

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4902925号
(P4902925)

(45) 発行日 平成24年3月21日 (2012. 3. 21)

(24) 登録日 平成24年1月13日 (2012. 1. 13)

(51) Int. Cl.	F I
C 1 2 N 15/09 (2006. 01)	C 1 2 N 15/00 Z N A A
C 1 2 N 9/00 (2006. 01)	C 1 2 N 9/00
C 1 2 P 19/62 (2006. 01)	C 1 2 P 19/62

請求項の数 18 (全 212 頁)

(21) 出願番号	特願2001-558462 (P2001-558462)	(73) 特許権者	502288816
(86) (22) 出願日	平成13年2月8日 (2001. 2. 8)		シンヴェント エイエス
(65) 公表番号	特表2003-528592 (P2003-528592A)		ノールウェー エヌー7465 トロンハ
(43) 公表日	平成15年9月30日 (2003. 9. 30)		イム ストリンヴァイエン 4
(86) 国際出願番号	PCT/GB2001/000509	(74) 代理人	100103218
(87) 国際公開番号	W02001/059126		弁理士 牧村 浩次
(87) 国際公開日	平成13年8月16日 (2001. 8. 16)	(74) 代理人	100110917
審査請求日	平成20年2月6日 (2008. 2. 6)		弁理士 鈴木 亨
(31) 優先権主張番号	0002840. 7	(74) 代理人	100115392
(32) 優先日	平成12年2月8日 (2000. 2. 8)		弁理士 八本 佳子
(33) 優先権主張国	英国 (GB)	(72) 発明者	ゾチェフ, セルゲイ ボリソビッチ
(31) 優先権主張番号	0008786. 6		ノールウェー エヌー7030 トロンハ
(32) 優先日	平成12年4月10日 (2000. 4. 10)		イム エルゲスター ゲート 51
(33) 優先権主張国	英国 (GB)		

最終頁に続く

(54) 【発明の名称】 ナイスタチンポリケチドシンターゼをコードする新規の遺伝子、およびそれらの取扱ならびに用途

(57) 【特許請求の範囲】

【請求項 1】

(a) SEQ ID No. 35に示されるヌクレオチド配列；または
 (b) SEQ ID No. 35の相補物であるヌクレオチド配列；または
 (c) SEQ ID No. 35の縮重物であるヌクレオチド配列；または
 (d) SEQ ID No. 35と少なくとも80%の配列同一性を有するヌクレオチド配列
 を有し、当該配列が、ナイスタチンPKS系をコードするヌクレオチド配列またはそれらをコードする配列と相補的であるヌクレオチド配列を有する核酸分子。

【請求項 2】

SEQ ID No. 35と少なくとも85%の配列同一性を有するヌクレオチド配列を有し、当該配列が、ナイスタチンPKS系をコードする配列またはそれらをコードする配列と相補的であることを特徴とする請求項1に記載の核酸分子。

【請求項 3】

SEQ ID No. 35と少なくとも95%の配列同一性を有するヌクレオチド配列を有し、当該配列が、ナイスタチンPKS系をコードする配列またはそれらをコードする配列と相補的であることを特徴とする請求項1または2に記載の核酸分子。

【請求項 4】

請求項1～3のいずれかで定義された核酸分子を有する発現ベクター。

【請求項 5】

請求項1～3のいずれかで定義された核酸分子を宿主細胞内に導入し発現させることを

10

20

含むナイスタチンまたはナイスタチン誘導体を産生する方法。

【請求項 6】

修飾ナイスタチン P K S 系の調製、またはナイスタチン誘導体の調製における請求項 1～3 のいずれかで定義された核酸分子の使用。

【請求項 7】

請求項 1～3 のいずれかで定義される核酸分子を修飾することを含み；当該修飾核酸分子が、ナイスタチン P K S の自然産物以外のポリケチドを製造する修飾 P K S 系をコードする、修飾 P K S 系または修飾ポリケチド分子の調製方法。

【請求項 8】

前記核酸分子が、該核酸分子にコードされる活性をコードするヌクレオチド配列を導入、突然変異、欠失、置換または不活性化することによって修飾される、請求項 7 に記載の方法。

【請求項 9】

請求項 1～3 のいずれかで定義される核酸分子が、

(i) Nys C のモジュール 5 内の E R ドメイン (SEQ ID No. 7 のアミノ酸 4953-5239) が不活性化される；または

(i i) P 4 5 0 モノオキシゲナーゼをコードする遺伝子である nysL (SEQ ID No. 40) もしくは nysN (SEQ ID No: 42) が不活性化されるように修飾される、請求項 7 または 8 に記載の方法。

【請求項 10】

Nys C のモジュール 5 内の前記 E R ドメイン (SEQ ID No. 7 のアミノ酸 4953-5239) が不活性化される、請求項 9 に記載の方法。

【請求項 11】

ナイスタチン誘導体を産生するために使用され、宿主細胞内で前記修飾核酸分子を発現することを含む請求項 7～10 のいずれかに記載の方法であって、

該修飾が、請求項 7～10 のいずれかに記載されている方法。

【請求項 12】

前記核酸分子が発現されるような条件のもとで前記宿主細胞を成長させること、および、該核酸分子の発現産物 (類) がナイスタチンまたはナイスタチン誘導体を合成するようにさせることを含む請求項 5 または 11 に記載の方法。

【請求項 13】

修飾ポリケチドシンターゼ酵素または酵素系をコードするヌクレオチド配列を有する核酸分子を調製する方法であって、

(a) 請求項 1～3 のいずれかで定義される核酸分子を足場として使用し、酵素活性または他の機能性活性をコードする核酸分子の一以上の部分を修飾することを含むことを特徴とする請求項 7 に記載の方法。

【請求項 14】

前記部分が、該部分を導入、突然変異、欠失、置換または不活性化することによって修飾される請求項 13 に記載の方法。

【請求項 15】

修飾ナイスタチン P K S をコードするヌクレオチド配列を有する核酸分子の調製方法であって、前記修飾ナイスタチン P K S が請求項 1～3 で定義される核酸分子によりコードされたナイスタチン P K S から誘導されるものであり、当該核酸分子が酵素活性または他の機能性活性をコードする第 1 領域と、足場となるアミノ酸配列をコードする第 2 領域とを有し、該方法は、

(a) 前記少なくとも一つの第 1 領域を修飾すること；あるいは

(b) 該ナイスタチン P K S から一以上のモジュールを欠失または不活性化すること；あるいは

(c) 該ナイスタチン P K S に一以上のモジュールを挿入することを含む、請求項 7 に記載の方法。

【請求項 16】

請求項 1～3 のいずれかで定義される核酸分子が、

(i) Nys C のモジュール 5 内の ER ドメイン (SEQ ID No. 7 のアミノ酸 4953-5239) が不活性化される；または

(ii) P450 モノオキシゲナーゼをコードする遺伝子である nysL (SEQ ID No. 40) もしくは nysN (SEQ ID No. 42) が不活性化されるように修飾される、請求項 13 または 15 に記載の方法。

【請求項 17】

Nys C のモジュール 5 内の前記 ER ドメイン (SEQ ID No. 7 のアミノ酸 4953-5239) が不活性化される、請求項 16 に記載の方法。

【請求項 18】

請求項 15～17 のいずれかで定義されたように調製された核酸分子を、修飾ナイスタチン PKS が発現するような条件下で宿主細胞内に発現させることを含む、請求項 15～17 のいずれかで定義された修飾ナイスタチン PKS の調製方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、マクロライド抗生物質ナイスタチン (nystatin) の生合成に関与する、モジュール型ポリケチドシンターゼ酵素をコードする遺伝子クラスターのクローニングおよびシーケンシングに関する。従って、本発明は、ナイスタチン生合成に関与する機能的活性を示すタンパク質／ポリペプチドをコードする新規な遺伝子および核酸分子、このような機能的タンパク質／ポリペプチドそのもの、さらにナイスタチン生合成の促進およびナイスタチン誘導体の合成促進の双方におけるそれらの使用、および新規なポリケチドまたはマクロライド構造物に関する。

【0002】

【発明の技術的背景】

ポリケチド (polyketide) は微生物により合成される天然産物であり、その多くは医薬として、または農業用もしくは獣医学用の製剤として利用される可能性がある。医学的治療に使用されるポリケチドの例としては、抗生物質エリスロマイシン (抗細菌剤)、ナイスタチン (抗真菌剤)、アベルメクチン (ivermectin) (抗寄生虫剤) ラパマイシン (rapamycin) (免疫抑制剤) およびダウノルビシン (抗腫瘍剤) が挙げられる。グラム陽性菌 *Streptomyces* 属はポリケチドの主要な生産菌であり、これらの微生物におけるポリケチド生合成の遺伝学および生化学は、かなりよく特性決定されている (Hopwood ら, Chem. Rev. v. 97: 2465-2497 (1997))。マクロライドポリケチド化合物は、モジュール型 (I 型) のポリケチドシンターゼ (PKS) によって単純カルボン酸を繰り返し縮合させることにより、脂肪酸の生合成と同様の様式で形成される。Donadio ら, Science, v. 252: 675-679 (1991) により提案されたモジュール仮説は、I 型 PKS が反復単位 (モジュール) で組織され、その各々はポリケチド鎖合成における 1 つの縮合サイクルを担当していることを示唆した。このことは、マクロライドの化学構造において予測可能な変更をもたらす I 型 PKS 遺伝子操作によって正しいことが証明された。伸長するポリケチド鎖上への次のカルボン酸の縮合 (β -ケトアシルシンターゼ (KS) ドメインの触媒活性が請合う) に加えて、I 型 PKS のモジュールは、 β -ケトレダクターゼ (KR) 活性、デヒドラターゼ (DH) 活性およびエノイルレダクターゼ (ER) 活性を有するドメインを含有することができ、これらが挿入された伸長単位の還元状態を規定する。各モジュールに存在するアシルトランスフェラーゼ (AT) およびアシルキャリアタンパク質 (ACP) ドメインは、それぞれ、伸長単位の選択および PKS 上での伸長ポリケチド鎖の保持のために関係する。合成が完了すると、ポリケチド鎖はチオエステラーゼ (TE) の作用を通じて PKS から放出され、これは最終生成物の環化にも関係している。従って I 型 PKS は、ポリケチド生合成の組立ラインであり、これはモジュールの数、カルボン酸に対するその特異性を変更することにより、あるいは還元的活性でドメインを不活性化するかまたは挿入することにより、操作

10

20

30

40

50

することができる (Katz, Chem. Rev., v. 97, 2557-2575, 1997)。マクロライドのポリケチド部分が合成され、環化されてマクロラクトン環が形成された後に、このものは、通常ヒドロキシル化、グリコシル化、メチル化および／またはアシル化によって修飾される。これらの修飾は、マクロライドの生物活性にとって決定的に重要であると信じられている。

【0003】

ストレプトマイセス属において、マクロライド系抗生物質を生合成するための遺伝子は1以上のクラスターに組織化され、そのような多数のクラスターが既に同定されている。組換え DNA 技術の開発のおかげで、PKSをコードする DNA プローブを用いて遺伝子ライブラリーをスクリーニングすることにより、抗生物質生合成用の完全な遺伝子クラスターを単離することが可能となった (Schwecke ら, Proc. Natl. Acad. Sci. USA, v. 92: 7839-7843, 1995)。分子クローニングおよび完全 DNA シーケンシングが、ストレプトマイセス属のマクロライド系抗生物質の遺伝子クラスター (アベルメクチン、ピクロマイシンおよびラパマイシンのためのものを含む) について、幾つか記載されている (Ikeda ら, Proc. Natl. Acad. Sci. USA, v. 96: 9509-9514 1999; Xue ら, Proc. Natl. Acad. Sci. USA, v. 95: 1211-12116, 1998; Schwecke ら, Proc. Natl. Acad. Sci. USA, v. 92: 7839-7843, 1995)。ポリエンマクロライド系抗生物質、ピマリシン (pimaricin) のための遺伝子クラスターの部分的なクローニングおよび DNA シーケンシングが、最近報告された (Aparicio ら, J. Biol. Chem., v. 274: 10133-10139, 1999)。しかしながら、抗真菌剤活性を有するポリエンマクロライド系抗生物質の生合成のための遺伝子についての完全な DNA 配列は、まだ開示されていない。利用可能な抗真菌剤抗生物質のレパートリーを増大させること、および／または現存する薬剤の諸特性 (例えば効力、毒性など) を改善することの必要性および要求が存在する。それ故に、新規な抗真菌剤治療法、特に新規な特性または改善された特性を示す処置法を提供することは、当技術分野における著しい進歩を表すであろう。

【0004】

本発明はこの目的に志向しており、ナイスタチン生合成遺伝子クラスターのクローニングおよび DNA シーケンシングに基づいている。これは、そのような抗真菌性抗生物質生合成の遺伝子を同定する最初の例を提供し、ならびにナイスタチンの諸特性および／または産生微生物を変更するための、または有用である可能性を有する新規化合物を獲得するための遺伝子操作のツールを提供する。

【0005】

ポリエン抗真菌性抗生物質、ナイスタチン A1 (その完全な立体構造 (図1参照) は Lancelin & Beau, Tetrahedron Lett, v. 30: 4521-4524, (1989) により決定されている) は、ストレプトマイセス・ヌールセイ (Streptomyces noursei) ATCC11455 によって産生される。マクロライド化合物のクラスに属するナイスタチン分子の構造から、本発明者らは、そのポリケチド主鎖が I 型 PKS 酵素により合成されると予想した。この仮定に基づいて、また下記の実施例に記載するように、特別に設計されたプローブ (これは、既知モジュール型 PKS 酵素の既知 β -ケトアシルシンターゼ (KS) およびアシルキャリアタンパク質 (ACP) ドメイン内における保存的なアミノ酸配列に基づくプライマーを用いた PCR で得られる) を使用して、ストレプトマイセス・ヌールセイ ATCC11455 のゲノムライブラリーをスクリーニングした。このことは多数のクローンおよび断片の同定に導いた。これらを本発明者らは配列決定し、これらがナイスタチン PKS 遺伝子クラスターの部分または一部を含有することを示した。本発明者らはさらに、コードされた産物を不活性化する断片配列の変更がナイスタチン生合成の廃止に導くことを示し (下記の実施例1参照)、これによってナイスタチン生合成のための同定された PKS の要件を確認した。クローン／フラグメントに関する後続の研究は、PKS 遺伝子クラスターのシーケンシング、およびその中の異なるモジュールおよび酵素ドメイン、調節領域などの同定に導いた。

【0006】

さらに詳細は以下に記載するとして、本発明者らが同定した新規ナイスタチン PKS

遺伝子クラスター内の機能的 DNA 配列を操作することが、新規な分子構造物、例えば改善された機能を有するナイスタチン誘導体の合成を導いたことを本発明者らは示した。このことにより有益な新規ナイスタチン誘導体を得るためだけでなく、生合成産生過程を改善して促進するため（例えば収率または産生条件を改善するか、または利用可能な宿主細胞の範囲を拡大するため）、あるいは新規な活性および／または諸特性を有する新規化合物を提供するためにも、ナイスタチン A1 PKS 遺伝子クラスターを操作する刺激的な可能性が開かれる。

【0007】

より具体的には、ナイスタチン PKS 遺伝子クラスターの2つの主要領域（または"部分"または"一部"）を最初に同定して配列決定した。これらは併せて、ナイスタチン PKS 遺伝子クラスターの約80%を表していた。また、これらの領域内の機能的配列（例えば PKS 遺伝子、調節領域など、ならびに機能する遺伝子産物、酵素ドメインなども）を同定して特性決定した。

【0008】

第一領域（"領域1"）（これを"Nys 1"と名付け、その完全 DNA 配列を配列番号1の配列に示す）は、多数の PKS または関連遺伝子または調節領域を含むことが示され、13の別個の"特徴"またはオープンリーディングフレーム（ORF）が同定された。その翻訳産物のアミノ酸配列をそれぞれ配列番号（SEQ ID No.）3～15に示す。

第二領域（"領域2"）（これを"Nys 2"と名付け、その完全 DNA 配列を配列番号2に示す）もまた、多数の"機能的"領域を含んでおり、5つの別個の"特徴"または ORF が同定された（その翻訳産物のアミノ酸配列をそれぞれ配列番号16～20に示す）。

【0009】

Nys 1 および Nys 2（すなわち配列番号1および2、およびこれらがコードする配列）は、2000年2月8日出願の英国特許出願第0002840.7号の対象である。

引続き配列決定に努めた結果、配列番号1および2の間隙（gap）にまたがる DNA 配列の決定、およびこの領域中にある新規な遺伝子の同定に至った。加えて、配列番号1および2に含まれており、遺伝子産物 NysI（配列番号20）および NysDII（配列番号3）をコードする部分的遺伝子配列（さらには下記参照）を完成した（それぞれ新しい配列番号36および37を参照 —さらには下記参照）。従ってこれらを配列決定する努力をした結果、ナイスタチン PKS 遺伝子クラスター全体を含む DNA 領域の同定および配列決定、およびこの領域内にある機能的配列の同定および特性決定に至った。

【0010】

ナイスタチン生合成遺伝子クラスターについての完全なコーディング配列（すなわち、このクラスターをコードする完全ヌクレオチド配列）を、配列番号35に示す。この配列は、多数の PKS または関連遺伝子または調節領域を含むことが示され、23の別個の"特徴"または ORF が同定された（その翻訳産物のアミノ酸配列をそれぞれ配列番号2～19および36～42に示す）。

【0011】

配列番号35に示した、ナイスタチン生合成遺伝子クラスターについての完全なコーディング配列（すなわち、このクラスターをコードする完全ヌクレオチド配列）は、2000年4月10日出願の英国特許出願第0008786.6号および2000年4月14日出願の英国特許出願第0009387.2号の対象である。

従って、一つの態様において、本発明は、(a) 配列番号35に示すヌクレオチド配列；または (b) 配列番号35の相補物であるヌクレオチド配列；または (c) 配列番号35による縮退物であるヌクレオチド配列；または (d) 配列番号35に、配列番号35の相補物に、または配列番号35から誘導したハイブリダイゼーションプローブまたはその相補物に、高ストリンジェンシー条件下でハイブリダイズするヌクレオチド配列；または (e) 配列番号35と少なくとも80%の配列同一性を有するヌクレオチド配列；または (f) 配列番号35と少なくとも65%の配列同一性を有するヌクレオチド配列（この配列は、ナイスタチン PKS 酵素をコードする配列またはその部分をコードするか、またはその

10

20

30

40

50

配列と相補性であることが好ましい)
を含む核酸分子を提供する。

【0012】

"ナイスタチン PKS 酵素" は以下で詳しく定義するが、上記 (f) 部の文脈で簡単に言えば、マクロライド系抗生物質またはポリケチド部分、好ましくはナイスタチンまたはナイスタチン誘導体またはナイスタチン関連分子の合成、輸送または転移において機能的である酵素またはタンパク質またはポリペプチドを意味する。

【0013】

もう一つの態様において、本発明はまた、(a) 配列番号 1 および／または配列番号 2 に示すヌクレオチド配列；または (b) 配列番号 1 および／または配列番号 2 の相補物であるヌクレオチド配列；または (c) 配列番号 1 および／または配列番号 2 による縮重物であるヌクレオチド配列；または (d) 配列番号 1 および／または配列番号 2 に、配列番号 1 および／または配列番号 2 の相補物に、または配列番号 1 および／または配列番号 2 から誘導されたハイブリダイゼーションプローブまたはその相補物に、高ストリンジェンシー条件下でハイブリダイズするヌクレオチド配列；または (e) 配列番号 1 および／または配列番号 2 と少なくとも 65% の配列同一性を有するヌクレオチド配列（この配列は、ナイスタチン PKS 酵素またはその部分をコードする配列をコードするか、またはその配列と相補性であることが好ましい）を含む核酸分子を提供する。

【0014】

本発明の核酸分子は、単離された核酸分子（換言すれば、天然に通常一緒に存在する諸成分から単離されまたは分離されたもの）であってもよく、または組換えまたは合成核酸分子であってもよい。

本発明の核酸分子は、ナイスタチン A1 PKS 酵素をコードするか、あるいはその一部、例えば単一ドメインをコードする配列をコードするか（またはコードするヌクレオチド配列を含むか）、あるいは機能的または非機能的な遺伝子要素であるナイスタチン A1 PKS 遺伝子クラスター中のヌクレオチド配列を含む。より正確には、本発明の核酸分子は、1 個またはそれ以上のポリペプチドをコードするか、またはマクロライド抗生物質またはポリケチド部分、好ましくはナイスタチンまたはナイスタチン誘導体またはナイスタチン関連分子の合成において機能性活性を有する 1 個またはそれ以上の遺伝子要素を含む。このような機能性活性は、酵素活性、例えばポリケチド部分またはマクロライド分子の合成または輸送または転移に関係する活性であってもよく（これは、マクロライド鎖または環の合成、またはそれに寄与するいずれの段階、またはマクロライド環またはポリケチド鎖の修飾などでありうる）、および／または調節活性、例えば合成に関係する遺伝子またはタンパク質の発現の調節、または合成過程の調節であってもよく、および／または "トランスポーター活性" であってもよい。従って、一般的には、ポリケチドまたはマクロライドの諸々の部分の転移または輸送に、例えば細胞内へ、または細胞から合成された分子の搬入または流出に関連する輸送タンパク質が含まれる。同様に、これに関しては、グリコシル化タンパク質類が含まれ、これらのタンパク質にはサッカライド（例えばマイコサミン(mycosamine)）の生合成および／またはマクロライドまたはポリケチドへのその付着に関係する分子が含まれる。

【0015】

本発明によれば、所望の産物をコードするヌクレオチド配列が好ましいが、プロモーター、プロモーター-オペレーター領域、エンハンサー、他の調節配列などのような機能性遺伝子要素を含むヌクレオチド配列も包含される。従って、本発明の核酸分子は、全 PKS 遺伝子クラスターを含む必要はないが、その一部もしくはその部分、例えば特定の機能または調節配列を有するポリペプチドをコードする部分を含んでもよい。このことは、1 個またはそれ以上の遺伝子、および／または 1 個またはそれ以上の調節配列、および／または 1 個またはそれ以上のモジュールまたは酵素ドメイン、または非コードの、またはコード機能性の遺伝子要素（例えば遺伝子の発現、転写、翻訳などを制御する要素）

を含んでいてもよいことである。

【0016】

一つの態様において、本発明は、配列番号35の上記ヌクレオチド配列（または上記（b）～（f）で規定したような、その変異体（variant））が ORF 1 を含む分子の諸々の部分を含有しないことを特徴とする、上記で定義した核酸分子を提供する（以下の表1参照）。換言すれば、この態様において、配列番号35のヌクレオチド配列は、ヌクレオチド 124026～125222 を含まない。

【0017】

他のそうした態様において、本発明は、配列番号35の上記ヌクレオチド配列（または上記（b）～（f）で規定したような、その変異体）が ORF 2 を含む分子の諸々の部分を含有しないことを特徴とする、上記で定義した核酸分子を提供する（以下の表1参照）。換言すれば、この態様において、配列番号35のヌクレオチド配列は、ヌクレオチド 122812～123876 を含まない。

【0018】

もう一つのこのような態様において、本発明は、配列番号35の上記ヌクレオチド配列（または上記（b）～（f）で規定したような、その変異体）が NysF を含む分子の諸々の部分を含有しないことを特徴とする、上記で定義した核酸分子を提供する（以下の表1参照）。換言すれば、この態様において、配列番号35のヌクレオチド配列は、ヌクレオチド 454～1191 を含まない。

【0019】

その代わりに、本発明は、ORF 1、ORF 2 および／または NysF 遺伝子の産物の発現または機能が除去されるように、ORF 1、ORF 2 および／または NysF の一部分、または ORF 1、ORF 2 および／または NysF の修飾された配列を含有する核酸分子を提供するものである。

配列番号1または2または35またはその相補物に、またはそれらの部分（すなわち、以下により詳細に論じる、配列番号1または2または35から派生したハイブリダイゼーションプローブ）に、高ストリンジェンシー条件下でハイブリダイズし、そして好ましくは、ナイスタチン PKS 酵素またはその部分をコードする配列をコードするか、またはその配列と相補的であるヌクレオチド配列が、本発明の範囲に含まれる。高ストリンジェンシーの条件は、例えば Sambrook ら, 1989, Molecular Cloning, A Laboratory Manual, 第2版に記載されたような当分野でよく知られた技術により、容易に決定することができる。本発明の範囲に含まれ、ハイブリダイズする配列は、非ストリンジェントな条件（室温で 6 x SSC/50% ホルムアミド）において結合する配列であり、高ストリンジェンシーの条件（例えば 0.1 x SSC, 68℃）においては洗浄されるものである。ここで、SSC = 0.15 M NaCl, 0.015 M クエン酸ナトリウム, pH 7.2 である。

【0020】

ハイブリダイゼーションプローブは、配列番号1または配列番号2または配列番号35の配列（または相補的配列）の部分であってもよく、この部分は、高ストリンジェンシー条件下でハイブリダイゼーションが起きるかどうかを決定するためのサンプルまたは検体の核酸配列にハイブリダイズするように機能するのに十分な塩基長および組成のものである。従って、このプローブは、少なくとも15塩基の長さ、好ましくは少なくとも30、40、50、75、100または200塩基の長さである。従って、代表的なプローブの長さとしては、30～500塩基、例えば30～300、50～200、50～150、75～100が挙げられる。

【0021】

ハイブリダイゼーションプローブは、配列（すなわち、配列番号1または2または35、またはその相補物）のコードまたは非コードの、機能性または非機能性部分から誘導することができる、例えば遺伝子またはモジュールに、または酵素ドメインに、またはその部分（例えば活性部位をコードする部分）に、または酵素ドメインまたはモジュールを連結する配列に対応していてもよい。従って、ハイブリダイゼーションプローブは、上記で定

義したようにポリケチド／マクロライド合成において機能的活性を有することができる。

【0022】

ヌクレオチド配列の同一性は、ウィスコンシン大学からの Genetics Computer Groups (GCG) Version 10 ソフトウェア・パッケージの BestFit プログラムを用いて決定することができる。このプログラムには、デフォルト値：ギャップ生成ペナルティー = 50、ギャップ延長ペナルティー = 3、平均マッチ = 10,000、平均ミスマッチ = -9.000 を用いる Smith および Waterman の局部相同性演算法(local homology algorithm)が使用される。

【0023】

本発明に係るヌクレオチド配列は、配列番号1または2または35と少なくとも65%、70%、80%、85%、90%、95%または98%の配列同一性を示すことができ、好ましくは、ナイスタチン PKS 酵素またはその部分をコードする配列をコードするか、またはその配列と相補的である。本明細書で定義した%配列同一性の基準を満たすヌクレオチド配列は、"実質的に同一な" 配列と考えることができる。

【0024】

従って、配列番号1および／または2に関して本発明の核酸分子を定義する場合、本発明の核酸分子は、配列番号1または"配列番号1-変異体" 配列（すなわち、配列番号1と相補的なまたは縮重した配列、または上記で定義したようなハイブリダイズする配列または実質的に同一の配列といった機能的に等価な変異体）、または配列番号2または"配列番号2-変異体" 配列、またはその両方を含むことができる。配列番号1／配列番号1-変異体の配列および配列番号2／配列番号2-変異体の配列の両方を含む核酸分子が好ましい。

【0025】

本明細書で言及する、"機能的に等価な変異体" または "機能的な等価物" は、これらが関係する本体 (entity)（またはこれらが誘導される本体）の少なくとも1つの機能を保持しており、例えば実質的に同一の特性を有するタンパク質をコードするか、あるいは実質的に同一の調節活性または他の機能的特性または機能性活性を示す。

【0026】

上記のように、配列番号35または配列番号1および／または2（またはそれらの相補的な、縮重した、または機能的に等価な変異体）のヌクレオチド配列の諸部分または一部（例えばフラグメント）を含む核酸分子もまた、本発明の範囲に包含される。

PKS 遺伝子クラスターのこのような諸々の部分または一部分は、有利には、完全な配列の機能上の等価物と考えることもできる。例えば、配列の一部またはフラグメントは、上記で定義した機能性活性、例えばポリケチドまたはマクロライド生合成における酵素的、調節的またはトランスポーター活性を保持していてもよい。

【0027】

好都合には、PKS 遺伝子クラスター（例えば配列番号35、または1および／または2）の諸部分または一部は、少なくとも15塩基の長さ、好ましくは少なくとも20、25、30、35、40、50、70、100、200、300、400、500、1000、2000、5000、10,000、15,000、20,000、30,000または50,000塩基である。かくして典型的なフラグメントの長さとして、15～50,000塩基、具体的には50～30,000塩基、あるいは100～20,000、100～10,000、200～5,000、もしくは200～2,000である。この諸部分または一部は、隣接もしくは非隣接のヌクレオチドまたはアミノ酸を含むか、またはコードすることができる。

【0028】

PKS 遺伝子配列の機能性部分の諸部分または一部は、以下でより詳細に論じられる。

PKS 遺伝子クラスターの諸部分または一部は、DNA 分子（またはヌクレオチド配列）の非コードのまたは非機能性の部分、例えばプロモーターまたはオペレーター配列、または個々の遺伝子、モジュールまたは酵素ドメインを連結するリンカー配列を含むこともでき

10

20

30

40

50

る。これらは隣接していてもよく、非隣接的であってもよいが、以下で詳細に論じる。

【0029】

上記のように、配列番号35、1および2内の多数の遺伝子および ORF を同定した。このような遺伝子（または上記で定義した、それらの相補的な、縮重した、または機能的に等価な変異体）は、配列番号35、1および2の好ましい "部分" またはフラグメントである。これらを下記の表1にまとめて示す。

【0030】

【表1】

表 1

配列番号 3 5、1 および 2 (それぞれ、全遺伝子クラスター配列
(125401 bp)、Nys 1 (65140 bp) および Nys 2 (27541 bp))
の分子特徴

配列番号 . 35			
始点	終点	遺伝子	説明
1191	454 C	<i>nysP</i>	推定-4'-ホスホパンテテイン トランスフェラーゼ
3092	1275 C	<i>nysG</i>	ABC トランスポーター
4824	3070 C	<i>nysH</i>	ABC トランスポーター
5122	6156	<i>nysDIII</i>	dGDP-マンノース4,6-デヒドラターゼ 同族体
6338	34771	<i>nysI</i>	NysI PKS, モジュール 9-14
34792	51097	<i>nysJ</i>	NysJ PKS, モジュール 15-17
51155	57355	<i>nysK</i>	NysK PKS (モジュール18 + TE)
57503	58685	<i>nysL</i>	P450 モノオキシゲナーゼ NysL
58980	58788 C	<i>nysM</i>	フェレドキシン NysM
60241	59047 C	<i>nysN</i>	P450 モノオキシゲナーゼ NysN
61296	60240 C	<i>nysDII</i>	推定アミノトランスフェラーゼ
62837	61317 C	<i>nysDI</i>	推定 UDP- グルクロノシルトランスフェラーゼ
63067	67167	<i>nysA</i>	NysA PKS, 負荷モジュール
67213	76791	<i>nysB</i>	NysB PKS, モジュール 1 および 2
76811	110101	<i>nysC</i>	NysC PKS, モジュール 3-8
110521	111276	<i>nysE</i>	推定チオエステラーゼ
111666	114566	<i>nysRI</i>	転写アクチベーター
114590	117451	<i>nysRII</i>	推定転写アクチベーター
117441	120224	<i>nysRIII</i>	推定転写アクチベーター
120676	121308	<i>nysRIV</i> (short)	推定応答レギュレーター
120628	121308	<i>nysRIV</i> (long)	推定応答レギュレーター
121997	122758	<i>nysRV</i>	推定リプレッサー
123876	122812 C	ORF2	推定転写レギュレーター
124026	125222	ORF1	推定ペプチターゼ (アミノヒドロラーゼ)

注: "C" はその遺伝子が相補DNAストランドでコード
されることを示す。

【0031】

【表2】

nys 1 (配列番号 . 1)

始点	終点	遺伝子	説明
1035	3 C	<i>nysD2</i>	推定 アミノトランスフェラーゼ
2447	1058 C	<i>nysD1</i>	推定 UDP- グルクロノシルトランスフェラーゼ
2806	6904	<i>nysA</i>	ナイスタチンPKS, 負荷モジュール
6952	16528	<i>nysB</i>	ナイスタチンPKS, モジュール 1 および 2
16550	49838	<i>nysC</i>	ナイスタチンPKS, モジュール 3-8
50227	51013	<i>nysE</i>	推定チオエステラーゼ
51405	54303	<i>nysR1</i>	推定転写アクチベーター 1
54329	57188	<i>nysR2</i>	推定転写アクチベーター 2
57180	59961	<i>nysR3</i>	推定転写レギュレーター 3
60367	61045	<i>nysR4</i>	推定応答レギュレーター
61736	62495	<i>nysR5</i>	推定リプレッサー
63615	62553 C	ORF2	推定転写レギュレーター
63765	64959	ORF1	推定ペプチターゼ (アミノヒドロラーゼ)

nys 2 (配列番号 . 2)

始点	終点	遺伝子	説明
1191	456 C	<i>nysF</i>	推定 4'-ホスホパンテテイン トランスフェラーゼ
3092	1277 C	<i>nysG</i>	推定 ABC トランスポーター
4824	3072 C	<i>nysH</i>	推定 ABC トランスポーター
5122	6154	<i>nysD3</i>	推定 GDP-マンノース 4,6- デヒドラターゼ
6338	27541	<i>nysI</i>	ナイスタチンPKS, モジュール 9-13 (不完全)

【 0 0 3 2 】

上記の表中の "C" は相補性ストランドを指す。

配列番号 1 の *nysD2*、*D1*、*A*、*B*、*C*、*E*、*R1*~*R5*、*ORF1*、*ORF2*、ならびに配列番号 2 の *nysF*、*G*、*H* および *D3*、および部分的 *nysI* 配列は、遺伝子クラスターについての完全コード配列の全体を表す配列番号 3 5 において、それらの指定対応物 (counterpart) に対応しており、配列番号 1 および 2 のヌクレオチド配列を含むと認識されよう。しかしながら、配列番号 2 および 3 5 中の対応する特徴部間にあるように、ヌクレオチドの番号付け

10

20

30

40

50

には相違がある；配列番号 1 および 2 の場合は、停止コドンにおける最初のヌクレオチドが遺伝子の終点と考えられる一方で、配列番号 3 5 の場合は、停止コドンの第 3 番目のヌクレオチドが遺伝子の終点と考えられる。NysD1、D2 および D3 は、NysDI、DII および DIII としても知られており、nysR1～R5 は nysRI～RV としても知られている。

【0033】

配列番号 1 (nys1) および遺伝子配列 nysRI に関して、表 1 は、これがヌクレオチド 51405～54303 を含むことを示している。しかしながら、その後の配列分析から、実際には 2 つの始点コドン、すなわち GTG をコードするヌクレオチド 51405-51407、および ATG をコードするヌクレオチド 51408-51410 があることが明らかになった。従って、配列番号 1 のヌクレオチド 51408 は、nysRI のための代わりの始点ヌクレオチドである。ATG

10

は GTG よりも始点コドンとして好ましく、その結果、51408 は 今後nysRI に関しての始点ヌクレオチドと見なされる。配列番号 3 5 における nysRI の始点は、上記の表 1 において ATG コドンとして示されており、下記の配列番号 9 に示す nysRI の翻訳産物は、配列番号 1 のヌクレオチド 51408-54303 から導き出される。[配列番号 1 のヌクレオチド 51405 が始点ヌクレオチドである NysRI 翻訳産物の代わりの表示において、配列番号 9 は追加の "最初の" アミノ酸、V を挿入することによって修飾される。]

配列番号 3 5 および 1 ならびに遺伝子配列 nysRIV に関して、表 1 は、これが配列番号 3 5 中のヌクレオチド 120676～123108 (nysRIV 短) を含むことを示している。しかしながら、さらなる配列分析により実際には 2 つの始点コドン、すなわち GTG をコードするヌクレオチド 120676-120678 および GTG をコードするヌクレオチド 120628-120630 があることが明らかになった。これらの始点コドンは、配列番号 1 においてnysR4 のためのヌクレオチド 60367-60369 (表 1 に示すとおり) および 60415-60417 での始点コドンに対応している。上流 GTG (配列番号 3 5 中の 120628-120630、これは配列番号 1 中の 60367-60369 に対応している) は、始点コドンとして好ましく (実施例 6 参照)、その結果として、120628 は 今後nysRIV に関しての始点ヌクレオチドと見なされる。配列番号 3 5 中の nysRIV の始点は、下流の始点コドン120676-120678として上記の表 1 に示されており (nysRIV 短)、本明細書で "NysRIV 短" と名付けた nysRIVの推定翻訳産物は、配列番号 1 2 に示されている。また表 1 には、配列番号 3 5 における nysRIV (nysRIV 長) の代わりの始点コドンが、上流始点コドン 120628-120630 として示されている。従って、本明細書で "NysRIV (長)" と名付けた、 NysRIV 翻訳産物の代わりの好ましい表示は

20

30

【0034】

配列番号 3 5、または 1 および／または 2 の配列のもう一つの代表的部分は、それぞれの "始点" および "終点" ヌクレオチド位置の間にあるヌクレオチドを、個別的にまたはひとまとめに含んでいる。

それぞれの "遺伝子" の翻訳産物が導き出され、アミノ酸配列は下記に示す次の配列番号に表わされる。

【0035】

【表 3】

<u>遺伝子産物</u>	<u>配列番号</u>	
NysD2 (NysDII) (部分)	3	
NysD1 (nysDI)	4	
NysA	5	
NysB	6	
NysC	7	
NysE	8	
NysR1 (NysRI)	9	10
NysR2 (NysRII)	10	
NysR3 (NysRIII)	11	
NysR4 (NysRIV) (短)	12	
NysR4 (NysRIV) (長)	43	
NysR5 (NysRV)	13	
ORF2	14	
ORF1	15	
NysF	16	20
NysG	17	
NysH	18	
NysD3 (NysDIII)	19	
NysI	20	
NysDII (完全)	36	
NysI (完全)	37	
NysJ	38	
NysK	39	
NysL	40	30
NysM	41	
NysN	42	

【0036】

配列番号4、10、12、13、14、16、18、41および43は、最初のアミノ酸としてバリン（V）を示す。この分野の慣行および細菌 DNA の概念上の翻訳によれば、タンパク質（翻訳産物）の最初のアミノ酸は、始点コドンと関係なく、常にメチオニン（M）である。従って、本発明の翻訳産物およびアミノ酸配列は、提示したように配列番号4、10、12、13、14、16、18、41および43だけでなく、最初のVがMと置換された上記配列の修飾物をも含んでいる。以下で、配列番号4、10、12、13、14、16、18、41および43についての言及は、提示した配列だけでなく、最初のVがMに置換された上記配列をも含むと理解されるであろう。

【0037】

代わりの態様からみると、本発明はまた、配列番号3～20または36～43から選択される1個またはそれ以上のアミノ酸配列をコードするヌクレオチド配列、またはそれと相補的であるか、またはそれと縮重したヌクレオチド配列を含む核酸分子をも提供する。

また、配列番号3～20または36～43の何れか1つと少なくとも60%の配列同一性を示す、1個またはそれ以上のアミノ酸配列（すなわちポリペプチド）をコードするヌクレオチド配列を含む核酸分子が提供される。

【0038】

本発明のもう一つの態様は、上記で定義した本発明の核酸分子によりコードされるポリペプチドが提供される。

より詳細には、本発明のこの態様は、(a) 配列番号3～20または36～43の何れか1個またはそれ以上に示されるアミノ酸配列の全部または部分；あるいは (b) 配列番号3～20または36～43の何れか1個またはそれ以上と少なくとも60%の配列同一性を有するアミノ酸配列の全部または部分を含むポリペプチドを提供する。

【0039】

特に、前記アミノ酸配列は、配列番号3～20または36～43の何れか1個のポリペプチドと少なくとも60%、65%、70%、75%、80%、85%、90%、95%または98%の配列同一性を示すことができる。あるいは前記アミノ酸配列は、配列番号3～20または36～43の何れか1個のポリペプチドと少なくとも70%、75%、80%、85%、90%、95%または98%の類似性を示してもよい。本明細書における%配列同一性または類似性の基準を満たすアミノ酸（ポリペプチド）配列は、"実質的に同一"と考えられる。本発明のポリペプチドは、単離、精製または合成されたポリペプチドであってよい。"ポリペプチド"という用語は、本明細書において、2個またはそれ以上のアミノ酸のアミノ酸配列を包含するために用いられ、すなわち、短いペプチドおよびより長いもの（すなわち、ポリペプチド）の両方が含まれる。

10

【0040】

アミノ酸配列の同一性または類似性は、ウィスコンシン大学からの Genetics Computer Groups (GCG) Version 10 ソフトウェア・パッケージの BestFit プログラムを用いて決定することができる。このプログラムには、デフォルト値：ギャップ生成ペナルティー = 8、ギャップ延長ペナルティー = 2、平均マッチ = 2.912、平均ミスマッチ = -2.003 を用いる Smith および Waterman の局部相同性演算法が使用される。

20

【0041】

配列番号3～20または36～43（または上記で定義した "実質的に同一の" 配列）の何れか1個のアミノ酸配列の "部分" は、少なくとも20の隣接アミノ酸、好ましくは30、40、50、70、100、150、200、300、400、500、1,000、2,000、5,000または10,000の隣接アミノ酸を含むことができる。

【0042】

ポリペプチド、および好ましくはその部分についても、上記の定義によれば機能的に活性であり、例えば酵素的に活性であるか、または調節または輸送の機能的活性を有する。この部分は、自らは機能的に活性でないことがあるが、場合によっては全体として機能的特性を有する領域を与えることができ、例えば、酵素活性に必要とされる活性部位または補因子の結合部位である。

