

(19) AUSTRALIAN PATENT OFFICE

- (54) Title
Genetically modified microorganism and process for production of macrolide compound using the microorganism
- (51)⁶ International Patent Classification(s)
C12N 15/09 (2006.01) 20060101ALI2006113
C12N 1/21 (2006.01) 0BHJP **C12P**
C12P 17/08 (2006.01) 17/08
C12R 1/465 (2006.01) 20060101ALI2006113
C12N 15/09 0BHJP **C12R**
20060101AFI2006113 1/465
0BHJP **C12N** 20060101ALN200611
1/21 30BHJP
PCT/JP2006/310835
- (21) Application No: 2006250304 (22) Application Date: 2006 .05 .24
- (87) WIPO No: W006/126723
- (30) Priority Data
- | | | |
|-------------|--------------|--------------|
| (31) Number | (32) Date | (33) Country |
| 2005-154114 | 2005 .05 .26 | JP |
- (43) Publication Date : 2006 .11 .30
- (71) Applicant(s)
Mercian Corporation, Eisai R & D Management Co., Ltd.
- (72) Inventor(s)
Machida, Kazuhiro, Aritoku, Yasuhide
- (74) Agent/Attorney
Davies Collison Cave, 1 Nicholson Street, Melbourne, VIC, 3000
- (56) Related Art
W02005/052152 A1

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2006 年 11 月 30 日 (30.11.2006)

PCT

(10) 国際公開番号
WO 2006/126723 A1

- (51) 国際特許分類:
C12N 15/09 (2006.01) C12P 17/08 (2006.01)
C12N 1/21 (2006.01) C12R 1/465 (2006.01)
- (74) 代理人: 古谷 聡, 外 (FURUYA, Satoshi et al.); 〒1030007 東京都中央区日本橋浜町 2-1 7-8 浜町花長ビル 6 階 Tokyo (JP).
- (21) 国際出願番号: PCT/JP2006/310835
- (81) 指定国 (表示のない限り、全ての種類の国内保護が可能): AE, AG, AI, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, LY, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SI, SM, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.
- (22) 国際出願日: 2006 年 5 月 24 日 (24.05.2006)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2005-154114 2005 年 5 月 26 日 (26.05.2005) JP
- (71) 出願人 (米国を除く全ての指定国について): エーザイ・アール・アンド・ディー・マネジメント株式会社 (EISAI R & D MANAGEMENT CO., LTD.) [JP/JP]; 〒1128088 東京都文京区小石川四丁目 6 番 10 号 Tokyo (JP). メルシャン株式会社 (MERCIAN CORPORATION) [JP/JP]; 〒1048305 東京都中央区京橋一丁目 5 番 8 号 Tokyo (JP).
- (84) 指定国 (表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), ヨーロッパ (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IS, IT, LT, LU, LV, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- (72) 発明者: および
- (75) 発明者/出願人 (米国についてのみ): 町田 和弘 (MACHIDA, Kazuhiro) [JP/JP]; 〒4380078 静岡県磐田市 中泉 1 7 9 7-3 3 6 Shizuoka (JP). 有徳 保秀 (ARITOKU, Yasuhide) [JP/JP]; 〒4380078 静岡県磐田市 中泉 1 7 9 7-3 3 3 Shizuoka (JP).
- 添付公開書類:
— 国際調査報告書
- 2 文字コード及び他の略語については、定期発行される各 PCT ガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(54) Title: GENETICALLY MODIFIED MICROORGANISM AND PROCESS FOR PRODUCTION OF MACROLIDE COMPOUND USING THE MICROORGANISM

(54) 発明の名称: 遺伝子組換え微生物およびそれらの微生物を用いるマクロライド系化合物の製造方法

(57) Abstract: A genetically modified microorganism capable of directly producing a 16-hydroxylated macrolide compound; and a process for producing a 16-hydroxylated macrolide compound using the microorganism. Specifically, a genetically modified microorganisms having DNA encoding a polypeptide involved in the biosynthesis of a macrolide compound pladienolide and DNA encoding a polypeptide having a pladienolide 16-hydroxylase activity; and a process for producing a 16-hydroxylated macrolide compound comprising the steps of culturing the genetically modified microorganism in a culture medium and collecting a 16-hydroxylated macrolide compound from the culture medium.

(57) 要約: 本発明は、16 位水酸化マクロライド系化合物を直接生産する遺伝子組換え微生物、およびそれらの遺伝子組換え微生物を用いた 16 位水酸化マクロライド系化合物の製造方法を提供する。詳しくは、マクロライド系化合物プラジエノライドの生成に関与するポリペプチドをコードする DNA およびプラジエノライドの 16 位水酸化酵素活性を有するポリペプチドをコードする DNA を有する遺伝子組換え微生物、およびそれらの遺伝子組換え微生物を培地で培養し、その培養液から 16 位水酸化マクロライド系化合物を採取することを特徴とする、16 位水酸化マクロライド系化合物の製造方法である。

WO 2006/126723 A1

Description

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

Field of the invention

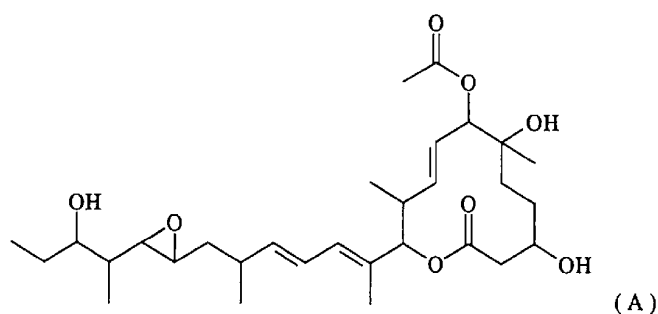
The present invention relates to recombinant microorganisms having the ability to produce 16-position hydroxyl macrolide compounds, and to methods for producing 16-position hydroxyl macrolide compounds using those microorganisms.

Related art

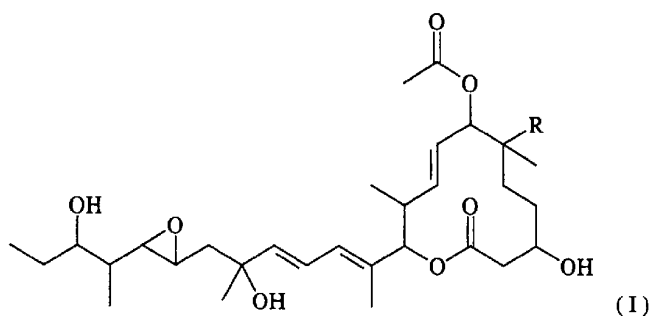
Important bioactive substances have been discovered among the various metabolites produced by the actinomycetes. In particular, many compounds have been found having a polyketide in the basic structure (hereunder called polyketide compounds). Compounds having various biological functions are known including 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

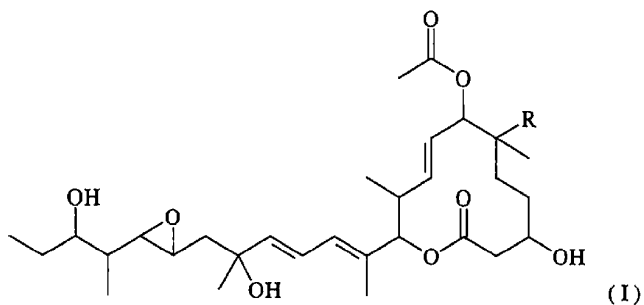
It is an object of the present invention to provide recombinant microorganisms having the ability to produce the 16-position hydroxyl macrolide compound represented by the above formula (I). Another object is to provide a method of producing a 16-position hydroxyl macrolide

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).

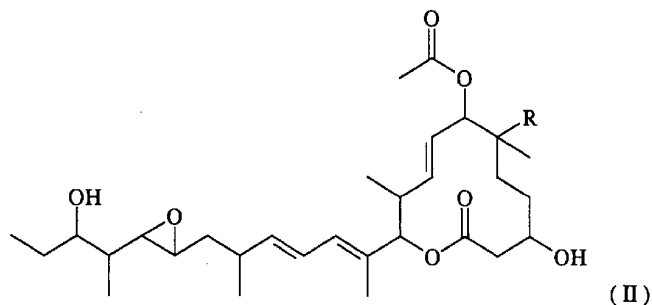
(1) A recombinant microorganism, which is a microorganism having ability to produce 16-position hydroxyl macrolide compounds represented by the formula (I):



(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).

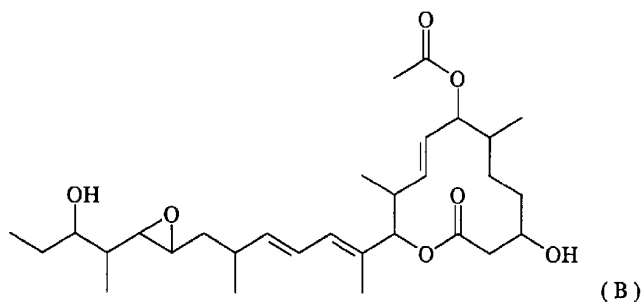
"DNA that hybridizes under stringent conditions" means for example DNA obtained by colony hybridization, plaque hybridization, Southern hybridization or the like using the original DNA as the probe, and specifically may be DNA which can be identified using a filter having DNA derived from a colony or 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

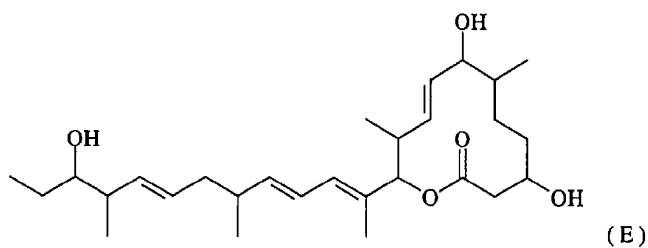
chain.

"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.

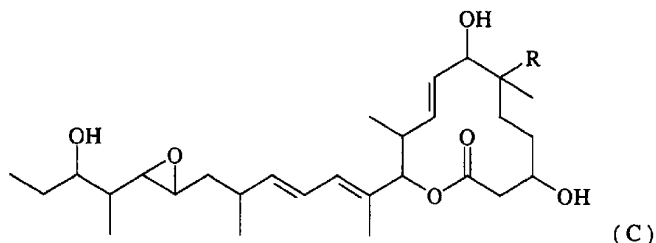


"Polyketide synthetase activity" means the activity of synthesizing the macrolide compound represented by the following formula (E):



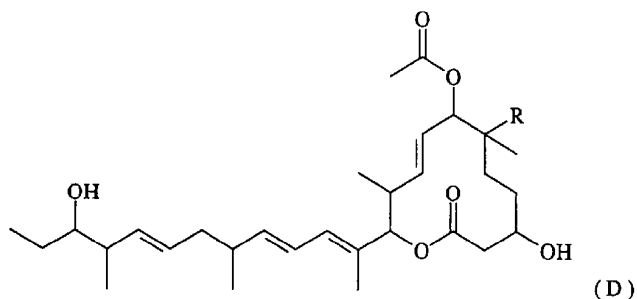
"7-position acetylase activity" means the activity of converting the 7-position hydroxyl group in the macrolide

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 cosmids 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 Depository (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 cosmids. 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%, CaCO₃ 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 × 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 of 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'-GTGCCCAGATGTTGGACTTCAACGA-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'-TCACGCGCTCTCCGCCGCCCTGCC-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'-GCCCCGCATATGGATCTGGAAACCAACTTCTC-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'-GGGAGATCTAGAGCCGGTTACCTCTACGAGTA-3'

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

pldB-R-Hind3F: 5'-GGGAAGCTTGAAGTGGCGGACAGTGTCTT-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 BglII and HindIII.

The DNA fragments L1 and R1 which had been digested with restriction enzymes BglII 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'-TTCGCSCTSCCSGTCCCSTCSATGGTSAT-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'-CTGTTCTCTCGAAGAACTCGTGGTCGGCGTA-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'-GCCCCATATGACGGAAGTACGGACATCA-3': see the Sequence No. 33) obtained by adding a NdeI site to the 5'-terminal and a primer DM-SpeR (5'-GGGCGGCTAGTCAGCCGGCCGGTTCGGTCA-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'-TTGCSCTSCCSGTCCCSTCSATGGTSAT-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 µM). 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 µg/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'-CCATGATCCTGCTGGTGGCCGCCATGAGA-3'

6PIN-2R: 5'-CTGTTCTCTGAAGAACTCGTGGTCGGCGTA-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'-TTGCGSCTSCCSGTCCSTCSATGGTSAT-3'

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

(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'-TATGCGTCACTAGTCGGGAGTGCCTTA-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'-GGGCATATGACTAGTAGCCGTGCTGCCCCGCC-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 *pldA* 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 (CD₃OD, 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 (17 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 pIdB::hyg::pAOC::PapsmAB strain		Mer-11107 pIdB::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 pIdB::hyg, ME-282 was produced by the Mer-11107 pIdB::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 pIdB::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 pIdB::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 (CD₃OD, 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 (CD₃OD, 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 µL

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

(Reference Example 2(1), 10 ng/µL)

(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::intphiC31 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::PoeppsmAB 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-1107 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-1107 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-1107::pAOC, pAOC::PepsmAB-transformed Mer-1107::pAOC::PepsmAB and pAOC::PopsmAB-transformed Mer-1107::pAOC::PopsmAB strains.

Example 13: Pladienolide production test of Mer-1107::pAOC strain, Mer-1107::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

pPampsmAB-CLM, pPampsmAB-CLLM, pPampsmAB-CLFM and pPampsmAB-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'-TTCGCGGCCGTCTGCGACCGCGGGTG-3'

R91C-R: 5'-CACCCGCGGTCGCAGACGGCCGCGAA-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 pPampsmAB-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'-CTCGAGCCGCTCGAGCGCCCTTCCGA-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 pPampsmAB-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'-ACGATCCAGGGGATGATGGAACCCCCGTGA-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 pPapsmAB-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'-AGGGGCGCGGAGTACCTCGAGGAGTACCT-3'

R195Y-R: 5'-GGTACTCCTCGAGGTACTCGCGGCCCTT-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 pPapsmAB-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'-GGGAGATCTGCAGGTCATCGGGGGAAGAACCACA-3'

pldout-L-Sph1R: 5'-TCTCCAGCCGCATGCGGTCCCGGTCACCGT-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 BglII 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'-GCCGCATGCGTCAGGCGCCGGGGCGGTGT-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 pPapsmAB-CLM, pPapsmAB-CLLM, pPapsmAB-CLFM and pPapsmAB-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 pldA promoter and psmAB or variants thereof (hereunder sometimes called PapsmAB, PampsmAB-CLM, PampsmAB-CLLM, PampsmAB-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, PampsmAB-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, PampsmAB-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-PampsmAB-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-PampsmAB-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::PapsmAB-CLM, Mer-11107-ZX::hyg2::PapsmAB-CLLM, Mer-11107-ZX::hyg2::PapsmAB-CLFM and Mer-11107-ZX::hyg2::PapsmAB-TY.

Example 18: Pladienolide production by Mer-11107-ZX::hyg2::PapsmAB, Mer-11107-ZX::hyg2::PapsmAB-CLM, Mer-11107-ZX::hyg2::PapsmAB-CLLM, Mer-11107-ZX::hyg2::PapsmAB-CLFM and Mer-11107-ZX::hyg2::PapsmAB-TY

500 μ L each of frozen seed of the Mer-11107-ZX::hyg2::PapsmAB, Mer-11107-ZX::hyg2::PapsmAB-CLM, Mer-11107-ZX::hyg2::PapsmAB-CLLM, Mer-11107-ZX::hyg2::PapsmAB-CLFM and Mer-11107-ZX::hyg2::PapsmAB-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

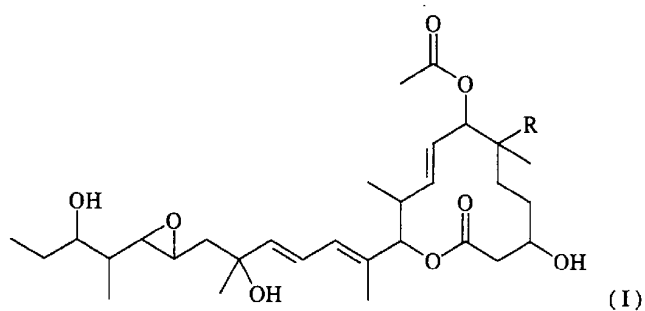
As a result, large amounts of the pladienolide degradation products pladienolide H13, pladienolide C2 and pladienolide C3 were produced together with pladienolide D by Mer-11107-Zx:hyg2::PapsmAB having introduced PapsmAB, while with Mer-11107-ZX::hyg2::PapsmAB-CLM, Mer-11107-ZX::hyg2::PapsmAB-CLLM, Mer-11107-ZX::hyg2::PapsmAB-CLFM and Mer-11107-ZX::hyg2::PapsmAB-TY introduced with PapsmAB-CLM, PapsmAB-CLLM, PapsmAB-CLFM and PapsmAB-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.

Throughout this specification and the claims which follow, unless the context requires otherwise, the word "comprise", and variations such as "comprises" and "comprising", will be understood to imply the inclusion of a stated integer or step or group of integers or steps but not the exclusion of any other integer or step or group of integers or steps.

The reference in this specification to any prior publication (or information derived from it), or to any matter which is known, is not, and should not be taken as an acknowledgment or admission or any form of suggestion that that prior publication (or information derived from it) or known matter forms part of the common general knowledge in the field of endeavour to which this specification relates.

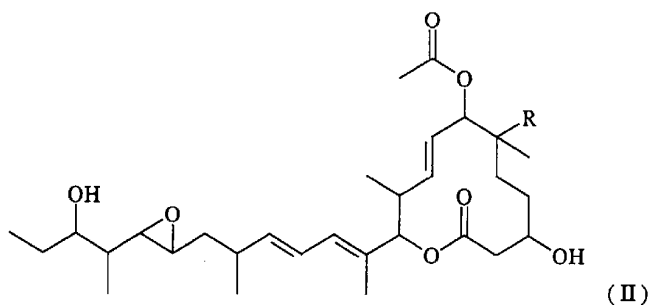
Claims

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 recombinant microorganism according to Claim 1, or a method according to Claim 15, substantially as herein described with reference to the Examples and the accompanying drawings.

Fig. 1

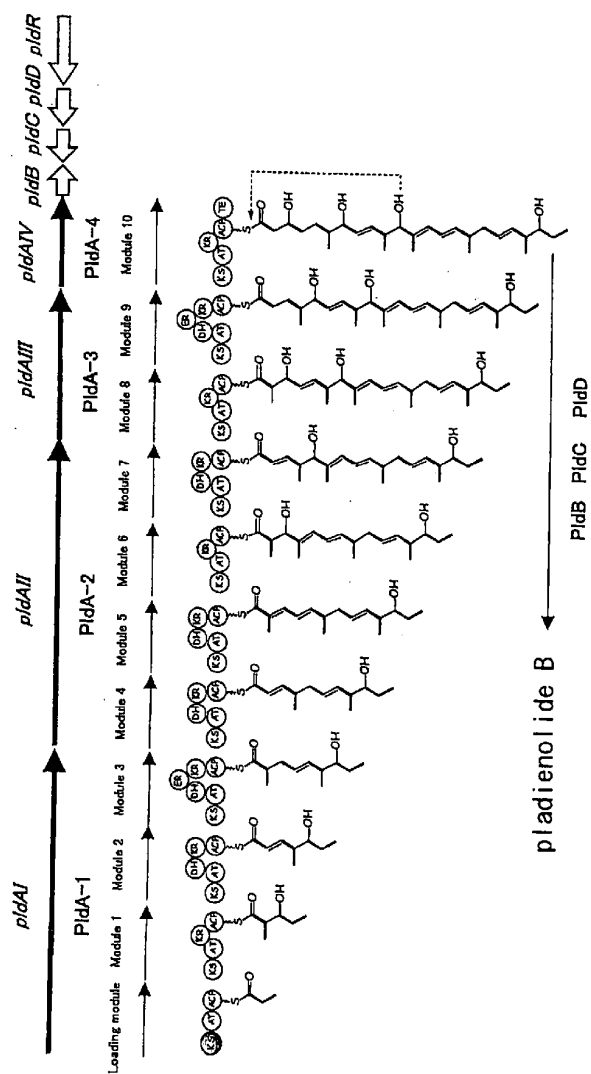


Fig. 2

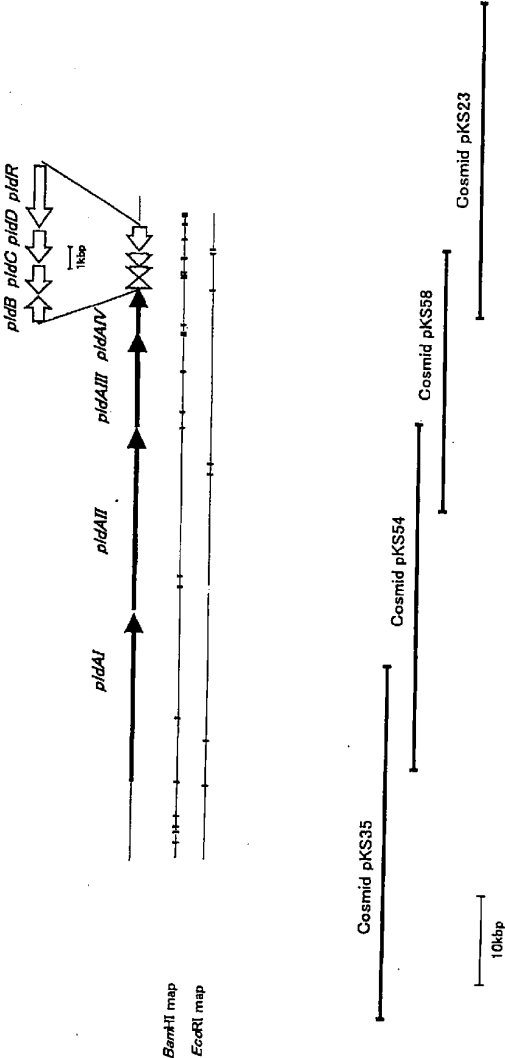
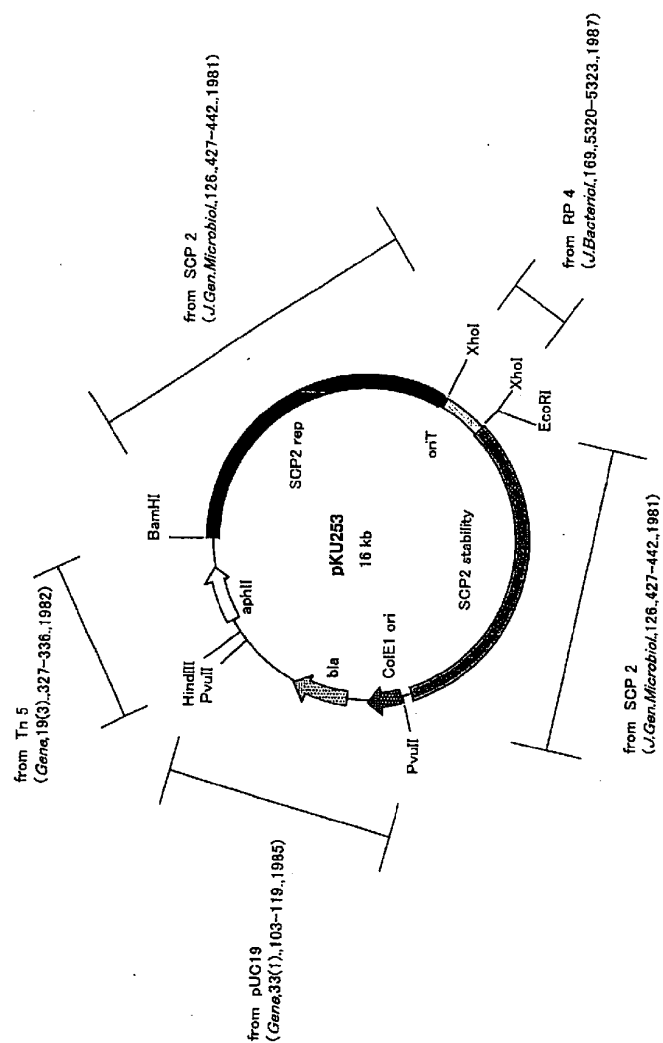
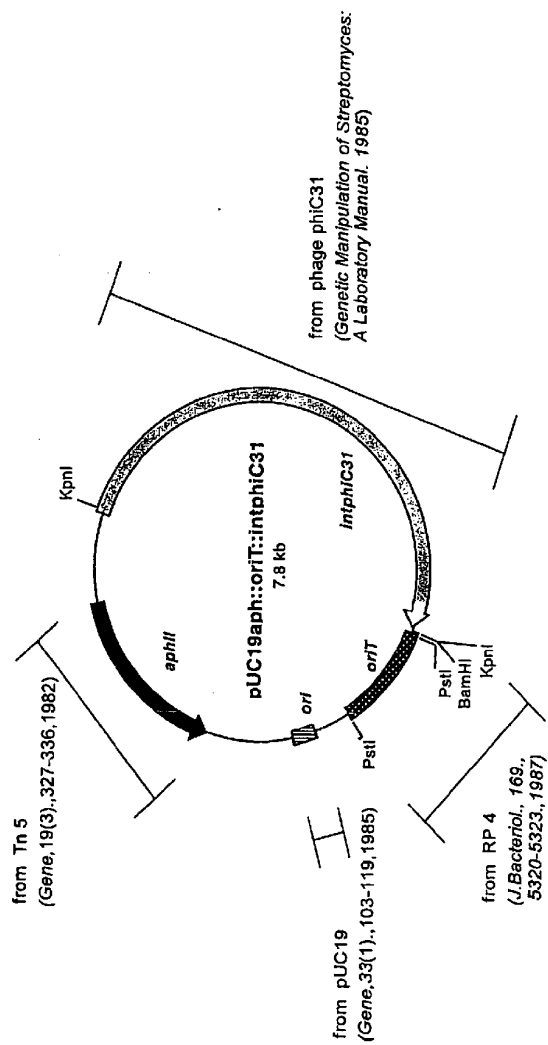


