,m), 1.60-1.64 (1H,m),
1.65 (1H,dd,J=6.2,13.9Hz), 1.77 (3H,d,J=1.1Hz),
1.86 (1H,dd,J=5.4,13.9Hz), 1.89-1.94 (1H,m), 2.00 (3H,s),
2.43 (1H,dd,J=5.5,13.9Hz), 2.50-2.60 (1H,m), 2.56
(1H,dd,J=3.3,13.9Hz), 2.66 (1H,dd,J=2.2,7.7Hz), 2.89
(1H,dt,J=2.2,6.2Hz), 3.52 (1H,dt,J=4.8,8.4Hz), 3.75-3.80
(1H,m), 4.90 (1H,overlapped with D₂O), 5.01 (1H,d,J=10.6Hz),
5.42 (1H,dd,J=9.2,15.0Hz), 6.13 (1H,d,J=10.6Hz), 6.52
(1H,dd,J=11.0,15.0Hz)

Example 9: Preparation of plasmid pPepsmAB comprising ermE promoter connected to psmAB

The ermE promoter region was ligated to psmAB in order to cause expression of psmAB using the ermE promoter. The two primers eDM-SphF (SphI site added to 5'-terminal: see Sequence No. 27) and eDM-SphR (SphI site added to 5'-terminal: see Sequence No. 28) comprising the sequences shown below were synthesized based on the nucleotide sequence information of Sequence No. 2 for purposes of amplifying psmAB.

eDM-SphF: 5'-TCCCCGCATGCCGGAAGTACGGACATCAC-3'

eDM-SphR: 5'-CCCGGGCATGCAGACGACGACGTAGAGGAA-3'

These primers were used to carry out PCR under the following conditions.

(PCR reaction liquid composition: using Takara LA Taq (Takara Bio Inc.))

Sterile purified water:

17 μ L

2x concentration GC buffer II	25 μ L
dNTP mixed solution (2 mM each dATP, dGTP, dCTP and dTTP):	
	5 μ L
LA Taq	0.5 μ L
eDM-SphF (50 pmol/ μ L)	0.5 μ L
eDM-SphR (50 pmol/ μ L)	0.5 μ L
<i>Streptomyces esp.</i> A-1544 genome DNA	1 μ L

(Reference Example 2(1), 10 ng/ μ L)

(Reaction conditions: using T GRADIENT (Biometra))

95°C, 3 minutes

(98°C 20 sec, 68°C 4 min) 30 cycles

68°C, 5 minutes

The 1570 bp DNA fragment comprising psmAB that was amplified as a result of this reaction was collected with a QIAquick PCR Purification Kit (QIAGEN Co.). The resulting 1570 bp DNA fragment was digested with restriction enzyme SphI. The plasmid pSK117 (*J. Bacteriol.* Vol. 2002 184, 6417-6423) was similarly digested with restriction enzyme SphI. The resulting 1570 bp DNA fragment digest comprising psmAB and the plasmid digest were connected using DNA Ligation Kit Ver. 2.1 (Takara Bio Inc.). A plasmid pPepsmAB comprising the ermE promoter connected to psmAB was constructed in this way.

Example 10: Preparation of recombination plasmid containing psmAB connected to ermE promoter

The plasmid pPepsmAB prepared in Example 9 was

digested with restriction enzymes KpnI and BglII, and blunt-ended using BKL Kit (Takara Bio Inc.). The resulting DNA was electrophoresed (Mupid-ex: Advance) on 0.8% Molecular Biology Agarose™ (BIO-RAD), and the isolated 2060 bp DNA fragment (hereunder called PepsmAB) comprising the ermE promoter and psmAB was excised and collected and purified by QIAquick Gel Extraction Kit (QIAGEN Co.). The recombination vector pUC19aph::oriT::intphiC31 shown in Fig. 4 was also digested with restriction enzyme BamHI, and blunt-ended with BKL Kit (Takara Bio Inc.). The resulting vector was connected to PepsmAB using DNA Ligation Kit Ver. 2.1 (Takara Bio Inc.). Thus, the recombination plasmid pAOC::PepsmAB having PepsmAB inserted into the recombination vector pUC19aph::oriT::intphaiC31 was constructed.

Example 11: Preparation of recombination plasmid containing natural psmAB

A primer DM-BglF (5'-CGCATAGATCTTCACCCGAGCGGGTGATCA-3': see the Sequence No. 35) having a BglII site added to the 5'-end and a primer DM-BglR (5'-TCCCGAGATCTTGAAGGTCCGCGTCACCGT-3': see the Sequence No. 36) having a BglII site added to the 5'-end were designed and prepared with reference to the nucleotide sequence of the Sequence No. 2. Next, these two primers (DM-BglF and DM-BglR) and the A-1544 genomic DNA obtained in Reference Example 2(1) as the template were used to carry out PCR

reaction. The PCR reaction was conducted by repeating a three-stage reaction including denaturing carried out at 98°C for 20 seconds, annealing carried out at 63°C for 30 seconds and elongation carried out at 68°C for 4 minutes 30 times by using Takara LA Taq (Takara Bio Inc.) and a PCR amplifier (T Gradient, Biometra Co.). A roughly 3.5 kbp DNA fragment (hereunder called PopsmAB) comprising natural psmA and psmB was collected from the PCR amplification reaction liquid by agarose gel electrophoresis and SUPREC 01 (Takara Bio Inc.). The resulting DNA fragment was digested with restriction enzyme BglII. The pUC19aph::oriT::intphiC31 recombination vector shown in Fig. 4 was also digested with restriction enzyme BamHI. The resulting vector and PopsmAB were connected using DNA Ligation Kit Ver. 2.1 (Takara Bio Inc.). A recombination plasmid pAOC::PopsmAB comprising PopsmAB inserted into the recombination vector pUC19aph::oriT::intphiC31 was thus constructed.

Example 12: Introduction of recombination plasmids

pUC19aph::oriT::intphiC31, pAOC::PepsmAB and pAOC::PopsmAB into Mer-11107.

The recombination vector pUC19aph::oriT::intphiC31 and the pAOC::PepsmAB and pAOC::PopsmAB obtained in Examples 10 and 11 were each transformed into conjugated *E. coli* S17-1 (ATCC47055) by electroporation to obtain S17-1/pUC19aph::oriT::intphiC31, S17-1/pAOC::PepsmAB and S17-

1/pAOC::PoepsmAB strains. The resulting three strains were inoculated into 10 mL of LB medium (1% bacto tryptone, 0.5% yeast extract, 0.5% NaCl) comprising 25 µg/mL kanamycin and shaking cultured at 30°C for 2 hours, after which the mycelia were collected, washed twice with 10 mL of LB medium and suspended in 5 mL of LB medium. These were the donor suspensions.

While the donor suspension was being prepared, Mer-11107 was inoculated into 10 mL of TSB medium (Trypto-Soya broth: Nissui Seiyaku) and shaking cultured at 30°C for 5 hours, after which the mycelia were collected, washed twice with 10 mL of sterile water and suspended in 1 mL of sterile water. This was the recipient suspension. 500 µL of each of the donor suspensions of the obtained three strains above was mixed with 10 µL of the Mer-11107 recipient suspension, and applied to Actino Medium No. 4 agar medium. After incubated at 30°C for 18 hours, these were covered with 2.5 mL of SNA (0.8% nutrient medium: Difco, 0.4% agar) comprising 2 mg/mL ribostamycin, and incubated at 30°C for 7 days to obtain ribostamycin-resistant pUC19aph::oriT::intphiC31-transformed Mer-11107::pAOC, pAOC::PepsmAB-transformed Mer-11107::pAOC::PepsmAB and pAOC::PopsmAB-transformed Mer-11107::pAOC::PopsmAB strains.

Example 13: Pladienolide production test of Mer-11107::pAOC strain, Mer-11107::pAOC::PepsmAB strain and Mer-

11107::pAOC::PopsmAB strain

The Mer-11107::pAOC strain, Mer-11107::pAOC::PepsmAB strain and Mer-11107::pAOC::PopsmAB strain obtain in Example 12 were tested for productivity of pladienolide B and its 16-position hydroxylate pladienolide D. 300 μ L each of frozen seed of Mer-11107::pAOC, Mer-11107::pAOC::PepsmAB and Mer-11107::pAOC::PopsmAB was inoculated into 20 mL of seed medium (soluble starch 2%, soy bean meal (ESUSAN-MEAT manufactured by Ajinomoto Co. Ltd.) 2%, yeast extract 0.3%, K_2HPO_4 0.1%, $MgSO_4 \cdot 7H_2O$ 0.25%, $CaCO_3$ 0.3%, pH not adjusted) and incubated at 25°C for 2 days. 300 μ L of the resulting seed culture broth was inoculated into 30 mL of main culture medium (soluble starch 5%, glucose 1%, Pharmamedia 3%, $CaCO_3$ 0.1%, pH 7.5) and incubated at 25°C for 4 days. After the completion of the cultivation, the culture broth was extracted by addition of 9 times the amount of acetonitrile. The amounts of pladienolide B and its 16-position hydroxide pladienolide D in the resulting extract were measured by HPLC. The measurement results are shown in Table 6. The HPLC measurement conditions were as follows.

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (ϕ 4.6 mm $\times$ 50 mm 3 μ m)

Mobile phase: 45% to 55% methanol (0 to 5 min)

55% methanol (5 to 13 min)

55% to 70% methanol (13 to 21 min)

70% methanol (21 to 25 min)

Flow rate: 1.2 mL/min

Detection: UV 240 nm

Injection volume: 5 µL

Column temperature: 40°C

Analysis time: 25 min

Retention time: Pladienolide B 12.4 min

Pladienolide D 4.3 min

Table 6

	Mer- 11107::pAO C strain	Mer- 11107::pAOC: :PepsmAB strain	Mer- 11107::pAOC: :PopsmAB strain
Pladienolide B (mg/L)	243.0	309.0	241.3
Pladienolide D (mg/L)	0.0	54.8	18.8

While almost no production of pladienolide D was seen with Mer-11107::pAOC in which only the vector was introduced, pladienolide D was produced by the Mer-11107::pAOC::PepsmAB and Mer-11107::pAOC::PopsmAB strains having introduced pAOC::PepsmAB and pAOC::PopsmAB. This shows that direct production of pladienolide D can be achieved by introducing the pladienolide B 16-position hydroxylase gene psmAB into the pladienolide B producing strain Mer-11107.

Example 14: Preparation of pPampsmAB comprising variant psmA

pPampasmAB-CLM, pPampasmAB-CLLM, pPampasmAB-CLFM and pPampasmAB-TY having mutations introduced using QuickChange™ Site-Directed Mutagenesis Kit with the pPapsmAB prepared in Example 1 as the template were prepared in order to add site-directed mutations to psmA.

Primers R91C-F and R91C-R (see the Sequence Nos. 37 and 38) having the sequences shown below were first synthesized:

R91C-F: 5'-TTCGCGGCCGTCTGCGACCGGCGGGTG-3'

R9AC-R: 5'-CACCCGCCGGTCGCAGACGGCCGCGAA-3'

and a mutation was introduced using QuickChange™ Site-Directed Mutagenesis Kit with pPapsmAB as the template, replacing the #91 arginine of PsmA with cysteine to prepare pPampasmAB-C.

Next, primers R195L-F and R195L-R (see the Sequence Nos. 39 and 40) having the sequences shown below were synthesized:

R195L-F: 5'-TCGCAAGGGGCGCTCGAGCGGCTCGAG-3'

R195L-R: 5'-CTCGAGCCGCTCGAGCGCCCCTTCCGA-3'

and a mutation was introduced using QuickChange™ Site-Directed Mutagenesis Kit with pPapsmAB-C as the template, replacing the #195 arginine of PsmA with leucine to prepare pPampasmAB-CL.

Next, primers L403M-F and L403M-R (see the Sequence Nos. 41 and 42) having the sequences shown below were synthesized:

L403M-F: 5'-ACGATCCAGGGGATGATGGAAC TCCCCGTGA-3'

L403M-R: 5'-TCACGGGGAGTTCCATCATCCCCTGGATCG-3'

and a mutation was introduced using QuickChange™ Site-Directed Mutagenesis Kit with pPapsmAB-CL as the template, replacing the #403 leucine of PsmA with methionine to prepare pPapsmAB-CLM.

Next, primers R236L-F and R236L-R (see the Sequence Nos. 43 and 44) having the sequences shown below were synthesized:

R236L-F: 5'-AGCTGGACCGACTCGACGTGGTGGCGCTGG-3'

R236L-R: 5'-AGCGCCACCACGTCGAGTCGGTCCAGCTC-3'

and a mutation was introduced using QuickChange™ Site-Directed Mutagenesis Kit with pPapsmAB-CLM as the template, replacing the #236 arginine of PsmA with leucine to prepare pPapsmAB-CLLM.

Next, primers I244F-F and I244F-R (see the Sequence Nos. 45 and 46) having the sequences shown below were synthesized:

I244F-F: 5'-TGGCGCTGGCCGTCTTCCTGCTCGTGG-3'

I244F-R: 5'-CCACGAGCAGGAAGACGGCCAGCGCCACCA-3'

and a mutation was introduced using QuickChange™ Site-Directed Mutagenesis Kit with pPapsmAB-CLM as the template, replacing the #244 isoleucine of PsmA with phenylalanine to prepare pPapsmAB-CLFM.

Next, primers M112T-F and M112T-R (see the Sequence Nos. 47 and 48) having the sequences shown below were

synthesized:

M112T-F: 5'-CAGCGGCGGATGACGATCCCGTCGTTC-3'

M112T-R: 5'-GAACGACGGGATCGTCATCCGCCGCTG-3'

and a mutation was introduced using QuickChange™ Site-Directed Mutagenesis Kit with pPapsmAB as the template, replacing the #112 methionine of PsmA with threonine to prepare pPampsmAB-T.

Finally, primers R195Y-F and R195Y-R (see the Sequence Nos. 49 and 50) having the sequences shown below were synthesized:

R195Y-F: 5'-AGGGGCGCGCGAGTACCTCGAGGAGTACCT-3'

R195Y-R: 5'-GGTACTCCTCGAGGTACTCGCGCGCCCCTT-3'

and a mutation was introduced using QuickChange™ Site-Directed Mutagenesis Kit with pPapsmAB-T as the template, replacing the #195 arginine of PsmA with tyrosine to prepare pPampsmAB-TY.

Example 15: Preparation of introduction plasmid by double homologous recombination

In the belief that psmAB or variants thereof could be introduced by double homologous recombination using a nucleotide sequence comprising DNA participating in pladienolide biosynthesis that had already been determined in Reference Example 1 (7) determining the nucleotide sequence of the pladienolide biosynthesis gene cluster, along with neighboring DNA thereof, an introduction plasmid was prepared by the following methods.

The Four primers, pldout-L-Bgl2F, pldout-L-Sph1R, pldout-R-Sph1F and pldout-R-Bgl2R (see the Sequence Nos. 51, 52, 53 and 54) having the sequences shown below, were synthesized based on the nucleotide sequence downstream from DNA participating in pladienolide biosynthesis in order to amplify a sequence for purposes of double homologous recombination:

pldout-L-Bgl2F: 5'-GGGAGATCTGCAGGTCATCGGGGGGAAGAACCACA-3'

pldout-L-Sph1R: 5'-TCTCCAGCCGCATGCGGTCCCGGGTCACCGT-3'

pldout-R-Sph1F: 5'-CCCGCATGCGGCTGGAGATCATCCAGAGCGCA-3'

pldout-R-Bgl2R: 5'-GGGAGATCTAGAACATGCCGGGCCAGAGGCTGAC-3'

These primers were used to run a PCR under the following conditions.

(PCR reaction liquid composition)

Sterile purified water	30 μ L
2x concentration GC buffer II	50 μ L
dNTP mixed solution (2.5 mM each dATP, dGTP, dTTP and dCTP)	16 μ L
pldout-L-Bgl2F or pldout-R-Sph1F (50 pmol/ μ L)	1 μ L
pldout-L-Sph1R or pldout-R-Bgl2R (50 pmol/ μ L)	1 μ L
Mer-11107 total DNA (100 ng/ μ L)	1 μ L
LA Taq polymerase (5u/ μ L, TAKARA HOLDINGS INC.)	1 μ L

(Reaction temperature conditions)

95°C, 3 minutes	30
cycles	

(98°C 20 sec, 65°C 30 sec, 68°C 3 min)

72°C, 5 minutes

A 2.08 kb DNA fragment (DNA fragment LZ) was amplified from the reaction using pldout-L-Bgl2F and pldout-L-Sph1R, while a 2.73 kb DNA fragment (DNA fragment RX) was amplified from the reaction using pldout-R-Sph1F and pldout-R-Bgl2R. DNA fragments LZ and RX were purified with QIAGEN PCR Purification Kit (QIAGEN Co.), and digested with restriction enzymes BglIII and SphI.

Next, the two primers hyg-SpeSphF and hyg2-SphR (see the Sequence Nos. 55 and 56) having the sequences shown below were synthesized based on the nucleotide sequences for purposes of amplifying the hygromycin resistance gene:

hyg-SpeSphF: 5'-GGGGCATGCGGACTAGTACACCGTCGCCTCGGT-3'

hyg2-SphR: 5'-GCCGCATGCGTCAGGCGCCGGGGGCGGTGT-3'

These primer were used to run a PCR under the following conditions.

(PCR reaction liquid composition)

Sterile purified water	30 µL
2x concentration GC buffer II	50 µL
dNTP mixed solution (2.5 mM each dATP, dGTP, dTTP and dCTP)	16 µL
hyg-SpeSphF (50 pmol/µL)	1 µL
hyg2-SphR (50 pmol/µL)	1 µL
<i>Streptomyces hygroscopicus</i> JCM 4772strain	1 µL
total DNA (100 ng/µL)	
LA Taq polymerase (5u/µL, TAKARA HOLDINGS INC.)	1 µL

(Reaction temperature conditions)

95°C, 6 minutes

30 cycles

(98°C 20 sec, 63°C 30 sec, 68°C 2 min)

72°C, 5 minutes

The 1.21 kb DNA fragment (DNA fragment hyg2) amplified as a result of this reaction was purified with QIAGEN PCR Purification Kit (QIAGEN Co.) and digested with restriction enzyme SphI.

The DNA fragments LZ and RX digested with restriction enzymes BglII and SphI, the DNA fragment hyg2 digested with restriction enzyme SphI and shuttle vector pKU253 (see Fig. 3) digested with restriction enzyme BamHI were all 4 connected together with DNA Ligation Kit Ver. 2.1 (Takara Bio Inc.). Thus, a roughly 22 kb plasmid pKU253-LZ-hyg2-RX for double homologous recombination was constructed comprising pKU253 having inserted therein a roughly 6.0 kb DNA fragment comprising DNA fragment hyg2 inserted between DNA fragments LZ and RX.

Example 16: Insertion of psmAB or variants thereof into introduction plasmid pKU253-LZ-hyg2-RX by double homologous recombination

The pPapsmAB constructed in Example 1 and the pPampsmAB-CLM, pPampsmAB-CLLM, pPampsmAB-CLFM and pPampsmAB-TY constructed in Example 14 were digested with restriction enzyme SpeI. Each kind of DNA digested with restriction enzyme SpeI was electrophoresed (Mupid-ex:

Advance) on 0.8% Certified™ Molecular Biology Agarose (BIO-RAD), and the isolated 1.8 kbp DNA fragment comprising the *pIdA* promoter and *psmAB* or variants thereof (hereunder sometimes called *PapsmAB*, *PampsmAB-CLM*, *PampsmAB-CLLM*, *PapmsmAB-CLFM* or *PampsmAB-TY*) was excised and collected and purified with QIAquick Gel Extraction Kit (QIAGEN Co.). The resulting DNA fragments (*PapsmAB*, *PampsmAB-CLM*, *PampsmAB-CLLM*, *PapmsmAB-CLFM* and *PampsmAB-TY*) were connected using DNA Ligation Kit Ver. 2.1 (Takara Bio Inc.) to the *pKU253-LZ-hyg2-RX* prepared in Example 15, which had been digested with restriction enzyme *SpeI*. The respective DNA fragments (*PapsmAB*, *PampsmAB-CLM*, *PampsmAB-CLLM*, *PapmsmAB-CLFM* and *PampsmAB-TY*) were thus inserted upstream from the *hyg2* of *pKU253-LZ-hyg2-RX* to construct the roughly 24 kb plasmids *pKU253-LZ-hyg2-PapsmAB-RX*, *pKU253-LZ-hyg2-PampsmAB-CLM-RX*, *pKU253-LZ-hyg2-PampsmAB-CLLM-RX*, *pKU253-LZ-hyg2-PapmsmAB-CLFM-RX* and *pKU253-LZ-hyg2-PampsmAB-TY-RX* for purposes of introducing *psmAB* or variants thereof by double homologous recombination.

Example 17: Introduction of *psmAB* or variants thereof into Mer-11107 by double homologous recombination

The resulting *pKU253-LZ-hyg2-PapsmAB-RX*, *pKU253-LZ-hyg2-PampsmAB-CLM-RX*, *pKU253-LZ-hyg2-PampsmAB-CLLM-RX*, *pKU253-LZ-hyg2-PapmsmAB-CLFM-RX* and *pKU253-LZ-hyg2-PampsmAB-TY-RX* were each transformed by electroporation into conjugated *E. coli* S17-1 to obtain S17-1/*pKU253-LZ-*

hyg2-PapsmAB-RX, S17-1/pKU253-LZ-hyg2-PapsmAB-CLM-RX, S17-1/pKU253-LZ-hyg2-PapsmAB-CLLM-RX, S17-1/pKU253-LZ-hyg2-PapsmAB-CLFM-RX and S17-1/pKU253-LZ-hyg2-PapsmAB-TY-RX. The resulting S17-1/pKU253-LZ-hyg2-PapsmAB-RX, S17-1/pKU253-LZ-hyg2-PapsmAB-CLM-RX, S17-1/pKU253-LZ-hyg2-PapsmAB-CLLM-RX, S17-1/pKU253-LZ-hyg2-PapsmAB-CLFM-RX and S17-1/pKU253-LZ-hyg2-PapsmAB-TY-RX were inoculated into LB medium (1% bacto tryptone, 0.5% yeast extract, 0.5% NaCl) comprising 25 µg/mL of kanamycin and 100 µg/mL of hygromycin B and shaking cultured at 30°C for 2 hours, and the mycelia were collected, washed twice with 10 mL of LB medium and suspended in 5 mL of LB medium. These were the donor suspensions.

While the donor suspensions were being prepared, Mer-11107 was inoculated into 10 mL of TSB medium (Trypto-Soya broth: Nissui Seiyaku) and shaking cultured at 30°C for 5 hours, and the mycelia were collected, washed twice with 10 mL of sterilized water and suspended in 1 mL of sterilized water. This was the recipient suspension. 500 µL each of the S17-1/pKU253-LZ-hyg2-PapsmAB-RX, S17-1/pKU253-LZ-hyg2-PapsmAB-CLM-RX, S17-1/pKU253-LZ-hyg2-PapsmAB-CLLM-RX, S17-1/pKU253-LZ-hyg2-PapsmAB-CLFM-RX and S17-1/pKU253-LZ-hyg2-PapsmAB-TY-RX donor suspensions was mixed with 10 µL of the Mer-11107 recipient suspension, and applied to Actino Medium No. 4 agar medium (Nippon Seiyaku). After 18 hours of culture at 30°C, this was covered with 2.5 mL of

SNA (0.8% nutrient medium: Difco, 0.4% agar) comprising 2 mg/mL of ribostamycin. These were incubated at 30°C for 7 days to obtain ribostamycin-resistant pKU253-LZ-hyg2-PapsmAB-RX, pKU253-LZ-hyg2-PapsmAB-CLM-RX, pKU253-LZ-hyg2-PapsmAB-CLLM-RX, pKU253-LZ-hyg2-PapsmAB-CLFM-RX and pKU253-LZ-hyg2-PapsmAB-TY-RX transformants.

The resulting pKU253-LZ-hyg2-PapsmAB-RX, pKU253-LZ-hyg2-PapsmAB-CLM-RX, pKU253-LZ-hyg2-PapsmAB-CLLM-RX, pKU253-LZ-hyg2-PapsmAB-CLFM-RX and pKU253-LZ-hyg2-PapsmAB-TY-RX transformants were each inoculated into 10 mL of TSB medium containing no ribostamycin, and shaking cultured at 30°C for 24 hours. Each of pKU253-LZ-hyg2-PapsmAB-RX, pKU253-LZ-hyg2-PapsmAB-CLM-RX, pKU253-LZ-hyg2-PapsmAB-CLLM-RX, pKU253-LZ-hyg2-PapsmAB-CLFM-RX and pKU253-LZ-hyg2-PapsmAB-TY-RX transformant culture liquid was collected and, washed twice with 10 mL of sterilized water and suspended in 10 mL of sterilized water. Each suitably diluted suspension was applied to YMS agar medium (0.4% yeast extract, 1% wheat germ extract, 0.4% soluble starch, 2% agar, 10 mM calcium chloride) comprising 200 µg/mL hygromycin B, and incubated at 30°C for 4 days. Single colonies growing in YMS agar medium comprising hygromycin B were transplanted to YMS agar medium comprising 200 µg/mL of hygromycin B and YMS agar medium comprising 200 µg/mL of ribostamycin, and incubated at 30°C for 2 days. After the cultivation, hygromycin B-resistant,

ribostamycin-sensitive strains were selected. The resulting strains were strains having the hygromycin B-resistance gene *hyg2* and *psmAB* or variants thereof inserted downstream in the genome (between nucleotides 73362 and 73369 in the Sequence No. 1) from DNA participating in pladienolide biosynthesis, and were called Mer-11107-ZX::*hyg2*::*PapsmAB*, Mer-11107-ZX::*hyg2*::*PampsmAB*-CLM, Mer-11107-ZX::*hyg2*::*PampsmAB*-CLLM, Mer-11107-ZX::*hyg2*::*PampsmAB*-CLFM and Mer-11107-ZX::*hyg2*::*PampsmAB*-TY.

Example 18: Pladienolide production by Mer-11107-

ZX::*hyg2*::*PapsmAB*, Mer-11107-ZX::*hyg2*::*PampsmAB*-CLM, Mer-11107-ZX::*hyg2*::*PampsmAB*-CLLM, Mer-11107-

ZX::*hyg2*::*PampsmAB*-CLFM and Mer-11107-ZX::*hyg2*::*PampsmAB*-TY

500 μ L each of frozen seed of the Mer-11107-ZX::*hyg2*::*PapsmAB*, Mer-11107-ZX::*hyg2*::*PampsmAB*-CLM, Mer-11107-ZX::*hyg2*::*PampsmAB*-CLLM, Mer-11107-ZX::*hyg2*::*PampsmAB*-CLFM and Mer-11107-ZX::*hyg2*::*PampsmAB*-TY obtained in Example 17 was inoculated into 25 mL of seed medium (soluble starch 2%, soy bean meal (ESUSAN-MEAT manufactured by Ajinomoto Co. Ltd.) 2%, yeast extract 0.3%, K_2HPO_4 0.1%, $MgSO_4 \cdot 7H_2O$ 0.25% $CaCO_3$ 0.3%, pH not adjusted) and incubated at 25°C for 3 days. 300 μ L of the resulting seed culture broth was inoculated into 30 mL of main cultured medium (soluble starch 5%, glucose 1%, Pharmamedia 3%, SoyPro 1%, adekanol LG-126 0.05%, β -CD 2.0%, $CaCO_3$ 0.1%, pH 7.5), and incubated at 25°C for 6 days.

After the completion of the cultivation, the resulting culture broth was extracted by addition of 4 times the amount of acetonitrile. The amounts of pladienolide D, pladienolide H13 corresponded the 20-position hydroxyl pladienolide D, pladienolide C2 corresponded the 21-position carbonyl pladienolide D and pladienolide C3 corresponded the 20-position hydroxyl and 21-position carbonyl pladienolide D in the resulting extract were measured by HPLC. The measurement results are shown in Table 7. The HPLC measurement conditions were as follows.

(HPLC analysis conditions)

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (4.6 mm × 50 mm 3 μm)

Mobile phase: 20% to 35% acetonitrile (0 to 13 min)

35% to 100% acetonitrile (13 to 17 min)

20% acetonitrile (17 to 22 min)

Flow rate: 1.0 mL/min

Detection: UV 240 nm

Injection volume: 5 μL

Column temperature: 40°C

Analysis time: 22 minutes

Retention time: Pladienolide H13: 5.5 min

Pladienolide C3: 8.3 min

Pladienolide D: 9.0 min

Pladienolide C2: 11.1 min

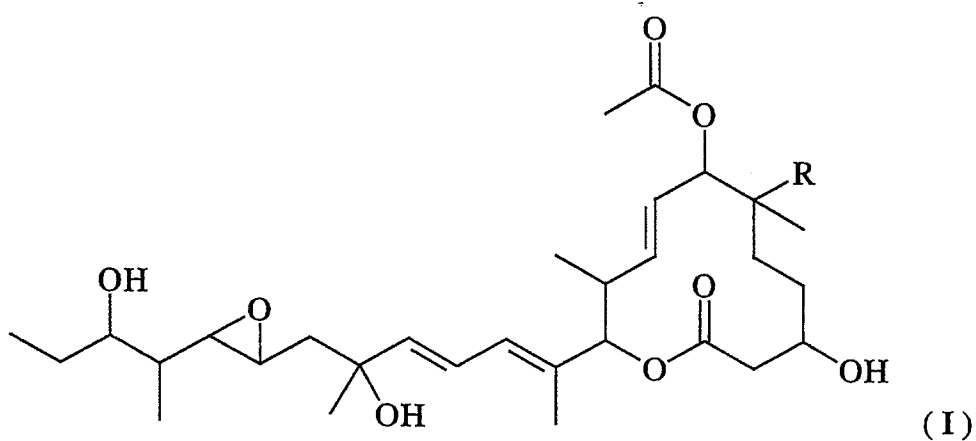
Table 7

mg/L	Mer-11107- ZX::hyg2:: PapsmAB strain	Mer-11107- ZX::hyg2:: PapsmAB-CLM strain	Mer-11107- ZX::hyg2:: PapsmAB- CLLM strain	Mer-11107- ZX::hyg2:: PapsmAB- CLFM strain	Mer-11107- ZX::hyg2:: PapsmAB-TY strain
Pladienolide D	1315.5	1412.8	1372.1	1371.3	755.6
Pladienolide H13	163.3	28.2	24.3	16.3	43.9
Pladienolide C3	93.8	17.5	50.4	28.5	10.6
Pladienolide C2	180.2	45.4	28.6	68.7	45.6
Pladienolide D degradation products (H13 + C3 + C2)	437.3	91.1	103.3	113.4	100.1

As a result, large amounts of the pladienolide degradation products pladienolide H13, pladienolide C2 and pladienolide C3 were produced together with pladienolide D by Mer-11107-Zx:hyg2::PapsmAB having introduced PapsmAB, while with Mer-11107-ZX::hyg2::PampsAB-CLM, Mer-11107-ZX::hyg2::PampsAB-CLLM, Mer-11107-ZX::hyg2::PampsAB-CLFM and Mer-11107-ZX::hyg2::PampsAB-TY introduced with PampsAB-CLM, PampsAB-CLLM, PampsAB-CLFM and PampsAB-TY, pladienolide D was produced but production of the pladienolide degradation products pladienolide H13, pladienolide C2 and pladienolide C3 was reduced to approximately one quarter. These results suggested that by introducing variants of the pladienolide B 16-position hydroxylating enzyme gene psmAB into the pladienolide B-producing strain Mer-11107 it is possible to achieve direct production of pladienolide D while reducing production of pladienolide D degradation products.