30

【0043】

以下の実施例に記載する研究では、本発明のヌクレオチドおよびポリペプチド配列が特性決定され、その中にある種々の機能性領域が同定された。このような機能性領域は、本発明の別の態様を形成する。種々の翻訳産物（すなわち配列番号3～20および36～43の遺伝子産物）について、これらの機能的領域を下記の表2にまとめて示す。

40

【0044】

【表4】

表 2

配列番号 3 ～ 2 0 および 3 6 ～ 4 3 の翻訳産物の分子特徴

(i) 配列番号 : 5

翻訳産物の名称 : NysA

始点 AA	終点 AA	名称	説明
8	430	KS ^a	KS ドメイン, 負荷モジュール
528	841	AT	AT ドメイン, 負荷モジュール
855	1055	DH	DH ドメイン, 負荷モジュール
1285	1359	ACP	ACP ドメイン, 負荷モジュール

10

(ii) 配列番号 : 6

翻訳産物の名称 : NysB

始点	終点	名称	説明
42	462	KS1	KS ドメイン, モジュール 1
578	897	AT1	AT ドメイン, モジュール 1 (mMCoA- 特異的)
911	1110	DH1	DH ドメイン, モジュール 1 (不活性)
1201	1447	KR1	KR ドメイン, モジュール 1
1484	1559	ACP1	ACP ドメイン, モジュール 1
1579	2004	KS2	KS ドメイン, モジュール 2
2117	2439	AT2	AT ドメイン, モジュール 2 (mMCoA- 特異的)
2453	2659	DH2	DH ドメイン, モジュール 2 (不活性)
2749	2996	KR2	KR ドメイン, モジュール 2
3025	3102	ACP2	ACP ドメイン, モジュール 2

30

40

【 0 0 4 5 】

【表 5】

(iii) 配列番号 : 7

翻訳産物の名称 : NysC

始点	終点	名称	説明	
35	455	KS3	KSドメイン,モジュール3	10
546	858	AT3	ATドメイン,モジュール3	
872	1073	DH3	DHドメイン,モジュール3	
1381	1628	KR3	KRドメイン,モジュール3	
1662	1735	ACP3	ACPドメイン,モジュール3	
1757	2180	KS4	KSドメイン,モジュール4	
2291	2603	AT4	ATドメイン,モジュール4	
2617	2818	DH4	DHドメイン,モジュール4	
3124	3371	KR4	KRドメイン,モジュール4	
3407	3480	ACP4	ACPドメイン,モジュール4	
3501	3924	KS5	KSドメイン,モジュール5	20
4032	4346	AT5	ATドメイン,モジュール5	
4360	4561	DH5	DHドメイン,モジュール5	
4953	5239	ER5	ERドメイン,モジュール5	
5248	5495	KR5	KRドメイン,モジュール5	
5528	5601	ACP5	ACPドメイン,モジュール5	
5623	6046	KS6	KSドメイン,モジュール6	
6166	6478	AT6	ATドメイン,モジュール6	
6492	6704	DH6	DHドメイン,モジュール6	
7038	7281	KR6	KRドメイン,モジュール6	30
7315	7388	ACP6	ACPドメイン,モジュール6	
7408	7831	KS7	KSドメイン,モジュール7	
7939	8253	AT7	ATドメイン,モジュール7	
8267	8470	DH7	DHドメイン,モジュール7	
8812	9086	KR7	KRドメイン,モジュール7	
9120	9193	ACP7	ACPドメイン,モジュール7	
9214	9637	KS8	KSドメイン,モジュール8	
9758	10072	AT8	ATドメイン,モジュール8	
10086	10289	DH8	DHドメイン,モジュール8	40
10657	10904	KR8	KRドメイン,モジュール8	
10939	11012	ACP8	ACPドメイン,モジュール8	

【0046】

【表6】

(iv) 配列番号 : 9

翻訳産物の名称 : Nys R1

始点	終点	名称	説明
42	49	P-ループ	ATP/GTP 結合部位モチーフ A
904	932	HTH	LuxR-型 ヘリックス-ターン-ヘリックス モチーフ (DNA 結合)

(上記 NysR1 の代わりの表示において、配列番号1の nt 51405 が開始コドンとみなされる場合、これらの開始およびアミノ酸終結の構成員はそれぞれ1だけ増えるであろう。)

10

(v) 配列番号.: 10

翻訳産物の名称 : NysR2

始点	終点	名称	説明
902	930	HTH	LuxR-型 ヘリックス-ターン-ヘリックス モチーフ (DNA 結合)

20

(vi) 配列番号.: 11

翻訳産物の名称 : NysR3

始点	終点	名称	説明
26	47	LZ	ロイシン ジッパー モチーフ (DNA 結合)
548	568	TM1	膜貫通ドメイン (推定)
583	610	TM2	膜貫通ドメイン (推定) n
884	912	HTH	LuxR ヘリックス-ターン-ヘリックス モチーフ

30

(vii) 配列番号.: 12

翻訳産物の名称 : NysR4 (短)

始点	終点	名称	説明
97	104	P-ループ	ATP/GTP 結合部位モチーフ A
149	177	HTH	LuxR ヘリックス-ターン-ヘリックス モチーフ (DNA 結合)

40

【0047】

【表7】

(viii) 配列番号 .: 13

翻訳産物の名称 : NysR5

始点	終点	名称	説明
6	40	HTH DeoR	ヘリックス-ターン-ヘリックス モチーフ (DNA 結合)

10

(ix) 配列番号 .: 14

翻訳産物の名称 : ORF2

始点	終点	名称	説明
186	202	HTH AsnC	HTH モチーフ 特徴

(x) 配列番号 .: 17

翻訳産物の名称 : NysG

始点	終点	名称	説明
31	313	TM	膜貫通領域 1e regions
392	399	P- ループ	ATP/GTP 結合部位
496	510	ABC	ABC トランスポーター特徴 cure

20

30

(xi) 配列番号 .: 18

翻訳産物の名称 : NysH

始点	終点	名称	説明
17	288	TM	膜貫通領域
368	375	P- ループ	ATP/GTP 結合モチーフ A
472	486	ABC	ABC トランスポーター特徴

40

【0048】

【表8】

(xii) 配列番号 : 20

翻訳産物の名称 : NysI (部分)

始点	終点	名称	説明	
34	448	KS9	KS ドメイン, モジュール 9	
572	890	AT9	AT ドメイン, モジュール 9	
904	1123	DH9	DH ドメイン, モジュール 9	10
1443	1686	KR9	KR ドメイン, モジュール 9	
1720	1793	ACP9	ACP ドメイン, モジュール 9	
1813	2236	KS10	KS ドメイン, モジュール 10	
2346	2664	AT10	AT ドメイン, モジュール 10	
2678	2890	DH10	DH ドメイン (不活性), モジュール 10	
2983	3229	KR10	KR ドメイン, モジュール 10	
3266	3339	ACP10	ACP ドメイン, モジュール 10	
3358	3780	KS11	KS ドメイン, モジュール 11	
3898	4217	AT11	AT ドメイン, モジュール 11	20
			(mMCoA- 特異的)	
4231	4432	DH11	DH ドメイン (不活性), モジュール 11	
4523	4770	KR11	KR ドメイン, モジュール 11	
4806	4879	ACP11	ACP ドメイン, モジュール 11	
4801	5325	KS12	KS ドメイン, モジュール 12	
5432	5754	AT12	AT ドメイン, モジュール 12	
5768	5977	DH12	DH ドメイン (不活性), モジュール 12	
6068	6315	KR12	KR ドメイン, モジュール 12	30
6348	6421	ACP12	ACP ドメイン, モジュール 12	
6454	6873	KS13	KS ドメイン, モジュール 13	

【0049】

【表9】

(xiii) 配列番号 : 37

翻訳産物の名称 : NysI

始点	終点	名称	説明	
34	448	KS9	KS ドメイン, モジュール 9	
572	890	AT9	AT ドメイン, モジュール 9	
904	1123	DH9	DH ドメイン, モジュール 9	10
1443	1686	KR9	KR ドメイン, モジュール 9	
1720	1793	ACP9	ACP ドメイン, モジュール 9	
1813	2236	KS10	KS ドメイン, モジュール 10	
2346	2664	AT10	AT ドメイン, モジュール 10	
2678	2890	DH10	不活性 DH ドメイン, モジュール 10	
2983	3229	KR10	KR ドメイン, モジュール 10	
3266	3336	ACP10	ACP ドメイン, モジュール 10	20
3355	3777	KS11	KS ドメイン, モジュール 11	
3898	4217	AT11	AT ドメイン (メチルマロニル-CoA-特異的), モジュール 11	
4231	4432	DH11	不活性 DH ドメイン, モジュール 11	
4523	4769	KR11	KR ドメイン, モジュール 11	
4806	4879	ACP11	ACP ドメイン, モジュール 11	
4901	5325	KS12	KS ドメイン, モジュール 12	
5432	5754	AT12	AT ドメイン, モジュール 12	
5768	5977	DH12	不活性 DH ドメイン, モジュール 12	30
6068	6315	KR12	KR ドメイン, モジュール 12	
6348	6421	ACP12	ACP ドメイン, モジュール 12	
6454	6873	KS13	KS ドメイン, モジュール 13	
6973	7293	AT13	AT ドメイン, モジュール 13	
7307	7448	DH13	不活性 DH ドメイン, モジュール 13	
7535	7774	KR13	不活性 KR ドメイン, モジュール 13	
7813	7886	ACP13	ACP ドメイン, モジュール 13	
7908	8323	KS14	KS ドメイン, モジュール 14	
8430	8741	AT14	AT ドメイン, モジュール 14	40
8755	8962	DH14	不活性 DH ドメイン, モジュール 14	
9050	9296	KR14	KR ドメイン, モジュール 14	
9319	9394	ACP14	ACP ドメイン, モジュール 14	

【0050】

【表10】

(xiii) 配列番号 : 38

翻訳産物の名称 : NysJ

始点	終点	名称	説明
41	464	KS15	KS ドメイン, モジュール 15
578	889	AT15	AT ドメイン, モジュール 15
903	1102	DH15	DH ドメイン, モジュール 15
1446	1731	ER15	ER ドメイン, モジュール 15
1740	1988	KR15	KR ドメイン, モジュール 15
2023	2096	ACP15	ACP ドメイン, モジュール 15
2117	2538	KS16	KS ドメイン, モジュール 16
2635	2953	AT16	AT ドメイン, モジュール 16
2967	3167	DH16	不活性 DH ドメイン, モジュール 16
3257	3500	KR16	KR ドメイン, モジュール 16
3539	3612	ACP16	ACP ドメイン, モジュール 16
3634	4057	KS17	KS ドメイン, モジュール 17
4153	4472	AT17	AT ドメイン, モジュール 17
4486	4725	DH17	不活性 DH ドメイン, モジュール 17
4997	5245	KR17	KR ドメイン, モジュール 17
5277	5350	ACP17	ACP ドメイン, モジュール 17

10

20

(xiii) 配列番号 : 39

翻訳産物の名称 : NysK

始点	終点	名称	説明
34	457	KS18	KS ドメイン, モジュール 18
568	881	AT18	AT ドメイン, モジュール 18
898	1102	DH18	不活性 DH ドメイン, モジュール 18
1416	1663	KR18	KR ドメイン, モジュール 18
1695	1769	ACP18	ACP ドメイン, モジュール 18
1849	2066	TE	チオエステラーゼ ドメイン

30

(xiv) 配列番号 : 43

翻訳産物の名称 : NysRIV (長)

始点	終点	名称	説明
31	85	PAS	PAS様 ドメイン
113	120	P-loop	ATP/GTP 結合部位モチーフ A
165	193	HTH	LuxR ヘリックス-ターン-ヘリックス モチーフ (DNA 結合)

40

【0051】

上述したように、「不活性」とは、他のPKS類におけるDHで見られる活性部位モチーフ、H(X₃)G(X₄)Pを表す保存アミノ酸配列を欠いているDHドメインを表示している。しかしながら、これらのドメインはある活性を有してもよいが、もっともそれはおそらく他のPK

50

S類におけるDHドメインの活性とは、異なっているかも知れないと理解されるだろう。

【0052】

配列番号5 (NysA)、6 (NysB)、7 (NysC)、20および37 (NysI)、38 (NysJ) および39 (NysK) は、実際の"PKS"酵素、すなわちポリケチド合成に関わる酵素を構成することはわかるであろう。これらの遺伝子産物は、同定可能な酵素のドメインおよびモジュールを含んでおり、これらは、上記の表2ならびに図4、8および9（下記の実施例1および表4、ならびにナイスタチン生合成遺伝子クラスターのDNA配列分析を詳細に記載している実施例4もまた参照のこと。）にも示されている。そうした個々のドメインおよび分子は、本明細書において同定されるように本発明の別の面を形成している。

【0053】

配列番号3および26 (NysDII)、4 (NysDI)、8 (NysE)、16 (NysF)、19 (NysDIII)、40 (NysL)、41 (NysM) および42 (NysN) は、ポリケチドまたはマクロライド合成、すなわち、PKSからのポリケチド放出、翻訳後のPKS修飾およびポリケチド修飾に機能する他の酵素類を表す。配列番号10~15および43 (NysRIからNysRVおよびORF2) はそれぞれ転写レギュレーターを表し、配列番号17および18 (NysGおよびNysH) は、細胞からのポリケチド輸送に関与していると推測されている「輸送 (transport)」タンパク質を表している。このことは、以下の実施例においてもさらに詳細に記載されている。そのような機能的タンパク質は本発明の別個の面を形成するものである。さらにそうしたタンパク質の機能性の部分またはフラグメント、すなわち、それらが由来する「親分子」（すなわちタンパク質全体またはポリペプチド全体の）の生物活性を（すなわち測定可能なレベルに示す）保持する活性な部分またはフラグメントもまた含まれる。

【0054】

本発明のヌクレオチド配列は、ナイスタチン生合成を操るため、新しいナイスタチン誘導体または新規なポリケチドもしくはマクロライド構造を合成するため、ならびに新規なもしくは改変PKS系を提供するための多くの方法において利用できる重要なツールと情報を提供する（ここでは「PKS系」によってポリケチド合成系、すなわち遺伝子クラスターまたはタンパク質の複合体、集合体または構築物を意味し、ポリケチド合成において機能するが必ずしもPKS酵素類または酵素上のドメインに限定されるものでなく、他の機能性活性、たとえば他の酵素的（例えば修飾的な）タンパク質、トランスポータータンパク質もしくは調節機能的なタンパク質をもまた含むものである）。

【0055】

したがって、たとえば、本明細書に提示されるナイスタチンPKS遺伝子クラスター全体またはPKS合成系、またはその部分は修飾を受けてもよく、それにより1以上の遺伝子、1以上のモジュール、酵素ドメインまたはその内部の機能的配列を修飾することとなる。かくして修飾された、または誘導化されたPKS系は、あとでさらに詳細に記述されるように、新規な、または修飾されたポリケチドの諸部分を合成するのに用いることができる。この場面では、本明細書に記載するナイスタチンPKS系、またはそのフラグメント、部分は、修飾のための「原型」または「鋳型」または「起源」の系または配列として機能するに違いない。

【0056】

さらに具体的には、あるそうした態様において、また後で述べるように、その系の機能的でない部分（たとえば、生物活性ではない部分）は、「足場 (scaffold)」として利用することができ、また機能性部分（たとえば、酵素部分をコードする配列）は、修飾されて派生的な、修飾されたPKS系を生み出すことができる。ある態様では、1個だけが選択された、または少数しか選択されない機能的（たとえば、酵素的な）領域は修飾されてもよく、残りの配列または構造は大部分がそのままに残されるであろう。

【0057】

代わりに、そのような機能性部分は、他の「足場」構造の修飾のためのツールまたは材料として利用することができる。たとえば個々のナイスタチン遺伝子、モジュール、またはドメインは、他のPKS足場構造への導入（たとえば挿入か置換）のために使用されても

10

20

30

40

50

よい。たとえば他のマクロライド抗生物質（たとえばエリスロマイシンまたはラパマイシンなど）用の、PKS系に由来するPKS足場系である。

【0058】

合成のもしくは遺伝子組換えのポリケチドシンターゼ酵素は本発明の範囲に含まれ、これはナイスタチン遺伝子クラスターによりコードされる「足場」から派生したものであり、該酵素は、他のモジュール型酵素から誘導される1以上の機能性単位を含むように修飾される。そのような機能性単位は、触媒性もしくは輸送のタンパク質ドメインをコードすることができるかも知れない。たとえば、PKS酵素からのケトレダクターゼドメイン、モジュール型ハイブリッドポリケチド／ペプチド合成酵素からのACPドメインである。そうしたドメインは、酵素ドメインDNA配列、たとえばポリケチド合成酵素、ペプチド合成酵

10

【0059】

したがって、ナイスタチン生合成遺伝子クラスターについて本明細書に提示される配列および活性の情報は、現存する公知の遺伝子クラスター、結果としてそれらの生産産物を改変するために使用することができる。特に、本明細書に記載される特定のドメインを選

20

【0060】

かくして、極めて一般的な意味において、本発明は、修飾されたPKS系の調製あるいは修飾を受けたポリケチド分子の調製において、本明細書に記載された本発明の核酸分子の利用を提供する。そうした新規なまたは修飾されたポリケチドまたはマクロライド分子は、本発明の別個の面を形成するものである。

【0061】

かようにして本発明によれば、ヌクレオチド配列は、ランダムに、または指定されて、または志向的なやり方で利用することができる。たとえば、特定の所定もしくは所与の構造を獲得してテストしたり、あるいはスクリーニング用の（これもあとでさらに詳細に記載する）ポリケチド構造のライブラリーといったランダム分子を創生する目的のためである。

30

【0062】

ナイスタチン－PKS足場内の修飾のためであろうと、あるいは代替の足場構造への導入のためであろうと、修飾を受け得る遺伝子または遺伝子要素は、実際のPKS遺伝子（NysA、NysB、NysC、NysI、NysJおよびNysK）、それらの個々の分子もしくはそのドメインだけでなく、ナイスタチンの生合成および輸送に関係している他の酵素もしくは機能的タンパク質をコードする遺伝子をも含むものである（本明細書では一括して「PKS遺伝子（類）」または「ナイスタチン遺伝子（類）」という）。

40

【0063】

実際のPKS遺伝子類については、あとでさらに詳細に記載するように、これらは修飾されて、それによりスターター単位（starter unit）の性格、モジュールの数、エクステンダー（extender）の性格を決定する負荷（loading）ドメイン分子の性質のみならず、ポリケチド鎖の構造を決定する様々なデヒドラターゼ、レダクターゼおよびシンターゼの活性を変更することとなる。

【0064】

修飾を受け得る他の遺伝子として、チオエステラーゼ遺伝子（配列番号8のNysEをコードする）が挙げられ、これはPKS系の効率が上昇するように修飾され得るものである（PKS

50

から不適切な基質を取り除く「編集機能」の活性を有するチオエステラーゼの場合）。該チオエステラーゼが単にPKSの最終産物を切り離すだけならば、該酵素はナイスタチンPKSの頭部を切断し、遺伝子工学によって、該エステラーゼをその頭部切断されたタンパク質末端に融合することによって、より小さいマクロラクトン環を有するナイスタチン誘導体を作成するために使用できる。

【0065】

ポリケチドまたは抗生物質の産生を亢進させるため、調節遺伝子のアクチベーターは、強発現され得るし、リプレッサーは不活性化され得る。このことは、組換え株において新しいナイスタチン誘導体を産生するためには特に重要である（それは極めて僅かな量しか生産されないかも知れない）。

10

推定の4'-ホスホパンテテイン（PPT）トランスフェラーゼ遺伝子（配列番号16のNysFをコードする）は、PKSの効率的な翻訳後修飾および完全な機能性を達成するために、強発現され得る。特異的なPPT活性を欠いている異種の宿主におけるナイスタチン（または他の）PKSの発現に使用することも可能である。そのような宿主として、大腸菌、サッカロマイセス・セレビシイなどが挙げられる。

【0066】

合成された分子、たとえば新規なナイスタチン誘導体のグリコシル化を亢進させるように、デオキシ糖遺伝子：グリコシルトランスフェラーゼ（配列番号4のNysDIをコードする）は、強発現され得る。それは、新しい基質に対する特異性を増加するためにインビトロの突然変異誘発によっても修飾を受け得る。このグリコシルトランスフェラーゼの不活性化は、非グリコシル化ナイスタチン（おそらく何らかの修飾も欠いている）を生産する組換え体菌株をもたらし得るであろう。それは、たとえば化学修飾または新しい修飾活性をスクリーニングするための酵素的アッセイに使用することができる。

20

【0067】

アミノトランスフェラーゼ（配列番号3および36のNysDII）は、不活性化されてナイスタチン誘導体を与え得る。この酵素は、デオキシ糖マイコサミン(deoxysugar mycosamine)にアミノ基を付着させると推定されている。この遺伝子もまた、通常はアミノ基を欠いている別のデオキシ糖との同じ反応を遂行するように、他のストレプトマイセス属においても発現され得る。

【0068】

30

ABCトランスポーター（具体的には、配列番号17および18のNysHおよびNysG）は、ナイスタチンの流出とその誘導体の流出をもっと効率的にするように強発現され得る。これらは、異なる化合物に対してその特異性をシフトするように突然変異を受け得る。もし、何らかの理由により細胞内にナイスタチンまたはその誘導体を蓄積することを所望するならば、これらは不活性化させることができる。

【0069】

モノオキシゲナーゼ、NysLおよびNysN（配列番号40および42）をコードする遺伝子は、ヒドロキシル化および酸化をともにされていないナイスタチン誘導体を得られるように、*S. noursei*において不活性化され得る。代わりにこれらの遺伝子は、ナイスタチン前駆体に対するそれらの特異性を変更することをねらって、変異を起こさせ得る。もし上記のヒドロキシル化および／または酸化の段階がナイスタチン生合成経路において律速であるならば、nysLおよびnysNの強発現は、おそらくナイスタチンまたはその誘導体の収率の上昇をもたらすようになるに違いない。ナイスタチン前駆体の修飾のためのインビトロP450ヒドロキシラーゼ系を確立しようと望むならば、フェレドキシン（配列番号41）をコードするnysMを用いる遺伝子操作も有用であるかもしれない。

40

【0070】

かくして、「ナイスタチン」遺伝子を使用することにより、ナイスタチンPKS系の修飾または他のPKS系の修飾に加えて、本発明の核酸分子は、たとえば宿主の範囲を拡張したり、あるいは収率もしくは生産効率の上昇によって、生合成過程を操るかまたは促進するためにも用いることができる。

50

上記の原理に従って、本発明の実施を可能とするために、本発明はさらに発現ベクター、本明細書で定義した核酸分子を含む宿主細胞を提供する。

【0071】

ポリケチドまたはマクロライド分子（具体的にはナイスタチンまたはナイスタチン誘導体）を産生するための方法もまた提供される。その方法は、上記のような核酸分子を宿主細胞内で発現することを含むものである。該核酸分子の発現結果として（すなわち、導入された「PKS合成装置機構」の発現）、宿主細胞内で、ポリケチドまたはマクロライド分子が産生され、宿主細胞により分泌されるか、または輸送され、その後回収される。

【0072】

したがって、本発明の方法は、核酸分子が発現されるような条件のもとで宿主細胞を成長させるか培養すること、および核酸分子の発現産物（類）がポリケチド／マクロライド分子を合成するようにさせること、換言すれば、ポリケチドまたはマクロライド分子が産生される条件下で、宿主細胞を成長させるか、または培養することを含むものである。

【0073】

本発明によれば、組換え核酸分子を調製する方法もまた提供され、その方法は、本発明のヌクレオチド配列を含む核酸分子を別の核酸分子、たとえばベクターの核酸、具体的にはベクターDNA内に挿入することを含むものである。

本発明の発現ベクターは、適当な調節配列を含めることができる。たとえば翻訳調節エレメント（たとえば開始コドンおよび停止コドン、リボソーム結合部位）ならびにリーディングフレームを本発明の核酸分子とマッチさせることに関係する転写調節エレメント（たとえばプロモーターオペレーター領域、終結停止配列）である。

【0074】

本発明のベクターは、当業界で周知であり文献に記録された諸技術によるプラスミドおよびウィルス（バクテリオファージおよび真核生物のウィルスの両方を含む）を含んでもよく、さらに、当業界で周知であり文献に記録された様々な異なる発現システムで発現させることもできる。

発現目的で、原核生物の細胞または真核生物の細胞内にそうしたベクターを導入するための、あるいは形質転換動物を作成するために生殖系細胞菌株または体細胞内に導入するための、様々な技術が知られており、そして用いることができる。適切な形質転換またはトランスフェクション技術が文献に十分に記載されている。

【0075】

上記で規定された本発明の核酸分子を含むトランスジェニック生物体用に、本発明は、形質転換された、あるいはトランスフェクションされた原核生物の宿主細胞または真核生物の宿主細胞をも含むものである。そのような宿主細胞には、たとえば、大腸菌、ストレプトマイセスおよび他の細菌といった原核生物細胞、酵母またはバキュロウィルスー昆虫細胞系といった真核生物細胞、形質転換された動物細胞およびトランスジェニック動植物細胞が挙げられる。

【0076】

上記の発現ベクター、宿主細胞および生物体等の内部に含まれる核酸分子は、あとでさらに詳細に記載するように、上記で定義された核酸分子に由来する派生的な核酸分子であってもよく、そうした分子またはその部分を修飾することによるか、あるいは他の核酸分子に導入したり、または結合するかによって得られる。

よって、ある面で本発明は、ポリケチド類のコンビナトリアル・ライブラリーの製造のため、組換え体の材料を提供する。該ライブラリーのポリケチド構成員は、天然に存在するナイスタチンA1系（本明細書において提示される）から、この系を足場として用いることにより誘導される修飾PKS系によって合成される。一般的には、該ライブラリーの多くの構成員自体が新規な化合物であり、本発明はこれらのライブラリーの新規なポリケチド構成員を含むものである。それゆえ、本発明の方法は個々のポリケチドの調製に適用されてもよい。ポリケチドが新規であろうとなかろうと、該調製方法はそれを調製するためのより便利な方法であってもよい。生成するポリケチドは、さらに修飾を受けてから抗生物

質に変換されてもよい。それは典型的にはグリコシル化を通じてである。本発明のライブラリーをスクリーニングすることによって、所望の結合活性を有する新規なポリケチドを取り出すための方法も本発明に含まれる。

【0077】

したがって、ある面で本発明は、修飾ポリケチドシンターゼ酵素または酵素系をコードするヌクレオチド配列から構成されるか、または含む核酸分子を調製する方法にも適用される。その方法は、本明細書において提示される配列をコードするナイスタチンPKSを足場として用いること、ならびに酵素的または他の機能的活性をコードするヌクレオチド配列の部分を、たとえば突然変異、不活性化（具体的には欠失もしくは挿入による）または置換により修飾することを含むものである。このようにして修飾されたヌクレオチド配列（修飾PKSをコードする）は、適当な宿主細胞を改変するために用いることができ、さらにこのようにして改変された細胞は、未変性のナイスタチンPKSにより産生されるポリケチドとは異なるポリケチドを産生するために用いることができる。未変性ナイスタチンPKSの「足場組み」は、酵素的活性の修飾を支援するために用いられてきた。

10

【0078】

代わりに酵素的または他の機能的活性をコードするヌクレオチド配列の1つ以上の部分は、代替的な（すなわち、異なる「二番目」の）PKS足場（すなわち、「二番目」のPKS系に由来する足場であり、ナイスタチンPKSシステムとは相違している）に導入されてもよい。

本発明は、このようにして産生されたポリケチドおよびそれらが次に変換され得る抗生物質に対しても適用される。

20

【0079】

本発明は、別の面で、細胞コロニーのライブラリーを含む多数の細胞コロニーにも適用される。そのライブラリーでは、各々のコロニーは、異なるモジュール型PKSを産生するための発現ベクターを含んでいるが、上記のように本発明のナイスタチンPKSに由来している。異なるモジュール型PKSのライブラリーは、ある酵素的活性をコードしている天然産「ナイスタチン」遺伝子または遺伝子クラスターの1つ以上の領域を修飾することにより得ることができ、その結果、その活性を変更するが、該天然産遺伝子の足場部分をもとのままにしておく。

【0080】

他の側面では、本発明は、コロニー・ライブラリーを含む細胞コロニーの多様性に関し、ここでライブラリーの各コロニーは、本発明のナイスタチンPKSから誘導された異なるモジュール型PKSを含む。また本発明は、PKS複合体ライブラリーの製造方法およびこれらコロニーの培養によるポリケチド・ライブラリーの製法、ならびにこのようにして製造されたライブラリーに関する。さらに、本発明は、このようにして得られるポリケチド・ライブラリーのスクリーニング方法およびそこに含まれる新規ポリケチドに関する。

30

【0081】

上述したように、ナイスタチンPKS遺伝子クラスター／系の構造的および機能的な配列解析は、下記実施例に与えられ、またDNA配列は配列番号1, 2および35に示され、また上記表1および2でさらに解析され、配列番号3~20および36~42は、また図2, 3, 7, 8および9でさらに解析された。「PKS」遺伝子のモジュールで「ドメイン」をコードする構造が見られるだろう。モジュールは典型的には、ケトシンターゼ(KS)、アシルトランスフェラーゼ(AT)およびアシルキャリア・タンパク質(ACP)を含有してもよい。これらの三つの機能は、伸長単位を活性化するのに十分であり、それを大きくなる分子の残余部に結合する。モジュールに含めてもよい付加的な活性は、クライゼン縮合以外の反応に関係しており、デヒドラターゼ活性(DH)、エノイルレダクターゼ活性(ER)およびケトレダクターゼ活性(KR)を含む。付加モジュールは初期の縮合反応を触媒作し、すなわち、それはATおよびACPによって表わされる「ローディング・ドメイン」で始まり、これはスターター単位の性質を決める。分子の「仕上げ」は、チ

40

50

オエステラーゼ活性 (TE) により調整されると信じられ、これはN y s Kに組み込まれたTE活性により達成されると信じられている。このチオエステラーゼは、マクロライド環の環化を触媒すると思われ、それによりポリケチド生成物の収率を上昇させる。N y s E TE活性は、「編集」活性であると信じられ、ナイスタチンPKS複合体からのある種の実体物の開裂に関与する。

【0082】

上記および後記の配列、図面、および表から、酵素活性をコードする遺伝子の領域およびモジュールは、リンカーまたは「足場」コード領域により分け隔てられていることがわかる。これらの足場領域は、適切な距離および正しい順序で、酵素的／機能性の活性を分けるアミノ酸配列をコードする。したがって、これらのリンカー領域は、全体として種々の活性が、特定の順序および空間的配置でその中に位置付けられている一つの「足場」をコードすると考えることができる。しかしながら言及すべきは、いくつかの場合に足場の領域が、生成するPKSの活性に深刻な影響を与えることなく、切除あるいは改変されることもあるということである。実際、機能上の活性を有するいくつかのドメインをコードする配列が、意味のある効果がないまま完全に、あるいは部分的に削除されることもある。したがって、ここで用いられている「足場」とは、該PKSの機能上の、たとえば酵素的活性に直接的に寄与しないが、その全体的な活性を維持すること、たとえば正しい空間的配向または構造を保持することを担うPKSクラスターのある部分を言及している。かかる足場は、PKSのすべてのリンカーおよび非機能性領域、あるいはその機能上活性である部分（すなわち構造保全の維持）を含んでもよい。このような組織化は、自然界に発生する他のモジュール型PKS遺伝子クラスターに類似している。

【0083】

本発明は、ナイスタチンPKS遺伝子クラスター中に修飾を発生させ、該クラスターにより生成するタンパク質複合が1以上の点で変更された活性を有するようにすることで、ライブラリーあるいは個々の修飾体、究極的にはポリケチドの改変体を与え、またこうしてPKSの自然産物（すなわちナイスタチンA1）以外のポリケチドを製造する。この方法を用いることで、新規ポリケチドが調製され、あるいはポリケチドが概して容易に製造される。自然界に見出されるナイスタチンPKS遺伝子クラスターから誘導される様々な遺伝子群または遺伝子クラスター、各々が未変性クラスターから種々の方法で修飾もしくは改変されたものを多数与えることによって、これらの活性における多様な変動の結果として、ポリケチドの効果的なコンビナトリアルライブラリーが製造される。あるいは、ナイスタチンPKSの「機能領域」（たとえば遺伝子、モジュールまたはドメイン）を、別の自然発生的なPKS系から得られた「足場」に導入するために用いてもよい。したがって、本発明で用いられる、修飾PKSをコードする配列および系は、自然発生ナイスタチンPKS「から誘導される」モジュール型ポリケチドシンターゼを示す。

【0084】

ナイスタチンPKS「から誘導される」モジュール型PKSとは、自然発生遺伝子の有用部分全体の足場を維持するモジュール型ポリケチドシンターゼ（あるいはその対応部がコードする遺伝子）を意味する。（その構成中にすべてのモジュールまたは遺伝子が含まれる必要はない）。定常の足場の上では、少なくとも1つの酵素的または機能性活性が、変異を受け、欠失されあるいは置換され、その結果活性が変化する。天然生成物以外のポリケチドがこれらの集合的な活性から生じるように、そうした活性が削除されたり、活性の異なる型に置換されたり、あるいは単純に変異を受けたりされる場合に、変化が起る。これが起るのは、生成物ポリケチド中の対応する位置においてスターター単位および／または伸長単位、および／または立体化学、および／または鎖長または環化および／または還元的もしくは脱水的な環生成に、結果として生じる変化があるためである。削除された活性が置換される場合、置換活性の起源は、異なる自然発生型ポリケチドシンターゼにおける対応活性から、あるいは同じPKSの異なる領域に由来する。あるいはまた、このような「誘導された」モジュール型PKSは、第2の異なるモジュール型PKS（またはその遺伝子）の足場において、本明細書に記載したナイスタチンPKSから得られるか、ま

たはこれに由来する1つ以上の酵素的または他の機能性活性（あるいはこれらがコードするヌクレオチド配列）を取り込む。

【0085】

モジュール型PKSの修飾または改変操作は、たとえば遺伝子またはドメインまたはモジュールの削除またはドメイン／遺伝子／モジュールの交換、付加、不活性化といった、挿入、削除を含む切断（truncation）を包含している。あるいはまた、ランダムまたは指定的な改変（すなわち突然変異）が、選択された部分（たとえば遺伝子／ドメイン／モジュール等において）のヌクレオチド配列で行われてもよい。

【0086】

前記誘導体は、好ましくは少なくともナイスタチンPKS遺伝子クラスターからのチオエステラーゼ活性を有する。

有利には、ナイスタチンPKS「から誘導される」ポリケチドシンターゼは、ナイスタチンシンターゼ遺伝子のすべてまたは利用される一部によってコードされる足場を含んでもよく、機能性である少なくとも2個のモジュール、好ましくは4個以上のモジュールを含み、またこれら機能性モジュールの諸活性のうち、1以上について突然変異、欠失または置換を含み、かくして得られるポリケチドの性質は変化する。この定義は、タンパク質レベルおよび遺伝子レベルの双方に適用される。特に好ましい態様には、KS, AT, KR, DHまたはERが不活性化または削除、あるいは異なるPKSから、または同じPKS中の別の部分からのある活性型により置換されたものが含まれる。また好ましいものとしては、少なくとも1つの非縮合サイクル酵索性活性(KR, DHまたはER)が削除されているか、あるいはこれらの活性の何れかが突然変異され、最終的に得られるポリケチドが変化するようにしたものが挙げられる。

【0087】

したがって、生成されるポリケチドに関し、ポリケチドシンターゼの構築には5つの自由度がある。第1には、ポリケチド鎖長は、PKS系中のモジュールの数により決定されるであろう。第2には、PKSの炭素主鎖骨格の性質は、各位置におけるマロニル、メチルマロニルまたはエチルマロニルのような伸長単位の性質を決定するアシルトランスフェラーゼの特異性により決定されるであろう。第3には、ローディング・ドメインの特異性はまた、得られるポリケチドの炭素主鎖骨格に影響を及ぼすであろう。したがって、ローディング・ドメインは、アセチル、プロピオニルなどのような異なるスターター単位を使用してもよい。第4には、ポリケチドの様々な位置における酸化状態は、モジュールのデヒドラターゼおよびレダクターゼ部により決定されるであろう。これは、ポリケチド中のケトン、アルコール、二重結合、または単結合の存在と位置を決定するだろう。

【0088】

最後に、得られるポリケチドの立体化学は、シンターゼの3つの側面の機能である。第1の側面は、伸長単位としての置換マロニルに関係したAT/KS特異性に関するもので、それは、デヒドラターゼがキラリティを廃止するため、還元サイクルが失われている場合か、あるいはケトレダクターゼのみを含む場合のみに、立体化学に影響を及ぼす。第2に、ケトレダクターゼの特異性は、いずれの β -OHのキラリティも決定するだろう。最後に、伸長単位としての置換マロニルについてのエノイルレダクターゼ特異性は、完全なKR/DH/ERが得られる場合に、結果に影響するだろう。

【0089】

このように、モジュール型ナイスタチンPKS系は、広範なポリケチドの合成を可能とする。芳香族PKS系と比較して、脂肪族単量体（アセチル、プロピオニル、ブチリル、イソバレリル等）、芳香族類（アミノヒドロキシベンゾイル）、脂環族類（シクロヘキサノイル）およびヘテロ環類（チアゾリル）を含む広範囲のスターター単位が、様々なマクロ環ポリケチド中に見出される。近年の研究によれば、モジュール型PKSは、これらスターター単位について緩和された特異性を有することが示されている。モジュール型PKSはまた、各縮合環中の伸長単位の選択に関して、かなりの多様性を示す。縮合反応に続く β -ケト還元の度合は、遺伝子操作により変更されることが示されている（Donadio, S

ら Proc. Natl. Acad. Sci. USA (1993) 90:7119-7123)。同様に、ポリケチド生成物の大きさは、適当な数のモジュールとともに変異体を設計することで変化させ得る (Kao, C. M.ら J. Am. Chem. Soc. (1994) 116: 11612-11613)。最後にこれらの酵素は、高度に制御された方式で、それらの生成物においてみごとな範囲の非対称中心を生み出すことでよく知られている。本発明の方法により製造されるポリケチドおよび抗生物質は、典型的には、単一の立体異性的な形態をとる。本発明の化合物は、立体異性体の混合物として生じ得るが、この系を使用して、個々の立体異性体を発生させることがより実用的である。ナイスタチン、PKS足場のような自然発生のモジュールに基づくモジュール型PKS経路内におけるコンビナトリアルの可能性は、実質的に制限されない。

【0090】

一般的に、PKSのポリケチド生成物は、抗生物質活性を示すようにするために、典型的にはグリコシル化によりさらに改変される必要がある。ポリケチドのグリコシル化の方法は当業界でよく知られており、グリコシル化は、適当なグリコシル化酵素を与えることにより細胞内で行ってもよく、また化学合成の手段でin vitroで行ってもよい。

【0091】

マクロライド系抗生物質、モジュール型PKSにより合成されるポリケチド部分は、多くの異なるデオキシ糖 (deoxysugars) のいずれかを含んでいてもよい。ナイスタチン分子は、マイコサミンデオキシ糖部 (mycosamine deoxysugar moiety) を含んでいる。デオキシ糖の生合成は、典型的には、グルコース-1-ホスフェートで開始され、dTDP-グルコースシンターゼおよびdTDP-グルコース-4, 6-デヒドラターゼの作用で進行する。後者の生成物、典型的にはdTDP-4, 6-ケト-6-デオキシグルコースには、さらに次の諸反応-エピマー化、異性化、還元、脱水、アミノ交換反応あるいはメチル化-のうちの少なくとも2つが施され、その結果dTDP-D-デオキシ糖を与える。次いで後者は、グルコシルトランスフェラーゼの作用によりマクロラクトン環に結合し、マクロライド化合物のグリコシル化を与える。

【0092】

グリコシル化は、出発物質として、非グリコシル化マクロライドを用い、放線菌 (streptomycetes) の突然変異体あるいはグリコシル化活性を与え得る他の生物 (すなわち、ミクソコッカス、シュードモナス、マイコバクテリウム等) を用いて行うこともできる。あるいはまた、上記の生物からの、グリコシルトランスフェラーゼをコードする遺伝子を、S. nourseiまたは天然または修飾ナイスタチンPKS遺伝子またはその部分を含む生物に導入することができ、これで所望のグリコシル化を与える。ナイスタチン遺伝子クラスターからのデオキシ糖生合成遺伝子は、別のデオキシ糖用の生合成経路を生み出すためのみならず、別のPKS生成生物における対応する活性の補充のために用いることができる。

【0093】

ナイスタチンPKS誘導体は、関係遺伝子の操作によって、あるいは他のPKS中にナイスタチン遺伝子またはその部分を導入することによって調製できる。多数のモジュール型PKS遺伝子クラスターが、マップ化されおよび／または配列決定されており、それらには完全にマップ化および配列決定されているエリスロマイシン、soraphen A、リファマイシン、avermectinおよびrapamycin、部分的に配列決定されているFK506およびオレANDOMycin (oleandomycin)、さらにマップ化され、部分的に配列決定されているカンジジシン、pimaricinおよびnemadectinが含まれる。追加のモジュール型PKS遺伝子クラスターが、時の経過とともに入手できるようになるであろうと期待される。これらの遺伝子は、標準的な諸技術を用いて操作することができ、その結果活性をコードする領域の削除または不活性化、同一または別のPKS系からの対応する諸活性をコードする遺伝子領域の導入、あるいは他にも、標準的手法による突然変異により遺伝的改変体を得られる。もちろん、所望の誘導体をコードする配列の一部または全部を、Jaye ら、J. Biol. Chem. (1984) 259:6331に記載されているような標準的な固相合成法で合成することもでき、さらにたとえばApplied Biosystems, Incから市販されている。