Fig. 3



3

Fig. 4



SEQUENCE LISTING

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

<110> Merckian 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 cgtgtgtgtg gtgaggcatg ctctatttgg aacataaaac      60
ctctgaacct ttaagaggtt atggcggagg ctctgacgc gacacgaggg agaagcggat      120
gagaatcgtg gggattcacc gggaggcgcg aggcatagag tggcccgcg tctcggagca      180
cgggcggcgg gcagtcgtgc tggcccgctc cgaagtcttc tggcccgacg ccaccggcca      240
tctggcgcgc ggggacggtg gaccagtcgt ccggtgttcc gcggtggagc tggtaaccgc      300
ggttctgcgc gacgcgcggg tgatctgcat cgggtccaac tacctcaagc atgtggccga      360
gggaacctac cgcgaccagg aagtcccgca gaccccgacg ctgttcgccc gctggacacg      420
gtgctgacc gtggacggag cggaggtccc ggtgcccctg gacgaggcgg gctggactg      480
ggagggtgag gtgtggcctt ggttggcgcg accactcgtg gacgccacgc cggaggaggc      540
gctgaccgcc gtcctcggtt actccctctt caacgacctc acctcccgcc gggctcagaa      600
gttcacctct cagtggaccc tgggcaagaa cggggacaac tccggcccgc tcgggcccgt      660
ggttccggct gccgaggttg gcgacctcgc cgacggcgctg cgggtacaga cccgggtcaa      720
cggggagacg atgcagatg gcagcacgga cgagatggtc tacaccgtgg gtgacacgct      780
cgccgacatc tcccgcacct tcatctctgc tcccggcgac ctgttgcgca cgggcacccc      840
tccggagatc ggtacgccc ggaccccgcc gcagctcctg cagccgggag acgtcgtcga      900
ggtggaggtc gaacggctcg gcgtgctgcg caaccccgctg gtgtccaacg acgcccggct      960
gcgcgcaccc aagtggagac gcaagaggcc ccgcccgcgc ccgccaagc cgggtgctcg      1020
ccctcggcca cagcccgag gacacacctg gtcacctccc tgcctcgcgg ggtccttgcg      1080
tggggcaccg ttgaccgttg tcaaggacta caccgagaa cagaagcgcc cggactgagc      1140
cgggcccgat cgtggggaca ttccgaagcg aggcggcgcc gctcccgccc gcttctgcta      1200
ggatgcgttt cctcccgggc gtgaacgccc ggggttcttc gcacggtcag gctgagactt      1260
ctctcacatg gggagtccag tctcggcgcg gatctgcggc cggagccggt accggcgccg      1320
agggctggaa cggcgaactc acgtactcct ccacgagcgg attacggtg cgggacacgaa      1380
aggtcggaga ccagggggcg gctcggcgcc gggattaggc gtagggttgg ctgatgtgtg      1440
gcttggcgcc ggcggtggcc agggacgtgt cagggaagcg ggcgcccgcg acgtagtggc      1500
ggcccttggc ctgccccacc gccgtcagca gctccgcctt gtcaggggcc tgcagggtcc      1560
gtgtggcctg ctgggtgttg agggcctcgg ctgcgtcata gcgcgagcgg cgaacccggc      1620
ccaccatcgc cacctcgtgc agcgcggtga tctgcgcgtc ggtgatcccg aggtctgtcg      1680
cgccgtccat gagctgcctc cagcagtcgt tggagcggtc gacgcggcgc tgcacccgct      1740
gggtctgtct gtggtaggcg agcaggtlga agcggatcca cggcccggtg tgcgcttgc      1800
gtgagtagac cgggcccgcg acctcgcgca gcgttttga gtactccag gtgttgcggc      1860
gcatcccgag ccaactcctc atggaggaga actccggcgc cagcaactcc cccggggcga      1920
tgaccagcgt ctgcagggag cgggacatcc gcccggtccc gtcggaccac ggggtggatc      1980
tcaccaggtt caggtgcgcc atcgccgccc gcaccaggac gtggcggtcc aggtcgcgt      2040
cgttgagcca gtccaccagc tcgcccata ggcggcgag caggtcggcg tccgggcctt      2100
cgtagtccgt ggccagctcg tcgcccgggg cgggtgatcg gatcgagtg cgacgccact      2160
ggccggccag ccgagcgga tgggtgtggc cctgcagcat ccagtcagc gagttgagca      2220
actccttctc gtagctgaag tcgctacgt cgtgcaggga ctggatgtag gccatcgct      2280
gctgtaggca cagcgtctcg gccttgttct cctcgtggc atccacggca tccggttcgc      2340
cgtccatcag gtccgcgaca tccctcgcac cgacctggtg acctcgtatg gtgttggacg      2400
ccgcatcgcg cgtggccgtc agcgccttgc gcagatcctg cgtccacttc gtcggcacct      2460
gctgcacgcg gtggcgcgag tgcctctgca gggatgtgat ctctccagg acccgcggt      2520

```

1/39

cgctcggtgt	gaggtgaggc	gtctgataca	gcataagtga	atgataacctg	actcctatca	2580
ttccttcaac	gctgcggctt	catcacccgtc	cgccaacgat	gaacatgaac	tcgccagct	2640
cgacgagcag	ggccaccggc	ctgcgcaccc	acgctgggtca	ctgatcaaca	caccaggtca	2700
acgagtagct	ctgcaggact	ggatggagtg	acaggcagcg	cggggcagcc	aagcaggcgg	2760
cgcgaccacg	taacggaccg	gcgaaccgga	gctgagcacc	ggtagctcc	gcgaccggct	2820
gactgocctc	agctacaaca	tcgagaacag	ctggccgagc	ttgcccgzag	ccggcggcgc	2880
ggatcacgtc	cgccgtgtgt	cgaggggcg	cgccgtttcc	ggcggtgtcg	tcgacgccag	2940
gcccccaag	ttgctgagcg	acgcggcggt	ccggacgtcc	gcacgtcgcg	gatccccgaga	3000
cgagggttga	gccaaagccg	cgggatcacg	ccgcgcgcga	tgatgcctac	cacggcttgc	3060
cgatgtggag	gctgggcaga	gatccccca	gcctccggtg	cgctgacggg	caggcacccg	3120
cctggccggg	gctccgtgcg	gggcccgttg	agcggagtgt	gcccccgcac	cggagaccgg	3180
ccaggactcc	gtgcgaccaa	aagactcttc	gtctcacctg	gtctcacaca	ttctcacaa	3240
acgagagcgc	acgggagcac	gcgagaacct	ctgacctgga	cattcgctgt	cgccacggac	3300
caggcgctac	caggtagaca	ttacataggg	taggactcga	ttctagtgt	caagtcaaat	3360
gccacaggtc	gaggcgtttt	ttgctgttcg	ccgagggtcg	atatctcaca	tatttattgc	3420
tggaaactgcg	gcacaaggctc	gagatccatc	ggatgaccgg	cgggctgctc	gagtgtctga	3480
tcgacgaggt	ggaccgcccc	cgggtatcca	aggacgcgcg	cctgcggatc	ccgcggaccg	3540
agctggaggt	caggcgctgt	tcacgcgctg	ggttggcgag	ctggcgactt	gccgcaagtt	3600
cgccctcacc	gcgcgggaact	acaccgcctc	gaaactgatg	tcgtaggctc	ggctcggggt	3660
gagcggagca	tgcaaaactcg	acctggccga	catcaagtgg	gacctggggc	gottcggcaa	3720
gctccatgtg	cgtaacggcg	aggcgccccc	cggctcgggt	ccgcgcgagc	ggatggtgcc	3780
ctgatcaaac	ggcgccgacc	gcctgctgcg	gtggttcacg	gaggacgtcc	ggggccagtt	3840
cgacgacgac	cgccaccgcc	ccggtgcccc	gctgttcccc	tcctgagcga	agaacgcogt	3900
cgctcctcgc	cgccctgtcg	gcgacgacgc	gttgcgcaac	ggcttggccg	ccgcggcgga	3960
ggctcgctcc	aaaagggtgt	gtaaaaggcg	acgagcgaaa	gtttcgtact	gaacagccaa	4020
gttcttacga	aaggaccacac	gctcgtgtag	tcacgagtgt	caccatcga	ggcgatccat	4080
gocggagatcg	atgccgtctt	cgctagtac	tgcaacgggtg	tttgccgtga	cggttgggca	4140
ggctgtgtgt	tggtctgagc	atgggtgggg	acctgtgtga	tgctggggtg	tgggccagtg	4200
aagtggagcg	tgctcacgag	aggttcgtgc	accggttttc	caggcgcgag	ccgcgggagt	4260
cgccgcttgc	ctatatcgcg	ggactgaccg	ctccgttgga	gcggaagaac	ggctggacac	4320
tggccgaaca	ggccgggtcat	gtcgtctcgg	accgtattca	tcgactgctg	aaccggatcg	4380
agtgggaagc	cgatgaggtc	ctggcccgatg	tgccgcgacta	cgctatggag	aacctcgcg	4440
accccgaggc	cctcatcgtg	gacgacaccc	gcttctgtaa	gaaggggacc	cgttcggcag	4500
ggggccggcg	tcagtactcc	gggaccgccc	gggctgttat	ttcaattact	gctcgggtga	4560
cgggtgtgtg	cgggacaatg	agatgtgact	gcgcgcacgg	ccgctcgggc	ctcagttgtg	4620
gctttgaggg	cgaagttcga	tcagttcctt	ccccatctcg	atgagcggcg	ccgtcggatc	4680
taoctggcca	ggcaggccgc	cgccgttggc	cacggcgggga	tcacgtggtt	ggccaccgct	4740
tcocggcgca	gtcggccac	catcgacgcg	gggatcgccg	agctgtccgg	gcacactctg	4800
ccggccgggc	ggatccgggc	tcggggagcc	ggccgcaagc	cggtcacggt	caccgacccc	4860
ggtctgtctg	ccgcgcttga	agctctgato	gagccgcaca	cccgggcgga	tcgggtctcg	4920
ccattgcgct	ggaccacgct	ctcgtgcccg	tcctggccct	cgccgctgac	cactcagggc	4980
caccgggtca	gcgcggcgac	cgctggagcc	ctgctacatg	ccctgggata	cagcctgcag	5040
ggcacccgca	agaccacgga	agggggcagt	catcccagcc	gggatgtcga	gttcacgcac	5100
atcaacgcca	ccgcgcggga	cttccctgaa	gacaaccagc	cggtagtcag	cgctgacacc	5160
aaggccaagg	agtggctcgg	caaccgcgac	cgaccgggac	gcacctggcg	accgggcaag	5220
aacctatccc	gtgtggactg	ccacacgttc	accaccagtg	accagccagt	agccatcccc	5280
taocggatct	acgacatcgc	tcgcaacacc	ggctgggtca	acgtcgggac	cgaccacgac	5340
acggcgaggt	tcgcggttga	atccatccgc	cgctggtggt	agcagcacgg	acgcggcgac	5400
cacccggagc	ccggccgact	gctcatcacc	gcgcactcgg	gtggttccaa	cgacccccgc	5460
cgctggacat	ggaagaagca	ttctgcgcgc	ttgcctctgg	aaagcggact	cagatcacgc	5520
gtctgccact	tcaccaccgg	aacatcgaag	tggaacaaga	tcgaacaccg	gatgttctgc	5580
cacatcaccc	cgaactggcg	cgccaggccc	ctgaccagct	accaggtcgt	catcgagacc	5640
atcgccgcca	cgaccactcg	cacccggctc	agcatcggcg	ccgaactcga	caccggccga	5700
tacgaccttg	gcaccacagt	cccaaccgcc	gagttccaag	ccctgccaat	cacacccccc	5760
accttccacg	gcgactggaa	ctacaccctg	gcaccactcg	caccccggtc	gcccggagcg	5820
gcaccagacc	gacaacggat	cgaccccgcc	ctgaccacga	tgctcaccga	cccgccctcg	5880
acggcgatgt	cacgctccgc	cttcgaccac	ctggtcgcca	ttctgggaacc	gtactgggac	5940
gcccgtggcg	aggcggtatt	ccaacgacgc	ttccaccgcc	cacgcagcta	cctccacccc	6000
cagaccagca	gcctcgacca	ctaccaccgc	ctgctgacgg	ccctgtttacg	ccgcgcgaga	6060
gcccgtacca	gcacactgct	ggcccagctc	ctgaacgtcg	gcgcgaccaa	cctgtccaac	6120
cagttccaag	acggccaccg	cctcctggac	ctgcaccgca	tcgcggtcac	tcgctatct	6180
ggagccccgg	cccgaccctt	cgcccaacta	caagcccgcc	taccgccaca	cgacgacacc	6240
cgcacagatc	aactctgaca	gttattcaga	cacaggcccc	cggcgggatc	gagaactccc	6300
aggtcgccgt	ctacctggtc	taggcaggtg	cccggggcca	cgcggcggtg	gaccgggaac	6360

tgtaactgcc	ccgttcctgg	acctgtgacc	agggccgctg	caggcgccgg	gggctcggcg	6420
aggacatcgt	cttcgccacc	aagccggagc	tggccgcac	gatgatcgaa	cggttcctgg	6480
acgccggaca	ccacgtgggc	tgggtcgtg	cgacagaggt	ctacggcgcc	aaccggaagc	6540
tgcatctgc	gctggaggtg	cgccgctcgc	gctatgtcct	cgcgttgccc	tgctcggccg	6600
aagtcaccac	caaggcaggc	aagttccgag	ccgacacgct	ggcggcgaag	gtaccgaagc	6660
gggcctggca	gaagctgtcg	gcaggcgcgg	gagccaaggc	caaccgcttc	tacgactggg	6720
ccgtcgtcga	cctggccgag	ccggcccccgc	gccaccggca	gctcgtgatc	cgccgcaacc	6780
gcgcaccgg	tgaactggcc	tactaccgat	gccactccac	ctcaccggtc	ccgctcgcca	6840
ccctggtcag	ggttgccgga	tcacgggtgc	gggtggagga	gacattccag	accgagaagg	6900
ccctggccgg	cctggacgag	caccagctcc	cccgtatccc	ctcctggccc	cgctggggga	6960
ccctcgccat	gctcgccccc	gctttcctcg	ccgtcgtccg	cgccgacgaa	cacaccggcc	7020
cgacccccga	cgacctcatt	ccgctgacct	gcaacgagat	ccagcaacct	ttcctcgccg	7080
tcgtcgtcca	ccgctgtcc	aacgtcgcgc	accgctcgc	ctggtccgag	tggagacgac	7140
gtcatcaagc	ccgatcacgc	accagtcact	accggcgaca	agccgcaact	cagacatgaa	7200
gatcacgac	tacagctgga	gtattaggtc	cccgaacaac	aaccggggac	agtcctaagt	7260
ggcgtcactt	tcgtatcgaa	ccagcaaaat	ttggaggga	cgatattcga	ctcgcatagc	7320
cttacggtcg	gagcatatta	cgaccaggtc	aatgaattgc	tcocccgcat	gcattctcag	7380
cggggatgcc	gtcggagggg	cgagggtctg	tatgatgtcg	ttgacatcgg	ccgcagtggt	7440
gatggccttg	agcggagtg	cggttcgtgc	acagatcagg	tgggtgtcgc	tggcgttcct	7500
gcgccggctg	accggcgag	gacggcgctc	ggctccccc	tttcgcgcgg	acatgagagc	7560
tgctccagca	cgcggtgac	cagtcgagtt	cgccggccgc	gttgagttcg	gcgagcggga	7620
cggtcgctga	ccctgggtcg	ctgcctatcg	tcacgcccgc	gcaccgctgc	ctcgcagatg	7680
cgctcgtata	gtccggccgg	gacgtccacc	cgacacctc	ttcggaggcg	agggtcacc	7740
gacgcccagc	ggccagcggt	actgcataca	ctogatctcc	ctgcggctga	ttcccgacga	7800
ggctgagatg	ccgtatcgac	acaccacgtc	cgtagacgtg	gcccacgggc	gtccggtcaa	7860
cggtgagttg	cgtagcagcg	atctcctcgc	cgctcagtag	ccgcgcgagg	cgcccgaggaa	7920
tcggcgagtt	ggacgggggt	ggggcgagcc	tcgttagctt	tcctggccgt	caactcgacg	7980
acgtggccgt	gctgctggcc	gtagccgtgt	cctgcccgc	ggcggggtgg	ctgtccgggg	8040
ctgccggcct	ccccggcgtc	gttgccgcca	ccaggccgca	cgccactgag	cagctgaccc	8100
ggcgccgact	cgacgagggc	gccttcgtca	ccgaactcgt	cctgagggca	ccgggtcacc	8160
aacgctctcc	gttctcctcc	agacctctct	gggaagatca	ataggggccc	gtaaaggggg	8220
ggtagagggt	gaaaaagggg	gggtattcaa	aaataggctg	agtcgcgtgc	aaaaatcttg	8280
agaccggctg	gaacgggtgt	agctgaatga	ctgaatcgaa	tgaattcacg	tcgaagcggg	8340
tgctcagtcg	ggctgatgat	gcaatagcca	tcacgcgcat	gtcttcggcg	ctgcgcgagc	8400
cagtcactcc	tcaggagtcc	tgggaactcc	tgaggaaatg	tgagagcggg	attaccgagg	8460
tgccgcccc	cggtggggac	gcgaactccc	tcctcgtatg	ggaaacgttc	acgcccggga	8520
cgatgaatac	acgttggggc	gggttcacgc	acggcgtgga	ccagttcgac	cccggtctct	8580
tcgggatcgc	ctcccgcgaa	gggttcgcca	tggtcccgca	gcaacggctc	gtactggagc	8640
tgagctggga	ggccctggag	gacgcgcgaa	tcgtcccgga	gcgccttcgc	cacaccgcta	8700
ccggtgtctt	gctggcgccg	atctgggacg	actacgcata	attgatgagc	gogcaggggc	8760
gagaagcggg	gacctatcac	accgtgaccg	gtacgcaccg	cagcatcatt	gccaaaccggg	8820
tgctgtacgc	cctcggccta	caggggccga	gcattggcgt	ggactccggg	cagtcgtcgt	8880
cactggtctc	cgctcatctg	gcctgcgaga	gcctgcgcag	gggggagttc	acgtcgcgcg	8940
tggccggcgg	ggtgaatctc	aaccttgctc	cggagagcac	catcgccatg	gcgaagtctg	9000
cggggtctct	ccccgatggc	cgctgcttca	ccttcgacac	ccgcgccaac	ggctacgtgc	9060
gggttgaggg	cgccggtgtg	gtcgtctcca	aaccgctggc	ggacgcgata	cgggaccagg	9120
accgatctca	ctgcgtcctc	cggtgcagcg	ccgtcaacaa	cgacggttcc	ggtgagaacc	9180
tgaccacgcc	gaactcccag	gcgcaggcag	ctgtgctgag	cagggcctac	cgccgcggcg	9240
cgctggagcc	ggcccagggt	cagtacgtgg	aactgcacgg	taccgggacc	cctgtcggcg	9300
accgatctga	agccgaggcc	ctcggcgccg	tgatcggctg	cgcccgcccg	ccgggtgacc	9360
ccctgtgggt	gggatccggc	aagaccaaca	tcggccatct	ggaggccgcc	gcccggcatc	9420
ccggcctcgt	caaggtcgtg	ctgtccatca	gccaccggga	gctcccggcc	agtcctcaact	9480
tcgccacggc	caatccggcg	attccactgg	actccctgaa	cctgcgcgtg	ggcgacgagc	9540
tcacatcgtg	gccgtctgcc	ggtcggccga	tgctgcggcg	tgtagcgcg	ttcgccatgg	9600
gcggtaccaa	cgcccacgcc	gtggtcgaa	aatctcccgt	agcagcgccg	cagattccgg	9660
ctcccgagg	cacgccgacg	gatcaggggg	ggccggtgcc	gtggttctg	tcgggtgggt	9720
cggtggcgcc	ggtgcggggg	caggcgccgc	ggttgttctg	gcattctggg	ggtcggctcg	9780
gtctcgtctg	ggtggatgtc	ggctggtcgc	tgccacgac	tcgttccgtg	ttccctcctc	9840
gttctgttgt	cgttgcgcag	gatggtggtt	acggccagag	tcctcccgcg	ctggccggcg	9900
gttccttgga	tgccgggggt	gttaggggcc	ttgccgatgt	gagtggaag	acggtgttgc	9960
tcctcccgcg	tcagggttcg	cagtggttgg	gtatggccgt	tgagctgctg	gacggctcgg	10020
aggttttgcg	cgagcatatg	cccgcctcgc	ccagggccct	ggaaccgttt	gtgggctggt	10080
ccctggaggga	tgtcctcgt	caggtggacg	gtacgtggtc	actggatcgt	gtggatgtgg	10140
tcacgctcgt	gctgtggcg	gtcatggtct	cgctcgcggg	actgtggcag	gcacatggcg	10200

ttgagcctgc	tgcggtgctg	ggccactccc	aagggtgagat	cgtgcggct	tgcgtggcgg	10260
gtgcgctgag	tctggaagac	ggagcccggg	tgggtgctct	tcgcagccgc	gccatgcgcg	10320
aggcccttgc	gggcatggc	gggatgctgt	cgatagccgc	ccccccacc	gaagtacagg	10380
ccctgatcac	ccccggggc	aggcagatca	ccattgccac	ggtaacagg	ccgattcgg	10440
tggtggtcgc	aggagaccct	gacgcgtcgc	aggcaactccg	cgccgaactg	gagaccggtg	10500
gtctccgcaa	tcgtgcacac	ccggtcgact	acgcctcaca	caccctcac	gtcagggcga	10560
tcogtgaacg	gctcctggcc	gacctggcag	tgatccagcc	acgtccgcg	agcattcccg	10620
tgctgtccac	cgtaaccggc	gcattggctcg	acaccaccgt	gatggacgcc	gagtactggt	10680
accgcaacct	acgtcagacc	gtggagtccg	aagcagccac	ccgcaactctc	ctcgaccagg	10740
accaccgcta	cttcgtcgag	atcagcccg	accccgctact	caccaccgcg	atccaggaaa	10800
ccctcgactg	cacagacacc	gccgctgctg	ccaccgggaa	cctgcgacgc	aacgaaggca	10860
gccctccggc	tttccagctc	gcccttgccg	aactcgtcac	ccgtggcctc	accccgcact	10920
ggcccgccct	ctatcccgac	gcccgcacac	cggaactccc	cacctatccc	ttccaacgcg	10980
agcgctactg	ggtcggcagc	tcctcgggtg	gggacgcggc	gcccgtccg	caaccggacc	11040
cgccgaactg	gagagccggc	gggtccggctt	cgcccgggcg	cgccgtcgat	ggcggcgacg	11100
ggcccgccga	gctgctggct	ctgggtcggtg	cccacgtggc	cgtgggtcgc	ggtgagacga	11160
cgccggacag	tgctgatccg	aaactgaact	tcaagcagct	cggtctcgac	tcggtcatgt	11220
ccctcgactg	ccggaacccg	ctgagctcgc	ccaccggatc	gtctctcgcc	agcacagtgc	11280
tggtcaacca	ccccacggcg	gaccggctcg	ccgcacatct	gtccgcggag	cgctccagcc	11340
aggtggaagg	cgccacggac	gcccgcggcg	cggttgccgc	cgacgagccg	atcgcgatcg	11400
tgggtatggg	atgcaggtag	cccggaggag	tcgctgcgcc	ggaggacttg	tggcggtcgg	11460
tgacatccgg	ggcgatgacg	atctccggct	tcgccacggg	ccgtggctgg	gacctcgagg	11520
tcattgacga	cccgaccat	cgccggcccg	gcaccagcag	taccgcggag	ggcgggttcc	11580
tgtagcaggc	cggtgacttc	gacgcgggtt	tcctcgccat	cagcccgccg	gaggcgctcg	11640
ccatggaccc	gcagcagccg	ctgctgctcg	agacttcctg	ggaggccgtg	gaaccggcgg	11700
gcacatgccc	gctgtcgtcg	caaggtagcg	ggccgggggt	tttctcgagg	gccatggccc	11760
aggagtagcg	cccgctgtcg	gacgaggcgc	cggaaggcta	tgagggttcc	ctgctgaccg	11820
gtggcctgac	gagcgtgttg	tcggggcgcc	tgccctacag	cctgggggtg	gaggggaccg	11880
cggtacacgt	ggacaccggc	tgctcgtcgt	cgtggttcgc	cgtgcacatg	gcccggcagg	11940
ctctccgtca	ggggcaggtg	tcctggcgcg	tgccaggcgg	ggtcacctgc	atgtccggcc	12000
ccgggatatt	cctggagtcc	agcaggcaga	gcccactggc	accggacggc	cgtgcgaagg	12060
cgttcgcgcc	cggagctgac	ggcacgggct	ggcccggaagg	cgtcggcgtg	cgtgtgctgg	12120
agcggctctc	cgacgcccgc	cgcaacggac	atccggtgct	ggcgtgggta	cgggggctcg	12180
cgatacaacca	ggacgggtgc	tcgaacggcg	tgacggcacc	gaacgggctc	gcccaggagg	12240
gggtgatccg	tgaggccctg	acggacgcag	ggctgtctcc	gcccgcagtc	gacctgggtg	12300
aggccaccgg	caccggcacc	accttgggtg	acccgatcga	ggccgacggc	ctgatcgcca	12360
cctacggaca	ggggcgtcgc	gcccggcggc	cgtcgcgact	gggctcgtcg	aagtccaaca	12420
tcgggacccg	ccaggcgcca	gcccggagtgg	gcccagtcac	caagacgggt	atggcggtgc	12480
ggcacgcaca	catgccccag	accctgcagt	tcgacgcgcc	gtcacccgat	gtggactggt	12540
cgtccggcca	ggtccggctg	ctgaccgagg	cagtcggctg	gcccagatcc	gaccaccccc	12600
ggagggcggc	ggtctcgtcc	ttcgggatca	gcccacacaa	cgtctacgtt	gtcgttgagc	12660
agccccggcg	ggagggtgac	gcccgtcaccg	ggccatcacc	tatggcggcg	gacgaggccg	12720
taccggcccg	ggggcagccg	gtgcctggcg	tgctgtggcg	caagtcaccg	gaagcgggtg	12780
gcgagcaagc	ggcgcgcccg	cggtcgtacc	tgcccgaccg	gcccggcgcc	ggtctcgccg	12840
acatcgctgt	gtccctggcg	tcgaccgggt	cgccgttcga	gcaccgtacg	gtgggtggtg	12900
cgccggacga	tgggcagttc	cgtgaggcgc	tgggcgccgg	cgccgggggt	tcggcggtatg	12960
cccgggtcgt	cgaggcggtg	gcccacatcg	acggcaagac	cgtcttcgtc	ttccccggcc	13020
aggcgccgca	gtggggcggc	atggccgggg	aaactcctgga	ctcctccgag	gtgttcggcg	13080
cccggatggc	cgactgcgcg	cgggctttgg	cccggttcgt	cggtgtggtg	ttgcaggatg	13140
tcgtccggca	ggccgagggc	gcccggcccg	tggaaccgggt	cgacgtcgtc	cagccgggtg	13200
tggtggcggt	catggtgtcg	ctggccgacc	tatggcgtgc	tcattggcgtt	gagccctcgg	13260
ccgtggtggg	ccactcgacg	ggtgagatcg	cgcccgcccg	cgtcgccggt	gggtgacgc	13320
tgggaagcgc	cgccgggggt	gtgtcgtcgc	ggagccgggc	catcgccgaa	gtactcgccg	13380
gacacggcgg	catgctgtcg	gtgaccgcgg	ccggggaaca	ggtcgaggag	tggtgctccc	13440
cctgggaggg	caggatttcc	ctcgcaacca	tcaacgggaa	ggaatccgtc	tggtcgccg	13500
gcgatcccca	cgccgtggcg	gaattccggc	cgtggttggg	gaaccgacag	atccgtagcc	13560
gcaccctgcc	ggtcgattac	gcctctcact	cgccgcagggt	cgaggctgtc	caccagcgac	13620
tgctggacga	cctggcgccg	atccgccccg	gtacgtgccc	taccccgctg	ctgtctcgg	13680
tcaccggcca	gtggctggac	accgcctcga	tggaacggcg	gtactggtac	cagaacctgc	13740
gcccggaccgt	ggagttcgcc	gcccggaccc	gcaccttggc	cgacgggggg	caccgcatct	13800