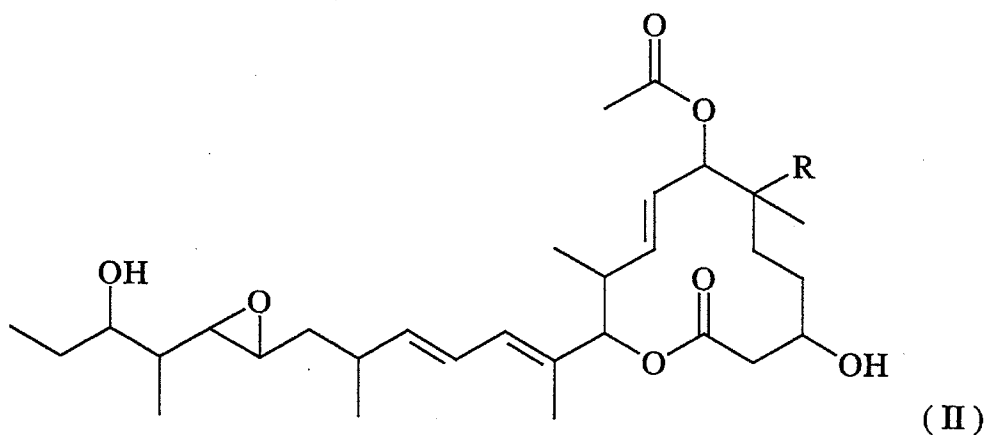
Claims

1. A recombinant microorganism, which is a microorganism having ability to produce 16-position hydroxyl macrolide compounds represented by the formula (I):



(wherein R represents a hydrogen atom or hydroxyl group), and, which comprises:

(a) DNA encoding wholly or partially a polypeptide participating in biosynthesis of the macrolide compounds represented by the formula (II):



(wherein R represents a hydrogen atom or hydroxyl group);

and

(b) DNA encoding wholly or partially a polypeptide having 16-position hydroxylating enzyme activity with respect to the macrolide compounds represented by the above formula (II).

2. The recombinant microorganism according to Claim 1, wherein the DNA described in the above (a) is DNA encoding wholly or partially a polypeptide having polyketide synthetase activity, a polypeptide having 7-position acetylase activity, a polypeptide having 18,19-position epoxidase activity and a polypeptide having transcriptional regulator factor activity.

3. The recombinant microorganism according to Claim 1 or 2, wherein the DNA described under (a) above is DNA comprising:

(a-1) the continuous nucleotide sequence of nucleotides 8340 to 27935 in the Sequence No. 1;

(a-2) the continuous nucleotide sequence of nucleotides 28021 to 49098 in the Sequence No. 1;

(a-3) the continuous nucleotide sequence of nucleotides 49134 to 60269 in the Sequence No. 1;

(a-4) the continuous nucleotide sequence of nucleotides 60269 to 65692 in the Sequence No. 1;

(a-5) the continuous nucleotide sequence of nucleotides 68160 to 66970 in the Sequence No. 1;

(a-6) the continuous nucleotide sequence of nucleotides

69568 to 68270 in the Sequence No. 1; and
(a-7) the continuous nucleotide sequence of nucleotides
72725 to 70020 in the Sequence No. 1, or a variant thereof.

4. The recombinant microorganism according to Claim 2
or 3, wherein the DNA described under (a) above further
includes DNA encoding wholly or partially a polypeptide
having 6-position hydroxylating enzyme activity.

5. The recombinant microorganism according to Claim 4,
wherein the DNA encoding wholly or partially a polypeptide
having 6-position hydroxylating enzyme activity is DNA
comprising the continuous nucleotide sequence of
nucleotides 65707 to 66903 in the Sequence No. 1, or a
variant thereof.

6. The recombinant microorganism according to any of
Claims 1 to 5, wherein the DNA described in the above (b)
above is DNA selected from the group consisting of:

(b-1) the continuous nucleotide sequence of nucleotides
1322 to 2548 in the Sequence No. 2;

(b-2) the continuous nucleotide sequence of nucleotides 420
to 1604 in the Sequence No. 3; and

(b-3) the continuous nucleotide sequence of nucleotides 172
to 1383 in the Sequence No. 4, or a variant thereof.

7. The recombinant microorganism according to any one
of Claims 1 to 6, wherein the DNA represented by (b) above
is a variant of DNA comprising the continuous nucleotide
sequence of nucleotides 1322 to 2548 in the Sequence No. 2

which has 1 or 2 or more modified sites selected from the group consisting of 1) to 6) below:

- 1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding an amino acid other than arginine;
- 2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding an amino acid other than methionine;
- 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding an amino acid other than arginine;
- 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding an amino acid other than arginine;
- 5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding an amino acid other than isoleucine; and
- 6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding an amino acid other than leucine.

8. The recombinant microorganism according to any one of Claims 1 to 6, wherein the DNA represented by (b) above is a variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the Sequence No. 2 which has 1 or 2 or more modified sites selected from the group consisting of 1) to 6) below:

- 1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding cysteine, methionine or proline;
- 2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding threonine or serine;
- 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding leucine, isoleucine, proline, tyrosine or phenylalanine;
- 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding leucine, isoleucine or proline;
- 5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding phenylalanine or valine; and
- 6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding methionine, cysteine or isoleucine.

9. The recombinant microorganism according to any one of Claims 1 to 6, wherein the DNA represented by (b) above is a variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the Sequence No. 2 which has 1 or 2 or more modified sites selected from the group consisting of 1) to 6) below:

- 1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding cysteine;

- 2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding threonine;
 - 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding leucine or tyrosine;
 - 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding leucine;
 - 5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding phenylalanine;
- and
- 6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding methionine.

10. The recombinant microorganism according to Claim 1, comprising a host having the DNA described under (a) above into which has been incorporated the DNA described under (b) above derived from a microorganism of a different species.

11. The recombinant microorganism according to Claim 1, comprising a host having the DNA described under (b) above into which has been incorporated the DNA described under (a) above derived from a different species.

12. The recombinant microorganism according to Claim 1, comprising a host having none of the DNA described under (a) and (b) above into which has been incorporated the DNA described under both (a) and (b) above derived from different species.

13. The recombinant microorganism according to any of Claims 1 to 12, wherein the recombinant microorganism having the ability to produce the 16-position hydroxyl macrolide compound represented by the formula (I) is a strain belonging to Streptomyces genus.

14. The recombinant microorganism according to any of Claims 1 to 13, wherein 50 mg or more of the 16-position hydroxyl macrolide compound represented by the formula (I) per 1 liter of culture broth can be produced when the recombinant microorganism is inoculated into a 250 mL Erlenmeyer flask containing 30 mL of the following medium, cultured at 25°C for 4 days with rotary shaking (220 rpm) and extracted by addition of 270 mL of acetonitrile, and the resulting extract is analyzed by HPLC under the following measurement conditions to assay the amount of the 16-position hydroxyl macrolide compound represented by the formula (I).

(Medium)

Soluble starch 5%, glucose 1%, Pharmamedia 3%, CaCO₃ 0.1%, pH 7.5

(HPLC measurement conditions)

Device: Shimadzu HPLC 10Avp

Column: Develosil ODS UG-3 (φ 4.6 mm x 50 mm 3 μm)

Mobile phase: 45%-55% methanol (0-5 min), 55% methanol (5-13 min), 55%-70% methanol (13-21 min), 70% methanol (21-25 min)

Flow rate: 1.2 mL/min

Detection: UV 240 nm

Injection volume: 5 μ L

Column temperature: 40°C

15. A method for producing 16-position hydroxyl macrolide compounds represented by the formula (I) or pharmacologically acceptable salts thereof, or hydrates of them, wherein the recombinant microorganism according to any of Claims 1 to 14 is cultured in nutritious medium, and the 16-position hydroxyl macrolide compounds represented by the formula (I) are collected from the culture broth thereof.

16. The method according to Claim 15, wherein a cyclodextrin is present in the culture broth.

17. The method according to Claim 16, wherein the cyclodextrin is a cyclodextrin selected from the group consisting of β -cyclodextrin, γ -cyclodextrin, partially methylated β -cyclodextrin, dimethyl- β -cyclodextrin, trimethyl- β -cyclodextrin, glycosyl- β -cyclodextrin and hydroxypropyl- β -cyclodextrin.

18. A variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the Sequence No. 2 which have 1 or 2 or more modified sites selected from the group consisting of 1) to 6) below:

1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding an amino acid

other than arginine;

2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding an amino acid other than methionine;

3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding an amino acid other than arginine;

4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding an amino acid other than arginine;

5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding an amino acid other than isoleucine; and

6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding an amino acid other than leucine.

19. A variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the Sequence No. 2 which have 1 or 2 or more modified sites selected from the group consisting of 1) to 6) below:

1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding cysteine, methionine or proline;

2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding threonine or serine;

- 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding leucine, isoleucine, proline, tyrosine or phenylalanine;
- 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding leucine, isoleucine or proline;
- 5) a modified site wherein the sequence atc of nucleotides 2054 to 2056 is modified to a codon encoding phenylalanine or valine; and
- 6) a modified site wherein the sequence ctg of nucleotides 2528 to 2530 is modified to a codon encoding methionine, cysteine or isoleucine.

20. A variant of DNA comprising the continuous nucleotide sequence of nucleotides 1322 to 2548 in the Sequence No. 2 which have 1 or 2 or more modified sites selected from the group comprising 1) to 6) below:

- 1) a modified site wherein the sequence cgc of nucleotides 1592 to 1594 is modified to a codon encoding cysteine;
- 2) a modified site wherein the sequence atg of nucleotides 1655 to 1657 is modified to a codon encoding threonine;
- 3) a modified site wherein the sequence cgc of nucleotides 1904 to 1906 is modified to a codon encoding leucine or tyrosine;
- 4) a modified site wherein the sequence cgc of nucleotides 2027 to 2029 is modified to a codon encoding leucine;
- 5) a modified site wherein the sequence atc of nucleotides

2054 to 2056 is modified to a codon encoding phenylalanine;
and

6) a modified site wherein the sequence ctg of nucleotides
2528 to 2530 is modified to a codon encoding methionine.

21. A polypeptide encoded by the variant of DNA
according to any of Claims 18 to 20.

SEQUENCE LISTING

<110> Eisai R&D MANAGEMENT CO., LTD.

<110> Mercian Corporation

<120> Recombinant microorganisms and methods for producing macrolide compounds using the microorganisms

<130> 06011PCT

<150> 2005-154114

<151> 2005-05-26

<160> 56

<210> 1

<211> 74342

<212> DNA

<213> Streptomyces sp.

<400> 1

cgattttgca	ccttgtccat	cgctggtggt	gtgaggcatg	ctcctattgg	aacataaaac	60
ctctgaacct	ttaagagggt	atggcggagg	ctttcgacgc	gacacgaggg	agaagcggat	120
gagaatcgtg	gggattcacc	gggaggggcg	aggcatagag	gtggcccggc	tgtcggacga	180
cgggcggcgg	gcagtcgtgc	tggccccgct	cgaagtcttc	tgggcccagc	ccaccggcca	240
tctggcgcgc	ggggacggtg	gaccagtcgt	cccgggtgtc	gogggtggagc	tggtaccgcc	300
ggttctgccg	gacgcgcggg	tgatctgcat	cgggctcaac	tacctcaagc	atgtggccga	360
gggaacctac	cgcgaccagg	aagtccccga	gcacccacac	ctgttcgccc	gctggacacg	420
gtcgtgacc	gtggacggag	ccgaggtccc	ggtgccctcg	gacgaggccg	ggctggactg	480
ggagggtgag	gtggtggcct	gggtggggcg	accactcgtg	gacgccacgc	cggaggaggc	540
gctgaccgcc	gtcatcggct	actccctctt	caacgacctc	acctcccggc	gggctcagaa	600
gctcacctct	cagtggaccc	tgggcaagaa	cggggacaac	tccggcccgc	tcgggcccga	660
ggtgccggct	gccgaggtgg	gcgacctgcg	cgacgggctg	cgggtacaga	cccgggtcaa	720
cggggagacg	atgcaggatg	gcagcacgga	cgagatggtc	tacaccgttg	gtgacacgct	780
cgcgcacatc	tcccgcacct	tcctcctgcg	tccggcgacg	ctgctggcga	cgggcacccc	840
gtccggagtc	ggctacgccc	ggaccccacc	gcagctcctg	cagccgggag	acgtcgtcga	900
ggtggaggtc	gaacggctcg	gcgtgctgcg	caaccccgtg	gtgtccaacg	acgcccggct	960
gcgcgcaccc	aagtgaggac	gcaagaggcc	ccgcgcccgc	ccgcggaacg	cgggtgctcg	1020
ccctgcggca	cacgcgcag	gacacacctg	gtcacctgcc	tgcgtcgccg	ggtcctgcgg	1080
tggggcacccg	ttgaccgtgg	tcaaggacta	caccgagaac	cagaagcgcc	cggactgagc	1140
gcggcccgat	cgtgggagca	ttccgaagcg	aggcggcggc	gctccgcgcc	gccttgctca	1200
ggatgcgttt	cctcccgggc	gtgaacgccc	ggggttcctc	gcacggtcag	gctgagactt	1260
ctctcacatg	gggagtcacg	tcgtcggccg	gatctgcggc	cggagccggt	accggcggcg	1320
agggctggaa	cggcgaactc	acgtactcct	ccacgagcgg	attacggtgc	gggacacgaa	1380
aggctcggaga	cggggggcgc	gctcggggcg	gggattagcg	gtaggggtgg	ctgatgggtg	1440
gctggcgccg	ggcggtgccc	aggacgttgt	caggggaagcg	ggagccggcg	acgtagtggc	1500
ggcccttggt	ctgccccacc	gccgtcagca	gctccgcctt	ggtcaggggc	tgcaggctcc	1560
gtgtggcctg	ctgggtgttg	agggcctcgg	ctcgtctata	gcgcgagcgg	cgaacccggc	1620
ccaccatcgc	cacctcgtgc	agcgcggtga	tctgccgctc	ggtgatcccg	aggctgtcgg	1680
cggcgtccat	gagctgcctc	cagcagtcgt	tggagcggtc	gacgcggcgc	tgcacccgct	1740
gggtctgctg	gtggtaggcg	agcaggttga	agcggatcca	cggcccgggtg	tcgcgttcg	1800
gtgagtagac	cgggcgcgcc	acctcgcgca	gcgctttgta	gtactcccag	gtgttgcccg	1860
gcatgcccag	ccactcctcg	atggaggaga	actccggcgc	cagcactccg	ccccgggoga	1920
tgaccagcgt	ctgcagggag	cgggacatcc	gcccgttccc	gtccgaccac	gggtggatct	1980
tcaccagggt	caggtgcgcc	atcgcgcgcc	gcaccaggac	gtgggcgtcc	aggtcgcggt	2040
cgttgagcca	gtccaccagc	tcgccatca	ggcccggcag	caggtcggcg	tccgggcctt	2100
cgtagtccgt	ggccagctcg	tcgccggggg	cggtagtcgc	gatcgcagtg	cgacgccact	2160
ggccggccag	ccgcagcgga	tgggtgtggc	cctgcagcat	ccagtgcagc	gagttagaca	2220
actccttgct	gtagctgaag	tcgcctacgt	cgtgcaggga	ctggatgtag	gccatcgoc	2280
gctggtaggc	cagcgtctcg	gccttgttct	cctcgttgcc	atccacggca	tcgcgttcgc	2340
cgtccatcag	gtccgcgaca	tccttcgcat	cgacctggta	accttcgatg	gtgttggaac	2400
ccgcgatcgc	gctggccgtc	agcgccttgc	gcagatcctg	cgtccacttc	gtcggcacct	2460
gctgcaccgc	gtggcgcagc	tgctcctgca	gggaattgat	ctcctccagg	accggcggt	2520

cgctcgggtggt	gaggtgaggc	gtctgataca	gcataagtga	atgataacctg	actcctatca	2580
ttccttcaac	gctgcggctt	catcacctgc	cgtccacgat	gaacatgaac	tcgcccagct	2640
cgacgagcag	ggccaccggc	ctgccgaccc	acgctggtoa	ctgatcaaca	caccaggtca	2700
acgagtacgt	ctgcaggact	ggatggagtg	acaggcagcg	cggggcagcc	aagcaggcgg	2760
cgcgaccacg	taacggaccg	gcgaaccoga	gctgagcacc	ggtgagctcc	gcgaccggct	2820
gactgccctc	agctacaaca	tcgagaacag	ctggccgagc	ttgcccggag	ccggcggcgc	2880
ggatcacgtc	cgccgtgtgt	gcgagggcg	cgccgcttcc	ggcgggtgtc	tcgacgccag	2940
gcccgcgaag	ttgctgagcg	acgcggcggt	ccggacgtcc	gcacgtcgcg	gatcccgaga	3000
cgaggggtga	gccaaagccg	cgggatcacg	ccgccgcgca	tgatgcctac	cacggcttgc	3060
cgatgtggag	gctgggcaga	gatccctca	gcctccggta	cgtcgacggg	caggcacccg	3120
cctggccgga	gctccgtgcg	gggcccgttg	agcggagtgt	gccccgcac	cggagacccg	3180
ccaggactcc	gtgcgaccaa	aagactcttc	gtctcacctg	gtctcacaca	tctctcaca	3240
acgagagcgc	acgggagcac	gcgagaacct	ctgacctgga	cattcgctgt	cgccacggac	3300
cacggcggac	caggtagaca	ttacataggc	taggactoga	ttctagtgt	caagtcaaat	3360
gcccagggtca	gaggctgttt	ttgctgttgc	ccgagggtcg	atatctcaca	tatttattgc	3420
tggaactgcg	gcacaaggtc	gagatccatc	ggatgacggg	gcgggtcgtc	gagtgtcga	3480
tcgacgaggt	ggaccgcccg	cgggtatcca	aggacgcgc	cctgcggatc	ccgccgaccg	3540
agctggaggt	caggcgctgt	tcacggtgt	ggttgccgag	ctggcgactt	gcccgaagtt	3600
cgcctcacc	gcgcggaact	acaccgcac	gaaactgatg	tcgtaggtcg	ggctcggggt	3660
gagcgaggca	tgcaaaactc	acctggccga	catcaagtgg	gacctgggcc	gcttcggcaa	3720
gctccatgtg	cgtaacggcg	agggcgcccg	cggctcgggt	ccgcgcgagc	ggatggtgcc	3780
gctgatcaac	ggcgccgacc	gcacgtctgc	gtggttcatc	gaggacgtcc	ggggccagtt	3840
cgacgacgc	cgcacccgcc	ccggtgcccc	gctgttcccc	tctgagcgca	agaacgccgt	3900
cggctcctcg	cgctcgtcgc	gcgacgacgc	gttgccgaac	ggtctggccg	ccgcggcgga	3960
ggctcgcgtcc	aaaagggtgt	gtaaaaggcg	acgagcgaaa	gtttcgtact	gaacagccaa	4020
gttcgtacga	aaggaccac	gctcgtgtag	tcacgagtgt	cacccatcga	ggcgatccat	4080
gcccagatcg	atgcctctti	cgcctagtac	tgcaacggtg	tttgccgtga	cggttgggca	4140
ggctggctgt	tggtctgagc	atgggtgggg	accttgcgtg	tgtcgggggt	tgggccagtg	4200
aagtggacgc	tgtgcacgag	aggttgcgtc	accggttttc	cagggcggag	ccgcgggagt	4260
cggcgcttgc	ctatatgcgg	ggactgaccg	ctccgctgga	gcggaagaac	ggctggacac	4320
tggccgaaca	ggccgggtcat	gtcgtccggg	accgtattca	tcgactgctg	aaccggatcg	4380
agtgggaagc	cgatgaggtc	ctggccgatg	tgccgcgacta	cgtcatggag	aaacctcggc	4440
accccagagg	cctcatcgtg	gacgacaccg	gcttcttgaa	gaaggggacc	cgttcggcag	4500
ggggccggcg	tcagtactcc	gggaccgcgg	gggcctgtat	ttcaattact	gctcggtgaa	4560
gggtgtttgg	cgggacaatg	agatgtgact	gcgcgcacgg	ccgctcgggc	ctcagttgtg	4620
gctttgaggg	cgaagttcga	tcagttcctt	ccccatctcg	atgagcggcg	ccgtcggatc	4680
tacctggcca	gcgaggccgc	cgcgcttggc	cacggcgggga	tcacgctggt	ggccaccgct	4740
tcgggcgcca	gtgcggccac	catcgacgc	gggatcgccg	agctgtccgg	gcacactctg	4800
ccggccgggc	ggatccgggc	tcggggagcc	ggccgcgaagc	cgggtcacggt	caccgacccc	4860
ggtctgctgc	ccgcgcttga	agctctgac	gagccgcaca	cccggggcga	tcgggtctcg	4920
ccattgcgct	ggaccacgct	ctcgtcgogg	tccttgccct	cggcgctgac	cactcagggc	4980
caccgggtca	gcgcggcgac	cgtcggaacgc	ctgctacatg	ccctgggata	cagcctgcag	5040
ggcaccgcca	agaccacgga	aggggccagt	catcccagacc	gggatgctca	gttcacgcac	5100
atcaaccagg	ccgcgcggga	cttcctcgaa	gacaaccagc	cgggtgatcag	cgtcgacacc	5160
aaggccaagg	agtggctcgg	caaccgcgac	cgaccgggac	gcacctggcg	accgggcaag	5220
aaccttatcc	gtgtggactg	ccacacgttc	accagccagt	accagccagt	agccatcccc	5280
tacgggatct	acgacatcgc	tcgcaacacc	ggctgggtca	acgtcgggac	cgaccacgac	5340
acgggcgagt	tcgcgggtga	atccatccgc	cgtggttgtg	agcagcacgg	acgcggcgac	5400
caccgggacg	ccggccgact	gctcatcacc	gccgactcgc	gtggttccaa	cgacccccgc	5460
cgttgacat	ggaagaagca	tctgcgcgc	ttcgccctgg	aaagcggact	cgagatcacg	5520
gtctgccact	tcacacccgg	aacatcgaag	tggaacaaga	tcgaacaccg	gatgttctgc	5580
cacatcaccc	cgaactggcg	cggcaggccc	ctgaccagct	accaggtcgt	catcgagacc	5640
atcgccgcca	cgaccactcg	caccgggctc	agcatcgcg	ccgaactcga	caccggccga	5700
tacgacctgg	gcaccacagt	cccacccgcc	gagttccaag	ccctgccaat	cacaccccac	5760
accttccacg	gcgactggaa	ctacaccctg	gcaccactcg	caccccggt	gcccagaccg	5820
gaccgagcg	gacaacggat	cgaccccgcc	ctgaccacga	tgctcaccga	ccgggccctg	5880
acggcgatgt	cacgctccgc	cttcgaccac	ctggtcgcca	tctcggaacc	gtactgggac	5940
gccctggccg	aggcggcatt	ccaacgacgc	ttccaccgcc	cacgcagcta	cctccaccgg	6000
cagaccagca	gcctcgacca	ctaccaccgc	ctgctgacgc	ccctgttacg	ccgcccgaga	6060
gccgtcacca	gcacactgct	ggcccagctc	ctgaacgtcg	gccgcaccaa	cctgtccaac	6120
cagttccaag	acggccaccg	cctcctggac	ctgcaccgca	tcgcggtcac	tcgctatct	6180
ggagccccgg	ccgcgaccc	cgcccaacta	caagcccgcc	taccgccaca	cgacgacacc	6240
cgcacagatc	aactctgaca	gttattcaga	cacaggcccc	cgggcggatc	gagaactccc	6300
aggtcgccgt	ctacctggtc	taggcagggtg	ccgggggcca	cgcggcggtg	gaccgggaac	6360

tgtagctgcc	ccgttcctgg	acctgtgacc	agggccgctg	caggcgccgc	gggctcggcg	6420
aggacatcgt	cttcgccacc	aagccggagc	tggcccgcac	gatgatcgaa	cggttcctgg	6480
acgccggaca	ccacgtgggc	tgggtcgcgtg	gcgacgaggt	ctacggcggc	aaccogaagc	6540
tgcgatctgc	gctggaggta	cgcggcctcg	gctatgtcct	cgcgggtggcc	tgctcggccg	6600
aagtcaccac	caaggcaggc	aagttccgag	ccgacacgct	ggcggcgaag	gtaccgaagc	6660
gggcctggca	gaagctgtcg	gcaggcgcgg	gagccaaggg	caaccgcttc	tacgactggg	6720
ccgtcgtcga	cctggccgag	cccggccccg	gccaccggca	gctgctgata	cgcgcgaacc	6780
gccgcaccgg	tgaactggcc	tactaccgat	gocactccac	ctcaccggtc	ccgctcgcga	6840
ccctgggtcag	ggttgccgga	tcacgggtggc	gggtggagga	gacattccag	accgagaagg	6900
gcctggccgg	cctggacgag	caccagctcc	gccgtacccc	ctcctggggc	cgttggggca	6960
ccctcgcatt	gctcggccac	gctttcctcg	ccgtcgtccg	cgcgcacgaa	cacaccggcc	7020
cgacccccga	cgacctcatt	ccgctgacct	gcaacgagat	ccagcacctg	ttcctcgcgc	7080
tcgtcgtcca	gccgtgttcc	aacgtcgcgc	accgcctcgc	ctgggtccgag	tggagacgac	7140
gtcatcaagc	ccgatcacgc	accagtcaact	accggcgaca	agccgcaact	gagacatgaa	7200
gatcacgata	tacagctgga	gtattaggctc	cccgaacaac	aaccggggac	agtctaagt	7260
ggcgtcactt	tcgtatcgaa	ccagcaaaat	ttggaggga	cgatattcga	ctcgcatagc	7320
cttacggctg	gagcatatta	cgaccaggctc	aatgaattgc	tccccgccat	gcattctcag	7380
gcgggatgcc	gtcggagggg	cgagggtctg	tatgatgtcg	ttgacatcgg	ccgcagtggg	7440
gatggccttg	agcggagtg	cgtgtccgtc	acagatcagg	tggtgtctgc	tgccggtcct	7500
gcgccggctg	accggcgcag	gaccggcgtc	ggctcccccc	tttcgcgcgg	acatgagagc	7560
cgtccacgca	cgcgcgtgac	cagtcgagtt	cgcggccgcg	gttgagttcg	gcgagcggga	7620
tgccgtgcga	ccctgggtctg	ctgccatcgt	tccagccgcc	gcaccgctgc	tctcgcagat	7680
ccgtcgcata	gtccggccgg	gacgtcacc	gcgacacctc	ttcgaggggc	agggtacccc	7740
gacgcgcgac	ggccagcggg	actgcatcaa	ctcgatctct	ctgcggctga	ttcccgacga	7800
gggttgagat	ccgtatcgac	acaccacgtc	cgtggacgtg	gcccacgggc	gtccgggtcaa	7860
gcgtgagtg	cgtggcatcg	atctcctgcg	cgtcagtgga	ccgcgcgagg	cgggcccga	7920
tcggcgagtt	ggacggggtc	ggggcgagcc	tgcttggtct	tcttgccgtg	caactcgacg	7980
acgtggccgt	gctgctggcc	gtagccgtgt	cctgcccgcc	gggcccgttg	ctgtccgggg	8040
ctgccgcgct	ccccgcggtc	gttgccgcca	ccaggccgca	cgcactgag	cagctgaccc	8100
ggcgcgggact	cgacgaggcc	gccttcgtca	cogaactcgt	cctgagggca	ccgggtcacc	8160
aacgccttcc	gttgctcccc	agacctctct	gggaagatca	atagggggcc	gtaaaggggg	8220
ggtgagggtt	gaaaaagggg	gggtattcaa	aatagggctg	agtccgctgc	aaaaatcttg	8280
agaccggctg	gaacgggtgt	agctgaatga	ctgaatcgaa	tgaattcacg	tccgaagcgg	8340
tgctcagtg	ggctgatgat	gcaatagcca	tcacggcat	gtcttgccgg	ctgccgcgag	8400
cagtcaatcc	tcaggagtgc	tgggaactcc	tgaggaatgg	tgagagcggg	attaccgagg	8460
tgccgccccca	gcgggtgggac	gcgaactccc	tcttcgatgc	ggaacgggtc	acgcccgga	8520
cgatgaatac	acgctggggc	gggttcacgc	acggcgtgga	ccagttcgac	cccgcttct	8580
tcgggatctc	ctcccgcgaa	gcggtcgcca	tggatccgca	gcaacggctc	gtactggagc	8640
tgagctggga	ggccctggag	gacgcgcgaa	togtcccgga	gcgccttcgc	cacaccgcta	8700
ccggtgtctt	cgtcggcgcg	atctgggacg	actacgcata	attgatgagc	gcgcgaggcc	8760
gagaagcggg	gacccatcac	accgtgaccg	gtacgcaccg	cagcatcatt	gccaacgggg	8820
tgctgtagcc	cctcggccta	caggggcccga	gcatggcggg	ggactccggg	cagtctcgt	8880
cactgggtctc	cgtccatctg	gcctgcgaga	gcctgcgcag	gggggagttc	acgctcgcgc	8940
tggccggcgg	ggtgaatctc	aaccttcttc	cggagagcac	catcggcatg	gcgaagttcg	9000
ggggctcctc	ccccgatggc	cgtcgttca	ccttcgacac	ccgcgccaac	ggctacgtgc	9060
ggggtgaggg	cggcggtgtg	ctcgtcctca	aaccgctggc	ggacgcgata	gcggaccagg	9120
acccgatcta	ctgcgtcatc	cgtggcagcg	ccgtcaacaa	cgacggttcc	ggtgagaacc	9180
tgaccacgcc	gaactcccag	gcgcaggcag	ctgtgctgcg	cgaggcctac	cgcgcgcgcg	9240
gcgtggaccc	ggcccaggtc	cagtacgtgg	aaactgcacg	taccgggacc	cctgtcggcg	9300
acccgattga	agccgaggcc	ctcggcgcg	tgatcgggtg	cgcgcggccg	ccgggtgacc	9360
ccctgtgggt	gggatcggcg	aagaccaaca	tggccatct	ggaggccgcc	gccggcatcg	9420
ccggcctgct	caaggtcgtg	ctgtccatca	gccaccggga	gctcccggcc	agtctcaact	9480
tcgccacggc	caatccgcgg	attccactgg	actccctgaa	cctgcgcgtg	ggcgacgagc	9540
tcacatcgtg	gccgtctgcc	ggtcggccga	tgctcgcggg	tgtagcgcg	ttcggcatgg	9600
gcggtaccaa	cgcacacgcc	gtggtcgaa	aatctcccg	agcagcgcgg	cagattccgg	9660
ctcccggagg	cagcccgacg	gatcaggggg	ggcgggtg	gtggtgtgtg	tcgggtgggt	9720
cgggtggcgg	ggtcggggg	caggcggcgc	ggtgtgtgtg	gcactcggag	ggtcgggtgg	9780
gtctgcgtgc	ggtggatgtc	ggctggtcgc	tggccacgac	tcgttcogtg	ttccctcact	9840
gtgctgttgt	cgttgccgac	gatggtgggt	acggccagag	tctcgcgcgc	ctggccgcgg	9900
gttccgtgga	tgccgggggtg	gttgaggggc	ttgcccgtgt	gagtggcaag	acgggtgtcg	9960
tcttccccgg	tcagggttcg	cagtgggtgg	gtatggccgt	tgagctgctg	gacggctcgg	10020
aggttttcgc	cgagcatatg	gccgcctcgc	ccaggccct	ggaaccgttt	gtgggtgtgt	10080
ccctggaggga	tgtcctgcgt	cagggtggacg	gtacgtgggt	actggatcgt	gtggatgtgg	10140
tccagcctgt	gctgtggg	gtcatggtct	cgtcgcggg	actgtggcag	gcacatggcg	10200