【0094】

天然発生のPKSの多様な誘導体をコードするヌクレオチド配列ならびに、ライブラリー構築のための多様なポリケチドを得るために、所望の数の構築物が、酵素性の活性コード部を「混合および適合」することによって、また変異を、天然のホスト（すなわチニスタチン）PKS遺伝子クラスターまたはその部分に導入することができる。

【0095】

突然変異は、従来法を用いて天然の配列に対してもたらしることができる。突然変異の基質は、遺伝子クラスター全体でもよく、またその1または2つの遺伝子のみでもよく、つまり突然変異の基質はこれら遺伝子の1つ以上の部分であってもよい。突然変異の技術は良く知られており、たとえばW098/49315に記載されており、その文献をここに引用により取り込む。このような技術は、変異を含む合成オリゴヌクレオチドを調製することと、制限エンドヌクレアーゼ消化を用いて、PKSサブユニットをコードする遺伝子内に該変異配列を挿入することを含む。（たとえばKunkel, T.A., Proc. Natl. Acad. Sci. USA (1985) 82:448; Geisselsoder ら BioTechniques (1987) 5:786.参照）あるいはまた、天然のヌクレオチド配列にハイブリダイズする不適合プライマー（一般的に鎖長は10～20ヌクレオチド）を、不適合な二重鎖の融点よりも下の温度で用いることで突然変異を起こさせることもできる。プライマーの鎖長および塩基組成を比較的狭い範囲に保持し、変異塩基を中央部に位置するように保つことで、プライマーを特異的なものにできる（Zoller and Smith, Methods Enzymol. (1983) 100:468）。プライマー伸長は、DNAポリメラーゼを用いて行われ、生成物はクローン化され、プライマーの伸長鎖の分離により誘導された変異DNAを含有するクローンが選別される。変異プライマーをハイブリダイゼーション・プローブとして用いることにより選別が達成される。この技術は、多重点突然変異を起こさせることにも適用できる。たとえば、Dalbie-McFarland ら Proc. Natl. Acad. Sci. USA (1982) 79:6409参照。PCR変異誘発は、所望の突然変異を起こさせるための利用の途もあるだろう。

【0096】

酵素性の諸活性をコードするヌクレオチド配列の選択された部分におけるランダムな突然変異誘発は、いくつかの異なる公知の方法によって達成される。たとえばプラスミド中にオリゴヌクレオチドリンカーをランダムに挿入すること、X線または紫外線を照射すること、in vitroのDNA合成中に不正ヌクレオチドを取り込ませること、誤りがちなPCRの突然変異誘発によること、合成変異体を調製することで、あるいは化学物質によってプラスミドにin vitroで損傷を加えることで達成される。突然変異誘発性化学物質としては、たとえば亜重硫酸ナトリウム、亜硝酸、ニトロソグアニジン、ヒドロキシルアミン、ヒドラジンあるいは蟻酸のように、塩基を損傷するかまたは除去して正常な塩基対の形成を阻害する試薬、5-ブロモウラシル、2-アミノプリンのようなヌクレオチド前駆体の類似体、あるいはプロフラビン、アクリフラビン、キナクリンなどのようなアクリジン挿入化剤等が挙げられる。一般的に、プラスミドDNAまたはDNAフラグメントは、化学物質により処理され、E. coliに形質転換され、変異プラスミドのプールもしくはライブラリーとして増殖される。

【0097】

酵素的または他の機能性の活性をコードする領域について変異形態を提供することに加えて、所望する諸々の機能性または活性をコードする領域が、同一のチニスタチンPKSの異なる部位から、たとえば適当なプライマーを用いるPCR技術によって回収され得る。「対応する」活性をコードする領域とは、活性の同一な一般的タイプをコードする領域を意味し、--たとえば、ある遺伝子クラスターの一つの場所におけるケトレダクターゼ活性は、その遺伝子クラスターの別の場所におけるケトレダクターゼをコードする活性に「対応」する。

【0098】

（ホストのチニスタチンPKS、あるいは「チニスタチン」配列が挿入されることとなる異なるPKSとするために）このホストのポリケチドシンターゼにおける特定のターゲット領域の置換を行おうとする場合には、この置換は、好適な制限酵素を用いてin vitroで行

うか、あるいはドナープラスミドに置換遺伝子を枠構成する相同配列、およびレシピエントプラスミド、ゲノムまたは染色体中にレセプター領域を含む組換え技術を用いてin vivoで行うことが可能である。このような系は、好ましくは異なる温度感受性を持つプラスミドを含むものであり、たとえばPCT出願WO 96/40968に記載されている。

【0099】

WO 00/77181には、配列においていくつかのDNAユニットを集合させて大きいDNA構築物とする方法が記載されている。これらの方法は、本発明の範囲内の組換えポリケチドシンターゼに対しても適用できる。ホストPKS遺伝子における酵素的活性を置き換えるか、あるいはホストPKS遺伝子のそのような領域の変異をサポートする目的で、種々の操作を実施するために使用されるベクターは、次のような調節配列を含むように選択してもよい。コード配列の発現が適当なホスト内で起きるような態様で、該調節配列が、生じたコード配列と作動可能に連結されている。しかしながら、単純なクローニングベクターも同様に使用できる。

10

【0100】

誘導PKSをコードしているPKS遺伝子を得るために用いられるクローニングベクターが、コードするヌクレオチド配列と作動可能に連結された発現調節配列を欠いているならば、そのヌクレオチド配列は好適な発現ベクターに挿入される。これは個別に実施される必要はなく、単離されたコードするヌクレオチド配列のプールを、ホストベクターに挿入することができる。得られるベクターは、宿主細胞内に形質転換され、または形質導入され、その結果、生成する細胞は播種されて各コロニーに培養される。

20

【0101】

好適な調節配列としては、真核および原核宿主細胞内で機能するものが挙げられる。好ましい宿主としては、原核細胞の宿主、および酵母のような菌類が挙げられるが、単一の細胞培養物、たとえば哺乳類の細胞もまた使用できる。しかしながら、そのような系を使用することに有利な点は特にない。Streptomyces sppと適合し得る調節配列を使用する酵母および原核細胞宿主が特に好ましい。種々のタイプの生物についての単一の細胞培養物に好適な調節配列は、当業界でよく知られている。酵母における発現系は、分泌をもたらす調節配列を含み、広く市販されるルーチンに使用される。調節因子として、プロモーター、所望によりオペレーター配列も含んでいてもよく、宿主の特質に応じて他の因子、たとえば、リボソーム結合部位などが挙げられる。特に原核細胞宿主に有用なプロモーターとしては、2次的な代謝産物としてのポリケチドを産生することとなるPKS遺伝子クラスターからのプロモーター、たとえば、芳香族（タイプII型）PKS遺伝子クラスターからのものが挙げられる。具体例として、actプロモーター、tcmプロモーター、スピラマイシンプロモーターなどが挙げられる。しかしながら、他の細菌のプロモーター、たとえば、ガラクトース、ラクトース（tac）およびマルトースといった糖代謝酵素から誘導されるものも有用である。追加の例としては、生合成酵素、たとえば、トリプトファンシンターゼ(trp)、β-ラクタマーゼ(bla)、バクテリオファージラムダ FL、T5およびT7などをコードしている遺伝子から誘導されたプロモーターが挙げられる。さらに、合成されたプロモーター、たとえば、tacプロモーター（米国特許明細書4,551,433号）を用いることもできる。

30

40

【0102】

また、他の調節配列、宿主細胞の成長に関してのPKS置換配列の発現調節を可能にするものが望ましい。調節配列は公知の技術であり、その例としては、化学的または物理的な刺激（たとえば調節性化合物の存在）に応じて、遺伝子の発現を生じさせたり、させなかったりするものが挙げられる。他のタイプの調節因子もまたベクター内に存在してもよく、たとえば、エンハンサー配列が挙げられる。

【0103】

選択可能なマーカーを組換え発現ベクター内に取り込ませることもできる。様々なマーカーが知られ、形質転換された細胞系統を選別する際に有用であり、一般にある遺伝子を含み、その遺伝子の発現は、好適な選択培地で成長させるとき、形質転換細胞に選別可能

50

な表現型を与える。このようなマーカーとしては、たとえば、抗生物質耐性または抗生物質感受性をプラスミドに与える遺伝子が挙げられる。または、いくつかのポリケチドは本来的に着色しており、この特性は、本構築物によって首尾よく形質転換された細胞をスクリーニングするための組み込みマーカーを与える。

【0104】

種々のPKSヌクレオチド配列またはそのような配列の混合物は、個別の制御因子とともに、または、たとえば単一のプロモーターの制御下に、個別のカセットとして1以上の組換えベクターへクローン化できる。前記PKSサブユニットまたは成分の混合物は、隣接する制限部位 (flanking restriction site) を含み、それにより、容易に欠失を可能とし、また他のPKSサブユニットまたは成分混合物を挿入させる。その結果、ハイブリッドPKSが生成できる。このような独特の制限部位の設計は当業者にとり公知の技術であり、上述した技術を用いて達成することができる。たとえば、位置指定突然変異導入法およびPCRである。

【0105】

上述したように、具体的に有用な調節配列は、自らまたは好適な制御系を使用して栄養菌糸体の増殖から定常期までの遷移の間、発現を活性化するものである。プラスミドRM5に含まれる系、すなわち、actI/actIIIプロモーターペアおよびactII-ORF4、活性化遺伝子が、特に好ましい (McDanielら、Science, v.262、P1546-1550,1993)。

【0106】

特に好ましい宿主 (ホスト) は、ポリケチドを産生する独自の手段を欠いており、それにより明確な結果が得られる。このタイプの宿主細胞の実例としては、PCT出願WO 96/40968に記載された、修飾*S.coelicolor* CH999培養物および*S.lividans* の同様の菌株が挙げられる。

修飾されたPKS系、または異なるポリケチドを生成するための種々のPKS系をコードしているヌクレオチド配列を含む発現ベクターは、その後たとえば、ライブラリーを構築するために好適な宿主細胞へ形質転換してもよい。ある率直なアプローチにおいては、前記ベクターの混合物を選択された宿主細胞へ形質転換し、得られた細胞を個々のコロニーに培養し、首尾よく形質転換体を選別する。各コロニーは、それぞれ特定のPKSシンターゼおよび最終的に特定のポリケチドを産生する能力を有するコロニーを表わすであろう。典型的には、いくつかのコロニーは複製されることになり、形質転換されたコロニー群のサブセットは、コロニーごとにそれぞれ異なるPKSを含有しており、ライブラリーとみなすことができる。または、発現ベクターをそれぞれ宿主を形質転換するように使用することができ、次いで形質転換された宿主は、ライブラリーを構築するために集められる。様々な戦略がコロニーの多様性を獲得するために工夫されるべきであり、各コロニーは、ナイスタチン宿主遺伝子クラスターから誘導されるPKS遺伝子クラスターを含有する。その結果、ライブラリーの各コロニーは、異なるPKS、および最終的に異なるポリケチドを生成する。前記ライブラリーによって生成される異なるポリケチドの数は、典型的には少なくとも3 (たとえば、PKS遺伝子中に2つの変異が別個にまたは組み合わせて現れる)、より典型的には少なくとも10、好ましくは少なくとも20、より好ましくは少なくとも50であり、異なる改変PKS遺伝子クラスターおよびPKS遺伝子産物の数をほぼ反映する。ライブラリーの構成員数は任意に選択できるが、スターターユニット、伸長ユニット、立体化学、酸化状態および鎖長に関して上記の自由度は、著しく大きい。

【0107】

本発明の組換えベクターを好適な宿主 (ホスト) へ挿入する方法は、当業者にとり公知の技術であり、典型的には、CaCl₂または他の試薬、たとえば2価のカチオンの使用、リポフェクション、接合、プロトプラスト形質転換およびエレクトロポレーションが挙げられる。

たとえ、いくつかの宿主が本来的に上記シンターゼのアシルキャリアータンパク質を活性化するための好適な翻訳後のメカニズムを含んでいないとしても、幅広い種類のホストを使用できる。これらのホストには、そのような修飾を実現させるのに好適な組換え酵素

、たとえば、NysFを補足することができる。

【0108】

ポリケチドを生成するコロニーは、公知の技術を用いて同定し単離することができ、産生されたポリケチドはさらに特徴付けられる。これらのコロニーによって産生されたポリケチドは、ライブラリーを表す枠内で集合的に使用でき、または個々に活性を評価してもよい。

したがって、前記ライブラリーには4つのレベルが考えられる；（1）異なるPKSクラスターを含有し、異なるPKSをコードする配列をそれぞれ有する多数のコロニー（しかしながら、すべてナイスタチンPKSクラスターから誘導される。）、（2）コード配列により生成された、PKSの一員であるタンパク質を含有するコロニー、（3）産生されたポリケチド、（4）ポリケチドから誘導された抗生物質。

10

【0109】

ライブラリーにおけるコロニーは、関連するシンターゼを生成するように誘導することができ、したがって、ポリケチド候補のライブラリーを得るために関連するポリケチドを生成するように誘導することができる。産生されたポリケチドは、抗菌活性、抗癌活性、抗寄生虫活性、免疫抑制活性に関してスクリーニングできると同様に、望ましいターゲット、たとえば、レセプター、シグナルタンパク質などへの結合に関してスクリーニングできる。上清または培養ペレット自体をスクリーニングに用いることができ、あるいはポリケチドの部分的もしくは完全な精製を最初に行うことができる。典型的には、このようなスクリーニング方法は、ライブラリーの各構成員とレセプターまたは他のターゲットリガンドとの結合を検出することを含む。結合は、直接的に検出するか、または競合アッセイを通じて検出することができる。結合に関してそのようなライブラリーをスクリーニングする手段は、当業者によく知られている。

20

【0110】

または、ライブラリーの各ポリケチド構成員を、望ましいターゲットに対して試験することができる。この場合には、ターゲットの生物学的応答を測定するスクリーニングがより容易に含め得る。

本発明の方法によって多数の新規なポリケチドを、後述する実施例で代表して説明されるようにして調製することができる。これらの新規なポリケチドは、グリコシル化または上記の反応を通じて、抗菌活性を有する化合物を形成する場合において、有用な中間体となるであろう。上記に示したように、個々のポリケチドは、適当な糖誘導体と反応させて、抗菌活性を有する化合物を得ることができる。抗菌活性は、典型的なスクリーニングアッセイ、たとえばLehrer, R.らJ. Immunol. Meth. (1991) 137 :167-173に提案されているようなものを用いて確認することができる。

30

【0111】

したがって、本発明はさらなる面として、修飾ナイスタチンPKSをコードするヌクレオチド配列を含む核酸分子を調製する方法を与える。ここで前記修飾ナイスタチンPKSは、本明細書で定義されたナイスタチンPKS（すなわち、天然に生じるナイスタチンPKS）から誘導されるものであり、酵素活性または他の機能性活性をコードする第一領域と、および足場となるアミノ酸配列をコードする第2領域とを含有する（本明細書で定義された）ヌクレオチド配列によってコードされており、該方法は、（a）前記少なくとも1つの第一領域を修飾すること、あるいは（b）前記少なくとも1つの第一領域を、異なるPKSをコードしているヌクレオチド配列からの足場をコードしている第2領域に取り込ませることを含んでなることを特徴としている。

40

【0112】

上述したように本発明の核酸分子に関して、第一領域は本発明のヌクレオチド配列（すなわち、配列番号1または2または35）のいずれかの部分（たとえば、ドメイン、モジュールまたはそれらの部分をコードしている）であってもよい。

また、（i）上記で定義した修飾ナイスタチンPKSを調製する方法も与える。該方法は、上記で定義したように調製した核酸分子を宿主細胞内で、修飾ナイスタチンPKSが発現

50

するような条件下で、発現させること（すなわち、そのような核酸分子を含む宿主細胞を培養または増殖させる）を含んでなる。また、(ii)そのようにして産生されたか、そのようにして得られうる、修飾ナスタチンPKSをも与える。

【0113】

そのような修飾されたPKSによって産生された新しいポリケチドもまた、本発明の範囲内である。一般に、これらのポリケチドのグリコシル化された形態である新しい抗生物質も同様である（もっとも、あるものは異なる翻訳後修飾、たとえばヒドロキシル化または酸化によりグリコシル化されていなくても活性を有する）。

【0114】

これより、下記の限定的でない実施例において図を参照しつつ、本発明をより詳細に述べることにする。

【0115】

【実施例】

【0116】

【実施例1】

ナスタチン（Nystatin）生合成遺伝子クラスターのクローニング：細菌株、プラスミドおよび増殖条件

この研究で用いられた細菌株およびプラスミドを表3に示す。この研究の中で発育させた新規の細菌株およびプラスミドをここで説明し、図3に示す。S.noursei ATCC 11455およびその変異体を、固体ISP2培地(Difco)および液体TSB培地（Oxoid）中で増殖させた。E.coli ET12567（pUZ8002）からStreptomyces菌株への属間接合を、「ヒートショック」時間を5分短縮した以外は、FlettらがFEMS Microbial Lett.、155巻、223～229頁（1997年）に記載したように実施した。50mg/mlの濃度のアプラマイシン（Apramycin）（Fluka）を用い、固体培地上のS.noursei被接合体を選択した。全DNA単離に先立って、S.noursei ATCC11455被接合体の接種のために、20 μg/mlのアプラマイシンを添加したTSB液体培地を用いた。E.coli菌株を、SambrookらがCold Spring Harbor Laboratory（1989年）に記載したように増殖させ、形質転換させた。E.coli ET12567（pUZ8002）を、20 μg/mlのクロラムフェニコールおよび50 μg/mlのカナマイシン（kanamycin）を含有する培地上に保持した。

【0117】

【表11】

10

20

30

表3 本研究で用いた細菌株、プラスミドおよびファージ

株、プラスミド またはファージ	特性	出所または文献
<i>E. coli</i> DH5α	一般のクローニング宿主	Sambrook et al, 1989, 上述
<i>E. coli</i> XL1-Blue MRA (P2)	遺伝子ライブラリー用宿主	Stratagene
<i>E. coli</i> ET12567	属間接合用株	MacNeil et al, Gene, v. 111: 61-68, 1992
<i>S. noursei</i> ATCC 11455	野生種、ナイスタチン産生株	ATCC
pGEM3Zf(-)	ColE1 レプリコン, Ap ^R , 3.2kb	Promega
pGEM72f(-)	ColE1 レプリコン, Ap ^R , 3.0 kb	Promega
pGEM11Zf(-)	ColE1 レプリコン, Ap ^R , 3.2kb	Promega
pUZ8002	RK2 誘導体, Km ^R , Tc ^R	D.H.Figurski
pWHM3	ColE1+pIJ101 レプリコン, Thio ^R , 7.2kb	Vara et al, J.Bacteriol v. 171: 5872-5881, 1989
pSET152	ColE1 レプリコン+ <i>oC31</i> int, <i>oriT</i> , Am ^R , 5.5kb	Bierman et al, Gene, v. 116: 43-49, 1992
pSOK101	3.1kb BamHI/SphI断片が ColE1, <i>oriT</i> and Am ^R , 7.1kbを含む pSET152 の3.0kb BamHI/SphI断片で置換された pWHM3 誘導体	本研究 (図2参照)
pGM11	pSG5 レプリコン, Neo ^R , 5.3kb	Wohlleben & Muth, In "Plasmids, a practical approach", Ed. Hardy, IRL Press, p 147-175, 1993
pSOK201	1.2kb EcoRI/HindIII断片が ColE1, <i>oriT</i> and Am ^R , 7.1kbを含む pSOK101 の3.0kb EcoRI/HindIII断片で置換された pGM11 誘導体	本研究 (図2参照)
DASHII	バクテリオファージ λ ベクター	Stratagene
pSOK804	ColE1 レプリコン, Am ^R , <i>OriT</i> , int _{WVS} , AttP _{WVS}	本研究 (図2参照) および Van Mellaert et al., Microbiol., v. 144:3351-3358, 1998
pGEM7ermELi	PermE* プロモータ含有 pGEM7ZF プラスミド	C.R. Hutchinson

【0118】

Amーアブラマイシン、Apーアムピシリ、Neoーネオマイシン、Thioーチオストレプトン、Kmーカナマイシン、Tcーテトラサイクリン

二次代謝産物の分析

培養は、グルコース水溶液を 90g l^{-1} 、 NH_4NO_3 を 2.5g l^{-1} 、コーンフラワーを 3.0g l^{-1} 、 $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$ を 0.4g l^{-1} 、 KH_2PO_4 を 0.2g l^{-1} 、 CaCO_3 を 7g l^{-1} 、微量元素溶液を3ml含む1.3 l SAO-23培地が最初に入っているApplicon 3-1培養器中で実施した。微量元素溶液 (g l^{-1}) : $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 5.0 ; $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0.39 ; $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 0.44 ; $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0.15 ; $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 0.01 ; $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$, 0.02 ; HCl , 50。培養は、 HCl (2M) および NaOH (2M) でpHを6.5~7.0に調節しながら、 28°C で実施した。溶存酸素は、攪拌 (300~900rpm) および通気 (0.25vvm) により飽和濃度の40%より高くなるように調節した。SAO-23培地を含む培養用の接種物 (3 vol%) を振盪フラスコ (100mlの培地を含む、500mlの邪魔板付き三角フラスコ、200rpm) 内のTSB培地 (TSB, Oxoid CM129, 37g l^{-1}) 中で、中、 28°C で増殖させた。各振盪フラスコに0.2mlの孢子懸濁液を接種し、18~20時間培養した。培養後、培養物のジメチルホルムアミド抽出物のHPLCによりナイスタチン (Nystatin) 産生を分析した (Raatikainen, J., Chromatogr., 588巻、356~360頁 (1991年))。

10

DNA操作

プラスミド、ファージおよび全DNAの調製、エンドヌクレアーゼ消化、核酸連結ならびに分画を、前述 (Sambrookら、1989年、上述 ; Hopwoodら、1985年、上述) したように実施した。QIAGENキット (QIAGEN GmbH) を用いてDNAフラグメントをアガロースゲルから分離し、ベーリンガー・マンハイムのdigoxigeninキットを用いて標識化し、供給メーカーの手順書に準拠してサザンブロット分析に用いた。DNA配列の決定は、QIAGEN GmbH (ドイツ) で実施し、そのデータをGCGソフトウェア (Devereuxら、Nucleic Acid Res., 12巻、387~395頁、1984年) で分析した。

20

S.noursei ATCC 11455ゲノム由来の、PSKをコードするDNAフラグメントのPCR-支援増幅およびクローニング

ナイスタチン生合成遺伝子をコードするDNAを得るために、S.noursei ATCC 11455遺伝子ライブラリーを、標識化されたPSKをコードするDNAを用いて探索した。DNAプローブを得るために、公知のモジュール型PKSの β -ケトアシルシンターゼ (KS) およびアシルキヤリアータンパク質 (ACP) ドメイン内の保存アミノ酸領域に相当する、縮重オリゴヌクレオチドプライマーを設計した。増幅に用いられる縮重プライマーは、公知PKSのACPおよびKSドメイン中の保存アミノ酸モチーフに相当し、Streptomyces用のコドン使用表 (Wright & Bibb, Gene, 113巻、55~65頁、1992年) に従って設計された。ACPヌクレオチドプライマー (センス) は、Glu (Asp) Leu Gly Phe (Leu, Val) Asp Ser Leuモチーフ (SEQ ID No. : 21) に相当し、5'-GAG/C CTG/C GGC/GT/CTG/C GAC TCC/G CTG/C-3'配列 (SEQ ID No. : 22) を有した。KSヌクレオチドプライマー (アンチセンス) は、Val Asp Thr Ala Cys Ser Serモチーフ (SEQ ID No. : 23) に相当し、5'-G/CGA G/CGA G/ACA/ G/CGC C/GGT GTC G/CAC-3'配列 (SEQ ID No. : 22) を有した。S.noursei ATCC11455から単離された全DNAを、KSおよびACPオリゴヌクレオチドプライマーを用いて、この生物のゲノム由来のDNAフラグメントについてポリメラーゼ連鎖反応 (PCR) -支援増幅用鋳型として用いた。モジュール型PKS上の該モチーフの相対位置から、得られたPCR生産物が約0.7kbの大きさであろうと想定された。50 μl のPCR混合液は、0.1 μg のS.noursei ATCC11455の全DNA、それぞれ25pmのACPおよびKSオリゴヌクレオチドプライマーdNTP (最終濃度350 μM)、Expand High Fidelity PCRシステム (ベーリンガー・マンハイム) の1 x PCR緩衝液、および同じ反応系の1.5 UのDNAポリメラーゼ混合物を含有していた。PCRは、以下のプログラムによりPerkin Elmer GeneAmp PCR System 2400で実施した。 96°C (4分) で1サイクルの変性、 94°C (45秒) および 70°C (5分) で35サイクルの変性/アニーリング/合成、および 72°C (7分) で1サイクルの最終アニーリング/伸長。この手法で得られた0.7kb DN Aフラグメントは、E.coli DH5 α 内のpUC18ベクターでSure Clone Ligationキット (Pharmacia) を用いてクローン化された。得られた組換えプラスミドの1つ3.5kbのpPKS72をDNA配列分析した。

30

40

【0119】

Escherichia coliベクター、pUC18中における後のクローニングおよび得られた0.7kbの

50

PCR生産物のDNA配列決定は、続いて概念的翻訳 (conceptual translation) およびデータベース探索が実施されそれはPKSタイプIの一部をコードすることが確認された。このDNAフラグメントは、ファージベクターDASHIII (下記参照) 中に構築されたS.noursei ATCC 11455遺伝子ライブラリーをスクリーニングするためのプローブとして用い、λ DASHII-N1と称される1つの組換えファージ (該プローブにハイブリダイズする) が単離された。以下、さらに説明するように、DASHII-N1をさらにプローブを形成するために用いた。

S.noursei ATCC11455遺伝子ライブラリーの構築およびスクリーニング

S.noursei ATCC11455遺伝子ライブラリーを、供給メーカーの指示書に従ってファージλベクターDASHII (Stratagene) に構築させた。S.noursei ATCC11455の全DNAを、上述のHopwoodら (1985年) に記載されているように単離し、特に上述のSambrookら (1989年) に記載されているように、Sau3 AI制限酵素を用いて部分的に消化し、スクロース勾配で分画した。13~17kbの大きさのフラグメントを含むS.noursei ATCC11455 DNAの画分を、BamHI制限酵素によって消化されたDASHIIベクターのアームとのライゲーション (ligation) に用いられた。E.coli XL1-Blue MRA (P2) (Stratagene) を遺伝子ライブラリーの構築および伝播のための宿主として用いた。

【0120】

メーカーの指示書に従って、遺伝子ライブラリーをスクリーニングするためのプローブとして用いられるDNAフラグメントを、QIAGENキット (QIAGEN GmbH、ドイツ) を用いてアガロースゲルから精製し、バーリンガーマンハイム (ドイツ) のdigoxigenin (DIG) キットを用いて標識化した。プローブを、ライブラリーのスクリーニングに用い、関係する組換えファージを発見した。

【0121】

【表12】

プローブ	このプローブを用いて見つけられた DASHII 組換え型ファージ
PKS72	N1, N14
E12.1: N1 由来	N41, N42, N44, N45, N48
E4.7.2: N1 由来	N58
L42E9.1: N42 由来	N64, N76
B1.0_58: N58 由来	N69

【0122】

プローブの説明

1. PKS72プローブ。制限酵素EcoRIおよびHindIIIを用いて、pPKS72プラスミド (上記参照) から単離された0.7kbのDNAフラグメントをPKS72プローブとして用いた。
2. E12.1プローブ。組換えファージN1の挿入断片 (insert) に由来する2.6kbのBamHIフラグメント (その左隣接領域 (flanking region) を表わす) をpGEM3Zf(-) にサブクローン化して、プラスミドpGEM(B2.6)-1を得た。このプラスミドをEcoRI/AvaIにより消化し、該N1のDNA挿入断片の左末端に対応する0.55kbのフラグメントを精製し、E12.1プローブとして用いた。
3. E4.7.2プローブ。組換えファージN1由来の4.7kbのEcoRIフラグメント (その右側の隣接領域を示す) を、pGEM3Zf(-) にサブクローン化して、プラスミドpGEM(E4.7)-1を得た。このプラスミドをEcoRI/HindIIIにより消化し、該N1のDNA挿入断片の右末端に対応する1.5kbのDNAフラグメントを精製し、E4.7.2と名づけた。
4. L42E9.1プローブ。組換えファージN42中のDNA挿入断片の左側隣接領域を示す9.0kbのEcoRIフラグメントを、pGEM3Zf(-) にサブクローン化して、プラスミドpH42E9.1を得た。このプラスミドをEcoRI/BamHIにより消化し、該N42のDNA挿入断片の左側隣接に対応する0.6

kbのフラグメントを精製し、L42E9.1プローブとして用いた。

5. L58B1.0プローブ。組換えファージN58のDNA挿入断片の右側隣接領域を示す3.0kbのEcoRI断片を、pGEM3Zf(-)にサブクローン化して、プラスミド pGEM(E3.0)-58を得た。後者のプラスミドから、該N58のDNA挿入断片の右末端に位置する1.0kbのBamHIフラグメントを精製し、L58B1.0プローブとして用いた。

ナイスタチン生合成遺伝子クラスターを用いた遺伝子分裂実験

DASHII-N1組換えファージから単離されたBamHIのDNAフラグメント（4.2kb）を、まずE. coli中のpGEM3Zf(-)ベクターにクローン化し、プラスミドpGEM4.2-1を得た。クローニングされたフラグメントの両末端のDNA配列を決定し、データベースの探索後、PKSタイプIをコードすることを見出した。pGEM4.2-1にクローン化されたS.nourseiのDNAフラグメントを、制限酵素EcoRIおよびHindIIIを用いてこのプラスミドから切り出して、これをベクターpSOK201（表3および図3を参照）に由来する3.0kbのEcoRI/HindIIIフラグメントと連結させ、E.coliに形質転換させた。pK0(4.2)-1と名づけられたプラスミドDNAを、形質転換体から回収した後、「細菌株、プラスミドおよび増殖条件」の実施例1に記載したような接合によってS.noursei ATCC 11455に移行させた。

【0123】

pK0(4.2)-1は、S.noursei ATCC 11455中では複製できないので、それは、相同組換え(homologous recombination)を介してS.nourseiのゲノムに統合される自殺ベクター(suicide vector)として機能するであろうと仮定した。そのような組換えの結果として、S.noursei ATCC 11455ゲノム中のPKS遺伝子に対して、pK0(4.2)-1にクローン化された4.2kbBamHIフラグメントが内部にあるものと想定し、該PKS遺伝子は、そのコード領域の分裂によって不活性化されているであろう。3つのS.noursei被接合体(transconjugant)ゲノムへのpK0(4.2)-1の統合を、pGEM4.2-1由来の標識化BamHIフラグメント(4.2kb)をプローブとして用いるサザンブロット分析により確認した。そのゲノムに統合されたpK0(4.2)-1を担持するS.nourseiの分裂変異体の1つを、親株のATCC11455（「二次代謝産物の分析」の方法に関する上記を参照）と平行して、ナイスタチン産生について試験した。後者は期待したレベルでナイスタチンを生産することを示したが、pK0(4.2)-1の分裂変異体においてはナイスタチン産生は検出されず、よってナイスタチン生合成に対する同定されたPKSの必要性が確認された。

ナイスタチン生合成遺伝子クラスターのクローニング

ナイスタチン生合成用の遺伝子クラスター全体をクローン化するために、組換えファージDASHII-N1(N1)内にあるS.noursei DNA挿入断片の末端から誘導され、次いで組換えファージと部分的に一致することが見出されたDNAフラグメントを、遺伝子ライブラリーをスクリーニングするためのプローブとして用いた（プローブに関する上記を参照）。このふるいわけが、図4に示すように、S.noursei ATCC11455ゲノム（全体で約98kb）の2つの領域（それぞれSEQ ID No.1および2）から構成される組換えファージN14、N41、N42、N44、N45、N48、N58、N64、N69およびN76の単離という結果になった。組換えファージN64から誘導されたEcoRI/BamHIのDNAフラグメント（4.3kb）を用いた遺伝子分裂実験（上述と同様に実質的に実施された）で、第二領域（SEQ ID No.2）もナイスタチン生合成遺伝子をコードすることを確認した。

ナイスタチン生合成遺伝子群のDNA配列分析

上述の組換えファージから由来する完全なるDNA挿入断片を、XbaIまたはEcoRIのいずれかのフラグメントとして、E.coli宿主内のpGEM3Zf(-)ベクターにサブクローン化し、これらフラグメントの両方のDNA鎖についてヌクレオチド配列を決定した。ナイスタチン生合成遺伝子クラスターの2つの領域（それぞれSEQ ID No.1および2）から構成されるDNA配列についてコンピュータ支援解析をした結果が、図4に示され、表4に列挙した遺伝子の同定となった。

【0124】

【表13】

表 4 S.Noursei のナイスタチン生合成遺伝子フラスター内で
同定された遺伝子

名 称	生産物	推定上の機能	
<i>nysR1</i>	転写アクチベーター	ナイスタチン生産の調節	
<i>nysR2</i>	転写アクチベーター	ナイスタチン生産の調節	
<i>nysR3</i>	転写アクチベーター	ナイスタチン生産の調節	10
<i>nysR4</i>	転写アクチベーター	調 節	
<i>nysR5</i>	転写リプレッサー	調 節	
<i>ORF1</i>	ペプチダーゼ	ペプチド代謝	
<i>ORF2</i>	転写アクチベーター	調 節	
<i>nysA</i>	P K S タイプ I	ナイスタチンポリケチド の主鎖合成 (負荷モジュール)	20
<i>nysB</i>	P K S タイプ I	ナイスタチンポリケチド の主鎖合成 (モジュール 1&2)	
<i>nysC</i>	P K S タイプ I	ナイスタチンポリケチド の主鎖合成 (モジュール 3-8)	
<i>nysI</i> (不完全)	P K S タイプ I	ナイスタチンポリケチド の主鎖合成 (モジュール 9-13)	
<i>nysE</i>	チオエステラーゼ	PKSからの最終生産物の放出	30
<i>nysD1</i>	グリコシルトランスフェラーゼ	ポリケチド主鎖への マイコサミンの付着	
<i>nysD2</i>	アミノトランスフェラーゼ	マイコサミンの生合成	
<i>nysD3</i>	GDP-マンノース-4, 6- デヒドラターゼ	マイコサミンの生合成	
<i>nysH</i>	A B C トランスポーター	ナイスタチンの流出	
<i>nysG</i>	A B C トランスポーター	ナイスタチンの流出	40
<i>nysF</i>	4'-ホスホパンテテイン トランスフェラーゼ	P K S 翻訳後修飾	

【 0 1 2 5 】

P K S タイプ I をコードする 3 つの完全遺伝子 (*nysA*、*nysB*および*nysC*) (SEQ ID NO: 1中) および 1 つの不完全遺伝子 (*nysI*) (SEQ ID NO: 2中) が同定された。当該 4 つの遺伝子でコードされる生成物のアミノ酸(aa)配列が、公知の P K S タイプ I の aa 配列と比較して解析された (分子の特徴は上記表 2 も参照)。4 つのタンパク質は、すべてリファマイシンおよびラパマイシン P K S に高水準の相同性を表すので (Aparicioら Gene, v. 169: 9-16,

1996、TangらGene, v.216: 255-265, 1998)、ナイスタチンPKSに推定される機能の解析は、前者に対する比較に基づいた。1366 aaのNysAタンパク質は、KS^S、AT、DH、KRおよびACPドメインから構成されるPKSの1つのモジュールをコードする。KS^Sドメインにおいて保存システイン残基を欠くことは、このモジュールが縮合反応を行うことができず、このため、ナイスタチン・ポリケチド主鎖骨格生合成開始のためのアセテート スターター単位を提供するローディングモジュール（負荷モジュール）を表していることを示唆している(Bisangら、Nature,v. 401:502-505, 1999)。NysBの3192 aa配列の解析は、KS、AT、KR、ACPおよび明らかに不活性なDHドメインを有する2つのモジュールを含むことを示した。両NysBモジュールに同定されたATドメインは、メチルマロニルCoA-特異的ATドメインについて特徴的な特性を表わしている (HaydockらFEBS Lett., v. 374: 246-248, 1995)。NysBタンパク質のこの特徴は、ナイスタチン分子中に組み込まれた2つの近接するメチルマロニルCoAのみが最初の2つの伸長単位であるために、ナイスタチン・ポリケチド主鎖骨格生合成経路に関係する第1番目および2番目のモジュールを含むことを示唆している。11096 aaのNysCタンパク質は、我々が知る限りこれまでに見出された最も大きい細菌性ポリペプチドであり、ナイスタチン・ポリケチド鎖生合成 (C21-C32の取り込み) における縮合段階3~8を明らかに担う6つのモジュールから構成される。NysCタンパク質のモジュール5はERドメインを含み、これは、C29とC28との間の二重結合の還元を説明できる (図1参照)。モジュール5の他に、NysCにおける他の全てのモジュールは、これらがすべてKS、AT、DH、KRおよびACPドメインを含む点において類似している。モジュール4および5におけるKRドメインは、aa配列レベルで100% 同一であり、これらのドメインをコードするDNA配列レベルで99.9%同一であることがわかった。こうして、モジュール4および5におけるKRドメインは、おそらく、進化過程において比較的最近になって重複化された一例を表す。NysIのC末端側で切断された7066 aaのタンパク質 (C末端でaa配列情報が未だ喪失している) は、少なくとも4つの完全モジュールから構成される。これらモジュールは、全てKS、AT、DH、KRおよびACPドメインを含んでいるが、3つのモジュールにおけるDHドメインは明らかに不活性であり、NysI PKSの知られた部分は伸長ステップ9~12 (C13-C20の組み込み) に関与することを示唆している。この推定は、さらにモジュール11中のATドメイン (配列類似性に基づく) が、メチルマロニルCoAエクステンダー (extender) に特異的であり、一方、NysA、NysCおよびNysIにおける他のすべてのATドメインはマロニルCoAに特異的であるという事実により支持される。こうして、モジュール11中のATドメインは、おそらくナイスタチン・ポリケチド主鎖骨格のC16位メチル基の存在 (後に酸化されてC16に結合したカルボキシルを与える) の原因となる (図1参照)。

【0126】

nysC遺伝子の下流に、チオエステラーゼ (おそらくナイスタチンPKSからの成熟ポリケチド鎖放出に関わる) のためのコード配列nysEが見出された。nysR1、nysR2およびnysR3遺伝子がnysEの下流に見つけられた。これらの遺伝子による生成物は、想定される転写レギュレーターに対して相同である。NysR1(966aa)、NysR2(953aa)およびNysR3(927aa)タンパク質は、すべてそれらのC末端で、LuxRタイプの推定ヘリックス-ターン-ヘリックス (HTH)-DNA結合モチーフを含むことが見出された。それに加えて、NysR1は明瞭なATP/GTP結合モチーフを含み、NysR3はN末端に「ロイシンジッパー」 (推定DNA結合) モチーフを含む。NysR1をコードする遺伝子は、遺伝子の始まりに近接した希TTAコドン (Leu11) を含むことが発見され、S. nourseiではNysR1発現が、bldA様の遺伝子によって、翻訳レベルで調節されているかも知れないことを示唆している (White&Bibb, J. Bacteriol, v. 179: 627-633, 1997)。

【0127】

ナイスタチン生合成におけるnysR1遺伝子の関与を確認するため、遺伝子分裂実験を、nysR1コード配列に対して内部にある1379bpApaI DNAフラグメントを使用して行った。nysR1コード配列に対して内部にあり、しかもSEQ ID NO 1のnt 51531-52910を表す1379bp ApaI DNAフラグメントは、pGEM11Zf(-)ベクターのApaIサイトにクローン化され、組換えプラスミドpNRD1を与えた。pNRD1から分離された1430 bp EcoRI/HindIII DNAフラグメントは、p

SOK201の3.0kb EcoRI/HindIII フラグメントと連結され、その結果としてpNRD2ベクターとなった。後者は、続いてE. coli ET12567(pUZ8002)から接合を介して、S. noursei ATCC11455におけるnysR1遺伝子分裂に使用された。

【0128】

nysR1分裂変異体によるナイスタチン産生を検討したところ、ナイスタチンを生成する能力がないことを示した。しかしながら、nysR1分裂変異体から観察された表現形質は、nysR1とともに転写しうるnysR2およびnysR3遺伝子上の変異の極性効果を反映するものであることがわかる。

nysR3の下流に、遺伝子nysR4、nysR5、ORF1およびORF2が見つかった。210 aaのnysR4遺伝子産物（これは、後にSEQ ID No. 35におけるヌクレオチド120628での上流開始コドンから266 aa発現であることを見出された）は、応答レギュレータータイプである転写アクチベーターとの類似性を示し、中心に位置したATP/GTP結合モチーフおよびC-末端に位置したLuxR型のHTH DNA結合モチーフを含む。253 aaのNysR5タンパク質は、転写リプレッサーとの類似性を示し、しかもN末端で推定DeoR型HTH DNA結合モチーフを含む。nysR4およびnysR5遺伝子の産物は、ナイスタチン生合成遺伝子近傍への局在に基づいてナイスタチン生合成の調節に関係している可能性があるようである。nysR5の下流に位置し、かつ反対方向へ転写されるORF2は、転写アクチベーターと類似性を示し、かつ中心に位置した、AsnC型の推定HTH DNA結合モチーフを有する354 aaペプチドをコードする。この遺伝子がナイスタチン生合成の調節に関係しているかどうか明白ではない。は、ORF2のすぐ上流に位置するORF1が推定ペプチターゼをコードするからである。ORF2はむしろORF1の調節に
10
20
関与することがありそうであるが、その確認のためにはORF2不活性に関する実験が要求されるようである。ペプチターゼをコードする遺伝子（ナイスタチン生合成における役割は何ら与えられていないが）が、配列決定された領域の右側隣接部上に見いだされたという事実は、ナイスタチン生合成遺伝子クラスターの右側境界が同定されたことを示唆する。