tcacgaggtg	gagctcgcat	ccggtgctgg	tcggcgcgat	acgggaaacc	ctcgaagccg	13860
tcgaggtcca	ggccgctgtc	gcccgggtcac	tcggcggtga	cgacggaggc	ctgcggcggtt	13920
tcgggctctc	gcttgcggcg	ctgtcacccc	gggggctggc	ccccgactgg	tcctatgctc	13980
gccccggggt	gagccgaacc	gacctcccca	cctacccttt	ccagcgcgag	cgttactgga	14040

tcaccgcctt	ctcggggctg	cggagcgccg	gtgaactcaa	cgtgcggac	tcacgtttct	14100
gggaggcggt	cgacagcgag	gaccccgggc	ggctggccga	ggtgctcagc	ctcagacgac	14160
acgcgttcgt	cgaaccggtc	ttcctggcac	tgctctcgtg	gcggcgacgg	caccgggtgc	14220
ggtccaccct	ggacgactgg	cgttatcggg	tgacctggca	gccgctgccc	ggggccgccc	14280
tcocgttgac	ggcggcaacc	ctcggaggga	cctggctggt	ggcctgcccc	cacgaggacg	14340
cctacgtctc	ccaggtgctg	cgcgggctgg	gcgaccgcgg	cgcgaccgtg	atcaccctac	14400
gagccgacga	cccgcgccac	ggcccgctcg	ccgagcgggt	ccggagggcg	ctggccggag	14460
cggggcgagat	caccggcgtg	ctgtcgtctg	tggcgttgga	cgagcggccg	caccgggaac	14520
atccggttct	tcccatgggc	ctggcgctca	acacggcgct	ggtgcgggca	ctggtggaca	14580
aggacgtccg	ggctccgttg	tgtgtcgcca	cgcggggcgc	ggtgtcgggt	ggccgatccg	14640
accggctggg	cagccctgcc	caggcgatgg	tgtgggggct	cggcctgggt	ggggccctgg	14700
aacacccgcg	gcactggggc	gggctggtgg	atctgcccga	aacogtggac	gagcgggtgc	14760
tgaaccggct	ggtgaccgtg	atctcggggc	aacgagtgca	cggacaggga	ggcccgggcc	14820
aggacggcga	aaaccggggc	gatgaggacc	agcttgccgt	gcgggcgtcc	ggagtgttgc	14880
cgcggcggtc	gtcgcacccg	cccgtgtcgg	gcagccgcaa	ccggagtggt	acgcccgggg	14940
gcacccgtct	ggtcaccgga	ggcaccgggt	gcgcggggcac	ccaggtggct	cgtgtgctgg	15000
ccgctaaagg	cgccgaacac	ctgctgtctg	ccagccgtcg	tggcaggagc	ggccgagggg	15060
ccgcgcgagct	ggcggccgaa	ctacggaag	ccggcgctcag	ggtcacggtc	ggccgctcgc	15120
acgtagcggg	ccgggacccc	ctggcccgcc	tgctgcggcg	cgtaccggac	gagctgcccg	15180
tgaccgccgt	gattcatgcc	cccgtgtgtg	tcaccaccgc	cccgtgggac	agcaccggtc	15240
cggaggaaact	ggccgaggtg	ctggcggggc	aggtggccgg	cgccgcccat	ctggagcgtc	15300
tgctcggcga	ccggcagttg	gacgccttcg	tactgttctc	ctccaacgcc	ggcgtgtggg	15360
gcagcggcgg	gcaggcggcc	tacgcgcggg	ccaacgccta	cctggacgcc	ctggcccagc	15420
agcgttcctc	tatgggcccag	accgcgacct	cagtgccctg	gggtgcctgg	ggcaggggcc	15480
ggatggcggg	cgaggaaagg	ttcaaggagc	ggctgcgcgg	gcggggcatc	atcgaatagg	15540
accggcgagct	ggccgtcacc	gcgctcgtcg	aggccgtcga	gtccggagag	gcgtcgatag	15600
ccgttgccga	cgtcgattgg	gcacgcttcg	tgcccggctt	cacctogaac	cggcccagtc	15660
cgtgatccgg	cgacctgcct	gaggtgcggg	acgcgtgcgg	ggaggccgac	agccggcccg	15720
ccgtcgatca	ggggcggtcg	gcgctcgcca	cgcggctggc	cgggctgtcc	gtgctcgaac	15780
gggagcgggt	cctgctcaac	ctgggtcgca	ccgaggtggc	ctcggtactc	ggtcacacca	15840
cggccgacat	ggtcgatgcc	cgtgcgccct	tcogtgaact	cgggttcgac	tcgctgatcg	15900
cgttgaggtt	ccgcggccgg	ttgaaccggc	cgaccgggct	gcgctgcctt	acctcggtcg	15960
ccttcgaccca	ccccaccccc	gcccagctcg	ccggccatct	cggggagtgg	ttgcgggat	16020
cccgcgggtga	caccgccatg	cccgtgtcgg	tgaccaccgc	cggggagcac	gaaccgatcg	16080
ccatcgtagc	gatgtcctgc	cgtlaccggg	gcggtgtcgg	cactccggag	gacctgtggc	16140
ggctggtggc	cgaggggccc	gacgcgatca	cggacttccc	caccgaccgc	ggctgggata	16200
tcgaagccct	gtatgacccc	gaccggggcc	ggtccggcac	ctctacaccc	cggcggggcg	16260
gcttctctga	cgacgcggcg	gccttcgatc	cggcgttctt	ccggatctcc	ccccgcgagg	16320
ccctggccat	ggaccccgag	cagcggctgc	tcctcgaaat	gacgtgggag	accctogaac	16380
gggcctctat	cgacccaaca	acgtggaagg	gcagccaggg	cggggtgttc	atcggcaccg	16440
cacaccccg	ctacggcgag	ggcatccacc	acgagtcgca	gggcgtcgag	ggccagcagc	16500
tgctcggcgg	ctcggccccc	gtggcccgag	gcgggatcgc	ctacacgttc	ggcctggaag	16560
ggccggcgat	gacggtggac	acctgtgtct	cgttcctcgt	ggtggcactc	catctggcct	16620
gccagtcctt	gcgcaccggc	gagtcctcga	tggcgtcgcg	cggcggggtc	acggtaatgg	16680
cacggccgac	cgttttcacc	gagttcagcc	ggcatcgggg	actgtccccc	gacggacggt	16740
gcaagtcctt	ctccgacgcc	gcgcagcgca	ccgctggggc	cgaggggccc	ggtgtgctcc	16800
ttctogaacg	gctctccgac	ggccgtcgaa	acggccaccc	cgtgtcggcc	gtcatccggc	16860
gcagccgcat	caaccaggac	ggcgccagca	acggccttac	cgcacccaac	ggcccccctc	16920
agcaacgcgt	catccaggag	ggcctggcga	acgcgtccct	gtgcggggcc	gacgtcgccg	16980
ccgtcgaggc	ccacggcacc	ggtaccaccc	tggcgacccc	gatcgaggcc	caggccctga	17040
tcgcgcgcta	cggacaggac	cgcggcgacg	accggccgct	acggctgggc	tcgctgaagt	17100
ccaacatcgg	ccacgcgcag	tcgcgagccg	cagtcggcgg	cgtgatcaag	atggtccagg	17160
ccatccggca	cggcctctct	ccgcgcacgc	tgacgcggga	gcagccctcc	cgcacagtgg	17220
actggtccgc	cggctcgggt	gaactgtcta	ccgaggcgat	ggcgtggccg	gacaacgacc	17280
aaccccgccg	ggcgggtgtc	tcggcggttcg	gcggcagcgg	caccaacgcc	cacatgatca	17340
tcgagcaggc	ggccgcggcg	gacgagccgg	agcacacoga	cggcacggag	aggaccagcg	17400
gcgagagcgg	cgcggaacag	ccaggccgcg	tgccgatggt	ggcctgggtg	ctgtccgcgc	17460
ggagtgacac	cgcgctgcgg	gcacaggccc	ggcgccctcg	gcctacgcg	ggcgccgcgg	17520
aggcggcgag	catctcgac	atcggtgggg	cgcigggcag	cacccgagcc	acgctggacg	17580
accggcccg	ggtcgtggcc	cgggaacggg	aaggattcct	caccgctctc	gacgcgctgg	17640
ccgaggaccg	gaccgcccc	ggtctggtcc	ggggggcggc	tggaaacagg	gtcggtcggg	17700
catctcgtt	ctcggccgag	ggctcacaaa	gactcggcat	ggggcgcgag	ctgtacgaca	17760
cgtccctcgt	gttcggcgag	gcgctggacg	aggtgtgcgc	ccagctcgac	gggcaacctg	17820
accggccctt	cctcgggggt	ctgttcggcg	cggagggttc	cgacgacgcg	tcgatgctgg	17880

accagaccgc	cttcaccag	gocgcgttgt	tgcgggtoga	ggtggcgtg	ttccgcctcg	17940
tctggctctg	gggcctcggg	cccgatttcc	tcacgggca	ttccgtgggc	gaagtcggg	18000
ccgcccatgt	ctcgggcttg	ctgtccctcg	ccgacggcg	gacactggtg	gtcggccgg	18060
ctggcgctat	gcaggcgtt	ccctccggcg	gcgcgatggt	ggccttgcaa	gcgggtgagg	18120
aggaaagtacg	ctgttccctg	gccccactgg	aggacgttgt	cggcgctgcc	gccctcaacg	18180
gccccgcctc	gaccgtgatc	tctggcgacg	aggagccgtt	cttcccggtg	gcccgcgact	18240
ggcgcgcca	gggcggcaag	acgcgtcgcc	tcaaggtgag	ccacgccttc	cactcaccgc	18300
gtatggaaac	catgttgac	cgtttccacg	ccgtgctcaa	aacgctttcc	ttcggcgagc	18360
cgccatttcc	cgtgtgtctg	aattgtaccg	gocgtccgcg	cgagcggacc	gaactgtgctg	18420
cggcggaacta	ctgggtgcgc	catgtccggc	atccggtgag	cttccatgac	ggcatccgg	18480
cgctggaggc	cgaaggcgct	agcgcatctc	tggagtgggg	gcccgcggcg	acactctcgg	18540
cgatggtccg	cgactgcctg	gaacaccagc	gcccgtggtt	cacggcaccg	gttttgagac	18600
gtgaccgtac	cgatgtgtct	gcccgcgtga	cgccactggc	cgaagcgac	gggcacgggg	18660
tgccgtggga	ctggcgctgc	ctcttcggcg	gctcgaccgc	ccggcgcgctc	gagctgcga	18720
gctaccggtt	ccaggcgga	cacttctggc	tggattccgt	cacggcgagc	agtgcacatga	18780
gcacggcgcg	actggcgctc	cccgatcatc	cgctgttggg	agccgtgacg	acgggtggccg	18840
gcagagggcg	ccctctcttc	acgggcaacc	tgtcggtacg	gacgcaccca	tggctggccg	18900
accacaggat	caccggttcc	gtcctgtctc	ccggcacggc	gttctctgaa	ctggccgtcc	18960
aggccggggg	ccaggccggc	tgcggcgggg	tcaggacact	gacgctgctg	gctccgctcg	19020
tactgcocga	agaggcgagc	gtcagggtcc	agatgaaggt	gggggagccc	gacgccacgg	19080
gcccgcgcac	catcgaggtg	tactctctcg	accagcaggc	ccccggccgg	gaacgttggg	19140
tcctcaacgc	gagcgggatg	cttggccggc	aaccggtgga	ggccccggcg	agtcacacca	19200
ccctggccccc	ggaaggcgct	gtccccgttc	cgctggacgg	cttcacgac	cggctggcgg	19260
cacgcggcta	cggtacggc	ccgacattcc	gcccgtgag	cgccgcgtgg	tcacgcgggtg	19320
acgagatctt	cccggaagcg	gcgcctccct	cgggccatcg	gcaggatgcc	gcccgcctatg	19380
gactccaccc	cgccctactc	gacgctgccc	tcacgcccat	ggaactccgg	gaaccccggc	19440
cggccggcga	cggagtccgg	cttccgttcc	cctggaaacg	cttctccctg	cacgcgtogg	19500
gtcccggaagc	ggtacggctg	cgccctcgcc	cgacggcgcc	cgacgctctg	tcgggtgacc	19560
tcggcgatgc	catcggtcgc	ccggttgctc	cagcccgctc	gctggccctg	cgggagctct	19620
cgctcgacct	gctgcggccc	gogtccgtct	cgtaacggga	ctcgctgttc	cgacccgctt	19680
ggataccgcg	cctcgtcgcc	ccggaggcgg	agtcggggcc	ggtgcgacgg	tcggccggct	19740
gggcgggtgct	ggggcccgat	ccgctcgccg	cgcccaacgc	ccgtaacctc	acgggaacct	19800
tcctgcctcg	ctatccggag	ctggcgcgcc	tgatcgccgc	cgtcgacggc	ggagcccgcg	19860
tgcccgaggg	cgtaactcgc	ccgtacggcg	cgagccagc	cccgacggcg	ggatctcccg	19920
cggacggcgt	acgggcctcg	accggccggg	cgctgcaact	gctgcaatcc	tggctgtccg	19980
aggaccgggt	ggagcgaagc	cggctgacg	tgctaccccg	ggggcggtg	gcccgtggta	20040
cgagcgaagg	cgtaaccgac	ctgggtgag	cgctggctcg	gggtctggtc	cggttggcgc	20100
aggccgagca	ccctggcagg	ttctccctgg	tcacatcgca	cgacccggag	gagtcctggg	20160
ccgtcctgag	cgccggcgcc	gtatccgatg	agccacaact	cgccctgcgc	tgccggccaga	20220
tgaagggtcc	ccgctcggcg	tcctgcgacg	ttcccacgac	cggtatccct	gagatgcccg	20280
acgtttgggg	tggtgacggt	accgtgttga	tcactggcgg	gaccggtgtg	ctgggtggcg	20340
tcgtcgcccg	tcactgtgct	gcccggcatg	gggtccgtcg	tcgtttgctc	tgacgacggc	20400
ggggccctga	tgccgggggt	cggttgagag	tggtogccga	gctcaccgct	ctgggtgccc	20460
atgtaccgct	tgccgcctgc	gacgcggccg	accgggatgc	gctggccgcg	ctcttgagaca	20520
ccgttcccg	cacgcaccct	ctgactgggt	tcgtgcatac	cgctggtgtc	atcgatgacg	20580
ccactgtcac	caccctcact	cccgagcgca	tcgacggggt	cctacgcccc	aaggtcgacg	20640
ccgcgcctaa	cctccatcag	ctgacggcgc	atctcggttt	gaccgcgttt	gtgctcttct	20700
cctccgcgcg	cgggctcttc	ggcgcccgcc	gcccgggtaa	ctacgcggcc	gccaacgcct	20760
tcctcgacgc	actggcccaa	cacggcgggg	ccaacggcct	caatgcccag	tccttgccgt	20820
ggggactgtg	ggcggaagcc	agcgggatga	ccgggcacct	ggacgcggcc	gacctcggcc	20880
gggtggcccg	ttccggcctc	accgcgatgc	ccaccgggga	cgggctggcg	ctgctcgaca	20940
ccgctcagcg	ggtggacgaa	gccaccctgg	tcacggccgc	gctggacacc	cgggcccctgc	21000
atgccggggc	cgacagcgcc	acgctgcggg	cgctgttcca	cgactcgtg	cccgtaccgc	21060
gcccgatccg	gacctcccc	gcggcccagg	ccggggggcc	ggatggactc	cgccagcggt	21120
tgctgggggt	ggtcaggggg	gagcgtcgag	cggcgctgct	gcatgtgggt	tggtgtcatg	21180
tcgcgggggt	gctggggcac	gcggaccgga	gcagcattga	ggagaccggg	cccttcaagg	21240
acacccggtt	cgactcattg	accgctgtgg	agctgcgcaa	tgctgtcac	ggtgcgaccg	21300
ggttgcgggt	gcccggccag	ctggcttctg	actaccggac	gcctgcagct	ctcaccgato	21360
accctotacga	cagacttctg	ggttcccggg	aggacgcggt	gctcgcccg	atcaccaggg	21420
ccgctacaga	cgagccgato	cgatcgtgg	ggatggcctg	ccgctatccg	ggcgggggtg	21480
agtcgccgga	ggacctgtgg	cagctggctg	ccgacggccg	tgacgccatc	tcgacttcc	21540
ccgcgcagcc	gggctggaac	gtcgagagcc	tcaccacccc	cgaccccgac	caccccgga	21600
ccagctacac	ccgtgcggga	ggcttctctg	acgacgcggc	ggacttcgac	ccggagtctt	21660
tcgggatctc	accgctgag	gcactggcca	ccgaccccca	gcagcgactg	ctgctggaga	21720

cgacgtggga	ggccttcgaa	cacgccgggg	tcggcccggc	gtcactgcgt	ggcagccgga	21780
ccggcgtctt	cgtcggcgtg	atgtacaacg	actacgcctc	gcgtatccgg	cacatcccag	21840
agagcgtoga	ggcgggtctg	accaccaaca	gcgcggggag	tgtagcgtcg	ggcgggtctt	21900
cgatcacgtt	cggtctggag	ggaccggccg	tcacggttga	taccggtgtg	tcgtctgcgc	21960
tggtggcgtt	gcatctggcc	gcgcaggcgt	tgcccaacgg	tgagtgactc	ctggctctgg	22020
cgggcgggtg	tgcggtgatg	tcacttctgt	ccacgtttgt	cgagticagc	cggcagcggg	22080
ggctggcagc	tgatggcggg	tgcaaaagct	tcggcgacgc	tcggcagggc	accggctggg	22140
gcgaaggcgt	cggtgtgctg	ctgggtggagc	gtttgtcgga	gcgcgcgcgc	aacgggcatc	22200
cggtgtctgg	ggtcgtttcg	ggcagtgtgt	tcaaccagga	cggggccagc	aatggcttga	22260
cggcgcccaa	tggtccttcg	cagcaacggg	tgatccaaca	ggcgtctggc	aatcgggggt	22320
tgccgggggg	ggatgtcgat	gccgtggagg	gcgcagggac	gggaaccggg	ctgggcgacc	22380
cgatcgaggg	gcgaagcgtt	atcgccacct	acggacaggc	ccggtcggcg	gaccggccgt	22440
tgtggctggg	ttcgctgaag	tccaacatcg	gtcacaccga	ggcgcgcgcg	ggcgtcggcg	22500
gcgtcatcaa	aatggtgcag	gcgatgcagc	acgggactct	gcgcgccacc	ctgcacatcg	22560
accagccccc	ggggcaggtc	gactgggcta	cggtgtcagt	ggagctgtgt	accgaggccg	22620
tgccctggcc	ggacagtgtg	cggcccccgc	gggtggctgt	ctctctgttc	ggtgtcagcg	22680
gtaccaacgc	ccacgtcatc	atcgaaacac	ccccacacac	ccccacacac	accgcacctt	22740
cccaatctct	ccaatccccc	caggccccgc	agactgtgca	ggcccatcgg	ccggtgcccgt	22800
ggctgctgtc	ggcgaagacc	tcgcaggccc	tgccgcgcga	ggcccgccgc	ctgtcagctc	22860
acttgcgagc	caaccccgat	ctgcgttcgg	ctgatgtggc	gcattccctg	ctcaccacgc	22920
ggctctgtca	cgccgagcgc	gccgtcttca	tcgccgggtg	ccgggatgag	gctcttggcg	22980
ccctggagcg	actggccgac	ggcacccctg	ccctccacct	cgctcagggc	cttgcgatcg	23040
tgagtggaac	gacgggttct	gtcttccccg	gtcagggttc	gcagtgggtg	ggtatggccg	23100
ttgagctgct	ggacggctcg	gaggttttct	ccgagcatat	ggccgcctgc	ggcaggggcc	23160
tggaacccaa	tgtagactgg	tccttgaggc	acgtcctacg	ccagacggag	ggtagctggc	23220
cactgggaacg	cgctgaagtg	gtccagcccc	tgctgtgggc	ggctcatgtt	tcgtctgcgg	23280
gactgttgca	ggcacatggc	gttagccctg	ctgcgggtgt	gggccaactc	caaggtgaga	23340
tcgctcggcg	ttcgttggcg	ggagccctga	gtctggaaga	cggagcccg	gtttcgcgc	23400
ttgcagcgca	agccatcgcc	gaaaccctcg	caggacacgg	cggaaatgct	tcaatcgccg	23460
ccccgcgcac	cgacatcgca	ccccgtatcg	cccgttgga	cgagcggatc	tcacatcgca	23520
cggtaaacgg	accgcattcg	gtgggtgtcg	caggagaccc	tgacgcgctc	gaggcactcc	23580
cgccgggaact	ggagaccctg	ggtctccgca	atcgtcgcat	cccgttcgac	tacgctccac	23640
acaccctcaa	cgctgaggcg	atccgtgaac	ggctcctggc	cgacctggca	gtgatccagc	23700
cacgtgcgcg	gagcattccc	gtgctgtcca	ccgtcaccgg	cgcatggctc	gacaccaccg	23760
tgatggactc	cgagtactgg	taccgcaacc	tacgtcagac	cgtggagtgc	gaagcagcca	23820
cccgcactct	cctcgaccag	gaccaccgct	acttcgtcga	galcagcccg	caccocgtac	23880
tcaccatcgg	tctacagcag	accatcgagg	aaaccaccgc	tcggcccgcc	accctctcca	23940
ccctcccgag	caacgaaggc	accctccggc	acctgttcac	ttccctcgcc	caggccccacg	24000
cccacggcct	gaccatcgac	tggacccccg	ccctcaccca	caccgagccc	cgcaccaccc	24060
ccctccccc	ctacccttcc	caacacgaac	gctactggct	ggaggacgga	gctccgaagt	24120
ccggggagct	ggcttcggcc	ggactcggct	cggcggaaca	tcgctgtctg	ggcgcgcgtg	24180
tgccgctgcc	cgattccggg	ggcttcctgt	tcaccggcca	gttgtcgtg	cggagtccac	24240
cctggttcgc	cgaccacgcg	gtacacggca	ccgtgctgct	gccggggacc	gcgttcgtgg	24300
aaactggcgt	ccaggccggt	ggcgtctctg	gctgcgggct	gctggaggaa	ctcaccctgg	24360
aggcaccgct	ggtgctcccg	gaaaacagct	ccgtccagct	ccaactcgtg	gtgaacgccc	24420
cggacgcoca	ggacgactcg	ggcggcagga	ccctcagcgt	gtactcgcgc	ccgcaggacc	24480
gtactcgga	cgcgccttgg	gtgcggcacg	ccaccggagt	ggtccggctc	ggaggcgccg	24540
cggagccgga	gggactgacc	gtgtggccgc	cgaccggagc	ggtcggggtg	ccggtcgagg	24600
acttttacc	ggtgctcggt	gaccgtggct	atgactacgg	acctgcgttc	cgtggggtaa	24660
gggcccgcgt	gcgccacggt	gaactgtgtg	atgccagggc	cgactggcc	gaggagcagc	24720
agtccgaacg	cgcgctgttc	cacctccacc	cggccctgct	cgactcggcg	ctgcacggga	24780
tgaggactgat	gccctcggcg	agcgcggagc	agaccgggct	gccgttcgcg	tgccgcgggtg	24840
tgacgctgca	tgccgtgggg	gcgtcggccc	ttcgggtgag	tcttagggcc	ggcggggccg	24900
acacgggtga	ggtctactgt	gccgatggcg	caggctcggc	ggtcgtctcg	gccgacgcac	24960
tggtggtccg	gccgctccga	caggagggaac	tggcggtctg	gcaggacgcg	taccgcgact	25020
ggctgtaccg	ggtcgactgg	cccaggttgc	cggagggtcc	cctggtggct	ccggccgggc	25080
catggggcgt	cctgggcggg	aaacggcgcg	ggatactcgg	caccgatggc	tcggccgggt	25140
tgctggccgg	ggtcccgatc	gacgcctatc	gggacctggc	ggagctcgcg	gaccggacgg	25200
gcccggagcag	cgcgttcccg	accgtgggtg	tcgcgcgggt	gcccacggga	accggtgccc	25260
cggcggagcg	ggtgcggggg	gtgacgtacc	agggtcgtga	catgatccag	tcatggctcg	25320
ccgacgatcg	ttccgcctcg	tcgacccttc	tcctggtgac	cgcggcgccg	gtgtccaccg	25380
gcttoggggg	cgacctgggt	gatctggggc	aggcgccggt	atgggggttg	gtaggggccg	25440
cgcatcgga	gaacccggac	cgcttcgtcc	tcctcgacct	cgacgggagc	gagccggtcg	25500
ggcctctccc	gacggcgccg	ctgctctcgg	gggagccgca	actggcgctc	cggaggggca	25560

aggtgtgac	cgcccggtg	gaccgggtgt	cgcccgacgc	gggaacgctg	ctgcccggcg	25620
ccggcgccga	ccgtggcga	ctcgacgtca	ccagccgggg	ccgctcgac	aacctcgccg	25680
tcctcgccgc	gcccaggtg	tcggcgccgc	tcgcccaggg	acaggtccgg	gtccgggtgc	25740
acgcccggcg	cctgaattc	cgcatgtgc	tgctcgctct	ggcatgtac	ccgggtgagg	25800
gttcgatggg	cagcgaaggc	gcccgggtgg	tgctggaggt	cgggcccgcc	gttgagcgcc	25860
tgccccgggg	cgaccgggtg	atgggcatgc	tcggggcgcc	cttcttcggg	ccggtcgccg	25920
taaccgacca	gcgcatgtg	accaagcttc	cgagcgctg	gtcgttcacc	ggggcgccat	25980
cggtaccgat	cgcttctct	accgctact	acggactgtt	cgacctgggc	ggcctgcgcg	26040
ccggccagtc	gctcgtgtg	catgcggoga	ccggtgtgtt	gggaatggcg	gctacgcagc	26100
tgccccggca	cctcgccgct	gagggtgttc	gcacggcgag	ccccggcaag	tgggaggcgc	26160
tgccggggat	gggattggac	gaggagcaca	tcgctctgtc	gcgggacctg	gacttcgaga	26220
agaagtcttc	ggcccgccgc	gttgcccgcc	gtgtcgacgt	ggtcgtgaac	tcgctggccc	26280
gggagttcgt	ggacgcgtcg	ctgcccgtgc	tgccgcggcg	cggtcgattc	gtggagatgg	26340
acaagaccga	catccgtgac	gcggaggcgg	ttgcccgccg	gcaccccgcc	gtcgtctacc	26400
gggcccctcga	cctcgtggac	gcccggggcg	cgagccgtat	ccaggagatg	ctggccgagt	26460
tgctcgccgt	cttcgaggcg	gggggtgatc	agccgctgcc	gctgacgacc	tgggacatcc	26520
ggcgtgcccc	ggaggcgctg	cgccacctga	gccaggcacg	gcacatcgcc	aagatgttct	26580
tcacctgtgc	gcccggcccg	gaccgggacg	gtacgttctt	gatcacgggt	gtgcccggcg	26640
cgctggggcaa	cctggtcgcc	gcccattctg	tgaccgaggg	tgccatacgg	aacctgtcgc	26700
tcgtcagccg	ccggggggcg	gcggcccccg	gcggcgaggg	gctggccacc	gagctggccg	26760
ggctggggcg	gacgggtgac	ctggcgccct	gtgactggcg	cgaccggcag	gcccgtggcg	26820
ggctgtctgc	cgacatcccg	cgggagcatc	cgctgacggg	tgtgtgtcac	gcccggcggt	26880
tgctggacga	cgggatcgtg	gcacccctga	cccgccgaac	gctggacgcg	gtctaccgcc	26940
ccaaggtgga	gcccgcctgg	aacctgcacg	agctgaccac	ggacagcgcc	ctggcccgct	27000
tcgtactgtt	ctctcggcgc	gcccgcgacg	tcggcagcgc	aggccaggcg	aactatgcag	27060
cgcccaaacg	cttctcgcac	gccttggccc	aattccgcga	ggcccaggcg	ctggcgccca	27120
gctccctcgg	ctgggggattc	tgggcccgaga	gcggtgagat	gaccggtcac	ctgggggcct	27180
ccgacctggc	acggatggca	cgttcgggca	tcgcccgcct	gacggtcgag	caggccctgg	27240
ccctgttcga	ctccgacggc	tcgggtgtct	gtgcgtcagt	gctgcccgtg	cggctggaac	27300
tcaccggggc	cggtgcggcg	gcccgggtcg	gaacgggtgc	ggcgctgatg	cggggcgctg	27360
tgccggcgacc	ggcccggcgg	gtggtggaaa	caaccacggg	cggtgcgcgc	acaggccctg	27420
gccaacggct	ggcgcccgct	tcggggcgcc	accgcgacgc	cgccctccaa	gagctgtgtt	27480
gctcgcatgc	ggccaccgtg	ctggggcaca	gccgttccgg	atcggtgccc	gcgcagcggg	27540
cgctcaaggga	gctcgggcttc	gattcgtcga	cagccgtcga	gttcgcgaac	cggctcaacg	27600
tgccgacggg	cctcgggctc	cccgcgactc	tggtgttcga	ccaccggacc	ccgctggcga	27660
tgccgggaaca	gctcgggaag	gagctgttgc	cgagacagat	cccggtggcg	ccgcaggttt	27720
tgagggaact	ggaccgtctg	gaggcgccgt	tcgcccgtct	ctccggcgcc	gacctccagc	27780