ttgagcctgc	tgcggtgctg	ggccactccc	aaggtgagat	cgctgaggct	tgcgtggcgg	10260
gtgcgctgag	tctggaagac	ggagcccggg	tggtggctct	tcgcagccgc	gcatcgccg	10320
aggcccttgc	gggcatggc	gggatgctgt	cgatagccgc	ccccccacc	gaagtacgg	10380
ccctgatcac	cccctggggc	aggcagatca	ccattgccac	ggtcaacgga	ccgcattcgg	10440
tggtggctgc	aggagaccct	gacgcgctcg	aggcactccg	cggcgaactg	gagaccctg	10500
gtctccgcaa	tcgtcgcatc	ccggtcgact	acgcctcaca	cacccctcac	gtcaggcgga	10560
tccgtgaacg	gctcctggcc	gacctggcag	tgatccagcc	acgtgccgcg	agcattcccg	10620
tgctgtccac	cgtcacccggc	gcatggctcg	acaccaccgt	gatggacgcc	gagtactggt	10680
accgcaacct	acgtcagacc	gtggagtctg	aagcagccac	ccgcactctc	ctcgaccagg	10740
accaccgcta	cttcgtcgag	atcagcccgc	accccgtact	caccaccgcg	atccaggaaa	10800
ccctcgacgt	cacagacacc	gcccgcgtcg	ccaccggaac	cctgcgacgc	aacgaaggca	10860
gcctccggcg	tttccagctc	gcccctggcg	aactcgtcac	ccgtggcctc	accccgcact	10920
ggcccgcctt	ctatcccgac	gcccgcacac	cggacctccc	cacctatccc	ttccaacgcg	10980
agcgctactg	ggtcggcagc	tcctcgggtc	gggacgcggc	gcccggctccg	caaccggacc	11040
cggcaactgg	gcgagccggc	ggtccggctt	cgggcggggc	cggcgctgat	ggcggcgacg	11100
ggcccgcgga	gctgtggct	ctggtgcgtg	cccacgtggc	cgtggtgctc	ggtgagacga	11160
cgcgggacag	tgatgatccg	aaactgacct	tcaagcagct	cggcttcgac	tcggtcatgt	11220
ccgtcgagct	ccgggaaccgg	ctgagctccg	ccaccggatc	gtctctgccc	agcacagtgc	11280
tgttcaacca	ccccacgcgc	gaccggctcg	cccgcctatc	gtccgcggag	gcgtccagcc	11340
aggtggaagg	cgcgcaacgac	gcggcgcgga	cgggtgccgc	cgacgagccg	atcgcgatcg	11400
tgggtatggg	atgcaggtac	cccggaggag	tcgcgtcgcc	ggaggacttg	tggcggctgg	11460
tgacatccgg	gggcgatgcg	atctccggct	tccccacgga	cgtggctgg	gacctcgagg	11520
tcattgtacga	cccggaccat	cggcggcccg	gcaccagcag	taccgcgag	ggcgggttcc	11580
tgtagcaggc	cggtagcttc	gacgcgggtt	tcctcggcat	cagcccgcgc	gaggcgtcgg	11640
ccatggaccc	gcagcagcgc	ctgctgctcg	agacttcctg	ggaggccgtg	gaacggcgcg	11700
gcacgaccc	gctgtcgtcg	cacggtagcg	gggcccgggt	tttcgtcggg	gcatggccc	11760
aggagtacgg	cccgcgtctg	gacgaggcg	cggacggcta	tgagggttcc	ctgctgaccg	11820
gtggcctgac	gagcgtgttg	tcgggcggcg	tggcctacag	cctgggggtg	gagggacccg	11880
cggtcaccgt	ggacaccgcg	tgctcgtcgt	cgttggtcgc	cgtgcacatg	gcccgccagg	11940
ctctccgtca	ggggcagtg	tccttgccgc	tggcaggcgg	ggtcaccgtc	atgtccggcc	12000
cgggatatt	cctggagttc	agcaggcaga	goggactggc	accggacggc	cgtgcaagg	12060
cgttcgcggc	cggagctgac	ggcacgggct	gggcccgaagg	cgtcggcgtg	ctggtgctgg	12120
agcgcctctc	cgacgcccg	cgcaaccggac	atccggtgct	ggcgggtgta	cgggggtcgg	12180
cgatcaacca	ggacgggccc	tcgaacggcc	tgaccggacc	gaacgggctc	gcgcaggagc	12240
gggtgatccg	tgaggccctg	acggacgcag	ggctgtctcc	cgcgcagctc	gacctggtcg	12300
aggccacagg	caccggcacc	accttgggtg	accgatcgga	ggcgcaggcc	ctgatcgga	12360
cctacggaca	gggcccgtcg	gcggaccggc	cgttcgcact	gggctcgtcg	aagtccaaca	12420
tcggccacgc	ccaggcggca	gcccggagtgg	gcggagtcat	caagacggtg	atggcgggtg	12480
ggcacgcac	catgccccag	accctgcatg	tcgacgcgcc	gtcaccgcat	gtggactggt	12540
cgtccggcca	ggtccggctg	ctgaccgagg	cagtgcctg	gcccaggtcc	gaccaccccc	12600
ggaggggcggc	ggtctcgtcc	ttcgggatca	goggcaccaa	cgtcacgtt	gtcgttgagc	12660
agcccccgcc	ggaggtgtcc	gcggtcaccc	ggccatcacc	tatggcgccg	gacgaggccg	12720
taccggccccc	ggggcagccg	gtgccttggc	tgctgtcggg	caagtacccg	gaagcgggtg	12780
gcgagcaagg	ggcgcggctg	cggctgtacc	tggccgaccg	gcccggcgcc	ggtctcggcc	12840
acatcggtcg	gtccctggcg	tcgaccggt	cggcttcga	gcaccgtacg	gtggtggtcg	12900
cggcggacca	tgggcagttc	cgtgaggcgc	tgggcgcggc	cgcggcggtg	tcggcggatg	12960
cccggtctgt	cagggcgctg	gcccacatcg	acggcaagac	cgtcttcgtc	ttccccggcc	13020
agggcgcgca	gtgggcggcg	atggccgggg	aactccttga	ctcctccgag	gtgttcggcg	13080
cccggtatgg	cgaactgcgc	cgggctttgg	ccccgttctg	cggctggtcg	ttgcaggatg	13140
tcgtccggca	ggccgagggc	gccccgcgc	tggaccgggt	cgacgtcgtc	cagccgggtg	13200
tgtagggcgt	catggtgtcg	ctggccgacc	tatggcgtgc	tcattggcgtt	gagccctcgg	13260
ccgtggtggg	ccactcgcag	ggtgagatcg	cggccgcctg	cgtcgcgggt	gggtgacgc	13320
tggaagacgc	cgcgcgggtg	gtgtcgtcgc	ggagccgggc	catcgccgaa	gtactcggcg	13380
gacacggcgg	catgctgtcg	gtgaccgcgg	cccgggaaca	ggtcgaggag	tggctgtctc	13440
cctgggaggg	caggatttcg	ctcgcaacca	tcaacggaac	cgaatccgtc	gtggtcgccg	13500
gcgatccgga	cgcgctggcg	gaattccgog	cgtggttggg	gaaccgacag	atccgtagcc	13560
gcacccctgcc	ggtcgattac	gcctctcact	cggcgagggt	cgaggctgtc	caccagcgac	13620
tgctggacga	cctggcgccg	atccgcccc	gtacgtgcgg	taccccgctg	ctgtcctcgg	13680
tcaccggcca	gtggctggac	accgcctcga	tggacgccga	gtactggtac	cagaacctgc	13740
gcccggaccgt	ggagttcgcc	gcggcgaccc	gcaccttggc	cgaacggggg	caccgcctct	13800
tcacgcagggt	gagctcgcac	ccggtgctgg	tcggcgcgat	acgggaaacc	ctcgaagccg	13860
tcgaggtcca	ggccgctgtc	gcccgggtcac	tccggcgtga	cgacggaggc	ctgcggcgtt	13920
tccggctctc	gcttgccgcg	ctcgtcaacc	gggggctggc	ccccgactgg	tccatgctct	13980
gccccggggt	gagccgaacc	gacctcccca	cctacccttt	ccagcgcagc	cgttactgga	14040

tcaccgcctt	ctcgggggtcg	cggagcgccg	gtgaactcaa	cgctgcggac	tcacgcttct	14100
gggaggcggg	cgacagcgag	gaccccgggc	ggctggccga	ggtgctcagc	ctcgacgacg	14160
acgcgtcgct	cgaaccgggtc	ttcctggcac	tgtcctcggtg	goggcgacgg	caccgggtgc	14220
ggtccaccct	ggacgactgg	cgttatcggtg	tgacctggca	gocgctgccc	ggggcgcccg	14280
tcccggtgac	ggcggcaacc	ctcggaggga	cctggctggt	ggcgtgccc	cacgaggacg	14340
cctacgtctc	ccagggtgctg	cgcgggctgg	gcgaccgcgg	cgcgaccgtg	atcacccctac	14400
gagccgacga	cccgcgccac	ggcccgctcg	ccgagcgggt	ccgggaggcg	ctggccggag	14460
cggcgagat	caccggcggtg	ctgtcgctgc	tggcggttga	cgagcggccg	cacccggaac	14520
atccggctct	tcccatgggc	ctggcgctca	acacggcgct	ggtgcgggca	ctggtggaca	14580
aggacgtccg	ggctccgttg	tggtcgcca	cgcggggcg	ggtgtcggtg	ggccgatccg	14640
accggctggg	cagccctgcc	caggcgatgg	tgtgggggct	cggcctgggtg	goggccctgg	14700
aacacccgcg	gcactggggc	gggctggtgg	atctgccga	aaccgtggac	gagcgggtgc	14760
tgaaccggct	ggtgacgtg	atctcgggcc	aacgagtcca	cggacaggga	gccccggggc	14820
aggacggcga	aaacccgggg	gatgaggacc	agcttgcggt	gcgggcgtcc	ggagtgttcg	14880
cgcggcggct	gtcgacgcg	cccgtgtcgg	gcagccgcaa	ccgggagtgg	acgccccggg	14940
gcaccgtgct	ggtcacggga	ggcacccgtg	gocggggcac	ccaggtggct	cgtgtgggtg	15000
cccgtaacgg	gcgcgaacac	ctgtctgtga	ccagccgtcg	tggcaggggac	gocgaggggg	15060
ccgcgagact	ggcggccgaa	ctcacggaag	cggcgctcag	ggtcacggtc	gocgcctgcg	15120
acgtagcgga	ccgggacgcc	ctggcccggc	tgtctgcggg	cgtaccggac	gagctgccgc	15180
tgaccgccgt	gattcatgcc	gocgggtgtg	tcaccaccgc	cccgtggac	agcacccgtc	15240
cggaggaaact	ggccgagggtg	ctggcgggca	aggtggcccg	cgcgcgccat	ctggacgctc	15300
tgctcggcga	cgggcagttg	gacgccttcg	tactgttctc	ctccaacgcc	ggcgtgtggg	15360
gcagcggcg	gcaggcggcc	tacgcgcggg	ccaacgccta	cctggacgcc	ctggcccagc	15420
agcggctcct	tatgggocag	accgcgacct	cagtggcctg	gggtgcctgg	ggcggggccg	15480
ggatggcgcg	caggaaggga	ttcaaggagc	ggctgcgcgg	gcggggcatc	atcgaaatgg	15540
acccgagact	ggccgtcacg	gcgctcgtgc	aggccgtcga	gtccggagag	gcgtcgatag	15600
ccgttgccga	cgtcgattgg	gcacgcttcg	tgcgcggctt	cacctcgaa	cggcccgatc	15660
cgtgatcg	cgacctgcct	gaggtgcggg	acgcgctcgc	ggaggccgac	agccggcccg	15720
ccgtcgatca	gggcgggtcg	gcgctcgcca	cgcggctggc	cgggctgtcc	gtgctcgaa	15780
gggagcgggt	cctgtctaac	ctggtgcgca	ccgaggtggc	ctcggtactc	ggtcacacca	15840
cggccgacat	ggtcgatgcc	cgtcgccctc	tccgtgaact	cgggttcgac	tcgctgatcg	15900
cgggtggagt	ccgcggccgg	ttgaacgcgc	cgaccgggct	gocggctgcct	acctcggtcg	15960
ccttcgacca	ccccaccccg	gocgagctcg	ccggccatct	cggggagtgtg	ttcgccggat	16020
ccgcgggtga	caccgccatg	cccgtgtogg	tgaccaccgc	cggggacgac	gaaccgatcg	16080
ccatccgtagc	gatgtcctgc	cggtagccgg	gcggtgtgcg	cactccggag	gacctgtggc	16140
ggctggtggc	cgagggccgg	gacgcgatca	cggacttccc	caccgaccgc	ggctgggata	16200
tcgaaagcct	gtatgacccc	gaccggggcc	ggtccggcac	ctcctacacc	cggcggggcg	16260
gcttcctcga	cgacgcggcg	gccttcgata	cggcgttctt	ccggatctcc	ccccgcgagg	16320
ccctggccat	ggacccgcag	cagcggctgc	tcctcgaaat	gacgtgggag	acctcgaa	16380
gggcgctcat	cgacccaaca	acgtgaagg	gcagccaggc	cggggtgttc	atcggcaccg	16440
cacaccccg	ctacggcgag	ggcatccacc	acgagtccga	gggcgtcgag	ggccagcagc	16500
tgttcggcgg	ctcggccgcc	gtggccgcag	gocggatcgc	ctacacgttc	ggcctggaag	16560
ggccggcgat	gacggtggac	accatgtgct	cgtcctcgct	ggtggcactc	catctggcct	16620
cccagctccct	gcgcaccggc	gagtcctcga	tggcgctcgc	cggcggggtc	acggtaatgg	16680
cacggcgac	cgttttcacc	gagttcagcc	ggcatcgggg	actgtcccc	gacggacggt	16740
gcaagtcctt	ctccgacgcc	gocgacggca	ccggctgggg	cagggcgccc	ggtgtgtccc	16800
ttctcgaacg	gctctccgac	gcccgtcgaa	acggccaacc	cgtgctggcc	gtcatctcgc	16860
gcagcgccat	caaccaggac	ggcgccagca	acggccttac	cgcacccaac	ggccctccc	16920
agcaacgcgt	catccagcag	gccctggcga	acgcgtccct	gtcgccggcc	gacgtcgccg	16980
ccgtcgaggc	ccacggcacc	ggtaccaccc	tggcgacccc	gatcgaggcc	caggccctga	17040
tcgcccgccta	cggacaggac	cgcgcgacgg	accggccgct	acggctgggc	tcgctgaagt	17100
ccaacatcgg	ccacgcgcag	tccgcagccg	cagtcggcgg	cgtgatcaag	atggtccagg	17160
ccatccggca	cggcctcctc	ccgcgcacgc	tgcacgcgga	gcagccctcc	cgccacgtgg	17220
actggtccgc	cggctcgggtg	gaactgctca	ccgaggcgat	gocgtggccg	gacaacgacc	17280
aaccccgggc	ggcgggtgtc	tcggcggttcg	gcggcagcgg	caccaacgcc	cacatgatca	17340
tcgagcaggc	gcccgcgccg	gacgagccgg	agcacaccga	cggcacgagc	aggaccagcg	17400
gcgagagcgg	gcgcgaacag	gccaggccgc	tgcgatggtg	gocctgggtg	ctgtccgcgc	17460
ggagtgcac	cgcgctcggg	gcacaggccc	ggcgctacgc	cgcctacgcg	gocgcgcccg	17520
aggcggggcag	catctgcgac	atcggttggg	cgtggcgac	cacccgagcc	acgctggacg	17580
accggggccgt	ggtcgtggcc	gcggaacggg	aaggattcct	caccgctctc	gacgcgctgg	17640
ccgaggaccg	gaccgcccc	ggtctggttc	ggggggcggc	tggaacagga	gtcggttcgg	17700
cattcctgtt	ctccggccag	ggctcacaaa	gactcgccat	ggggcgcgag	ctgtacgaca	17760
cgtccctcgt	gttcgcccag	gcgctggacg	aggtgtgcgc	ccagctcgac	gggcacctgg	17820
accggccctt	cctgcgggtg	ctgttcgcgg	cggagggttc	cgacgacgcg	tcgatgctgg	17880

accagaccgc	cttcacccag	gccgcgttgt	tcgcggctga	ggtggcgctg	ttccgcctcg	17940
tctggtcctg	gggcctgcgg	cccgatttcc	tcacggggca	ttccgtgggc	gaagtgcggg	18000
ccgcccattg	ctcgggcgtg	ctgtccctcg	ccgacgcggc	gacactgggt	gtcggccggc	18060
gtcggctgat	gcaggcggtg	ccctccggcg	gcgcgatggt	ggccttgcaa	gcgggtgagg	18120
aggaagtacg	gctgtccctg	gcgggactgg	aggacgttgt	cggcgctgcc	gcccacaacg	18180
gccccgcctc	gacctgatc	tcctggcgacg	aggaggccgt	cctcccggtg	gccgcgcact	18240
ggcgcgcgca	gggcccgaag	acgcgtcgcc	tcaaggtgag	ccacgccttc	cactcacccc	18300
gtatggaacc	catgttgca	cggttccacg	ccgtgtctaa	aacgctttcc	ttcgccgagc	18360
cggccattcc	cgtggtctcg	aatgtgaccg	gccgtcccgc	cgagcggacc	gaactgtgcg	18420
cggcggacta	ctgggtgcgc	catgtccggc	atacgggtcg	cttccatgac	ggcatccggc	18480
cgtcggaggc	cgaaggcgct	agcgcattcc	tggagttggg	gcccgcaggc	aaactctcgg	18540
cgatggtccg	cgactgcctg	gacaccagcc	gcccgggtgt	cacggcaccg	gttttgcgac	18600
gtgaccgtac	cgatgtgtct	gccgcgttga	cggcactggc	cgaagcgcac	gggcacgggg	18660
tgccgggtga	ctgggcgtcg	ctcttcggcg	gctcagaccg	cggggcggtc	gagctgccga	18720
cgtacccgtt	ccagcgggaa	cacttctggc	tggattccgt	cacgggcagc	agtacatga	18780
gcacggccgg	actggcgctc	cccgatcatc	cgtgtttggg	agccgtgacg	acggtggccg	18840
gcgaggacgg	cctcctcttc	accggcaacc	tgtcggtagc	gacgcaccca	tggctggccg	18900
accacaggat	caccggttcg	gtcctgctgc	cgggcacggc	gttccctgga	ctggccgtcc	18960
aggccggggg	ccaggccggc	tgcgggcggg	tcgaggacct	gacgctgctg	gtccgctcgc	19020
tactgcccga	agagggcagc	gtcagggtcc	agatgaaggt	gggggagccc	gacgccacgg	19080
gccgcgcgac	catcgagggt	tactcctcgg	accagcaggc	ccccggccgg	gaacgctggg	19140
tcctcaacgc	gagcgggatg	cttgccggcg	aaccgggtga	ggccccgcgg	agtctacca	19200
cctggccccc	cgaaggcgct	gtccccgttc	cgttgacagg	cttccacgac	cggctggcgg	19260
cacgcggcta	ggcctacggc	ccgacattcc	gcgggctgag	cgcgcgctgg	tcacgcggtg	19320
acgagatctt	cgcggaagcg	gcgctccctc	cgggcatcgc	gcaggatgcc	gcccgcctatg	19380
gactccaccc	cgccctactc	gacgctgccc	tgcacgccat	ggaactccgg	gaaccccgcc	19440
cggccggcga	cggagtccgg	cttccgttcg	cctggaacgg	cttctccctg	cacgcgtcgg	19500
gtccggaagc	ggtacggctg	cgcctcgcgc	cgcggggcgc	cgcgcctctg	tcggtgaccc	19560
tcgcgatgac	catcggctgc	ccggttgctc	cagcccgcct	gctggccctg	cgggagctct	19620
cgtccgacct	gctgcgcccg	gcgtccgtct	cgtacgggga	ctcgtctgtc	cgcaccgctt	19680
ggatacccg	cctcgtcggc	ccggaggcgg	agtccggggc	ggtgcgaccg	tcggccggct	19740
gggcgggtgt	gggccccgat	ccgctcggcg	cggccaacgc	cctgaacctc	acgggaacct	19800
cctgctcctg	ctatccggac	ctggcggcgc	tgatcgcggc	cgtcgcggc	ggagccgcgg	19860
tgcccgaggc	cgtactcgcg	ccgtacgcgg	cggagccagc	cccgacgcgc	ggatctcccg	19920
cggacgcgct	acgggcctcg	accggccggg	cgttgcaact	gctgcaatcc	tggtgtccg	19980
aggaccggtt	ggagcgaagc	cggctgatcg	tgtcaccccg	gggggcgggtg	gccgtcggtg	20040
cggacgaagg	cgtcacgcac	ctgggtgagt	cgtcggctcg	gggtctgggtc	cgttcggcgc	20100
aggccgagca	ccctggcagg	ttctccctgg	tcgacatcga	cgaccgggag	gagtccctgg	20160
ccgtcctgag	cgcggcggcg	gtatccgatg	agccacaact	cgcctcgcgc	tgccggccaga	20220
tgaagggtgc	ccgctcggcg	tcgctcgacg	ttcccacgac	cggtatgcct	gagatgcccg	20280
acgtttgggg	tgttgacggt	accgtgttga	tcactggcgg	gaccggtgtg	ctgggtgggc	20340
tcgtcggccg	tcctctgggt	gccgggcatg	gggtccgtcg	tcgtgttgctc	tcgagcaggc	20400
ggggccctga	tgcgccgggt	gcggtggagc	tggtcggcca	gctcacccgt	ctgggtgcgg	20460
atgtcacctg	tgcgcctcgc	gacgcggccg	accgggatgc	gctggccgcg	ctcttgga	20520
ccgttcccg	cacgcaccc	ctgactgggt	tcgtgcatac	cgtggtgtc	atcgatgacg	20580
ccactgtcac	caccctcact	cccgagcgca	tcgacgggt	cctacgcccc	aaggctgacg	20640
ccgcgtcaa	cctccatcag	ctgacggcgc	atctcggctt	gaccgccttt	gtgctcttct	20700
cctccgccc	cgggctcttc	ggcggcgccg	gccagggtaa	ctacgcggcc	gccaacgcct	20760
tcctcgcagc	actggcccaa	caccggcggg	ccaacggcct	caatgcccag	tccttgccgt	20820
ggggactgtg	ggcgggaagc	agcgggatga	cgggacacct	ggacgcggcc	gacctgcgcc	20880
gggtggcccc	ttccggcctc	accgcgatgc	ccaccgggga	cgggctggcg	ctgctcgaca	20940
ccgctcagcg	ggtggacgaa	gccaccctgg	tcacggccgc	gctggacacc	cgggccctgc	21000
atgcccgggc	cgcagacggc	acgctgcggg	cgtgtttcca	cgcactcgtg	cccgtaaccg	21060
gocgatccgc	gacctccccg	gcggcccagg	ccgcggggcc	ggatggactc	cgcagcgggt	21120
tgtcgggggt	ggtcaggggg	gagcgtcgag	cggcgtgct	ggatttggtg	tgtggtcatg	21180
tcgcgagggt	gctggggcac	cgggaccgga	gcagcattga	ggagaccggg	cccttcaagg	21240
acaccggctt	cgactcattg	accgctgtgg	agctgcgcaa	tgtgctgcac	ggtgcgaccg	21300
ggttgccgct	gccggccacg	ctggtcttcg	actaccggac	gctgcagct	ctcaccgatc	21360
acctctacga	cgagcttctg	ggttcccgcg	aggacgcgct	gctcggcccc	atcaccaggg	21420
ccgcgtacga	cgagccgatc	gcgatcgtgg	ggatggcctg	ccgctatccg	ggcgggggtg	21480
agtccccgga	ggacctgtgg	cagctggctg	ccgacggccg	tgacgccatc	tcgacttcc	21540
ccgcccagcg	gggctggaac	gtcgagagcc	tctaccaccc	cgaccccgac	caccccgga	21600
ccagctacac	ccgtgcggga	ggcttctcgc	acgacgcggc	ggacttcgac	ccggagttct	21660
tcgggatctc	accgcgtgag	gcactggoca	ccgaccccca	gcagcgactg	ctgctggaga	21720