【0129】

上記遺伝子を包含する配列決定された領域の左側に、nysAの上流に位置し、かつ、nysAと反対方向に転写される2つの遺伝子が同定された（図4）。506 aaのnysD1遺伝子産物は、ほ乳類からのUDP-グルクロノシルトランスフェラーゼに対してかなりの相同性を示す。この酵素は、UDP-グリコシルトランスフェラーゼ系統群に属し、毒性であり得る生体異物の除去プロセスにおいて、それらのグリコシル化の方式で関与している。nysD1は、ナイスタチン ポリケチド主鎖骨格へのデオキシ糖部分（ミコサミン）の付加に関わっているグリコシルトランスフェラーゼを示すようである。nysD2の産物は、デオキシ糖の生合成プロセスにおいてデオキシ糖へのアミノ基の付加に関与するペロサミンシンターゼとトランスアミナーゼに対して高度の相同性を示す。こうしてnysD2は、おそらくミコサミン生合成に係るアミノトランスフェラーゼを示す。

【0130】

PKSタイプIをコードするnysIに加えて、4つの他の遺伝子が、S. noursei ATCC11455のナイスタチン生合成遺伝子クラスターの第2の配列決定済み領域に（SEQ ID NO.2）同定された（図4）。nysD3遺伝子によってコードされた344 aaタンパク質は、デオキシ糖形成に必要とされるGDP-マンノース-4,6-デヒドラターゼに対して高い相同性がある。NysD3タンパク質が、S. noursei ATCC11455におけるナイスタチン ミコサミン部分の生合成に関与していることはあり得る。nysHおよびnysG遺伝子産物である、584 aaおよび605 aaは、それぞれ、ABSファミリーのトランスポーターに高度の類似性を示す。NysHおよびNysGポリペプチドは両方とも、それらのC末端に明瞭なABCトランスポーターの標識、中心に局在したATP/GTP結合モチーフ、およびN末端に局在した膜貫通領域を含む。これらの2つのタンパク質は、おそらくナイスタチンを産生する生物からATPに依存して活発に放出することに責を負い、こうして、疎水性ナイスタチン分子によって膜が塞がれる危険を排除する。nysF遺伝子の産物である245 aaポリペプチドは、4'-ホスホパンテテイントラン
40
50
フェラーゼへの相同性を示す。後者の酵素は、PKS類のACPドメインにおける翻訳後修飾

の原因となり、かつ、その十分な機能性が要求される (CoxらFEBS Lett., v. 405:267-272, 1997; KealeyらProc. Natl. Acad. Sci. USA, v. 95:505-509, 1998)。このため、NysFタンパク質が、ナイスタチンPKSの翻訳後修飾において機能し、かつ、ナイスタチン合成に要求されることはありそうなことである。

【0131】

【実施例2】

新規ポリエン抗生物質の生成に導くナイスタチンPKS遺伝子の遺伝子的処置

ナイスタチンは、3～8個の共役二重結合をマクロラクトン環内に有することを特徴とするポリエン マロクライド化合物の群に属する。ナイスタチン自体は、テトラエンであり、C20とC27の間に4つの共役二重結合を有する。またナイスタチン分子上には、前記4個の共役二重結合セットからC28-C29により離されている、2個の共役二重結合のセットがC30とC33との間にある (図1参照)。NysC PKSのコンピューター支援解析から、このタンパク質中のモジュール5におけるERドメインは、C28とC29との間における二重結合の還元に関与していることが明らかになった。かくして、この特定ドメインが不活性となることで、C28とC29との間に二重結合を有する化合物を得ることは理論的に可能となり、ナイスタチン分子中に共役二重結合の2セットが加入し、よってヘプタエン マロクライド化合物が創造される。

【0132】

NysCのモジュール5におけるERドメインを不活性にするために、nysC遺伝子のなかで"インフレーム"欠損方法が選択された。S.nourseiのNysCをコードするゲノム領域における遺伝子置換のため、ベクターpERD4.2の構築は、以下のようにされた：

NysCのモジュール5におけるERドメインの不活性

NysCのモジュール5におけるERドメインC末端部分についてのコード配列、およびモジュール5 (nt32174-32559)におけるKRドメインN末端部分についてのコード配列を表わしている394bp DNAフラグメントについてのPCRに基づく増幅が実施された。オリゴヌクレオチドプライマーERD1 (5'-GTTGGTACCCCACTCCCGTCCGCAC-3'、センス) (SEQ ID NO.25)はKpnI制限酵素切断部位を創出するために、5'エンド上の付加ヌクレオチドとともに、nt 32174-32190を含むnysC遺伝子のヌクレオチド配列から選択された。オリゴヌクレオチドプライマー、ERD2 (5'-CCAGCCGCATGCACCACC-3'、アンチセンス (SEQ ID NO.26))は、nysCをコードするDNA配列から選択され、かつSphI制限酵素切断部位を有する、nt32559-32542 (SEQ ID NO.1)間のDNAセグメントを含んでいた。結果として得られるPCRフラグメントは、KpnIおよびSphIで消化され、1828 bp BamHI/KpnI DNA フラグメント (SEQ ID NO.1のnt 29224-31052)および1273 bp SphI/EcoRI DNA フラグメント (SEQ ID NO.1のnt 32548-33821)とともに、EcoRI/BamHIで消化されたpGEM3Zf(-)ベクター中に連結された。その連結混合体はE.coli. DH5αに形質転換され、pERD4.1と命名される6.7kbの組換えプラスミドが形質転換体の1つから回収された。後者は、nysCコード領域のnt 31052とnt 32174との間に内部欠損を有する、nt 29224-33821のSEQ ID NO. 1DNA配列を表すハイブリッドDNAフラグメントを含んでいた。この欠損は、DH-ERドメイン間リンカーの部分および推定NADP(H)結合サイトを含むERドメインのC末端部分のため、コード領域 (nysCのaa 4837から5208)を除去した。

【0133】

S.nourseiにおけるNysC ER4ドメインの不活性化用のベクターを構築するため、組換えDNAフラグメントがpERD4.1からEcoRIおよびHindIII制限酵素を用いて切り出されさらにこのものはpSOK201からの3.0kb EcoRI/HindIII DNAフラグメントに連結され、該連結混合体はE.coli. DH5αに形質転換された。6.5kbのプラスミドpERD4.2が、その形質転換体の1つから分離され、S. noursei ATCC11455における遺伝子置換手順を実施するために使用された (以下参照)。この手順の後に選択された組換えS.noursei菌株はERD44およびERD48と命名された。

【0134】

後者のプラスミドは、S.noursei ATCC11455中に、属間接合によって導入され、1つの

トランス接合体、*S.noursei*(pERD4.2)は、さらなる操作のために選択された。*S.noursei*(pERD4.2)のゲノム中にpERD4.2の正しい様式による組込みが、サザンブロット解析によって確認された後、二次乗換えイベント(second crossover event)用の選別が、Sekurovaら FEMS Microbil Lett, v.177:297-304, 1999 (および下記参照)に記載されたようにして行われた。

遺伝子置換操作 この方法は、前記Sekurovaら、1999に記載されたようにして行われる。

【0135】

上記のように遺伝子置換用に構築されたプラスミドは、*S.noursei*中に*E.coli* ET12567(pUZ8002)からの接合によって導入された。相同的組換えを通して、染色体内に組み込まれたプラスミドを担持するクローンの1つは、抗生物質フリーのISP2寒天培地の上で、3巡の孢子形成を行い、3巡後の子孫は、抗生物質抵抗マーカーの欠損がテストされた。適当なプローブを用いて、いくつかの抗生物質感受性菌株から分離された全DNAのサザンブロット解析が、所望の変異を確認するために使用された。

【0136】

選別マーカーを喪失し、かくして二次乗換えイベントが行われた8つのコロニーのうち、4つはサザンブロット解析により、野生型遺伝形質に戻ることが示された。2つの菌株は、ナイスタチン遺伝子クラスターの実質的部分を明らかに欠いている大きな欠損を含むことが示され、他の2つの変異体は、所望の1116bp欠失(ERD44)、あるいは期待したより多少大きい欠失(ERD48)のいずれかを含んでいた。菌株ERD44およびERD48によるポリエン抗生物質産生の解析は、それらがナイスタチンを産生しないことを示した。その代わりに、ERD44は、期待していたように、ヘプタエンに特徴付けられるUVスペクトルピークを有するポリエン化合物を生じることが示された。驚くべきことに、ERD48変異体は、NysCタンパク質からのモジュール5全部のインフレーム欠損と一致するのであるが、(分光学的解析により)ヘキサエンのポリエンを生じることが示された。ERD48変異体中で起ったる事象をさらに詳細に調べるために、DNAフラグメントが、*S.noursei* ERD48変異体のゲノムからPCR増幅された(これは、かかる欠失の推定産物を包含するであろう。)(以下参照)。

S.noursei ERD48ゲノムからのPKSをコードするDNAフラグメントのPCRアシスト増幅およびクローニング

S.noursei ERD48における変異体nysC遺伝子部分の増幅を目的としたPCR反応は、それぞれSEQ ID NO.1のDNA配列のnt 24744-24760およびnt 33818-33833に相当する、オリゴヌクレオチドプライマーKR48.1(センス、5'-CCG CGT CGG ATC CGC CGA C-3')(SEQ ID NO: 27)およびKR48.2(アンチセンス、5'-AGC CTT CGA ATT CGG CGC C-3')(SEQ ID NO: 28)を用いて行われた。50 μ lのPCR混合物は0.1 μ gの*S.noursei* ERD48全DNA、25pmの各KR48.1およびKR48.2オリゴヌクレオチドプライマー、dNTP(最終濃度350 μ m)、「エキスパンド・ハイフィデリティPCRシステム」(Boehringer Mannheim)からの1 x PCRバッファー、および同システムからの1.5UのDNAポリメラーゼ混合物を含む。PCRは、下記プログラムを用いるパーキンエルマーGeneAmpPCRシステム2400で行った: 96°C(4分)で1サイクル変性、94°C(45秒)および70°C(10分)での35サイクルの変性/アニーリング/合成および、および72°C(10分)での1フルサイクルの最終アニーリング/伸長。この手順で得られた2.7kb DNAフラグメントは、EcoRIおよびBamHI制限酵素を使用して消化し、これを、同じ酵素を用いて消化されたpGEM3Zf(-)ベクターDNAに連結させた。該連結混合体は、形質転換によって*E.coli* DH5 α に導入され、5.9kbのプラスミドpKR48が形質転換体の1つから分離され、DNA配列解析が行われた。

【0137】

pKR48における挿入物のDNA配列は、SEQ ID NO. 29(ここではERD48 seqとして同定される)中に存在している。その翻訳産物は、SEQ ID NO. 30に示され、899aaタンパク質である—SEQ ID NO. 29および30の分子特徴は以下に示される:

【0138】

【表14】

10

20

30

40

50

配列番号 : 29 (DNA: ERD48.seq)

始点	終点	名称	説明
1	254	DH4	DH4 ドメイン コード化領域 C-末端
1170	1913	KR4/5	ハイブリッド ケトリダターゼ ドメイン, モジュール 4/5
2010	2231	ACP5	ACP5 ドメイン コード化領域
2295	2700	KS5	KS5 ドメイン コード化領域

10

配列番号 : 30 (AA: 翻訳産物)

始点	終点	名称	説明
1	84	DH4	DH4 ドメイン, モジュール 4, C- 末端
390	637	KR4/5	ハイブリッド KR ドメイン
670	743	ACP5	ACP ドメイン, モジュール 5
764	899	KS5	KS ドメイン, モジュール 5, N- 末端

20

【0139】

この2700bpフラグメントのクローニングおよびDNA配列決定により、該フラグメントがハイブリッドPKSモジュールの一部をコードすることが確認され、このことは、NysCのモジュール4および5の相同性の高いKRドメインをコードするDNA配列間の組換えと矛盾しない。この組換えによって、モジュール4のKRおよびACPドメインのC末端部、KS、AT、DHおよびERドメイン、およびモジュール5のKRドメインのN末端部をコードするDNA配列が明らかに削除され、結果としてNysC PKSからモジュールが一つ完全に失われる。

30

【0140】

S. noursei ERD44およびERD48から生成された化合物を予備的に分析した。ERD44変異体により生成されたヘプタエン系化合物は、抗真菌活性試験で使された菌であるカンジダアルビカンス (Candida albicans) に対して高い抗真菌活性を有することがわかった。この時点では、上記化合物 (仮にS44とする) は、まだ十分に精製されておらず、またその正確な濃度も決定することが困難であるため、該化合物の抗真菌活性を正確にアッセイすることは不可能である。しかし、ナイスタチン、S44、およびアムホテリシン (ヘプタン系マクロライド) に特有な波長における紫外線吸光度に基づいて概算すると、S44はナイスタチンと比較して4~5倍高い抗真菌活性を有しているようである。ERD48変異体から産生された化合物を高速液体クロマトグラフィーで分析すると、この菌株 (混合体をS48とする) によって、保持時間の異なるヘキサエン系マクロライドが少なくとも5つ生成されることがわかる。これは恐らく、C19でのグリコシル化、C10での水酸化、またはC16でのメチル基の酸化による、マクロラクトン環の修飾状況の違いを示していると思われる。この予想の根拠は、マクロラクトン環のサイズ縮小が、修飾する酵素の新しい基質に対する低い親和性につながることは大いにあり得るためである。

40

【0141】

S48混合体の抗真菌活性を試験した結果、該抗真菌活性はナイスタチンの抗真菌活性

50

の約10%であることがわかった。検出された抗真菌活性が、環修飾酵素によって完全に修飾されたERD48によって生成されるS48混合体中の化合物の一つのみから由来するようである。よって、精製前にこの化合物の相対的な抗真菌活性を評価することは不可能である。いくつかのヘキサエン抗生物質は抗真菌活性とともに抗菌活性を有することが知られている(Ciftciら、J. Antibiot, v. 37: 876-884, 1984)。よって、ERD48変異体により生成されるヘキサエン系化合物を、抗菌剤の製造に使用することが可能である。ERD44変異体およびERD48変異体における新ポリエン化合物の産生につながるNysCタンパク質の変化、ならびに該変異体のマクロラクトン環の推定構造を図5に示す。

NysCのモジュール8におけるDH8ドメインの不活性化

ナイスタチンPKSを遺伝子的に操作する可能性を、さらにNysCのモジュール8におけるDHドメインを不活性化することにより実証した。プラスミドpNPR1.1とNysC遺伝子とを組み換えると、DH8ドメインをコードするDNA領域がフレーム単位で削除される結果となる。上記プラスミドpNPR1.1は下記のように構成した。

【0142】

3989bpのKpnI/BclI DNAフラグメント(領域1 DNA配列(SEQ ID No.1)のnt43004~46993)および2409bpのBamHI/EcoRI DNAフラグメント(同上のnt47680~50089)を、組換えファージN1のDNAから切り出し、EcoRIおよびKpnIで消化したpGEM3Zf(-) DNAベクターと連結した。該連結混合体をE. coli DH5αに形質転換し、形質転換体の一つから組換えプラスミドpGEM-NPR1を分離した。該組換えプラスミドpGEM-NPR1は、領域1 DNA配列(SEQ ID No.1)のnt43004~50089を表すハイブリッドDNAフラグメントを含んでおり、nt46993からnt47680の間は欠失していた。この欠失により、NysCポリペプチドのaa10150から10378をコードするDNA領域が除去され、それによりモジュール8のDH8ドメイン、およびモジュール8のDH8-KRドメイン間リンカーが影響される。S. nourseiにおけるNysCのDH8ドメインを不活性化するためのベクターを構築するため、制限酵素EcoRIおよびHindIIIを用いて、組換え型DNAフラグメントをpGEM-NPR1から切断し、pSOK201からの3.0kbのEcoRI/HindIII DNAフラグメントと連結し(図3および表3参照)、該連結混合体をE. coli DH5αに形質転換した。9.4kbのプラスミドpNPR1.1を形質転換体の一つから分離し、上記Sekurovら、1999に準拠した方法によりS. noursei ATCC11455の遺伝子組換えを実施する際に使用した。

【0143】

上記方法の後に選択された組換えS. noursei菌株は、NPR1.1と名づけた。このS. noursei組換え菌株NPR1.1は、サザンブロット法で分析した結果、NysCのDH8をコードする配列において好ましい欠失の形態を有することがわかった。前記S. noursei組換え菌株NPR1.1の培養物から抽出した抽出物中にある二次代謝生成物を薄層クロマトグラフィー(TLC)で分析したところ、予想されていたマクロライド化合物の存在が明らかになった。これらの化合物の相対移動度はナイスタチンとは異なり、上記抽出物からは、ナイスタチンに特有な紫外線スペクトルは検出されなかった。分析により、NPR1組換え体によって新しく産生された分子内において、ナイスタチン・アグリコン上の4つ一組の二重結合が乱されており、さらにマクロラクトン環がC23に結合した水酸基を含有することが示された(表5)。NPR1培養抽出物を用いてなされたカンジダ・アルビカンスに対するバイオアッセイが極めて低い抗真菌活性を示したため、上記化合物をNPR1から精製する試みはなされなかった。しかし、このNPR1変異体は、ナイスタチンPKSを用いた更なる操作に有用に使用され得る可能性がある。

モジュール7(NysC)におけるKRドメインの不活性化

4404bpのDNAフラグメントを、制限酵素EcoRIおよびSmaIを用いて、

組換えファージN1のDNAから切り出した。EcoRI部位は、ファージN1のポリリンカーにおいて、領域1 DNA配列 (SEQ ID No. 1) のnt 38398において開始するS. noursei DNA挿入フラグメントの左側に位置し、一方SmaI部位は領域1 DNA配列のnt 42802に対応する。3303bpのDNAフラグメントを、制限酵素SmaIおよびBamHIを用いて、同じ組換えファージのDNAから切り出した。SmaIおよびBamHI制限酵素部位は、それぞれ領域1 DNA配列 (SEQ ID No. 1) のnt 43099およびnt 46402に位置していた。両DNAフラグメントを、制限酵素EcoRIおよびBamHIで消化したpGEM3Zf (−) DNAと連結し、該連結混合体を大腸菌DH5αに形質転換した。10.7kbのプラスミドpGEM-NPR2を形質転換体の一つから回収した。該プラスミドpGEM-NPR2は、領域1 DNA配列 (SEQ ID No. 1) のnt 38398~46402を表す組換えDNAフラグメントを含有しており、またnt 42802からnt 43099の間は欠失していた。この欠失により、モジュール7のKRドメインにおいて推定NADP (H) 結合部位を包含するNysCタンパク質の、aa 8753から8851をコードするDNA領域が除去される。S. nourseiにおけるNysCのKR7ドメインを不活性化するためのベクターを構築するため、組換えDNAフラグメントを、制限酵素EcoRIおよびHindIIIを用いてpGEM-NPR2から切り出し、pSOK201からの3.0kbのEcoRI/HindIII DNAフラグメントと連結し、該連結混合体をE. coli DH5αに形質転換した。10.7kbのプラスミドpNPR2を形質転換体の一つから分離し、上記Sekurovaら、1999に準拠した方法によりS. noursei ATCC 11455の遺伝子組み換えを実施する際に使用した。上記方法の後に選択された組換えS. noursei菌株は、NPR2. 1と名づけた。

【0144】

前記S. noursei組換え菌株NPR2. 1の培養物から抽出した抽出物中にある二次代謝生成物をTLCで分析したところ、予想されていたマクロライド化合物の存在が明らかになった。これらの化合物の相対移動度はナイスタチンとは異なっており、さらにNPR1変異体から産生された代謝生成物とも異なっていた。上記抽出物からは、ナイスタチンに特有な紫外線スペクトルは検出されなかった。NPR2. 1組換え体によって新しく生成された分子内において、ナイスタチン・アグリコン上の4つ一組の二重結合が乱されており、さらにマクロラクトン環がC25においてケト基を含有することが示された (表5)。上記化合物をNPR2. 1から精製する試みはなされず、また、NPR2. 1培養抽出物を使用しての抗真菌活性を検査するバイオアッセイは行われなかった。上記変異体は、生成する代謝物の利点を有するだけでなく、ナイスタチンPKSを用いた更なる操作に使用することができる。

変異NysCタンパク質NysC__DH8のモジュール5におけるERドメインの不活性化

第二の突然変異体を、モジュール8において不活性DHドメイン (NysC__DH7) を有するNysCタンパク質に導入するため、プラスミドpERD4. 2をS. noursei変異体NPR1. 1に導入し、上記文献Sekurovaら、1999の記載に準拠して遺伝子組換え作業を行った。これにより、NysCの配列をコードするER5およびDH8中に変異を有する、組換えS. noursei菌株ERDH9が生成された。この二つの変異の組合せにより、C23に水酸基を有するペンタエン系ナイスチン誘導体が生合成されると推測される (表5)。ERDH9培養抽出物を予備分析したところ、数量的にはその紫外線スペクトルを正確に同定することは困難であるが、上記菌株によってポリエン化合物が産生されていることが確認された。この予備分析の結果により、さらに上記化合物が好ましくは培養上清に蓄積され、野生型S. nourseiにより産生されるナイスタチンは大部分が菌糸体と結合していることがわかった。このことは、ナイスタチン分子上に水酸基を付加することが、そうした化合物の水溶性増加の要因であるという仮説と合致する。上記新規化合物の精製はなされず、また、バイオアッセイも行われなかった。

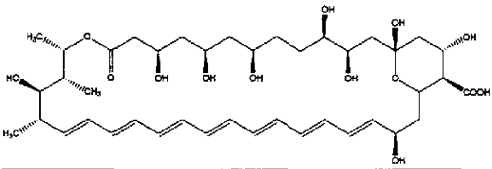
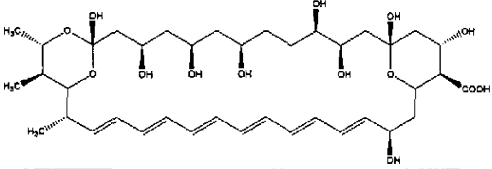
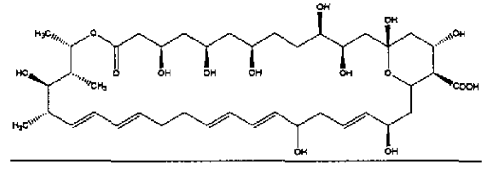
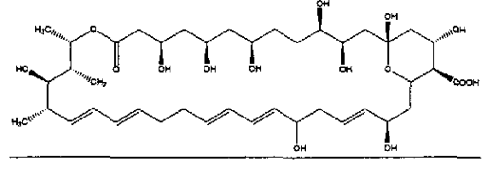
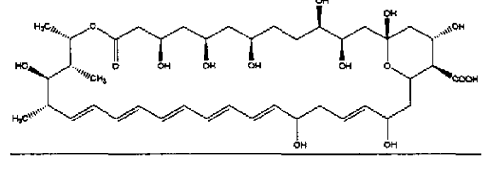
【0145】

本発明者は、現在のところ突然変異体NPR1.1、NPR2.1およびERDH9の生成物の構造を確認できていない。

【0146】

【表15】

表5：S. nourseiのNysC遺伝子を遺伝子操作することにより調製されるナイスタチン誘導体の予想構造（詳細は実施例2を参照）

突然変異体	予想構造 (ポリケチド部位のみ)	紫外線スペクトル nm (DMSO)	抗真菌性
ATCC11455 ERD44	ナイスタチン (図1) 	299, 312, 327 375, 395, 419, 444	普通 高い
ERD48		336, 352, 370, 391	低い?
NPR1		-	低い
NPR2.1		-	?
ERDH9		?	?

【0147】

【実施例3】

10

20

30

40

50

ナイスタチンの大量生成につながる調節遺伝子の操作

S. noursei ATCC11455における、Perme*プロモーター制御下のNysR1の発現

NysR1遺伝子がナイスタチン生合成遺伝子の転写アクチベーター（活性化因子）をコードすることをさらに確認するため、S. noursei ATCC11455においてPerme*プロモーター制御の下、NysR1遺伝子を発現させた（Bibbら、Mol. Microbiol.、v. 14: 433-545、1994）。まず、S. noursei ATCC11455ゲノム内に安定的かつ効率的に組込むためのプラスミドpSOK804を調製した（図2）。上記調製は、pSET152（上記文献Biermanら、1992）からの3.0kbのSphI/HindIII DNAフラグメントと、部位特異的組込みに必要な機能を有するバクテリオファージVWB（上記文献Van Mellaertら、1998）からの2.3kbのSphI/HindIIIフラグメントとを連結することによりなされた。E. coli ET12567（pUZ8002）からのpSOK804をS. noursei ATCC11455と接合したところ、このプラスミドは、特に $3 \sim 10^{-6}$ の頻度において、S. nourseiゲノムの1サイト中に特異的に組込まれることが示された。

【0148】

pSOK804において、Perme*プロモーター制御の下にNysR1遺伝子をクローンするため、以下の手順を採用した。NysR1のN末端部を、組換えファージN1 DNAテンプレートから、オリゴヌクレオチドプライマーNR1.1（センス）5'-C G C C G C A T G C T G T T C T C A C C C C A C G T - 3'（SEQ ID No. 31）およびNR1.2（アンチセンス）5'-G G C G C G A C C C G G T T C G G C C T - 3'（SEQ ID No. 32）を用いて、PCR増幅した。オリゴヌクレオチドプライマーNR1.1の配列は、配列標識番号1のnt 51376～51391に対応しており、5'位エンド部においてヌクレオチドC G C C G C A T G Cが付加されているため、SphIエンドヌクレアーゼ用部位を形成している。オリゴヌクレオチドプライマーNR1.2は、配列標識番号1のnt 51964～51982に対する相補的な配列と対応しており、AgeIエンドヌクレアーゼ用制限部位を包含していた。0.6kbフラグメントを、実施例1に記載の条件下で、ファージN1 DNAをテンプレートとして使用してPCR増幅し、SphIおよびAgeIで消化した。消化した0.6kbのフラグメントを、ファージN1挿入フラグメント（ファージのポリリンカーに起源を持つEcoRI部位）からの2.8kbのAgeI/EcoRI DNAフラグメントと共に、SphIおよびEcoRIで消化したpGEM7Zf（-）ベクターに連結した。上記連結において使用した2.8kbフラグメントは、配列標識番号1のnt 51971～54814に対応しており、NysR1のC末端部およびNysR2のN末端部（NysR2の最初の162aaをコードする）を包含していた。この連結によって得られた組換えプラスミドをpNRE1と名づけた。pNRE1からの3.5kbのSphI/HindIII DNAフラグメントを、Perme*プロモーターを含有するpGEM7ermELi（表3参照）からの0.3kbのEcoRI/SphIフラグメントと共に、EcoRI/HindIIIで消化したpSOK804ベクターに連結した。得られたプラスミドpNRE2を、接合（実施例1参照）によりS. noursei ATCC11455中に導入し、組換え菌株S. noursei（pNRE2）を生成した。低減グルコース培養液の入った振盪フラスコ内での、上記菌株によるナイスタチン生成を分析したところ、上記菌株によるナイスタチン生成は野生型菌株と比較して50%多いことがわかった。これは、Perme*プロモーターからのNysR1遺伝子の過剰発現が原因であると思われる。以上のことから、NysR1遺伝子は、S. noursei菌株でのナイスタチンおよびその誘導体の産生を増強させるために使用される正のアクチベーター（活性化因子）をコードすると思われる。

S. noursei ATCC11455におけるNysR5遺伝子の部分的除去

ナイスタチン生合成の調節において翻訳領域のDNA配列の分析（実施例1参照）を通して推測された、NysR5遺伝子の機能を確認するため、S. noursei ATC

C11455ゲノムに特定の突然変異を導入した。ここで報告されるヌクレオチド配列の nt62037～63360を包含する *S. noursei* ATCC11455ゲノムからのDNAフラグメントを、プライマーNR5D1 (5'-GCGAGCGGCGCTTCACCCCGCAACTCA-3') (配列認識番号33) およびNR5D2 (5'-CGCGAAGCTTGCGCGACTGCTCGACGTC-3') (配列認識番号34) を用いてPCR増幅した。PCR増幅の条件は上記の通りである。1341bpのPCR生成物を、NotIおよびHindIIIで消化し、1688bpのEcoRI/NotI DNAフラグメント (nt60301～61989) およびpSOK201からの3.0kbのEcoRI/HindIIIフラグメントと連結した。得られたプラスミドpNR5Dは、上記*S. noursei* DNAフラグメントを含有しており、NysR5遺伝子の翻訳領域における43bpが欠失していた。この欠失により、NysR5の翻訳領域においてフレームシフト変異が発生し、生成物の切断(truncation)につながる。そのような切断の結果、NysCR5の165個C末端アミノ酸は除去され、別の読み枠でコードされる14個のアミノ酸で置換される (よって、NysR5とは異質となる)。pNR5Dを用いて、*S. noursei* ATCC11455における遺伝子組換え作業を、上記文献Sekurovára、1999の記載に準拠して行った。

【0149】

遺伝子組み換えによって導入された変異により、野生型*S. noursei*と比較して、得られる組換え菌株NR5Dによるナイスタチン生成が5～15%増加した。ナイスタチンの生合成に対し、NysR5のC末端側部を除去することの、小さいながらも再現性のあるプラス効果は、コンピュータによる分析 (実施例1) に基づき推定される、このタンパク質が担うリプレッサーとしての機能と充分に関連している。NysR5遺伝子に対して除去を行っても、このタンパク質においてN末端の位置に確認される推定上のヘリックス・ターン・ヘリックスモチーフは除去されないことから、切断されたNysR5ポリペプチドの残っているリプレッサー作用が、ナイスタチン産生に対し、上記変異の効果が比較的小さいことの説明になり得る。だが、この結果は、ナイスタチンおよびその誘導体の産生を向上させるためのさらなる手段として、リプレッサーをコードする遺伝子に変異を導入することの利点を確認するものである。

【0150】

【実施例4】

ナイスタチン生合成遺伝子クラスター配列決定の完成

SEQ ID No. 1とSEQ ID No. 2との間隙に及ぶDNA配列は、この領域をカバーする組換えバクテリオファージN20, N32, N95, N98およびN99の重複した挿入体の両DNA鎖に対して決定された。

【0151】

配列決定に用いた手順、プローブ生成およびスクリーニングは実施例1に記載したとおりである。ライブラリー・スクリーニングに用いたプローブおよび発見された関係する組換えバクテリオファージは下記のとおりである：

PKS 72	N20, N32
L44ES3.5	N90
L76SN0.5	N95
L20S0.64	N98, N99

プローブの記載

6. L44ES3.5プローブ. 制限酵素EcoRIおよびSalIを用いて、バクテリオファージN44から分離された3.5kbのDNAフラグメントが、L44ES3.5プローブとして使用された。

【0152】

7. L76SN0.5プローブ. 制限酵素SacIおよびNotIを用いて、組換えバクテリオファージN76から分離された0.5kbのDNAフラグメントが、L76SN0.5プローブとして使用された。

8. L20S0.64プローブ. 制限酵素SalIを用いて、プラスミドpL20EB3.7から分離された0.64kbのDNAフラグメントが、L20S0.64プローブとして使用された。

【0153】

新しい配列情報が、完全なnysIおよびnysDII遺伝子のみならず新しい遺伝子であるnysJ、nysK、nysL、nysMおよびnysNの同定と相成った(図7参照)。新しい情報に基づけば、9477aaのNysIタンパク質(SEQ ID No. 37)は、6つのモジュールからなるPKSを表し、ナイスタチンポリケチド主鎖骨格生合成の伸長ステップ9から14に対し役割を果たす。これら全てのモジュールは、KS、AT、DH、KRおよびACPドメインを含む。モジュール11のmATドメインの存在は、この伸長ステップにおけるメチルマロニル-CoAエクステンダーの取り込みと一致する。モジュール10、11、12および14のDHDドメインは、活性部位モチーフH(X₃)G(X₄)Pが存在しないために不活性であるようだ。NysIモジュール13のKRドメインは、保存モチーフaSRrGを欠き、よって不活性である。後者の特徴は、モジュール11の不活性DHドメインとともに、ナイスタチン分子の6員ケタリック環(ketalic ring: C13とC17との間)の存在を最もよく説明する。

10

【0154】

nysJ遺伝子は、PKSをコードし、nysIの下流に位置して同じ方向に転写される。NysJ(SEQ ID No. 38)におけるモジュールの組織化から判断して、後者はナイスタチンのマクロラクトン環組み立てにおいて伸長ステップ15から17にとり必要である。NysJ内のモジュール16および17におけるDHDドメインは不活性らしく、またモジュール15に局在化したERドメインは、C8およびC9間の二重結合の還元におそらく最も関与している。

20

【0155】

ナイスタチンPKS系の、最後の18番目のモジュールは、nysJの下流に見出されるnysKによってコードされたNysKタンパク質(SEQ ID No. 39)によって表される。2066aaのNysKタンパク質は、KS、AT、不活性DM、ACPおよびTEDドメインから構成される。NysKタンパク質は、KRドメインを欠き、明らかに不活性DHドメインを含む。TEDドメインは、NysKのC末端に同定され、このタンパク質は最後のエクステンダー単位の縮合に加えて、PKS系からの成熟ナイスタチンポリケチド鎖の放出にも関与することを示唆する。

30

【0156】

クラスターの遺伝子組織化は、図9に示されており、図8は、ナイスタチンの生合成経路においてコードされた種々のタンパク質の提案された関与を表している。

ナイスタチンの生合成において、nysIおよびnysJの関与を確認するため、これらの遺伝子は、接合自殺ベクター(conjugative suicide vector) pKNI1およびpKNJ1を用いて、相同組換えを介して、S. noursei中で分裂させた。分裂実験のための上記ベクターの構築は下記の如くであった。nysIの内部にある3.8kb EcoRI/BamHI-フラグメントは、SEQ ID No. 35のヌクレオチド23711-27541に対応しており、N64 DNAから切り取られ、プラスミド pGEM3Zf(-)の相当部位に連結され、プラスミド pL64EB3.8となった。pL64EB3.8内へクローン化されたS. nourseiフラグメントは、このプラスミドから制限酵素EcoRIおよびHindIIIを用いて切り出され、ベクターpSOK201からの3.0kb EcoRI/HindIII-フラグメントに連結され、プラスミド pKNI1となった。

40

【0157】

nysJの内部にある3.7kbのEcoRI/BamHI-フラグメントは、SEQ ID No. 35のヌクレオチド43287-46992に対応し、バクテリオファージN20 DNAから切り出され、プラスミド pGEM3Zf(-)の相当部位に連結されて、プラスミド pL20EB3.7となった。pL20EB3.7内へクローン化

50

された *S. noursei* フラグメントは、このプラスミドから制限酵素 *EcoRI* および *HindIII* を用いて切り取られ、ベクター *pSOK201* からの 3.0 kb *EcoRI*/*HindIII* フラグメントと連結され、プラスミド *pKNJ1* となった。*pKNJ1* および *pKNJ1* 構築物の両者とも、*E. coli* ET 12567 (*pUZ8002*) に形質転換され、接合により、Zotchev ら、(2000) *Microbiology* 146: 611–619 に記載されるように、さらに *S. noursei* ATCC 11455 に転移された。ナイスタチンの生成は、*pKNJ1* および *pKNJ1* 分裂変異体のいずれにおいても検出されず、したがって、ナイスタチン生合成について同定された PKS の必要性が確認された。

【0158】

ナイスタチン分子の修飾に恐らく関与するであろうタンパク質をコードする 3 つの遺伝子が、*nysK* と *nysDII* との間に同定された。*nysL* および *nysN* の両者の遺伝子ともに、それぞれ 394 aa および 398 aa の P450 モノオキシゲナーゼ（それぞれ SEQ ID No. 40 および 42）をコードし、これらのタンパク質は恐らくナイスタチン ポリケチド部分の C10 位の水酸化および C16 位メチル基の酸化に役割を果たす。どのタンパク質がいずれの反応に関与するかは未だ明確ではなく、ナイスタチン生合成経路における *nysL* および *nysN* の正確な配置には、追加実験が必要とされる。*nysM* 遺伝子は、明らかに 64 aa のフェレドキシン (SEQ ID No. 41) をコードし、それは多分、一方または両方の P450 モノオキシゲナーゼ系の一部を構成し、電子供与体として働く [O'Keefe, D. P. & Harder, P. A. (1991). Occurrence and biological function of cytochrome P450 monooxygenases in the actinomycetes. *Microbiol.* 5, 2099–2105]。

【0159】

図 7 (SEQ ID No. 35) に表示された領域をおよそ 10 kb 左側（組換えバクテリオファージ N90）および 5 kb 右側（組換えバクテリオファージ N69）に伸長している DNA 配列が決定された。ナイスタチンの生合成において、それらしく思われる作用を有する遺伝子は見出せなかったが、これはナイスタチン遺伝子クラスター全体が同定されたことを示唆する。

【0160】

このように、ナイスタチン生合成遺伝子クラスターの完全な配列情報に基づき、以下の遺伝子が同定され、それらの役割を以下に記載する（図 8 および 9 も参照）：

【0161】

【表 16】

10

20

30

表 6. *S. noursei* ATCC 11455 のナスタチン
生合成遺伝子クラスター内にて同定された遺伝子

名 称	生産物	推定上の機能	
<i>nysP</i> 推定	4'-ホスホパンテテイン トランスフェラーゼ	PKS 翻訳後修飾	
<i>nysG</i>	ABC トランスポーター	ナスタチンの流出	10
<i>nysH</i>	ABC トランスポーター	ナスタチンの流出	
<i>nysD3</i>	GDP-マンノース-4,6- デヒドラターゼ	マイコサミンの生合成	
<i>nysI</i>	PKS タイプ I	ナスタチン PKS (モジュール 9-14)	
<i>nysJ</i>	PKS タイプ I	ナスタチン PKS (モジュール 15-17)	20
<i>nysK</i>	PKS タイプ I	ナスタチン PKS (モジュール 18 + TE)	
<i>nysL</i>	P450 モノオキシゲナーゼ	C-10でのヒドロキシル化	
<i>nysM</i>	フェレドキシン	P450系での電子伝達	
<i>nysN</i>	P450 モノオキシゲナーゼ	C-16位でのメチル基の酸化	
<i>nysD2</i>	アミノトランスフェラーゼ	マイコサミンの生合成	
<i>nysD1</i>	グリコシルトランスフェラーゼ	マイコサミンの付着	
<i>nysA</i>	PKS タイプ I	ナスタチン PKS (負荷モジュール)	30
<i>nysB</i>	PKS タイプ I	ナスタチン PKS (モジュール 1および 2)	
<i>nysC</i>	PKS タイプ I	ナスタチン PKS (モジュール 3-8)	
<i>nysE</i>	チオエステラーゼ	PKSからのポリケチド鎖の 放出	
<i>nysR1</i>	転写アクチベーター	ナスタチン生産の調節	40
<i>nysR2</i>	転写アクチベーター	ナスタチン生産の調節	
<i>nysR3</i>	転写アクチベーター	ナスタチン生産の調節	
<i>nysR4</i>	転写アクチベーター	調 節	
<i>nysR5</i>	転写リプレッサー	調 節	
ORF2	転写アクチベーター	調 節	
ORF1	ペプチダーゼ	ペプチド代謝	

【実施例 5】

ナイスタチン生合成遺伝子の操作

以下に、ナイスタチン分子における特異的な化学変化を目標とする、ナイスタチン生合成遺伝子の合理的な遺伝子操作に対する指針を予測的な実施例として表わす。これらの操作は、ポリエン系抗生物質の構造－機能の関係、共役二重結合の数および二つのイオン化可能な基の存在（デオキシ糖部ミコサミンに属するエキソサイクリックなカルボキシル基およびアミノ基）についての現時点での理解に基づく。

共役二重結合の数および位置の変化

ナイスタチンのマクロラクトン環内の共役二重結合は、ケトレダクターゼ（KR）およびデヒドラターゼ（DH）両活性を有するPKSモジュールによって行われる二つの還元ステップの結果として形成される。これらの二重結合のさらなる還元は、このようなPKSモジュールにエノイルレダクターゼ（ER）活性を導入することによってもたらしうことができる。これによりナイスタチン生合成の特定の段階において、二重結合に代えて完全な飽和結合を形成することとなる。以下の操作を提示することができる。（これらの操作の結果として、理論的に生成され得る化合物を図10に示す。）

－モジュール3へのERドメインの組込み（1）；

－モジュール4へのERドメインの組込み（2）；

－モジュール5でのERドメインの同時的失活およびモジュール3へのERドメインの組込み（3）；

－モジュール5でのERドメインの同時的失活およびモジュール4へのERドメインの組込み（4）；

－モジュール5でのERドメインの同時的失活およびモジュール7へのERドメインの組込み（5）；

－モジュール5でのERドメインの同時的失活およびモジュール8へのERドメインの組込み（6）；

－モジュール5でのERドメインの同時的失活およびモジュール9へのERドメインの組込み（7）；

－モジュール5でのERドメインの同時的失活およびモジュール8および9へのERドメインの組込み（8）；

－モジュール5でのERドメインの同時的失活およびモジュール7および8へのERドメインの組込み（9）；

これらの操作は、S. noursei（Sekurovaら、FEMS Microbiology Letters, 177: 297-304、1999）に対して記載された遺伝子置換の技術を用いることにより行うことができる。操作材料は、ナイスタチン遺伝子クラスター自身、あるいは他のPKS遺伝子クラスターから供給できる。後者が組換え菌株の遺伝的安定性の観点より優先されう。