agtccgggagc	cgccgcacgg	ctgagggcac	tgctgaggcg	gacgggcacc	gtcactccgg	27840
cgggaggggga	cgctgcccgc	ggcctcggcg	tagagctcga	aacagccacc	cacgacgaga	27900
tccttcgccct	tatcgacgag	gaggtagggg	acgtgtgacc	ggctggctgc	tcgccctcca	27960
ccctaccctc	ccgcccggac	tagcagcatg	gatgagatca	cgatgactga	tgagaccgct	28020
gtgcccacaaa	cagagaccac	cgaggagaag	ctcttctcct	acctgaagaa	ggccacctcc	28080
gaactccagc	agagccggcg	ccgggtggca	gagctggagg	cgccggaggc	ggagcccatc	28140
gcgatctggg	gcacggccctg	ccggtaccgc	ggtggagtac	gttcccggga	ggacctgtgg	28200
cggttggtcg	cgagggggca	gcacgcgatc	tcagcttcc	cgacggaccg	cggctgggat	28260
ctcgaagacc	tcacgaccc	ggaccgggac	cgcccgggca	agtccctacg	ccgggacggc	28320
ggcttctcgc	acggtgcgcg	ccagttogac	gcgggttct	tcgggatctc	gccacgtgag	28380
gcgctggcca	tggaaccgca	gcagaggctg	ctgctcgaga	cgacgtggga	ggtcttcgag	28440
cgcgccggga	tcgaccgcac	atcgctccgt	ggcagccggga	ccgggtgttt	cgccggcatc	28500
agccaccagg	actacgctgc	cgacacgcgc	ccgtcgcccg	aggtctcga	ggggcacctg	28560
atgaccggca	cccggttcag	ctggtgttcc	ggcggggtcg	cctatgcctt	cggcctggaa	28620
gggcccggca	tgacgggtga	cacggccctgc	tcctcgtcgc	tggtggcggt	gcacctggcc	28680
ggcagggcgt	tgcccaatgg	tgagtgcacg	ctggcggtgg	ccggcgcgct	cacctgcatt	28740
ggcacggccg	gcgccttcac	caggttcacg	cgggagcggg	gcctggcccc	ggacggggcg	28800
tgcaaggccct	tcagctcgga	cgccgacggc	accggttcca	gcgagggtgt	gggtgtgctg	28860
ctgtgggagc	gtttgtcgga	cgccgcggcg	aacgggcacg	cggctgctgc	ggctgtttcg	28920
ggcagtgctg	tcaaccagga	cgggggccagc	aatggtctga	cgccgcccaa	tggtccttcg	28980
cagcaacggg	tgatccaaca	ggcgctggcc	aatgcggggg	tgccggggcg	ggatgtogat	29040
gcccgtggag	gcacggcgac	gggaaccccg	ctgggtgacc	cgatcgagcg	gcaggcggtg	29100
atcgcgacgt	atggacaggc	ccggtcggcg	gaccggccgt	tgtggtggg	ttcgctgaag	29160
tcacacatcg	gccacaccca	ggccgcccgc	gggtcgcgcg	gogtcatcaa	gatgatccag	29220
gccaatgggtc	acgggacgct	gccccgtacg	ctgatgttca	accagccctc	gccccaggct	29280
gactggggcg	caggcgcggt	ggagctactg	accgaagcca	tgccctggcc	cgggggtgac	29340
cggcccccgc	gggcccgaat	ctcctcttcc	ggaatcagcg	gtaccaacgc	ccacgtcatc	29400

atcgaacagg	gggccccgc	acggacagcg	tccgaccccg	gtgaaagtcg	tgctgacgag	29460
cccggcggtac	ggggcggggc	tcccgccct	gccaccacgg	agtcggccac	cgaaccgcag	29520
ccggtccot	ggctgtgtc	cgggcacagc	gcgacccgc	tgccggcgca	ggcggtatcg	29580
ttgaagtcgt	acggggccaa	caacaccggc	atccgtccgg	ccgacatcgg	cttctcgtcg	29640
gtcaccaccc	gggcccgcgt	ggaacaccgc	gctgtcgtcg	tggcagccga	ccatgccgggt	29700
ttcacggctg	gtctcgacgc	gctggccgag	ggccggacag	ctcccgaggt	ggtgagcggga	29760
acggctgtcg	ccggtgcccc	gagcgcgttc	ctctctcccg	gtcaggggctc	gcagcgggtc	29820
ggcatggggc	gcgagctcca	gcaggcggttc	ccggttttcg	ccgagggttt	cgaagcagtc	29880
tgcccccagg	tcgaccgcta	cctggagcac	ccacttctcg	atgtcgtact	cggccgcggc	29940
gcagcgcact	tcggcgcgtt	gctccatcag	accgcctaca	cgcagccggc	actgttcgcc	30000
ctcgaagtgg	ccctgttccg	gctggtcgaa	tcctgggggtg	tcaggccgga	ttacgttgcc	30060
gggcatttcg	tcggtgagat	cggcgccggc	catgtggcgg	gggtgttctc	gctggaggat	30120
cgccgtctgc	tgggtgggac	gcgcggacag	ttgatgcagg	cgttgccggc	tgaaggcggc	30180
atggatggcg	tcagggtgtc	cgaggacggg	gtccgtccgt	ccctgactcc	ttgctgggag	30240
caggaccggg	tggatgtcgc	ggcggtcaac	ggcgagcat	ccacagtgtt	gtcggcgcat	30300
gaggaggcgg	tcctggcggt	tgccgagcac	tggcaggcgc	ggggccgcaa	ggttcgtcgg	30360
ctcactgtca	gccatgcctt	ccactcacct	cgtatggacc	cgatgtcga	ccagtctcgt	30420
gtggtgttgg	agggtatccg	tttcgaggag	ccggccatcc	cggctcgtctc	cagcgtcacc	30480
ggtcgtcttg	ccgagccggg	gcagttgacc	actgcggact	actgggtcgc	ccacgtccgt	30540
caaacggctc	gcttccacga	cggcctccag	accctccaga	ccgagaatgt	gaccgcgttt	30600
ctggagatcg	gtcccagagg	gcaactctcg	gcaatgacct	gcgacttctt	gaccgatacc	30660
ggggccacag	ccgcgtcgc	accctctcgt	cggcgcgaac	gtcccagggc	accagcgcgc	30720
ctcaccgcaa	tcgcccggct	gcacaccac	ggcgtctcga	tcgactggcg	cacgtacttc	30780
accagaccga	gcaccagcac	cagcaccagc	accggtacgc	gtaccggtag	ggggcaggcc	30840
actgcgcaga	cgcccgctca	gctgcccacg	facgccttcc	agcaccagtc	cttctggctc	30900
ggcccccagg	ccctgttcgg	cgacgtcagc	accgcggggc	tcacctcgcc	cgaccacccc	30960
ctgctcagcg	cagccaccac	caccgctgtc	gacggcagcc	tcctgtctac	cggcaggctg	31020
tcgcagcggt	cgcccgcgtg	gatcggcgac	caccgcctcg	gcggtgttgt	cctgctccca	31080
ggcacgcctc	tcgtggaact	cgtcgtacgc	gcccggggacc	aggccgggtg	cagccgcctc	31140
gagcaactca	tcagtctcac	gcgcgtgacg	ctgcccagagc	atggtgcctg	gcggtaccag	31200
gtcgcgtctg	gcggcccggc	ccacgacggc	cgcgcggggc	tgccatccca	ctccagcacc	31260
tcggacacga	ccggcgacga	acagtgagcc	ctcaacccca	gcggtctgct	caccgtcgag	31320
atgaccgctc	cgcccgccga	ttccacccc	tgcgcggcg	agcaccgcac	ccgcataccg	31380
ctcgacggcc	ttacgacggc	gctcggccga	agcggtacgc	gatacggccc	ggtcttcag	31440
ggcctgcggc	ctgctgggac	actcggcgac	gacacctacg	ccgaggtoga	gatcccccgc	31500
ggcgaccaga	ccgacaccca	ccgctacgaa	ctccaccocg	cgtcctcga	cgccgcgctg	31560
cagcgctcct	ccctccaggc	cgacgaggcc	ggggccggggc	agctcgtgcc	gttcgccttg	31620
accgggggtg	cgtgtacgc	ggccggcgcc	tcggccctgc	tcgtcaaggt	gtccgttacc	31680
ggtccggaca	ccatggcgct	gctcgtggcc	gacaccgagg	gcccaccggt	cgccaccgtc	31740
gactcactga	ctgtccggcc	gatggccatc	gaccagaccg	cccgagcacc	cagccaccct	31800
gacgcgtgtg	tcaccgtggg	gctggagttg	gcccgaagcc	gggaggcgaa	ccgaccatc	31860
ccctctgccc	actgcggcat	gctggctccg	gacgaaccgg	acctcacctc	cggcccggcc	31920
tgcccccgggt	cctccgcgca	gcggtacggc	ggcctcggcg	cgtcgtcga	gatctgcgga	31980
acggacgggg	cggtagcctg	cgtggtactg	gcgccccttc	tccccgccga	tgccgcggcc	32040
cgcgacaccg	ccgcgcggac	gcacgcgacg	acgcgcggcg	ccgcgcgtct	catcaagggc	32100
tggtcggcgc	acgaccgttt	caccgactcg	cgtctggtct	tcgtcaccgg	tgcccggttg	32160
gcccaccagcg	gcggggacga	actgcacgac	ctgggaacct	ccaaggtctg	gggtctgggtc	32220
cggctcggcc	agaccgagaa	ccccggcagg	ttcgcgtcgc	tcgatctcga	cgaccgggac	32280
accgtcaccc	aactgcggga	agccatcctg	gcccgtcagg	cacagctggt	cctgcggggac	32340
gggcggcgtg	gaaacctccg	gctggccaag	ggcgtcgcga	tacaggatcc	cgaccggggt	32400
tgggtgtgtg	acgggtaccgt	gttgatcact	ggcgggacgc	gtgtgctggg	tgggtcgtgc	32460
gcccgtcatc	tggctgcggg	gcctgggggtc	cgtcgtctgt	tgctctgcag	caggcggggc	32520
cctgatgcgc	cgggtgcgggt	ggagctgggtc	gcccagctca	ccgctctggg	tgccgatgtc	32580
accgttgccc	cctgcgacgc	ggctgacggg	gatgcgctgg	ccgcgctctt	ggaccaccgt	32640
ccgcgcacgc	accctctgac	tgggtctgtg	cataccgctg	gtgtcatcga	tgaccgcact	32700
gtcaccaccc	tcactcccga	gcgcctcagc	gcggtcctac	gcccgaaggt	cgaccgcggc	32760
ctcaacctcc	atcagctgac	ggcgcatctc	ggcttgaccg	gctttgtgct	cttctcctcc	32820
gcccgcgggc	ttctcggcgg	cgcggggcag	ggttaactacg	cggccggcaa	cgccttcttc	32880
gacgcactgg	cgcagctgcg	gaagcggcag	ggactgcggg	gcgtgtcgtc	ggcctgggggt	32940
gcctgggttc	aggacggcgg	aatgaccgca	acgctggacg	cggcgagcgt	cgagcgggat	33000
gcgcgcggcg	gtgtgctgcc	gctcagccac	gagcaggggc	tgaacctgtt	cgacctggca	33060
gtggcagggt	ccgagccgct	ggtggcaccg	atcgggctgg	acaccaccgc	gctcgcggag	33120
tcgggtgcca	cgtgcggga	gatgctgcgc	gggttggtgc	gtgagcggtc	acgcgcggcg	33180
gtcggaccct	cgcacacgac	gtcgcggcg	atggcgctgg	aacaacggtt	gtcgggggtg	33240

gtcagggggg	agcgtcagc	ggcgtcgtc	gatttgggt	gtggcatgt	cgcaggggt	33300
ctggggcagc	cggaccgag	cagcattgag	gagaccggc	ccttcaagga	caccggcttc	33360
gactcattga	ccgtctgga	gctgcgcaat	gtgctgcacg	gtgcgacgg	gttcgggtg	33420
ccggccacgc	tggtcttoga	ctaccggacg	cctgcagctc	tcaccgatca	cctctacgac	33480
gagcttctgg	gttcccgoga	ggacgocgtg	ctcggcccga	tcaccagggc	cgcgtacgac	33540
gagccgatcg	ccatcgtagc	gatgtcctgc	cggtaaccgg	cgggtgtctg	cactccggag	33600
gaactgtggc	ggctgggtgc	cggggggcgg	gacacgatca	cggacttccc	ggacgaccgc	33660
ggctgggata	tcgacgacct	gtatgacccc	gacccgggoc	accocggcac	ctcctacacc	33720
cggcggggcg	gcttctgtgc	cgaacggggc	ggtttcgatc	cgaagtcttt	cgggatctcc	33780
cccccgagg	cgtcggccat	ggaccggcag	cagcggctgc	tgctcgaaat	gacgtggggg	33840
atgttcgaac	ggcgctcat	cgacccaaca	acgtgaagg	gcagccaggc	cggggtgttc	33900
atcgccaccg	ccggccccgg	ctacggcggc	cgcattccacc	acgagtgcga	gggctogag	33960
ggccagcagc	gtttcggcgg	ctcggccggc	gtgacctcag	cggcgatctc	gtacacgttc	34020
ggctgggaag	ggccggcgat	gacgggtggc	accatgtgct	cgtcctcgct	ggtggccctg	34080
caactggcgg	tcacgtccct	gcgcaacggc	gagtcctcga	tggcgtctgc	cggcggggtc	34140
acgggtgatgt	ccggccggcg	cgcgttcacc	gagttcagcc	ggcagcgggg	gctgtccccc	34200
gacgtggcgt	gcaagtctgt	cggcgacggc	ggcgacggca	cgggttgggg	cgaggcgccc	34260
ggcgtgctcc	tcctcgagcg	gctctccgac	ggcgtcgca	acggccaccc	ggtgctggcc	34320
gtcatccggc	gcagcgccgt	caaacaggac	ggcgccagca	acggcctcac	ggcaccacac	34380
ggccccctgc	agcaacgcgt	catccggcag	ggcctggcga	acggttccct	gtcccgggcc	34440
gacgtggagc	ccgtcgaggg	ccacggcacc	gggaccccc	tggcgagccc	gatcgagggg	34500
agggccctga	tcgccacota	cggacaggac	cgccccggcg	accggccgct	cgggctgggc	34560
tcggtgaagt	ccaacatcgc	ccacgcgcag	ggcgacggcg	cagtggcggg	cgtcatcaag	34620
atggctcagg	cgtacggcca	cggcctcctc	ccgaagaccc	tgacgtgga	gcagccctcc	34680
cgcctggcgt	actggtccgc	cggtctgggt	gagctgctca	ccgaggcgat	gcggtggcgg	34740
gagaccgacc	aaccccgggc	ggcgggtgtc	tcggcggtcg	gcggcagcgg	caccaacgcc	34800
cacatgatca	tcgagcaggc	ggccgcggcg	gacgaggagg	acacgcaggg	cacgagcagg	34860
accagcgagg	agagcgcgcg	cgaacaggcc	aggccgctgc	cgatgggtgc	ctggctgctg	34920
tcggcggaaga	cctcgcaggc	cctggcggcg	caggccggcg	gcctgtcagc	tcacttgcca	34980
ggcaaccccc	atctgcgttc	ggctgatgtg	ggcattcccc	tgctaccacc	gcggtctgtc	35040
cacgcggagc	gcgcgctctt	catcgcgggt	gacggggatg	aggctcttgc	cgcctgggac	35100
gcactggcgg	acggcacccc	tgccccctac	ctcgttcagg	gccttgccga	tgtgagtggc	35160
aagacgggtg	tcgtcttccc	cgttcagggt	tcgcagtggt	tgggtatggc	cgttgagctg	35220
ctggacggct	cggaggtttt	cgccgagcat	atggccgcct	gcgccagggc	cctggaaccg	35280
tttgtggact	ggtccctgga	ggacgttcta	ggccagacgg	acggctacgt	ggcactggaa	35340
cgcgtcgagg	tgttcagacc	cgtgctgtgg	gcggctatgg	tctcgtcgcg	gggactgtgg	35400
caggacacatg	gcgttgagcc	tgctgcgggt	ctgggcccact	cccaagggtg	gatcgtcgcg	35460
gcttgcgtgg	cgggagccct	gagtcctgga	gacggagccc	gcgttgtcgc	gcttcgcagc	35520
caagccatcg	ccgaaccctt	cgcaggacac	ggcggaatgc	tctcaatcgc	cgcgcccgcc	35580
accgacatcg	caccocctgat	cgcccgctgg	aacgagcgga	tctccatcgc	cacggccaac	35640
ggaccgcatt	cgggtgggtg	cgcaggagac	octgacgcgc	tcgaggcact	cgcggcgcaa	35700
ctggagagccc	gtggtctccg	caatcgtcgc	atcccggctg	actacgcctc	acacacccct	35760
cacgtcgagg	cgtaccgtga	acggctcctg	ggcgacctgg	cagtgatcca	ggcactgtcc	35820
gcgagcattc	ccgtgctgtc	caccgtcacc	ggcgcatggc	tcgacaccac	cgtgatggac	35880
ggcgagttact	ggtaccgcaa	cctacgtcag	accgtggagt	tcgaagcagc	caccgcgact	35940
ctcctcgacc	aggaccaccg	ctacttcgtc	gagatcagcc	ggcaccocgt	actctcggcg	36000
atggtcccg	actgcctgga	caccagccgc	ccggtgggtca	cggcaccac	cctccgacgt	36060
gaccgtaccg	atgccactgc	cgcgttgacg	gacttgccgg	aagcgacgg	gcacgggggt	36120
ccggtcgact	ggcgctcgtc	cttcgcggcg	tcgaccggcc	ggcggttcca	cctgcggacg	36180
taccoccttc	agcggaaca	ctactgggtg	gattccggta	cggcgacgag	tgacatgagc	36240
acggccggac	tggcgtcccc	cgtatcccg	ctgttgggag	cgtgacgac	ggtggccggc	36300
gaggacggcc	acctcttcac	cggccgggtg	tcggtacgga	gcacccatg	gctggccgac	36360
caccagatca	ccggttcggt	cctgttgccg	ggcacggcct	tcgtogaact	ggcctgcggg	36420
ggcggggacc	aggccggctg	cggcggggtc	gaggagctga	cgtgctgagc	tccgctcgta	36480
ctgcccgaag	aggcgacggt	cagggtccag	atgaagggtg	gggagcccca	gcgcacgggc	36540
gcgcgcacca	tcgaggtgta	ctcctcggac	cagcaggccc	ccggccggga	acgctgggtc	36600
ctcaacgcga	gggggatgct	tgcggcgaaa	ccggtggagg	ccccgcggag	tctcaccacc	36660
tggcccccg	aaggcgtgtg	ccccgttcgg	ctggacggct	tccacgaccg	cgtggcgcca	36720
cgcggcttcg	gctacggctc	gacattccgc	gggctgagcg	ccgcgtggtc	acgcggtgac	36780
gagatcttcg	ccgaagcggc	gctccctcgc	ggccatcggc	aggatgccgc	ccggttcgga	36840
ctccaccogg	cgtactcga	cgtgcctctg	cacgccatgg	aactccggga	accccgcccc	36900
gcggcgacg	gagtcgggtc	tccgttcggc	tgggaacggc	tctccctgca	cgcgtcgggt	36960
gcccgaagcg	tacggctcgc	cctcgcggcg	acggcgcccg	acgctctgtc	ggtgacctc	37020
gcccgatgca	tcggtcggcc	ggttgctcca	ggccgtcgc	tggccctcgc	ggagctctcg	37080

tccgacctgc	tgcccccggc	gtccgtctcg	tacggggact	cgctgttccg	caccgcttgg	37140
ataccgccc	tcgtcggccc	ggaggccggag	tccggggcgg	ggcgaccgtc	cgccggctgg	37200
cggtgtctgg	gccccgctcc	gctcggcgcg	gccaaacgccc	tgaacctcac	gggaacctcc	37260
tgctcctgct	atccggacct	ggcggcgctg	atccggcgcc	tcgacggcgg	agccgcgctg	37320
cccagggccg	tactcgcgcc	gtacgcggcg	gagccagccc	cggacgcggg	atctcccgcg	37380
gacgcgctac	ggccctcgac	cggccggggc	ctgcaactgc	tgcaatcctg	gctgtccgag	37440
gaccggttgg	agcgaagccg	gctgatcgtg	ctcaccgggg	ggcggttggc	gctcggctacg	37500
gacgaaggcg	tcaccgacct	ggtgagtgcg	toggccgggg	gtctggtccg	ttcggcgccag	37560
gcccagcacc	ctggcagggt	ctccctggtc	gacatcgacg	accggggagg	gtcctggggc	37620
gtcctgagcg	cggcgccggg	atccgggtgag	ccgcaggctc	ccctgcgctg	cggccagatg	37680
aagggtcccc	gctcgggctc	cgctgacggt	cccacgacgg	gtatgcctga	gatgcccgcg	37740
gtttgggggt	ttgacgggtac	cggtttgatc	actggcgggg	ccggtgtgct	gggtgggctc	37800
gtgcgcccgc	atctggttgc	cgggcctggg	gtccgtcggt	tggtgtctg	cagcaggcgg	37860
ggcctgtgat	cgccgggtgc	gggtggagctg	gtggccgagc	tcaccgctct	gggtgcggat	37920
gtcaccgttg	ccgcctgtga	tgccggccgac	cgggatgcgc	tgcccgctct	cttggacacc	37980
gttccgcca	cgcacccctct	gactggtgtc	gtgcataccg	ctggtgtcat	cgatgacgcc	38040
actgtaccca	ccctcactcc	cgagcgcatc	gacgcggtcc	tacgccccaa	ggtcgacgcc	38100
gcgctcaacc	tccatcagct	gacggcgcat	ctcggcttga	ccgccttgtg	gctcttctct	38160
tcggccgccc	ggctcttcgg	cggcgccggg	caggccaact	acgcggcgcc	caacgccttc	38220
ctcgacgcac	tgccccaaca	ccgcggggcc	aacggcctca	atgcccgctc	cctggcgttg	38280
ggactgtggg	cgggaaggcag	cgggatgacc	gggcaccttg	acgcggccga	cctcgccggc	38340
atggggcgtt	ccggcctcac	cgcgatcccc	accggggacg	ggctggcgct	gctcgacacc	38400
gcccagcggg	tggaagaaag	cacctgtgtc	acggccgcgc	tggaaccccg	ggccctgcat	38460
gcccggggcg	cagacggcac	gctgccggcg	ctgttccacg	cactcgtgcc	cgtaccgcgc	38520
cgatccggga	ccctcccggc	ggcccaggcc	gcggggccgg	atggactccg	ccagcgggtt	38580
tcggggcgtg	tcgtggggga	gcccggagcg	gcgctgctgg	atttggtgtg	tggtcatgtc	38640
gcgaggggtg	tggggcacgc	ggaccggagc	agcattgagg	agaacaaggg	cttcaaggac	38700
acggccttcg	actccttgag	cgcgggtggg	ttccgcaacc	ggctgcacgg	tgccaccggg	38760
ttcggtctgc	cggccacgct	ggtcttcgac	taccgcagcg	ctgcagctct	caccgatcac	38820
ctctacggcg	agcttctggg	ttccgcgag	gacgcgctgc	tcgcccgcg	caccaggccc	38880
gcgtacgacc	cgttggaact	cgactaccgc	acgcctgcag	ctctcacoga	tcacctctac	38940
gacgagcttc	tggtttcccg	cgaggacgcc	gtgctcgccc	cgatcacagg	ggccgcgtac	39000
gacgagccga	tcgcgatcgt	ggggatggcc	tgccgctatc	cgggcggggg	ggagtccccc	39060
gaggacacct	ggcagctggt	cgccgacggc	cgtgacgcca	tctccgaact	ccccgcgcac	39120
cggggcttga	acgtcgagag	cccttaccac	cccgaaccgg	accacccggg	caccagctac	39180
accctgtccc	gaggcttccc	gcacgacgcg	gcggacttgc	accgggagtt	cttcgggata	39240
tcaccgcgtg	aggcactggc	caccgacccc	cagcagcgac	tgctgctgga	aaccagctgg	39300
gaagccatgg	aacggggcgg	aatcaacccc	tcacccctga	agggaacccc	caccggcgctc	39360
ttcctcggcg	tcattgtacaa	cgactacggc	actgccatgc	agcaggcgcc	agaggcttcc	39420
gaggcccata	tgccagcggg	tagcgcgggg	agtggtggcg	cgggcggggg	ctcgtacacg	39480
ttcggtctgg	agggaccggc	cgtaacgggt	gataccgctg	gttcgtcgtc	gctggtggcg	39540
ttgcatctgg	ccgcgcaggc	gttcgcgaac	ggtgagtgcg	ctctggtctc	ggcggggcgg	39600
gttgcggtga	tgtccactcc	tgccacgttt	gtcaggttca	gcggcgagcg	ggggctggca	39660
gctgatgggg	ggtgcgaagg	cttcgcgggc	gctgcgagcg	gcaccggctg	gggcgaaggc	39720
gtcggtgtgc	tgctggttga	gcgtttgtcg	gacgcgcggc	gcacggcgga	tcgggtgctg	39780
cgggtcgttt	cgggcagtg	tgtcaaccag	gacggggcca	gcaatggtct	gacggcgccc	39840
aatggtcctt	cgcagcaacg	ggtgatccaa	caggcgctgg	ccaatgcggg	gttgccgggg	39900
gcggatgtcg	atgccttggg	ggcgacggcg	accgggaacc	ggctggcgga	cccgatcgag	39960
gcgcaagcgt	tgatgcccac	ctacggacag	gcccggctcg	cggaccggcc	gttggtgctg	40020
ggttcgctga	agttccacat	cggtcacacc	caggccggcg	cgggcgtcgc	cggcgtcctc	40080
aaaaatggtc	aggcgatgca	gcacgggact	ctgcggccca	ccctgcacat	cgaccagccc	40140
acggggccagg	tcgactgggc	tacgggtgca	gtggagctgc	tgaccgaggg	cgtgccctgg	40200
ccggacagtg	accggccccc	cgggtgggct	gtctcctcgt	tcggtgtcag	cggtaaccaac	40260
ccccacgtca	tcattgaaca	caccccacac	accccacaca	ccacccgcac	ctgcccaatc	40320
ctcccaatcc	ccccaggccc	cgagactgtg	gcaggcccat	cggccgggtg	gtggctgctg	40380
tcggcgaaag	cctcgcaggc	cctggccggc	caggccggcg	gcctgtcagc	tcacttgcca	40440
gccaaccccg	atctgcgttc	ggctgatgtg	gcgcattccc	tgctcaccac	gcggtctgtc	40500
cacgcccagg	gcgcggtctt	catcgccggt	gaccgggag	aggctcttgc	cgccctggac	40560
gcactggccg	acggcaccgc	tgcccccac	ctcgttcagg	gccttgccga	tgtgagtggc	40620
aagacgggtg	tcgtcttccc	cggtcagggt	tcgcagtggg	tggttatggc	cgttgagctg	40680
ctggacggct	cggaggtttt	cgcgagcat	atggccgcct	gcgccagggc	cctggaaccg	40740
tttgtgact	ggtcccttga	ggacttcta	cgccagacgg	acggtacgtg	gcccactgaa	40800
cgcgtcgaag	tggtccagcc	cgtgctgtgg	gcggtcattg	tctcgtcgc	gggactgtgg	40860
caggcacatg	gcgttgagcc	tgctcgggtg	ctggccact	cccaagtgga	gatcgtcgc	40920

gcttgcgtgg	cgaggagccct	gagtcctgaa	gacggagccc	gcgttgtcgc	gcttcgcagc	40980
caagccatcg	ccgaaccct	cgaggacac	ggcggaaatgc	tctcaatcgc	cgcccccgcc	41040
accgacatcg	caccccctgat	cgcccgttg	aacgagcgga	tctccatcgc	cacggtaaac	41100
ggaccgcatt	cggtggtggt	cgaggagac	cctgacgcgc	tcgaggcact	ccgcggcgaa	41160
ctggagaccg	gtgtctcctg	caatcgtcgc	atcccgtcgc	actacgcctc	acacaccctc	41220
cacgtcgagg	cgatccgtga	acggctcctg	gcccacctgg	cagtgcacca	gccacgtgcc	41280
gcgagcattc	ccgtgctgtc	caccgtcacc	ggcgcatggc	tcgacaccac	cgtgatggac	41340
gccgagtagt	ggtaccgcaa	cctacgtcag	accgtggagt	tcgaagcagc	caccgcact	41400
ctcctcgacc	aggaccaccg	ctaactcgtc	gagatcagcc	cgcaccocgt	actcaccatc	41460
ggtctacagc	agaccatcga	ggaaaccacc	gctccggccc	ggaccctctc	caccctcoga	41520
cgcaacgaag	gcaccctcgc	gcacctgttc	acttccctcg	cccaggccca	cgcccacggc	41580
ctgaccatcg	actggacccc	cgccctcacc	cacaccgagc	cccgcaccac	ccccctgccc	41640
acctaccctc	tccaacacga	acgtacttgg	ctggacacgg	cggagccgcc	tggtggcgag	41700