cgacgtggga	ggccttcgaa	cacgcccggg	tgggcccggc	gtcactgcgt	ggcagccgga	21780
ccggcgtcct	cgtcggcggt	atgtacaacg	actacgcctc	gcgtatccgg	cacatcccag	21840
agagcgtcga	gggcggtctg	accaccaaca	gcgcggggag	tgtggcgtcg	ggccgggtct	21900
cgtacacgtt	cggctctggag	ggaccggccg	tcacggtgga	taccgcgtgt	tcgtcgtcgc	21960
tggtagcggt	gcacgtggcc	gcgcaggcgt	tgcgcaacgg	tgagtgcact	ctggctctgg	22020
gggcggtgtg	tgcggtgatg	tccactcctg	ccacgtttgt	cgagttcagc	cggcagccgg	22080
ggctggcagc	tgatgggcgg	tgcaaagcct	tcgcggacgc	tgccgacggc	accggctggg	22140
gcgaaggcgt	cgggtgtgct	ctggtggagc	gtttgtcggg	cgcgcgccgc	aacgggcatc	22200
cgggtgctgg	ggtcgtttcg	ggcagtgctg	tcaaccagga	cggggccagc	aatggtctga	22260
gggcccacaa	tggtccttcg	cagcaacggg	tgatccaaca	ggcgtcggcc	aatgcggggg	22320
tggcgggggg	ggatgtcgat	gccgtggagg	cgcacggcac	gggaaccogg	ctgggogacc	22380
cgatcgaggc	cgaagcgttg	atcgccacct	acggacaggc	cgggtcggcg	gaccggccgt	22440
tgtggctggg	ttcgtggaag	tccaacatcg	gtcacaccga	ggccgccgcg	ggcgtcgccg	22500
gcgtcatcaa	aatggtgcag	gcgatgcagc	acgggactct	gccgccacc	ctgcacatcg	22560
accagcccac	gggcccaggc	gactgggcta	cgggtgcagt	ggagctgctg	accgaggccg	22620
tgccctggcc	ggacagtgc	cggccccggc	gggtggctgt	ctcctcgttc	ggtgtcagcg	22680
gtaccaacgc	ccacgtcatc	atcgaacaca	ccccacacac	ccccacacac	accgcacact	22740
cccaatcctc	ccaatcccc	caggccccgc	agactgtgca	ggcccatcgg	ccggtgccgt	22800
ggctgctgtc	ggcgaagacc	tcgcaggccc	tggccgcgca	ggccccggcg	ctgtcagctc	22860
acttgcgagc	caaccocgat	ctgcgttcgg	ctgatgtggc	gcattccctg	ctcaccacgc	22920
ggtctgtcca	cgcgagcgcc	gccgtcttca	tcgccgggtg	ccgggatgag	gctcttgccg	22980
ccctggacgc	actggccgac	ggcaccctcg	cccctcacct	cgttcagggg	cttgccgatg	23040
tgagtggaac	gacggtgttc	gtcttccccg	gtcagggttc	gcagtgggtg	ggtatggccg	23100
ttgagctgct	ggacggctcg	gaggttttcg	ccgagcatat	ggccgcctgc	ggcaggggccc	23160
tggaaaccgt	tgtggactgg	tccttgagg	acgtcctacg	ccagacggag	ggtacgtggc	23220
cactggaacg	cgtcgaagtg	gtccagcccg	tgtgttggcg	ggtcatggtc	tcgctcggcg	23280
gactgtggca	ggcacatggc	gttgagcctg	ctgcggtgct	gggccaactcc	caaggtgaga	23340
tcgctgcggc	ttgcgtggcg	ggagccctga	gtctggaaga	cggagcccgc	gttgtcgcgc	23400
ttcgcagcca	agccatcgcc	gaaaccctcg	caggacacgg	cggaatgctc	tcaatcgccg	23460
ccccgccac	cgacatcgca	cccctgatcg	cccgttgga	cgagcggatc	tccatcgcca	23520
cggtaaacgg	accgcattcg	gtggtggctg	caggagaccc	tgacgcgctc	gaggcactcc	23580
gcggcgaact	ggagaccogt	ggtctccgca	atcgtcgcac	cccggctgac	tacgocctac	23640
acacccctca	cgtcgaggcg	atccgtgaac	ggctcctggc	cgacctggca	gtgatccagc	23700
cacgtgccgc	gagcattccc	gtgctgtcca	ccgtcacogg	cgcattggctc	gacaccaccg	23760
tgatggacgc	cgagtactgg	taccgcaacc	tacgtcagac	cgtggagttc	gaagcagcca	23820
cccgcactct	cctcgaccag	gaccaccgct	acttcgtcga	gacagcccg	caccccgctac	23880
tcaccatcgg	tctacagcag	accatcgagg	aaaccaccgc	tccggcccgg	accctctcca	23940
ccctccgacg	caacgaaggc	accctccggc	acctgttcac	ttccctcgcc	caggcccacg	24000
cccacggcct	gaccatcgac	tggacccccg	ccttcaccca	caccgagccc	cgcaccaccc	24060
ccctgcccac	ctaccccttc	caacacgaac	gctactggct	ggaggacgga	gctccgaagt	24120
cgggggacgt	ggcttcggcc	ggactcggct	cggcggacca	tcgctgctg	ggcgcgcgtg	24180
tgccgctgcc	cgttcgggg	ggcttcctgt	tcaccggcca	gttgtcgtg	cggagtacc	24240
cctggttcgc	cgaccacgcg	gtacacggca	ccgtgctgct	gccgggcacc	gcttctgtg	24300
aatggcgct	ccaggccggt	ggcgtctcgc	gctgcgggct	gctggaggaa	ctcaccctgg	24360
aggcacgcgt	ggtgctgcgc	gaaaaacagc	ccgtccagct	ccaactcgtg	gtgaacgcc	24420
cggacgcccc	ggacgactcg	ggcggcagga	ccttcagcgt	gtactcgcgc	ccgcaggacc	24480
gtactgcgga	cgcgccctgg	gtgcggcacg	ccaccggagt	ggtccggctc	ggaggcgcgc	24540
cggagccgga	gggactgacc	gtgtggccgc	cgaccggagc	ggtcgcgggtg	ccggtcgagg	24600
acttctacca	ggtgctcggt	gaccgtggct	atgactacgg	acctgcgttc	cgtggggtaa	24660
gggcccgcgt	gcgccacggt	gacgtgggtg	atgccagggc	cgcactggcc	gaggagcagc	24720
agtcggacgc	cgcgctgttc	cacctccacc	cggccctgct	cgactcggcg	ctgcacggga	24780
tgggactgat	gccctcggcg	agcgcggagc	agaccggcgt	gccgttcgcg	tggcgcgggtg	24840
tgacgctgca	tgcggtgggg	gcgtcggccc	ttcgggtgag	tcttagggcc	gccggggccc	24900
acacggttga	ggtcctactg	gccgatggcg	caggctggcc	ggtcgtctcg	gccgacgcac	24960
tgggtgtccg	gccgctccga	caggaggaac	tggcggtctg	gcaggacgcg	taccgcgact	25020
ggctgtaccg	ggtcgactgg	cccagattgc	cggaggtccc	ccttgttggt	ccggccgggc	25080
catgggccgt	cctggggcgg	aacgcggcgc	ggatactcgg	caccgatggc	tggccgggtg	25140
tgttgcccg	ggtcccgatc	gacgcctatc	gggacctggc	ggagctgcgc	gaccggacgg	25200
gcccagacag	cgcgttcccg	gccgtggtgg	tcgcgcgggt	cgccacggga	accggtgccc	25260
cgcgggacgc	ggtgcgggag	gtgacgtacc	aggtgctgga	catgatccag	tcattggctcg	25320
ccgacgatcg	ttcgcctcgc	tcgacccttc	tcctggtgac	cgcgggcgcg	gtgtccaccg	25380
gcttcgggga	cgacctgggt	gatctggggc	aggcggcggg	atgggggttg	gtgagggccc	25440
cgcagtcgga	gaacccggac	cgtctcgtcc	tcctcgacct	cgacgggagc	gagccggtcg	25500
ggcctctccc	gacggcggcg	ctgctctccc	gggagccgca	actggcgttc	cgggagggca	25560

aggtgctgac	cgcccggtg	gaccgggtgt	cgctccgacgc	gggaacgctg	ctgccgcccg	25620
ccgggcccga	cccgtggcga	ctogacgtca	ccagccgggg	cacgctcgac	aacctcgccg	25680
tcctcgcggc	gcccaggtg	tcggcgccgc	tcgccgaggg	acaggtccgg	gtccggtg	25740
acgcggccgg	cctgaacttc	cgcgatgtgc	tggtcgctct	gggatgtac	ccgggtgagg	25800
gttcgatggg	cagcgaaggc	gcccgcgtgg	tgctggaggt	cgggcccgcc	gttagccggc	25860
tggccccggg	cgaccgggtg	atgggcatgc	tcgccggcgg	cttcttcggg	ccggtccggc	25920
taaccgacca	gcgcattgtg	accaagcttc	cggacggctg	gtcgttcacc	gagggcgcat	25980
cggtaaccgat	cgtcttcctc	acgcgctact	acggactggg	cgacctgggc	ggcctcgccg	26040
ccggccagtc	gctgctggtg	catgcggcga	ccggtgggtg	gggaatggcg	gctacgcagc	26100
tggccccgga	cctcggcgct	gaggtgttcg	gcacggcgag	cccggcaag	tgggaggcgc	26160
tgcgggggat	gggattggac	gaggagcaca	tcgcctcgtc	gcgggacctg	gacttcgaga	26220
agaagtcttc	ggccgcgacc	ggtggccgcg	gtgtcgacgt	ggtgctgaac	tcgctggccc	26280
gggagttcgt	ggacgcgtcg	ctgcggctgc	tgccgcgcgg	cggtcgattc	gtggagatgg	26340
gcaagaccga	catccgtgac	gcccagggcg	tgccgcgcgg	gcatcccgcc	gtcgtctacc	26400
gggctctoga	cctgctggac	gcccgcgggg	cggacgtgat	ccaggagatg	ctggccgagt	26460
tgctcgcgct	cttcgaggcg	ggggtgatcg	agccgctgcc	gctgacgacc	tgggacatcc	26520
ggcgtgcccc	ggaggcgctg	cggcacctga	gccaggcacg	gcacatcgcc	aagatggtct	26580
tcacctgccc	gcccgcggcg	gaccgcggac	gtacgttcct	gatcacgggt	gtgcccgagg	26640
cgttgggcaa	ccttggtcgcc	cgccatctgg	tgaccgaggg	tggcatacgg	aacctgctgc	26700
tcgtcagccg	ccggggggccg	gcccgcggcg	gcgcggaggg	gctggccacc	gagctggccg	26760
ggctgggggg	gacggtgacc	ctggcgccct	gtgacgtggc	cgaccggcag	gcccgtggcg	26820
ggctgctcgc	cgacatcccg	gcccagcatc	cgtgacggg	tgtggtgcac	gcccgcgggtg	26880
cttgggacga	cgggatcgtg	gcatccctga	cccgcgaacg	gctggacgcg	gtctaccgcc	26940
ccaaggttga	cgcgcgctgg	aacctgcacg	agctgaccaa	ggacagcgcc	ctggcccgct	27000
tcgtactgtt	ctcctcggcc	gcccgcgacg	tcggcagcgc	aggccagggc	aactatgcag	27060
cggccaacgc	cttctcgcac	gccttggccc	aattccgcca	ggcccagggc	ctggcgccca	27120
gctccctcgg	ctggggattc	tgggcccaga	gcggtgagat	gaccggtcac	ctgggggcct	27180
ccgacctggc	acggatggca	cgttcgggca	tcgccgccct	gacggtcgag	cagggcctgg	27240
ccctgttcga	ctccgcacgg	tcgggtgtct	gtgcgtcagt	gctgccggta	cggctggaac	27300
tcaccggggc	cgggtgcgcg	gcccgggtcgg	gaacggtgcc	ggcgtgatg	cgggggctgg	27360
tgcgggcacc	ggcccggcgg	gtggtggaaa	caaccacggg	cggtgccgtc	acaggccctgc	27420
gccaacggct	ggcgcgcgtg	tcgggcgcgg	accgogaccg	cgccctccaa	gagctggtgt	27480
gctcgcattg	ggccaccgtg	ctggggcaca	gcccgttcgg	atcggtgccc	gcgcagccgg	27540
cgttcaaggga	cgttcggctc	gattcgctga	cagccgtcga	gttgcgcaac	cggctcaaac	27600
tggcgaccgg	cctccggctc	cccgcgactc	tggtgttcga	ccaccgcacc	ccgctggcga	27660
tggcggaaca	gctccgggaag	gagctgttcg	cggacagagat	cccggtggcg	ccgaggtttt	27720
tggaggaact	ggaccgtctg	gaggcgccgt	tcgcccgtct	ctccgcgggc	gacctccagc	27780
agtcgggagc	cgcggcacgg	ctgagggcac	tgctgaggcg	gatcggcacc	gtcactccgg	27840
cgggagggga	cgttcggcac	ggcctcggcg	tagagctcga	aacagccacc	cacgacgaga	27900
tcttcgccc	tatcgacgag	gaggtagggg	acgtgtgacc	ggtcggtcgc	tcggccctca	27960
cccctacccc	cgcggcgac	tagcagcatg	gatgagatca	cgatgactga	tgagaccgct	28020
gtgccccaaa	cagagaccac	cgaggagaag	ctcttctcct	acctgaagaa	ggccacctcc	28080
gaactccagc	agagccggcg	ccgggtggca	gagctggagg	cggcgagggc	ggagcccatc	28140
gcgatcgtgg	gcacggcctg	ccggtaccgg	ggtggagtac	gttccccgga	ggacctgtgg	28200
cgttctgtcg	cgcaggggca	gcacgcgac	tcacgcttcc	cgacggaccg	cggctgggat	28260
ctogaagacc	tctacgaccc	ggaccgggac	cggcccgcca	agtccctacg	ccgggacggc	28320
ggcttctctg	acggtgcgcg	ccagttcgac	gcccgcgtct	tcgggatctc	gccacgtgag	28380
gcgctggcca	tggaccggca	gcagaggctg	ctgctcgaga	cgacgtggga	ggtcttcgag	28440
cgcgcgggga	tcgaccggac	atcgctccgt	ggcagccgga	ccggggtgtt	cgccggcatc	28500
agccaccagg	actacgctgc	cggacagcgc	ccgtcggccg	aggtctccga	ggggcacctg	28560
atgaccggca	ccgcggtcag	cgtggtgtcc	gggcccggct	cctatgcctt	cggcctggaa	28620
gggcccggca	tgacggtgga	cacggcctgc	tcctcgtcgc	tgggtggcgt	gcacctggcc	28680
gcgcaggcgt	tgcgcaatgg	tgagtgcacg	ctggcggtgg	ccggcgccgt	caccgtcatg	28740
gccacggcgg	gcgccttcac	caggttcagc	cgggagccgg	gcctggcccc	ggagggcgcc	28800
tgaaggcctt	tcagctcgga	cgcgcagggc	accggcttca	gcgagggtgt	gggtgtgctg	28860
ctggtgggag	gtttgtcgga	cgcgcgcggc	aacgggcac	cgggtgtggc	ggtcgtttcg	28920
ggcagtgctg	tcaaccagga	cggggccagc	aattggtctg	cggcgcccaa	tggctcctcg	28980
cagcaacggg	tgatccaaca	ggcgtggcc	aatgcggggt	tggcgggggc	ggtatgctgat	29040
gccgtggagg	cgcacggcac	gggaaccggg	ctgggtgacc	cgatcgaggc	gcaggcggtg	29100
atcgcgacgt	atggacaggc	ccggtcggcg	gaccggccgt	tgtggctggg	ttcgtgaag	29160
tccaacatcg	gccacaccca	ggccgcggcg	ggcgtcggcg	gcgtcatcaa	gatgatccag	29220
gccatgggtc	acgggacgct	gccccgtacg	ctgcatgtca	accagccctc	gccccaggtc	29280
gactgggccc	caggcgcggt	ggagctactg	accgaagcca	tgccctggcc	cagggtgac	29340
cggccccgcc	gggcccgaat	ctcctccttc	ggaatcagcg	gtaccaacgc	ccacgtcatc	29400

atcgaacagg	gggccccgcc	acggacagcg	tccgaccccg	gtgaaagtcg	tgctgacgag	29460
cccggcgta	ggggcgggcg	tcccgtccct	gccaccacgg	agtcggccac	cgaaccgcag	29520
ccggttccct	ggctgtgtgc	cgggcacagc	gcgaccgcgc	tgccggcgca	ggcggatcgc	29580
ttgaagtgt	acgcggccaa	caacaccggc	atccgtccgg	ccgacatcgg	cttctcgtg	29640
gtcaccaccc	gggcccgcgt	ggaacaccgc	gctgtcgtcg	tggcagccga	ccatgccggt	29700
ttcacgggtg	gtctcgacgc	gctggccgag	ggccggacag	ctcccggagt	ggtgagcgga	29760
acggtcgtcg	ccggtgcccg	gagcgcgttc	ctcttctccg	gtcagggctc	gcagcgggtc	29820
ggcatggggc	gcgagctcca	gcaggcggtc	ccggttttcg	ccgaggcttt	cgaagcagtc	29880
tgcgcccagg	tcgacccgta	cctgggagcac	ccacttctcg	atgtcgtact	cgccgcgcgc	29940
gacagcgact	tcggcgcggt	gctccatcag	accgcctaca	cgagccgggc	actgttcgcc	30000
ctcgaagtgg	ccctgttccg	gctggtcgaa	tcctgggggtg	tcaggccgga	ttacgttgcc	30060
gggcattcgg	tcggtgagat	cgcggcgggc	catgtggcgg	gggtgttctc	gctggaggat	30120
gggctcgtc	tgggtggtgc	gcgcggaacg	ttgatgcagg	cgttgccggc	tgaaggcgcg	30180
atggtggcgc	tccaggtgtc	cgaggacgag	gtcctgcggt	ccctgactcc	ttggtggag	30240
caggaccggg	tggatgtcgc	ggcgggtcaac	ggcgccagcat	ccacagtggg	gtcgggcat	30300
gaggaggcgg	tcctggcggt	tgcggagcac	tggcaggcgc	ggggccgcaa	ggttcgtcgg	30360
ctcactgtca	gccatgcctt	ccactcacct	cgtatggacc	cgatgctcga	ccagttccgt	30420
gtggtcgtgg	agggtatccg	tttcgcggag	ccggccatcc	cggctgtctc	cagcgtcacc	30480
ggtcgtcttg	ccgagcccgg	gcagttgacc	actgcggact	actgggtgcg	ccacgtccgt	30540
caaacgggtc	gcttccacga	cgccctccag	accctccaga	ccgagaatgt	gaccgcgttt	30600
ctggagatcg	gtcccagcgg	gcaactctcg	gcaatgaccc	gcgacttcct	gaccgatacc	30660
ggggcccacg	ccgccgtcgc	acccctcctg	cggcgcgaa	gtcccagggc	acccagcgcg	30720
ctcaccgcaa	tcgccgggct	gcacaccac	ggcgtctcga	tcgactggcg	cacgtacttc	30780
accagacaca	gcaccagcac	cagcaccagc	accggtaccg	gtaccgggtac	ggggcaggcc	30840
actgccgaca	cgcccgtcca	gctgcccacg	tacgccttcc	agcaccagtc	cttctggctc	30900
ggccccacgg	cccctgtcgg	cgacgtcagc	accgcggggc	tcacctcgcc	cgaccacccc	30960
ctgctcagcg	cagccaccac	caccgctgtc	gacggcagcc	tcctgctcac	cggcaggctg	31020
tcgcagcggg	cgcccgcggt	gatcggcgac	caccgcacgc	gcggtgtggt	cctgctgcca	31080
ggcaccgctc	tcgtggaact	cgtcgtacgc	gcccggggacc	aggccgggtg	cagccgcac	31140
gacgaactca	tcatgctcac	gcccgtgacg	ctgcccagac	atggtgccgt	gcggatccag	31200
gtcgcggtcg	gcccggccgg	ccacgacggc	cgccggccgg	tcacatcca	ctccagcacc	31260
tcggacacga	ccggcgacga	acagtggacc	ctcaacgcca	gcggtctgct	caccgtcgag	31320
atgaccgata	cgcccggcga	tctcaccccc	tggccggccg	agcacgccac	ccgcataacc	31380
ctcgacggcc	tctacgagcg	gctcggcgaa	agcggctacg	gatacggccc	ggtcttccag	31440
ggcctgcgcg	ctgcccgtgg	actcggcgac	gacacctacg	ccgaggtcga	gatccccgcc	31500
ggcgaccaga	ccgacaccga	ccgctacgaa	ctccaccccg	cgctcctcga	cgccgcgctg	31560
cacgcgtcct	ccctccaggg	cgacgaggcc	ggggccgggg	agctgctgcc	gttcgctggg	31620
accgggggtg	cgctgtacgc	ggccggcgcc	tcggccctgc	tcgtcaaggt	gtcccgtaac	31680
ggtccggaca	ccatggcgct	gctcgtggcc	gacaccgagg	gccaccgggt	cgccaccgtc	31740
gactcactga	ctgtccggcc	gatggccatc	gaccagaccg	cccgagcac	cagccaccct	31800
gacgcgctgt	tcaccgtggg	gctggagtgg	gcccgaagcc	gggagggcaa	ccggaccatc	31860
cccctgtcgg	actgcgccat	gctggctccg	gacgaaccgg	acctcacctc	cgcccgggcc	31920
tggcccggtg	cctccgcgca	gcggtacgcc	ggcctcgcgg	cgctcgtcga	gatctgogga	31980
acggagcggc	cgttacctgc	cgttggtactg	gcgcccttcc	tcccggcgga	tgccgcgccc	32040
gcgacacccg	cgcccgcgac	gcacgcgacg	acgcgcgcgc	ccgcgcctct	catcaagggc	32100
tggctggggc	acgacgggtt	caccgactcg	cgtctggtct	tcgtcaccgg	tggcgcggtg	32160
gccaccagcg	gcccgggacga	actgcacgac	ctggaacact	ccacggctcg	gggtctggtc	32220
cggtcggccc	agaccgagaa	ccccggcagg	ttcgcgctgc	tcgatctcga	cgaccgggac	32280
accgtcaccg	aactgccgga	agccatcctg	gcccgtcagg	cacagctggt	cctgcggggc	32340
ggcgcggtgg	gaaacctccg	gctggccaag	ggcgctgcga	tacaggatcc	cgaccggggt	32400
tggggtgttg	acggtaccgt	gttgatcact	ggcgggaccg	gtgtgctggg	tgggctcgtc	32460
gcccgtcatc	tggtcgcccg	gcatggggtc	cgtcgtctgt	tgctctgcag	caggcggggc	32520
cctgatgcgc	cgggtgcggg	ggagctggtc	gcccagctca	ccgctctggg	tgccgatgtc	32580
accgttgccg	cctgcgacgc	ggctgaccgg	gatgcgctgg	ccgcgctctt	ggacaccggt	32640
cccgccacgc	accctctgac	tgggtgctgt	cataccgctg	gtgtcatcga	tgacgccact	32700
gtcaaccacc	tcactcccga	gcgcacgcac	gcggtcttac	gccccaaagg	cgacgccgcg	32760
ctcaacctcc	atcagctgac	ggcgcatctc	ggctttgacc	gctttgtgct	cttctcctcc	32820
gcccgcgggc	tcttcggcgg	cgcggggccg	ggtaactacg	cggccgcca	cgccttcttc	32880
gacgcactgg	cgcagctgcg	gaagcggcag	ggactgcggc	gcgtgtcgtc	ggcctggggg	32940
gcctgggtcc	aggacggcgg	aatgaccgca	acgctggacg	cgggcgacgt	cgagcggatg	33000
gcgcgcgggc	gtgtgctgcc	gctcagccac	gagcagggcc	tgaacctgtt	cgacctggca	33060
gtggcagggg	ccgagccgct	ggtggcaccg	atgcggctgg	acaccaccgc	gctgcgcgag	33120
tccggtgcc	ccgtgocgga	gatgctgcgc	gggttggtgc	gtgagcggtc	acgcgcgcgg	33180
gtcggaccct	cgacacgac	gtccgcgcgc	atggcgctgg	aacaacggtt	gtcgggggtt	33240

gtcagggggg	agcgtcgagc	ggcgcgtgctg	gatttgggtgt	gtggtcatgt	cgcgaggggtg	33300
ctggggcacg	cggacccgag	cagcattgag	gagaccggc	ccttcaagga	caccggcttc	33360
gactcattga	ccgctgtgga	gctgcgcaat	gtgctgcacg	gtgcgaccgg	gttgcggctg	33420
ccggccacgc	tggtcttcga	ctacccgacg	cctgcagctc	tcaccgatca	cctctacgac	33480
gagcttctgg	gttcccgcga	ggacgccgtg	ctgcgcccg	tcaccagggc	cgcgtacgac	33540
gagccgatcg	ccatcgtagc	gatgtcctgc	cggtagcccg	cgggtgtctg	cactccggag	33600
gacctgtggc	ggctgtgtgg	caggggccc	gacacgatca	cggacttccc	ggacgaccgc	33660
ggctggggata	tcgacgccct	gtatgacccc	gacccggggc	accccgccac	ctcctacacc	33720
cggcggggcg	gcttctctgt	cgcgcggcg	ggtttcgatc	cggcgttctt	ccggatctcc	33780
ccccgcgagg	cgctggccat	ggacccgcag	cagcggctgc	tgtcgaat	gacgtgggag	33840
atgttcgaac	ggcgctcat	cgacccaaca	acgctgaagg	gcagccaggc	cggggtgttc	33900
atcggcaccc	ccggccccgg	ctacggcggc	cgcattccacc	acgagtcgca	ggcgctcgag	33960
ggccagcagc	tggtcggcgg	ctcggccgcc	gtgacctcag	gocggatctc	gtacacgttc	34020
ggccttgaag	ggccggcgat	gacggtggac	accatgtgtc	cgtcctcgct	ggtggccctg	34080
cacctggccg	tccagtccct	gcgcaacggc	gagttcctga	tggcgctcgc	cggcggggtc	34140
acggtgatgt	cccgcccggc	cgcgttcacc	gagttcagcc	ggcagcgggg	gctgtccccc	34200
gacgggcggg	gcaagtctgt	cgcgcgacgc	gcccagggca	ccggctgggg	cagggcgccc	34260
ggcgtgctcc	tcctcgagcg	gctctccgac	gcccgtcgca	acggccaccc	ggtgctggcc	34320
gtcatccgcg	gcagcgcggt	caaccaggac	ggcgccagca	acggcctcac	ggcaccacac	34380
ggccccctcg	agcaacgcgt	catccggccag	gcccgtggcg	acgcgtccct	gtcgcgggcc	34440
gacgtcgacg	ccgtcgaggc	ccacggcacc	gggaccccc	tggcgacccc	gatcgaggcg	34500
caggccctga	tcgccacct	cggacaggac	cgcgcggcg	acggcgcgct	gcggctgggc	34560
tcgggtgaagt	ccaacatcgc	ccacgcgcag	gcccagccg	cagtcggcgg	cgtcatcaag	34620
atggctcagg	cgtaccggca	cggcctcctc	ccgaagacc	tgcacgtgga	gcagccctcc	34680
cgccacgtcg	actggtccgc	cggctcgggt	gagctgtctc	ccgaggcgat	gocgtggccg	34740
gagaccgacc	aacccccggc	ggccggtgtc	tcggcggttc	gcccagcggg	caccaacgcc	34800
cacatgatca	tcgagcaggc	gcccgcggcg	gacgaggagc	acaccgacgg	cacgagcagg	34860
accagcggcg	agagcggcgc	cgaacaggcc	aggccgctgc	cgtatggtgc	ctggctgctg	34920
tcggcgaaga	cctcgcaggc	cctggccgcg	caggcccgcc	gcctgtcagc	tcacttgcca	34980
gccaaccccc	atctgcgttc	ggctgatgtg	gogcattccc	tgtcaccac	gcggtctgtc	35040
cacgcccagc	gcgcccgtct	catcgccggt	gaccgggatg	aggctcttgc	cgcctggac	35100
gcaactggcc	acggcacccc	tgcacctcac	ctcgttcagg	gccttgccga	tgtgagtgcc	35160
aagacgggtg	tcgtcttccc	cggtcagggt	tcgcagtggg	tgggtatggc	cgttgagctg	35220
ctggacggct	cggaggtttt	cgcgcgacat	atggccgcct	gcgccagggc	ccttgaaccc	35280
ttgttggaact	gggtccctgga	ggacgtccta	cgcacagcgg	acggtacgtg	gccactggaa	35340
cgcgtogaag	tggtccagcc	cgtgctgtgg	gcggctcagg	tcctcgtcgc	gggactgtgg	35400
caggcacatg	gcgttgagcc	tgtctcgggt	ctggggccact	cccaagggtga	gatcgtcg	35460
gcttgogtgg	cgggagccct	gagtcctggaa	gacggagccc	gcgttgtcgc	gcttcgcagc	35520
caagccatcg	ccgaaaccct	cgcaggacac	ggcggaatgc	tcctaatcgc	cgcccccgcc	35580
accgacatcg	caccctgat	cgcgcgctgg	aacgagcggg	tcctcatcgc	cacggtcaac	35640
ggaccgcatt	cgggtgtgtg	cgcaggagac	cctgacgcgc	tcgaggcact	ccgcggcgaa	35700
ctggagaccc	gtggtctccg	caatcgtcgc	atcccggctg	actacgcctc	acacaccctt	35760
cacgtcgagg	cgtccgtgga	acggctcctg	gcccacctgg	cagtgatcca	gccacgtgcc	35820
gcgagcattc	ccgtgctgtc	caccgtcacc	ggcgcatggc	tcgacaccac	cgtgatggac	35880
gcccagttact	ggtaccgcaa	cctacgtcag	accgtggagt	tcgaagcagc	caccgcact	35940
ctcctcgacc	aggaccaccg	ctacttcgtc	gagatcagcc	cgcaccccgt	actctcggcg	36000
atggctccgcg	actgcctgga	caccagccgc	ccggtgggtc	cggcacccac	cctccgacgt	36060
gaccgtaccg	atgccactgc	cgcgttgacg	gcactggccc	aagcgcacgg	gcacgggggtg	36120
ccggtogact	gggctcgtct	cttcgccggc	tcgaccgccc	gggctgtcca	cctcccgacg	36180
taccccttcc	agcggcaaca	ctactggctg	gattccggta	cgggcagcag	tgacatgagc	36240
acggccgggac	tggcttcccc	cgtatcatccg	ctgttgggag	cgtgacgac	ggtggccggc	36300
gaggacggcc	acctcttcac	cggccggctg	tcggtacgga	cgcacccatg	gctggccgac	36360
caccagatca	ccggttcggg	cctgttgccg	ggcacggcct	tcgtcgaact	ggcgtccgg	36420
gcccggggacc	aggccggctg	cggcggggtc	gaggagctga	cgtgctggc	tcgctcgtga	36480
ctgcccgaag	aggcgagcgt	cagggtccag	atgaagggtg	gggagcccga	cgcacggggc	36540
cgcgcgacca	tcgaggtgta	ctcctcggac	cagcaggccc	ccggccggga	acgctgggtc	36600
tgcccccccg	cgggatgtct	tgcggcgaa	ccggtggagg	ccccgccgag	tcaccacc	36660
cgcggcttcg	aaggcgtgtg	ccccgttcg	ctggacggct	tcacgaccg	gctggcggca	36720
gagatcttcg	gctacgggtc	gacattccgc	gggctgagcg	ccgctgggtc	acgcggtgac	36780
ctccacccgg	ccgaagcggc	gctccccctg	ggccatgcgc	aggatgccgc	ccggttcgga	36840
gcccggcgacg	cgtactcga	cgtgccctg	cacgccatgg	aactccggga	accccgccc	36900
gcccgaagcgg	gagtcggct	tccgttcgcc	tggaaacggct	tcctccctgca	cgcgtcgggt	36960
gcccgaagcgg	tacggctcgc	cctcgcgccg	acgggcggcg	acgctctgtc	ggtgacctc	37020
gcccgaagcgg	tcggctcgcc	ggttgctca	gcccgtcgc	tggccctcgc	ggagctctcg	37080