COOH基の脱離／再配置

ポリエン系抗生物質のエキソサイクリック・カルボキシル基は、これらの化合物の選択的毒性に決定的な役割を果たすものと信じられている。より具体的には、イオン化可能なエキソサイクリック・カルボキシルとミコサミン部分のアミノ基との間の分子内および分子間の相互作用が特に重要と考えられる（Resat, H., Sungur, F. A., Baginski, M., Borowski, E., Aviyente, V. 2000. Conformational properties of amphotericin B amide derivatives—impact on selective toxicity. Journal of Computer-Aided Molecular Design, v. 14 p689-703）。ナイスタチン生合成遺伝子の操作を介してエキソサイクリック・カルボキシルを脱離させ、あるいはナイスタチン分子上に再配置することは、いずれも理論上可能である。このような操作を経て生成されたナイスタチン誘導体は、それ自身にとって、あるいはさらなる化学修飾のための基質として作用し有益である。

【0163】

以下の操作を提示する（結果としての化合物は、図11に示す）。

－ナイスタチンPKSのモジュール11のメチルマロニル特異的アセチルトランスフェラーゼ（AT）ドメインとマロニル特異的ATドメインとの置換（10）；

－モジュール11のメチルマロニル特異的ATドメインとマロニル特異的ATドメインとの同時的置換を伴う、モジュール12のマロニル特異的ATドメインとメチルマロニル特異的ATドメインとの置換（11）；

－モジュール11のメチルマロニル特異的ATドメインとマロニル特異的ATドメインとの同時的置換を伴う、モジュール10のマロニル特異的ATドメインとメチルマロニル特異的ATドメインとの置換（12）；

－P450 モノオキシゲナーゼをコードする遺伝子のnysLまたはnysNの失活（ナイスタチン分子のC16位メチル基の酸素化に役割を果たすことが見出されるいずれか）（13）。

【0164】

エキソサイクリック・カルボキシルの発現に関与するP450 モノオキシゲナーゼの特異性は、新規な基質に対しその作用を充足するように操作され得ることに注意すべきである。部位特異的、またはランダムな変異生成のような方法は、誤りがちなPCRおよびDNAシャッフリングとともにこの目的にとって有益であることが判明するであろう。

1. 3 付加的な水酸基の導入（水溶性の増加）

ポリエン系抗生物質は、主として高度に疎水性である共役二重結合の組により、水に極めて難溶である。ある場合において薬剤の薬理学的な性質を拡張することから、水溶性を増加させることは有利となり得る「Golenser, J., Frankenburg, S., Ehrenfreund, T. & Domb, A. J. 1999. Efficacious treatment of experimental leishmaniasis with amphotericin B-arabinogalactan water-soluble derivatives. Antimicrob Agents Chemother., v. 43: 2209-2214）。ナイスタチンの水溶性を増加させるために、ナイスタチン分子に付加的な水酸基（親水性）を導入することを提案する。ナイスタチン生合成遺伝子の以下の修飾は、所望する効果をもたらすことができる（結果としての化合物を図12に示す）。

【0165】

－ナイスタチンPKSモジュール3のデヒドラターゼ（DH）ドメインの不活性化（14）；

－モジュール4のDHドメインの不活性化（15）；

－モジュール5のERドメインの同時的不活性化に伴うモジュール3のDHドメインの不活性化（16）；

－モジュール5のERドメインの同時的不活性化に伴うモジュール4のDHドメインの不活性化（17）；

－モジュール5のERドメインの同時的不活性化に伴うモジュール7のDHドメインの不活性化（18）；

－モジュール5のERドメインの同時的不活性化に伴うモジュール8のDHドメインの不活性化（19）；

－モジュール5のERドメインの同時的不活性化に伴うモジュール9のDHドメインの不活性化（20）；

マクロラクトン環のエクステンション（伸長）、トランケーション（頭部切断）もしくは再構築

ポリエン系抗生物質ナイスタチンの新規誘導体が、やはりナイスタチンPKSのトランケーション（truncation）を通じて得ることができ、より小さいマクロラクトン環を有する誘導体へと導かれる（ERD48変異体について実施例2で例示したとおり）。これは、ナイスタチンPKSからの一つまたはそれ以上のモジュールの欠失を通じて達成される。こ

10

20

30

40

50

のようなトランケーションは、潜在的に新規薬物のさらなる修飾と合成に有用かもしれない、36員-ないし6-員ラクトン環を有するポリケチドの生成へと導くことができる。ナイスタチン・マクロラクトン環の伸長は、付加的なモジュールをナイスタチンPKSへ挿入することを通じて達成される。このような操作は、また薬剤用途のためのリード化合物の生成へと導くこともできる。

【0166】

このナイスタチン分子は、ナイスタチンまたは他のPKSの間で、PKSモジュールをシャッフルする方法により完全に再構築されることができ、よって完全に新しい誘導体が生成する。この点については、W0 00/77181の特許に開示されている方法が、この目的を果たす組換えDNA構築物を作製するのに有用であることを示せる。

10

【0167】

最終的に、ナイスタチン生合成遺伝子は、他のマクロライド抗生物質生合成経路の操作に有用であることを証明できる。PKSおよび修飾酵素の両方が、そのような目的に有用であることを証明できる。ナイスタチン生合成遺伝子は、ピマリシン (Aparicio, J.F., Fouces, R., Mendes, M.V., Olivera, N., Martin, J.F. 2000. 5つのポリケチドシンターゼ遺伝子でコードされた複合多酵素システムが、ストレプトマイセス・ナタレンシス (Streptomyces natalensis) において、26員ポリエン系マクロライド・ピマリシンの生合成に参与する。Chem Biol, v. 7: 895-905) 用のものののように、他のポリエン系抗生物質生合成クラスターの操作にも最も有用になることは確実のようである。ナイスタチンとピマリシン生合成酵素との間の、タンパク質レベルにおける高度の類似性は、それらのハイブリッドが機能することを、最も確実にするであろう。他方では、異種の修飾酵素の異なる特異性は、遺伝子操作がなされた菌株によって製造された分子についてさらなる構造上の変更のための新しいツールを提供するかもしれない。

20

【0168】

【実施例6】

ナイスタチンの産生増加を誘導する調節遺伝子の更なる操作

S. noursei ATCC 11455におけるPermE*プロモーターの制御下でのnysRIIの発現

nysRII遺伝子は、そのコード配列 (実施例1参照) の分析を通じて予測される機能を確認するために、S. noursei ATCC 11455中で、PermE*プロモーターの制御下で発現させた。ファージN58からの2168bpであるSalI-BclIフラグメント (nysRIIのC末端部を示す) は、をSalI-BamHIで消化したpGEM11Zf (-) 中にクローン化され、結果として構築物pC1A1となる。nysRII遺伝子の811bpのN末端部は、ファージN58テンプレートから、オリゴヌクレオチドプライマーNSR2. 1 (5'-GCCGGCATGCGACGAA CAG GACGA GAGGT-3') (SEQ ID NO. 44.) およびNSR2. 3 (5'-GCCGTGGTCGACGAA GGC-3') (SEQ ID NO. 45) を用いてPCR-増幅させた。PCRのための条件は上述の通りである。そのPCRフラグメントは、SphIおよびSalIを用いて消化され、pC1A1からの2168bpのSalI-HindIIIフラグメントと一緒に、をSphI-HindIIIで消化したpGEM3Zf (-) ベクター中に連結させ、その結果、プラスミドpC2A1を生じさせた。後者からは、3.0kbのSphI-HindIIIフラグメントが分離され、このものはPermE*プロモーターを含む0.3kbのEcoRI-SphIフラグメントとともに、をEcoRI-HindIIIで消化したpSOK804ベクター (プラスミドpC3A1を生成する)、あるいはpSOK201由来の3.0kbのEcoRI-HindIIIフラグメント (プラスミドpC3E1を生成する) のどちらかと連結させた。pSOK804をベースとするベクターは、S. noursei 染色体中に部位特異的に組み込まれるので、pC3A1プラスミドは、“トランス”でのnysRII発現の構築物と見なし得た。他方、プラスミドpC3E1は、クローン化されたnysRII遺伝子を通じて相同的組換えのみを介してS. noursei ゲノム中に組み込まれ得る自殺ベクターであり、よって後者

30

40

50

の“シス”の発現をもたらす。プラスミド pC3A1 および pC3E1 の導入は、結果としてそれぞれ組換え菌株 C3A1 および C3E1 となる。前者の菌株における *PermE**プロモーターは、*ny s R I I* および *ny s R I I I* 遺伝子の両方の上流に位置し（シス）、一方、後者の菌株中の *PermE**プロモーターは、*ny s R I I* 遺伝子の上流に位置する（トランス）。該 C3A1 および C3E1 菌株によるナスタチン産生は、野生型の *S. noursei* と比較して、それぞれ 18% および 22% だけ増加した。さらに、培養実験の間、該 C3E1 菌株からのナスタチン産生が野生型菌株よりも 24 時間早く最大値に到達したことがわかった。これらのデータは、*ny s R I I* 遺伝子が、*S. noursei* 菌株においてナスタチンおよびその誘導体の産生を増大させるために使用できるプラスのアクチベーターをコードするという仮説を裏付ける。

10

S. noursei ATCC 14455 中における *PermE**プロモーターの制御下での *ny s R I V* の発現

ny s R I V の開始コドンは再特定され、もともと提示された開始ヌクレオチドの 48 nt 上流に位置する可能性がある。このため前に示唆したように、*ny s R I V* は、おそらく 210 aa（短い）よりもむしろ 226 aa（長い）のタンパク質をコードすると推定される。この *ny s R I V* 遺伝子の長い種類および短い種類は、それぞれ、オリゴヌクレオチドプライマー NR4P3 (5'-CTCAGCATGCGGAAAGGATGGCGG-3') (SEQ ID NO. 46) および NR4P5 (5'-AGGCAAGCTTTCGGCGACACGGGCGT-3') (SEQ ID NO. 47) を用いてあるいは、NR4P4 (5'-CTCAGCATGCGGTACGACCGGCGGG-3') (SEQ ID NO. 48) および NR4P5 を用いて、N58 組換えファージ DNA から PCR 増幅された。PCR の条件は、上述の通りである。0.78 kb および 0.73 kb の対応する PCR 生成物は、*SphI* および *HindIII* を用いて消化され、*PermE**プロモーターを含む 0.3 kb の *EcoRI*-*SphI* フラグメントとともに、*EcoRI*-*HindIII* で消化した pSOK804 ベクターに連結させた。結果としてそれぞれプラスミド pNR4EL および pNR4ES になった。両プラスミドは、*S. noursei* ATCC 11455 中に導入され、変異菌株 NR4EL および NR4ES をそれぞれ生成した。NR4ES 菌株 (*PermE**プロモーターからの 210 aa タンパク質を発現する) によるナスタチン産生は、pSOK804 のみを係留する野生型 *S. noursei* 由来のものと明確には異なっておらず、他方、NR4EL 組換え体 (*PermE**プロモーターからの 226 aa タンパク質を発現する) は、野生種のレベルよりも 36% 高いレベルでナスタチンを産生した。これらのデータは、長い方の種類、すなわち 226 aa の長さである種類のものは、実際の *Ny s R I V* ポリペプチドを表すことを示唆する。さらに、これらのデータは、*ny s R I V* 遺伝子が、*S. noursei* 菌株中でナスタチンおよびその誘導体の産生増加させるために使用できるプラスのアクチベーターをコードするという仮説を裏付ける。

20

30

【図面の簡単な説明】

【図 1】 図 1 は、ポリエン系抗真菌性抗生物質であるナスタチン A1 の構造を示す。

【図 2】 図 2 は、実施例 1, 2 および 3 で使用した *E. coli* - *Streptomyces* シャトルベクター pSOK101、pSOK201 および pSOK804 の物理地図および機能地図を表す。

40

【図 3】 図 3 は、ナスタチン生合成遺伝子 (SEQ ID NO. 1 および 2 に対応する) をコードしている *S. noursei* ATCC11455 ゲノムの 2 つの領域を示す模式図である。提示される DNA 配列を含む重複する組換えファージは、領域図の上に重ねて示される (実施例 1 参照)。

【図 4】 図 4 は、ナスタチン PKS NysA (SEQ ID NO.5)、NysB (SEQ ID NO.6)、NysC (SEQ ID NO.7)、NysI (SEQ ID NO.20) の機能の組織化およびナスタチン生合成におけるそれらの役割を示す模式図である。

【図 5】 図 5 は、組換え *S. noursei* 菌株によって新しいポリエン化合物の産生に至る NysC PKS のモジュール 5 に対する遺伝子操作を示す模式図である (実施例 2 参照)。

【図 6】 図 6 は、ナスタチン、ならびに遺伝子改変した NysC PKS を有する組換え *S. no*

50

ursei菌株から得られた新しいポリエン化合物S44およびS48に関するUVスペクトル（DMSO中）を示す（実施例2参照）。

【図7】 図7は、ナイスタチン生合成遺伝子クラスター（SEQ ID NO.35に対応する）をコードしているS.noursei ATC11455ゲノムの領域を示す模式図である。遺伝子クラスター内の遺伝子組織化が示されている。クローン化領域を包含している重複組換えファージからの挿入物は、上記物理地図／遺伝子地図に示されている。nys遺伝子はイタリック体の大文字で示され、他の翻訳領域（ORF）は番号を付けられている。

【図8】 図8は、S.noursei内でのナイスタチン生合成について提唱されているモデルを示す。

【図9】 図9は、ナイスタチンPKS NysA（SEQ ID No.5）、NysB（SEQ ID No.6）、NysC（SEQ ID No.7）、NysI（SEQ ID No.37）、NysJ（SEQ ID No.38）およびNysK（SEQ ID No.39）タンパク質の機能組織化を示す模式図である。KS^Sは、活性部位においてCysからSerへの置換を有するケトシンターゼである。KSは、ケトシンターゼである。ATは、アセテート特異的アシルトランスフェラーゼである。mATは、プロピオネート特異的アセチルトランスフェラーゼである。DHは、デヒドラターゼである。DHiは、不活性型デヒドラターゼである。ERは、エノイルレダクターゼである。KRは、ケトレダクターゼである。KRiは、不活性型ケトレダクターゼである。ACPは、アシルキャリアータンパク質である。

【図10】 図10は、ナイスタチン遺伝子クラスター内における下記の操作から理論的に生成され得る化合物を示す。

モジュール3へのERドメインの挿入（1）；

モジュール4へのERドメインの挿入（2）；

モジュール5のERドメインの不活性化と同時に行われるモジュール3へのERドメインの挿入（3）；

モジュール5のERドメインの不活性化と同時に行われるモジュール4へのERドメインの挿入（4）；

モジュール5のERドメインの不活性化と同時に行われるモジュール7へのERドメインの挿入（5）；

モジュール5のERドメインの不活性化と同時に行われるモジュール8へのERドメインの挿入（6）；

モジュール5のERドメインの不活性化と同時に行われるモジュール9へのERドメインの挿入（7）；

モジュール5のERドメインの不活性化と同時に行われるモジュール8および9へのERドメインの挿入（8）；

モジュール5のERドメインの不活性化と同時に行われるモジュール7および8へのERドメインの挿入（9）。

【図11】 図11は、ナイスタチン遺伝子クラスター内における下記の操作から理論的に生成され得る化合物を示す。

ナイスタチンPKSのモジュール11内のメチルマロニル特異的アセチルトランスフェラーゼ（AT）ドメインの、マロニル特異的ATドメインによる置換（10）；

モジュール12内のマロニル特異的ATドメインのメチルマロニル特異的ATドメインによる置換と、同時に行われるモジュール11内のメチルマロニル特異的ATドメインのマロニル特異的ATドメインによる置換（11）；

モジュール10内のマロニル特異的ATドメインのメチルマロニル特異的ATドメインによる置換と、同時に行われるモジュール11内のメチルマロニル特異的ATドメインのマロニル特異的ATドメインによる置換（12）；

P450モノオキシゲナーゼをコードしている遺伝子であるnysLまたはnysN（いずれかが、ナイスタチン分子上でのC16位メチル基の酸素化の原因である）の不活性化（13）。

【図12】 図12は、ナイスタチン遺伝子クラスター内における下記の操作から理論的に生成され得る化合物を示す。

ナイスタチンPKSのモジュール3内のデヒドラターゼ（DH）の不活性化（14）；

モジュール 4 内の DH ドメインの不活性化（15）；

モジュール 3 内の DH ドメインの不活性化と同時に行われるモジュール 5 内の ER ドメインの不活性化（16）；

モジュール 4 内の DH ドメインの不活性化と同時に行われるモジュール 5 内の ER ドメインの不活性化（17）；

モジュール 7 内の DH ドメインの不活性化と同時に行われるモジュール 5 内の ER ドメインの不活性化（18）；

モジュール 8 内の DH ドメインの不活性化と同時に行われるモジュール 5 内の ER ドメインの不活性化（19）；

モジュール 9 内の DH ドメインの不活性化と同時に行われるモジュール 5 内の ER ドメインの不活性化（20）。 10

【配列表】

SEQUENCE LISTING

SEQ ID NO. 1

DNA sequence: (Nys1; region 1)

1	gatccggccg gtgatgaact cctgctcctg tgcggtcagt ccggtgtgcg	
51	tcggcagata gaagccgtcg gcgctgagcc gggccgcgtt caacgcgggc	20
101	cagtcggcgg agaagtaacc ggctgacgg ctcctcggt tgaagaagac	
151	ccgggtctcg atgcctcgc cggcgaggtc cgcgcacagc tcgtcccgcc	
201	gctcggcccg caggctgtac atccacagca cgtcgcggcg cggcatcagc	
251	gtgatgcggg ggaatgtcgc cagcgccctc tcgtagcgct tctcgatgtc	
301	gcggcgcagg gcgaggatgg tgtccagctg ctgggtctgc gcgagcgcca	
351	ccgcggcctg catgtttgtc atccggaagt tgtaggcagc ettcctgtgc	
401	aggaagctgt ggtccttggg gaacgccatc gcccgagat gggccatctg	
451	ctcggccagg tgcgggtcgt gggtcaggca gacgcggccc tcgccggccg	
501	agatgatctt gttggcgaag agcgagaaac aggcgatgtc gccccgcggc	
551	cgcaccccggt gcgcctcggc ggagtcctcc accaccgca ggttgactc	30
601	gtacgccagg ttacgacgg cgtccatgtc gcactgccgg ccgtagatgt	
651	gcacggcat gatcactttg gtgcgcgggg tgatcttctc ctcgatgcgg	
701	gacacgtcga tgttcaggtc gtgcgcggcg tccacgaaca ccggcgtggc	
751	accggtgtag gtcaccgccc aggcggacgc gatcatcgtg aactccggga	
801	cgatcacctc gtcaccgggg ccgacgcccc gcgcgcggcg cgcagcgtc	
851	agcgcggtgg tgcgggagga gcaggcgacg ccgaacggca cgtcgttgta	
901	cgcggcgaac gctcctcga accgcctgac gtacggcccc tgcgaagaga	
951	tccagccgcc gccgacggcc tccgtcacat agtcgagctc gcggccctgg	
1001	agccacggca tggacaccgg atacgtaaag gacatgggtt ttgagtcttc	
1051	ctcggtcagt cgggtgccag ggcgggaagg ccgagcagca ggtccgcgc	
1101	cgcggcccg cgcgcggcgt cccgcagcag cccggcgaag tgctccgcgc	40
1151	gctcggtgaa ggagggttgg tcgagcacgc ggggtgatctt gtccaggacg	
1201	tcctcgggtg ccacgggtctc cggccggtec agggtcaggc tcaccccgaa	
1251	gtcctggccc cggatcgctt ggtcgtcgca gtccaccac aacggccgga	
1301	ccaccagcgg ctttcggaag tacaggccct cgtggtagcc gttgccaccg	
1351	gcatgggtga agaacgcctt cacgttcgga tgggccagca cgtccagctg	
1401	cgacggcacc cagccctcga tccgcaggtt gtccggcagc tcggcggccg	
1451	gcggcagcaa ctctgttgg ccgcgggga gtttcacaa cacctggtgg	
1501	ccccggcgt ccagtcgcgg ggcgacctc accagcgacg ccacctgtc	

1551 acgggtcagc cgggtgatcg tgccgaagcc catgtacacc acggacttct
 1601 gcgccgacag ccagtccgac aggccgtcgt cgtccggtgc ctggggcagc
 1651 ggccggcacca tcgtgcccac cagccgcagc ttccgatgca tcgggaacgg
 1701 gtagtccaac tcccttacgg agtagcaca gacctgctcc gcatgggtcg
 1751 tccgcgccat catctgccgc gcctggggcg ccatgcccag ctcggtgcgg
 1801 acccggttgt cctcctcgac gaccttgccg acgtccgacg tcaggaacat
 1851 cccgagcgtc cgcagccgga acagctggtt ctcgatccgc tgagccaggg
 1901 acatcgccgc cggcagcccc gagtgcggca cgggaaacc cgacggggtg
 1951 taggacttgg cgaacgggac gtgcgaggtg aggacgttgc tcggcacgaa
 2001 cggcaccctg agcacgaacg gaatgccctt ggtgatcgcc agctcgtacc
 2051 cgaactggca catgctctcg atcaccatca gcgccggctc gacctcctcg
 2101 acgatctcct ccaggcggcg gtacttegcc atccgcgact ccggcgcgaa
 2151 cgaatgccga atcacccgcg cgtgcgcctt gaaccgcgac cgtgcgtca
 2201 cctccgcata cgtcgcgtcg tcccatgtga ccgccgacat ctgcgagacg
 2251 gtgtcgccga gcgacgcgaa ccgaaccggg ctgccgtcca ccacggccgc
 2301 cacctcgtcg cgcgctttct cgtcgggtgg gaaccacagg tccgccacgt
 2351 cgcgccggga caattccccg gccagcacga gcagcggatt gagcaggccg
 2401 ctttcggcat aactgacgaa caggatcggc cgcggattcg cggccatgga
 2451 caacaccctc cggaatgtgg cgggccgcgc gggccgcgcg ccacgcaccc
 2501 gcccggcccg gtcgcggggt gagtgcattc gccgacgcgc ccacccgagg
 2551 cgcgtgttgc cggaaaggaag ggtcacccgc cggcaccgcg aacgcgcgcg
 2601 gtggaaaacg ggtcggttac ttggtctcat gccacggacc ggggaatcac
 2651 tagtcttcgg cgcgcgacgg ccttttcggg gccgtgtggc caatgcccgt
 2701 ccccgccgcc cgtcattcct tagggaaaag tacagcgttt gcgaacgtac
 2751 gatccggcac gcagagggtga cctgaggcca acttttcgcg aggggtgagc
 2801 aaggcatgac gatcggagcc gacgaggacc cgggtggtgg cgtcggaatg
 2851 gcctgcctgt atccgggtgg ggtcgcgcgc ccggaggacc tgtgggaact
 2901 ggtccgcacc ggcgcgcgac cgaccaccgc cttcccgac gaccgcggct
 2951 gggacctggc cgcactggcc ggcgacggac ccggccgcag cgcgacccgc
 3001 gagggcggat tcttcaccgg cgcgcgcgac ttcgacgccg ccttcttcgg
 3051 catgtcgcgc cgcgaggccg tctccacgga ccgcgaacag cgcctcgtcc
 3101 tggagaccgc ctgggaagcc ctggagcgcg ccggcatcga ccgcactcc
 3151 ctgcgcggca gccgcaccgg ggtcttcgtc ggcgcacgcg gccaggacta
 3201 cgcgcgcgtc acccagcctc cgcgcgacga cctggacgga cagccctca
 3251 ccggcctggc ccccgccgtc gcctccggtc gcctggcgta cgtcctgggc
 3301 ctggaaggcc ccgcgcgcac cgtcgacacc acgtcctcct cgtcgtgggt
 3351 cgcgctgcac tgggcgggtc gcgccctgcg cgcgggggag tgcagaccg

10

20

30

3401 ccttgcccg cgccgtcacg gtgatgtcca ccccgccgc cttcgtcggc
 3451 cacacccgac agggcgccct cgcgccgac ggccgctgca agccgttctc
 3501 cgacgacgcc gacggcaccg cctgggcgga gggcgtcggc atcgtcgtcc
 3551 tggagcacct gtccaccgcc cgcgccgcg gcaacccgt cctcgccgtg
 3601 ctgcgcggct cggccgtcaa ccaggacggc gcctccgacg gcctcaccgc
 3651 acccagcggc cccgccagc aacgcgtcat ccgcgccgc ctcgccgacg
 3701 cccgactcgc cccgccgac atcgatctcg tcgaggcgca cggcaccggc
 3751 acccggtctc gcgacccgt cgaggcccg gcgctgctcg ccgcctacgg
 3801 ccaggaccgg gaccggacc gaccgtcgc cctcggttcc ctgaagtcca
 3851 cctcggcca cgcacaggcc gccgccgga tcggcgaggt gatcaagacc
 3901 gtcttgacc tgccgacgg cctgatgcc cgcctccgc acctggccac
 3951 cccacccgc caagtgcact ggtcccagg cgcgtggcc cccctcaccg
 4001 accacacgcc ctggccacc gccgaccgac cgcgccgcg cggcgtctcc
 4051 tccttcggca tcagcggcac caacgcccat gtgatcctcg aagaggcgcc
 4101 gcccgccgac gtccccgtca cccggcccg caccctccgc cccagcaccg
 4151 tcccttgcc ggtctccgcc gccacgccg aagcctcga cgcaccaactc
 4201 gcccggtctc gcgccaccc gccgaccac tcggacctgg acccgttgga
 4251 cgtcggctac tccctggcca ccggccgcg cgcgtccgc caccggcgcg
 4301 tcctctgcc gcccgccgac ggccaccgc cggacgccgt cgagcaccgc
 4351 cgcggtgccc cccaccagc cgcaccgcc gtctcttct cggccaggg
 4401 cagccagcgc cgggcatgg gcccggaact cgcgccccg ttcgggtgt
 4451 tcgccgacgc actggacgac gcgctgcgc cctggaccg gcacctggac
 4501 ggcccggtgc gcgaggtgat gtggggcacc gacgcgcgc tctggaccg
 4551 gaccggttg acccagccc cctgttcgc cgtcgaggt gccctccacc
 4601 gcctggtcgc gtccctcggc gtcaccccgc acttcgtcgc cggccactcc
 4651 gtcggcgaga tcgccgcgc ccacgtcgc ggcgctctga gcctggagga
 4701 ccgctgcgc ctggtggcc cccgcgccac gctgatgcag gcgctcccg
 4751 ccggcggcgc gatggccgc ctggaggcca ccgaggacga agtggccccg
 4801 ctgctcggcg cacacctgc gctggccgc gtcaacggcc ccaccgcgt
 4851 cgtcgtcgc ggagccgag acgcgtgcg gcaactgacc gcccgcttcg
 4901 ccgaccgcg ccggcgacc agccggttg cgtctcga cgccttcac
 4951 tcgccgctga tggagcccat gtcgacgcc ttcgggacg tcgtgagccg
 5001 actgacctc caccagccgt cgatcccgct ggtctccaac ctcaccggtg
 5051 aactcgccg cagttagatc accagcgcg agtactgggt ccggcacgtc
 5101 cgcgacacc tcgcttcgc cgacggcat accgactgg ccaaggccg
 5151 cgcgacgtc ctgatcgaac tcggccccc cggcgtgctg tccgcgatgg
 5201 cccgcgacac cctcgcccc gacagacca ccgacgtcgt ccccgccctg

10

20

30

5251 agcaaggagac ggcccagagga gaccgccttc gccggcgccc tcggccgcct
 5301 gcacaccctc ggcgccccg tcgactggcc cgcttctac gccggcaccg
 5351 gcgccccgcg cgtcgaactg cccacctacg ccttccagca cgtgcgccac
 5401 tggcccaccc cgcgccgcc gaaaggcgcc ggccccggcg cctcggcca
 5451 cccctgctc ggctccgcg tcgaactcgc cgacggcgcc ggcaccgtct
 5501 gctccggcg cctctccctc cgcaccacc cctggctcgc cgaccacacc
 5551 gtgcggggc gggctcgtgt gccggccacc gcgtcgtgg aactcgccgt
 5601 gcgcgcggc gacgaggcg gctgcgacgt cctgcacgaa ctccacctca
 5651 ccaccccgcc ggccctgcc gacgacgcg cctgcacgt ccaggtgcac
 5701 gtccggcccc cgcacaccac gggcgccgc gccgtcaccg tccacaccg
 5751 ccccgaccac caccggcgcc gcgactggac ccgatgcgc accggcacc
 5801 tcggcagcac ccgcgcgtcc gcagccgaag ccgccacggg cggcaccgcc
 5851 gccgcctggc cgcggcgca cgccgaaccc ctgcacctc ccgaccacta
 5901 cgaacggctc gccgacgcg gcttcgacta cgccccgacc ttccggcgcc
 5951 tgccggccgc ctggcgacgc ggcgcggaga tcttcgcgga cgtggaatgc
 6001 ccgcccggca ccgcgcgca cgcgccgac cagggactgc accccgcct
 6051 gctcgacgc gccccggcac ccgccatggc ggtggacggc accgtgccg
 6101 tcgcctggca cgcgctccg ctgcacgcg tcggcgccac cgcgctcgg
 6151 gtccgcatcc gccccaccac gaccggcacg ctgacctca ccgcggtcga
 6201 cgtgcacggc gcgcgggtc tcaccgtcga ggccctcacc gcccgccgc
 6251 tgaccgacga ggaacgcgc gcccccgga cgcgcggca ggcccgggc
 6301 gagacgccc cgcacgccc ccggccccg ccgcggcgcc ccgccccg
 6351 ccggccggc gaacccctc cgcacaccac cgggtccac ccacccgcg
 6401 gccacctcgc cgcgctgcg ccggcgccc gggagcgcca gctgctggac
 6451 ctggtgcga cccaggccgc cgcgctcctg ggccacccc gccccgaggc
 6501 cgtcggcacc cgcagcgtct tcaaggagct gggcttcgac tcgttgccg
 6551 gcgtcgaact cgcgaccgc ctcacgccc gcaccggact gcgcctgcg
 6601 gccacctcgc tcttcaactt cccaccccc gaacgtgctg cccaccgct
 6651 cggggaactc ctgcgcgca cgcgccctc cgaccccggg gcgtacggag
 6701 aggaactcac caggttcgag gcgatcgtga cgaacctgcc gcaggacggc
 6751 ccgaacgcc gggcgctcgc ggaccggttg gacgccatcg tctccgact
 6801 ccgccagaac tcgcctgcag aggtgccctc ctcgacgag gacatcgaca
 6851 cgggtgctggc cgacagactg ctcgacatca tcgatgaaga gttcgaaacc
 6901 acatagagaa attgttgctt tcgttcgca cccgatgacg aggacggacc
 6951 gatgcaggaa cccagcaag gccagccgga ccagcaggag aaaatcgtc
 7001 actatctccg gcgggtcact tcagatcttc gccgtgccc ccgcccatt
 7051 ggcaactgg aatccaagga caacgagccc atcgccatcg tcggaatggg

10

20

30

7101 ctgccgaactt cccggcggcg tcaattcgcc ggaatccctg tgggacctgg
 7151 tgcgttcctg cggcgacgcc atttcggat tccccgtcga ccgcggctgg
 7201 gacctgga aa cctcaccgg aaacggcgac ggcagcagcg cccccacga
 7251 aggcggattc ctctacgacg ccgcggaatt cgacgcgcgc ttcttcggca
 7301 tctcgccgcg cgaggcgact gctatggacc ccagcagcg cctcctcctc
 7351 gaagtcgcct gggaggcgct ggagcgcgcc ggcatcgccc ccacagccct
 7401 gcgcggcagc cggtcggcg tggtcgctcg ctctaccac tggggcgcg
 7451 cctcggccga cgcgccacc gaactgcacg gccacgccc caccggcacc
 7501 gccgccagcg tgctgtccg ccgcctggcc tacaccctcg gcctcgaagg
 7551 cccggccgtc accgtcgaca ccgcctgctc ctctcctc gtcgcctcgc
 7601 acctggcgcg ccagtcctcg cgcgctggcg aatcctcgt cgccgtgatc
 7651 ggcggcgctc cgaatcctc cgagccgtcc gtcttcgtcg agttcagcgc
 7701 ccaggcgcg ctcggcaccg acggccgctg caaggcgctc tcgcagcgcg
 7751 ccgacggcac cgggtggggc gaggcgctcg gtgtcctcgt cgcgagcgg
 7801 ctctccgacg cgcagcgcaa cggccatccg gtctcgcgcg tgctgcgcg
 7851 ctccgcgctc aaccaggacg gcgcctccaa cggcctgacc gccccaacg
 7901 gcccctccca ggaacgggct atccagcagg cctcgcgccg gaccggcctg
 7951 accccgcgcg acatcgacgc cgtcgaggcg cagggcaccg gcacccggct
 8001 cggcgacccc atcgaggccc aggcctgct cgccacctac ggccagggac
 8051 acacccccga ccagccgctg tggctcggt cctgaagtc caacatcggg
 8101 cacaccagc cggccgcgcg cgtcgccggt gtcacaaaga tggcatggc
 8151 gctgcgcac gccacctgc cgcgacct gcacgcgcg gcgcctcct
 8201 cgcacgtgga ctggtccgc ggatcggtac gcctgctgac cgagggccag
 8251 cagtggccgg agaccggacg tcgcgcgcgg gccgcggtgt cctcgttcgg
 8301 catcagcggc accaacgcgc acgcctgct ggaacaggca cccaccccg
 8351 cggacaccgc ggacgcggc gacgacgcgc cgcacaccga accggccggc
 8401 gcgcgcgcg cgtgcctcgt gatcgtctcc ggacactccc cgcaggcgct
 8451 gcgcgaccag gccgcgcgc tggccgcgcg ggtcgagacc gacccgcgc
 8501 tcgccccca ggacatcggg cacaccctgc acaccgccg gcgcctgctc
 8551 gaacgacgc cgtcgtcgt cgcgccgac cgcgcgaac tcctcgcggc
 8601 taccacagc ttggccgcgc gccggtccgc gaacgcgctc gtcgaggcc
 8651 tcgcggacgt cgagggtcgc acggtgttcg tggtccccg tcagggttcg
 8701 cagtgggtgg gtagggggc ccaactcctc gatgagtcg cgggtgttcg
 8751 ggagcggatt gccgagtgct cggcggaact cgcgagttc accgactggt
 8801 cgtggtcga tgtgctgcg ggtgtggtg gtgcgcgcg gttggagcgc
 8851 gtcgatgtgg tcgagccgc gtcgttcgc gtgatggtgt cgttggtcgc
 8901 gttgtgggt tccgtggtg tggtgcgga tcggtggtg gggcattcgc

10

20

30

8951 aggggtgagat cgctgccgcg gtggtgtcgg gtgcgctgtc gttgcgggac
 9001 gggggcgcggt tggtagcgct gcggagtcag gccattggtc gtgcgttggc
 9051 gggggcgggc gggatgatgt ccgtcgcgct gtcggtggac gtgctcgaac
 9101 cgcggttgggt cgagttcgag gggcggggtgt cgggtggccgc cgtcaacggc
 9151 ccgcgctccg tcgtggtcgc cggcgagccc gaggcgctgg acgcgctgca
 9201 cgccccggtg accgccgacg acatccgggc ccgcgggacg gcggtggact
 9251 acgcctcgca ctgcaccag gtcgaggacc tgcacgagga actgctggag
 9301 gtgctggcgg agctggcgcc gcgcacgtcg gaggtgccgt tcttctcgac
 9351 cgtgaccggc gactggctgg acaccgcgcg gatggacgcc ggctactggt
 9401 tccgcaacct gcgcggacgg gtgcggttcg cggacgcggt ggcggacctg
 9451 ctggcgcgcg agtaaccgac gttcgtcgag gtcagctcgc acccggtgct
 9501 gacgatggcg gtcttggaac tgatcgagga ggcgggggtc acggccgtcg
 9551 cgaccggcac cctgcgcgct gaccaggggt gcgcggggcg cttcctgctg
 9601 tcggccgccc aggtcttcgt gcgcggtgtg gacgtggact gggcgggggc
 9651 gttcgagggg accggtgcgg cccgggtcga cctgcccacc tacgccttcc
 9701 agcgcgagcg gtactggaac acccgaccgc ccgcggaccg caccgccggc
 9751 gacgccccga tggacgccga attctgggccc gccgtcgaa agggcgacgt
 9801 ctccgcgctg accgccgcgc tcggcaccga cgaggactcc gtcgcgcga
 9851 tcctgcccgg cctcacctcc tggcgccggg cccgctccca gcgcaccacc
 9901 ctcgactcct ggcgctaccg cgtcacctgg acgcccctcg cccaggtgcc
 9951 ccgcgccacc ctgaccggca cctggctgct ggtcaccacc gacggcatcg
 10001 acgacaccga tgtggcaggg gcgttggaga gctacggcgc cgaggtgcgc
 10051 cggctggtcc tggacgagga gtgcaccgac cgcgcgctcc tgcgggagcg
 10101 gctggccggc gcggaggacg tgaccggcat cgtctccgtc ctgcgcccg
 10151 ccgaggacga cgcgcacgc caccgccggc tcaccggggg actcgcgctc
 10201 accgtctccc tcgtccaggc cctgggagac gccgaggcga ccgcgccgct
 10251 gtggttccctg acccgcgggc ccttcgccac cggcccgctc gaccccgta
 10301 cccggccctt gcagagccag atcgcgggcg tcggctggac caccgcgctg
 10351 gagcaccgcg agcgctgggg cggcaccgtg gacctgcccg acacctcga
 10401 cgccccggcc gccagcggc tcgcccgcgc gctgtccggc gccctcggcg
 10451 ccgaggacca gctcgcgctc cgcgcgcgcg gggtaactgg ccgcgcgac
 10501 gtgcgtgcgg gacaccgcgc cggacgacgg gcacggacct gggcgccgcg
 10551 cggcaccacc ctgatcaccg gcggctccgg caccctcgcc ccgcagctcg
 10601 cccgctggct ggcgaacgc ggcgccgagc acgtggtgct ggtcagccgg
 10651 cgcggtgccc acgcccccg agcggccgaa ctcatcgcg aggcagccga
 10701 gtccggcacc gaggtgaccg tcgcccctg cgacatcacc gaccgcgacg
 10751 cggctcgcgc gctgctggcc gacctacgg ccgacggcg caccctgcgc

10

20

30

10801 accgtcatcc acgcccgcgc cgccatcgag ctgtccgcgc tcgccgacac
 10851 caccgtggcg gagttcgccg acgtcgtgca cgccaaggtc accggcgcac
 10901 ggatcctcga cgaactgctc gacgacgcgg aactggacga cttcgtcctg
 10951 tactcctcca ccgcccgcac gtggggcagc ggcgtgcacg ccgcctacgt
 11001 cgccggcaac gcctatctgt ccgcgctcgc cgagcagcgc cgccccgcg
 11051 gactgcgcac cactccatc cactggggca agtggcccga cgaccgggca
 11101 cgcgagctgg ccgaccgcga ccggatccgc cgcagcggtc tggagtacct
 11151 cgaccccgag ctggcgctca ccgcgctcca gcacgtcctg gacgacgacg
 11201 agaccgtcat cggcctcatg gacatcgact gggacaccta ccacgacgtg
 11251 ttacccgcgg gccggccgcg gcacctcttc gaccagatcc ccgaggtgcg
 11301 gcgcgcctc gaccaggcat ccgtcccga ccccgcgggc ccggccgcg
 11351 acggcctggc cgcccgctg caccgctcgc ccgcccgcga acaggaccgg
 11401 ctgctgctca ccttgggtcc caccgaggcc gccgcccgtc tcggccacgc
 11451 ctgggcccga tcttccccg agcgcgcgcg cttccgtgac ctcggttcg
 11501 actcggtcac cgccgtggac ctgcgcaacc ggctcgtggc cggcaccgga
 11551 ctgcggtgc ctcgacgat ggtcttcgac caccctaaact gcgcggcgct
 11601 cgcccgcttc ctgaagacga cggcgctcgg cgtccccggc gccgcaccgc
 11651 agcagcacgc cgctaccggc accccggccg acgacgaccc gatcgccgtg
 11701 atcggcata gctgcgcga ccccgggggc gccgccacc ccgaggaact
 11751 gctgcggctc gccctcgacg gcgcgcgacg catctcgag tccccgcgg
 11801 acccgggctg ggacgcccgc ggcctgtacg acccggaacc cgaccgccc
 11851 ggccacacct actccgtcca gggcggttc ctcacgagg ccgcggctt
 11901 cgatcccgcg tcttcggga tctccccgcg cgaggcggtc gccatggacc
 11951 cgcagcagcg gctcctgctg gagacctcct gggaggcggt cgaacgcgc
 12001 ggtatcgacc ccgcgtcact gcgcggcagc gccgcggca ccttcttcg
 12051 cgccagctac caggactact cctccaccgt gcagaacggc acgggggagt
 12101 ccgaggcgca catggtgacc ggcaccgcgg ccagtgtcct gtccggccg
 12151 gtctcctacc tgctcggcct ggaggggccc gcggtcaccc tggacaccgc
 12201 ctgctcctcc tactgggtc cctgcacct ggcctgccag tccctgcgcg
 12251 acggcgagag ctccctcgcg ctggccggcg gtgcggccgt gatggccacc
 12301 ccgcacgcgt tcgtcggctt cagccggcag cgtgccctgg ccaaggacgg
 12351 ccgctgcaag ccgttctccg acaccgcca cggcatgacg ctgcgcgagg
 12401 gcgtcggcgt cgtcctgctg gagcgctgt cccacgccc cgccaacggg
 12451 caccgggtac tggccgtgat ccgcgggttc gccgtcaacc aggacggcgc
 12501 ctccaacggc ctgaccgcgc ccaacggccc gtcccagcag cgcgtcatcc
 12551 gccaggcgct cgccaacgcc ggcctgaccg gcgcgcgacg cgacgcggtc
 12601 gagggcgacg gcaccggcac caagctgggc gaccccatcg agggccaggc