ggagccggga	ccgacaccgt	cgagagcggt	ttttgggacg	cgctcgaggg	cgaggagtgg	41760
cagacgttgg	ccgacacgct	cgccgttacc	gcccagcgcc	cgttcgactc	cgtgatgtcc	41820
gcccgtcgtg	cctggcgggt	ccgacagcgt	gagcagtcct	tggtggacgg	ctggcggttac	41880
cggtatcgat	ggaagccgtt	ccgcgccccc	gtgtcggcac	cggattccgt	gtcggggacc	41940
tggtgggtgg	tcgttccccg	ccatgccggc	gacgcggacc	gggagaggcg	gcaagccgtg	42000
cgggggcacgc	tgaggtcctc	cgccgttgcg	cggacgatcc	tggtcggcgt	ggaccggccc	42060
gcccagcacc	gggggtcgtc	cgaactgaaa	ctcaggagac	ccgcgaccca	ggcgggtccc	42120
ccggccgggg	tgctgtccct	gctggccacc	gacgaacgtc	ccctcccccg	gcatgatgtg	42180
gtgcccgggg	ggctggcagc	caaccctggc	ctcgtccagg	cactggcgca	cgccgagatc	42240
gatgccccc	tcgtgggtgg	cacctgcggg	gcgtctctcg	ccggccggtc	cgaccggcgt	42300
gcgaaccocg	ggcagggccc	ggtctggggg	ctcggacggg	tggtccgccc	ggagcacccc	42360
gaacgctggg	gcggtctgat	cgatctgccc	gtggtcctcg	acccgcgcgc	tggtgaacgg	42420
ctggtgacag	tacttgcgcg	gtcgggcgag	gaggaccagc	tcgcctagcg	ggcgtcgggg	42480
gtcctcgtgc	gcaggctcgt	gcggttacc	gcacgccaa	tgccggacgg	cgtgcagtgg	42540
aagcccgagg	ggacggtcct	ggtgaccggt	gggaccggag	cgctggcgcg	ggaggtcgcg	42600
cggtggcctg	ctcatggcgg	cgccgaacac	ctggtgctga	ccagccctcg	cgccggctcg	42660
cgcccccgtg	cgcccgagct	gacggacgag	ctgcttgcgc	tcgggacgga	agtgacactg	42720
gcccgcctgt	acatggcaga	ccgggacgcg	gtcgcgcgcg	tgctcgcoga	gcacgcgcgc	42780
agctcgtgtg	tgccacacgc	cgccgtcctc	gacgacgggt	tactggacag	cctggaccgc	42840
ggcgccgctg	agtcggttct	gctgccgaag	tgggccgcgc	ctcggcacct	gcacgagtgg	42900
acgaagggag	cgaacgtgtc	ggcattcgtg	ttgttctcgt	ccgcccgagg	cgtgctcgcc	42960
agcgccaggc	aggccaacta	cgacgcgcgc	aacgcctacc	tggaacccct	ggccgaacag	43020
cgagggccgc	atggactggt	cgcccatcgc	atcgccctgg	ggcgttggga	cgccggtggg	43080
ctggccgctg	gcgacacgct	ggtcgaggaa	cgctcgcgcc	acggaggagt	ggtcccatg	43140
cgcccccagc	tgccgatcac	ggcgtccag	cagacgttgg	acggggcgga	gaccgcgggtg	43200
gtcatcgctg	acgttgactg	gcccgcgtac	ctcaccgcgg	tcacaccgcg	cccatggctg	43260
gcggacctcg	cgagagtgcc	ccaggccctt	aacgccgacg	acgcggctgg	tgcccttgc	43320
ggcacagccg	ggcagggtcc	gtccccgcgt	gcccagcgtc	tctccggcgc	cccggcaccc	43380
gagcagcgcc	gactggtgct	cgacctggtc	cgtacgaacg	tgccggcggt	gctcggccac	43440
ccgggtgcgg	agtcgatcga	gtccggccgg	gccttccgcg	agctgggctt	cgactctctg	43500
accgcgctcg	agctgcgcaa	caggctggct	gcggccacgc	ggctgcggct	gcccaccacc	43560
ctggtgttgc	actaccgcag	cgctgccgtg	ctcggcgatc	acctgtacgc	gcaggcgatc	43620
ggttcggagc	aggggcccg	ggcggatctg	tcctccggcg	ccgatccggc	ggccggaccg	43680
gacgacggag	ccatcgccat	cggtgcgatg	agctgcggtt	tccccgcgcg	tgtctctctc	43740
ccggaggagc	tgtaggcagc	gctgctggcc	ggtgaggaca	cgattaccgg	gttcccgagc	43800
gaccgggact	gggatgtcga	cgccctgtac	gaaccggacc	cggaccaccc	ggggaccacg	43860
tattcccgca	gogggcggtt	cctgtccgac	cgccggcggt	tcgacgcgac	gctgttcggg	43920
atctcccgcc	tgtaggcgct	ggccatggac	ccgcagcagc	ggctgctgct	ggagaccgca	43980
tgggagggtg	tcgagcgggc	gggcatcgat	cccacctcgg	tacgtggcag	ccgggcccgc	44040
gttttcgtcg	ggaccaacgg	ccaggactac	gcccgcctatg	tgcccacgga	accgatcgcc	44100
gtggaggggg	atctgctggc	gggcaatgcg	gccagcgtea	tctccggcgc	tctgtcgtac	44160
acgtttggic	tgagggggccc	ggccgtcacg	gtggacaccg	cgtgttcgtc	ctcgtcgttc	44220
ggcctgcacc	tcgccctcca	ggccttgcgc	aacggcgaa	gctccatagc	cctggcgggg	44280
ggcgtgtcgc	tgatgtccac	ccggcgccgc	ttcgtggaat	tcagccggca	gcgggggctg	44340
gcggctgacg	ggcgttgcaa	ggcgttcgcg	gacgcggcgg	acggcacccg	ctggggcgag	44400
ggggttggcg	tgctcctcgt	ggagcgtctg	tcgcagcgcg	gcccgaacgg	tcaccgggtg	44460
ctggccgctcg	tacgcggcag	cgccgttaac	caggacggcg	ccagcaacgg	cctcacggcg	44520
cccaacggac	cctgcagca	acgcgtcctc	cgccaggcac	tcgttgacgc	cgccgtgacc	44580
ggtagcgaca	tcgacgcgct	cgaagccacc	ggcacgggga	cccggttggg	tgaccgcatc	44640
gaggcgagg	ccctgatcgc	cacctacggt	caggacccgc	cgccgaacgg	gcccctgtgg	44700
ctgggctcgg	tcaaatccaa	catgcacac	acgcaggccg	ccggggcgct	cgccggcgctc	44760

atcaagatgg	tccaggcgat	ccgcatggc	gtacttccca	agaccctgca	cgtaggacgg	44820
ccgaccagcc	acgtcgactg	ggaggcaggc	gcggtggagt	tgctgaccga	ggccatgccc	44880
tgcccgagga	ccgaccggcc	gcgtcggggc	ggcatctctt	ccttcggcgt	cagcggcacc	44940
aacgcacaca	ccatcgtgga	gcaggcacct	cgccgggaag	acgagccgga	aacggggcca	45000
ccgcccgatg	ctccggccac	ggtggtgccc	tggtgctct	ccgctgccac	cgaggacgcg	45060
ctcgagagcg	aggccgcacg	cctcgccacg	tacctcgacg	agcggcccgga	gccaagcccg	45120
gcgacatcgg	ggtcctccct	ggtcaccacg	cgtagcagcc	ttgaccacgg	ggcgggtgtg	45180
ctcgggtagg	accgcgacgc	tctcgggggc	gggctgggtc	tgctggcgaa	cgggaagtcc	45240
gggtcccgctg	tcgtccgtgg	cctcgccagg	cccgacaga	aggtggcgtt	cctgttcacc	45300
gggcagggca	gcccagcact	gggcatgggc	aggagctccc	atcgccacct	ggcgggtttc	45360
cggcagttct	tcgacgaggc	gtgcgcggcg	ctcgacgcac	acctgcccgt	accgatagcg	45420
gcccgcctgt	tcgcgcaggc	ggatggggcg	gatcgggggc	tgatcgatgg	gacggaattc	45480
gcgcagccgg	cgttgttcgc	gctggagggtg	gcgttgtgcc	ggacgttggga	gttctgcggt	45540
gtcaggccgg	tttacgtttg	cgggcattcg	gtcgggtgaga	tcgcccggcg	ccatgtggcg	45600
gggggtttct	cgtcggagga	tgccgctcgt	ctggtgtgtg	cgccgggaca	gttgatcgag	45660
gcgttgccgg	cgggtgggtg	gatggtcgcg	ctccaggtgt	cgaagacga	cctcctgcca	45720
tccttgactc	cttggttggg	gcaggaccgg	ctgggtatcg	cggcggtcaa	cggcgacaga	45780
ttccacagtg	tgctggcgga	tgaggaggcg	gtcctggcgg	ttgccgagca	ctggcaggcg	45840
cggggccgca	aggttcgtcg	gctcactgtc	agccatgcct	tcactcacc	tcgtatggac	45900
ccgatgctcg	accagttccg	tgtggtcgtg	gagggtatcc	gttcgcggga	gccggccatc	45960
ccggtgtgtc	ccagcgtcac	cgttcgtctt	ggcgagcccg	ggcagttgac	cactgcggac	46020
tactgggtgc	gccacgtccg	tcaaacggtc	cgcttccacg	acgccctcca	gacctccag	46080
accgagatg	tgaccgcgtt	tctggagatc	ggtcccagac	ggcaactctc	ggcaatggcc	46140
caggagacgc	tcaccgcccc	ggtccatacc	atccccacc	tcgaaagaa	cgggtctgag	46200
accacggcct	tgctcacgcg	actggcgcaa	ctccacacca	ccggcacctg	ccccgactgg	46260
accgcttacc	tcaaacacca	ccccacacc	tccacaccgc	tgcccactca	cccttccaa	46320
caccacact	actggtatga	cggcggtacc	caggccaccg	atgtcagctc	cgccggcctg	46380
tcaggagcca	accacccgct	gctggggggc	gcggtcccgc	tgccgggtgg	ggaggcccac	46440
ctgttaccgc	gcccgctgtc	ggtgoggacc	caccgctggc	tgcccgacca	ccaggtcggc	46500
agcaccgtcg	tggtccggcg	cactgccttc	gtcgaactgg	cggtagggcg	cggtagaccag	46560
gtcggctcgc	gccacgtgga	ggagctgacg	ctggaagcgc	cgctcgtgct	gcccagagac	46620
ggcgccgtac	agatacagct	ccggctgcgc	cgggcggagc	aatccggagc	gctgtaactc	46680
gtcgtgtacg	ggcgctcgcg	gacggaccgt	gaggacctgt	ggtccgagga	ggaatggacc	46740
cggcacgccca	gcggtgtcgt	cgtcgacaga	gcgcctcggc	cccccagacc	cgtccaaactg	46800
accgtatggc	cccgggaagg	cgccaccgag	ctcatcgtga	aggacctcta	cgaacggatc	46860
gcccggcacca	gcttoggcta	cgggtccggc	ttccaagggc	tgccgcgcgc	ctggcggctg	46920
gacgacggcg	tggttcggga	ggtcgtgtgt	ccacaggatc	agtacgcgtg	cgcgagccgg	46980
ttcggactcc	accggcgctt	gctcgacgcg	gcctccacg	gggtcgcgct	ggggcagccg	47040
gcggctgaca	ccgccagacc	gcacaccgac	cggatgocct	tctcctggag	cggcgttacc	47100
ctctacgccg	ccggtgcccac	cgcactgcgg	gtcgggttgg	acatcgtctc	gcccaggagc	47160
gtgtcgtcgc	tcgtcgccga	tggtcggggg	gctccgggtg	cccggtgaa	ctcgtctgaag	47220
ctgcgcccgg	tcgcccggga	cctggccagt	gccggtgtcg	ccgactcgtc	gttccggctg	47280
gagtggctga	aggcggtcga	cgacgagccc	ggccggggcg	aaccggggca	atgggcccctg	47340
atcggaacgc	cgcccggctg	cgacttcacg	ccgggcgagg	acggcgtcat	catcggaagc	47400
taccgggaca	tgcccgcggt	gaccgacgcg	ctcgacaagg	gagtcgcgtg	cccgcagccg	47460
gtgtttgtgt	ccgcccgcgt	ggaggaggag	caggaccagg	cgcacgatct	cgcgagcgcg	47520
gtggacaagg	ccacgaacgc	gctgctcgca	gtgctccagc	agtggtgtgc	cgacgaccgg	47580
ttcgactcct	ccaggctggc	tgtgtgacc	cgtcacgcgg	tgtccacggc	tgggcaggag	47640
gacgtgacgg	accttgccca	cgccctcgtg	tggggactcg	ttcgtcggcg	cgagtccgaa	47700
catcccgacc	ggttcgtgct	ggccgacacc	gacggcacc	agatcagtca	cgttgccttt	47760
ctgcccgttt	tgctgtccgg	tgagcccgag	gtcgcgtcgc	tgacgggaac	ccglatgtgt	47820
ccgcggctgg	ccaggcccg	tgcttcgggg	gacgggccc	tgccgcgggt	ggaccggcg	47880
gggacgggtg	tggtgaccgg	tggtacgggg	actctggggt	cttcgctggc	caggcatttg	47940
gtggttgagc	acgggggtgc	gcggttgttg	ctggtgagcc	gtcggggcgg	ggagtcggag	48000
ggcgccggcg	agttggtggc	tgagctgacc	gggctcgggg	cggatgtcac	ggtggcgcg	48060
tgtgatgtgg	gggaccgggg	ggccgtggcg	gagttgctgg	cggggattcc	ggccggtcat	48120
ccgttgacgg	cgggtggtga	tgccctgggg	gtcactgatg	acgcggtgat	cgaggcgttg	48180
actgcggagc	aggtcggccg	ggtgctggcg	tcgaaggctg	atggggcggt	caatctgcat	48240
gagttgacgc	gggggctgga	tcgtcggcg	tttgtgtgt	tctcttcggc	ggccggtgtg	48300
ttcgggaatc	cggggcaggg	caactacgcg	gcggccaatg	cctttctgga	tgcggtggcg	48360
gtcggcgccc	gggcggaggg	cctggctgcg	cgtcgcgtgg	cctggggtct	gtgggaggag	48420
gccagcgcca	tgacgagccg	gctggccggg	gocgatctgg	tcgggatggg	cgtgtcggcg	48480
ctgcttcccc	tcaccaccgg	gcaagggtcg	gcctcttcg	acgcgcgcca	cggacagac	48540
gagccctctg	tactgcccgt	gaggctggac	accacggccc	tgccgtccac	caccggacag	48600

ccgccggcgc	tgctgcgcaa	cctggtccgg	gtccaggctc	gcggagcggc	ggcgccggcc	48660
ccgggaccgg	acggcgccgc	caccttccag	cagcagctca	tcagccgtgc	cgtcgccggag	48720
cgcgccgggg	tgctgctgga	gaccgtacgc	ggccacgcgg	ccgccgtgct	cgggcactcc	48780
ggcccggaag	ccgtcagatg	cgacaagggc	ttcatggaag	cgggcttcga	ctccttgagc	48840
gcgggtggagt	tcgggaaccg	gctgacgtcc	accacggggc	tgccgatgcc	ggccaccgtc	48900
acgttcgact	acccgagccc	ggccgcgctg	ggcgagcacc	tgctgacgcg	gttggttccc	48960
gaggtcgcca	tgcccgcgga	ggagcagcac	ccgcacaccc	ggcccggaag	cgggccgggtg	49020
gacaggcccg	gagacgaaca	ggggggcgcg	atcgacgaca	tggacgtcga	cagcctcgta	49080
gaactcgcgc	tcggcgaatg	attcctgatg	ccgcacatgc	ccgggaggac	agcatgagca	49140
agcccccata	aaaagtagtc	ggcgcgctcc	ggcgctgcgt	gaaggccaac	gaacgcctgc	49200
gggagctcaa	cgacgagctc	gcctcggcgt	ccgcggaacc	ggtcgccatc	gtcgccatgg	49260
cgctccggta	tcccgccggg	gtgacgtccc	ccgaggaact	gtgggacctg	gtcgccggcg	49320
gcaccagcgc	ggtgtcggag	ttccccgcgc	accgtggctg	gaacgtcgag	gagctctacc	49380
accgggaccc	ggaccactcg	ggcacctcct	acgtgaggga	gggagggttc	ctgcatgagg	49440
cgccggagtt	cgatccgggtg	ttcttcggca	tgtccccacg	ggaggcgctg	ggccaggatc	49500
cgcaacagcg	gctgctcctg	gaaacggcat	gggaggcctt	cgagcggggc	ggtatcgacc	49560
cactccggct	ccggggcagc	cgggaccggc	tattcgtcgg	cgctatgtac	aacgactacc	49620
tcaccgcgct	ccagccggcc	cccgccgact	tcgaggggca	gctcggcaac	ggcagcgccg	49680
cgagcgtcgc	caccggccgg	ctggccctaca	cgctcgggct	ggaggggccg	gcggtcacgg	49740
tggacacggc	gtgttcgtcc	tcactggtcg	ctctgcacct	cgccgccag	gcgctgcgca	49800
acggcggaatg	caccatggcg	ctggcagggc	gggtcgccgt	gatggccacc	ccggggccct	49860
tcaccggagt	cagccggcag	cgcggtctcg	cggtggacgg	ccggtgcaag	ccgttcggcg	49920
cgccggcgga	cgccaccggc	tggcgagagg	gcgtcgccct	gctgctggtc	gagcggtcti	49980
cgagcgcgcc	ggcgaacgga	caccgggtgc	tggtctcat	acggggaacg	gcggtgaacc	50040
aggacggacc	cagcagcgcc	ctgaccgtgc	ccaatggccc	ctcgacgag	cgctcatcc	50100
ggcaggcact	ggcgaaacgc	ggcctgtcgg	ccgcggacgt	cgacggcggtg	gagcacacg	50160
gcacggggac	cccgctgggg	gacccgatcg	aggcccgagg	cctgatcgcc	acctacgggc	50220
aggaccggcc	ggccggccgg	ccgttgtggc	ttggttcgct	gaagtccaac	atcgccca	50280
cccgagccgc	cgccggcgcc	gccggagtca	tgaagatggt	ccaggccatg	cgccacggga	50340
ccctcccgaa	gagcctgcac	atcgacggcc	ccacggccca	gttcgactgg	gaggccgggg	50400
cggtggaact	gctcaccgag	gcgtgcccgt	ggcacgagac	cgacggccgc	cgagggcggg	50460
gcgtgtcttc	cttcggggtc	agtggcacc	acggccacgt	gatcatcgag	gaggtctccc	50520
cgaccggaag	tcgggagggc	gtgacggcgc	ggcgcccgct	caacggcgag	accttgccgt	50580
gggtgtcttc	ggcccgctgc	gtagaggccg	tcggggcgca	ggccggcgag	ctgcgtctct	50640
atctgtcgga	gcgtcaggac	tcgtcactgg	aggccatcgg	acitctctgt	ggcaccacgc	50700
ggtcggcggt	ccagcaccgg	gcccgtgtac	tggcgccgca	ccacgatggc	ttcatggcgg	50760
ggctggacgc	gctggccacc	ggggaaacgg	cgaagggtct	ggtcgtatgg	gaggccgtat	50820
cgccggcggg	agtcgcccgt	gtcttccccg	gccaggcgctc	ccaatggggc	ggaatggcgc	50880
tcgaactgct	ggactcctca	tcctgtttca	gagacggat	ggaagcctgc	gcgcaggcgc	50940
tgagccctta	catcgactgg	tcactgaccg	aggctcctgc	ctcctgcgaa	ggcgagctgg	51000
aacgggttga	cggtgtccag	cccgccgtgt	ggcgctgat	ggtctcgtg	ggcgaactat	51060
ggcgttccct	cgaggtccgg	cccgcccgcg	tcctggggca	ctcgcaggcg	gagatagcgg	51120
cgccctgtgt	ggccggcgcg	ctcagcttgg	aggacggcgc	gctgggtgct	gcgctgcgca	51180
ggcaggccat	cgcgaccgag	ctggccggcc	ggggcgcaat	gctgtccgtc	ggcctgcgca	51240
aggcaccggc	ccaggactgg	atgacggggc	ggcgggaaag	gctgtcggtc	ggcgcggtca	51300
acggggcccg	atcagttgtg	gtctcggggg	acgtggacgc	gggtgaggag	ctgcggcgcg	51360
agctggccgc	cgagggggtg	cggttcggca	ggcttcgggt	cgactacgcc	tcgcacagct	51420
cgcatgttga	cgggatccgc	acacgtctgc	tggcgccgct	cgcccccgtc	tcgccgcgcc	51480
cttcggagat	caccctgtac	tcgtccgtga	ccgggtggtc	catcgacacc	acgacctatg	51540
acgccgagta	ctggtaccgg	aacctgcggc	agaccgtgga	gttcgagcgg	cggttcggca	51600
cctcgatgtc	cgacggctac	cggtttctca	tcgagtcag	ccgcaccccg	gtgctgacga	51660
cgggcatcga	ggagaccgcg	gaggacgctg	accggttcgc	ggcgcggtgc	ggttcgctgc	51720
gcccgttcga	cggtggcccc	gacaggttcc	tgactgcgct	cgcgaggctc	caogtgcgcg	51780
gcgtgcgggt	ggagtggggc	gtgatgttcg	ccggccggcc	cgtagtcag	cccgatctcc	51840
cgactgtact	cttcagcggg	cagcggtatt	ggctggcccc	cgacacgtcc	cccgcgagcg	51900
acggcgccgg	cgacgaacgc	tcggagacgc	ggttctggga	ggctgtcgag	cgccaggacc	51960
tcggcgaaat	gagcgagacc	ctgcggatcg	gtgacgcgga	ccggcaggcg	tcgttgggtg	52020
agttgttgcc	ggccctgtgg	acgtggcggt	agcagaacgc	gtccgcggcc	gtcctggaca	52080
gclggcgga	ccgggtctca	tggcgccggc	ttccccggc	gtccgatcca	gccttgcggg	52140
gcacctggct	gatcgtggtc	ccggcgggga	cgcgggacca	gcagtggggc	gaagcgctct	52200
cccgagccgc	cgagggcctg	ggagaccagg	ctgtccgggt	cgaactgggc	agggcgggaag	52260
ccggcgggga	ggagtacgcg	gccaggctcg	ccgaggcggc	ggccggcggt	ccggtggcgc	52320
gcgtgcttcc	cctgctcgcc	ctggccgagg	agccggcgga	cgccgacccg	gtgtggcgcc	52380
cgatatgtac	cagcacgctg	gcacttatgc	aggcgctggg	cgacgcgggg	atcgccgcgc	52440

cgtgtgtgct	ggccaccggg	ggcgcggtct	cgatcgggcg	gtccgacaag	ccggtcccgt	52500
cgacacggcg	acaggcccag	ctgtggggcc	tgggcccggg	catgggactc	gaacaccccg	52560
aacgggtggg	tgggctcgtt	gatctgcggg	agacggccga	cgctcgtcgg	acggcgcgcg	52620
tggccggcat	cctggccggc	ggtctcggcc	ccgaggacca	gtcgcgggtg	cggctctccg	52680
gctgttacgt	acggcgtctg	gtccgogcac	cgctcgaccg	gcgagcggcg	aggccgtcct	52740
ggcacacgic	ccgtacggcc	ctggtcaccc	gtggcacccg	cggtctcggg	gcgcacgtcg	52800
cccgatggct	ggcggagcacc	ggcgcgggaa	acctggtgct	caccagcagg	cgcggcccg	52860
acgcccctgg	gacggagcag	ctgtgcggcg	aactgtccgc	cctcgggggtg	cgggtgagcg	52920
tgggtggcctg	cgtgtgttcc	gaccgggacc	aactggccgc	cacattggca	cgattgaccg	52980
ccgacggcca	caccgtccgt	acgggtgtac	atgcggccgg	ggtcagtagc	ccggcgcccg	53040
tggccgacct	cgggcccggc	gagttcggcg	aggccgtcgc	gggcaaggcg	cggggcgccg	53100
cgcacctcga	cgaactgtct	ggcgacggcg	agctggagcg	cttcgtgctc	ttctcatcca	53160
acgcccggct	gtggggcgcc	ggcgcccgag	gcgcctatgc	cgcgcgcaac	gcctacctgg	53220
acgcgtggc	caaacggcgc	cgggtcccgcg	ggcgcgctgc	gaacctccgtc	gcattggggg	53280
cttggccggg	cggcgccatg	ggcgccggagc	gtaccggcca	cgagcagctg	cgcggccggg	53340
gggtgcgggc	gatggaccca	ggcgtggcga	ttctccgact	ccaggaggcg	ctggagcagc	53400
aggagacgtt	ttctcggcgt	ggcgacatgg	actgggaccg	tttctcccg	tccttcacca	53460
tggcccgccg	ccgccccctc	ctcgacgacc	tgccggaggt	ccagcggcag	cggctgagcg	53520
cggccccgtc	atggggccac	ggcgagaccg	acggcccggc	actcgcgcag	cagctcggcg	53580
gggtctttga	acgggagcgc	ggcgggcgcc	tgctcgacct	ggtgcgcaag	cacgcggcgg	53640
cggtgtctcg	ctacggccgg	ccgaacggag	tcgaggcgga	acggcccttc	cgggagctgg	53700
gcttcgactc	cctcaccggc	gtggagatgc	gcacccgact	ccagccggcg	accgggctga	53760
cgtgcggcg	cacctgtgtc	ttcgaccacc	cgacgccccg	cgtctgtggc	gcgcactctc	53820
gggtgagct	gttcgggtg	caggagcaca	cgccgggaac	ggcgggcgcg	tcggcacccg	53880
acgacgacc	gatccgcatc	gtgtcgatgg	gctgcggttt	ccccgggtgt	gtctctcccg	53940
cggaggggct	gtggagctg	ctgctgtccg	ggcgtgacgc	catgtctctg	ttccagtggt	54000
acaggagctg	ggacctggac	agccttgccg	gtgacggccc	cggacagatc	ggcgccggtt	54060
acaccttga	ggcgggcttc	ctcgatgacg	cggccgggtt	cgacgcggcg	ctgttcggga	54120
ttctgcggcg	tgaggcgctg	ggcatggacc	cacagcagcg	gctgctgctg	gaggcttcgt	54180
gggaggcctt	cgaacgagcg	ggcatccctc	cggccgacct	gcggtccagc	cggaccgggg	54240
tggtcatcgg	cgttctctca	caggagatac	cccaggtcgc	cgcggagttc	gcggaaggag	54300
tcaggggaca	tggtgtgacc	ggtgacgcgg	ccagcgctcat	gtccggccgt	ctgtctgtaca	54360
cgttcggctc	ggaggggacc	ggcgtcacgg	tgataaccgc	gtgttcgtcg	tcgtgtgtgg	54420
cgttcgacct	ggctgcgcag	gggttgcgca	acgggtgagtg	cactctggct	ctggccggcg	54480
gggtgcgggt	gatgtgtacc	ccggcgcgct	ttgtcgagtt	cagccggcag	cggggcgctg	54540
cagctgagtg	gcgggtgcaa	gccttcggcg	acgctgcgga	cggcacccgg	tggggcgcaag	54600
gcgtcgggtg	gctgtgtgtg	gagcgtttgt	cggacgcggc	ccgcaacggg	catccggtgc	54660
tggcgggtct	ttcgggcagt	gctgtcaacc	aggacggggc	cagcaatggt	ctgacggcgc	54720
ccaatgggtc	ttcgagcaaa	cgggtgatcc	aacaggcgct	ggccaatgca	gggttggcgg	54780
ggcgcgatgt	cgtatccgtg	gaggcgccag	gcacgggaa	ccggtcgggc	gacccgatcg	54840
aggcgccagg	gttgatcgcc	acctacggac	aggcccggtc	ggcggaccgg	ccgttgtggc	54900
tgggttcgct	gaagtccaac	atcgccacca	cccaggccgc	cgcggggcgtc	gcggggctca	54960
tcaagatgat	ccaggccatg	ggtcacggga	cgtgcggccg	tacgtgcgat	gtcgaccggc	55020
cctcgtccca	ggtggattgg	gaagccggcg	cgggtggagct	gctgaccgaa	gcgatgcctc	55080
ggcccgaggc	cgaccggccc	cgccgggcag	cagttctctc	gttcggtgtc	agtggtacga	55140
acgcgcacgt	catcatcgaa	cacgccccgc	aggtcactcc	cgcctcccag	gccccgggaa	55200
cgggtgaagt	cccggatgct	gtggaggctg	atcgaccggg	cccgtggctg	ctgtcggcgg	55260
gcagtgacgc	ggcgttggcg	gaggtggccg	aacggctggc	cgccctacgc	gaatcgccac	55320
cggaggctcag	tgccggcgag	gtcgcgttct	cgtcgcgac	caccgggtcc	ctgttgccgt	55380
ggcgcggccg	cgtcgttggc	ggcgaccggc	acgagctggg	ccagcgatc	cgggtccgtg	55440
cgggggggcac	caccggcccg	ggcgtcttct	cggggacggc	gagttcgagg	tgacacacgg	55500
cgttctctgt	ctccgggcag	ggcagccagc	gactgggcat	ggggcatgag	ctgtacggcg	55560
cgcacccggga	gttcggcgag	gcgctcgacg	aggtctcgcg	tcacctcgac	gtgttcgggg	55620
accggccggt	gaaggagggt	ctgttcggcg	aggcgagatg	ggcgagatgc	gggtgatcgc	55680