tccgacctgc	tgcgcccggc	gtccgtctcg	tacggggact	cgctgttccg	caccgcttgg	37140
ataccgccc	tcgtcgccc	ggaggcggag	tccgggccc	ggcgaccgtc	cgccggctgg	37200
gcggtgctgg	gccccgatcc	gctcggcgcg	gccaacgccc	tgaacctcac	gggaacctcc	37260
tgctcctgct	atccggacct	ggcggcgctg	atcgcgcccg	tcgacggcgg	agccgcgggtg	37320
cccagaggccg	tactcgcgcc	gtacggcgcg	gagccagccc	cggacggcgg	atctcccgcg	37380
gacgccgtac	gggcctcgac	cggccgggcg	ctgcaactgc	tgcaatcctg	gctgtccgag	37440
gaccggttgg	agcgaagccg	gctgatcgtg	ctcaccgggg	ggcggtggc	cgtcggtagc	37500
gacgaaggcg	tcaccgacct	ggtgagtgcg	tcggtccggg	gtctggtccg	ttcggcgag	37560
gcccagcacc	ctggcagggt	ctccctggtc	gacatcgacg	accgggagga	gtcctgggcc	37620
gtcctgagcg	cggcggcggg	atccgggtgag	ccgcaggctg	ccctgcgctg	cggccagatg	37680
aaggtgcccc	gcctcgggtc	cgtcgacgtt	cccacgaccg	gtatgcctga	gatgcccggc	37740
gtttgggggtg	ttgacgggtac	cgtgttgatc	actggcgggg	ccggtgtgct	gggtgggctc	37800
gtcgcccgctg	atctggtcgc	cgggcatggg	gtccgtcggt	tggtgtctg	cagcaggcgg	37860
ggccctgatg	cgccgggtgc	ggtggagctg	gtggccgagc	tcaccgctct	gggtgcggat	37920
gtcaccgttg	ccgccttgta	tcggccgcac	cgggatcgcg	tgcccgcgct	cttggacacc	37980
gttcccggcca	cgcacctctt	gactgggtgc	gtgcataccg	ctggtgtcat	cgtatgagcc	38040
actgtcacca	ccctcactcc	cagcgcgcatc	gacgcggtcc	tacgccccaa	ggtcgagccc	38100
gcgctcaacc	tccatcagct	gacggcgcat	ctcggcttga	cccgccttgt	gctcttctct	38160
tcggccgccc	ggctcttccg	cggcgccggg	cagggcaact	acgcggcggc	caacgccttc	38220
ctcgacgcac	tggcccaaca	ccgcccggcc	aacggcctca	atgcccagtc	cctggcggtg	38280
ggactgtggg	cgggaagccag	cgggatgacc	gggcaacctg	acgcggccga	cctcgcccgg	38340
atgggcccgtt	ccggcctcac	cgcgatgccc	accggggacg	ggctggcgct	gctcgacacc	38400
gcccagcggg	tggacgaagc	caccctggtc	acggccgcgc	tggacaccgg	ggccctgcat	38460
gcccgggccc	cagacggcac	gctgccggcg	ctgttccacg	cactcgtgcc	cgtaccgcgc	38520
cgatccggga	cctcccgggc	ggcccaggcc	gcggggccgg	atggactccg	ccagcggttg	38580
tcggggcctgg	tcgtggggga	gcgcgcgagc	gcgtgtctgg	atttgggtg	tggtcatgtc	38640
gcgaggggtgc	tggggcacgc	ggacccgagc	agcattgagg	agaacaaggg	cttcaaggac	38700
accggcttcg	actccttgag	cgcggtggag	ttccgcaacc	ggctgcacgg	tgcgaccggg	38760
ttgcggctgc	cggccacgct	ggtcttcgac	tacccgacgc	ctgcagctct	caccgatcac	38820
ctctacgacg	agcttctggg	ttccgcgag	gacgccgtgc	tcgccccgat	caccagggcc	38880
gcgtacgacc	cgggtggactt	cgaactaccg	acgctgcag	ctctaccga	tcacctctac	38940
gacgagcttc	tgggttcccg	cgaggacgcc	gtgctgccc	cgatcaccag	ggccgcgtac	39000
gacgagccga	tcgcgatcgt	ggggatggcc	tgccgctatc	cgggcggggg	ggagtccccg	39060
gaggacctgt	ggcagctggt	cgcgcacggc	cgtgacgcca	tctccgactt	ccccgccgac	39120
cgggctggga	acgtcgagga	cctctaccac	cccagccccg	accaccccg	caccagctac	39180
acccgtggcg	aggcttccct	gcacgacgc	gcggacttcg	accggaggtt	cttcgggatc	39240
tcaccgcgtg	aggcactggc	caccgacccc	cagcagcgac	tgctgtgga	aaccgctgg	39300
gaagccatgg	aacgggcggg	aatcaacccc	tcacccctga	agggcacccc	caccggcgtc	39360
ttcctcggcg	tcatgtacaa	cgaactacgg	actgccatgc	agcaggcggc	agaggtcttc	39420
gagggcata	tggccagcgg	tagcgcgggg	agtgtggcgt	cgggcggggg	ctcgtacacg	39480
ttcggctcgtg	agggaccggc	cgtcacgggtg	gataccgcgt	gttcgtcgtc	gctggtggcg	39540
ttgcactcgtg	cgcgcagggc	gttgogcaac	ggtgagtgca	ctctggctct	ggcgggagg	39600
gttgccgtga	tgtccactcc	tgccacgttt	gtcgagttca	gccggcagcg	ggggctggca	39660
gctgatgggc	ggtgcaaagc	cttcggcgac	gctgccgacg	gcaccggctg	gggcgaaggc	39720
gtcgtgtgtg	tgctggtgga	gcgtttgtcg	gacgcgcgcc	gcaacgggca	tcgggtgctg	39780
cgggtcgttt	cgggcagtg	tgtcaaccag	gacggggcca	gcaatggtct	gacggcgccc	39840
aatggtcctt	cgcagcaacg	ggtgatccaa	caggcgctgg	ccaatgcggg	gttggcgggg	39900
gcggatgtcg	atgccgtgga	ggcgacggc	acgggaaccc	ggctgggcga	cccgatcgag	39960
gcgcaagcgt	tgatcgccac	ctacggacag	gcccggctcg	cggaccggcc	gttgtggctg	40020
ggttcgctga	agtccaacat	cggtcacacc	caggccggcg	cgggcgtcgc	cggcgtcatc	40080
aaaatggtgc	aggcgatgca	gcacgggact	ctgcgccc	ccctgcacat	cgaaccagccc	40140
acgggcccagg	tcgactgggc	tacgggtgca	gtggagctgc	tgaccgaggc	cgtgccctgg	40200
ccggacagtg	accggccccg	ccgggtggct	gtctcctcgt	tcggtgtcag	cggatccaac	40260
gcccacgtca	tcatcgaa	cacccacac	acccacaca	ccaccgcac	ctgcccacat	40320
ctcccaatcc	ccccaggccc	cgcagactgt	gcaggcccat	cggccgggtg	gtggctgctg	40380
tcggcgaaga	cctcgaggc	cctggccgcg	caggcccggc	gcctgtcagc	tcacttgcca	40440
gccaacccc	atctgcgttc	ggctgatgtg	gcgcattccc	tgctcaccac	gcggtctgtc	40500
cacgcccagc	gcgccgtctt	catcgccggt	gaccgggatg	aggctcttgc	cgccctggac	40560
gcactggccg	acggcacccc	tgccctcac	ctcgttcagg	gccttgccga	tgtgagtgcc	40620
aagacggtgt	tcgtcttccc	cggtcagggt	tcgcagtggg	tgggtatggc	cgttgagctg	40680
ctggacggct	cggaggtttt	cgcgagcat	atggccgcct	gcgccagggc	cctggaaccg	40740
tttgtggact	ggtccctgga	ggacgtccta	cgcagacgg	acggtacgtg	gccactggaa	40800
cgcgtcgaag	tgggtccagcc	cgtgctgtgg	cgggtcatgg	tctcgtcgc	gggactgtgg	40860
caggcacatg	gcgttgagcc	tgctgcgggtg	ctgggccact	cccaagggtga	gatcgtcgcg	40920

gcttgcgtgg	cgaggagccct	gagtctggaa	gacggagccc	gcgttgcgc	gcttcgcagc	40980
caagccatcg	cggaaaccct	cgcaggacac	ggcggaatgc	tctcaatgc	cgccccgc	41040
accgacatcg	caccctgat	cgcccgctgg	aacgagcgga	tctccatgc	cacggtcaac	41100
ggaccgcatt	cgggtggtgg	cgcaggagac	cctgacgcgc	tcgaggcact	ccgcggcgaa	41160
ctggagaccc	gtggtctccg	caatcgtcgc	atcccggtcg	actacgcctc	acacaccctt	41220
cacgtcgagg	cgatccgtga	acggctcctg	gccgacctgg	cagtgatcca	gccacgtgcc	41280
gcgagcattc	ccgtgctgtc	caccgtcacc	ggcgcatggc	tcgacaccac	cgtgatggac	41340
gccgagtact	ggtaccgcaa	cctacgtcag	accgtggagt	tcgaagcagc	caccgcact	41400
ctcctcgacc	aggaccaccg	ctacttcgtc	gagatcagcc	cgcaccccg	actcaccatc	41460
ggtctacagc	agaccatcga	ggaaaccacc	gctccggccc	ggaccctctc	caccctccga	41520
cgaacgaag	gcaccctccg	gcacctgttc	acttccctcg	cccaggccca	cgcccacggc	41580
ctgaccatcg	actggacccc	cgccttcacc	cacaccgagc	cccgcaccac	ccccctgcc	41640
acctaccctt	tccaacacga	acgctactgg	ctggacacgg	cggagccggc	tgttgggcag	41700
ggagccggca	ccgacaccgt	cgcagcggt	ttttgggacg	ccgtcgaggg	cgaggagtgg	41760
cagacgttgg	ccgacacgct	cggcggttacc	gccgacgcgc	cgttcgactc	cgtgatgtcc	41820
gccctgtcgt	cctggcggct	ccgacagcgt	gagcagtcct	tggtggacgg	ctggcgttac	41880
cggatcgagt	ggaagccgtt	ccgcgcccc	gtgtcggcac	cggattccgt	gtcgggcacc	41940
tggtgggtgg	tcgttcccgc	ccatgccggc	gacgcggacc	gggagagggc	gcaagccgtg	42000
cggggcacgc	tggagtccct	cggccgtgcg	cggacgatcc	tggtcgoggt	ggaccgggcc	42060
gccgacgacc	gggggtcgtc	cgaactgaaa	ctcagggacg	ccgcgaccga	ggcgggtccg	42120
ccggccgggg	tgctgtccct	gctggccacc	gacgaacgtc	ccctccccgg	gcatgatgtg	42180
gtgccggggg	ggctggcagc	caacctggct	ctcgtccagg	cactgggcga	cgcgacagtc	42240
gatccccgc	tctgggtggg	cacctgcggg	gcggtctccg	ccggccggtc	cgaccggctg	42300
cgaaccccc	ggcaggccgc	ggtctggggg	ctcggacggg	tggtcgccct	ggagcaccog	42360
gaacgctggg	cggtctgat	cgatctgccc	gtggtcctcg	accgcgcgc	tgtggaacgg	42420
ctggtgacag	tacttgccgc	gtcgggcgag	gaggaccagc	tcgccgtacg	ggcgtcgggg	42480
gtcctcgtgc	gcaggctcgt	gcgggtaccc	gcacgccaa	tgccggacgg	cgtgcagtgg	42540
aagcccaggg	ggacggtcct	ggtgaccggt	gggacccgag	cgctgggcgc	ggaggtcgcg	42600
cggtggtcgt	ctcatggcgg	cgccgaacac	ctggtgctga	ccagccgtcg	cggcggctcg	42660
gcgcccgggt	cggccgagct	gacggacgag	ctgcttgccc	tcgggacgga	agtacactg	42720
gccgcctgtg	acatggcaga	ccgggacgcg	gtcgcgcgc	tgctcgccga	gcacgcgcgc	42780
agctcgggtg	tgacacccgc	cggcgtcctc	gacgacggtg	tactggacag	cctggaccgc	42840
gggcggctgg	agtcggttct	gctgccgaag	gtggccgcgc	ctcggcacct	gcacgagttg	42900
acgaaggacg	cgaacgtgtc	ggcattcgtg	ttgttctcgt	ccgcgcagg	cgtgctcggc	42960
agcgcaggcc	aggcgaacta	cgcagccgcc	aacgcctacc	tggaacccct	ggccgaacag	43020
cgcaggggcc	atggactggg	cgccattcgc	atcgcttggg	gcgcgtggga	ggcgggtggg	43080
ctggccgtgg	gcgacagcgt	ggtcgaggaa	cggctgcgcc	acggaggagt	ggtcccatg	43140
cgcccgacgc	tggcgatcac	ggcgctccag	cagacgttgg	accgggcgga	gaccgcggtg	43200
gtcatcgctg	acgttgactg	gocgcgctac	ctcaccgcgg	tcacaccgcg	cccatggctg	43260
gcggacctgc	cggaggtcgc	ccaggccctt	aacgcgcagc	acgcggctgg	tgcccttgc	43320
ggcacagccg	ggcagggtc	gtccccgctg	gccgagcgtc	tctccgggcg	cccggcacc	43380
gagcagcggc	gactggtgct	cgacctggtc	cgtacgaacg	tgccggcggt	gctcggccac	43440
gccggtgcgg	agtcgatcga	gtccggccgg	gccttcgcgc	agctgggctt	cgactctctg	43500
accgcgctcg	agctgcgcaa	caggctggct	gcggccaccg	ggctgcggct	gcccaccacc	43560
ctggtgttgc	actaccgag	cgtgcgctg	ctcgcgcatc	acctgtacgc	gcaggcgatc	43620
ggttcggacg	aggggccccg	ggcggatctg	tctcgcggc	ccgatccggc	ggcgggaccg	43680
gacgacgagc	ccatcgccat	cgtgtcgatg	agctgcccgt	tccccggcgg	tgtctcctcc	43740
ccggaggagc	tgtggcagct	gctgctggcc	ggtgaggaca	cgattaccgg	gttcccgac	43800
gaccgggact	gggatgtcga	cgccctgtac	gaccgggacc	cggaccaccc	ggggaccacg	43860
tattcccgcg	gcggcgcggt	cctgtccgac	gcggccggtt	tcgacgcgac	gctgttcggg	43920
atctcgcgcg	gtgaggcgtc	ggccatggac	ccgcagcagc	ggctgctgct	ggagacggca	43980
tgggagggtg	tcgagcgggc	gggcacgat	cccacctcgg	tacgtggcag	ccgggcgggc	44040
gttttcgtcg	ggaccaacgg	ccaggactac	gcccgcctatg	tgccccagga	accgatcggc	44100
gtggaggggg	atctgctggc	gggcaatgcg	gccagcgtca	tctccggccg	tctgtcgtac	44160
acgtttggtc	tggaggggcc	ggcgtcacg	gtggacaccg	cgtgttcgtc	ctcgttggtc	44220
gccctgcacc	tcgctgcaca	ggccttgccg	aacggcgaat	gctccatagc	cctggcggga	44280
ggcgtgtcgg	tgatgtccac	cccggcggcg	ttcgtggaat	tcagccggca	gcgggggctg	44340
gcggctgacg	ggcgggtgcaa	ggcgttcgcg	gaocggcgcg	acggcaccgg	ctggggcgag	44400
gggggttggc	tgctcctcgt	ggagcgtctg	tccagcgcgc	gccgcaacgg	tcaccgggtg	44460
ctggccgctc	tacgcggcag	cgccgttaac	caggacggcg	ccagcaacgg	cctcacgcgc	44520
cccaacggac	cctcgcagca	acgcgtcatc	cgccaggcac	tcgttgacgc	cgcgctgacc	44580
ggtagcgcga	tcgacgcggt	cgaagcccac	ggcaccggga	cccggctggg	tgacccgatc	44640
gaggcgcagg	ccctgatcgc	cacctacggt	caggaccgcc	cggcgaaccg	gcccctgtgg	44700
ctgggctcgg	tcaaatccaa	catcgcacac	acgcaggccg	ccgcgggcgt	cgccggcgctc	44760

atcaagatgg	tccaggcgat	ccgccatggc	gtacttccca	agacctgca	cgtggaccgg	44820
ccgaccagcc	acgtcgactg	ggaggcaggc	gcgggtggagt	tgctgaccga	ggccatgccc	44880
tggccggaga	ccgaccggcc	gcgtcgggcc	ggcatctctt	ccttcggcgt	cagcggcacc	44940
aacgcacaca	ccatcgtgga	gcaggcacct	gcggcggaag	acgagccgga	aacggggcca	45000
cccgcgatg	ctccgcccac	ggtggtgccc	tgggtgctct	ccgctgccac	cgaggacgcg	45060
ctgcgagagc	aggccgcacg	cctcgccacg	tacctcgacg	agcggcccga	gccaagcccg	45120
gcccagatcg	ggtcctccct	ggtcaccacg	cgtgcagccc	ttgaccaccg	ggcgggtggtg	45180
ctcggtgagg	accgcgacgc	tctgcggggc	gggctggttc	tgctggcgaa	cggggaagtcc	45240
ggtcccgctg	tcgtccgtgg	cctcgccagg	cccggacaga	aggtggcgtt	cctgttcacc	45300
gggcagggca	gccagcgact	gggcatgggc	agggagctcc	atcgccacct	gccggtgttc	45360
cggcagttct	tcgacgaggc	gtgcgcccg	ctcgacgcac	acctgcccgt	accgatagcg	45420
cccgcgctgt	tcgcgcaggc	ggatggggcg	gatgcggggc	tgatcgatgg	gacggaattc	45480
gcgcagccgg	cgttgttcgc	gctggagggtg	gcgttgtgcc	ggacgttgga	gttctgcggg	45540
gtcaggccgg	tttacgttgc	cgggcattcg	gtcgggtgaga	tcgcggcggc	ccatgtggcg	45600
ggggtgttct	cgctggagga	tgcggctcgt	ctgggtggtg	cgcgcgga	gttgatgcag	45660
gcgttgccgg	ccggtggtgc	gatggtcgcg	ctccagggtg	ccgaagacga	cctcctgcc	45720
tccttgactc	cttggctgga	gcaggaccgg	ctgggtatcg	cgccggtcaa	cgccgcagca	45780
tccacagtgg	tgtcggcgga	tgaggaggcg	gtcctggcgg	ttgcccagca	ctggcaggcg	45840
cggggccgca	aggttcgtcg	gctcactgtc	agccatgcct	tccactcacc	tcgtatggac	45900
ccgatgctcg	accagttccg	tgtggtcgtg	gagggtatcc	gttccgcgga	gccggccatc	45960
ccggtcgtct	ccagcgtcac	cggctcgtct	gccgagcccg	ggcagttgac	cactgcggac	46020
tactgggtgc	gccacgtccg	tcaaaccgtc	cgcttccacg	acgcccctca	gacctccag	46080
accgagaatg	tgaccgcgtt	tctggagatc	ggtcccgcacg	ggcaactctc	ggcaatggcc	46140
caggagacgc	tcaccgccca	ggtccatacc	atccccaccc	tccgaaagaa	ccggtctgag	46200
accaccggct	tgtcacccgc	actggcgcaa	ctccacacca	ccggcaccgt	ccccactgg	46260
accgcttacc	tcaaccacca	ccccacaccc	tccacaccgc	tgcccacct	ccccctccaa	46320
caccaccact	actggatgca	cggcggtacc	caggccaccg	atgtcagctc	cgccggcctg	46380
tcaggagcca	accaccgcct	gctggggggc	gcggtcccgc	tggccggtgg	ggaggggccac	46440
ctgttcaccg	gccggctgtc	ggtgcggacc	caccgcctggc	tggccgacca	ccaggctcggc	46500
agcaccgtcg	tgttgccggg	cactgccttc	gtcgaactgg	cggtacgggc	cggtgaccag	46560
gtcggctgcg	gccacgtgga	ggagctgacg	ctggaagcgc	cgctcgtgct	gcccagagagc	46620
ggcgccgtac	agatacagct	ccggctgcgc	cgggcggacg	aatccggacg	gcgtgaactc	46680
gtcgtgtacg	ggcggctcgc	gacggaccgt	gaggacctgt	ggtccgagga	ggaatggacc	46740
cggcacgcca	gcgggtgtcg	cgtcgcagca	gcgcctcgg	cccccgagcc	cgtccaactg	46800
accgtatggc	ccccggaagg	cgccaccgag	ctcatcgtga	aggacctcta	cgaacggatc	46860
gccggcacca	gcttcgggta	cgggtcccgc	ttccaagggc	tgccgcgcgc	ctggcggctg	46920
gacgacgcgg	tgttcgcgga	ggtcgtgctg	ccacaggatc	agtacgcgct	cgcgagccgg	46980
ttcggactcc	accggcgct	gctcgacgcc	gccctccacg	gggtcgcgct	ggggcagccg	47040
gcggctgaca	ccgcccagcc	gcacaccgac	cggatgccct	tctcctggag	cggcgttacc	47100
ctctacgcgc	ccggtgccac	cgcactgcgg	gtcgggttgg	acatcgcttc	gcccagggac	47160
gtgtcgctgc	tcgtcgccga	tggctcgggg	gctccggttg	ccgcggtgaa	ctcgtgaag	47220
ctgcgcccgg	tcgcggccga	cctggccagt	gccggtgtcg	ccgactogct	gttccggctg	47280
gagtggctga	aggcggtcga	cgacgagccc	ggccggggccg	aaccggggca	atgggcccctg	47340
atccgaacgc	cgcccgggtc	cgacttcacg	ccgggcgagg	acggcgcat	catcggaagc	47400
taccgggaca	tggccgcgtt	gaccgacgcg	ctcgacaagg	gagtcgcgct	cccgcagcgg	47460
gtgttgttgt	ccgccccgtc	ggaggaggag	caggaccagg	cgacgatct	cgcgagcgcg	47520
gtggacaagg	ccacgaacgc	gctgctcgca	gtgctccagg	agtggctgtc	cgacgaccgg	47580
ttcgactcct	ccaggctggc	tgtgctgacc	cgtcacgcgg	tgtccacggc	tggcgaggag	47640
gacgtgacgg	accttgccca	cgctcgttgg	tggggactcg	ttcgtcggc	gcagtccgaa	47700
catcccagcc	ggttcgtgct	ggccgacacc	gacggcaccc	agatcagtca	cgctgccctt	47760
ctgcccgttt	tgtgtccgg	tgagccgcag	gtcgcgctgc	gtgacggaac	ccggtatgtg	47820
ccgcggctgg	ccaggggcgt	tgcgtccggg	gacgggcccg	tggcgcgggt	ggaccggcg	47880
gggacgggtg	tggtagccgg	tggtaggggg	actctggggg	cctcgtggc	caggcatttg	47940
gtggttgagc	acggggtgcg	cgggttgttg	ctggtgagcc	gtcggggcgg	ggagtcggag	48000
ggcgccggcg	agttggtggc	tgagctgacc	gggctcgggg	cggatgtcac	ggtggcggcg	48060
tgtgatgtgg	gggaccgggg	ggcgtggcg	gagttgctgg	cggggattcc	ggccggtcat	48120
ccgttgacgg	gggtggtgca	tgcctcgggg	gtcactgatg	acgcggtgat	cgaggcgttg	48180
actgcggagc	aggtcggccg	ggtgctgcgg	togaaggctc	atggggcggt	caatctgcac	48240
gagttgacgc	gggggctgga	tctgtcggcg	tttgtgttgt	tctcttcggc	ggccggtgtg	48300
ttcgggaatc	cggggcaggg	caactacgcg	cgggccaatg	cctttctgga	tgcgttgccg	48360
gtcggcgcc	gggcggaggg	cctggctgcg	cggctcgttg	cctggggtct	gtgggaggag	48420
gccagcgca	tgacgagccg	gctggccggg	gccgatcttg	tccggatggg	ccgtgcgggc	48480
ctgcttcccc	tcaccaccgg	gcaagggtc	gccctcttcg	acgcccacca	ccggacagac	48540
gagcccctgg	tactgccgat	gaggctggac	accacggccc	tgctctccac	caccggacag	48600

ccgccggcgc	tgctgcgcaa	cctgggtccgg	gtccaggctc	gccggacggc	gggcgcggcc	48660
cccggaccgg	acggggccgc	caocttccag	cagcagctca	tcagcctgtc	cgtcgcggag	48720
cgccggcggg	tgctgctgga	gaccgtacgc	ggccaacggc	ccggcgtgct	cgggcactcc	48780
ggcccgaag	ccgtcgatgt	cgacaagggc	ttcatggaag	cgggcttoga	ctccttgagc	48840
gcggtggagt	tccggaaccg	gctgacgtcc	accaccgggc	tgccgatgcc	ggccaccgtc	48900
acgttcgact	acccgagccc	ggccgcgctg	gccgagcacc	tgctgacgcg	gttggttccc	48960
gaggtgccca	tgcccgcgga	ggagcagcac	ccgcacaccc	ggccogaaga	cgggcccgtg	49020
gacaggcccg	gagacgaaca	ggggggcgcg	atcgacgaca	tggacgtcga	cagcctcgta	49080
gaactcgccc	tcggcgaaatg	attcctgatg	ccgcacatgat	ccgggaggac	agcatgagca	49140
agccccatga	aaaagtagtc	gcggcgctcc	ggcgctcgct	gaaggccaac	gaacgcctgc	49200
gggagctcaa	cgacgagctc	gcctcggcgt	cccgcgaacc	ggtcgccatc	gtcggcatgg	49260
cgtgcgggta	tcccggcggg	gtgacgtccc	ccgaggaact	gtgggacctg	gtcggccggc	49320
gcaccgacgc	ggtgtcggag	ttcccgcgcg	accgtggctg	gaacgtcgag	gagctctacc	49380
acccggaccc	ggaccactcg	ggcacctcct	acgtgaggga	gggaggcttc	ctgcatgagg	49440
cggcggagtt	cgatccggtg	ttcttcggca	tgccccacg	ggaggcgctg	gccaccgatc	49500
cgcaacagcg	gctgctcctg	gaaacggcat	gggaggcctt	cgagcggggc	ggtatcgacc	49560
cactccggct	ccggggcagc	cggaccggcg	tattcgctgg	cgatcatgtac	aacgactacc	49620
tcaccgcct	ccagccggcc	cccgcggact	tcgaggggca	gctcggcaac	ggcagcgcg	49680
gcagcgtcgc	caccggccgg	ctggcctaca	cgttcgggct	ggaggggccg	gcggtcacgg	49740
tggacacggc	gtgttcgtcc	tcactggctg	ctctgcacct	cgccgccag	gcgctcgca	49800
acggcgaatg	caccatggcg	ctggcaggcg	gggtcgccgt	gatggccacc	ccggggccct	49860
tcaccgagtt	cagccggcag	cgcggtctcg	cgggtggacgg	ccggtgcaag	ccgttcgccg	49920
cggcggcggg	cggcaccggc	tgggcagagg	gcgtcggcct	gctgctggtc	gagcggctct	49980
cggacggccg	gcgcaacgga	cacccggtgc	tggctgtcat	acgcggaacg	gcggtgaacc	50040
aggacggtgc	cagcagcggc	ctgaccgtgc	ccaatggccc	ctcgacgag	cgcgtcatcc	50100
ggcaggcact	ggcgaacgcg	ggcctgtcgg	ccgccacagc	cgacgcggtg	gaggcacacg	50160
gcacgggcac	cccgtgggg	gaccgatcg	aggcccaggc	cctgatcgcc	acctacgggc	50220
aggaccgccc	ggccggccgg	ccgttgtggc	ttggttcgct	gaagtccaac	atcggccaca	50280
cccaggccgc	cgcgggcgcc	gccggagtca	tgaagatggt	ccaggccatg	cgccacggga	50340
ccctcccgaa	gagcctgcac	atcgacgccc	ccacgcccc	ggtcgactgg	gaggccgggg	50400
cggtggaact	gctcacccag	gccgtgccgt	ggcacgagac	cgaccggccc	cgcaggcggg	50460
gcgtgtcctc	cttcggggtc	agtggcacca	acgccacagt	gatcatcgag	gaggctcccc	50520
cgaccgaagc	tcccaggggc	gtgacggcgc	gggcgcgct	caacgcgag	accttgccgt	50580
gggtggcttc	ggccgtggc	gtagaggccg	tccgggcgca	ggccgggcag	ctgcgtcct	50640
atctgtcgga	ggctcaggac	tcgtcactgg	agggcacatcg	actctctctg	gccaccacgc	50700
ggtcggcggt	ccagcacccg	gccgtcgtac	tggcggccga	ccacgatggc	ttcatggcgg	50760
ggctggacgc	gctggccacc	ggggaaccgg	cgaagggctt	ggatgatggg	gaggccgtat	50820
cgggcggcgg	agtcgccctg	gtcttccccg	gccagggctc	ccaatggggc	ggaatggcgc	50880
tcgaactgct	ggactcctca	tccgtgttca	gagaccggat	ggaagcctgc	gcgcaggcgc	50940
tgagccocta	catcgactgg	tcactgaccg	aggctctcgc	ctcctgcgaa	ggcgagctgg	51000
aacgggtgga	cgtgggtccag	cccgcgctgt	gggcctgat	ggtctcgtcg	gccgaactat	51060
ggcgttccct	cggagtccgg	cccgcgcggc	tcctggggca	ctcgaggggc	gagatagcgg	51120
cggcctgtgt	ggccggcgcg	ctcagcttgg	aggacgccgc	gctggtggtc	gcgtgcgca	51180
gccaggccat	cgcgaccgag	ctggccggcc	ggggcgcaat	gctgtccgtc	gccctgccga	51240
aggcacgggc	ccaggactgg	atgacggggc	gggcggaaac	gctgtcggtc	gcggcggtca	51300
acgggcccgg	atcagtgtg	gtctccgggg	acgttgacgc	gggtggaggag	ctgcgggcgg	51360
agctggccgc	cgagggggtg	cgggtccgca	ggcttcgggt	cgactacgcc	tcgcacagct	51420
cgcattgtga	gcggatccgc	acacgtctgc	tggcggcgct	cgccccggtc	tccccgcgcc	51480
cttcggagat	cacctgttac	tcgtccgtga	ccggtgggtc	catcgacacc	acgaccatgg	51540
acgccgagta	ctggtaccgg	aacctgcggc	agaccgtgga	gttcgagcgg	gcggtccgca	51600
cctcgatgtc	cgacggctac	cggttcttca	tcgagtccag	cccgcacccg	gtgctgacga	51660
cgggcatcga	ggagaccgog	gaggacgctg	accggttcgc	ggcggcggtc	ggttcgctgc	51720
gccgttcgga	cgggtggccc	gacaggttcc	tgactgcgct	cgcggaggct	cacgtgcgcg	51780
gcgtgccggt	ggagtgggcg	gtgatgttcg	ccggccggcc	cgtgagtcag	cccgatctcc	51840
cgacgtactc	cttcacgagg	cagcggattt	ggctggcccc	cgacacgtcc	cccggcgacg	51900
acggcggcgg	cgacgaacgc	tcggagacgc	ggttctggga	ggctgtcgag	cgccaggacc	51960
tcggcgaact	gagcgagacc	ctgcggatcg	gtgacgcgga	ccggcaggcg	tcgttgggtg	52020
agttgttgcc	ggccctgtgg	acgtggcgtg	agcagaaccg	gtccgccgcc	gtcctggaca	52080
gctggcggta	ccgggtctca	tggcggcccg	tctccccggc	gtccgatcca	gccttgcggg	52140
gcacctggct	gatcgtggtc	ccggcgggga	cggcggacca	gcagtggggc	gaagcgtctt	52200
cccgagccgc	cgagggcctg	ggagaccagg	ctgtccgggt	cgaactgggc	agggccgaag	52260
ccggccggga	ggagtacgcg	gccaggctcg	ccgaggcggc	ggccggcggt	ccggtggccg	52320
gcgtgctttc	cctgctcgcc	ctggccgagg	agccggcgga	cgcgcacccg	gtgtggcgcc	52380
cgtatgtcac	cagcacgctg	gcacttatgc	aggcgtggg	cgacgcgggg	atcggcgcg	52440