10

20

30

12651 cctgctcgcc acctacggcc aggacogcga cgcogaacgg ccgctgctgc
 12701 tgggctcggg gaagtccaac atcggccaca cccaggccgc ggccggcgtc
 12751 gccggcgtea tcaagatggg gctggccatg gacgcccggc aactgcccgg
 12801 caccctgcac ctcgacgcgc cctccagcca cgtcgactgg accgccggcg
 12851 ccgtggaact gctgcgcggg cgcaccccggt ggcccagagc cgggcccccc
 12901 cgcggggccg gtgtctcttc gttcggcatc agcggcacca acgcccacct
 12951 gatcctcgaa caggccccgg ccaccgagcc gccagccgac cccgaccgcc
 13001 tccgggaacac cgcaccgac accgtcgcc cctggccgct cgcgcgaag
 13051 tccccggccg cctgcgcgc ccaggccgcc cggctcctcg ccaccgtcga
 13101 gcacgacccc gacctccgc cgcgcccggt gggccacgcc ctggccacca
 13151 cccgcgccgc cctcgaacac cgcgcgctcg tcgtcggcga gcgcgcgag
 13201 gacttctcgc gcggcctggc cgcctgtcc accggcgct cgcgcgcg
 13251 cctggtcagc ggcatcgccg gcccgcacc cgcgggagcg gtcttcgtct
 13301 tccccggcca gggatcccag tgggtgggaa tgggcccga actcctcgc
 13351 acgtccgagg tgttcgcac cgcgatcgat gactcgcga cggccctcgc
 13401 cccgtacgtc gactggtcgc tgcacgacgt cctggccggc gaggcgacc
 13451 ccgccctgct ggagcgggtg gacgtggtcc agcccgcgt gttcgccatg
 13501 atggctgggc tgtccgcgt ctggcgctcc cagggcgctg tcccggcggc
 13551 cgtggtcggc cactcgcagg gcgagatcgc cgcggcctgc gtcgcggag
 13601 cctcagcct ggcgcgcgc gccgcgtgg tggcgctgc cagccaggca
 13651 ctgccgaac tgtccggacg cgcggcatg atgtcggct cgcccccgt
 13701 agagcgggtc accgcactcc tcgccccgtg gcaggaggcg ctgtcgtcgc
 13751 ccgcggtcaa cggccccctg tccgtggtcg tctccggcga caccgacgc
 13801 ctcgacgccc tgcacaccgc ctgccaggaa caggcgctgc gggcccgcaa
 13851 ggtgtcgtg gactacgcct cgcacgggcg gcacgtcgag gccgtccgc
 13901 acgaactcgc ccgcgtctc gcgcgggtcg accgcgcgc ccccgaggtg
 13951 ccgttctact cgcaggtcac cggcgaccgc gtggacgac cgccttcga
 14001 cggcgccctac tggtaacca acctccgcca gaccgtccgc atggaggagg
 14051 ccaccgcgc cctcctcgc gccggacacc gcgtcttcac cgcggtcagc
 14101 ccgcaccccg tgcgcgcgc cccgatccag gagacgcagg aggcctagc
 14151 ggaggccacc ggcggttcg cgtggtcct cggctcgtc cgcgcgacg
 14201 agggcgggcc gcggcgcttc ctgacgtcgc tcgcgaggc ccacaccac
 14251 ggcgcccccg tcgactggac caccaccttc gccgggtcc cctaccagcc
 14301 ggtggacctg ccgacctacc cctccaaac acaggacttc tggcccgagg
 14351 cccggccgc caccgggccc gccggcgcc agcgtccga cgcgcgttc
 14401 tggcaactgg tcgagaacca ggacctcgc gcgctcgcg acgcgctcg
 14451 cgtccccgcc gacgacgagc acaccgcgt cggcacctg ctgcggccc

10

20

30

14501 tgtccgcctg gcgcgccaag gcccaggccc gcacccggat cgacgaactc
 14551 cgctaccacg tccagtggac ccgggtcgcc gagcccgagg cggccccac
 14601 caccggccgg ctgctggtcg ccgtcccgcc ggaccacgcc gacgccccct
 14651 gggtcgcccg ggcgctcgac gccctgggca ccgacaccgt ccgcttcgag
 14701 gccaaaggca ccgaccgcgc gggatgggcc gcacagatcg cccaactcgt
 14751 cgaggacggc gaggagttca ccggcgtagt gtgctgctg ggcgcgccc
 14801 aggatctcca cccggaactc ggctcggtag cgctggggct ggggcagacc
 14851 ctgctcctcg tccaggccct cggcgacgcc ggctgaccg cggccctgtg
 14901 gtgcctgacc cgcggcgccg tcgccaccgg ccgcgacgac gccctcgaca
 14951 gcccgaccca gggcgccctg tggggcctcg gccgggtcgt ggccctggaa
 15001 caccocgacc gctggggcgg cctgatcgac ctgcccgcc cctcgacgc
 15051 ccgcgcgcg gcccgccctca ccggcctgct cggcgacccc gccggtgagg
 15101 accaactcgc cgtccgcgcc accggcgtgc tcgcccgcg catggtgcac
 15151 gccgcgcctg ccgcgccccg caccggcgcc cgctggcggg gccgcggcac
 15201 ctgcctgatc accggcgcca ccggcgccat cggcgcccg gtgcgccgt
 15251 ggatggccga gcacggcgcc gccacactgg tctgaccag ccggcgcgcc
 15301 ccggacgcac ccggcgccgc cgcactccgg gccgaactgg aggcctggg
 15351 cggccgcgtc accctcgccg cctgcgacgt cggcgaccgc gacgcccgtg
 15401 ccgcctgct ggccgacctc cccgcgacc agccgctcac ctccgtcttc
 15451 cactccgcg gcgtggccga cggggacgcc cgggcagccg acctgacct
 15501 agatcagctc gacgcgctgc tgcgcgcaa actgaccgcc gccaccacc
 15551 tgacagagct gaccgcccc ctgcacctg acgcgttcgt gctcttctcc
 15601 tcggcgccg cggctcgggg cagcgcgccg cagcccggt acgcgcgcg
 15651 caacgcctac ctgcagccc tcgccgccc ccgcaggtcc ctgcacctgc
 15701 ccggcgcgct cgtcgccctg ggcacctgg gcgaggtcgg catggccacc
 15751 gtccccgagg tccacgagcg actgcaccgc caaggggtcc gcgccatgga
 15801 accggaccac gcgatcgccg cgctccagca gatgctggag gacgacgaca
 15851 ccacctcgc cgtgacctc atggactggg aggcgttcgc accgagctc
 15901 accgcgacc gccccagcgc cctgttcagc acggtgccc aagccgtccg
 15951 cgcggtgacc ggcgacccg gcaccacggc cggcgacgac gtggactccg
 16001 cgaccccgcc gctccgcgc caoctggagg agctgtccg cgcgagcgc
 16051 ggccggggcc tggtcgaggc ggtccgcgcc gaggcgtccg cgacctcgg
 16101 ccacgacacc cccgacgcca tccccgcgg ccgtgccttc cgcgacgtc
 16151 gcttcgactc ggtcaccgcc gtcgaactgc gcaaccggct gcgcaccgcc
 16201 ctgggctgc cgctgcggc cgcgctcgtc ttcgaccacc cccccccac
 16251 ggccctcgcc ggccacctg gcgcgctgt cttcggcacc gccccgagg
 16301 acgcgggcac cggccgcccc gacgacccc acgcccgat ccgcgaggcg

10

20

30

16351 ctcgccaccg tccccatcgg acggctgcgc aaggcgggcc tcctcgacat
 16401 ggtgctgaaa ctcgccgacg gagacgcgac cgacgcccc gccccgagg
 16451 ccgacgcccc ctcggaatcc ctcgacgaca tggacgccga agccctgctg
 16501 cggctggcca ccgagaactc ggcgaactga aagagagctg gagccacca
 16551 ttagcacgaa ccccgacaag tacgtcgagg cactccggtc gtccctgaag
 16601 gagatcgagc ggctgcgccc gcagaacgaa cagctggtgg cggcggcgg
 16651 cgagcccgctc gcggctcgctc gcacgcgctc ccgcttcccc ggcggcgctc
 16701 cctccccga ggacctgtgg gagctggtcg ccgaggggag cgacgtcatc
 16751 gggccgttcc cgcaggaccg cggctgggac ctggagaagt tggcggcgg
 16801 cggcggaggc ggcagcctcg cgcaggctcg cggcttcgtc gaggacgcc
 16851 ccggttcga ccccggttcc ttcggcatct ccccgcgaga agcggctcgc
 16901 atggaccgcg agcagcgcat cctgctggag atcacctggg aagccctgga
 16951 acgcgcgcgc atcgaccggt ccaccctgcg cggtaacccc accggcgtct
 17001 tcgtcggcac caccggccag gactacggcg aggtcatcaa ggcgtccgcc
 17051 gaggacgtcg aggtctactc gaccacggcg caccgcccga gcgtcatctc
 17101 cggccggctc tcctacaccc tcggcgccga gggcccgccc gtcaccgtcg
 17151 acaccggctg ctctcgtcc ctggctgccc tgcactgggc cgtccaggcg
 17201 ctgcgcggcg gcgagtgtc catggccctg gccggcggcg cgtccatcat
 17251 ggccaccccg ggcccgctcg tcgccttcac cgcgcagagc ggccctggcc
 17301 ccgacggcgg ctgcaagccc ttctccgacc gggccgacgg caccggctgg
 17351 ggcgagggcg ccggcatgct ggtcctgatg cggctctccg acgcccagcg
 17401 cgagggccgc ccggtcctcg ccgtgctgag cggctccgcc atcaaccagg
 17451 acggcgctc caacggcctg accgtcccca acggccctc ccagcagcg
 17501 gtcacccgcg ccgcgctgga cagcgcacac ctaccgcgg ccgacatcga
 17551 cgccgtcgag gccacggca ccggcaccaac cctcggcgac ccgacgagg
 17601 cccaggcgct cctggcgacc tacggacagg accggcggcg gcccctgtgg
 17651 ctgggctcgg tgaagtccaa catcgccac acccaggccg cctccggtgc
 17701 cgccggcgtg atcaaatga tcatggcggt gcagcgcggc gtgctgccgc
 17751 gcagcctgca cgccaccgaa cccaccacgg acgtcgactg gaccgcccgc
 17801 tccgtcgacc tcctcgacga gacggtcgcc tggcccgaga ccggacgcgc
 17851 ccgcgcgccc ggcgtctcct ccttcggcat cagcggcacc aacgcccacg
 17901 tcacctcga acaggccccc accgccccg aagagccac caccgaaccc
 17951 accgtccgcc ccgcgctcgt ccgctgggag ctctccgccc gaccgcccgc
 18001 cgccctcgac gcccgcgcg cccgcctcac cggccacctc gccgacccc
 18051 ccgacgccga cccctcgac gtcggctacg cgctcgccga cggacgcgcc
 18101 accttcgaac accgcgcgct cctgtcccc gacggcaccg aactcgcccc
 18151 cggaaccgcc ggcgaaggcc cctgcgcgct cctcttctcc ggccagggct

10

20

30

18201 cccagcgccc gggcatggga cgcgaactcc acgcccgtt cccggtgttc
 18251 gccgcgccct tcgacgagat cacagcgctc ctgacaccc acctcgaccg
 18301 cccgctgctc gaggtcgtct ggggcaccga cgcgacctg ctgaacgaca
 18351 ccggtcgggc ccaaccggcc ctgttcggcg tcgaggtcgc cctctaccgc
 18401 ctggctcgcgt ccctcggcgt cacccccgcac ttcgtcggcg gccactccat
 18451 cggcgagctc gccgcgcgc acgtcgcgg ggctctctcc ctgaagacg
 18501 cctgcacct cgtcgcgcgc cgcgcgcgc tcatgcaggc cctgcgcgc
 18551 ggcgggcgga tgctcgcgat ccgcgccacc gaggacgagg tcacccccca
 18601 cctcaccgac gacgtctcga tcgccgccgt caacgggccc acctccgtcg
 18651 tcgtcgcgcg caccgaggaa gccgtcgcgc cgatcggggc gcgtttcacc
 18701 gccaggacc gcaagaccac ccggtcgcgc gtcagccacg ccttccactc
 18751 gccgctcatg gaccgatgc tggcggaatt ccgcgcgcgc gcccggggc
 18801 tgacctacca cgagccgcgc atcccggtec tctccaacct caccggcacc
 18851 gtgcgcgcgc tcgcgcacct gtgtcgcgc gactactggg tcgccacgt
 18901 ccgcgaggcg gtccgtctcg ccgacggcgt caccgcctc accgaccgcg
 18951 gcgtgaccac gctcgtcgaa ctggcccg acggcgtgct gtccgccatg
 19001 gccaggaat ccctgcgcga cggcgccgc gccgtgcgc tgctgcgaa
 19051 ggaccgcccc gaggagctct ccgcgcgcac cggcctggcc cgcgcacag
 19101 tcgcgcgcgt caccgtcgc tgggccggcc tcttcgacgg caccggcgcg
 19151 cgcgcgcgc acctgcacc ctaccccttc cagcaccagc ggttctggcc
 19201 gaccgcgcgc cgcgcgcgc aggacgtcac cgcgcggga ctgggcgcgc
 19251 ccgaccaccc gctcgtcggc gccaccgtcg aactcgcga cggggccggc
 19301 tacttgttca ccagccggct ctccgtcgcg acccaccct ggctcgcga
 19351 ccacggggtc caggccggg ccctgctgcc cggcaccgc ttcgtcgaac
 19401 tggcgcctcg cgcgcgcgc gaggccggct gcgaccgcgt cgaggaaactg
 19451 accctggcgc cccctcgtg gctgcccag cgcggcgcg tccaactcca
 19501 ggtcgcgcgc ggcgcgcgc acgcgcgcgc cgcgcgcac ctcgcatct
 19551 tctccgcgt cagggaaggc ttcgacctgc cctggtcga acacgccacc
 19601 ggcgtcctga ccgcgcgcgc cgcgcgcgc gacccacct tcgacgccac
 19651 cgtctggccc ccagcggcgc cgaacccgt cgacctcacc ggcgctacg
 19701 agcgcctggc cgactcggc ttccagtac gcccgcctt ccagggcctg
 19751 cgcgcgcct ggcgccgga caccgaggtc tacgccgaag tggccctgcc
 19801 cgacggcgcg gacaccgac ccgcgcctt cggactgcac ccggccctgc
 19851 tggacgcgc acaacacgc gccgcctac ccgacctcg cgcctcagc
 19901 cgcggcgccc tgccgttcgc ctgggaaggc gtctcgtcgc ccgcgcgcg
 19951 cgcaccacac gtccgcgcgc ggatcgcgc ggccggcgag gacaccgtca
 20001 ccacgcgcgt ctacgacgc gccggcgga ccgtgctgct cgtcgcactc

10

20

30

20051 ctggtctccc gcgaggtccc cgccgacgca cccggcgccg ccggcaccgt
 20101 ccaccgcgac tccctcttcc acgtcgagtg gaacccgctc caggggcgcc
 20151 cgggccccgc accggccacc gtcgcccgtc tcggccccga cccggacgcc
 20201 ctgcccgcaca cctccgcgc caccggcctc cggaccaccg cccccgcga
 20251 cctggccgccc ctgcccgcgc ccgaaggggc cgtccccgac ctggtcgta
 20301 ccacccctcac caccaccccg ggcgcccccg tccccgacgc cgcgcacgcc
 20351 accaccgccc ccgtcctcgc cctcgcccaa cagtggctcg ccgacgaccg
 20401 cttcgcccgc gcccgccctg tctctgtcac ccgcgcgccc accgacggca
 20451 ccgacccccg cgcgcgggccc gccggcgccc tgatccgcac cgcgcgacc
 20501 gagaaccccg gcggtttcgc cctcctcgac ctgcccccg acaccggccg
 20551 gcccgacccc gagaccctgg ccaccgccct ggccgcccagc cagcagcagc
 20601 ccgacctcgc cgtccgcggc accgacgtgc acgcccgcg cctggcccgt
 20651 gtcccgcctc ccaccgaacc caccacctgg aaccgggacg gcaccgtcct
 20701 gatcacccgc ggcacggcg gccctggcgc ggtcctcgc cgcacactgg
 20751 tcgccaccca cggcgtccgc cacctgctgc tcgccagccg ccgcgggccc
 20801 gcgcgcgacg gcgcgcgaca cctgacggcc gaactcaccg ggctcggcgc
 20851 caccgtccac atcgccgcct gcgacgtcgc cgaacccggc gccctcgcg
 20901 acctgctcgc caccgtcccg gccgggcacc cgctcaccgt cgtcgtccac
 20951 accgcccggc tcgtcgacga cggcgtcctc ggctccctca cccgcagcg
 21001 cctggacacc gtctgcggc ccaaggccga cgcgcctgg cactgcacg
 21051 aggcgacccc ccacctcgac ctggacgcct tcgtcctctt ctgctccgtc
 21101 gccgccaccc tcggcagccc cggacaggcc aactacccg ccggcaacgc
 21151 ctctctggac gccctgcgc cccggcgcg cgcacccggc ctgcccgcga
 21201 cctccctcgc ctggggcccg tggacccaga gcgtcgcat gacaagcagc
 21251 ctgtccgacc tcgacgtcga gcgcctcgcc cgctccggca tgccccgct
 21301 gacctggaa cagggcaccg cctctctcga cgcggccctg gccgcgggc
 21351 ccgcgcctc cgcgccggtc cgctcgcacc tgcccgtcct gcgcacccag
 21401 ggcgacatcg ccccgctgct gcgcggcctg atccgcaccc ccgtgcgcg
 21451 caccgcgcgc caggctcgc agaccgccga cggcctcgcc cagcggctcg
 21501 ccggcctcga cgcgcgcgc cgcgggaag cctcctgga actcgtccgc
 21551 acccagatcg ccaggctcct cggccacgcg gacgccaccg aggtggagac
 21601 cggcgcacg ttccaggacc tcggcttcga ctccctcacc gccgtcgaa
 21651 tcgcgaacgc cctgaacacc gccaccggcc tgccgctgccc cgcacccatg
 21701 gtgttcgact acccgacacc acacgccctc gccgaccacc tcgcgcacga
 21751 actcctgggc accgaggccg agtcgaccac cgcgtcccc gtgcccagcc
 21801 gtaccgcccg caccgacgac ccgatcgtca tcgtcgcat ggctgcccgc
 21851 taccgcggcg gcctcgctc acccgaggac ctctggcgcc tggtcagcca

10

20

30

21901 gggcgccgac gccactggcc cgttccccac caaccgcggc tgggacctgg
 21951 acaacctcta cgaccccgac cccgaccgcc cgggcccgcac ccacgtccgc
 22001 gccggcggct tctgcacga cgcgggtccc ttcgacgcgc acttcttcgg
 22051 gatgagcccg cgcgaggcga tggccaccga ctcccagcag cgcctgctgc
 22101 tcgaactctc ctgggaagcc gtcgaacgcg ccggcatcga ccccgccctc
 22151 ctgcgcgact ccggcaccgg cgtcttcgcc ggcgatcatg acaacgacta
 22201 cggcaccacc ctgaccggcg acgagtaaga ggcgttccgc ggcaacggca
 22251 gcgccccgag cgtgcctccc ggccgcgtct cctacacctt cggcctggaa
 22301 ggccccggcg tcacggtgga caccgcctgc tcttcctctc tggtcgccct
 22351 gcaactggcg gcgcaggcgt tgcggggcgg ggagtgctcg ttggcgttgg
 22401 ccggtggtgt gacggtgatg tcgacgccga gcacgttcgt ggagttctcg
 22451 cggcagcggg gtctggcgcc tgatggtcgt tcgaaggcgt tcgccgaggc
 22501 cgcggacggc gtggcctggt ccgaggcggt cggcatgctg gtccctggagc
 22551 ggcaagtccga cgcggtgcgc aacggtcacg agatcctggc cgtggtgcgc
 22601 ggctcggcgg tcaaccagga cgggtgcgtc aacggtctga ccgcgcccc
 22651 cggcccgctc cagcagcggg tgatccgtca ggcgttggcc agtggcgggc
 22701 tgtccacggc cgacgtggac gccgttgagg cgcacggcac gggtacgacg
 22751 ctcggtgacc cgatcgaggc ccaggcgctc ctggccacct acggtcgcga
 22801 ccgcgacccc gagaaaccgc tgctgctcgg ctcgatcaag tccaacatcg
 22851 gtcacaccca ggcaaggccc ggtgtgcgg gtgtcatcaa gatggtcatg
 22901 gcgatgcggc acggcggtgt gccgcagacc ctgcatgctg acgcgcctc
 22951 ctgcgcgctc gattggagcg tcggcgccgt cgaactgctc accgagcaga
 23001 ccgcctggcc ggagaccggc cgggcccgtc gcgcgggtgt ctctccttc
 23051 ggcatcagcg gcaccaacgc ccacgtcgtc atcgagcagt ccccgaccgc
 23101 cgtccccgcc acgcccgcgt ccgccgaacc gtccgtcgag gaaccgcgg
 23151 ccgtccctcg ggccctgtcc ggcaagaccc ccgacgccct ccgcgaccag
 23201 gccgcccgcc tcttcgccca cgtcgaggcc caccgccac tgcgcccgt
 23251 cgacatcagc tactccctga tcgccacccg caccgccttc gaccaccgcg
 23301 ccgtcgtcct cggcaccgac cgcgcgagg ccctgcgcgc cctcaccgcc
 23351 ctgcgcgcgc gcgagaccga cccggccgcc ctaccggca ccgtccgcac
 23401 cggccgcacc gccttcctct tctccggcca gggctcccaa cggctcggca
 23451 tgggcccgcgt cctctacgag cggttccccg ccttcgccga agccctcgac
 23501 accgtcctca ccgcccctga cgcggaactc ggccaccccc tccgcgacat
 23551 catctggggc gaggacgctc aactcgtcga ccggaccggc tacaccaaac
 23601 ccgcccgtgt cgcacatcga gtggcactct tccgcctcct ggaagcctgg
 23651 ggcatcacac cggacttcgt ggccggccac tccatcggcg agatcgcgc
 23701 cgcacacgct gccggcgctg tctccctcgg cgacgcctgc cgcctcgtcg

10

20

30

23751 tggcccgcg cgtgctgatg cagtcgctgc ccgaaggcgg cgcgatgatc
 23801 gccgtccagg ccaccgagga cgaggtccctg cccctcctca ccgacgacgt
 23851 ctcgatcgcc gccgtcaaca gcccgacctc cgtcgctcgtc tccggctacg
 23901 agaacgccac cctcgccgtc gcccggaact tcgcccacca gggccgcccgc
 23951 accacgcggc tgcgcgtcag ccacgccttc cactcgccgc tgatggcgcc
 24001 gatgctcgac gacttccgag ccgtcgctga gagcctcacc ttcaccgccc
 24051 ccacgacccc cgtcgctctcc aacctgaccg gcgaactggc cccggccgag
 24101 gcgctctgct cggccgacta ctgggtccgg cactccgag aggcgggtccg
 24151 ctccgcccgc ggcatccgca ccctcgccga ccgcccgcgtc accaccttcg
 24201 tcgaactcgg ccccgacagc gtgctgtccg ccatggccca ggagtccgccc
 24251 cccgaaggcg ccggcaccat cccgctcctg cgcgcgacc gggccgagga
 24301 acaggccgtc ctggccgccc tctgccacct ccaggtgctc ggcgtcgagg
 24351 ccgactggtc cgccaccttc cgcggcctcg acccgcctcg cgtcgacctg
 24401 ccgacctacg cctccagca ccgctgggtc tggcccgcgg cccgacccgc
 24451 ccgccccgac gacgtccgag cgcgcggcct gggcgccgccc gaacaccccc
 24501 tcctcggcgc cgcggtgcaa ctccccgacg acgacggcgc actcttcacc
 24551 ggcgcgctct ccctgcgcac ccacccgtgg ctggccgacc acacgcctct
 24601 gggcacccgtc ctgctcccg gcaccgcact ggtggaactc gccgtccgag
 24651 cgggcgacga gaccggcagc ggccacctcg aagaactcac cctcgccgag
 24701 cccctgaccc tccccgagga cggcgccacc ctctccagg tccggtcgg
 24751 atccgcccgc gacaccggcc gccgcaccgt caccgtccac gcccgccccg
 24801 acgacaccgc cgaccgcacc tggacgctgc acgccaccgg tgtgctcgcc
 24851 accacgccac cggccgcccgc ggcgttcgac accacggtct ggcgcccgc
 24901 cgacgccgaa cccctcacca ccgacgactg ctacgcacac ttcaccaccc
 24951 accgcttcgc ctacggcccc gccttcagg gcctgcgggc cgcctggcgc
 25001 gccggcgagc tgctgtacgc cgaggtcgc ctgccggagt ccgccaccga
 25051 cgaagcgggc gccttcggcc tgcacccggc gctcctggac gccggcctgc
 25101 acgcccgcct cctcgccgac gaccgcgaca ccggactccc gttctcctgg
 25151 gaaggcgctc ctctgcagc ctccggcgcc accgcgctac gcgtccggct
 25201 cgccccgaac ggccccaacg gcctgtccgt caccgcgcgc gaccggccg
 25251 gcaacccgct cgccaccgtc accgcctgc tcgccccccc cctggacgcc
 25301 gagcagttga ccatccacag cgccctgacc cgcgacgcgc tcttccacct
 25351 ggactggacc ccggtcccg ttcggacac cgccaactcc gcgcccggc
 25401 ccctcctcgg cccggacacc gccgtgctcg ccgacgccct cggcgaccgc
 25451 gccgtcgac gccacgcaac cctcgacgac ctctggccg gggacaccac
 25501 cccgcccgc accgtcctcg tccccctcg cgcgccactc gacggcgaca
 25551 ccgcgacgca cgcgacgcc ctacccgca gcgcgctgac cctcgtccag

10

20

30

25601 cagtggctcg ccaccgaccg cctcgccgac tcccgcctgg tcttcgtcac
 25651 ccacggagcc gtcgccaccg acgacgcgcc ccccaccgac ctggccgcgc
 25701 ccgcggctctg gggcctgacg cgctccgcgc agaccgagaa ccccggcacc
 25751 ttcacctccg tcgaacctga caccgagccc gactcgacca ccgcgctcag
 25801 ccgcgcctcg accctcgacg aaccacagct cctcctccgc gcgcggccgc
 25851 cccgcgcgcg ccgcctcacc cgcacccccc ccccaccac caccaccac
 25901 acgccgtggt ccgcggacgg aacgggtgtg gtgacgggtg gtacgggtgg
 25951 tctgggtggg ttggtggccc ggcatctggt gcggctgtgt ggggtgcggc
 26001 atttgttgtt gaccagtcgt tctggtgtgg gtgctgcggg tgcggccggg
 26051 ttggtcgcgg agttggagtc gttgggcgcg cgggttgttg ttgcggcgtg
 26101 tgatgtgggt gatggctcgg ctgttgccga gttggttgcc ggtgtgtcgg
 26151 agtcgtatcc gttgtctgcg gtggtgcatg cggtggtgt gtggtatgac
 26201 ggtgtggtgg gttcgttgac gccggagcgg ttggctgcgg tgttgctcc
 26251 gaaggtggat ggtgcgtgga acctgcatga ggcgacgcgt ggtctggatc
 26301 tggacgcgtt tgtgtcttc tcgtctgttg cgggtgtgtt cgggggtgcg
 26351 ggtcaggcca actatgcggc gggtaatgcg tttttggacg cgttgatggt
 26401 tcacgcgggt gctggtgggt tgcctggtgt gtcgttgccg tggggtgctt
 26451 gggatcaggg tgtgggatg acggcggggc tgacggagcg ggatgtccgt
 26501 cgtgctgctg agtcgggtat gccgttgttg acggttgatc aggggtgtggc
 26551 gttgttcgat gcggcgttg cgacggggag tgccgcgttg gtgcgggtcc
 26601 gtctggacct ggccgcactg cgcacccggg gcgacatcgc accgctcctc
 26651 cgcggcctcg tcgcgcacc gctgcgcgc accgcggcca ccggcctcgc
 26701 caccggcgcg gacacccggc tcgtccaacg gctcggccga ctcgaccacg
 26751 cccaacgcca cgaggcactg ctcgacatgg tcgcagcag ccgcgcgtc
 26801 gtccctcgcc acgcgcacgg caacgccatc gacgccgaac gcgccttccg
 26851 cgacctcggc ttcgactcgc tcaccgcggg cgaactccgc aaccgtctgc
 26901 gcaccgccac cggcctgcac ctgtcggcca ccatggtctt cgaccacccc
 26951 accctgtccg cctcgcggga gcacctgcgg gacgagttgt tcggcgcggg
 27001 cgagagcag gtgcgggtgc cggctcaggc actgccgcg accgcgcag
 27051 atcccatcgt ggtggtgggc atggcctgcc gtttccccg ttggtgtgac
 27101 tcgcccaggg acctgtggcg cctggtcgac gacggcaccg acgccatcac
 27151 caccctccc accaaccgcg gctgggacct ggacaacctc tacgaccgg
 27201 accccgagca cttcggcacg tcgtacaccc gctccggttg ctctctgcac
 27251 gaggggggg agttcgaccc ggcgttcttc ggaatgagcc cgcgtgaggc
 27301 gctggcaacc gactcccaac agcgtctcct gctggaatcc tcctgggagg
 27351 cgatcgagcg ggcggcatc gaccgcgtga cctgcgcgg cagcgcacc
 27401 ggcgtcttcg ccggcgtgat gtacagcgac tacgggagca tcctcggcgg

10

20

30

27451 caaggagttc gagggcttcc aaggccaggg aagtgcgggc agcgtggcct
 27501 cgggccgcgt ctctacgcc ctgggttcg agggcccggc cgtcacggtg
 27551 gacacggctt gctcttcctc tctggtcgcc ctgcactggg cggcgcaggc
 27601 gttgcgggcg ggggagtgct cgttggcgtt ggccggtggt gtgacggtga
 27651 tgtcgacgcc gagcacgttc gtggagtctt cgcggcagcg gggctctggcg
 27701 cctgatggtc gttccaaggc gttcgccgag gccgcggacg gcgtcggctg
 27751 gtccgagggc gtccgcatcc tcgtcctgga gcgccagtcg gacgcggtgc
 27801 gcaacggcca cgagatcctc gccgtgatcc gcggctcggc ggtcaaccag
 27851 gacggtgctg ccaacggcct gaccgcgcc aacggcccgt cccagcagcg
 27901 cgtcatccgt caggcgttgg ccagtggcgg cctgtccacg gccgacgtgg
 27951 acgccgtcga ggcgcacggc acgggtacga cgctcgtgta cccgatcgag
 28001 gcccaggcgc tcctggccac ctacggccgt gaccgcgacc ccgagaaccc
 28051 cctgtggctg ggctccctga agtccaacat cgggcacacc caggcagcgg
 28101 ccggtgtcgc cgggtgtcatc aagatggtca tggcgatgcg gcacggcgtg
 28151 ctgccgcaga ccttgcattg cgacgcgcgg tcctcgcacg tcgattggag
 28201 cgtcggcgcc gtcgaactgc tcaccgagca gaccgcctgg ccggagaccg
 28251 gccgggtccg tcgcgcgggt gtctcctcct tcggcatcag cggcaccaac
 28301 gcccacgtca tcgtggaaca gccggcgctc gtcgaaagcc cggccgcgga
 28351 gccgagcgga cgcgaaccgg gcgtcgttcc gctgccgctg tccggaaagt
 28401 cccccgaggc cctgcgcgac caggccgcac gcctgctggc cgggttggcg
 28451 gagcggcccc cgctgcgcgc gctcgacctc ggctactcgc tggcgacgac
 28501 ccgttcggcg tcgaccacc gggcgggtgt gctcgccacc gaccgcgcgg
 28551 atgcgggtccg cgcgctgacg gcgctcgccg ccgcgcagcg ggatctctcc
 28601 gccgtcgtcg gcgacacccg caccgggtcgt caccggtgc tgttctcggg
 28651 tcagggtcctg caacgcctgg gcatggggcg tgagttgtac gagcgtttcc
 28701 cggctcttcgc cgaggctctc gatgtcgcca tcgaccacct ggacgcgcgc
 28751 ttgcccgcgc aggccagtct gcgtgaggtg atgtggggcg acgatgtcga
 28801 gctgctggac gagacgggtt ggacgcagcc ggctctgttc gccgtcgagg
 28851 tggccctgtt ccggctggtg gagagttggg gtgtccgtcc ggacttcgtg
 28901 gccggtcatt ccacgggtga gatcgccggc gcgcattgct tcgggggtgtt
 28951 ctcgctggag gacgcctgcc gtctgggtgg ccgccgtgcg acgctgatgc
 29001 aggcgctgcc gaccggcggc gcgatgatcg cgatccaggc cgcgcaggac
 29051 gaagtcaccc agcacctgac tgacgacgtc tcgatcgccg ccgtcaacgg
 29101 cccgacctcc gtggtcgtct ccggggccga gagcgtgcc cgcacgggtg
 29151 ccgaccggct cgcggagaac ggccgcaaga cgaccggct gcgggtctcg
 29201 catgcttcc actcgcgtt gatggatcc atgctggcgg agttccgtgc
 29251 ggtggccgag ggctgtcct acgccacccc gacctcccc gtgctctcga

10

20

30

29301 acctgacggg ccggtctggcc acggccgatg acctctgtct gcccgagtac
 29351 tgggcgcgcc acgtccgcga ggccgtccgc ttccgcgacg gcgtcagcac
 29401 cctggagaac gagggcgtca ccacgttcct ggaactggga ccggacggcg
 29451 tgctgtccgc catggccag cagtcgtca ccggcgacgc cgcaccgtc
 29501 ccggccctcc gcaaggaccg cgcgaggag acgtccgcgc tcaccgcct
 29551 cgcacacctc cacacggcag gtctccgcgt cgactggcg gcgttcttcg
 29601 ccggcagcgg cgcacccgc gtggacctgc cgacctacgc cttccagcac
 29651 gccacctact ggcacaccgg caccctgccc accgcgcacg ccgcggccgt
 29701 cggcctcacc gccgcggagc acccgtgtct gaacggttcg gtcgaactcg
 29751 ccgaaggcga aggggtgttg ttccacggac ggctgtcact gcagtcacat
 29801 ccgtggctgg ccgaccacgc cgtcatggga caggctcctgc tgcgcggcac
 29851 cgcactgtcg gaactggcgt tccgggccgg cgacgaggcc ggttcgcacc
 29901 gcgtcgagga actgacgtc gccgcaccgc tcgtctgcc cgagcgcgt
 29951 gcggtacaga cccaggtccg ggtcggcgtc gccgacgaca ccggccgccg
 30001 taccgtcacc gtccactccc ggcccagca cgcgaccgac gtgtcgtgga
 30051 cccagcacgc gaccggcacc ctgaccatgg gctccgcccc ggccgacacc
 30101 ggtttcgacg cactgcctg gccgcccgcc gacgcgaac cctcgcacc
 30151 cgacgactgc tacgcgcgt tcacgacgt cggcttcgcc tacgggccgg
 30201 tcttcaggg cctgcgggcc gcctggcgcg ccggtgacgt gctgtacgcc
 30251 gaggtggccc tggcgagtc caccggcgac gaggcgaccg ccttcggtct
 30301 gacccccga ctgtcgacg ccgccctgca cgcctccctc gtgcgccacg
 30351 agggcgagga gagcaacggc ggactgcgt tctcctggga gggcgcgacc
 30401 ctctacgcga ccggcgccac cgcgctgcgc gtccggctga ccccgacggg
 30451 caccgacggc cgttcgggtg ccacgcctgt ggccgacacc gccggtcgtc
 30501 cggtcgcgc catcgacaac ctctgtctcg gccgggtctc cggcgaccag
 30551 ttgaccggcg ccgcgggact ggcccgcgac gccctgttca cctggactg
 30601 gaaccccgta ccggagaacc tcgtaccgga gaacctgta ccggagaaca
 30651 ccggcggggg ccacgcccag gaccaggacg gccggccgc cgcgccacc
 30701 gtgcgcgtgg tcggcgcgga cggcaccgc atcgccgcg acctgaccgc
 30751 cgcggcatc cacaccacc tccacccga cctcaccacc ctgcaccga
 30801 ccgacgcga cgttcgaag acggtcctca tccccctcac cggaacggga
 30851 accggaaccg gcaccgggac tgagtcgacg gacggaatcg ggacggggg
 30901 cgcgagtcg gacgcgtccg cccctcccc ggccgaggtc gccacaccc
 30951 tgteccacgc cgaactcgc ctctccagg agtggaccgc acaggagcgc
 31001 ttgcgcggt cccgcctggc gttcgtcacg accggggcga cggccgcgg
 31051 cggtaaccgac gtcattggac tggccgcgc cgcggtctgg ggcctggtcc
 31101 gatccgcca gtccgaagc ccggacacct tcgtctgat cgacctgac

10

20

30

31151 cccggcccg cgggcacgca cgaccgcaca gccgccgccc aacggggcca
 31201 actgctccta cgggcaactgc acaccgacga accgcagctc gccctgctg
 31251 acggcggcgt gctcgccgcc cgcctggccc gcttcgacac cgcggccgcg
 31301 ctcaccccg cggccgaccg ggccctggcg ctcgacagca cggccaaggg
 31351 cagcctcaac ggcctcgccc tgaccccgta tccggcgcca ctggcgccgc
 31401 tcacccggca cgaggtgagg gtcgaggtgc gtgcccggg cctgaacttc
 31451 cgtgacgtgc tcaacgcgtt ggggatgtat ccgggtgatg atgtcggatc
 31501 gttcggttcg gaggcgcccg gtgtggtcgt cgaggtcgga ccggaggtga
 31551 ccggcctggc ccccgccgac caggatcatg gcgatgatcac cggcagcttc
 31601 ggctcgctcg ccgtggacga cgcgcggcgc ctcgcccgcc tgcgcgagga
 31651 ctggctcctg gagacgggtg cgtcggtgcc gttggtgttc ctcaccgctg
 31701 actacgcctt gaaggagttg ggtggtctgc gggcggggga gaagggtgctg
 31751 gtgcatgccg gtgccggtgg tgcggtatg gcggcgatcc agatcgcccg
 31801 gcgatgtcgt gccgaggtgt tcgccacggc cagtgagggc aagtgggacg
 31851 tgttgcgctc gctcgcgctg gccgacgacc acatcgccct ctcgccacc
 31901 ctcgacttcg aggcggcctt cgcggaagtc gccggcgacc gcgccctgga
 31951 cgtcgctcct aactccctcg ccggtgactt cgtcgacgcc tcgatgcggt
 32001 tgctcgccga cggcgcccg ttcttgaga tgggcaagac cgacatccgc
 32051 gcccgcgact ccgttccga cggcctctcc taccagtcct tcgacctgc
 32101 ctgggtggtg ccggaacca tcggcaccat gctggccgag ctgatggacc
 32151 tcttcgcac cggcgcaactg cggccactcc cggctccgac ctgggacgtc
 32201 cggcacgcca aggacgcgtt ccgcttcagt agcatggcca agcacatcgg
 32251 caagatcgtg ctcacccctg cccgctcctg gaagcccgag ggaacgggtg
 32301 tggtagcggg tggtaggggt ggtctgggtg ggttgggtgc ccggcatctg
 32351 gtgcggtcgt gtgggtgctg gcatttggtt ttgaccagtc gttctggtg
 32401 gggtagctgc ggtgcggccg ggttggctgc ggagttggag tcgttgggcg
 32451 cgcgggttgt ggttcggcg tgtgatgtgg gtgatggctc ggctgttcg
 32501 gagttggtg ccggtgtgtc ggagtcgtat ccgttgctc cgggtggtgca
 32551 tgcggtggt gtgttgatg acggtgtggt gggttcgttg acgccggagc
 32601 ggttggctgc ggtgttcgt ccgaagggtg atggtgcgtg gaacctgcat
 32651 gaggcgacgc gtggtctgga tctggacgcg tttgttgtct tctcgtctgt
 32701 tgcgggtgtg ttccgggggtg cgggtcaggc caactatgag gcgggtaatg
 32751 cgttttttga cgcgttgatg gttcatcggg tggctggtgg gttgcctggt
 32801 gtgtcgttgg cgtgggtgct ttgggatcag ggtgtgggga tgacggcggg
 32851 gctgacggag cgggatgtcc gtcgtgctgc tgagtcgggt atgccgttgt
 32901 tgacgggtga tcagggtgtg gcgttgctc atgcggcgtt ggcgacgggg
 32951 agtgccgctg tggtagcgggt ccgtctggac ctggccgcac tgcgcacccg