acggggcggg	gttcggcgag	ccggcggtgt	tcgcactgga	ggtcgcgctg	taccggaccc	55740
tgggaagcatg	gggcatcacc	cccgaactatc	tggccgggca	ctcccttggg	gagatcgcg	55800
cggctcatgt	cggcggggtg	ttcagcctgg	aggacggcgc	tcgcctgggt	acggcgccgg	55860
ggcagctcat	gcaggccctg	cccggcggtg	gcgcgatggt	ggccgtccag	gcctccgagg	55920
acgagatcct	ggccatctcg	gcgcggtggc	tggaggggga	cggggctggc	atcgcccgcg	55980
tcaacgggtcc	cgccctcggtc	gtcgtctcgg	gggacgagga	agccgtcctg	gcgatcgccg	56040
ggcaactggcg	ggcaccgggg	cgaagagacc	gtcggctcag	cgtcagccac	gccttcact	56100
caccgccacat	ggatcccatg	ctcgacgggt	tcggccgggt	cgtcagcgcc	atgcaccttg	56160
tcgagccgggt	cattccgggtc	atctccaacc	tcaccggtcg	cctcgccgat	cccgggcagc	56220
tgaccagcgc	cgactactgg	gtccggccag	tcgccaagc	cgtccggttc	cacgacggcc	56280

tacagaccct	gcacgatcag	ggcgtcacca	cctacctgga	aatcgccctt	gacgcccagc	56340
tcacggccat	ggctcaggag	gccctgagcc	cccagtcaca	caccgtctcc	accctgcgca	56400
ggaaccagcc	cgaaccacc	agtcgtctca	ccacgctcgc	gcgactccac	accaccggta	56460
ccacccccga	ctggatcacc	tacctcaacc	accgaccctc	atccccgaca	ccgctgccca	56520
cctaccctct	ccaacaccac	cgtactggc	cgcgcggoga	tgctcaggcc	gccgatgtca	56580
gctccgcggc	cctgtccggt	gcgaaccatc	cactgtctgg	agccgcggtc	ccgtggccg	56640
acggcgacgg	ccatctgttc	accggggcgc	tgctcgccag	gacgcaccgc	tggtggccg	56700
accaccaggt	cggcggcaac	gtcgtactgc	cgggcaccgc	cttcgtggaa	ctggcggtac	56760
gggctggtga	ccaggctggc	tgacgacagg	ttgaagaact	gacgctggaa	gcgcgcgtgg	56820
tgctgcccga	gagcggcgcg	gtccaggtag	agctccggct	gggcccggcg	gacgagtccg	56880
gccgacgtga	cctcaccgtc	tacggggcgac	tgccgggggg	cggcgaggac	ctgtggctcg	56940
aggagggatg	gaccggcgac	gccagcgggg	tcctctccag	cgctcggcc	cccgaaccgg	57000
tcgcactgac	cgtatggccg	ccgtccggcg	ccgagggcgt	gcccgtggag	ggcttctaca	57060
ccgctctggc	cggagagcgg	tacggctacg	gccccgcctt	ccaggggcgt	cgggcccgt	57120
gggctcaggg	cgacacggct	ttcggccagg	tcacactccc	tgagggtgta	cgggagagg	57180
ccgcttctta	caccatccac	ccggctctcc	tgatgcccgc	cctccaagcc	gtcggtttcg	57240
tcacggacgg	gagcgacaac	cccgtggtag	ggatgcggtt	cgctggttcc	ggcgtgtcca	57300
tgatcccgct	cggcgccctc	gagctgcggg	tgccgctcgc	ccggacagga	ccggagacgg	57360
tcaccttccg	cgtcaccgac	cccaccggcc	gccccgtggc	ctcggtcgcc	tcgctgtcca	57420
tgcccccggg	cgccaccgga	gtaccgcgcc	tgacacgcaa	cgggctccac	gagggtgtct	57480
gggagcaact	cctcgatgcg	ccggccaccc	ccgcgaccga	gtgcgcgctc	atcggggacg	57540
ccgctctggc	ggcgtgctgc	ggcgccggag	cgacccggga	cctggcgtcg	ttgggggaag	57600
cggtgcccc	gctgggtggg	gcccgtggcg	cgccgacagg	tacacggcgg	gcactggagc	57660
gcgcccctgg	ctgggtgcag	ggatggatgg	cggaggagcg	gttcgcccgt	tcggcgctcg	57720
ccgtgctcac	ccgtggctgc	gtggcggtcg	gtgcggcgga	gggtgctggc	gacgctcgcg	57780
gtcccgccgt	gaccggccgt	gtgaagtcgg	cggagtcgga	gaaccggggc	cgcttccctg	57840
tggtggatgt	ggacggcacc	accgagtcct	ggcggcgct	gocgactctc	ggcggcgccg	57900
acgagccgca	gatcgcgctc	cgcgacgggc	aggcgtacgt	ccccgcctg	gtgctgccc	57960
gtgaggacgg	cggctcgctg	ctgccccggg	ccggggcgga	cgcttgcgcg	ctggagacag	58020
cggagccggc	cagcctggag	gggctccggc	tcgccccctg	cagggaacgc	caggcgccgc	58080
tgctgcccgg	gacggtgcgg	atcgcggtcc	gtgcgcgagg	cctcaacttc	cgtgacgtcc	58140
tcgggtcgct	cggcatgtac	cccggcgagc	tcgacctcct	cggcagcgag	atcgccggcg	58200
aggtgtctga	gaccggcgat	ggggtgaccg	gcctcgccgt	ggcgacacgg	gtcatggggc	58260
tggtcccgcg	cggcttcggt	ccgatggccg	tcgcggacag	ctggcggttc	gtacggatcc	58320
cgtccggctg	gaccttcacc	cgcgcggccg	gtgttcgggt	gccttctctc	accgcccgtg	58380
acggactcgc	tgaactgggt	gggctggcgg	cgggccaagc	gggtcttggt	cacgcggccg	58440
ccggtggcgt	gggtacggcg	gcgggtcaac	tcgcccggt	attgggggct	gagggtacag	58500
ccacggccag	cgcgcccaag	caggagtatg	tgccggatct	ggcgctggac	cgcgcocgta	58560
tcgcctctct	ccgacccctg	gacttcgctt	ccagcttccc	tgaggtcgac	gtcgtgctga	58620
actccctggc	cggggagtac	gtggagccct	cgctggggtt	gttgccgag	ggcgcccggt	58680
tcgtggagat	gggcaagacc	gatgttcggg	atgctgcgcg	gtacgacggt	gtgacgtacc	58740
ggacgttcga	cctggggcag	gcccgtccgg	agctgatcgc	cgaatgctg	gatgagtgg	58800
tgagatgggt	cgaggccggg	gaactcactc	cggtcgcgac	agccgcctgg	gatgtccggc	58860
gcgcggctgg	gcggttcggt	tgatgagcc	aggcccgcca	cacaggcaag	atcgtcctga	58920
cggtgccgcg	gcacctggac	gcccagggca	cggctcctgat	caccggcggc	accggcacgc	58980
tgggcgggtc	gctcgcccg	cacctgggtca	ccgaacacgg	cgtacgacac	ctgctgctgg	59040
tcctccgcac	gggagaaagg	gocgctctcc	gtcgtgaact	ggaggagctg	ggcgccgagg	59100
taoggatcgc	ggcctcgac	atggctgacc	gcgcggcggt	ggccgaactc	ctcgacggca	59160
tcctctcgga	gcaccccgct	accgggtgtg	tcacgcggcg	gggtgtcctg	gaacgaggcg	59220
tggtcaccgg	cctcgactcc	gctcggtcgg	cacgggtgct	ggctccgaag	gtggacggcg	59280
ccctccacct	gcacgaactg	acggcgggag	tgacctctc	ggcgttctgc	ctgttctcct	59340
ctatgtcggg	tctcctcgcc	gcctccggcc	aggccgggta	cgcggcgggc	aacatgttcc	59400
tcgacgcgct	cgcgccagag	cggcgtgccc	aggccctgcc	cgcgctgctg	ctggcgtggg	59460
gtttgtggga	gaccgcgagc	gcgatgaccg	cgcacctgag	cgacaacgac	ctgocccgca	59520
tgggcgggat	cggcatgctc	gggctcacc	gcaacgaggg	catggaaact	ctcgacggcg	59580
cctggcagag	cggcgaggcg	ctgctggctc	cggctcgcgt	ggaccaccgg	gtgctcgggg	59640
agcggggcct	ctcggggcgc	cgggtgcctc	ccctgctgcg	gaggctgggt	cggcgccgga	59700
ggcgcggctac	ggtgocggag	agcgccaagg	gcgcggggcg	cgggtcggcg	gagcggctgg	59760
gcagcgtgcc	ggaggccggg	cgcgggggca	tgctcatcga	gctggtggcg	gggcacgtgg	59820
ccgcggtgct	gggccaatgc	ggcaccgatg	cgggtgtcgg	ggaccgcccc	ttcaaggagc	59880
tcggcttcga	ctcgtgaccc	tcgtggaggt	tcgcgaacgg	gctgaacgaa	gcgaccgggc	59940
tgccgctgoc	ttcgacctg	gtgttgaccc	acccacaccc	taccacgctg	gcggcccgcc	60000
tcgacggcct	gctgcccggg	gcagagagcg	cgacaacggt	tgctgcccc	acctgcggcg	60060
acgaggaact	cgaccgactg	gcaacgggtg	tgctgtcacc	cgcgttgaa	atggcggatc	60120

ggagagccct	cgccgcccgg	ctccgagccc	tggcttccca	gcttggcgag	ccgactggtc	60180
cgcccgatgg	cagcaccgtc	gcccagccga	tcagtcggcg	caccgatgac	gagctcttcg	60240
agttgtcga	cgacaggttc	gagaactcat	gagcaaacac	gacgatgttt	ctgacgcgct	60300
gaggacggcg	gatgttcga	tgacacagtt	tcggacgaac	gaggacaagc	tcgcgacta	60360
tcgaagcgg	gcggtcacog	acctgcacca	caccogtgag	cagctggccg	cgcccgaggc	60420
caagaaccgg	gaaccogtgg	cgatcgtgtc	gatgagctgt	cgcttccctg	gcggagtcag	60480
gtcgcccgaa	gccttggggc	agctgggtcg	tgccggtgaa	gacgtgatct	cgctgtttcc	60540
caccgacggt	ggatggggacc	tcgacggcct	ctacaacccg	gacccgggga	acagtgccac	60600
cacctacgtg	cgagagggcg	ggttcctgtc	cgacgcgacg	gagttcgacc	ccgccgtgtt	60660
cgggatctcc	ccgctgtagg	cgctgggaat	ggaccocgag	cagcggttga	tgctggagag	60720
ctcgtgggag	gccttcgagc	gggcccggcat	cggtccggca	tcggcacgog	gcagccggac	60780
cggtgtgttc	atcgccgcct	ccgcccaggg	ctacagcttg	ctgttcaga	actcgcggga	60840
ggaggccgag	ggcctcctgg	ccaccgggtga	ctcgcccgagc	gtgatctccg	gcccgggtctc	60900
ctacaccttc	ggcctcgaag	gacctgcggt	cacactcgac	accgcgtgct	cctcgtccct	60960
ggctcgtctt	cacctggccg	tgccctcgtg	tcggcaggcg	gagtgctcca	tggcgttggg	61020
ggcgccgctc	tcggtgatgt	gcacgcggcg	gatcttcctc	gagttcagcc	gcccagcagg	61080
tcctgcggcg	gacggccggt	gcaagccgtt	cgccgcggcg	cgggacggca	ccagctgggg	61140
ggaagccgcc	ggagtcgtcc	tcctcgagcg	gctggaggac	gcccgcagca	acgggcaacc	61200
ggtcgtggcc	gtcatccgcg	gcagtgcctc	caaccaggac	ggtgccagca	acggcctgac	61260
tgccccgcac	ggcccgctcg	agccgcggct	gatccagcag	cgctggcgcg	acgcccagct	61320
gtcgcccgcc	cagatcgaca	tggtcgaggc	acacggcacc	ggcaccctcg	tggggatccg	61380
gtcgaggcg	caggcactcg	tggaaacgta	cggtgccaac	cgcccgccgg	accgcccgtc	61440
ctggctcggt	tcctgaagt	ccaacatcgg	acacaccagc	cgcccgccgg	gtctcgcgtc	61500
cgctcatcag	accgtacagg	cgctgcgaca	cgcccaccctg	gcccaggacac	tgacgtcgca	61560
ccggccgacc	ccgcgcgtgg	actggtgtcg	gggtgggggtg	gaactgctgg	ccgacgacca	61620
ccgctggccc	gagacggggc	agcccccggc	agccgcgctg	tcctcgttcg	gggtcagcgg	61680
caccacggcg	cagctcgtcc	tcgaacaggc	gcccgcctcg	gagaaccggc	ccctccgccc	61740
tcggggaggg	gaccgcctcg	cgccgcggcg	ggtaactccc	ctggtgatct	ccggcaagac	61800
cgccgaagcc	ctcgccggctc	agcgggggaa	cctggtgtcc	catgtgcgcg	agcaccggga	61860
cctccggctg	gaggacctcg	ggtaactcgt	ggccacacac	aggtcggccc	tcggacaccg	61920
ggccgtcgtc	gtggcggaca	ccccccgacgg	attcctccgt	ggctgcgagg	cggtggagcg	61980
cgccgagacc	ccggcgtcgg	tggaccgggg	cgtggtccgg	ggcgccggca	cgaccgcgtt	62040
cctgttcacg	ggcgaggggc	cccagcgggt	cgccatgggg	cgccagctct	acgcggcgat	62100
ccccgcgttc	gcgcggttcc	tcgacgaggc	ctgctcccat	ctcgaccgct	ttacgaagca	62160
gccccgtagg	gacgtgctgt	tcgctgcgga	ggcgacggcc	gaggcagcgc	tcctggaccg	62220
taccggattc	gcccagccgg	ccctgttcgc	cctggagggtg	gcgctgttcc	gcacccctga	62280
gtcctggggg	gtgaccggcg	actacctcgc	cgccacactcc	atcggtgagc	tcgctccgac	62340
ccatgtggcc	gggtgtcctc	cgctgggaga	cgccaccggg	ctggtgacgg	cgctgtggcaa	62400
cctcatggaa	cagctccccc	cgggggggcg	catgctcggc	ctgcaagctt	ccgaagccgg	62460
ggtgtctccg	ctcctcgacc	gcgcggatgg	cctggtgtcc	gtgcgcggcg	tcaacagccc	62520
ccgctccacc	gtggttgcg	gagacagcga	cgccctcgcc	gcccctcgcc	gcccagcccg	62580
ctctcaggcg	atcaaggccc	gccaacctcac	tgtcagccac	gccttccact	ccccgctgat	62640
ggaccccgtc	ctcgacgcct	accgcgagac	cgccgagcag	ctctcctacc	accgcgcggc	62700
tatcccagtc	atctcgaccg	tcaccggccg	gtccgtcacc	accgagatgt	ccgaacccgg	62760
ctactgggtc	cggcacggcc	gcgaggccgt	ccggttcacc	gatgcgctgg	ccacgctccg	62820
gcagcacggc	accacgcgct	acctgggaact	cgcccccagc	gcccgtccta	ctgccaatgac	62880
ccgcgaacac	ctggcggggc	acggcacctc	gggcaaggag	tcacacttcg	cgccggtgat	62940
gcgcaggaac	cgcccgggag	cgagggtcct	gaccagcgcc	gtgtcccgag	tggtccggcg	63000
gggcacccgc	gtcgactggc	ggcccggtgt	cgcgatgtg	gatgggcagg	tcgtccagct	63060
gcccagctac	gccttcagcg	gcagccggta	ctggcccgag	gcatacactga	cccggccggc	63120
cgggggcgcc	tcgcgcagct	cgctgttcca	cctgcgctgg	gtgcccgtga	cgcccccagg	63180
cacggcgccg	gcggacgact	ggcgcttgct	cgccggggcg	gacgcgctgc	ccggccagggg	63240
cttcgccgac	ctggcgttcc	tgggggagac	gatcgacggc	ggatcgccg	caccccgcac	63300
ggtgtgtgtg	cogtltgtcg	ctccggccga	cgccgcccag	gattccggcg	ccacgcacga	63360
cgccgcccac	cgggcgctgg	cgctggctoa	ggcttggctc	gcccagcgtc	gcttcacctc	63420
ctcccgctg	gtgttcctca	cccggtgtgc	ggtggccgtg	accgacgagg	aataccccga	63480
ggactccgtc	gacgccttcg	catagccctc	cgltgggggt	ctgctgcgtt	cgcccccagac	63540
ggagaaccgg	ggccgggttg	gcctggttga	cctcgacccc	gacggcgacc	cgagcgcggc	63600
cgcccgagcg	tgcccggttc	cgcccgccgc	cctggacggc	gacgaacggc	agctggcgat	63660
gcgcggagcg	gtggtccacg	ctcccggtct	caccggggct	acggccgcgc	ccaaggaccc	63720
ggaccgggca	cccgcgggtt	tcgaccacgg	cggaaccgtg	ctgatcacgg	gcgccacggg	63780
tggactcgga	ccgctgctgg	ccgcacatct	ggtcgtcgag	cacggcgtag	gccacctgct	63840
gctcagagcg	cgctcgggcg	cgccggccag	cgccggccag	gcactgctgg	acgagctcgc	63900
cgacctgggt	gcccaggcca	ccgtggtctc	ctgcgacctg	gctgaccggg	aggcgggtggc	63960

cgccctgctg	gcccagggtg	cgcccgcgcg	tccgctgacc	gcggtgggtg	acgcccgggg	64020
cgctccggac	gacggcggtg	tccgctccct	gagcccgaa	cgctcgacg	gggtactgcg	64080
gcccgaaggcg	gacggggccc	tgcaacctga	tgagctgacc	aaggatctgg	acctggccca	64140
cttcatcctg	ttctcctcga	ccgcccgtgt	cctcgccagt	gcccggccagg	gcaactacgc	64200
ggccgcgaac	acgttccctg	acgcgctggc	ccagcaccgg	cgggcagcag	ggctggccgc	64260
tgctctgctc	gcctggggaa	cgtgggaacc	gagccggcgc	atgaccggcg	ggctgacgcg	64320
cgagacctg	gagcgcatga	cgaaggagg	catgccaccg	ttgtccccc	gggacgggct	64380
ggcgctcttc	gatcccgcca	tcgcttcggg	gcgggcccgt	gtgtgcccgg	ccgtgctcga	64440
ttctgacctg	ctgogttccc	ggatcggggc	gaacgtaccg	gcgtgctgct	gcggccctcat	64500
cgagccccgg	ccgttgaggc	cgtctgcccc	aggggaggca	gcccaggcac	tcgcccctgcg	64560
gatggccctc	tgctccggcg	cggagcgcac	ggcgctactc	ctggacctgg	tcggcgccga	64620
cgccggccacg	gtgctgggac	atgacggtcc	gcacgccatc	gaccgggagc	gtggactgct	64680
cgaagcgggc	ttcgaactcc	tgacgaogct	ggagctgcgc	aaccggctgg	ccgaggccac	64740
cgactggcc	gtcccggccg	gttacctcta	cgagtaccgc	acccgaacc	tgcttgccga	64800
acacctggcg	gcccgcgttg	ccgagtcgcc	gcagtccggc	gcggcgaccg	gagccgacgg	64860
accggccgag	ccgctgagcg	tgctcttcca	gcaggcgat	gaacctggca	aggtcaccca	64920
gggcatgacc	ctgctcaggc	gcgcgtccgc	gctccggccg	acctacgaca	ccccttcgga	64980
cctcagtgaa	ctgcgcgagc	ccactcgcc	ggcccggtgg	ccgaacgtg	ccacgtgctg	65040
gtgcttctcc	gccatcggtg	cactcgggg	ctgcaccag	tactcgogct	tcgctcgtc	65100
cttccgcgag	gaacgggagc	ttctgggtcc	ctacgcgcgc	gggttcttcg	ccggggagct	65160
cctgcggacc	agcctcga	cgttcacga	caccaggtg	gaaccgctgc	ggcagcaggc	65220
cgcgacgggt	ccggtggtgc	tcgtcggcgc	gtcttccggc	ggctggctcg	cccatgccgc	65280
cgccgcggcg	ctggaggcgc	tggaacacc	accggcagcc	gtgctcgtgc	tggaacaccta	65340
cctgcgggac	gaccagttcc	tcgcccgtga	ccaggaccgt	ttcatcgccg	gagtcttcga	65400
ccggcaggac	cggttctcca	tcggggaggc	cgtagccctg	tcggcgatgg	gctggtatct	65460
gcacctgttc	gacggctgga	agcccaccgc	gatctccgtc	ccggaactgc	tggtccgggc	65520
gagtgcggcg	ctgcccagcc	cttccggccg	cccgccagg	gcccggact	ggcggaacct	65580
atggcatgtg	gcacagcaca	gcgtcgaggt	gcccggcgat	cacttcacga	tgctggaggc	65640
attcaacgac	gcccagcccg	acgcccgtcc	acgtgggtt	ctgcacattg	actgaaaggc	65700
ctgtcccttg	atctggaac	ccaaacttct	tcgccggcat	acotacggaa	ccgcaccccg	65760
ctcaacgccg	cattgcgttc	gcgcgacct	gttcaacgtg	ccgtggcttc	ggggggccctg	65820
tcgctctggg	tggtgacccg	ctacgaggac	gtgcgcgcgc	tgctcgcga	ttccaggctg	65880
ggcaaaaggc	tcacgcagct	ccgcgaggcg	gtactgctca	acgggggtga	cgacgagccg	65940
atcagccagt	tcacgcagct	cctcacccag	cacatgtcca	acagcgacc	accgaccac	66000
accggctgcg	gcccgcctgt	cggaaggcg	ttcacccgcg	gcccgataga	acagcttccg	66060
cccaggatca	cggagatcgt	cgacaatcta	ctggaccggc	tgagtcggg	tcaggaggctc	66120
gacctcgtcc	ctgtcttcgc	cctgcccatg	ccgaccactg	tgatctcga	actgctcggc	66180
gtccgctccg	tcgaccggct	gtcgttcagc	cactggtcca	atgtgctgtg	gtcgaccgcg	66240
gaagtccggc	aactggccga	ggccggcgga	gcgatggtcg	cctatctagg	acagctcctc	66300
cgggacaacc	ggcccaacc	ctgtgacgac	ctgctcacc	agctgggtga	agccaccgac	66360
aacggcgacc	agctctccga	gacgggaact	gtggcgacgg	ccttctgct	gctgtccgcc	66420
gggcacgaga	ccacgggtga	cctcattgcc	gcccgtacgc	tcactctgct	ccagaaccgc	66480
gaccagctcg	cccggttgcg	ctccgacctc	acgctgctgc	ccggcgccat	cgaggagctc	66540
atagggtagc	acggggcccg	cgcatggtg	ctccggcaca	ccctggagcc	ggtcgaggctc	66600
gcgggtgtga	ccatcccggc	ccagcaggct	gtcctgctct	cgtgtctcct	ggcgggcccgc	66660
gactccaccc	ggttcagcga	cgccgaccgg	ctgcacatcg	gcccgtccat	cgggggcagc	66720
gtgggggttc	ggcacggtat	ccaccactgc	atcggcgcgc	cgctcgccag	gctggagggc	66780
gagatcgctg	tcggggccct	gctcacccgc	ttcccggacc	tgccgctcgc	ggtcccggcg	66840
gaggagctga	actggcgccga	cagtgtcttc	atccgcggcc	cggaaatcgt	gcccgtggtg	66900
ctgtgacgcg	catggggaga	ggggaccgac	ccgtcagtg	cggtccctc	tcgccatacc	66960
cgccggctac	gcccacatga	ggatccgacg	ggtgctttcg	actacctcgg	cgatctgccc	67020
atgcccgaag	cagctggggc	aatacgggac	ttccaggggc	agccggccct	ccacggtgat	67080
cgccgacacc	accagcgccc	cccgcctcgt	ctccggcgac	cagttctccg	gtaccgggat	67140
ccaccgcaaa	ccgcccagct	cgaagccggg	gggtgccacc	aggtcccgga	cacggcccag	67200
gttggtgagc	acgagactgg	ccgcagcgag	agcgggattc	tcaagaagt	ggcgacccgc	67260
gaggatctcg	cgttcgggat	cgcccgcctc	gacacccgcg	cgaagccggt	cgtagaccag	67320
ccggccgagc	gtgcgcacat	ccgcccggcg	ggagacttcg	acaatgctgt	agaacgagc	67380
ggcgccagc	accagggtct	ccctgcccag	cggtggagtg	acccggtgcc	ggaagtgcag	67440
gggggacgcc	aaggccagcg	aaaggggtgc	gtcggtcgt	tcgagggcgc	ggcgaccgcg	67500
gatgagcagt	ccgcccgcga	ccaggccctg	taccgagatc	ccggcgccgc	ggcgccgatcc	67560
ggcgagccgc	gtcgtctcgt	cggaaagtgc	ccggagggtc	cttactgtga	ttctgcccctg	67620
ctccggcgct	tccacgccc	ggtcccccag	atagggcagc	agcacggggg	gaaggcgctt	67680
cgctgctcgc	gcccggcggg	ccgcgtaggc	gaggacatcc	gctccgggt	gatggccgag	67740
acgggtctcg	atgggcggcg	ggtagctgtc	ggccacgtgt	gocgaggagc	ccatcgccgc	67800

ttcggccagc	gcgccgtacg	tccgccacac	cgccgacagc	aaggcgacca	cgctgcccgc	67860
gtcgcagatc	cggtgatcga	catggagtat	gaagggtgtc	tccgccgcgc	cgccgagcag	67920
gggtggccgg	accagcgggc	cgccagggtc	cagccggctc	cgcatctccc	ggctgaggtc	67980
ccaggagccg	gcccggcgga	cgacgagttc	ggcgccggcc	tcgtccagcg	ggtcgagcac	68040
caactcggta	cggtccggcg	atatccggct	ccgcaaggag	gggtgcgcgg	ccacggacga	68100
ggcgaaggcc	cgttggagaa	gctgttcgtt	ggatatcccc	cgtaccgtgc	acagcgccat	68160
gattctcgtg	agacctggcc	cggagagcat	gagctctccc	gtggacagct	cagctcgggt	68220
attcggttgc	atgtttgctt	gcgcctttcc	gttcccgggg	gcccagctac	tggggcgagc	68280
cgacgatgga	ttcccgttcc	cgggaggtga	gagggggccc	ggtagcggcg	ggccgggtag	68340
ggccggagaag	ggtctcgaag	acgatccggg	scgtatcagc	ccgggtcagc	ggggccgaca	68400
gggtgaacgt	gtccgacacc	gcggcggcca	ccggggggcg	gtcggctcgg	gtcttgggtca	68460
cccgggttgc	atagccgcgc	tgcatccgcg	cgcccgccgc	cgcccgccgg	ccactccat	68520
tccggtagaa	gatgtcctgt	cctgtggcca	tcgcccacgc	attgtttacg	gcgcccgcga	68580
cgccggccctg	tggtggcggg	ctcgtcccg	ccaccaacc	gtcgtccgc	aggacgtcgc	68640
gcagcccgga	agcgtctatg	gcggccaccg	acatgccgtg	cccgtacacc	gggttcaggg	68700
cgcccgccgg	gtcgcggagc	accacgaagc	ccttcggcca	gtcggccagt	tcctcgtagt	68760
agcggcgccg	gttgaccgtg	gtgcgactgc	tgtagatcgg	cccgatccgc	tcggcgttcg	68820
cgatgaggtc	cccgatcacc	gaatgcgcga	gtctccggcg	gaacgccagc	aagccttccg	68880
gatcacgcgg	cggtctgcac	ccacgggtcc	cggtcagcgt	gacgatccac	tgtccgtcct	68940
cgatccggag	cagcaccgcg	ccctggcccg	gctggctcgt	ctccgggtcc	ggcagcacgt	69000
tcacgatggg	aaagccgctc	tcccggcccg	ccggcgccgc	gtaccggcgg	gtggcgtagg	69060
agagcccgat	gtcgatcttc	acctcgcgca	cgccggggcag	accgagccgc	tggagccagg	69120
tggtccggcc	ggatccacgg	ccgggtggct	ccaccacgaa	gtccgcgtcc	agccggagcg	69180
attccccgga	cgcccggctc	tgggcttggg	ccccggctac	ccgggtggca	tcgcccgtga	69240
gcccctggag	gtcaaacacc	ctccgcaggg	tgatgcggtc	gtcttccagg	acgagggccc	69300
gcagcgtcca	gtccagcagc	ggacgaccac	aggtgacct	gaactggcgc	ccgggcatcc	69360
ttcggggcca	ccctgcctgt	gagcaggaga	cgagcccgct	gggtacctcg	gtccgggtcg	69420
caccggcgccg	cagcaggcgg	tggaggctcc	cgggcaccag	cgagtccatg	gtccgtgccc	69480
cgctcgacat	caggatgtgc	gagtgagggg	tcgtgggggt	gccccttctg	acttccgggc	69540
cgctccggtag	ctgggtccgc	tccagcatca	ctacctcgtc	cacgaacttc	gcgagcaccg	69600
atgccgttag	ggcaccagcc	agaccgctgc	cgagaactat	tgccgcgattc	accatatcca	69660