cgctgtggct	ggccacccgg	ggcgcggtct	cgatcgggcg	gtccgacaag	ccggtcccgt	52500
cgacagccgc	acaggcccag	ctgtggggcc	tgggcggggt	catgggactc	gaacaccccg	52560
aacggtgggg	tgggctcgtt	gatctggccg	agacggccga	cgctcgtgcg	acggcgccgg	52620
tggccggcat	cctggccggc	ggtctcggcc	cgagggacca	gtgcgcgggtg	gggtcctccg	52680
gcgtgtacgt	acggcgctctg	gtccgcgcac	cgctcgaccg	gcgagcgcgg	aggccgtcct	52740
ggcacacgtc	ccgtacggcc	ctggtcaccg	gtggcacccg	cggctctggg	gcgcacgtog	52800
cccgatggct	ggcgcagcacc	ggcgcggaac	acctgggtgt	caccagcagg	cgcgccccgg	52860
acgcccctgg	gacggacgag	ctgtgcgcgc	aactgtccgc	cctcgggggtg	cggttgagcg	52920
tgggtggcctg	cgatgtgtcc	gaccgggacc	aactggccgc	cacattggca	cgattgaccg	52980
ccgacggcca	caccgtccgt	acgggtggtac	atgccgcgcg	ggtcagtacg	ccgggcgcgc	53040
tggccgacct	cgggcccggc	gagttcgccg	aggccgtcgc	gggcaaggcg	gcgggcgcgc	53100
cgcacctoga	cgaactgctc	ggcgacgcgg	agctggacgc	cttcgtgctc	ttctcatcca	53160
acgcccgcgt	gtggggcggc	ggcggccagg	gcgcctatgc	cgccgccaac	gcctacctgg	53220
acgcgtggc	caaacggcgc	cggctcccgc	gcccgcctgc	gacctccgtc	gcatgggggg	53280
cttggcccg	cggcgccatg	gcccgcggagc	gtaccgcga	cgagcagctg	cgccgcggag	53340
gggtgcgggc	gatggacca	gogatggcga	ttccgcact	ccaggaggcg	ctggagcag	53400
aggagacgtt	tctcgcgggtg	gcccacatgg	actgggaccg	tttcctcccg	tccttcacca	53460
tggcccgcc	ccgccccctc	ctcgacgacc	tgccggaggt	ccagcggcag	cggctgagcg	53520
cggccccgtc	atgggcccacc	gcggagaccg	acggcccggc	actcgcgcag	cagctcgcgc	53580
gggtctttga	accggagcgc	gggcggcgcc	tgctcgacct	ggtgcgcaag	cacgcggcgg	53640
cgggtgctcgg	ctacgccggc	ccgaacgagg	tcgaggcgga	acgggccttc	cgggagctgg	53700
gcttcgactc	cctcaccgcg	gtggagatgc	gcaacgcact	ccagccggcg	accgggctga	53760
cgtgccccgc	cacctggtc	ttcgaccacc	cgacgccccg	cgtctggcc	gcgcactctgc	53820
gggatgagct	gttcggtgtg	caggacgaca	cgccggaacc	ggcgcgggcg	tcggcaccgg	53880
acgacgaccc	gatcgccatc	gtgtcgatgg	gctgccgttt	ccccggtggt	gtctcctccc	53940
cggaggggct	gtgggagctg	ctgctgtccg	gcccgtacgc	catgtcgtcg	ttcccagtg	54000
accgaggctg	ggacctggac	agccttgccg	gtgacggccc	cggacagatc	ggcggcgggt	54060
acacccttga	gggcggcctt	ctcgatgacg	cggccgggtt	cgacgcggcg	ctgttcggga	54120
tctcgcgcgc	tgaggcgctg	gccatggacc	cacagcagcg	gctgctgctg	gaggcttcgt	54180
gggaggcctt	cgaacgagcg	ggcatcccct	cggccgacct	gcggtccagc	cggaccgggg	54240
tgttcatogg	cgcttcctca	cagggatacg	cccaggctgc	cgccggagtcc	gcggaaggag	54300
tcgagggaca	tgtggtgacc	ggtgacgcgg	ccagcgtcat	gtccggccgt	ctgtcgtaca	54360
cgttcggtct	ggagggaccg	gcccgtcacgg	tggataccgc	gtgttcgtcg	tcgctggtgg	54420
cgttgcaact	ggctgcgcag	gcgttgcgca	acggtgagtg	cactctggct	ctggcgggog	54480
gggtcgcgggt	gatggtgacc	ccggcggcgt	ttgtcgagtt	cagccggcag	cgggggctgg	54540
cagctgatgg	gcggtgcaaa	gccttcgcgg	acgctgccga	cggcaccggc	tggggcgaa	54600
gcgtcgggtg	gctgctggtg	gagcgtttgt	cggacgcgcg	ccgcaacggg	catccgggtg	54660
tggcggctgt	ttcgggcagt	gctgtcaacc	aggacggggc	cagcaatggt	ctgacggcgc	54720
ccaatggtcc	ttcgagcaa	cgggtgatcc	aacaggcgct	ggccaatgcg	gggttgccgg	54780
gggcggatgt	cgatgccgtg	gaggcgcacg	gcacgggaac	ccggctgggc	gaccgatcgc	54840
aggcgcaggc	gttgatcgcc	acctacggac	aggcccggtc	ggcggaccgg	ccgttgtggc	54900
tgggttcgct	gaagtccaac	atcgccaca	cccaggccgc	cgcgggcgtc	gccggcgta	54960
tcaagatgat	ccaggccatg	ggtcacggga	cgtgcgcccg	tacgtgcat	gtcgaccggc	55020
cctcgtccca	ggtggattgg	gaagccggcg	cgggtggagct	gtgaccgaa	gccatgccct	55080
ggcccagggc	cgaccggccc	cgccgggcag	cagtctcctc	gttcggtgtc	agtggtagca	55140
acgcgcacgt	catcatcgaa	cacgccccgc	aggctcactc	cgcctcccag	gccccggaac	55200
cgttgaagtc	cccggatgct	gtggaggctg	atcgaccggt	cccgtggctg	ctgtcggcgg	55260
gcagtgcgc	ggcgttgggc	gaggtggccg	aacggctggc	cgcctacgcc	gaatcgacc	55320
cggaggtcag	tgccggccgag	gtcgcgttct	cgtctcgac	caccgggtcc	ctgttgccgt	55380
gccgcgcgc	cgtcgttggc	gcggaccgcg	acgagctggt	ccagcgcata	cggctccgtg	55440
gcggggggcac	caccgccccg	ggcgtcttct	gcgggacggc	gagttcggag	tgaccacagg	55500
cgttcctgtt	ctccgggcag	ggcagccagc	gactgggcat	ggggcatgag	ctgtacgcgc	55560
cgcacccgga	gttcgcccag	gcgctcgacg	aggctctcgg	tcacctcgac	gtgttcgggg	55620
accggccggt	gaaggagggtg	ctgttcgcgc	aggcggatgg	ggcggatgcg	gggctgatcg	55680
acggggcggg	gttcgcgcag	ccggcggtgt	tcgacttgg	ggtcgcgctg	taccggaccc	55740
tggaagcatg	gggcatcacc	cccactatc	tggccgggca	ctcccttgg	gagatcgccg	55800
cggctcatgt	cgccgggggtg	ttcagcctgg	aggacgccgc	tcgcctgggt	acggcgccgg	55860
ggcagctcat	gcaggccctg	ccggcggtg	gcgcgatggt	ggcgtccag	gcctccgagg	55920
acgagatcct	ggccatctcg	gcgcgtggc	tggaggggga	cggggctcgg	atcgccgcg	55980
tcaacggtcc	cgcctcgggtc	gtcgtctccg	gggacggaga	agccgtcctg	gcgatcccg	56040
ggcactggcg	ggcacagggc	cgcaagaccc	gtcggctcag	cgtcagccac	gcctccact	56100
cacccacat	ggatcccatg	ctcgacgggt	tcgcggggt	cgtcagcggc	atgcaccttg	56160
tcgagccggt	cattccggtc	atctccaacc	tcaccggtcg	cctcgccgat	ccggggcagc	56220
tgaccagcgc	cgactactgg	gtccggcacg	tcgcccaagc	cgtccggttc	cacgacggcc	56280

tacagaccct	gcacgatcag	ggcgtcacca	cctacctgga	aatcggccct	gacgcccagc	56340
tcacggccat	ggctcaggag	gcccgtgagc	cccagtocca	caccgtctcc	accctgcgca	56400
ggaaccagcc	ogaaaccacc	agtcgtgcta	ccacgtctgc	gcgactccac	accaccggtg	56460
ccacccccga	ctggatcacc	tacctcaacc	accagccctc	atccccgaca	ccgctgcccc	56520
cctacccctt	ccaacaccac	cgctactggc	cgccggcgga	tgctcaggcc	gccgatgtca	56580
gctccgcccg	cctgtccggt	gcgaaccatc	cactgtctgg	agccgcggtc	ccgctggccg	56640
acggcgacgg	ccatctgttc	accgggcccgc	tgctggcacg	gacgcaccgc	tggtggcccg	56700
accaccaggt	cggcggaac	gtcgtactgc	cgggcaaccgc	cttcgtggaa	ctggcggtac	56760
gggctggtga	ccaggtcggc	tcagccagg	ttgaagaact	gacgtctgaa	gcgccgctgg	56820
tgctgccoga	gagcgccgcg	gtccagggtac	agctccggct	gggcccggcg	gacgagtccg	56880
gccgacgtga	cctcaccgtc	tacgggcccgc	tgccgggggg	cggcgaggac	ctgtggctcg	56940
aggaggagtg	gacccggcac	gccagcgggg	tcctctccag	cgccctcgcc	cccgaaccgg	57000
tcgactgac	cgtatggccg	ccgtccgccc	ccgaggccgt	gccggtggag	ggcttctaca	57060
ccggtctggc	cgagagcggg	tacggctacg	gccccgcctt	ccaggccctg	cgggcccgcct	57120
ggcgtcaggt	cgacacggtc	ttcgccgagg	tccaactccc	tgaggtggta	cgggaggagg	57180
ccgcctccta	cccatccac	ccggtctccc	tggtgcccgc	cctccaagcc	gtcggtttcg	57240
tcacggacgg	gagcgacaac	cccgtggtac	ggatgcogtt	cgctgtgtcc	ggcgtgtcca	57300
tgtacgcgtc	cggcgccctcc	gagctgcggg	tgccgctcgc	ccggacagga	ccggagacgg	57360
tcaccttcgc	cgtcaccgac	cccaccggcc	gccccgtggc	ctcggtccgc	tcgctcgtca	57420
tgcccccggg	cggcaccgga	gtaccgcgcc	tgacacgcaa	cgggctccac	gaggtggtct	57480
gggagcaact	cctcgatgcg	ccggccaccc	ccgcgaccga	gtgcgccgtc	atcggggacg	57540
cggacgcggc	ggcgtctgtg	ggcgccggagg	cgccaccgga	cctggcgctg	ttgggggaag	57600
cgggtccccc	gctggtggtg	gccgtggccg	gcggcgacgg	tacacggcg	gcactggagc	57660
gcgcccttgg	ctgggtgcag	ggatggatgg	cggaggagcg	gttcgccggt	tcggcgctcg	57720
ccgtcgtcac	ccgtggtgcg	gtggcggtcg	gtgcggcgga	ggtgctggcg	gacgtgcggg	57780
gtgcgcccg	cgacggccctg	gtgaagtccg	cggagtcgga	gaaccggggc	cgcttctctg	57840
tggtggatgt	ggacggcacc	accgagtcct	ggcgggcgct	gccgactctc	ggcggcggcg	57900
acgagccgca	gatcgcgctc	cgcgacgggc	agccgtactg	ccccgcctg	gtgcgtgcgg	57960
gtgaggacgg	cggctcgtcg	ctgcccccg	ccggggcgga	cgccgtggcg	ctggagacag	58020
gcgaggccgg	cagcctggac	gggctccggc	tcgccccctg	cgaggacggc	caggcgccgc	58080
tgctgccggg	gcaggtgcgg	atcgccggtc	gtccgcagg	cctcaacttc	cgtgacgtcc	58140
tcggtgcgct	cggcatgtac	cccgccggac	tcgacctcct	cggcagcgag	atcgccggcg	58200
aggtgctgga	gaccggcgat	ggggtgaccg	gcctcgcggt	ggcgaccggg	gtcatggggc	58260
tggtgcggcg	cggcttcggg	ccgatggccg	tcgcccagag	ctggcgggtc	gtacggatac	58320
cgtccggctg	gaccttcacc	cgcgcggccg	gtgttcgggt	cgcccttctc	accgccctgt	58380
acggactgcg	tgaactgggt	gggctggcgg	cgggocagcg	ggtgcttggt	cacgcggccg	58440
ccggtgctcg	gggtacggcg	cgggtgcaac	tcgcccggct	attggggggc	gaggtgtacg	58500
ccacggccag	cgcccccaag	caggagtatg	tgccggtact	ggcggtggac	cgcgccgcta	58560
tcgctcctc	ccgcaccctg	gacttcgctt	ccagcttccc	tgaggtcgac	gtcgtgctga	58620
actccctggc	cggggagtac	gtggacgcct	cgtcgggggt	gttgcccgag	ggcgcccggt	58680
tcgtggagat	gggcaagacc	gatgttcggg	atgctgccgc	gtacgacggt	gtgacgtacc	58740
ggacgttcga	cctggggcag	gccggtccgg	agctgatcgc	ccgaatgctg	ggtgagttgg	58800
tggagtgggt	caggccgggg	gaactcactc	cggctccgcac	agccgcctgg	gatgtccggc	58860
gcgcggtggg	cgcgttcggt	tggtatgagc	aggcccggca	cacaggcaag	atcgtcctga	58920
cgggtccgcg	cgacctggac	gccgacggca	cggctcctgat	caccggcggc	accggcacgc	58980
tggcggtctc	gctcgcccgg	cacctggtca	ccgaacacgg	cgtacgacac	ctgctgctgg	59040
tctcccgcac	gggagaacgg	gccgctctcc	gtcgtgaact	ggaggagctg	ggcgccgagg	59100
tacggatcgc	ggcctgcgac	atggctgacc	gcgcggcggt	ggccgaactc	ctcgacggca	59160
tcccgtcgga	gcacccgctg	accggtgtgt	tccacggcgc	gggtgtcctg	gacgacggcg	59220
tggtcacogg	cctcgactcc	gctcggctgg	cacgggtgct	ggctccgaag	gtggacggcg	59280
ccctccacct	gcacgaactg	acggcgaggc	tggaacctct	ggcgttcgtc	ctgttctcct	59340
ctatgtcggg	tctcctcggc	gcctccggcc	aggccgggta	cgcggcgccg	aacatgttcc	59400
tcgacgcgct	cgcacagcag	cggcgtgccc	aggccctgcc	cgcgtgtcgc	ctggcggtgg	59460
gtttgtggga	gaccgcgagc	gcgatgaccg	cgcacctgag	cgacaccgac	ctgcgcggca	59520
tgggcgggat	cggcatgctc	gggctcaccc	gcaacgaggg	catggaactc	ctcgacgcgg	59580
cctggcagag	cggcgaggcg	ctgctggtcc	cggctccgctg	ggaccaccgg	gtgctgcggg	59640
agcgggcctc	ctcgggcgcc	cgggtgccct	ccctgctgcg	gaggctggtg	cgggcccoga	59700
ggcgccgtac	ggtgccggag	agcgccaagg	gcgcgggcgg	cgggctgcgg	gagcggctgg	59760
cgacgctgcc	ggaggcggag	cgcgggggca	tgctcatcga	gctggtggcg	gggcacgtgg	59820
cccggtgtgt	ggcccatgcg	ggcaccgatg	cgggtgctgg	ggaccgcccg	ttcaaggagc	59880
tcggcttcga	ctcgtgacc	tccgtggagt	tccgtaaccg	gctgaacgaa	gcgaccgggc	59940
tgcggctgcc	ttcgacctg	gtgttcgacc	acccacacac	taccacgctg	cgggcccggc	60000
tcgacgccct	gctgcggggg	gcagagacgg	cgacaacggt	tgctgcccc	acctcgccgc	60060
acgaggaact	cgaccgactg	gcaacggtgc	tgctgtcacc	cgcgttgaa	atggcggatc	60120

gggacggcct	cgccgcccgg	ctccgagccc	tggcttccca	gcttggcgag	ccgactggtc	60180
cgcccgatgg	cagcacccgtc	gccgaccgga	tccagtcggc	caccgatgac	gagctcttcg	60240
agttgctcga	cgacagggttc	gagaactcat	gagccaacac	gacgatgcit	ctgacgcgct	60300
gaggacgggg	gatgttccga	tgacacagtt	tccgacgaac	gaggacaagc	tccgcgacta	60360
tctgaagcgg	gcggtcaccg	acctgcacca	caaccgtgag	cagctggccg	cggccgaggc	60420
caagaaccgg	gaaccgctgg	cgatcgtgtc	gatgagctgt	cgcttccctg	cgaggatcag	60480
gtcgcocgaa	gccttgtggc	agctgggtcg	tgccggtgaa	gacgtgatct	cgctggttcc	60540
caccgaccgt	ggatgggacc	tcgacggcct	ctacaaccgc	gatccgggga	acagtggcac	60600
cacctacgtg	cgagaggggc	ggttcctgtc	cgacgcgacg	gagttcgacc	ccgccgtgtt	60660
cgggatctcc	ccgcgtgagg	cgctgggaat	ggaccgcgag	cagcggctga	tgctggagac	60720
ctcgtgggag	gccttcgagc	gggcccggcat	cggctccggca	tcggcacgcg	gcagccggac	60780
cgggtgtgtc	atcggcgccct	ccgccccagg	ctacagcttg	ctgttccaga	actcgcggga	60840
ggaggccgag	ggcctcctgg	ccaccggtga	ctcggccagc	gtgatctccg	gccgggtctc	60900
ctacaccttc	ggcctcgaag	gacctgcggt	cacactcgac	accgcgtgct	cctcgtccct	60960
ggctcgtctt	cacctggccg	tgcgctcggg	tggccagggc	gagtgctcca	tggcgttggt	61020
ggcgggcgct	tcggtgatgt	gcacgccggc	gatcttcatc	gagttcagcc	gccagcgagg	61080
tctcgcggcg	gacggccggg	gcaagccgtt	cgccgcggcg	gcggacggca	ccagctgggg	61140
ggaaggcgcc	ggagtctgtc	tcacgagcgc	gctggaggag	gcccgcagca	acgggcaccc	61200
gggtgtggcc	gtcatccgcg	gcagtgccat	caaccaggac	ggtgccagca	acggcctgac	61260
tgccccgcac	gggcccgtgc	agcggcggct	gatccagcag	gcgctggcgg	acgcccagct	61320
gtcgcocggc	cagatcgaca	tggtcgaggc	acacggcacc	ggcacctcgc	tgggggatcc	61380
gatcgaggcg	caggcactgc	tggaaacgta	cggtgccaac	cgccccgcgg	accgcccgtc	61440
ctggctcggg	tccgtcaagt	ccaacatcgg	acacaccacg	gcggcggccg	gtctcgcgtc	61500
cgatcatcaag	accgtacagg	cgctgcgaca	cgcccacctg	gccaggacac	tgcaogtcga	61560
ccggccgacc	ccgcgcgtgg	actggtcgtc	gggtgggggt	gaactgctgg	ccgacgacca	61620
ccgctggccc	gagacggggc	agccccgcgc	agccgcctg	tcctcgttcg	gggtcagcgg	61680
caccaacgcg	cacgtcgtcc	tcgaacaggc	ggccgcctcg	gagaaccgcg	ccctccgcgc	61740
tcggggaggg	gaccgcgtcg	cggcgcgcgc	ggtactcccg	ctggtgatct	ccggcaagac	61800
gcccgaagcc	ctgcgggctc	aggcggggaa	cctggtgtcc	catgtgcgcg	agcaccggga	61860
cctccggctg	gaggacctcg	ggtactcgtc	ggccaccacc	aggctcggcc	tcggacaccg	61920
ggccgtcgtc	gtggcggaca	cccccgacgg	attcctccgt	ggctgcgagg	cggtggagcg	61980
cggcgagacc	ccggcgtcgg	tggaccgggg	cgtggtccgg	gggcgcggca	cgaccgcgtt	62040
cctgttcacg	gggcaggggc	cccagcgggt	cggcatgggc	cggcagctct	acgcggcgat	62100
ccccgcgttc	gcgcggttcc	tcgacgaggc	ctgctcccat	ctogaccgct	ttacgaagca	62160
gcccctgagg	gacgtgctgt	tcgctgccga	gggcagcgcc	gaggcagcgc	tcctggaccg	62220
taccggattc	gcccagccgg	ccctgttcgc	cctggagggt	gcgctgttcc	gcaccctgga	62280
gtcctggggg	gtgaccocgg	actacctcgc	cggacactcc	atcgggtgag	tcgctgcgcg	62340
ccatgtggcc	ggtgtgctct	cgctgggaga	cgccaccgcc	ctggtgaccg	cgcgtggcaa	62400
cctcatggaa	cagctccccg	cggggggcgc	catgctcgcc	ctgcaagctt	ccgaagccgg	62460
ggtgtctccc	ctcctcgacg	gcgccgatgg	cctggtgtcc	gtcgcgcgcg	tcaacagccc	62520
ccgctccacc	gtggttgccg	gagacagcga	cgccctcgcc	gcccctcgcc	gccaggcccg	62580
ctctcagggc	atcaaggccc	gccacctcac	tgtcagccac	gccttccact	ccccgcgtat	62640
ggaccccgct	ctcgcgcctt	accgcgagac	cgcgcgagcag	ctctcctacc	accgcgcgcg	62700
tatcccgatc	atctcgaccg	tcaccggccg	gtccgtcacc	accgagatgt	ccgaaccocg	62760
ctactgggtc	cggcacgccc	acctggaaact	cgggttcacc	gatgccgtgg	ccacgctccg	62820
cgagcacggc	accaccgcct	acctggaaact	cggccccgac	gcccgtcctca	ctgccatgac	62880
gcagcaaac	ctggcggggc	acggcacctc	gggcaaggag	tcacacctcg	cggcgggtgat	62940
gcgcaggaac	ggcccgaggc	gagggtcct	gaccagcgcc	gtgtcccagc	tgctcgcgcg	63000
gggcacccgc	gtcgactggc	gggcccgtgt	cgcggatgtg	gatgggcagg	tcgtccagct	63060
gcccacctac	gccttcacgc	gcagccggta	ctggccgcag	gcataactga	cccggccggc	63120
cggggggcgc	tccgcgacgt	cgctgttcca	cctgcgctgg	gtgccggtga	cggcccagga	63180
cacggcgccg	gcggacgact	gggcgttgct	cggcggggcc	gacgcgctgc	cgggccaggg	63240
cttcgcccgc	ctggcgtccc	tgggggagac	gatcgacggc	ggatcggccg	caccccgcac	63300
ggtgtgtgtg	ccgttgctgc	ctccggccga	cggcgcgccag	gattccgcgc	ccacgcacga	63360
cgcgcgccac	cgggcgcgtg	cgctggctca	ggcttggttc	gccgacgatc	gcttcacctc	63420
ctcccggctg	gtgttctcca	cccgtgggtc	ggtggccgtg	accgacgagg	aataccccga	63480
ggactccgtc	gacgccttcg	catacgccct	cgtgtggggg	ctgctgcgtt	cggcccagac	63540
ggagaaccgc	ggccgggttc	gcctgggtga	cctcgacccc	gacgccgacc	cggacgcggc	63600
cggcgacggg	tgcccgggtc	cggccgcgcg	cctggacggc	gacgaaccgc	agctggcgat	63660
gcgcgaggc	gtggtccacg	ctcccgggct	caccggggtc	acggccgcgc	ccaaggaccc	63720
ggaccgggca	cccgcocggg	tcgaccacgg	cggaaaccgt	ctgatcacgg	gcgccaccgg	63780
tggactcgga	ccgctgctgg	cccgccatct	ggtcgtcgag	cacggcgtag	gccacctgct	63840
gctgacgagc	cgctgcggcg	cggcggccgag	cggcgcgccag	gcactgctgg	acgagctcgc	63900
cgacctgggt	gccgaggcca	ccgtggtctc	ctgcgacctg	gctgaccggg	aggcgggtgg	63960

cgccctgctg	gcccaggtgc	cgcccgcgcg	tccgctgacc	gcggtggtgc	acgccgcggg	64020
cgtcctggac	gacggcgtga	tcccgtccct	gagcccggaa	cgcgtcgacg	gggtactgcg	64080
gccgaaggcg	gacggggccc	tgacactgca	tgagctgacc	aaggatctgg	acctggccca	64140
cttcacctcg	ttctcctcga	ccgccggtgt	ccctggcagc	gcccggccagg	gcaactacgc	64200
ggccgcgaac	acgttcctgg	acgcgctggc	ccagocaccg	cgggcagcag	ggctggccgc	64260
tgtctcgctc	gcctggggaa	cgtgggaacc	gagcggcggc	atgaccggcg	ggctgacgcg	64320
cgcagacctg	gagcgcctga	cgaaggagg	catgccaccg	ttgtccccc	gggacgggct	64380
ggcgtctctc	gatgcccca	tcgcttcggg	gcggggccctg	gtggtgcggg	ccgtgctcga	64440
ttcgcacctg	ctgcgttccc	ggatcgggac	gaacgtaccg	gcgctgctgc	gcggcctcat	64500
cgagccccgg	cccggtggagc	cgtctgcccc	aggggaggca	gccgaggcac	tcgcccctgcg	64560
gatggcctcc	tgctccggcg	cggagcgcac	gggcgtactc	ctggacctgg	tccgcgccga	64620
cgcggccacg	gtgctgggac	atgacggctc	gcacgccatc	gaccgcggagc	gtggactgct	64680
cgaagcgggc	ttcgactccc	tgacgacgct	ggagctgcgc	aaccggctgg	ccgagggcac	64740
cggactggcc	gtcccggccg	gttacctcta	cagtagcccc	accccgaaac	tgcttgccga	64800
acacctggcg	ccgcggttgg	ccgagtcgcc	gcagtcggcg	gcggcgaccg	gagccgacgg	64860
accggccgag	ccgctgagcg	tgctcttcca	gcaggcgtat	gacctcggca	aggtaccoga	64920
gggcatgacc	ctgctcagga	gcgcgtccgc	gctccggccc	acctacgaca	ccccttcgga	64980
cctcagtga	ctgccgcagc	ccactcgcct	ggcccggtggc	cccgaaactg	ccagctgct	65040
gtgcttctcc	gccatcgtgg	cactcgcggg	ctcgcaccag	tactcgcgct	tcgctcgtc	65100
cttcgcgcag	gaacgggacg	ttcgggtcct	ctacgcgcgc	gggttcttcg	ccggggagct	65160
cctgccgacc	agcctcga	cgggtcatcga	caccaggtg	gaaaccgtgc	ggcagcaggc	65220
cgcggacggg	ccggtggtgc	tcgtcggcgc	gtcttcggcg	ggctggctcg	cccatgcgcg	65280
cgcggcccg	ctggaggcgc	tggaacacc	accggcagcc	gtggtcctgc	tggacacct	65340
cctgccggac	gaccagttcc	tcgcccgtga	ccaggaccgt	ttcatcggcg	gagtcttoga	65400
ccggcaggac	cggttctcca	tcggggagga	cgtcagcctg	tcgcgatgg	gctggtatct	65460
gcacctgttc	gcaggctgga	agcccaccgc	gatctccgtc	ccggaactgc	tggtccgggc	65520
gagtgagccg	ctgcccagcc	cttcggccgc	ccgcgcgagg	gcccgcgact	ggcggacctc	65580
atggcatgtg	gcacagcaca	gcgtcgaggt	gcccggcgat	cacttcacga	tgctggagga	65640
attcaacgac	gccacggccg	acgccgtccg	acgctggctt	ctcgacattg	actgaaaggc	65700
ctgtccatgg	atctggaaac	ccaacttctc	tccccggcat	acctacggaa	cccgaccccg	65760
ctcaacgcgc	cattgcgttc	cgcgcaccct	gttcaacgtg	ccgtggcttc	ggggggcctg	65820
tcggtctggg	tggtgaccgc	ctacgaggac	gtgcgcgcgc	tgctcgcgca	ttccaggctg	65880
ggcaaaggcg	tcacgcagct	ccgcgaggcg	gtactgctca	acgcgggtga	cgacgagcgg	65940
atcagccagt	tcaccgactc	cctcaccgag	cacatgctca	acagcgaccc	acccgaccac	66000
acccggctgc	gcccgcctgt	cggcaaggcg	ttcacccgcg	gcccgataga	acagcttcgc	66060
cccaggatca	cggagatcgt	cgacaatcta	ctggaccggc	tgagtcccgg	tcaggaggto	66120
gacctcgtcc	ctgtcttcgc	cctgcccattg	ccgaccactg	tgatctgcga	actgctcggc	66180
gtgccgtccg	tcgaccggto	gtcgttcagc	cactgttcca	atgtgctggt	atcgaccgcg	66240
gaagtcggcg	aactggccga	ggccggcgga	gcgatggctg	cctatctggc	acagctcatc	66300
gcggacaaac	gcgcaaaccc	ctgtgacgac	ctgctcacca	agctggtgca	agccaccgac	66360
aacggcgacc	agctctccga	gacggaactc	gtggcgacgg	ccttcctgct	gctgtccgcc	66420
gggcacgaga	ccacggtgaa	cctcattgcc	gcccgtacgc	tcactctgct	ccagaacccg	66480
gaccagctcg	cccggttgog	ctccgacctc	acgctgctgc	ccggcgcgat	cgaggagctc	66540
atacggtagc	acggggccgg	cggcatggtg	ctccggcaca	ccctggagcc	ggtcagggtc	66600
ggcgggtgta	ccatcccggc	ccagcaggtc	gtcctgctct	cgctgtcctc	ggcgggcgcg	66660
gactccacc	ggttcagoga	cgcgcaccgg	ctcgacatcg	gccgtcccat	cgggggcagc	66720
gtgggggttcg	ggcacgggtat	ccaccactgc	atcgcgccgc	cgctcggcag	gctggagggc	66780
gagatcgcgt	tcggggccct	gctcaccgcg	ttcccgcacc	tgccgctcgc	ggtcccgcgc	66840
gaggagctga	actggcgoga	cagtgtcttc	atccgcggcc	cggaatcgct	gcccgtggtg	66900
ctgtgacgcg	catggggaga	ggggaccgac	ccgtcagtg	cgggtccctc	tcccataacc	66960
cggcggtac	gccgacatga	ggatccgacg	ggtgctttcg	actacctcgg	cgatctgcg	67020
atgcccgag	cagctgggcg	aatacgggac	ttccaggggc	agccggccct	ccacggtgat	67080
cgcgcgaccc	accagcggcc	cccggccctg	ctccggcgac	cagttctccg	gtaccgggat	67140
ccaccgcaaa	ccggccagct	cagacccggg	gggtgccacc	aggtccgcga	cacggcccag	67200
gttggtgagc	acgagactgg	ccggccagcag	agcgggattc	tcaaagaagt	ggcgaccgcg	67260
gaggatctcg	cgttcgggat	cgcgccgcgc	gacacccgcg	cgaagccggg	cgtagaccag	67320
cgggcgcagc	gtgcgcacat	ccgcccgcgg	ggagacttcg	acaatgtcgt	agaacgacgc	67380
gggggcagc	accagggtct	cctctgccag	cgggtggagt	acccggtgcc	ggaagtgcac	67440
gggggacgcc	aaggccagcg	aaagggtg	gtcggctcgt	tcgagggcgc	ggcgaccgcg	67500
gatgagcagt	gcccgcgcca	ccaggccctg	taccgagatc	ccggcgggcc	ggcgagatcc	67560
ggcgagccgc	gtcgtctcgt	cgggaagtga	ccggagggtc	cttacgtgta	ttcgcctctg	67620
ctccggcgct	tcacgcccgc	ggtccccag	atagggcagc	agcacggggg	gaaggcgctt	67680
cgcctgctcg	gcccggcggg	ccgcgtaggc	gaggacatcc	gcctccgggt	gatggccgag	67740
acgggtctcg	atgggcgcgc	ggtagctgtc	ggccacgtgt	gccgaggacg	ccatcgccc	67800