10

20

30

33001 gggcgacatc gcaccgctcc tccgcggcct ggtcaaggcg cccatccgcc
 33051 gcgcagccgc caccacaccc ggcgacaccg gactcgccga gcagctcacc
 33101 cggctccagc gcgcgagcg acgggacacc ctctcgcgc tcgtccgcga
 33151 ccaggcccg atggtcctcg gccacacctc gggcgacggc gtcgaccgt
 33201 cccgcgcctt ccgcgacctc ggcttcgact cgctcaccgc ggtcgaactc
 33251 cgcaaccgca tcggcgggc caccggcctg cggctaccgg ccacggccgt
 33301 cttcgactac cccaccgccc atgccctcgc cgcacacctg ctaccgaac
 33351 tgctcgcccc cgacgcccag tcggaccccg acgagcccgg cgacccacc
 33401 gcgggaccga ccgacgaccc catcgtcac atcggcatga gctgccgctt
 33451 ccccgccgac atcggtcgc cggaggacct gtggcgctg ctccggcagc
 33501 gcgccgacgt cgtcaccgac ttcccgacca accgcggctg ggacctggac
 33551 aacctctacg accccgaccc cgcgcacgcc ggcacctgt acgcccgcac
 33601 cggcggtttc ctgcacgacg ccgccgactt cgacgccgac ttcttcggca
 33651 tgagcccccg cgaggccatg gccacggact cccagcagcg cctgctgtg
 33701 gagtcctcgt gggaggcgat cgagcgggccc ggcacgacc cgctgacctt
 33751 gcgcgacagc cgcaccggcg tcttcgccgg cgtcatgtac agcggctacg
 33801 gcaccgcctc cgacggcgcc gaattcgaag gcttcagggg gcagggcagc
 33851 gcactgagcg tggcctccgg ccgggtctcc tacaccttcg gcttcgaagg
 33901 cccggccatg acggtcgaca ccgcctgctc ctctcgcgtg gtcgccctgc
 33951 acctcgccgc acaggcactc cgcggcggtg agtgacacct cgccctcgcc
 34001 ggtggtgtca ccgtgatgtc catcccgac accttcacg agttctcccg
 34051 gcagcgcgga ctggcccccg acggccgctc caagccgttc tcgagtcgg
 34101 ccgacggcgt cggctggtcc gagggcgctg gaatgctgt cctggagcgc
 34151 cagtcggacg ccgtgcgcaa cggccaccag atcctggccg tggcgccgg
 34201 ctcgccggtc aaccaggacg gtgcgtccaa cggcctgacc gcgcccaacg
 34251 gcccgctcca gcagcgggtg atccgtcagg cgttgccag cggcgccctg
 34301 tccacggccg acgtggacgc cgtcgaggcg caccggcagc gcaccacgct
 34351 cggtgaccgg atcgaggccc aggcctcctt ggccacctac ggccgcgacc
 34401 gcgaccccca gaaccgcgtg ctgctcggtt cgatcaagtc caacctcgcc
 34451 cacaccagc cagccgcccg tgtcgccggc gtcacaaaga tggctcatggc
 34501 gatgcgccac ggcgtgctgc cccgcagcct gaacatcacc gagccgtcct
 34551 cgcacgtcga ttggagcgcc ggcgcgctg aactgctcac cgagcagacc
 34601 gcctggccgg agaccggccg ggcgcgctg gccggtatct cctccttcgg
 34651 catcagcggc accaacgccc acgtcactct ggagcagccg gagggccgac
 34701 ggcactcggc gccggaagaa gccgacacgg cggaggcagc cgccaaggcg
 34751 ccggccaccg cgcacctgcc cgtaatgccg tgggcaactgt ccggcaagac
 34801 gccggaggcc ctgcgtgccc aggcgcgacg cctcctcgcc cacctccagc

10

20

30

34851 agcgccccga actcgacccc gccgacatcg cctgtccct cgccaccag
 34901 cgctcccagt tcacccaccg ggcagtcgtc ctgagcaccg accgtgacga
 34951 ggcgaccgc gcgtgtccg cctcgccac caccgcgcg tccgaccct
 35001 cggccctcac cggcacggtc accatgggac gttgcgcggt gctgttctcg
 35051 ggtcagggct cgcaacgtct gggcatggg cgtgagttgt acgagcgttt
 35101 cccggtcttc gccgaggctc tcgatgtcgt gatcgatcac ctggacgccg
 35151 ccttgcgcgc ccaggccggt ttgcgtgagg tgatgtggg cgacgatgtc
 35201 gagttagctga acgagacggg ttggaccag ccgcgcctct tcgccatcga
 35251 ggtggcgctg tttcggttg tggagagttg ggggtgtccgt ccggacttcg
 35301 tggccggtca ttccatcggg gagatcgcg cggcgcatgt cgtcggggtg
 35351 ttctcgttg aggacgcgtg ccgtctggtg gccgcgcgg cgacgtgat
 35401 gcaggcggtg ccggccggtg gcgcgatgat cgcggtccag gcgaccgagg
 35451 acgaagtcac ccgcacctg accgacgagg tggcgatcgc ggccgtcaac
 35501 ggcccgcact ccgtggtgat ctggggcgca gaagaggcca cgagaccgt
 35551 ggcacaacac ttgcgcgacc aggggcgcgc gacgaccgc ctgcgggtct
 35601 cgcattgctt cactcgccg ctgatgatgc tggcgagtt ccgtgcggtg
 35651 gccgagggcc tgtctacgc caccgcgacc ctcccgtcg tctcgaatct
 35701 gacgggccag gtggccacgg ccgacgaact ctgctcgcc gactactggg
 35751 tgcgccacgt ccgtgaggcg gtccgcttcg ccgacggtgt gacggccctc
 35801 gaagccgagg gcgtgcggac cttcttgaa ctggcccg acggcgtct
 35851 cgcgcctatg gccaggaaa ccgtcgcca cgacacggtc accgtccccg
 35901 tcctccgag gaacatgcc gaggaacgga cctgtctac cgactcggc
 35951 cggtccaca ccaccggaac ccgatcgac tgggcgcgc tcctggcccc
 36001 gaccggcgcc cgccgggtg acctgccgac atacgcgttc caacaccgtc
 36051 cttctgccc ctccggcccc cgcgacacc cggatgccgc cgcgctcggc
 36101 atcgccggcg cgagccacc gctctcaac ggcactcgtc aactcgcca
 36151 cgaagagggc ctgttgttca ccgacggct gtcactgcag tcgcatccgt
 36201 ggctggccga ccacgccgtc atgggacagg tcctgctgcc cggcaccgca
 36251 ctgctggaac tcgcctcgc cgcggcgac gaggtcggt gtgaccacgt
 36301 cgagggaactg acgtcgccg caccgctcgt cctgcccag cgcggcgcg
 36351 tacagacca ggtccgggtc ggcgtcgcc aacaccagg gcgcgcacc
 36401 gtcacgatcc actcggtcc cgcacgcgc acgaccacc acagtgcac
 36451 ccacaccggc accgacacc cgtggacca acacgcaacc ggcgtcctc
 36501 tcgcggcct gccggcgac gcaaccgtc cgttcgatgc caccgtgtg
 36551 ccgcggcgcc acgccaacc cgttgacct gcggacttct acggtcccc
 36601 ggccggcgaa ggattcggt acggcccg tttccaggc ctgcgagcg
 36651 cctggcgcc cgacggcgag gtgttcgcc atgtcgact gccggaggcc

10

20

30

36701 ggccgtaccg aagccgaggc gtacgggctg catccggcac tgctcgacgc
 36751 cggactgcac gcagcctggc tcgtcgcccc ggacggggag cccacacgga
 36801 cgggcagcgt gccgttctcg tggcgcggcg tcttcctggc cgcttcgggt
 36851 gcctcctcgg tcgcgctccg actcgcccg gactccgacg gaacgctgag
 36901 cctggccatc gccgacacca ccggtgcacc ggtcgcgctc gtacaggccc
 36951 tctccatgcg caccgtctcg gtgacggccc tgagcgccac cgcgggcctc
 37001 gcccgcgacg cgctgttccg cctggactgg gcctcgggcc cggagccggc
 37051 gtgccagccg gacgacacgg tgaccgtgat cccggcggtc gcggtcgtcg
 37101 gtacggaaac ctccgaactc acctccgagc tcaccgcggc cctcgctgcg
 37151 gcgggcgccg acgtcgacgt ccgcacgacc ctgtcgaccg atgaaccgac
 37201 gcccgcgctg atcgcgctgc cgctcgtcgc ctccgaccag accggcaccg
 37251 cagaagcggc acccgctccc gggcgcgctg acgacctcac ccgccgagcc
 37301 ctggccctcg tacagacccg cctgcaagag cagcacttcg cggacacgaa
 37351 gttcgtcttc gtgaccctg gtgcgacggc cgggcgtgat gtggctgctg
 37401 ctgcggtgtg gggctctggt cgcttcggcg agtcggagaa tcggggtgtg
 37451 ttgctctggt tcgatctgga tccggtggt gcggtgggtg cggctgcgct
 37501 cgtcgctgcg ttggtcagtg gtgagccgca gcttgcggtg cgcggtgatg
 37551 tgttgcggtg cgcgctctg gtgcggcggc cgctcaccga ggtcggtgcg
 37601 ggtgctgatg gcaccgggga tggcgctcggg ggtggctctg gtgtgtcgtt
 37651 ctcggggtgag ggtcggttcc tggctactgg tggtaggggt ggtctgggtg
 37701 cgggtgttgc gccgtcatct gtggccgagt atggggtgcg ggatctgctg
 37751 ttggtcagtc gcagtggtga acgtgccgtg ggtgctgggg agttggtggc
 37801 ggagcttgcg ggtgtgggtg cgcgggtgcg ggtgggttgc tgtgatgtga
 37851 ccgatcgtgc cgcggtggtg gagttggtt gcgggcatgc ggtgtccgcg
 37901 gtggttcatg cggctggtgt gctggatgac ggcattggtg gtgcgttgac
 37951 cggggagcgg ttgtccgcg tgcctcgccc gaagggtgat gctgtctggc
 38001 atctgcatga ggcgacccgc ggcctggatc tggacgcgtt cgtcgtcttc
 38051 tcctctctcg ccggggtctt cggcagtcce ggccaggcca actacgcggc
 38101 cgcgaacgcc ttcttggaag cgctgatgac gcggcgccgg gcggagggac
 38151 tgcccgccct gtcactcgca tggggaccgt ggtcgctgac cgatggcaca
 38201 tcgggcatgc tcgcggacgc cgaggccgat cgcctgaccc gttcgggagt
 38251 gccaccgctg accgcggagc aaggactggc actgttcgac gcggccctgg
 38301 cgaccggtga cgcacactgc gtcccggctc gcctggacct gtcggcactg
 38351 cgtgcccagg gtgaggtgcc gcccttgcgt cggctccctga tccgaggccg
 38401 ctgcgcgcgc gccgccgcgc cggaatcggc aaccgccacc gggctgcggg
 38451 aacggctcgt cggactgaac ccggtcgagc gacaggaagt cctcctggac
 38501 ctcgtgcgcg gccagggtgc cctggtcctc ggccacgccg acgccgacga

10

20

30

38551 cgctcatccg gcacgtgcct tcaggaggtt gggcttcgac tccctcacct
 38601 cgggtggaact gcgcaaccgc ctcaacaccg tcaccggtct gcggctcccc
 38651 gcgaccatgg tggtcgacta tccgaccgtc gaagtgcctg tctcctacgt
 38701 cctggacgag ttgctgggca cggatgccga ggtggcgacc gtgcagccgg
 38751 ccgcgggttc ggtggcggac gatccgatcg tcactcgtgg catggcctgc
 38801 cgctaccccc gcggggtcgc ctctccggac gacctgtgga ggctcgtcac
 38851 ggacggcgct gacgcgggtgt ccccttccc gaccaaccgt ggttgggacg
 38901 tcgaatccct ctatcaccgc gaccccgacc atctcgttac ctctacacg
 38951 cgctcgggtg ggttcctgca tgaggcgggg gagttcgatc cggggttctt
 39001 cgggatgagt ccgcgggagg cgttgccgac cgattcccag cagcgggtgt
 39051 tgttgagtc gtctgggag gcgatcgagc gggccggtat tgatccggtg
 39101 agtttgccgg gtagtcggac ggggtgtgtc gcgggggtga gtacagcga
 39151 ttacagcgcg atgttggcga gtccggagtt cgagggttcc cagggcagtg
 39201 ggagttccgc gagtttgccg tcgggtcggg tggcctacac gttggggttg
 39251 gaaggcccgg cggtagcggg ggatacggcg tggtcgtcgt cgttgggtgg
 39301 gatgcactgg gcgatgcagg cgttgctgag tggtgagtg ggttggcgt
 39351 tggccggttg tgtgacggtg atgtcgacgc ctgcggtgtt tgtggacttt
 39401 gctcggcagc ggggtttgtc gccggatggt cggcgcaagg cgttgccgga
 39451 tgcggccgat ggtgtgggct ggtccgaggc cgtcggcgtg ttggtcctgg
 39501 agcggcagtc ggacgcgggt cgcaatggtc acgagatttt ggctgtggtg
 39551 cggggttcgg cggtcaccca ggatggtgcg tccaatggtt tgacggcgcc
 39601 caatggtccg tcgcagcagc ggggtgatcc gcaggcgttg gccagtgtg
 39651 gtctgacggc cggtagcgtg gatgtggtgg aggcgcagtg tacgggtacg
 39701 acgctcggtg atccgatcga ggcgcaggcg ttggtggcga cgtatggcg
 39751 ggatcgtgag cctgagcggc cgttgtgtgt gggttcggtg aagtcgaatc
 39801 tggggcatac gcaggctgct gcgggtgtgg cgggcgttat caagatggtg
 39851 ttggcgtatg ggcatggtgt ggtgccgcgg acgttgcatg tggatgcgcc
 39901 ttcttcgcat gtggactggt ccgaggggtg ggtggagctg ctcatgagc
 39951 aggcggcttg gccggagacg ggtcgggtgc ggcgggcggg tgtctcctcc
 40001 ttccgcatca gcggcaccac ccgcacgtg attctggagc gtccggaggc
 40051 cgcgcggcgt ccggtgatgg agacgaacac cgtggagccg tccacggtgc
 40101 cgtgggtgct ttccggcaag acaccggagg ccctgcgcgc ccaggcggcg
 40151 aaactgttgt cgtcgatcga ggaacgccc gagcttcgcc tcgtcgacgt
 40201 gggaaatgtc ctggtcaccg gccgctcgac ctttgaacac cgtgcggtgg
 40251 tactcggcgc cgaccgtgcc gacgccgcc gcgcattgtc ggcgattgct
 40301 gccgacgagg ccgatgctgc cgctgccacc ggcgcgctcg gcgcgggtcg
 40351 tcacgcggtg ctgttctcgg gtcagggtgc tcaacgtctg ggcatggggc

10

20

30

40401 gtgagttgta cgagcgtttc ccggtcttcg ccgaggctct cgatgtcgtg
40451 gtcgaccacc tggacgccgc cttgcccgcc caggccggtt tgcgtgaggt
40501 gatgtggggc gacgatgcgg agttgctgaa cgagacgggt tggacgcagc
40551 cggctctgtt cggcatcgag gtggcgctgt tccggctggt ggagagttgg
40601 ggtgtccgtc cggacttcgt ggccggctcat tccatcggtg agatcgccgc
40651 ggcgcgatgc gccggggtgt tctcgctgga ggacgcctgc cgcctggtgg
40701 ccgcccgtgc gacgctgatg caggcggttc cggccggcgg tgcgatgatc
40751 gcggtccagg cgaccgagga cgaagtcacc ccgcatctga ccgacgacgt
40801 agcgatcgcc gccatcaacg ggccgaacgc actggtcgtg tccgggtgtgg
40851 aggatgccgc cgtcgagatc ggggcgcgggt tcgcggccga ggggcgtcgc
40901 acgaccgcac tccatgtgtc gcatgcgttc cactcgccgt tgatggatcc
40951 gatgctggcg gagtccgtg tgggtggcga gggcctgtcc tatgctgctc
41001 cgtccctccc cgtcgtctcg aatctgacgg gccagggtgg caccggccgac
41051 gaactgtgct cggccgagta ctgggtgcgc cactgccgcg aggcggtccg
41101 cttcgccgac ggggtgacgg ccctcgaagc cgagggcgtg cggaccttc
41151 tggaaactcg ccggacggc gtccctgcgc ccatggcagg agcctcgctc
41201 accgaatcct ccctcgcggt accgctgctc cgtaaggacc ggccggagga
41251 accggcggca ctgcgcgcc tggcccagtt gcacatcgcc ggcgcgcgcg
41301 tcgattggcc cgtgctcttc gctggtgtgg gtgcggggcg ggtggagttg
41351 ccgacgtatg cgtccagcg tgggtggttc tggccggtt gtcgggttgg
41401 tgttgggtgt gatgtgggtg ctgtggggt tgggtctgcg gggcatccgt
41451 tgttgggtgc tgcggtggag ttggctgcgg gtgcgggggt ggtgttgacg
41501 ggtcgtctgt cgttgtctc gcatggttgg ttggctgatc atgcggtgat
41551 ggggcgggtg tttgttctg gtacggcggt gctggagatg gtgatgcgtg
41601 ctggtgatga ggtggggtgt ggtcgtgtt aggagctgac gttggcggcg
41651 ccgttggtgt tgcctgagcg tgggtggggt cgggttcagg ttgctgtgga
41701 tgctcctgat gtcgcggtc gtcgtggtgt ggggtgtat tcgtgtcctg
41751 atggtgtggg tcaggcgggt tggctgcagc atgctgtcgg tgtgttggcc
41801 tctggtgtgg ctgaccaggt cgggtgggtc ggtgacggtg gtgtgtggcc
41851 gccgcagggt gcggtgtcgg tggatgctga gggctgctac gagctgtttg
41901 cggatgctgg gttcggttat ggcccgtgt tccagggtt cgtgcggtg
41951 tggcgtcgcg gcgaggaact cttcgagag gtccctgt cggacgaggt
42001 tgctgagagc gctgatacgg cgaccggtt cgggttgac ccggcgttgc
42051 tggatgcctc gtcctatgcc tcgcttctct cctccctga aggtcaatcc
42101 gccgatggtg ggcctgcgtt gccgttcgcg tgggaggggt tttccctctt
42151 cgcctcgggt gcgacggctt tgcgctgcg gttggcgccg gcgggtgagc
42201 atgcggtgtc ggtgaccgcg gtggatccga ccggtgcgcc ggtgatttcc

10

20

30

42251 atcgacgcgc ttctgaccgc tgcctcacc ctcgatgagg tcaacgcac
 42301 tcacacccag ctgagcgatg cgctcttcgg cgtecaatgg accacgggcc
 42351 cgagcaccgc ggccgcccgc caccgcgcg tagccatcat cggaaaccgac
 42401 caccctggggc tgcgcgaagc gctcagcagt tcctctgctg gcgcgaccac
 42451 gaccacgacc gcggccgcgt acgagagcct tgacgcgcgt atcgcggcgg
 42501 ggcccgaaagt gtccgtccct gacgtcacac tcacggtct caccaccgag
 42551 gacgccatcg ctcagtacgt gaacgaccac gacgccacgg tggccggcca
 42601 aggcaccatc ggagccggcg cggcgccgt agacgcggct cgcgcctca
 42651 ccgcccagggc cctgcgcacg atccaggcat ggctggccga cgagcgtctg
 42701 gcggcaaggc gcctggtctt cgtgaccgc ggccggccg acgggcagga
 42751 tgcgcgcgcg gcggccgtcc agggcctggt gcgctccgcc cagaccgaga
 42801 acccgggcac ctccggcctc ctcgacctcg acgggaccga ggcgtcgacc
 42851 gcggtcctcg gcgaggtctt caccctcgac gaaccgcaac tgcttctgcg
 42901 cgacggacac ctgcacgccg ccaggctgac ccgcctggcg tcgcccgcg
 42951 acacggccgt gccacggag tggaaacgcg acggcacggg gctgatcacc
 43001 ggtggtaccg gcgggctcgg cgcgcagttc gcacggcacc tcgtcgacag
 43051 gtacggcgtc cgcaatctcc tgctcgtcag ccggcgccgg cccgatgccc
 43101 cgggaaccac ggagttggtc gccgagctga cggcgcacgg tgcgaggtg
 43151 gccgtgcagg catgtgacgt ggccgatggc gatgcggtgg cggcggtggt
 43201 cgcgggctg ccgatgagc acccgtcag ggcggtcgtc cacacggccc
 43251 gtgtgttga cgacggagt atcggctcgc tcaccgagga gcggctcgcc
 43301 accgtcctgc ggcccaaggc ggatgcggcc tggcatctgc acgaagcgac
 43351 ccgcggcctg gacctggacg cgttcgtcgt ctctcctcc gtcgcagggg
 43401 tcttcggcgg cgcgggccag gccaaactac ccgcggccaa cgccttctg
 43451 gatgcgctga tggcccagcg ccgggcagcg ggcctgcccg gactctcgtt
 43501 ggcctggggg ccgtgggacc agaccggcgg aatgacgggc atgctgtcgg
 43551 acgcgagggc cgaccgcctc gccgcctcgg gcaccccgcc gctctccgcg
 43601 gagcagggcc tcgcccctct cgacgcggca ctcgctcttg ccggaaccag
 43651 cagccagac agggcagccg gcagcgcgc cgcagcacg tcgggaaccg
 43701 gcgacacgat cgcctcccg gccgcggccc tcgtcgacc ggtccggtc
 43751 gacctggcgg cgtggccgc gcagggtgag gtcccgcga ttctgcgcg
 43801 actggtgcgg acccgcacc gccgtacggc ggccggcggg tccgtcaccg
 43851 tggccggact cgtcaaccgc ctgtccgggc tcaccgccga cgacggcgc
 43901 caggaaactc tcgaactggt ccgactcag gcagccctc tcctcgggca
 43951 cgcgatccg gcgtccgtg actccaccgc acagttccgt gacctcggct
 44001 tcgactcgt gaccgcgtc gagctgcga accggtgag cagggccacc
 44051 ggcctgcgcc tgaccgcaac cctggtcttc gactaccgga acaccgatgc

10

20

30

44101 cctcgcggag cacctgcggg acgagtgtt cggcgcggtc gagagcgagg
 44151 tgcgggtgcc ggtccaggca ctgccgccga ccgccgacga tcccatcgtg
 44201 gtggtgggca tggcctgccg tttccccggt ggtgtgacct cgcgcgagga
 44251 cctgtggcgc ctggtcgacg ccggcaccga cgccatcacc accttcccga
 44301 ccaaccgcgg ctgggacctc gaatcgctct acgaccggga ccccgcacac
 44351 ctccggcacct cctacacccg ctccggtggc ttcttgacag aggcggggga
 44401 gttcgacccg gcgttcttcg gaatgagccc gcgtgaggcg ctggcaaccg
 44451 actcccaaca gcgtctcctg ctggaatcct cctgggaggc gatcgagcgg
 44501 gccggcatcg acccgctgac cctgcgcggc agcgccaccg gcgtcttcgc
 44551 cggcgtgatg tacagcgact acgggagcat cctcggcggc aaggagtctg
 44601 agggcttcca aggccaggga agtgccggca gcgtggctc gggccgcgtc
 44651 tcctacaccc tcggcttcga aggtcccgcc gtcaccgtgg ataccgcctg
 44701 ctctctctcc ctggctgccc tgcacctggc agcccaggcc cttcgggccc
 44751 gtgagtgcac gctcgcgtc gccggtggg tgacggtgat gtccacgcca
 44801 ggcacgttcg tggagtctc gcggcagcgg ggtctggcgc ctgatggtcg
 44851 ttccaaggcg ttccgccagg ccgcggacgg cgtcggctgg tccgaggcgg
 44901 tcggcatcct cgtcttgag ccgcagtcgg acgccgtgcg caacggccac
 44951 gagatcctcg ccgtgatccg cggctcggcg gtcaaccagg accgtgcgtc
 45001 caacggcctg accgcgccc acggcccgtc ccagcagcgc gtcacccgtc
 45051 aggcgttggc cagtggcggc ctgtccacgg ccgacgtgga cgcgctcgag
 45101 gcgcacggca cgggtacgac gctcggtagc ccgatcgagg cccaggcgtc
 45151 cctggccacc tacggccgcg accgcgaccc cgagaaccgg ctgctgctcg
 45201 gttcgatcaa gtccaacctc ggccacaccc aggcagcggc cgggtgtcgc
 45251 ggtgtcatca agatggtcat ggcgatggcg caggcgtgc tgccgcagac
 45301 cctgcatgtc gacgcgccgt cctcgacagt cgattggagc gtcggcgccc
 45351 tcgaactgct caccgagcag accgtctggc cggagaccgg ccgggtccgt
 45401 cgcgcgggtg tctctctctt cggcatcagc ggcaccaacg cccacgtcat
 45451 cctggaacag ccggaggccg tgcagcgctt ggcaccggga gcagcagaga
 45501 ccgtggagcc ggtcgccatc aagccgtcgg cggaaaccgt cctggtgccc
 45551 tgggcgctgt ccggcaagtc acccgaggcc ctgcgcgccc aggcgcgacg
 45601 cctcgtgac ttcttgccg aacggcccga accgcgctcg atcgacatcg
 45651 gccactcact ggccgtcaca cgctcgcagt tcgaccaccg cgcgatcgtg
 45701 ctggctgacg atgcgaaggc cccggcggac agcctggccg ccctcgcggc
 45751 cctggcctcc ggtgtggcgg atccgcgcgt cgtctccgac gcggtatcga
 45801 ccggcggttc ggcagtgtg ttacaggctc aggggtgctca acgtctgggc
 45851 atggggcgtg agttgtacgg ccgtttcccg gtcttcgccc aggtctcga
 45901 tgctcgtggtc gaccacctg acgccgcctt gcccgcccag gcgggtttgc

10

20

30

45951 gtgaggtgat gtggggcgac gatgtcagat tgetgaacga gacgggttgg
 46001 acccagcccc cgctcttcgc cgtcgaggtg gcgctgttcc ggctggtgga
 46051 gcgttggggt gtccgtccgg acttcgtggc cggtcattcc atcgggtgaga
 46101 tcgcggcggc gcatgtcgcc ggggtgttct cgctggagga cgctgcccgt
 46151 ctggtggccg cccgtgcgac gctgatgcag gcgctgccga ccggcgccgc
 46201 gatgatcgcg gtccaggcca ccgaggacga agtgaccccc cacctgaccg
 46251 acgaggtggc gatcgcgccc gtcaacggcc cgacctccgt ggtgatctcg
 46301 ggccgcagaag agggccacga gaccgtggca caacacttcg ccgaccaggg
 46351 gcgcgggacg accgcgctgc gggctctcgca tgcgttcac tcgcgcgtga
 46401 tggatccgat gctggcggag ttccgtgcgg tggcggaagg actgtcctac
 46451 gccaccccgt cccctcccgt cgtctcgaat ctgacgggct ggctggccac
 46501 ggccgacgaa ctgtgctcgg ccgagtaactg ggtgcgccac gtccgcgagg
 46551 cggtcgcgtt cgccgacggc atcaccaccc tcgaagccga gggcgtgcgg
 46601 accttccctg aactcggccc ggacggcatc ctgtccgcgc tggctcagca
 46651 gtccctcgcc ggccaagccg tcaccgtgcc cgtcctgcgc aaggaccgcg
 46701 gtgaggagtc cacggccctg acggcccagc cgcattctca caccgcgga
 46751 ctgatcgaag actggcagga cttcttcgct ggtgtgggtg cggggcgggt
 46801 ggagttgccc acgtatgcgt tccagcgtgg gtggttctgg ccggttggtc
 46851 gggttggtgt tgggtggtgat gtgggtgctg tggggcttgg gtctgcgggg
 46901 catccgttgt tgggtgctgc ggtggagttg gctgcgggtg cgggggtggt
 46951 gttgacgggt cgtctgtcgt tgtcgtcgca tgggtggtt gctgatcatg
 47001 cggtgatggg gcgggtgttt gctcctggta cggcgttctt ggagatggtg
 47051 atgcgtgctg gtgatgaggt ggggtgtggt cgtgttgagg aactgacgtt
 47101 ggccggcgcc ttggtgttgc ctgagcgtgg tggggtgcgg gttcaggtt
 47151 ctgtggatgc tcctgatgct gcgggtcgtc gtggtgtggg ggtgtattcg
 47201 tgtcctgatg gtgtgggtca ggcggtgtgg tcgcagcatg ctgtcggtgt
 47251 gttggcctct ggtgcggctg accaggtcgg tgggttcggg gacggtggtg
 47301 tgtggccgcc gcagggtgcg gtgtcggtgg atgctgaggg ctgctacgag
 47351 ctgtttgcgg atgctgggtt cggttatggc ccggtgttcc aggggttgcg
 47401 tgcggtgtgg cgtcgcggcg aggaactctt cgcagaggtc gccctgtcgg
 47451 acgaggttgc tgagagcgt gatacggcga ccggtttcgg gttgcacccg
 47501 gcgttgctgg atgcctcgt ccatgcctcg cttctctcct cccttgaagg
 47551 tcaatccgcc gatgggtggc ctgcgttgcc gttcgcgtgg gagggtgttt
 47601 cctctctcgc ctcggtgcg acggttttgc gcgtgcggtt ggccgcggcg
 47651 ggtgagcatg cgggtgcggt gaccgcgggt gatccgaccg gtgcgccggt
 47701 gatttccatc gacgcgcttc gtaccgctcg cctcaccctc gatgaggtca
 47751 acgcatccca caccagctg agcgatgcgc tcttcggcgt ccaatggacc

10

20

30

47801 acggtcccca gcaccccgcc cgcgcaccac ccgtcggtag ccatcatcgg
 47851 aaccgacccc ttccggcctcg cagacggcct ttccgacgcc ttgcccctgg
 47901 tcgaggagcg cggtagacctc gcggcgctcg cagcgtcggg gcacccggta
 47951 ccggacacctc tcctcgtccc ggtagcgggc acccggcgca caggcgtacc
 48001 tgcggacgcc gaaggacaca ccgacgcggg gacatccgac atgctccgat
 48051 ccgtgcgtga ggccaccgca caggtagctgg agcagatcca gcagtgggtg
 48101 gcggacgacc gggttcgaggc ggcgcgctcg gtgttcgtga cgcgcggggc
 48151 gggttcctcg ggtagggggc gcacgcgccg cctggcggcc tcggccgtct
 48201 ggggtctggt gcggtcggcg cagtcggaga atccgggctg cttcggctct
 48251 ctccgacctc acctcgacct cgccttgac tccgacctg ccccgagggt
 48301 cgacatcgag cgcgacgtg accgcgatcc ggtcgggtgg accgtgcagc
 48351 ccgcgctcgc cgcggccctg cagcgcaccg ccgacgagcc gcagtggga
 48401 ctgcgcggcg ggaccgtgca ggccgcccga ctgacccgaa tccccgcgcc
 48451 gcagaccgac cgtgccgaga ccgacctgc cgagaccgac cgtccggaga
 48501 tcgacacccg gcggcccgcc accgtgctca tcaccggtgg taccggtggc
 48551 ctccgtgggt tgcctgcccg gcacctcgtc gccgagcggg gggtaccgag
 48601 cctggtgctc gccagccgga gcggtctcgc ggccgaggga gcggagaagc
 48651 tggtcgccga cctcgaagcg ctccgtgccg tggtagccgt gcagacgtgt
 48701 gatgtggccg atggcgatgc ggtggcgcg ttggtcgccg gcgtgtcggg
 48751 cgagtacccg ctgacggcgg tgcctcacac ggccggtgtg ttggacgagc
 48801 gagtgatcgg ctgcctcacc gaggagcggc tcgccacctg cctgcggccc
 48851 aaggcggatg ccgcctggca tctgcacgag gcgaccgag atctggacct
 48901 ggaacgcttc gtcgtcttct cctccctcgc cggcgctctc ggtggcgccg
 48951 gtcaggccaa ctacgcggcg gcgaacacgt tcctggagcg cttgatggcg
 49001 cagcgtcgcg ccgcggggtt gccgggtgtg tcgctggcgt ggggtccgtg
 49051 ggaccggggc ggccggcatga cggggaccct gtcggacgcc gaggcgacc
 49101 gcctcgcgcc ctccggtgtt ccgccgatct cggcgagca ggccttgcg
 49151 ctgtacgacg cggcgaccgc cggtagcgg ccgctggtgg tgccggtgag
 49201 gctggacctc gccgcgctcc gcgggctcgg tgatgtcccg gcgctgctgc
 49251 gcggactggt ccggaacccc gcgcggcgga ccgcggcgcc cggtagcgcg
 49301 ccgtcggcgg atgtgctcac ccggcagttg gccgggctcg gcggggcgga
 49351 gcaggaggag gtccgtgctga ggctggtgag cggtcaggcc gcggtggtgc
 49401 tcgggcacgc cgacggctcg gcgatcgggt cggggcgaca gttccaggag
 49451 ttgggcttcg actcgtgac cgcggtggag ttccgcaacc gactcaacgc
 49501 ggccaccgga ctgcggtgc cggccacct gctgttcgac taccgacgc
 49551 cggccgacgt cgtcggggac ctgcgcggcc ggctcggcac cggggagggtg
 49601 tcgggtgagg gctcgggtgt ggcgcgctg gacaacctg aggcggtgat

10

20

30

49651 cgccgggctg tccctcgacg acgcggggga gcaccagttg gtggccggcc
 49701 ggctggaggt cctcagggcg aagtgggcgg acatgcgaag cgcggaggga
 49751 gctgtggacg gcggtgcgga cgtcgacatc gaggaggcgt cggacgacga
 49801 catgttcgcg ctgtcggacg acgagctggg gctgaactga gccgctccgc
 49851 atgagcagtt ccgcaccaga agttccacag tggcctggtc cacagtgacc
 49901 tgggccacag tgagctgttc cacagcgacc tgttccgcgg cggttccgta
 49951 gtgagccgtt ccgcttaggc gggaaaaccgg gggcccggtg gtcggacgct
 50001 tgattccggc caccgggagc gtcaaaccgg cctcttctaa agaaagggaa
 50051 agaactaggg aaacacccac agcccgatca tgtgcaatga attccgtggc
 50101 tcggaaaacc attcgcgggc catggagtta cgggtgtgac acgtgtccgt
 50151 acggcagtgga cggcgtgcac gggttgccac ggcaggtccc cggccgaggc
 50201 gtctccccgc tgctgcccgg cgggtgcgtgc cgttgggttc gtccctagga
 50251 ggtcccccca tgaccacgtc caccgaggag agcctgtggg ccggtgctt
 50301 ccatccagca cggcccgccc ccgtccggct cttctgttcc ccacatgcgg
 50351 gcggctcggc ctcttcttac ttcccggtgt cggcccaact gtctcgggt
 50401 gccgaggtgt tcgccatcca gtacccgggg cgccaggacc ggcgcaagga
 50451 agccgggtgc agtgacctcg cgaccttggc cgaccaggtc tacgacgcgc
 50501 tcgccccctt gctgaaggag cggccgagca cgttcttcgg gcacagcatg
 50551 ggcgcgacgc tggccttcga ggtggcccgg cgttcgagg ccgacgacgg
 50601 tgacctggtc cggctgttcg cctccgggcg ccggggcccc tcccgcgtgc
 50651 gtgaagaggc cgtgcaccgg cggtcgacg acggcatcgt cgaggagctg
 50701 aagctgctcg ccggcaccaa caccgcgctg ctgggcgacg aggagatcct
 50751 gcggatgacg ctgcccgcga tccgcagcga ctaccaggcc atcgagacct
 50801 accgctgccc gcccgacgtc accgtccggg cgccgctgac cgtcctcacc
 50851 ggcgaccgcg acccgaagac ctccctggac gaggccgagg cgtggcgcg
 50901 ccacaccacc ggggaacttcg acctcaagggt gcttccgggt gggcaacttc
 50951 tcgtcagctc cgaggccccg gcgatcatcg atctgctccg ggcgcacctc
 51001 gccggcaacg gctagcgggc gactgcggc aggcgggcgg tgccggtctg
 51051 ccgcacggcc cccgcgcgcg ctgagacggc accatgccac ccggcgacgc
 51101 gcgttgctg cgcctgtgga gcaccgtccc gcgtgcgtac gcgggcgcgc
 51151 ggaggacgtg ctcgccggc cggatgtcct ggaggagacc attcgaaaaa
 51201 cgtcaaagtgt ctgacagctg ctgctcattt actccagccc gcactatatg
 51251 atcttgtggt ggggcatgta cggagcgtgc cactcgttcg ctccgtgtgc
 51301 atccatcccg accgtccgga tttgtgtgca ccggacggcc gtggtcctgg
 51351 tacgcgtcgg cgcgaggtcg gctcctgttc tcacccacg tggaggaacg
 51401 cgtgatgcgg aagcagtcgg gttcttccgg cttactgacc acgctggtgg
 51451 gacgcgacga cgaactgcgc accctcgccc ggcacgcgcg ggccgcccgc

10

20

30

51501 gacggggcggg ccggcctggt cctgctccac gggcccgcgg gcatgggcaa
 51551 gaccagtctg ctgcggtcct tcacggcgag cgatgtctgc cgcggcatga
 51601 cgggtgctgta cggcacctgc ggcgagaccg tcgccggcgc cgggtacggc
 51651 ggtgtgcgcg aactcctcgg cgggctcggc ctgagcggcg gcgacgcccg
 51701 gcgctcggcc ctctggagg gcctggcggc ccgcgcgtg ccgcactca
 51751 ccgcagaccc cgcgggtccc gacgcgcga cgggtgccta ccgggtgctg
 51801 cacggcctgt actggctcgc cgcgcgcctc atggcccaac ggccgctggt
 51851 cctcgtcttc gacgacgtcc actggtgcga cgaacgtcc ctggcctgga
 51901 tcgacttctt gctgcgacgc gccgaggacc tgccgctgct ggtcgtcctg
 51951 gcctggcgca gcgaggccga accggtcgcg ccgcgggtgc tcgcgacat
 52001 cgcgcgccag cgcgcgccca ccgtgctcgg cctgcacccg ctcgccccg
 52051 acgacatagg cgaatggtg cgtcgcgtct tcgggaccac ggccgcaccg
 52101 tcgttcgtca gccgggtcgc cgcggtgtcc ggccgcaatc cgtcggccct
 52151 cgcgcgcctc ctgcagcaac tcgcgcga gggcgtccgg ccggaacgcg
 52201 ccggggagcg ccgggcgcgc gaggtcgga gtcacgtcct cgcgcgtcgc
 52251 gtgcgctgcc tgttgagcg ccggccgcc tgggtgcgcg gcgtggcccg
 52301 tgccatcgcc gtactcggcc cggagtgcac cgagttgctg gcggcgctcg
 52351 ccggcgctcc gccgcgcacc gtcgacgagg cctgttgggt gctgcgcagg
 52401 gccggcatcc tggcgcga ccgcgtggac ttcgtccatg acgtcgtccg
 52451 ctccgcggtg ctgcagacg tcgcgccgc caccctggcc gaactgcga
 52501 ccaacgcgc gctgttgctg agcgcgcgc gccgcgcctc cgaggagctc
 52551 gccggccagc tcagtctgct gccggtgctc gaccagccgt ggatggccgc
 52601 ggtgctgcgc gacgcgcgc ccagggcgga gagccgcggc gccccggagg
 52651 ccgggtgctg ctgcctctac cgggtgttgg aggtggagcc ggacaacgtt
 52701 gccgtccgca tcagatggc ccgcgcgtc gccgagatca accgcgccga
 52751 ggcgatgcgc ctctcaagg aagcgtctc cctcgcggg gacgtccgca
 52801 ccgcgcacca ggtcgcgtc cagtacggct tcacctgcct cgcgctgcag
 52851 gaatccccgt ccggggtgct gatgctggag gacgcgtcgc ccgagctgac
 52901 cgcgaactg gcccccgaac caggggccgt ggaccgggag ttgcggaccc
 52951 tcgtggagtc cgtgctgctc atcgtcggg ccgacgagaa ggtgacgatc
 53001 ggtgcggtcc gcgaccgggc ggcccggctc accatgccgc ccggcgacac
 53051 accggcccag cggcagatgc tggccatgac caccgtgctg accgcgatgg
 53101 acggccggga cgcgcggtc gccgtcgacc aggccgcgc cgcctcgcgc
 53151 gcacccggcg tcgagctgga accctggctc ctgctgtccg cctccttcgc
 53201 cctctccctg gccgacgagg tcgcgcgc gcagtacgca ctggacctca
 53251 tgctccagta cggccaggac aacgcggcg tgtggacgta cgtcctggcg
 53301 ctctccacgc gcgccctgct ccaccacggg gtgggcgcct tccccgaggc