tcgtcctttc	aacgtccgag	aacacgtaat	taataccccc	aattcaagcc	gtgattttct	69720
cgacttagtg	cgacggggat	cgccctccaa	tactcctggg	ccccctcggt	ctcccgtctc	69780
gtcgagagtg	gatgggttta	tatccgttgc	cgttcgatga	ggtcggcaag	gaaacaatg	69840
tggtttgcgc	gacgctaccc	gtgtttccct	ctccgggggtg	atcgcccttcg	tcgtgcggtt	69900
gattggagggt	tccctgaag	attcttttgc	gtaatcgcta	tggcccccgc	caggccctgcg	69960
gcggcgatgg	tcattccgcc	ggggcgccgc	tgcaaccagg	tcctgtccgg	gtctcctcat	70020
gcgctctcgg	gcaggccagc	ctggagctcg	agaggcaact	cctccggcgc	ggtagatgtc	70080
agcttgccgt	acacgcgggt	cagggtgctgc	tccaccgtgc	tcactgtgat	gaacaaactg	70140
gcggagattt	cacggttggg	cagacccttg	gcgcgagtg	cggaactcgc	gcgctcggag	70200
tcgctcagat	tggcgccgac	gtccgctccg	cggaattcga	gcgcgagga	catgccacc	70260
tcgggggagg	gacgatggct	ggagccgatc	ggctcggagc	ggacctccgc	actgcaatcg	70320
cctgcgatct	gcccgcgcat	atggcgggat	gcgtcggcac	gcctgccacc	gcccagttcc	70380
tcgtaggttg	acgccaggtc	agcgaggacc	ttggccagtt	gcagccggtc	cccggctccc	70440
tgtaaccgct	cgccagcctg	gatgagcagt	cgtgtccgtt	gcgccgcctc	ggcgaacatc	70500
gcgcggagcg	gcagcactgc	accgtccgtc	gcgcgcctta	tgcggccggt	cgcgctgtcg	70560
tactcggcca	gcataccgtt	cgctgtctcc	cgaicgttga	gcccggagcca	ggcatggggc	70620
gagtcacact	gcccaggcag	ttctgccgac	ggagccagac	cccagcgctc	cgccagccgg	70680
ccgatgtga	ggaagtcgga	gagcgcgagg	tggggccggg	tgacggccag	ggcgtagtgg	70740
ccgcggggcac	ggagatacgg	gagcccgtag	acactccgga	acagcgctcc	gggaaccggg	70800
cggtccagca	gatgagcgac	gtccttgaat	cgtcccactc	cgtgtagac	agtcacaggg	70860
acggtcagcg	ggccgcccgt	cagccagggt	ctggacgggt	cgccgagggc	gtccaggggc	70920
atccatcgga	acgtctcggc	ctcggtcagc	ttccctgttc	tcagcgcgat	atcggtctgc	70980
acggcgccga	acagccgttg	ccagcccggg	atacccccga	cggtcgcgtt	cttcaggaaa	71040
acgtccgacc	acgtagccgc	cagatccagc	cgcccgacc	gcgtgagcga	gttcagctcg	71100
gtgaggatga	ggctgagcgt	catgtcgagc	agcggcgttg	tccgcagcag	ttctcccgag	71160
tcctggatat	ccccgggctt	gcgtccagc	tccttgatcc	aagtggccag	cgccccagc	71220
gcgtcgtctc	gggagggtgc	atcgccgcag	tcctttccgg	acagaccgtc	cgccagagcg	71280
cggtatccgg	cacaccagcg	cgccggcatg	ccggtcatcg	gggggaagaa	ccacagccac	71340
gtgttgcga	cgaccgtcag	gtccgttga	ctgcgaagac	cgccgagggg	ggccgttacc	71400
tcccgcagca	gttctccgc	ttcctccacg	cgccccaggg	tgaccagcag	ttggatcagc	71460
agaacggcat	cgacggggga	gagctcgggt	cccggtgccc	gctccccgca	gtagccgtcg	71520
agatgggtgc	gttcgacgga	acaggggtcg	acgcgccacc	tgacgagtgc	ccgcttcagg	71580
cggatgtggc	ccgctcccca	ggagtcgggt	gagcgctcat	ggccagtttc	gaggtacgcg	71640

cgggcctgtt	cggcgctgtc	cgagtcacac	gcctctctcg	cggcgctgcg	caggcgctccc	71700
acatgccagg	gttcgctgc	cgagcccgct	tcgagcaggt	ggcgggcgat	cgtacggctg	71760
ccgagccggt	gcccactcag	cagttccgcg	gcccgggtggc	gcagttcggc	gcgctgcttg	71820
gggcccagta	gtttcagggt	cgcgcgctcg	accaaggggt	gctggaaacg	gtagccgctg	71880
accagtcagg	cgagggccaa	cgaaggata	ccgcgcgcga	tctcgccggc	gttcagccc	71940
agcagctcct	ccaacagttc	cggccctggg	tcctcgccca	ggaccgcgat	gcccgtggcg	72000
agggacacca	cgcgggggtc	gttgccttgc	acgcagttga	ccgcggcctg	cgcgaagagg	72060
ccgtcgccgg	caggccacgg	ggcggtttgg	cctcgccggt	tgccgaccgg	gtgttcttcc	72120
aacagcgccc	gcacagcag	gggtttggcg	ccgctcagcc	ggaacacgtc	gtccaggaag	72180
gtgtcctcgg	cggccgggcc	ctccaggggc	ccgaccaggt	cgacgacatg	gtcccgggtc	72240
attggggcga	gcgcgatccg	gtggagattg	ggctgcgcga	ggagctcgca	gtggaaactcc	72300
ggcccagtg	atgtcgggag	tgctgtacg	acgatcagca	tcagcctgct	ggaccggagc	72360
ctggcccggt	tggcctccag	cagccagcgc	cagctcaggg	tgctgagatc	ctgtaggtcg	72420
tcgaggcaga	cgaccacggg	tgaccggctg	gccaggggct	cgagtcggcc	gcagaactcc	72480
acgaactcgg	cgttttgtgc	cgaactcatc	gagctcatcc	tgggaacatt	gtcgaatcca	72540
aggtcggggg	cgttaccac	gaccgctcgc	gaggccttca	catgctcgcc	gaaatttacc	72600
agttaattcg	ttttcccgca	gtaggcaccg	ccctccaata	ccacggtcac	agccttgccg	72660
atctcgcatt	cgacaagcaa	ggatttcagt	agatcaagtt	ccgagtcggc	cccgaagaga	72720
tgcatccgaa	ttgaatcccc	aatctccacc	acgaatgag	tgccacgac	gactccgggt	72780
gcaaatccgg	ccggccggcg	ggttggttgc	catacgggtc	gcctcgctcg	tggccggaatc	72840
taagcgctgt	cacgcggcga	ttggggcact	acacggggca	agtaagggc	aactcaggga	72900
ggtgaactgc	cgcccggtc	gactggacgc	ccgggttgca	gaacacccgc	cgaatgtggc	72960
cggaggactc	gcttttcag	aagacgcct	ggtgtcggcc	gagcccgccg	ttgagcgctt	73020
ccggccaact	ctggccggcg	cccgaatcgt	cacgaccgaa	accggtccgc	accaggatcc	73080
atccgagacc	caggcgcgct	gcacgatcgg	gcggagtga	cgtctccgtc	gactccaacc	73140
gcccccgcga	ccgttcaagg	caccgagtc	gtgggtcgg	tcagccggtt	gatcagcacc	73200
tcocgcacaa	agggggagct	cacgcggacg	gctcggttac	tcctgctcgg	ccgtcaggac	73260
ggacaggcag	tgccggaggt	ccocggcctc	cgacggccgc	catgggattc	gacgacaagg	73320
agcgccatc	ccgcctcgc	ggcgccagcg	tgaccgggga	ccagatctgg	ctggagatca	73380
tcagagcgcg	actcgaccgt	gcaccggctc	ctacggcgcg	gcgctgacac	gctacgtgca	73440
caccgaaggc	gtcacggta	tcagggtcaa	ccagccggac	taggccaccc	gcccggcag	73500
cgccagagcc	gacgctcgac	gcgcgcggcg	cgcccaagcg	gtgtgttccg	gcccggccac	73560
cgccaacgcc	aagaccggcg	gcagagacaa	ccctgaacgg	atggtcagcg	aggcatcctt	73620
cgccgcactc	ggcgcggtca	gcccgttgga	gtatcctca	gccaggaccc	aacgccgag	73680
gttcaaccgc	ggcgcgagcc	gccaagccaa	cgggtcgacg	tcacatgtcg	ggaacacccc	73740
gcgttcagcc	acgtactggc	ggagggtcag	ggccggaagg	cactggttct	ggctcccgga	73800
aacgatgtga	cggtgcccg	ggtcagggtc	tggatccggg	cacctgcgtt	gtcacctcg	73860
ggagtgttgc	ggcactccca	tcgcggcgaa	caacgcgtcc	gccgcaccac	gtacatccga	73920
tgctcgcac	gcgcggagag	cctgagcggc	ctgggcccata	cacgocacgc	actcgactgg	73980
gtgagctgca	aggcgcgctc	cgctcgtcc	gtggcctcgc	tcggagcgcc	gagacggtga	74040
aggacgtgca	ccgaacgtcg	cgagggcag	cgcgacgttg	cccgggatcc	cgcgtctcgt	74100
cgcaacaact	ccgggcccga	accgctctt	ggaccacgc	atggtagcgc	cttgccgggc	74160
tcagggggag	gagtggtcgg	cacgcgctct	tggattgcc	cgcgatgctg	attctggtgc	74220
ctcgacccaa	ccttcgagcg	gatccaggac	gttgaacctc	agctgctcct	tctgtgtgtc	74280
cgctgtcaag	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	acgtgcgggt	cggaacttcac
accgcccagc	gaggcgaccg	ccgcgttaact
		cccctcgtgc
		tcgaccgcga
		tcagcgcgct
		60
		120

cttgcggatc	tttccgtgc	gcgcgcccc	ggtggtgagc	acgatgaccg	gcagcccgg	180
gtccccagc	gtggtgccc	tggtgcccc	ggaactctc	tacagctcga	cctgctcgcg	240
caccacatgc	gtcgggctgg	gctcgtactc	gccctcaagt	ggcaagggat	cgtctctctt	300
cgtcgggtccg	gcggatggtg	ctccggacgg	tcccaactcc	cgcggccgcc	cggatcatcc	360
gtaccgcatg	ccccttcgcc	cgagcgggtg	atcaccggtc	cggccatccg	gtcgtccgca	420
ccgcgagcac	caggatcacg	gcgctggaga	gcagggccgt	gaccagccgc	cccgggtggc	480
ccgtcagggc	gcgacccagc	agcgcgcccc	cgcccgccag	cagtagctgc	cagctcggcg	540
acgcggcgaa	ggccgcccgc	gcgaacaccg	cccgtttcag	cggccgtgcc	gcaccggcgg	600
cgccgctgcc	gagcaccagc	gccacgaagt	agaccaccgt	catgggattg	agcagggtga	660
tcccgagaag	gcgagataaa	gcccctgccg	cgccctggaac	cggccgttcc	ggcgggtgg	720
tgagccgatg	ggcgcggtac	tgccgcaggg	cgagcagcgc	cgcccgacgc	gcgagcaccg	780
cggggaccag	cgccgaggcc	cagcgcagcg	ggtccagcac	cggccgcagc	tgtccgcgca	840
ggcggcgccc	gcccacggtc	gcgagcagcg	cgtacagccc	gtcggccgtg	gcgacgccga	900
gcgccgccga	ggcgcgggtg	cgccgcaggg	tgccggcggt	gaggagagcc	agataggtcc	960
cgaccgcgcc	gacgggcacc	gcgatgccgt	accggcgag	caggcccgcg	agcagccgcg	1020
ccgtcacggg	cgtgcgggac	tggttctctc	ggggacggcg	ggcgtgctgt	cggcccgcca	1080
cgcggggggc	ggtggcagcg	ggcgtcgcca	ggaggagggc	tgtaggaggc	atggccgat	1140
cctggggccg	ccgcgcccg	accggcaaat	gaattacggc	gcgttcacgc	cccggccggc	1200
ctcgtctctc	ggccacttca	ccgcgtacgg	cgatctggcc	gaacttgctg	tcgcccata	1260
ggtgcctcgg	gcactaatg	aagatcgcca	cgacgcacct	cttctctcgc	gaggctcttc	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 ccg gac ccc cgg ctg tcc acc gac cgc acc ctc ccg gcc						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 tgc 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 ctg gac gog 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 tgc 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 ccg acg ctg ctg cgc ggt ccc acg gcc gcc gac tgc						1894
Glu Glu Gln Ser Arg Thr Leu Leu Arg Gly Pro Thr Ala Ala Asp Ser						
180 185 190						
caa ggg cgc cgc gag cgg ctc gag gag tac ctc ggc 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 ccg 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			2086
Ala Leu Ala Val Ile Leu Leu Val Ala Gly His 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			2134
Met Ile Ser Leu Gly Thr Tyr Thr Leu Leu Arg His Pro Gly Arg Leu			
260	265	270	
gcc gag ctg cgc gcc gac cgc ggc ctg ctg ccc gcc gcc gtg gag gag			2182
Ala Glu Leu Arg Ala Asp Pro Ala Leu Leu Pro Ala Ala Val Glu Glu			
275	280	285	
ctg atg cgg atg ctc tgc atc ggc gac ggg ctg ctg cgc ctg gcc ctg			2230
Leu Met Arg Met Leu Ser Ile Ala Asp Gly Leu Leu Arg Leu Ala Leu			
290	295	300	
gag gac atc gag atc gcc ggc gcc acg atc cgg gcc ggc gag ggc gtc			2278
Glu Asp Ile Glu Ile Ala Gly Ala Thr Ile Arg Ala Gly Glu Gly Val			
305	310	315	
ctg ttc tcc acc tgc ctg atc aac cgc gac gag tcc gtg ttc gac gac			2326
Leu Phe Ser Thr Ser Leu Ile Asn Arg Asp Glu Ser Val Phe Asp Asp			
320	325	330	335
ccc gac acc ctg gac ttc cac cgc tcc acc cgc cac cac gtg gcc ttc			2374
Pro Asp Thr Leu Asp Phe His Arg Ser Thr Arg His His Val Ala Phe			
340	345	350	
ggt ttc ggc atc cac cag tgc ctg ggc cag aac cta gcc cgc gcc gag			2422
Gly Phe Gly Ile His Gln Cys Leu Gly Gln Asn Leu Ala Arg Ala Glu			
355	360	365	
ctg gag atc gcc ctg ggc acg ctc ctg gag cgg ctc ccc ggc ctc cgg			2470
Leu Glu Ile Ala Leu Gly Thr Leu Leu Glu Arg Leu Pro Gly Leu Arg			
370	375	380	
ctg gcc ggc ccc gcc gag gag atc ccg ttc aaa ccc ggc gac acg atc			2518
Leu Ala Ala Pro Ala Glu Glu Ile Pro Phe Lys Pro Gly Asp Thr Ile			
385	390	395	
cag ggc atg ctg gaa ctc ccc gtg acc tgg taa gaggtctgtg tc atg cac			2569
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			2617
Ile Asp Ile Asp Lys Asp Arg Cys Ile Gly Ala Gly Gln Cys Ala Leu			
415	420	425	
gcc gcc cgc ggc gtg ttc acc cag gac gac gac ggc tac agc acc ctg			2665
Ala Ala Pro Gly Val Phe Thr Gln Asp Asp Asp Gly Tyr Ser Thr Leu			
430	435	440	
ctc ccc ggc cgc gag gac ggc ggc ggc cgc atg gtc cgg gag ggc			2713
Leu Pro Gly Arg Glu Asp Gly Gly Gly Asp Pro Met Val Arg Glu Ala			
445	450	455	
gcc cgc gcc tgc cgc gtg agc gcc atc cgc gtg acc gaa cgc gcc gcc			2761
Ala Arg Ala Cys Pro Val Ser Ala Ile Arg Val Thr Glu Pro Ala Gly			
460	465	470	475
tga ggcggggccc ggcggccgcg gcccgctgcc gggaccgcg ttcccagttc agtagg			2820
gtcgtcgat gacctcacag gccgggaagc ccttcctcta cgtcgtcgtc tgcgcggccg			2880
ggaccgccgc cggagtcacc acgctgatcg gcgcgcgccca ggcgcggggc tgggaggtgg			2940
gggtcctggc cagcccggtg gcgatggcgc ggttcttcga cacggctcgc gtccaggaga			3000
tgaccggccg gcccatccgc tggcctggc gctcgcggcg cgtccgcgc cgttcccg			3060
cgccggggcg cgtgtgtgtg gcgcgcgccca ccttcaaac cgtcaacaag tggcgcccg			3120
gtctcgcga cagctcgcg gtccggcagc tctgcgaggc ggcgggcctc ggcgtgcga			3180
tcgccgtcct gccctgcgtg gcggacgcgc tggccgccca cccgcgtac cgggaggcc			3240
ttctcggct gcgtgggatg ggcgtccgct tgcgcgagcc gtacgcggc cgcggggg			3300
aggacggcga ggcggacggc gcacggcccg ggttcgcctg ggagaacgcc ctggacctgc			3360
tggagcgggc ctgaaccgc tcccgcacc gtaggcctg tctgacactg tcagacaggc			3420
cctaaccgca ggtcagcgcc ggcgcggcca gcatccgcg ggtgtagagg tcttgcccc			3480
gcggcagcca gtacccagc ctggagacca ccgtggagca gtcaggcccg acggtgacgc			3540
ggaccttcac cgtctcggga cgcggggct gcagcgcggt cagcgcgag tccaggaggt			3600
acgcgagcgc ggtcttcgag gtaccggccg accagcgggt tgacgcagcg ggcgtcgtcc			3660

gtggcgatcc gcaccccggt gcccgcccg ccatgagtc cgagccggc gctgccgtcg 3720
ccctcgtcgc tgcctccgcg gaccgtgtag gtcagcgtgg tggtagcgc 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 cgtccgccc ggtgaccgac cggcgtatcg gctatgtcgc 60
cgcgctcttc gcggcgctgg gcttccccga gggcgaggcg cgggaccgcg gcctgctggc 120
gtacaccgcc tacctcggcc acaccagct cggacatgcc gtccgacaga gcctgccggc 180
cgaggcgcca caccaccgct atctggatgg cgtgatcgac accctcgtac ggccgcggga 240
cggaggcgat gaagccgaac atgtcacaat ctgaacgagg ttggcggaac tgcgcgcaga 300
acatgcccg tatccgccc atgaggtgag atcggcgccc cgaacacagg tgcgccacag 360
cgttgccatc tcacacacga gcaactcgag ccacttgaga ctgtagggg aggaattc 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
ccg 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 gcc 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 cgg 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 ago 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

23/39

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 gcc 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 cgc 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 cgc 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 gcc cgt tgc 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 gcc 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 cgc 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 cgc gcc gcc gag att cgc 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 cgc ctg gcc tgg tag cagccaggac ggcaga 1623
 Leu Gln Gly Met Ile Glu Leu Pro Leu Ala Trp
 385 390 395
 ccaagaaag ggtccgga 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 gcc cgg gag gac gcc gcc gcc 1771
 Asp Asp Gly Phe Ser Ala Leu Leu Pro Gly Arg Glu Asp Gly Ala Gly
 425 430 435
 gac cgc 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
 ggc gtc acc gac gat tag cagcaccgcc ggcgacgacc cggcagacgc gcgcggcc 1875
 Ala Val Thr Asp Asp
 455
 ccggctgaca ccggcgccc gagcgcgccc cggcgcgtcc gccctccac ttgtccctac 1935
 ggcacccacc ccacccgcta ccgcaacacc ccttgggtga cgggcagttt caggaccgcc 1995
 ggtgtgccc gggcgtactg gtgaccgtca ccggcttcac ccgcgattg ccacatagg 2055
 cgtcgtcgtc cgcggcgatc acgaagcgcg gtcggtgcc cggctcgtaa cgggtcacga 2115
 tgcccgccag ttccacgggt aaccgcccgg ccacatcggg caccggggcc ggggccacca 2175
 acaggcgac cagcgtcttc ctgcccgttc gcgcgacatc gtagagcttg gcgaacagca 2235
 ccagcttgtc cgcgcacatc cgggaccgct gcgcccgccc ggcctgcgc gaggcaacct 2295
 tcagcgtcac cctcggcgcc cccaccacgt cgac 2329

<210> 4

24/39

<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 tggatggtc   60
ccagccagat ccgcaggta gccgatctgg ccgaacttga tgcgtgcac tggatgcctc   120
gggcactctaa tgaagatcgg cagcagcat ccttcgtctg 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 Asp Leu Thr Glu Leu Ser Asp Pro Val Ser Phe Pro Gln
                                     5
                                     10
                                     15
gac cgg agc tgc ccc tac cac ccg ccc acc ggg tac gac ccg ctg cgc   273
Asp Arg Ser Cys Pro Tyr His Pro Pro Thr Gly Tyr Asp Pro Leu Arg
                                     20
                                     25
                                     30
acc gaa cgg ccg ccc gcc cgc atc cgg ctc tac gac ggc cgc ccc gcc   321
Thr Glu Arg Pro Pro Ala Arg Ile Arg Leu Tyr Asp Gly Arg Pro Ala
                                     35
                                     40
                                     45
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
                                     60
                                     65
cgc ctg tcc acg gac cgc acc cgc tgc ggc ttc ccg gcc aca act ccc   417
Arg Leu Ser Thr Asp Arg Thr Arg Ser Gly Phe Pro Ala Thr Thr Pro
                                     70
                                     75
                                     80
cgc ttc gcc ggc 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
                                     90
                                     95
gac ccc aag cac cgc acc cag cgg 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
                                     105
                                     110
ctc agg cgc gcc acc gag ctc agg ccg cgc atc cag gag atc gtc gac   561
Leu Arg Arg Ala Thr Glu Leu Arg Pro Arg Ile Gln Glu Ile Val Asp
                                     115
                                     120
                                     125
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
                                     140
                                     145
cgt tcc ttc gcg ctg ccg gtc 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
                                     155
                                     160
ggc gtg ccc tac gcc gac cac gag ttc ttc gag gac cag tcc agg cgg   705
Gly Val Pro Tyr Ala Asp His Glu Phe Phe Glu Asp Gln Ser Arg Arg
                                     165
                                     170
                                     175
ctg ctg cgc gga ccg gcg gcc gag gac acg cag gac gcc ccg gac cgg   753
Leu Leu Arg Gly Pro Ala Ala Glu Asp Thr Gln Asp Ala Arg Asp Arg
                                     180
                                     185
                                     190
ctc gcc gcg tac ctg gag gac ctg atc gac gag aag cgg cgc cgg ccc   801
Leu Ala Ala Tyr Leu Glu Asp Leu Ile Asp Glu Lys Arg Arg Arg Pro
                                     195
                                     200
                                     205
ggt gac ggc ctg ctg gac gaa ctc gtc cag cgt ctg aac gaa ggc   849
Gly Asp Gly Leu Leu Asp Glu Leu Val Gln Gln Arg Leu Asn Glu Gly
                                     215
                                     220
                                     225
gag ctc gac cgg 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
                                     235
                                     240
gcg ggc cac gag acc acc gcc aac atg atc tcc ctg gcc acc tac acg   945