ttcgcccagc	goggcgtagc	tcggccacac	cgccgacagc	aaggcgacca	cgctgcggcc	67860
gtcgagatc	cgggtgatcga	catggagtat	gaagggtgtcc	tccgcccgcgc	cgcgagcag	67920
ggtggcgcgg	accagcgggc	cgcagcggtc	cagccggctc	cgcatctccc	ggtcgaggtc	67980
ccaggagccg	gcccggcgga	cgacgagttc	gggcccgcgc	tcgtccagcg	ggtgcagcac	68040
caactcggta	ccgtccggcg	atatccggct	ccgcaaggag	gggtgcgcgg	ccacggacga	68100
ggcgaaggcc	cgttggagaa	gctgttcgtt	ggtatcgccc	cgtaccgtgc	acagcgccat	68160
gattctcgtg	agacctggcc	cggagagcat	gagctctccc	gtggacagct	cacgtcgggt	68220
attcggttgc	atgtttgctt	gcgcctttcc	gttgcccggg	gccgagctac	tggggcgagc	68280
cgacgatgga	ttcccgttcc	cgggaggtga	gagggggccc	ggtgagcggc	ggccgggtag	68340
ggccgagaag	ggtctcgaag	acgatccggg	gcgtcatcag	ccgggtcagc	ggggccgaca	68400
gggtgaacgt	gtccgacacc	goggcggcca	cccgggggcg	gtcggctgcg	gtcttgggtca	68460
cccgggtgac	atagcgccgc	tgcatccgcg	cggccagccc	cggccgcggg	ccactcacat	68520
tcgggtagaa	gatgtcctgt	cctgtggcca	tcgcccacgc	attgttgacg	gcgcccgcga	68580
cggcgccctg	tgtggcgcg	ctogtcccgg	ccaccaaccc	gtcgtctcgc	aggacgtcgc	68640
gcagcgccga	agcgtctcat	gcggccaccg	acatgccgtg	cccgtacacc	gggttcaggg	68700
cggccgcggc	gtcgcggagc	accacgaagc	ccttcggcca	gtcggccagt	tcctcgtagt	68760
agcggcgggc	gttgaccgtg	gtgcgactgc	tgtggatcgg	ccgatcggc	tcggcgttcg	68820
cgatgaggtc	cccgatcacc	gaatgccgca	gtctccgggc	gaacgccacg	aagccttcog	68880
gatcacgcgg	cggctcgcac	ccacgggtcc	cggtcagcgt	gacgatccac	tgtccgtcct	68940
cgatcggcag	cagcaccgcg	ccctggcccg	gctggctcgt	ctccgggtcc	ggcagcacgt	69000
tcacgatggg	aaagccgctc	tccgccccgg	ccggcgcgcg	gtaccggcgg	gtggcgtagg	69060
agagcccgat	gtcgatcttc	acctcgcgca	cggcgggcag	accgagcggc	tggagccagg	69120
tggtcgcggc	ggatccacgg	cgggtggcgt	ccaccacgaa	gtccgcgtcc	agccggagcg	69180
attccccgga	gcggcggtcc	tgggcctgga	cccgggtcac	ccgggtggca	tcgcccgcga	69240
gcccctggac	gtcaacaccg	ctccgcaggg	tgtatccgtc	gtcttcagg	acgaggcgc	69300
gcagcgtcca	gtccagcagc	ggacgaccac	aggtagccat	gaactgggcg	ccgggcaccc	69360
ttcggggcca	cccctgccgt	gagcaggaga	cgagcccgtc	gggtacctcg	gtccggtgcg	69420
caccggcggc	cagcaggcgg	tggaggctcc	cgggcaccag	cgatcgatg	gtccgtgcgc	69480
cgttcgacat	caggatgtgc	gagtggaggg	tcgttggggg	gccctttctg	acttccgggc	69540
cgtccggtag	ctggtcccgg	tccagcatca	ctacctcgtc	cacgaacttc	gcgagcacgg	69600
atgccgtaag	ggcaccagcc	agaccgctgc	cgagaactat	tgcgcgattc	accatatcca	69660
tcgtcctttc	aacgtcgcag	aacacgtaat	taatacgccg	aattcaagcc	gtgattttct	69720
cgacttagtg	cgacggggat	cgcgcctcaa	tactcctggg	ccccctcggg	ctccccgtgc	69780
gtgcagagtg	gatgggttta	tatccgttgc	cgttcgatga	ggtcggcaag	gaaaacaatg	69840
tggtttgccg	gacgctaccg	gtgtttccct	ctccgggggt	atcgccctcg	tcgtgcogtt	69900
gattggagggt	tcccctgaag	attcttttcc	gtaatcgcta	tggcgcgcgc	caggcctgcg	69960
gccggcatgg	tcattgccgc	ggggcggccg	tgcacccagg	tcctgtccgg	gtctcctcat	70020
gccgtctggg	gcagggccag	ctggagctcg	agaggcaact	cctcccggcg	ggtgatgtcg	70080
agcttgccgt	acacgcgggt	cagggtgctgc	tccaccgtgc	tcactgtgat	gaacaacttg	70140
gccgagattt	cacggttggt	cagacccttg	gccgccagtg	cggcaactcg	gcgctcggac	70200
tcgtcagat	tggcgccgac	gtccgctccg	cggcaattcg	gcgccgagga	catgccaccc	70260
tcgggggagg	gacgatggct	ggagccgatc	ggctcggacg	ggacctccgc	actgcaatcg	70320
cctgcgatct	gccgcgccat	atggcggatg	gcgtcggcac	gcctgcccac	gccagttcc	70380
tcgtaggttg	agccagggtc	agcaggagacc	ttggccagtt	gcagccggtc	cccggctctc	70440
tgtaacgcgt	cggcagcctg	gatgagcagt	cgtgtccgtt	cgcccggctc	ggcgaacatc	70500
gcgcggacgc	gcagcactgc	accgtccgtc	gcccgcctta	tgcgggcctg	cgcgtgtcgc	70560
tactcggcca	gcattccgtt	cgcctgtctc	cgtatcgttg	gcccggagcca	ggcatggcg	70620
gagtccacct	gccagggcag	ttctgccgac	ggagccagac	cccagcgctc	cgcgagccgg	70680
ccgatgctga	ggaagtccga	gagcgcgagg	tggggccggg	tgacggccag	ggcgtagtgg	70740
ccgcgggcac	ggagatacgg	gagcccgtag	acactccgga	acagcgcctc	gggaaccggg	70800
cggtcagca	gatgagcgac	gtccttgtaa	cgtcccatct	cgggtgtagac	agtcacagg	70860
acggtcagcg	ggccgccgtg	cagccagggt	ctggacgggt	cggcgaggcc	gtccagggac	70920
atccatgcga	acgtctcggc	ctcggtcagc	ttgccctgtc	tcagcgcgat	atcggtcgc	70980
acggcggcga	acagccgttg	ccagcccggg	ataccccgca	cggtcgcgtt	cttcaggaaa	71040
acgtagcacc	acgtagccgc	cagatccagc	cggccgaccc	gcgtgagcga	gttcagctcg	71100
gtgaggatga	ggctgagcgt	catgtcggac	agcggcggtg	tcgcgagcag	tttctccgag	71160
tcctggatat	ccccgggctt	gcgtcccagc	tccttgatcc	aagtggccag	cgccccagc	71220
gcgtcgtcgc	gggaggtgcc	atcggcgcag	tcctttccgg	acagaccgtc	cgccagagcg	71280
cgggatccgg	cacaccaggc	cgccggcatg	ccggtcatcg	gggggaagaa	ccacagccac	71340
gtgttgccga	cgaccgtcag	gtccgttgta	ctgcgaagac	cgcgagggt	gggcccgtacc	71400
tcocgcagca	gttctcccgc	ttcctccacg	cggcccaggc	tgaccagcag	ttggatcagc	71460
agaacggcat	cgacggggca	gagctcgggt	cccgggtgcc	gctccccgca	gtagccgtcg	71520
agatggtggc	gttcgacgga	acaggggtcg	acgcgccacc	tgacgagtgc	ccgcttcagg	71580
cggatgtggc	cccgtcccca	ggagtcgggt	gaggcgtcat	gggcccagttc	gaggtacgcg	71640

ccggcctgtt	cggcgtcgtc	cgagttccaac	gcctcctcgg	cggcgtgccg	cagggtcccc	71700
acatgccagg	gttccgtcgc	cgagcccgtc	tcgagcagggt	ggcgggcgat	cgtacggctg	71760
ccgacgccgt	gccgactcag	cagttccgcg	gcccgggtggc	gcagttcggc	gcgtgtcttg	71820
gggcccgatga	tggtcagggt	cgcgcgctcg	accaaggggt	gctggaaacg	gtagccgtcg	71880
accagttccgg	ccgaggccaa	cgcaaggata	ccgcgcgcga	tctcggcggc	gttcagcccg	71940
agcagctcct	ccaacagttc	cggcctggag	tcctcgccca	ggaccgcgat	gccgggtggcg	72000
agggacacca	ccgcggggtc	gttgccctgc	acgcagttga	ccgcggcctg	cgcgaagagg	72060
ccgtcggcgg	caggccacgg	ggcggtttgg	cctgcggcgt	tcgggacccg	gtgttcttcc	72120
aacagcgccc	gcacgagcag	ggggttgccg	ccgctcagcc	ggaacacgtc	gtccaggaag	72180
gtgtcctccg	ccggccggcc	ctccaggggcg	ccgaccagggt	cgacgacatg	gtcccgggtc	72240
attgggcgca	gcgcgatccg	gtggagattg	ggctgcgcga	ggagctcgca	gtggaactcc	72300
ggcccagatg	atgtcgggag	tgccctgtacg	acgatcagca	tcagcctgct	ggaccggagc	72360
ctggcccgtg	tggcctccag	cagccagcgc	cagctcaggc	tgctcgagatc	ctgtaggctcg	72420
tcgaggcaga	cgaccaccgg	tgaccggctg	gocaggggcct	cgagtcggcc	gcagaactcc	72480
acgaactcgg	cgttttgtgc	cgaactcatc	gagctcatcc	tgggaacatt	gtcgaatcca	72540
aggtcgcggg	cgttcaccac	gaccgctccg	gaggccttca	catgctcgcc	gaaatttacc	72600
agtaattcgc	ttttcccgca	gtaggcaccg	ccctccaata	ccacggtcac	agccttgccg	72660
atctcgcatt	cgacaagcaa	ggatttcagt	agatcaagtt	ccgagtcacc	cccgaagaga	72720
tgcatccgaa	ttgaatcccc	aatctccacc	acgaaatgag	tgccacgatc	gactccgggt	72780
gcaaatcggg	ccggccggcg	ggttggttgc	catacggttc	gcctcgtctg	tggccgaatc	72840
taagcgctgt	cacgcggcga	ttggggcact	acaccgggca	agtaagcggc	aactcaggca	72900
ggtgacgtgc	cgcccggctc	gactggacgc	ccgggttgca	gaacaccgcg	cgaatgtggc	72960
cggaggactc	cgtttccag	aagacgccct	ggtgtcggcc	gagcccgcgg	ttgagcgctt	73020
ccggccactt	ctggccggcg	cccgaatcgt	cacgaccgaa	accggctccg	accaggatcc	73080
atccgagacc	caggcgcgct	gcacgatcgg	goggagtga	cgtctccgtc	gactccaacc	73140
gcccgcgcga	ccgttcaagg	caccgagttc	gtggggctcg	tcagccgggt	gatcagcacc	73200
tcgccacaaa	agggggacgt	cacgcggacg	gctcgggttac	tcctgctcgg	ccgtcaggac	73260
ggacaggcag	tgccggagtt	ccccggcctc	cgacggccgc	catgggattc	gacgacaagg	73320
agcgcgccatc	cccgcctcgc	gggcccagcg	tgaccgggga	ccagatctgg	ctggagatca	73380
tcagagcgcg	actcgaccgt	gcaccggctc	ctacggcgcg	gcgctgacac	gctacgtgca	73440
caccgaaggc	gtcacggtea	tcgagggtcaa	ccagccggac	taggccaccc	gccgccgacg	73500
cggcaagacc	gacgctcgac	cgggcgcgcg	cgcccaagcg	gtgctgtccg	gccgcgcac	73560
cgccaacgcc	aagaccggcg	gcagagacaa	ccctgaacgg	atggtcagcg	aggcatcctt	73620
cggccgactc	ggcggcgtea	gcccggtgga	gtcatcctca	gccaggaccc	aacgccgcag	73680
gttcaaacgc	ggcggcgacc	gccaagccaa	cgggtgcgacg	tcacatgtcg	ggaacacccc	73740
gcgctcagcc	acgtactggc	ggagggtcag	gggcggaagg	cactgggtct	ggctcccgga	73800
aacgatgtga	gcgtgccccg	ggtcagggtc	tggatccggg	cacctgcgtt	gtcacctgcg	73860
ggagttgtcg	ggcactccca	tcgcggcgaa	caacgcgtcc	gccgcaccac	gtacatccga	73920
tgctcgcatac	gcgccgagag	cctgagcggc	ctgggccata	cacgccacgc	actcgactgg	73980
gtgagctgca	aggcgcgctc	cgcgtcgtcc	gtggcctccg	tcggacggcc	gagacgggtga	74040
aggacgtcga	ccgaacgtcg	cgaggggccag	cgcgacgttg	ccccggatcc	cgcgtctcgt	74100
cgcacaactg	ccgggcggta	acccgctctt	ggaccaccgc	atggtagcgc	cttgccgggc	74160
tcagggggag	gagtggctgg	cacgcgctct	tggattgccc	cgcgatgctg	attctgggtg	74220
ctcgacccaa	ccttcgagcg	gatccaggac	gttgaacctc	agctgctcct	tcgtgtggtc	74280
cgtgtcaag	cccctctgca	cgtgacgtgc	atgaaatctt	ccgatgcagg	gtgcgtttga	74340
ac						74342

<210> 2
 <211> 3793
 <212> DNA
 <213> Streptomyces sp.

<220>
 <221> CDS
 <222> (1322).. (2548)

<220>
 <221> CDS
 <222> (2564).. (2761)

<400> 2	ctgcagctcg	acgtgogggt	cggacttcac	gttgaagtac	cagaccggat	gcttggggcg	60
	accgcccagc	gaggcgaccg	ccgcgtaact	cccctcgtgc	tcgacccgca	tcagcggcgt	120

cttgcggatc	tttccgctgc	gcgcgccccg	ggtggtgagc	acgatgaccg	gcagccccgt	180
gtcccgcagc	gtggtgccct	tgggtccccc	ggaactctcg	tacagctcga	cctgctcgcg	240
caccactgc	gtcgggctgg	gctcgtactc	gccctcaagt	ggcaagggat	ccgtctcctt	300
cgctcggtcg	gcggatggtg	ctcggacagg	tcccaactcc	cgcggccgcc	cggatcatcc	360
gtaccgcatg	ccccttcgcc	cgagcgggtg	atcaccggtc	cggccatccg	gtcgtccgca	420
ccgcgagcac	caggatcacg	gcgctggaga	gcaggggcgt	gaccagccgc	ccccggtggc	480
ccgtcagggc	gcgacccagc	agcgcgcccc	cgcccgccag	cagtagctgc	cagctcgcgg	540
acgcggcgaa	ggcgcgcccg	gcgaacaccg	cccgtttcag	cggccgtgcc	gcaccggcgg	600
cgccgctgcc	gagcaccagc	gccacgaagt	agaccaccgt	catgggattg	agcagggtga	660
tcccagagaag	gccgagataa	gcccctgccg	cgcttggaac	cggccgttcc	gggcggttg	720
tgagccgatg	ggcgcggtac	tgccgcaggg	cgagcagcgc	cgcccgagc	gcgagcaccg	780
cgaggaccag	cgccgaggcc	cagcgcagcg	ggtccagcac	cggccgcagc	tgtgccgcga	840
gggcgggccc	gcccacggtc	gcgagcagcg	cgtacagccc	gtcggccgtg	gcgacgccga	900
gcgcgcgcga	ggcgcggtg	cgagcagagg	tgcggcggtg	gagggagacc	agataggtcc	960
cgaccgcgcc	gacgggcacc	gcgatgccgt	accggcgag	caggcccgog	agcagcgccg	1020
ccgtcacggg	cgtgcgggac	tggttcctcc	ggggacggcg	gggctgctgt	cggcccggca	1080
ccgcgggggc	ggtggcagcg	ggcgtcggca	ggaggaggcg	tgtaggaggc	atgggcccgt	1140
cctggggccg	ccgcgcccgc	accggcaaat	gaattacggc	gcgttccagc	ccccggccgg	1200
ctcgtcttcc	ggccacttca	ccgcgtacgg	cgtctggcc	gaacttctg	tgcgcccata	1260
ggtgcctcgg	gcattctaag	aagatcggca	cgaacacact	cttcgtctgc	gaggtcttcc	1320
c atg acg gaa ctg acg gac atc acc ggc ccg ggg acc ccg gcc gaa						1366
Met Thr Glu Leu Thr Asp Ile Thr Gly Pro Gly Thr Pro Ala Glu						
1 5 10 15						
ccc gtc gca ttc ccc cag gac cgc acc tgc ccc tac cac ccc ccc acc						1414
Pro Val Ala Phe Pro Gln Asp Arg Thr Cys Pro Tyr His Pro Pro Thr						
20 25 30						
gga tac ggc ccg ctg cgc gac ggg cgc agc ctg tcc cgc gtc acc ctc						1462
Gly Tyr Gly Pro Leu Arg Asp Gly Arg Ser Leu Ser Arg Val Thr Leu						
35 40 45						
ttc gac ggc cgt gag gtc tgg atg gtc acg ggc cac gcc acc gcc cgc						1510
Phe Asp Gly Arg Glu Val Trp Met Val Thr Gly His Ala Thr Ala Arg						
50 55 60						
gcg ctg ctc ggc gac ccc cgg ctg tcc acc gac cgc acc ctc ccg ggc						1558
Ala Leu Leu Ala Asp Pro Arg Leu Ser Thr Asp Arg Thr Leu Pro Gly						
65 70 75						
ttc ccc gtg tcc acg gcc cgc ttc gcg gcc gtc cgc gac cgg cgg gtg						1606
Phe Pro Val Ser Thr Ala Arg Phe Ala Ala Val Arg Asp Arg Arg Val						
80 85 90 95						
gcg ctg ctc ggc gtg gac gac ccg gtc cac cag acc cag cgg cgg atg						1654
Ala Leu Leu Gly Val Asp Asp Pro Val His Gln Thr Gln Arg Arg Met						
100 105 110						
atg atc ccg tcc ttc acc ctc aag cgc gcg gcc ggg ctg cgg ccc acc						1702
Met Ile Pro Ser Phe Thr Leu Lys Arg Ala Ala Gly Leu Arg Pro Thr						
115 120 125						
atc cag cgg acc gtc gac ggg ctg gac gcg atg atc gag aag ggg						1750
Ile Gln Arg Thr Val Asp Gly Leu Leu Asp Ala Met Ile Glu Lys Gly						
130 135 140						
ccg ccg gcc gag ctg gtc tcc gcc ttc gcc ctg ccc gtg ccc tcc atg						1798
Pro Pro Ala Glu Leu Val Ser Ala Phe Ala Leu Pro Val Pro Ser Met						
145 150 155						
gtc atc tgc ggc ctg ctc ggc gtg ccg tac gcc gac cac gag ttc ttc						1846
Val Ile Cys Gly Leu Leu Gly Val Pro Tyr Ala Asp His Glu Phe Phe						
160 165 170 175						
gag gaa cag tcc cgc acc ctg ctg cgc ggt ccc acc gcc gcc gac tcc						1894
Glu Glu Gln Ser Arg Thr Leu Leu Arg Gly Pro Thr Ala Ala Asp Ser						
180 185 190						
caa ggg gcg cgc gag cgg ctc gag gag tac ctc gcc ggg ctg atc gac						1942
Gln Gly Ala Arg Glu Arg Leu Glu Glu Tyr Leu Gly Gly Leu Ile Asp						
195 200 205						
gac aag gag cgg cag gcc gaa ccc ggc gac ggc gtc ctg gac gac ctc						1990
Asp Lys Glu Arg Gln Ala Glu Pro Gly Asp Gly Val Leu Asp Asp Leu						
210 215 220						
gtc cac cag cgg ctg cgc acc ggc gag ctg gac cga cgc gac gtg gtg						2038
Val His Gln Arg Leu Arg Thr Gly Glu Leu Asp Arg Arg Asp Val Val						

225	230	235	
gcg ctg gcc gtc atc ctg ctc gtg gcc ggg cac gag acg acc gcc aac	ctg ctc gtg gcc ggg cac gag acg acc gcc aac	gag acg acc gcc aac	2086
Ala Leu Ala Val Ile Leu Leu Val Ala Gly His Glu Thr Thr Ala Asn	Leu Leu Val Ala Gly His Glu Thr Thr Ala Asn	Glu Thr Thr Ala Asn	
240	245	250	255
atg atc tcc ctc ggc acc tac acg ctg ctg cgg cac ccc ggc cgg ctg	atg atc tcc ctc ggc acc tac acg ctg ctg cgg cac ccc ggc cgg ctg	ctg ctg cgg cac ccc ggc cgg ctg	2134
Met Ile Ser Leu Gly Thr Tyr Thr Leu Leu Arg His Pro Gly Arg Leu	Met Ile Ser Leu Gly Thr Tyr Thr Leu Leu Arg His Pro Gly Arg Leu	Gly Arg Leu	
260	265	270	
gcc gag ctg cgc gcc gac cgg ggc ctg ctg ccc gcc gcc gtg gag gag	gcc gag ctg cgc gcc gac cgg ggc ctg ctg ccc gcc gcc gtg gag gag	gtg gag gag	2182
Ala Glu Leu Arg Ala Asp Pro Ala Leu Leu Pro Ala Ala Val Glu Glu	Ala Glu Leu Arg Ala Asp Pro Ala Leu Leu Pro Ala Ala Val Glu Glu	Val Glu Glu	
275	280	285	
ctg atg cgg atg ctc tcg atc ggc gac ggg ctg ctg cgc ctg gcc ctg	ctg atg cgg atg ctc tcg atc ggc gac ggg ctg ctg cgc ctg gcc ctg	ctg ctg gcc ctg	2230
Leu Met Arg Met Leu Ser Ile Ala Asp Gly Leu Leu Arg Leu Ala Leu	Leu Met Arg Met Leu Ser Ile Ala Asp Gly Leu Leu Arg Leu Ala Leu	Leu Ala Leu	
290	295	300	
gag gac atc gag atc gcc ggc gcc acg atc cgg gcc gag ggc gtc	gag gac atc gag atc gcc ggc gcc acg atc cgg gcc gag ggc gtc	gag ggc gtc	2278
Glu Asp Ile Glu Ile Ala Gly Ala Thr Ile Arg Ala Gly Glu Gly Val	Glu Asp Ile Glu Ile Ala Gly Ala Thr Ile Arg Ala Gly Glu Gly Val	Glu Gly Val	
305	310	315	
ctg ttc tcc acc tcg ctg atc aac cgc gac gag tcc gtg ttc gac gac	ctg ttc tcc acc tcg ctg atc aac cgc gac gag tcc gtg ttc gac gac	gac gac	2326
Leu Phe Ser Thr Ser Leu Ile Asn Arg Asp Glu Ser Val Phe Asp Asp	Leu Phe Ser Thr Ser Leu Ile Asn Arg Asp Glu Ser Val Phe Asp Asp	Asp Asp	
320	325	330	335
ccc gac acc ctg gac ttc cac cgc tcc acc cgc cac cac gtg gcc ttc	ccc gac acc ctg gac ttc cac cgc tcc acc cgc cac cac gtg gcc ttc	gtg gcc ttc	2374
Pro Asp Thr Leu Asp Phe His Arg Ser Thr Arg His His Val Ala Phe	Pro Asp Thr Leu Asp Phe His Arg Ser Thr Arg His His Val Ala Phe	Val Ala Phe	
340	345	350	
ggt ttc ggc atc cac cag tgc ctg ggc cag aac cta gcc cgc gcc gag	ggt ttc ggc atc cac cag tgc ctg ggc cag aac cta gcc cgc gcc gag	gcc gcc gag	2422
Gly Phe Gly Ile His Gln Cys Leu Gly Gln Asn Leu Ala Arg Ala Glu	Gly Phe Gly Ile His Gln Cys Leu Gly Gln Asn Leu Ala Arg Ala Glu	Ala Glu	
355	360	365	
ctg gag atc gcc ctg ggc acg ctc ctg gag cgg ctc ccc ggc ctc cgg	ctg gag atc gcc ctg ggc acg ctc ctg gag cgg ctc ccc ggc ctc cgg	ctc cgg	2470
Leu Glu Ile Ala Leu Gly Thr Leu Leu Glu Arg Leu Pro Gly Leu Arg	Leu Glu Ile Ala Leu Gly Thr Leu Leu Glu Arg Leu Pro Gly Leu Arg	Leu Arg	
370	375	380	
ctg gcc ggc ccc gcc gag gag atc cgg ttc aaa ccc ggc gac acg atc	ctg gcc ggc ccc gcc gag gag atc cgg ttc aaa ccc ggc gac acg atc	gac acg atc	2518
Leu Ala Ala Pro Ala Glu Glu Ile Pro Phe Lys Pro Gly Asp Thr Ile	Leu Ala Ala Pro Ala Glu Glu Ile Pro Phe Lys Pro Gly Asp Thr Ile	Thr Ile	
385	390	395	
cag ggg atg ctg gaa ctc ccc gtg acc tgg taa gagggctctggt tc atg cac	cag ggg atg ctg gaa ctc ccc gtg acc tgg taa gagggctctggt tc atg cac	atg cac	2569
Gln Gly Met Leu Glu Leu Pro Val Thr Trp	Gln Gly Met Leu Glu Leu Pro Val Thr Trp	Met His	
400	405	410	
atc gac atc gac aag gac cgc tgc atc ggc gcc ggc cag tgc ggc ctg	atc gac atc gac aag gac cgc tgc atc ggc gcc ggc cag tgc ggc ctg	ctg	2617
Ile Asp Ile Asp Lys Asp Arg Cys Ile Gly Ala Gly Gln Cys Ala Leu	Ile Asp Ile Asp Lys Asp Arg Cys Ile Gly Ala Gly Gln Cys Ala Leu	Ala Leu	
415	420	425	
gcc gcc cgc ggc gtg ttc acc cag gac gac gac ggc tac agc acc ctg	gcc gcc cgc ggc gtg ttc acc cag gac gac gac ggc tac agc acc ctg	acc ctg	2665
Ala Ala Pro Gly Val Phe Thr Gln Asp Asp Asp Gly Tyr Ser Thr Leu	Ala Ala Pro Gly Val Phe Thr Gln Asp Asp Asp Gly Tyr Ser Thr Leu	Leu	
430	435	440	
ctc ccc ggc cgc gag gac ggc ggg ggc gac cgc atg gtc cgg gag ggc	ctc ccc ggc cgc gag gac ggc ggg ggc gac cgc atg gtc cgg gag ggc	gag ggc	2713
Leu Pro Gly Arg Glu Asp Gly Gly Gly Asp Pro Met Val Arg Glu Ala	Leu Pro Gly Arg Glu Asp Gly Gly Gly Asp Pro Met Val Arg Glu Ala	Ala	
445	450	455	
gcc cgc gcc tgc cgc gtg agc gcc atc cgg gtg acc gaa cgc gcc ggc	gcc cgc gcc tgc cgc gtg agc gcc atc cgg gtg acc gaa cgc gcc ggc	gcc ggc	2761
Ala Arg Ala Cys Pro Val Ser Ala Ile Arg Val Thr Glu Pro Ala Gly	Ala Arg Ala Cys Pro Val Ser Ala Ile Arg Val Thr Glu Pro Ala Gly	Gly	
460	465	470	475
tga ggccggggccc ggccggccgc gcccgctgcc gggaccgccg ttcccagttc agtagg	tga ggccggggccc ggccggccgc gcccgctgcc gggaccgccg ttcccagttc agtagg	agtagg	2820

gtcgtgcgat gacctcacag gccgggaagc ccttcctcta cgtcgtcgtc tgcgcggccg	2880
ggaccgccgc cggagtcacc acgctgatcg ggcgcggcca ggcgcggggc tgggaggtgg	2940
gggtcctggc cagcccggtg cgcgtggcg ggttccttga caccgctgcg gtcgaggaga	3000
tgacggggcg gccatccgc tcggcctggc gctcgcggc cgatccgcgc ccgttcccgc	3060
cgcggggcg cgtggtggtg gcgccgccca ccttcaacac cgtcaacaag tggcgggccg	3120
gtctcgccga cagctcgcc gtcggcacgc tctgcgaggc ggccggccctc ggctgcga	3180
tcgcgctcct gccctgcgtg gcggacgcgc tggccgccca cccgcgtac cgggagggcc	3240
ttctccggct gcgtgggatg ggctgcgtc tcggcgagcc gtacgccggc ccgccggggg	3300
aggacggcga ggccggacggc gcacggcccg ggttcgcctg ggagaacgcc ctggacctgc	3360
tggagcgggc ctgaaccgc tccccgaccc gtagggcctg tctgacactg tcagacaggc	3420
cctaaccggca ggtcagcgcc ggcccgcca gcatgccgc ggtgtagagg tcctggcccc	3480
gcggcagcca gtagcccgcc ctggagacca ccgtggagca gtcaggcccg acggtgacgc	3540
ggaccttcac cgtctcgga cggccgggct gcagcgcggt cagcgcgcag tccaggaggt	3600
acgcgagccg ggtcttcgag gtaccggccg accagcgggt tgacgcagcg ggcgtcgtcc	3660

gtggcgatcc	gcaccccggt	gcccgcgccg	ccgatgagtc	cgagccgggc	gctgccgtcg	3720
ccctcgtcgc	tgtcccgccg	gaccgtgtag	gtcagcgtgg	tggtggcgtc	gcgccgcagg	3780
gtgtccggtc	gac					3793

<210> 3
 <211> 2329
 <212> DNA
 <213> Streptomyces sp.