```

25/39

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 cgc	993
Leu Leu Leu His Pro Glu Arg Leu Thr Glu Leu Arg Ala Asp Pro Ala	
260 265 270	
ctg ctg cgg gcc gcc gtc gag gaa ctg atg cgg atg ctg tcc atc cgc	1041
Leu Leu Pro Ala Ala Val Glu Glu Leu Met Arg Met Leu Ser Ile Ala	
275 280 285 290	
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 gcc gtg gtc ttc tcc acc tct gtc atc aac	1137
Thr Ile Arg Ala Gly Asp Gly Val Val Phe Ser Thr Ser Val Ile Asn	
310 315 320	
cgc gac gag gac gtc tac cgg gcc ccc gac acc ctc 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 ctc gcc cgc acc gaa ctg gag atc gcc ctg cgc acc ctc	1281
Gly Gln Asn Leu Ala Arg Thr Glu Leu Glu Ile Ala Leu Arg Thr Leu	
355 360 365 370	
ctc gaa cgg ctg ccc acc ctc cgg ctc gcc gcc cca cgg gag gaa atc	1329
Leu Glu Arg Leu Pro Thr Leu Arg Leu Ala Ala Pro Pro Glu Glu Ile	
375 380 385	
ccc ttc aaa ccc gcc gac acc atc cag ggg atg ctg gaa ctc 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 to 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 gcc gcc gga cag tgc gcc ctg acc gcc cgg 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 gcc ttc agt gac ctg ttg ccc gcc cgg gag gac gcc gcc	1524
Asp Asp Asp Gly Phe Ser Asp Leu Leu Pro Gly Arg Glu Asp Gly Ala	
415 435 440 445	
ggc gac cgg 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 gggccgagc cgcgccgcc gccggtccgc	1626
Ile Thr Leu Ser Glu Asp Gly	
465	
tgcccgccgc ccgtgccgac gccgccgcc gccgcccggt ccggtgccgc tcgctgcgc	1686
ccgtgccccc gccgccggct gattgactag ggttcccggt tgagcgaaca gggccagaag	1746
ccctccgggg ccgcccccgc gaaagacacc gggacggcgc ccgggaaacc ctttctcta	1806
cgctgctgtc tgcgccgcc gcctcgcga 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 ttggacttoa 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 cctgcc 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 aaccocactt 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
 28/39

```

<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 gaggccggtt acctotacga 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 acagtgtott
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 ctt 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
 ttgcscstsc csgtccstc 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
 gctgccctg gccctggagg acatcgagat 30
 <210> 20

30/39

<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
 aggtgcccag 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
 ccatgatcct 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 tcsccsggyt 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 cagtgccctg 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

ccgggcatg cagacgacga cgtaggaggaa

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

33/39

<400> 30
gaggatcctc 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 tccgactag 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 ttcacccgag 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 gcgtcacgt 30

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

 <220>
 <223> R91C-F primer

 <400> 37
 ttcgcggccg tctgcgaccg gcgggtg 27

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

<213> Artificial sequence
 <220>
 <223> R91C-R primer
 <400> 38
 caccgcgcgg 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 ctccga 27

 <210> 41
 <211> 31
 <212> DNA
 <213> Artificial sequence
 <220>
 <223> L403M-F primer
 <400> 41
 acgatccagg ggatgatgga actcccgtg a 31

 <210> 42
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <220>
 <223> L403M-R primer
 <400> 42
 tcacggggag ttccatcacc ccctggatcg 30

 <210> 43
 <211> 30
 <212> DNA
 <213> Artificial sequence
 <220>
 <223> R236L-F primer

36/39

<400> 43 agctggaccg actcgacgtg gtggcgtgg	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 tggcgtggc cgtttcctg ctctgtg	27
<210> 46 <211> 30 <212> DNA <213> Artificial sequence <220> <223> I244F-R primer	
<400> 46 ccacgagcag gaagcggcc agcgccacca	30
<210> 47 <211> 27 <212> DNA <213> Artificial sequence <220> <223> M112T-F primer	
<400> 47 cagcggcgga tgacgatcc gtcgttc	27
<210> 48 <211> 27 <212> DNA <213> Artificial sequence <220> <223> M112T-R primer	
<400> 48 gaacgacggg atcgtcatcc gcgcgtg	27

37/39

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

 <220>
 <223> R195Y-F primer

 <400> 49
 aggggcgcgc gactacctcg aggagtacct 30

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

 <220>
 <223> R195Y-R primer

 <400> 50
 ggtactcctc gaggtactcg cgcgcccctt 30

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

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

 <400> 51
 gggagatctg caggtcatcg gggggaagaa ccaca 35

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

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

 <400> 52
 tctccagccg catgcggtcc cgggtcaccg 31

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

 <220>
 <223> pldout-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	
<400> 54	
gggagatcta gaacatgccg gccagaggo tgac	34
<210> 55	
<211> 33	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> hyg-SpeSphF primer	
<400> 55	
ggggcatgccg gactagtaca cogtcgccto ggt	33
<210> 56	
<211> 30	
<212> DNA	
<213> Artificial sequence	
<220>	
<223> hyg-SphR primer	
<400> 56	
gcgcgatgccg tcaggcgccg gggcggtgt	30