<220>
 <221> CDS
 <222> (420).. (1604)

<220>
 <221> CDS
 <222> (1643).. (1834)

<400> 3

ggatccacgg	gtggccgccc	cgctcgcccc	ggtgaccgac	cggcgtatcg	gctatgtcgc	60
cgcgctcttc	gcgccgctgg	gcttccccga	ggcgaggcgc	cgggaccgcg	gcctgctggc	120
gtacaccgcc	tacctcgcc	acaccagct	cggacatgcc	gtccgacaga	gcctgccggc	180
cgaggcgcca	cacgaccgct	atctggatgg	cgtgatcgac	accctcgtac	ggccgcggga	240
cggaggcgat	gaagccgaac	atgtcacaat	ctgaacgagg	ttggcggaac	tgccgcgaga	300
acatgcccg	tatccgcggc	atgaggtgag	atcgccgcgg	cgaaacacgg	tgccgccacag	360
cgttgccatc	tcacacacga	gcaactcgag	ccacttgaga	ctcgtacggg	aggaaattc	419
gtg acc gaa gcc atc ccc tac ttt cag aac cgc acc tgt ccc tac cac	467					
Val Thr Glu Ala Ile Pro Tyr Phe Gln Asn Arg Thr Cys Pro Tyr His						
1 5 10 15						
cgg ccc gcc gcc tat cag cca ctg cgc ggg gcc ggc cgg ctg agc cat	515					
Pro Pro Ala Ala Tyr Gln Pro Leu Arg Gly Ala Gly Pro Leu Ser His						
20 25 30						
gtc acg ttc tac gac ggc cgg aag gtg tgg gcg gtc acc ggc cac ccc	563					
Val Thr Phe Tyr Asp Gly Arg Lys Val Trp Ala Val Thr Gly His Pro						
35 40 45						
gag gca cgg gcg ctg ctg acc gac cag cga ctc tcc gcc gac cgg cag	611					
Glu Ala Arg Ala Leu Leu Thr Asp Gln Arg Leu Ser Ala Asp Arg Gln						
50 55 60						
aac ccg gcc ttc ccg gtc ccc ttc gaa cgc ttc gcg gcc atc cgc cgg	659					
Asn Pro Ala Phe Pro Val Pro Phe Glu Arg Phe Ala Ala Ile Arg Arg						
65 70 75 80						
gtc cgg acg ccg ctg atc ggg gtc gac gac ccg gag cac aac acc cag	707					
Val Arg Thr Pro Leu Ile Gly Val Asp Asp Pro Glu His Asn Thr Gln						
85 90 95						
cgc cgg atg ctg atc ccc agc ttc agc ctc aag cgg acc gcc gca ctg	755					
Arg Arg Met Leu Ile Pro Ser Phe Ser Leu Lys Arg Thr Ala Ala Leu						
100 105 110						
cgg ccg gag atc cag cgg atc gtc gac ggg ctg ctc gac cgg atg ctg	803					
Arg Pro Glu Ile Gln Arg Ile Val Asp Gly Leu Leu Asp Arg Met Leu						
115 120 125						
gat cag ggc ccg ccc acc gag ctg gtc tcc gcg ttc gcc ctg ccg gtc	851					
Asp Gln Gly Pro Pro Thr Glu Leu Val Ser Ala Phe Ala Leu Pro Val						
130 135 140						
ccg tcg atg gtg atc tgc gca ctg ctc gga gtc tca tac gcc gac cat	899					
Pro Ser Met Val Ile Cys Ala Leu Leu Gly Val Ser Tyr Ala Asp His						
145 150 155 160						
gag ttc ttc gag gag gag tcc cgc cgc atc ctg cgc ggc cgg tcg gcc	947					
Glu Phe Phe Glu Glu Glu Ser Arg Arg Ile Leu Arg Gly Arg Ser Ala						
165 170 175						
gag gag gcg gag gac gcc cgg ctg aag ctg gag gag tac ttc acc ggg	995					
Glu Glu Ala Glu Asp Ala Arg Leu Lys Leu Glu Glu Tyr Phe Thr Gly						
180 185 190						
ctg atc gcc gcc aag gag aag aac ccg ggc gac ggg ctg ctg gac gag	1043					

Leu Ile Ala Ala Lys Glu Lys Asn Pro Gly Asp Gly Leu Leu Asp Glu	
195 200 205	
ctg atc gag gac cgg ctg cgg acc ggc ggc ctc acc cgc gac gag ctg	1091
Leu Ile Glu Asp Arg Leu Arg Thr Gly Ala Leu Thr Arg Asp Glu Leu	
210 215 220	
gtc cgg ctc gcc atg atc ctg ctg gtg gcc ggc cat gag acc acc gcc	1139
Val Arg Leu Ala Met Ile Leu Leu Val Ala Gly His Glu Thr Thr Ala	
225 230 235 240	
aac atg atc tcg ctc ggc acc ttc acc ctg ctg gac cac ccc gag cag	1187
Asn Met Ile Ser Leu Gly Thr Phe Thr Leu Leu Asp His Pro Glu Gln	
245 250 255	
ctg ggc cag ctc aag gcc gac gag ggc ctg atg ccg gcc gcc atc gag	1235
Leu Ala Gln Leu Lys Ala Asp Glu Gly Leu Met Pro Ala Ala Ile Glu	
260 265 270	
gag ctg ctg cga ttc ctg tcc atc ggc gac ggc ctg ctg cgg gtg ggc	1283
Glu Leu Leu Arg Phe Leu Ser Ile Ala Asp Gly Leu Leu Arg Val Ala	
275 280 285	
acg gag gac atc gag atc ggc ggt cag gtg atc cgg gcc gac gac ggc	1331
Thr Glu Asp Ile Glu Ile Gly Gly Gln Val Ile Arg Ala Asp Asp Ala	
290 295 300	
gtc ctg ttc ccc gcc tca ctg atc aac cgg gac gag gcc gcc tat ccg	1379
Val Leu Phe Pro Ala Ser Leu Ile Asn Arg Asp Glu Ala Ala Tyr Pro	
305 310 315 320	
gca ccc gac gag ctg gac ctc ggc cgt tcg gcc cgc cat cac gtg ggc	1427
Ala Pro Asp Glu Leu Asp Leu Gly Arg Ser Ala Arg His His Val Ala	
325 330 335	
tcc ggc ttc ggg atc cac cag tgc ctg ggg cag aac ctc gcc cgc ggc	1475
Ser Gly Phe Gly Ile His Gln Cys Leu Gly Gln Asn Leu Ala Arg Ala	
340 345 350	
gag atg gag atc ggc ctg cgc tca ctg ttc acc agg atc ccg cag ctg	1523
Glu Met Glu Ile Ala Leu Arg Ser Leu Phe Thr Arg Ile Pro Gln Leu	
355 360 365	
cgg ctc gcc gtg ccg gcc gcc gag att ccg ttc aag gac gga gac acc	1571
Arg Leu Ala Val Pro Ala Ala Glu Ile Pro Phe Lys Asp Gly Asp Thr	
370 375 380	
ctg caa ggc atg atc gaa ctg ccg ctg gcc tgg tag cagccaggac ggcaga	1623
Leu Gln Gly Met Ile Glu Leu Pro Leu Ala Trp	
385 390 395	
ccaaagaaag gggtccgga atg cgg atc ggc atc gac acc gac cgc tgt atc	1675
Met Arg Ile Ala Ile Asp Thr Asp Arg Cys Ile	
400 405	
ggc gcc ggc cag tgt gcc ctg acc ggc ccc ggg ggt ttc acc cag gat	1723
Gly Ala Gly Gln Cys Ala Leu Thr Ala Pro Gly Gly Phe Thr Gln Asp	
410 415 420	
gac gac ggt ttc agt gca ctg ctg ccc ggc cgg gag gac ggc gcc ggc	1771
Asp Asp Gly Phe Ser Ala Leu Leu Pro Gly Arg Glu Asp Gly Ala Gly	
425 430 435	
gac ccg ctg gtg cgg gaa gcc gcc cgc gcc tgc ccc gtg cag gcc att	1819
Asp Pro Leu Val Arg Glu Ala Ala Arg Ala Cys Pro Val Gln Ala Ile	
440 445 450	
gcg gtc acc gac gat tag cagcaccgcc ggcgacgacc cggcagacgc gcgcggcc	1875
Ala Val Thr Asp Asp	
455	
ccggctgaca cccggcgccc gaggcgcgcc cgagccgtcc gccctccac ttgtccctac	1935
ggcatccacc ccattcgcta ccgcaacacc ccttgggtga cgggcagttt cgaggacccc	1995
ggtgtgcccg gggcgacttg gtgaccgtca ccggcttcac gccgcgattg cccacatagg	2055
cgctgctgct cgcggcgatc acgaagcgcg gtcgggtgcc cggctcgtaa cgggtgcacga	2115
tgcccggcag ttccacggtg aaccgccggg ccacatcggg caccgcggcc ggggccacca	2175
acagggtgcac cagcgtttc ctgccgttcg gcgcgacatc gtagagcttg gcgaacagca	2235
ccagcttgct cgcgcacatc gcggaccgct gcgcccgcgc ggctgcggc gaggcaacct	2295
tcagcgtcac cctcggcgcg cccaccacgt cgac	2329

<211> 1860
 <212> DNA
 <213> A-1560 strain

<220>
 <221> CDS
 <222> (172).. (1383)

<220>
 <221> CDS
 <222> (1399).. (1593)

<400> 4

cggggatcgt acgccgtacc gtttcggggc aaccgaatta cgatgcggaa tggatggttc	60
ccagccagat cccgcaggta gccgatctgg ccgaacttga tgtcgtgcac tggatgcctc	120
gggcatctaa tgaagatcgg cacgacgcat cttcgtctg cgaggtctcc c atg aca	177
Met Thr	
1	
gac acg aca gac ctg acc gag ctg tca gat ccc gtc tcc ttc ccc cag	225
Asp Thr Thr 5 Asp Leu Thr Glu Leu Ser Asp Pro Val Ser Phe Pro Gln	
10	
gac cgg agc tgc ccc tac cac ccg ccc acc ggg tac gac ccg ctg cgc	273
Asp Arg Ser Cys Pro Tyr 25 Pro Pro Thr Gly Tyr 30 Asp Pro Leu Arg	
35	
acc gaa cgg ccg ccc gcc cgc atc ccg ctc tac gac ggc cgc ccc gcc	321
Thr Glu Arg Pro Pro Ala Arg Ile Arg Leu Tyr Asp Gly Arg Pro Ala	
40	
tgg ctc gtc acc ggc cac gcc gtc gcc cgt gac ctg ctg gtc gac ccc	369
Trp Leu Val Thr Gly His Ala Val Ala Arg Asp Leu Leu Val Asp Pro	
55	
cgc ctg tcc acg gac cgc acc cgc tcg ggc ttc ccg gcc aca act ccc	417
Arg Leu Ser Thr 70 Asp Arg Thr Arg Ser Gly Phe Pro Ala Thr Thr Pro	
75	
cgc ttc gcc gcg gtc cgc gac cgc aag ccg gcg ctc ctc ggc gtc gac	465
Arg Phe Ala Ala Val Arg Asp Arg Lys Pro Ala Leu Leu Gly Val Asp	
85	
gac ccc aag cac cgc acc cag ccg tgg atg atg atc ccg agc ttc acc	513
Asp Pro Lys His Arg Thr Gln Arg Trp Met Met Ile Pro Ser Phe Thr	
100	
ctc agg cgc gcc acc gag ctc agg ccg cgc atc cag gag atc gtc gac	561
Leu Arg Arg Ala Thr 120 Glu Leu Arg Pro Arg Ile Gln Glu Ile Val Asp	
115	
gaa ctg ctg gac gtg atg atc gcc cag gga ccc ccg gcc gac ctg gtg	609
Glu Leu Leu Asp Val Met Ile Ala Gln Gly Pro Pro Ala Asp Leu Val	
135	
cgt tcc ttc gcg ctg ccg gtg ccg tcc atg gtg atc tgc gcc ctg ctc	657
Arg Ser Phe Ala Leu Pro Val Pro Ser Met Val Ile Cys Ala Leu Leu	
150	
ggc gtg ccc tac gcc gac cac gag ttc ttc gag gac cag tcc agg ccg	705
Gly Val Pro Tyr Ala Asp His Glu Phe Phe Glu Asp Gln Ser Arg Arg	
165	
ctg ctg cgc gga ccg gcg gcc gag gac acg cag gac gcc ccg gac ccg	753
Leu Leu Arg Gly Pro Ala Ala Glu Asp Thr Gln Asp Ala Arg Asp Arg	
180	
ctc gcc gcg tac ctg gag gac ctg atc gac gag aag ccg cgc ccg ccc	801
Leu Ala Ala Tyr Leu Glu Asp Leu Ile Asp Glu Lys Arg Arg Arg Pro	
195	
ggt gac ggc ctg ctg gac gaa ctc gtc cag cag cgt ctg aac gaa ggc	849
Gly Asp Gly Leu Leu Asp Glu Leu Val Gln Gln Arg Leu Asn Gln Gly	
215	
gag ctc gac ccg gag gaa ctg acc gcg ctg gcg atg atc ctg ctg gtc	897
Glu Leu Asp Arg Glu Glu Leu Thr Ala Leu Ala Met Ile Leu Leu Val	
230	
gcg ggc cac gag acc acc gcc aac atg atc tcc ctg ggc acc tac acg	945
235	
240	

Ala Gly His Glu Thr Thr Ala Asn Met Ile Ser Leu Gly Thr Tyr Thr	
245 250 255	
ctc ctg ctg cac ccc gaa cgg ctg acc gag ctg cgc gcc gac ccc gcg	993
Leu Leu Leu His Pro Glu Arg Leu Thr Glu Leu Arg Ala Asp Pro Ala	
260 265 270	
ctg ctg cgc gcc gcc gtc gag gaa ctg atg cgg atg ctg tcc atc gcg	1041
Leu Leu Pro Ala Ala Val Glu Glu Leu Met Arg Met Leu Ser Ile Ala	
275 280 285	
gac gga ctg ctg cgg cag gcc acc gag gac atc gag atc gcc ggg acc	1089
Asp Gly Leu Leu Arg Gln Ala Thr Glu Asp Ile Glu Ile Ala Gly Thr	
295 300 305	
acc atc agg gcc ggg gac ggc gtg gtc tcc acc tct gtc atc aac	1137
Thr Ile Arg Ala Gly Asp Gly Val Phe Ser Thr Ser Val Ile Asn	
310 315 320	
cgc gac gag gac gtc tac cgc gcc ccc gac acc ctg gac ttc cac cgc	1185
Arg Asp Glu Asp Val Tyr Pro Ala Pro Asp Thr Leu Asp Phe His Arg	
325 330 335	
tgc acc cgc cac cac gtc gcc ttc ggt ttc gga atc cac cag tgc ctc	1233
Ser Thr Arg His His Val Ala Phe Gly Phe Gly Ile His Gln Cys Leu	
340 345 350	
ggc cag aac ctg gcc cgc acc gaa ctg gag atc gcc ctg cgc acg ctc	1281
Gly Gln Asn Leu Ala Arg Thr Glu Leu Glu Ile Ala Leu Arg Thr Leu	
355 360 365	
ctc gaa cgc ctg ccc acg ctg cgg ctg gcc cca ccg gag gaa atc	1329
Leu Glu Arg Leu Pro Thr Leu Arg Leu Ala Pro Pro Glu Glu Ile	
375 380 385	
ccc ttc aaa ccc ggc gac acc atc cag ggg atg ctg gaa ctg ccc gtc	1377
Pro Phe Lys Pro Gly Asp Thr Ile Gln Gly Met Leu Glu Leu Pro Val	
390 395 400	
agc tgg taa gaggtgccg tc atg cat atc gag atc gac aag gac cgc tgc	1428
Ser Trp Met His Ile Glu Ile Asp Lys Asp Arg Cys	
405 410	
atc ggc gcc gga cag tgc gcc ctg acc gcc ccg ggt gtg ttc acc cag	1476
Ile Gly Ala Gly Gln Cys Ala Leu Thr Ala Pro Gly Val Phe Thr Gln	
420 425 430	
gac gac gac ggc ttc agt gac ctg ttg ccc ggc cgg gag gac ggc gcc	1524
Asp Asp Asp Gly Phe Ser Asp Leu Leu Pro Gly Arg Glu Asp Gly Ala	
435 440 445	
ggc gac ccg atg gtc cgg gag gcc gcc agg gcc tgc ccc gtg agt gcc	1572
Gly Asp Pro Met Val Arg Glu Ala Ala Arg Ala Cys Pro Val Ser Ala	
450 455 460	
atc acg ctg tcc gag gac ggg tag ggggccgagc cgcgccgcc gccggtccgc	1626
Ile Thr Leu Ser Glu Asp Gly	
465	
tgccgaggcg ccgtgccgac gcggcgcccg gccggcccg cgggtgcccg tcgctggcc	1686
ccgtggcccc ggccggcggt gattgactag ggttcccggt tgagcgaaca ggcccagaag	1746
ccctccgggg ccgccccgc gaaagacacc gggacggcgc ccgggaaacc ccttcctcta	1806
cgctgctgctc tgcgccgcgc gcatcgccga aggcgtcagc aagctgatca ccgc	1860

<210> 5
 <211> 23
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> STRANDNESS : single

<220>
 <223> TOPOLOGY : linear

<220>
 <223> Description of Artificial Sequence : KS-3F Primer

<400> 5
 gaccgcggct gggacgtgga ggg 23

<210> 6
 <211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> STRANDNESS : single

<220>
 <223> TOPOLOGY : linear

<220>
 <223> Description of Artificial Sequence : KS-4R Primer

<400> 6
 gtgcccgatg ttggacttca acga 24

<210> 7
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> STRANDNESS : single

<220>
 <223> TOPOLOGY : linear

<220>
 <223> Description of Artificial Sequence : CB-1F Primer

<400> 7
 atgacagctt tgaatctgat ggatccc 27

<210> 8
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> STRANDNESS : single

<220>
 <223> TOPOLOGY : linear

<220>
 <223> Description of Artificial Sequence : CB-2R Primer

<400> 8
 tcagagacgg accggcagac tcttcagacg 30

<210> 9
 <211> 27
 <212> DNA
 <213> Artificial Sequence

<220>

<223> STRANDNESS : single
 <220>
 <223> TOPOLOGY : linear
 <220>
 <223> Description of Artificial Sequence : PKC-1F Primer
 <400> 9
 gtgcgccgta cccagcaggg aacgacc 27

 <210> 10
 <211> 27
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> STRANDNESS : single
 <220>
 <223> TOPOLOGY : linear
 <220>
 <223> Description of Artificial Sequence : PKC-2R Primer
 <400> 10
 tcacgcgctc tccgcccgcc ccctgcc 27

 <210> 11
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> STRANDNESS : single
 <220>
 <223> TOPOLOGY : linear
 <220>
 <223> Description of Artificial Sequence : PDL58-1F Primer
 <400> 11
 gccccgcata tggatctgga aacccaactt ctc 33

 <210> 12
 <211> 31
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> STRANDNESS : single
 <220>
 <223> TOPOLOGY : linear
 <220>
 <223> Description of Artificial Sequence : PDL58-2R Primer
 <400> 12
 gcactagtca gccgcgctcg acgaggaggt g 31

<210> 13
<211> 33
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single
<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : pldB-L-Bgl2F Primer

<400> 13
gggagatcta gaggcggtt acctctacga gta

33

<210> 14
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : pldB-L-Hind3R Primer

<400> 14
gggaagcttg cgatgagctg tgccagatag

30

<210> 15
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : pldB-R-Hind3F Primer

<400> 15
gggaagcttg aactggcgcg acagtgttt

30

<210> 16
<211> 34
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>

<223> TOPOLOGY : linear
 <220>
 <223> Description of Artificial Sequence : pldB-R-Bgl2R Primer
 <400> 16
 gggagatctg cagcggatcg tcttcgagac cctt 34

 <210> 17
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> STRANDNESS : single
 <220>
 <223> TOPOLOGY : linear
 <220>
 <223> Description of Artificial Sequence : 5Dm-3F Primer
 <400> 17
 ttgcscctsc csgtccctsc satgtsat 29

 <210> 18
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> STRANDNESS : single
 <220>
 <223> TOPOLOGY : linear
 <220>
 <223> Description of Artificial Sequence : 5Dm-3R Primer
 <400> 18
 gttgatsays gasgtsgaga a 21

 <210> 19
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> STRANDNESS : single
 <220>
 <223> TOPOLOGY : linear
 <220>
 <223> Description of Artificial Sequence : 6PIN-2F Primer
 <400> 19
 gctgcgcctg gccctggagg acatcgagat 30

 <210> 20

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

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : 6PIN-2R Primer

<400> 20
ctgttcctcg aagaactcgt ggtcggcgta

30

<210> 21
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : 5D-1R Primer

<400> 21
aggtgccag cgagatcatg tt

22

<210> 22
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : 7PIN-2F Primer

<400> 22
ccatgacct gctggtggcc ggccatgaga

30

<210> 23
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
 <223> Description of Artificial Sequence : 5Dm-2R Primer
 <400> 23
 ctggatsgtg tcscscggyt t 21

<210> 24
 <211> 30
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> STRANDNESS : single

<220>
 <223> TOPOLOGY : linear

<220>
 <223> Description of Artificial Sequence : 5PIN-2F Primer
 <400> 24
 cggaatccac cagtgcctcg gccagAACCT 30

<210> 25
 <211> 33
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> STRANDNESS : single

<220>
 <223> TOPOLOGY : linear

<220>
 <223> Description of Artificial Sequence : pldApro-SpeNdeF Primer
 <400> 25
 gggcatatga ctagtagccg tgcctgccc gcc 33

<210> 26
 <211> 29
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> STRANDNESS : single

<220>
 <223> TOPOLOGY : linear

<220>
 <223> Description of Artificial Sequence : pldApro-NdeR Primer
 <400> 26
 gggcatatgt tcggacgtga attcattcg 29

<210> 27
 <211> 30
 <212> DNA

<213> Artificial Sequence

<220>

<223> STRANDNESS : single

<220>

<223> TOPOLOGY : linear

<220>

<223> Description of Artificial Sequence : eDM-SphF Primer

<400> 27

tccccgcatg ccggaactga cggacatcac

30

<210> 28

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> STRANDNESS : single

<220>

<223> TOPOLOGY : linear

<220>

<223> Description of Artificial Sequence : eDM-SphR Primer

<400> 28

cccgggcatg cagacgacga cgtagaggaa

30

<210> 29

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> STRANDNESS : single

<220>

<223> TOPOLOGY : linear

<220>

<223> Description of Artificial Sequence : PRR-1F Primer

<400> 29

gccccccata tgaacgcaaa cgacaacgtg gtcac

36

<210> 30

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> STRANDNESS : single

<220>

<223> TOPOLOGY : linear

<220>

<223> Description of Artificial Sequence : PRR-2R Primer

<400> 30
goggatcctc aggcactact cagttcagct ttggc

35

<210> 31
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : SP-1

<400> 31
tatgcgtcac tagtcgggag tgcgtta

27

<210> 32
<211> 27
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : SP-2

<400> 32
tataacgcac tcccgactag tgacgca

27

<210> 33
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : DM-NdeF

<400> 33
gcccccatat gacggaactg acggacatca

30

<210> 34
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : DM-SpeR

<400> 34
gggccactag tcagccggcc ggttcggtca

30

<210> 35
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : DM-BglI

<400> 35
cgcatagatc ttcaccogag cgggtgatca

30

<210> 36
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> STRANDNESS : single

<220>
<223> TOPOLOGY : linear

<220>
<223> Description of Artificial Sequence : DM-BglI

<400> 36
tcccgagatc ttgaaggtcc gcgtcacct

30

<210> 37
<211> 27
<212> DNA
<213> Artificial sequence

<220>
<223> R91C-F primer

<400> 37
ttcgcgccg tctgcgaccg gcgggtg

27

<210> 38
<211> 27
<212> DNA

<213> Artificial sequence	
<220>	
<223> R91C-R primer	
<400> 38	
cacccgccgg tcgcagacgg ccgcgaa	27
<210> 39	
<211> 27	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> R195L-F primer	
<400> 39	
tcgcaagggg cgctcgagcg gctcgag	27
<210> 40	
<211> 27	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> R195L-R primer	
<400> 40	
ctcgagccgc tcgagcgccc cttccga	27
<210> 41	
<211> 31	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> L403M-F primer	
<400> 41	
acgatccagg ggatgatgga actccccgtg a	31
<210> 42	
<211> 30	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> L403M-R primer	
<400> 42	
tcacggggag ttocatcatc ccctggatcg	30
<210> 43	
<211> 30	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> R236L-F primer	

<400> 43	
agctggaccg actcgacgtg gtggcgctgg	30
<210> 44	
<211> 29	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> R236L-R primer	
<400> 44	
agcgccacca cgtcgagtcg gtccagctc	29
<210> 45	
<211> 27	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> I244F-F primer	
<400> 45	
tggcgctggc cgtcttcctg ctcgtagg	27
<210> 46	
<211> 30	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> I244F-R primer	
<400> 46	
ccacgagcag gaagacggcc agcgccacca	30
<210> 47	
<211> 27	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> M112T-F primer	
<400> 47	
cagcggcgga tgacgatccc gtcgttc	27
<210> 48	
<211> 27	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> M112T-R primer	
<400> 48	
gaacgacggg atcgatcatcc gccgctg	27

<210> 49
 <211> 30
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> R195Y-F primer

 <400> 49
 aggggcgcgc gagtacctcg aggagtacct 30

 <210> 50
 <211> 30
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> R195Y-R primer

 <400> 50
 ggtactcctc gaggtactcg cgcgccctt 30

 <210> 51
 <211> 35
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> plout-L-Bgl2F primer

 <400> 51
 gggagatctg caggtcatcg gggggaagaa ccaca 35

 <210> 52
 <211> 31
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> plout-L-Sph1R primer

 <400> 52
 tctccagccg catgcggtcc cgggtcaccg 31

 <210> 53
 <211> 32
 <212> DNA
 <213> Artificial sequence

 <220>
 <223> plout-R-Sph1F primer

 <400> 53
 cccgcacg gctggagatc atccagagcg ca 32

 <210> 54
 <211> 34
 <212> DNA

<213> Artificial sequence

<220>

<223> pldout-R-Bgl2R primer

34

<400> 54

gggagatcta gaacatgccg ggccagaggc tgac

<210> 55

<211> 33

<212> DNA

<213> Artificial sequence

<220>

<223> hyg-SpeSphF primer

33

<400> 55

ggggcatgcg gactagtaca ccgtcgccctc ggt

<210> 56

<211> 30

<212> DNA

<213> Artificial sequence

<220>

<223> hyg-SphR primer

30

<400> 56

gccgcatgcg tcaggcgccg ggggcggtgt