10

20

30

53351 cctggccgac gccagaccg ccgtcgagat cctcggcgag gagcgctggg
 53401 cggacggcgc cgtgctgccc cgtgtcgcg tggtccaccgc cctcgtcgac
 53451 cgcggcgagc ccgagcgcg cgaacacgtc ctcgacggca tcacccgccc
 53501 ccgcctcgaa cgcttcgtca tcgaatacca ctggtaacct caggcccgcg
 53551 cctacgcccc ctgggtccgc ggggatttcc aaggagccct ggacctcctt
 53601 ctgcctcgcg gtcggtcctt ggaggagtcg cgcttcagca acccggcgtt
 53651 cgtgccgtgg tgggcccgatg gcgcggtgct cctggcgacc ctggaccgcc
 53701 acgaccaggc gcgcgaactc gccgcatacg gaagcgagtt ggccgagcgt
 53751 tgggggacgg cgcgcggcct cggactggcc ttcattggcc agggcgtcgc
 53801 cgcacccggc cgcgcgggca tcgatcacct caccgaggcg gtctcgctgc
 53851 tcgcggactc cccggccccg gccatggagg cccgggcca acttcttctc
 53901 ggacacgccc acctgaagcg cgcgcacctg cgggccccgc gggaacacct
 53951 gcgcgcccc gccgacctcg cccagcgctg cggcgccgtg aagctcggcg
 54001 tcgacgccag aaaactgctg gtcaccgcgg gtggtcgggt acgcaggatg
 54051 accgcctccc cactcgacat gctgaccggg atggaacgca cggtgccgga
 54101 cctcgcggtg accggcgcg gcaaccgggc catcgcgga gccctcttcg
 54151 tgactgtaag gaccatcgaa acccatctca cgagcgtcta ccggaagctc
 54201 ggggtcggcg ggcgtgcgga gctgtccgcc gtccctggaga ccaggaccgc
 54251 cacctccggt cggcagccgc cggcctgggt ctcccaggca cgcggacgcg
 54301 cttgagacga acaggacgag aggtagcggg gccccggagc aaggcgcgga
 54351 accaaccgac caccctgcag ccgcagtcgc cacccgacgc ccacggcgac
 54401 cccaccatgc tcctcgaatg cggccgggaa cagcggtcga tcggcgacct
 54451 cctgcaccgc ctcgccaggg ggccggccatc ggtgctcagc ctgaccggcc
 54501 ggccccggca cgcgcgaac gccctggctc gctggggcg gtgccgggcc
 54551 aggcacgacg ggctgcgct cctgcgagcc caggcgacgc ccgcggaacg
 54601 ggaactccgc tacggcgccg ttctccaact gctggcgctc ctcgacggcc
 54651 cgcacggcag caccctggac gccgcgatcc gccacgacgg tccccgccca
 54701 ctgcccgtgc ccggcatcga ggaggtgctg cggcgcaacc gcacggcacc
 54751 caccctggtc gtggtcgaag acgtccagtg gttggacccg gcctcgctga
 54801 cgtggttgca gatcctgctg cgcacacctg ggccggacac ccgcctcgcc
 54851 gtccctggca gcagctgcgg tgacaccacg gccttcgaca ccgacccgaa
 54901 ggcccccgcc gtccccgggc cgcgggacac cgtgcccgtc gcgcgcttcg
 54951 tggtgcccgc gctcaccgac cgcggggctg ccgccaccgt ccgcgcccgc
 55001 tgcggcacc ccggcgacga ggagttcatc gccgcgctca cctccgccac
 55051 cgcgggcaac cccgccatcc tgcgggacgc cctgcgcgcc ttcgtcgacc
 55101 acggcctccc cgcgcgacgc gaccacctcc cggagctgca cgcctcacc
 55151 gctggcgctg tcggcgacca caccgtgcgc gccctggacg gcctgcccgc

10

20

30

55201 cgaagtcac ggcgtcctgc gggccctggc cgtctgcggc gacctgctcg
 55251 acttccaccg agtccggggc ctgcgccggc cgcactcgct gtccgaggac
 55301 cggatccgca cctgctggc gagcgtcggc ctgaccgtgt ccgtcggcga
 55351 caagggtgac atccgcttcc ccgcctccaa ggcacgcgtc atcgaggaca
 55401 tgcccgcgcg ggagcgcgcg gatctgtacg tccgcgcggc cgaactcacc
 55451 cacagttagc ggcgtcaacga cgaggacgtc gcccatctgc tgctgcgctc
 55501 gtcgcgcgctc ggcgcaacct gggtcgtgac cctgctccgc cgcggattcg
 55551 ccgcgcgcgtc ggcgcgggag gaccaccacc gggcctgtgc ctgcctctcc
 55601 cgcgcctcgc aggaacccct cgacccccgg gaacgcagcc tgctgacgtt
 55651 ggaactggcc ggcggccgaag ccgtcgcggc gccggaggcg ggggatcgac
 55701 gcctggggga actcgtccgc agcacgctc cggacaccga cccacgctc
 55751 tccggtgagg ggggtggggg ccgcgccatc gacctggggg tgcgccggg
 55801 caacagcgaa tgggtccgcg gcaccgcggg cgaggccctg ccgtacgcg
 55851 ggccggccga ccgggaggaa ctggtcgcgc tgttctgggt ggccgcgctg
 55901 cgggacgacg acgcgcgat gatccccgtg gtgccccggt tgcccgaccg
 55951 gcgggtgccc ccggcccagg ccggcgcccc tgccctggcag ctggccacgg
 56001 cgggggagga cgcggacaag gccagggaagc tcgcccggat cgcctcacc
 56051 ggcgggggtga acgagagcct gatgatgccc aaactggcgg cctgcgccgc
 56101 gctgttcgcc accgacgaca acgacgaggc ggtgcacggc ctggacacca
 56151 tgctcaccgc cgcgcgcagt gccacacctg gcagcatggc cgcgcgatt
 56201 ttcaacctac gggcccgat acacctgtgc gcggccccggc tggaggccgc
 56251 cgaacgcgat ctggacagcg ccgagcgcgc cctgccgcg acgagttggc
 56301 acccccgctc gctgccccaa ctgatcgcca cccgcacct cgtcagcatg
 56351 gagacgggccc gcccgagccg ggcccgccga ctgcgcgagg ccccggtccc
 56401 cgccggcggc gaggagggtg tgtggtggc cgccctgctg ctgcgccgcg
 56451 cccgtgtggc cgcgcgacg ggtgactggg aggaggccct gcggtgtcg
 56501 cgggagtgcg ggcgctggct ttttcgcgg cactgggcca acccgccat
 56551 gctcagttgg cgcgcgctgg ccgcgcaggc gtgtctgaag ctcggcgacg
 56601 tgacggaggc gcgcggctg cgggacgagg agctgttctt cgcgaccgc
 56651 tggggcaccg cgagcgccc cgggatcgcc cgctgacga cgcggcgact
 56701 cttcgacgac gacggcgacc gggccgtccg gcggatccgc gaggccgccc
 56751 ccctgctccg cgcctgccc gccgcctgg cctacctgtg gagccggctg
 56801 agccaggccg gtgcccagac gggccacggc gacaccgccc cggccgacg
 56851 ctcttgccag gcggtcgccc ggatgaccgc cgccacccc gccagccgc
 56901 tcgccaccgc cgcgcgacc ctgaccgtcc cgtccgttcc ggtcgccacc
 56951 gcgcgcacca ccgcgctcgt cccaccgga tggcgcgacc tgtrcgaggc
 57001 ggagaaggac accgtgctgc tcgccgcccg cggccacggc aaccgccaga

10

20

30

57051 tcgccgaaca actcgccgtc agcaggcgca ccgtggagct ccggctgagc
 57101 aacgcctacc gcaagctgag gatcggcgga cgcaaggagc tgtacctgct
 57151 cctggaggcg ctggaaggac cggtcgcgga tgtttcttga gcgggagAAC
 57201 gaactggccc ggatccgggc cgcctcgac gccgcggaag cggcgactc
 57251 ctgcctcctc ctgatcaacg gtccctcggc cagcgggctg tcggcgctgc
 57301 tgcgccggat accggagctg gccggcgacg gcacccgctg cctcggggac
 57351 agcgcgcctt ggcggaacg cgacttcccc ttccggatcg cccgccact
 57401 ctctgaccac ctgctgtccg gggcgggcgg cgcaggggcg gccgaacgca
 57451 ccgcgggggc agagcacttc agccgactga tggacaccgg cgaccgccct
 57501 accgggacgg gccccgccct ggaggtctcc caggcagtgc tccaggggcg
 57551 ccaggcgctg ctcgccgacg cgtccgcgga gcggcgctg ctgacctg
 57601 tcgacgacct ccagtgggac gacggccctg cctgcgctg gctggccac
 57651 ctaccccggc ggctgcacgg cctgcgggcg ctgctggtgt gcacgctggc
 57701 cgacggcgac caccggggca ggtacccctt ggtccgggag gtcccgcg
 57751 ccgcgcacac cgtcctgcgc ctggcgccgc tgtcccgga gccacccgc
 57801 gtctctgctg ccgggccccca gggccggcgg ccgcaggacg cactggtg
 57851 cgcctgttac gaggcgctca ggggcaaccc gctgttctct accgccttcc
 57901 ggagcgctct gcgcgccacc ggaaggccgc ccggcgcgga ccacttcggc
 57951 gccgtccggg agctgagccc gacggtgctg cgcgacggc tcgcgggcca
 58001 tctgcggatc cagccgcagc cggtcgcgga ggtcgcggtg gcggtggccg
 58051 cgttggcgga ccacagcgat ccggtgctgc tcgccagct cgcgggggtc
 58101 gatgaaatcg gtttcgccgg tgcgcgcgc gcgctggtgg acgcccgcct
 58151 gttggcccg ggacgggacg tccgcttctt ccacggcgtc gtcccgatg
 58201 cgggtggactc cctgctcacc ctcgacgagc gggaacgctc gcacgacgac
 58251 gccgcgacg tgcgtgacgg ctgcgggccc ccggccgagc aggtcgccgg
 58301 ccatctgctg gccgtgggtc acccgggcgg gccctggtcg gaggcggtcc
 58351 tgcgctccgc ggcccaaca gcgctgcgc ccggccggcc cgcgacgagc
 58401 gcccggtacc tgcgcgcgc cctgctgac caccgcaccc aggaaggctg
 58451 ccgcgcccgc atcttggctg atctggccac cgcgagcgc gccctcgacc
 58501 ccgatgctg tgtacgccac gtcagccagg cggtcgcgct gctggacacc
 58551 tcgcgggacg gggcgccgc cgtgttgcc atcccgccgt cctgctgc
 58601 cgcggccgag ccgtccgccc tcgagttggt gcggcaggcc gccgcggggc
 58651 tcgacgaacc gggcgagcgg gacgaggagg gagccgacga actcgccctg
 58701 cgcttggagg cgtggctgag gactccggc cagcagaacc cgtcgagct
 58751 ggcgtcctcg gtggcgcggc tgcggcgcat gggggcacgg ccgcgggtg
 58801 acagcgtcgc cgaacgcgag ttggtcgcg tgcgttgag tcggggcgcg
 58851 ctacgaggcc ggctcagcgc cgcggagatc gccacaccg gcaaccgcat

10

20

30

58901 cctcgaacgt gagccggcca ccgcccacca tgcccacacc ccgctgccgc
 58951 tggatgatgct ctcgctgttc gtcgccgagt ccgtgcaggg cgtggcctcc
 59001 tggctggcca gcgaacagca caccggcgcc cggtacgca ccggcgcgga
 59051 cgatgtgctg ctgaccgccc agcgggccc tctcctggtg acgcagggcc
 59101 gcccgccgccc cgcgccggag cacgtcgaac gcgccctggt catggacgcc
 59151 ggcgactggt cggaaacccc cgtcatgatg ttccgcccgg tgccttcga
 59201 actgcgcgac ccggccttga gcgaacgcat cctggaacgg atccgcgacc
 59251 gccggccggc cggactggcg ctcaccgcca ccggtcagat gctccaggcc
 59301 gccgtcgacg tgcacttcgg ccggggggcg gacgccctgg acacgctgct
 59351 ggcctgtggc cgacgcctgg agaccgtggg atggcgcaat tccgactgct
 59401 tgccctggcg tccgtatgca atcgggctgc accagcggct cggcgagacc
 59451 gatgccgccc tgcaactcgc cgaggacgag ctgagggtgg ccggggagtg
 59501 ggggtgcgac acgaacctgg gccgggccc gcgcctgaag ggctggctgt
 59551 tgcaggacga gggactggat ctgctgcgtg agagcgtcga gatcctccgc
 59601 gcttcgtcct acgcgacgga actggcgccc accctcgtcg tctcggggcg
 59651 tcggctgccc ggtggggccc agggcgaggc ggtgttcgcg gagggcgccc
 59701 ggatcgccgc cgcctgtggt gttccctggc tcgccgaacg ggccgaactc
 59751 ggtctgggca gtgccatcgt gccgccggtc gccaccctga ccccagcga
 59801 gcggcgagtg gcgtcgctgg tgagccgggg cctgaccaac caggccatcg
 59851 cgaccgaact cgggtgtgagc tcccggggcg tggagaagca cctcaccagc
 59901 gcctatcgca agctcggcgt ctcgggccc cgcgagttgg tgaatgcgct
 59951 cccgggtcgt tgacgcccgc ggccgtccc ttccggacga tcacaatcct
 60001 ctgtgcctgt cagccctgtg ggtgcggcgt cgttgtcgta cgcgcgtccg
 60051 gtcattgacag caatcctcaa caaaagctca aacggaagcg ggggcagggtg
 60101 catccacccc atcaccaatg tctcgtatgt tacttttggc ccaccgtagt
 60151 tgggactctg cggaccgtac ctcgtaagt acggaccacg gatttagggg
 60201 tgcgtagatg tgcgcggtg acgtgccgag ttgtggctaa ctacgctccg
 60251 ttccacggaa atcgaaactc ttccggctga atcgcgcgga acttcgatcg
 60301 aattctggat tccgtttctc agaaggctaa aagacgacgg gggttatccg
 60351 aaaggatggc ggtctcgtga ctatcactca cctcactgac tcgaatcacc
 60401 gtacgaccgg cggggtgata tccgctcaga ccgcaccggc cggggagtcc
 60451 gtcggccccc ggctgatggc gtccctcgac cgtgacctca ccataagca
 60501 cgccaaccag gagtccgcc gccgcttcga cgattccgcc ggtgacgtct
 60551 gcggctcgag ctccggggac ctgatgcacc cgagcgtgca gcagccgctg
 60601 atgcgccagt tctccggcgt gatcgagggc aagcggcacc gcttcgcctc
 60651 gcacgtggtc gccgtgggag cccaggacgc ggccctcggc ggcaccctca
 60701 ccgcctccgc ggtcacccgc aagaccccgc acatcgccgg gatcctggtg

10

20

30

60751 ctcatggact cgtccggcgc cgcgcagccc gccgatgcgc gcgtcgtgac
 60801 cagccagaag aagttcctga ccgagatcga tgcgcgcac ctcgaaggca
 60851 tcgcggccgg gctgtcgacc attccgctcg cctcgcggct ctacctcagc
 60901 cgccagggcg tggagtaacca cgtcaccggg ctgctgcgga agttgcgggt
 60951 gcccaaccgc gccgcgctcg tctcgcgcgc ctactccatg ggcatcctga
 61001 acgtcggcac ctggcccccg aaggctcgtg acgacttcat caagtgcgc
 61051 ccggccggcc tgtcccgcac cggccggcca ccgtcggacc gctgtcccgc
 61101 acgcccgtgt cgcgcagccc cgctctcac gggcgcatgc ccggcggcac
 61151 gggcgtttcg tgcctcgtgc gatgtttatc cattctcagc ttacgaacct
 61201 gcaatgctca tgcgcgcgat tccggccact tttcgttcgc tcgtaaccac
 61251 actgcccgtc atatcgtttc cactgaccgt ttctgtggtc cgcggtgtgc
 61301 cacctgggccc ccactccgac gcgcccgga cgggtcgcga gcggcctccg
 61351 aatactcacg gtaacccccg ttctcctcgtg tggagttccc cgcgcgtcac
 61401 attcctccgt gcctcgtgcg gcgtggctcg ccgctttcc ccgccgcgc
 61451 gacagccatt gttacaacgg cagcacagag cgggggtggt cgaacctcc
 61501 gcggtgcggc gcagtcaagc ggaacctctg ctcgatgttc gtgccacccc
 61551 gaactctgtg ttcttggtga cgttttcac caatgccccg ccggttcggt
 61601 gggaatccg ccgccggttc ctggccacaa cggatcgctt cggccgcca
 61651 ccacggccgg cgcagccgag gcttgttgga atcaccccca ccagtcggg
 61701 gcgccgttgt ggcccgga cgcaggagga acacctgga tgcagagggc
 61751 cgccgcaggg acatgctgga gctgatccgg cgcagcggat ccgccgatgt
 61801 cgtgcgtctc gccgaggagt tcgccgtcag caaggaaacg gtccgcggg
 61851 acctgaacgt cctggagggc catgggctga tacggcggcg gcacggtggc
 61901 gcctatccca tgggtcgggc gggctccgag gccgtcttcg tctcccggac
 61951 cgcacagccg atcccgagg agtcccgat cgcacccgcg gccgccgaac
 62001 tcctcagcga ggccgagacg gtcttcacgc acgagggtt caccgccaa
 62051 ctcatgcgcg acgcctgcc gcgcgaccgg ccgctgacca tcgtcaccgc
 62101 ctgcctcccg gtggtcagcg ccttcgcgac gagccacag gccaacgtgc
 62151 tgctcctggg cggccgggtc cggcgggga cgacggccac cgtcgaccac
 62201 tgggcgttcc acatgctgtc cggcttcgtc atcgacctgg ccttctcgg
 62251 cgcggagggg atctcgcgca ggtacggcct gaccacccc gacccggcg
 62301 tcgccgaggt caaggcccag gccatccgcg tcgcgcgcg cccggctctc
 62351 gccggggtgc acaccaagtt cggcacggcg agcttctgcc ggttcggaga
 62401 ggtgggcgac ctggagacga tcgtcaccgg cggcgccctg ccgctcgcg
 62451 agggccaccg ctaccacctc atgggcccc aagttttacg ggtgtgacgc
 62501 cgccggggcg tcccgcgcgc cggccacggc ctggcgcgcg ccggtgcggg
 62551 tcagccgcgt cgcgcgtccg tcggttccc cagagtgcg agcgggtgtca

10

20

30

62601 cccccacccg gtagccctcc tcgtcgagga gatggccgcc gtgcttcatg
 62651 tgccagaggg cgaccgcgcg ttccgtgacg gtcagggtccg cggcccgttc
 62701 ggcgagtcgg acctcgaagc gctgggcgtc gtcgaggag cgcaccagc
 62751 ccaccaggtg caggttggtg cggctgatga cgctcgcga caggcggacc
 62801 tcgcgcaccc cggtgacccg ggggtgcacc tcgcgcagcc tcgcggccgg
 62851 gacctgcccc cagaagctca ccgtgaccgg ccaactcggac agcgggcggg
 62901 cgacctcgca ccgggcgtgc aacatgtccg cggcgaagag ccgttgacg
 62951 cgacggcgca cgggtgtcggg gccggcgccg cactgtcgg ccagcgcgcg
 63001 gtaggtggcc cggccgtcca cggacagcgc ggtgaccagc tgttggtcga
 63051 gctcgtcgag gacgaaggcg ggggtgtccg tcgggtgtct ggacgcgtcg
 63101 gcggcgagcc gggcgacctg gtgccggccg agggcccgca ggcgccaccg
 63151 gctgccctcg gtgtgcaccg ggcgggccag gtgcgtccgg gcggcccgga
 63201 cgccgtccaa cgcggcgagg tcgtgggtca cccaacggga gagcatggcc
 63251 ggatcgccgg ccatgacgtt gagttggagg tcacggtccc cggtgacgtg
 63301 ggagagcgcc accacgtgcg gcgcggcgcc caacgcccgc gcgacgtcga
 63351 gcagtcggcc gggagcgcag tcgatctcga cgaacgccag gcacccctgg
 63401 cccgacgcgg ccagcaccgg ggcgggatgg cagctgatcc aggcggcgcc
 63451 ggtctcgacc agccggttcc agccctggc caccgtcacg gcgtccagtc
 63501 cgaggacgga gccgatccgg gtccagctgg cccgcggcgt gatctgcagc
 63551 gcgtgcacca gggcctgac cacatggtcc agggaccggg gggctctgcc
 63601 ggaatcctgc gccacgggcc acctccttgc gtgtttccgg cggattcggg
 63651 ccgccggtcg gctcaacctt cagcctggac tcgggtacgg ccggaccgta
 63701 ccaggcaacc cccggagcaa caggagtggg attcggcatg tcggcagtgg
 63751 aacgcgcaa ggaatgcaa cccgagttgg agcggctgcg ccggtcgcgtg
 63801 caccgggagc cggagttggg ccttgcgctg ccgcgcaccc aggagaaggt
 63851 gctcgcggcg ctcgacgggc tgccgttggg gatcaccccg ggcaccggcc
 63901 tcacctcggg gaccgccgtc ctgcgcggcg gccgccccgg cggcgcagtg
 63951 ctgctgcgcg gcgacatgga cgcgtgccg gtcaccgagg agagcgggtg
 64001 gccgtacgcc tcggagatcc ccggccggat gcacgcctgc ggacacgacc
 64051 tgacacaccg cggactcgtc ggcgcggcca ggctgttggc cgaacgccgg
 64101 gaacgcctgc acggcgacgt ggtgttcatg ttccagcccg gcgaggaggg
 64151 ccacaacggg gccggcgcgga tgatcgagga gggcgtgctg gacgcggcgg
 64201 gcaagccgct cgacgcggcg tacgccctgc acgtggcgtc caaccaggtg
 64251 cccggcggcg tggteatcac ccgtccggc accatcacct cggcctccga
 64301 cctgctgacg gtgacgggtg ggggcgaggg cggccacggc tcgacgcccg
 64351 ccacggccaa ggaccgggtc ccggcggcct gcgagatggc caccgcgggtg
 64401 cagaactggg tgaccgcgc cttcgacatc ttcgaccgg tcgtggtcac

10

20

30

64451 cgctcgccacc ttccacgccc ggacgaaggc gagtgtcatc cccgacaacc
 64501 gccgagttcc aggccacggt gcgcagttat tccgaatccg cccgggaccg
 64551 gctcgccgaa gggaaactgcc ccgtctggtg cgggccatcg cggccgcca
 64601 cggcctggaa gcggaggtcg actacctccg gcaactatccg gtcaccgtca
 64651 acgacacgga cgagaccgcc ttcgcggtgg ccaccgcccg cgacctgttc
 64701 gcgcccggcg aggtgtggga gtcgccgata cccaccaacg gctcggagga
 64751 cttcgcgtag gtctctccgc gcgtccccgg cgcgatgctg ctggtgggcg
 64801 ccgccccgga ggggagcgac tggcagcacg cgccgatgaa ccaactcgccg
 64851 cgcgcgctct tcgacgaccg ggtgctgtac cggcaggcgg cgtgctgac
 64901 cgaactcgcc gcgcgcccgc tcgcgccac cgcccgccg gagcccgccg
 64951 tggcgggatg acaccgccgc cgcacgcgtg gggcccgcac gcgtgcgcg
 65001 tacgggacct ggccggtctc ctccccgct acttgccgcg ggggcaactc
 65051 ttccacgca agtggttagg ggtgtgatg ctgccgtcgg tggagtcatg
 65101 gccatgtagc tgggtgacct ggggtcagcc ccgctgcctg

10

SEO ID NO. 2DNA sequence: Nys2

1 gaattcgat cgatgctgct cgtggtgtgt gccgcgggag gtgcgcagcg
 51 tgccgtgctt gaatcggggc gtggtcgccg tcgcgatcgc ggtggtgtcg
 101 gtgctgctcg tctgaccgt ggtggtggcg accgcctact tgcgtagc
 151 ccggtccttg cgcgcgacc tgggcagcgg atgcacctct tcggtcgacg
 201 ccaaggggag gtggcgagg tcgagatagg cgcggacgga gacgcgggac
 251 tgcggcgccg tctgccaggc cgcggagggg ccgctggaac gccacagacg
 301 gcccagctcg gccgcggcaa ggggtggcgt ggcttcgctg gcggcggtga
 351 cgtcgatgac gaccatgccg ggttcgctga catgttcctc gtggatatgc
 401 acagttctgt atgacctatt tcgccccgtg gcgtaaggag tagccggcag
 451 gtttcatccg aagggtgggc gcgggagcgg cgagttgacc gtgaggtgcg
 501 tgccgtgttc ccggtcgccg cggggggtgg gggccggcag gtggaccgcc
 551 gtggtcagca gggcgggggt gctgtgccag cggccgggga agaccgtgac
 601 ccggcgaccg cccagcaggg ggcccgggac caggagtcgg gcggtgaagg
 651 cgcctgtgtt cgggtcgatg acgatctcgg cctcgagaaa gtccagttcg
 701 cgctgggtca ggggatacca cgcctgaag acgctctcct tggcgtgaa
 751 gagcagccgg tcccagtgga cgtccggacg atgggcccgc agggccacga
 801 ggtgcggccg ttccgagggc agggcgatgg cgttcaggac gccggccggc
 851 agcggaccgt tcggttcggc gtcgatgctc acccgggccg acagctccgc

20

30

901 gggggagacg gcggcgccgc ggtagccggc gcagtgcgtc atgctgccga
 951 cgatgcccgg cggccactgc ggggcgccgc gccgattggg cagtatggcc
 1001 cggtcggggt ggccgagccg gcgcagggcc cggcgggcga gatggcggac
 1051 ggtggtgaac tcgcgtgcc gggaactccac ggcccgggcg atgacctcgc
 1101 gttccgagga gagcagccgg tcgccggggc gcggccggtc gtcgtacgcc
 1151 gcctcgggtg cgaccgtggc gggcaggatc agttcgatca ccgcatacct
 1201 ccggcgaaac gtaagaacag ggggttgctt cggggcacgg ttccgtcctt
 1251 gacgggggcg cgtggggcgc cgggtcagcc ggccgcggcg cccgcgggtg
 1301 gagtgtgggt gcgggctgcg tgcagccggg cgtacaggcc ctgcgcgcac
 1351 aggagttgat cgtgggtgcc ctgctcgacg atgggcccgg cgtccatcac
 1401 gacgatcagg tccgcgtcgc ggatcgtgga caggcgggtc gcgatcacga
 1451 agctggtccg gcccgcccgc agggagttca tggcgcggtg gatcaggacc
 1501 tcggtccggg tgtccacgga gctggtggcc tcgtccagca cgaggacggc
 1551 cgggtcggcg aggaaggccc gggccacggc cagcagctgc ttctcgccgg
 1601 cgctgacggg gcccgactcg tcgtccagca cgggtgtcgt gccctgcggc
 1651 aggggtgcga tgaagcggtc ggcacaggtc gcgcggggcg cctcctcgat
 1701 gtccgcacgg caggcaccgg gtgcgcgta cgcgatgttc tccgcgatgg
 1751 tgccgccgaa cagccagggt tcctggagca ccagcccga gggggaccgc
 1801 aggtcgtcgc ggggtcatcgt cgcggtgtcg gtgccgtcca ggaggatgcg
 1851 gccggagtcg ggttcgtaga agcgcacag gaggttgccg aggggtggtt
 1901 tgccggcgcc ggtggggccg acgatcgcca cgggtgctgc cggttccacg
 1951 gtcagcgaga ggttcctgat gagggcgctg tcgggggagt agcggaagga
 2001 cacgtcggtg aactcgacgc ggccctcggc gcgggcgggc gtgccgggcc
 2051 ggagcgggtc cggggcctgc tcggggggcg cgagcagggt gaagacgcgc
 2101 tgggcggagg cgatgccgga ctggagccgg cccgccaccg aggcgatctc
 2151 cacgatcggc tggctgaact ggcgggcgta gaggatgaac gcctgcacgt
 2201 caccgagggc cagggtgccg tttatgacct tcaggcgcc gatgacggcc
 2251 accagcacat agccgaggtt ggcgacgaac atcatgaccg gttccatggc
 2301 accggaggcg aactgcgcct tggccgcagc ccggtagacc gcgtcgttgc
 2351 aggcgtcgaa gcgctcctcg gccgccgcgc gccggtcgaa gcccttgatc
 2401 agcgcacgac cgggtgcacac ctctccaca tgggcggtga gggtgccgtt
 2451 cgcggaccac tgcgcggcgt agtggggctg cgcgcgcttg ctgatccggg
 2501 ccgcgatcaa cgcgagacc ggcacgctga gcagcatcac caccggccagc
 2551 gacggcgaga tcaccagcat cagcaccagc atcgtcaaca gcgagaagat
 2601 cgagggtgatc agctcggcga gggctctgct gagggctctg tggaggttgt
 2651 cgatgtcgtt ggtggtgcgg ctgagcagct caccggccgg ctgccggtcg
 2701 aagtggcgca gcggcagccg ggtcagcttc tcccggcggt cgcggcgacg

10

20

30

2751 ttctgtggatg gtgcgccaca ccgcggacgc caccagccgg ccctgcgcca
 2801 gcatgaacag cgacgccacg acgtagagcg ccagcaggac cagcagcagc
 2851 cggccgatcg cggcgaagtc gatcccgggc gccggggccg ggacgcgcgc
 2901 gagcacgcgc tcggcgatca gatcggtgac ccggccgagc agcagcgggc
 2951 cgaacgcggt gagcacgac ccgccgacgc ccacgacac ggcgagtgc
 3001 acggagcggc ggtgcggacg cagcaggccg acgagccgac gtgccggccg
 3051 gggcgcggtc cgctcctcct caaggctcgc cggcgaggcc atggcgcggt
 3101 tcctcctcgc tcagctgcga gagcgcgac tcgcggtagg tcgggctggt
 3151 gcgcagcagc acgtcgtggg tgccctgggc gaccacgcgc ccgcgggtcca
 3201 gcaccacgat ccggtccgcg tcgcggccgg cggagatccg ctgcgccacg
 3251 gtgatcaccg tggcgcccg gcgtgtacgc accagcgcgc tcgcagcgc
 3301 cgcatcggtg gcctggtcga gcgcgagaa acagtcgtcg aagagataga
 3351 tctccggccg gcgcgcaac gcccgggcga gggacaggcg ttggcgctgg
 3401 ccgcgggaga cattgccgcg gccctgggtg atctccgcgt cgaggccgtc
 3451 cggcatccgc gccacgaagt cggccgcctg ggcgacccgc agcgcctccc
 3501 acagctcctc gtcggtggcg tcggccgcgc cgaagcgagc attgctcgcc
 3551 acggtgcgcg agaacaggta cggccgctgc ggcaacgaacc cgacggcggc
 3601 ggcgagcgtg gcccggtca gctgcggac gtcggtgccg ccgacccgca
 3651 ccgcgcctc ggtggcgtcg gccagccgca gcaccaagtt caacagggtc
 3701 gtcttgccgc tgccggtgct gccgagcagc gcgatccgct cgcgggctc
 3751 gacggtcagg tcgacgtccc gcagcacggg ctctcggcg ccggggtagc
 3801 ggtaccgggc cgcgcacagt tcgatccggc cggcgggccc gcgcaccggc
 3851 tcggcgcgcg ccggcggcgc cagctcgac ccggtgtcca ggacctccgc
 3901 gatccggccg gcacagaccc gggcccgcg caccgacagg aacacgaagg
 3951 cgagcatcac gacggacatc aggatcagc agagatagct caggaggcg
 4001 ctgagcgagc cgatcggcat ccggcccgcg tcgatccggt gggagccggt
 4051 ccacagcagg gctacggtga aaccgttcat cagcagcagc acgaccggca
 4101 gcatcgtcgc gatcagccga cccaccgcgc gcgacaccac gaggaacgcg
 4151 tcgttggtct gcgcgaaccg cgcgcgctcg tggctgctcg ggacgaagga
 4201 ccggaccacc cgcaccccg tgatcgctc gcgcagcagc cgcgccagcc
 4251 ggtccagggt cagctgcac cgcgcgtaca gggtgcccat ccgggcccagc
 4301 agcaggccga agcagaccgc caccaccagc accagcgcca ccagcagcag
 4351 tgccagcggc acgtcctggc gcagcgccag cagcacgctg ccagggcaca
 4401 tcagcggcgc gcagacgacg atgccgaagc cggctctggc gaggttctgc
 4451 acctgctgca cgtcgttcac cgaccgggtc agcagggagg gggtgccgaa
 4501 ccggccgatc tcgcggggcg agaagtccag gatgcggcgg aagagcgcg
 4551 accgcagatc gcggcccatc gccgtcgcgg tcggggcggc cagcgcgcc

10

20

30

4601 gcaccgagcg cggcgggcgat ctgcaccagc gccaccacgc ccatcaccac
 4651 acccagctcg gtgatgcgcc cgccgtcgcc gcgcaccacc ccctggtcga
 4701 taagtgcggc gccacgcgtc ggcagcagca aagtgccag gatctggacg
 4751 agttgaaggc cgacgagagc ggcgggtggc caggcgtagg ggcgcagctg
 4801 tgcccgaga agtctcaaca gcacggagga acaccccggt tgacggcggc
 4851 gtgcggccgc ggcgtggacg gagtcggccg gcccgccgac ccggtcattg
 4901 caccgcagtc cgctccgaaa ttctactagt gttgggggtg gcaacggact
 4951 tttagcggcc cggtaactcg gaattcccta gaaagcccg gatgcgttga
 5001 cagcatcttc cgcggcttgc gagcgtgcgg gtgtctatcg tccggactgc
 5051 gtattcgacg ccaggtcgtg gccgagctca agcgtttgga cgtctcttt
 5101 cgtgttcgag aagggtttgc catgtccaaa cgagcgtga tcaccggaat
 5151 caccggccag gacggctcct atctcgcgga gcacctgctg tcccagggct
 5201 accagggtgtg gggctctgac cgcggccagg ccaatccccg caagtcccgg
 5251 gtcagccgcc tcgcctccga actcgacttc atcgacgggg acctgatgga
 5301 ccagggcagc ctggtctccg ccgtcgacac cgtgcagccc gacgaggtct
 5351 acaacctcgg cgccatctcg ttctgtccga tgtcctggca gcaggccgag
 5401 ctggtcaccg aggtcaacgg catgggcgtg ctgcgcatgc tggaagccat
 5451 ccgcatggtc agcggactgt ccacctccg cacggtcagc ccgcgcggcc
 5501 agatccgctt ctaccaggcg tccagctcgg agatgttcgg caaggccgcc
 5551 gagacgccgc agcgcgagac caccctcttc caccgcgca gccctacgg
 5601 cgcggcaaaag gcgtacgggc actacatcac ccgcaactac cgcgagtcct
 5651 tcggcatgta cgcggtctcc ggcgtgctct tcaaccacga atccccgcgc
 5701 cgcggccagg aattcgtcac ccgcaagatc agcctggcgg tcgcccgcct
 5751 caagcaaggc ctccaggaca agctggcact cggcaacctc gacgcggtgc
 5801 gcgactgggg ctatgccggc gactacgtcc gcgccatgca cctgatgtct
 5851 cagcaggacg ccggcgacga ctacgtcatc ggcaccgggc agatgcactc
 5901 ggtgcgcgac ggcgttcgga tcgcgttcga acacgtcggc ctgaactggg
 5951 aggactacgt cgtcatcgac cccgacctgg tgcggcccgc cgaggctcag
 6001 gtgctgtgcg ccgacagcgc caaggcccag gaccgcctcg gctggaagcc
 6051 ggacgtcgac tccccaccc tcatgcgcat gatggtcgat tccgacctgg
 6101 cgcagggttc ccgcgaaaac caatacggcg acgtgctgct cgcgcgcaac
 6151 tggtagcagt tctcaagctt tcgaaaacta gtgaattcct gccggaatc
 6201 cgacgacact cgccacatgg attcccccg tgagtggcga atccagggtg
 6251 cgaatccgaa cgtaccgccg aaacggcgtg gagaagtcgg acgccattca
 6301 cgtgcggccg tcggctgtcg agatgggttg agttgagatg gacaacgaac
 6351 aaaaactccg ggattacctc aagcttgca cggccgacct tcgacgcacc
 6401 cggcggcgcg tccacaagct ggagtcggcg gccaggaac cggtgcccat

10

20

30

6451 catcggcatg acctgtcgct accccggcgg cgtccgcagc cccgaagacc
 6501 tctggcgcat ggtcgaggcc ggcgagcacg gcgtcacccc gttccccacc
 6551 gaccgcgggtt gggacctgga ggcgctggcc gccgcgccga ccgcctccgg
 6601 cggattcctg cacgacgcac ccgacttcga cgcggacttc ttcggcattct
 6651 cgccgcgcga ggcggtcgcc atggaccgcg aacagcgcggt cgtcctggaa
 6701 tccgcctggg aggcgttcga acgcgcgggc atcgaccgga cgtccgtgaa
 6751 gggcagccgc accggagtct tcctcggcgc gatggcccag gactaccggg
 6801 tcggccccgc cgacggcgcc gagggcttcc aactcacggg caacaccggc
 6851 agcgtgctgt ccggcgcgat ctctacacc ttcggcacgg tcggccccgc
 6901 cgtcaccgtc gacacgcct gctcctctc cctcgtcgcc gtccacctcg
 6951 ccaccaggcc gctgcgggcc ggcgagtga ccctcgccct cgccggcgcc
 7001 gtcaccatca tgtccggccc cggcaccttc atcgaaatgg gccgccaggg
 7051 cgggctctcc gccgacggcc gctgccgctc cttcgggcag accgccgacg
 7101 gcaccggctg ggcggaaggc gtcggcatcc tcgtcctgga acggtgtcc
 7151 gacgccgtcc gcaacggcca cgagatctc gccgtcgtcc gcggcaccgc
 7201 cgtcaaccag gacggcgcc ccaacggcct gaccgcccc aacggccct
 7251 cccagcagca ggtcatccag caggccctgg tcaacgcccg actcgccgcc
 7301 ggggacatcg acgtcgtcga ggcgcacggc accggcacca ccctcggcga
 7351 ccccgctcag gcccgagccc tgctcgccac ctacgggcag aaccgcccgg
 7401 cggaccggcc gctgctgctg ggctcggtca agtccaacct cagccacacc
 7451 caggccgccg ccggcgctgc cggcgtgatc aagatggtea tggcgatgcg
 7501 gcacggcacc ctgccgcga ccctgcacgc cgaggagccc acccaccacg
 7551 tcgactggtc gcaggcgcc gtgcggctgc tgaccgacac caccgactgg
 7601 ccgcccaccg gggcgccgcg ccgcgccgcc gtctcctctc tcggcatcag
 7651 cggcaccaac gccacacca tcctcgagca ggccccgaa ccgcagcccg
 7701 aggacggcgc gaccgcgag gacgacgcg ccggcagcac gccggccacc
 7751 gccccgtag tgccggcgt cgtaccggtc ctgctctccg gccgcacccc
 7801 ggacgccctg cgcggccagg ccgcggccct gcgcgccgcc ctcgacaccg
 7851 gccggcggcc cgacctgctc gacctcgac actccctcgc caccaccgcg
 7901 gccgggttcg agcacgcgc cgtcctctc gccaccgacc accccgcctc
 7951 gaccgacggc ctaccgccc tcgcccagc cgacgacccg gccgcggccc
 8001 ccgcctggat caccggcacc acccgggccg agaccggct cgcgctctg
 8051 ttcaccggcc agggcgccca acgcctcggc gcgggacggg aactcgccgc
 8101 ccgtttcccg gcgctcgcca ccgcctcga cgcggcgctc gacgccttca
 8151 ccccgacact cgaccgccc ctgcgcgagg tcctgtgggg caccgacgcc
 8201 gccctgctcg accgcaccgc ataccccag ccggccctct tcgccgtcga
 8251 agtggcgctc taccggctga tcgaatcgtt cggcgctccg cccgaccacc

10

20

30

8301 tcgccggcca ctccgtcggc gagatcgtcg ccgcgcacct cgcgggggtc
 8351 ctctccctgg ccgacgcgc caccctcgtc gccgcccgcg gtgcctgat
 8401 gcaggcgctg cccgacggcg gggcgatgat cgcggtccag gcgtcggaag
 8451 ccgacgtcgc cccgtgctc gccgggcacg aggaccaggt cgcgatcgcc
 8501 gccgtcaacg gccctccgc cgtcgtcctg tccggcgccg aagccaccgt
 8551 caccgcgctc gccgaacagc tcgccgccga cggccgcaag acccgccggc
 8601 tgccggtctc gcacgccttc cactcgccgc tcatggagcc gatgctcgac
 8651 gccctccgcg ccgtcgtcga agacctcacg ctccagccgc cgtcctgcc
 8701 ggtcgtctcc aacctgaccg gcaagccgc caccgtcgcc caactcacct
 8751 ccgccgacta ctgggtcgac cagtcgggc acgcggtccg cttcgccgac
 8801 ggcacgcact ggctcgcccg gcacgacacc accgccttc tcgaactcgg
 8851 ccccgacggc gtgctgtccg ccattggcca ggactgcctg gacgccggcg
 8901 acgcagacgc cgtcaccctc cccgccttcg gcgcggggcg ccccgaggag
 8951 cacaccctca ccaccgcct cgcgggtctg caggtccacg gccccacct
 9001 ggactggacc ggctgcttcg ccggcacccg cgcgcgcgc accgacctgc
 9051 cgacctacgc cttccagcgc cgcgctact ggccaaggc cctccagagc
 9101 ggcaaccgcg acctgcgctc ggtcggcctc ggtgcgcgc accacccgct
 9151 gctctccgcc gccgtctccc tcgccgacgc aggcggcacc ctgctcaccg
 9201 gccgcctctc ccggcagacc caccctggc tcgccgacca caccgtccgc
 9251 ggcaaccacc tgcgtcccggt taccgccttc ctgaactcg ccgtccgcgc
 9301 cggcgacgag gtccggtcgc accgcgtcga ggaactcac ctccgcgcac
 9351 cgtcctctgt gccgaacag ggcggcgctc aggtccagtt gtggatcggc
 9401 aaccccgacg tgcctggctg ccgcaccgct aacgtccacg cccgccccga
 9451 caccggcgac gacacccctt ggaccgcca cgcaccggc gtctcacca
 9501 ccgccgacgc ctcccgccag ctcccggtt cgtccgagca gggcggcacc
 9551 cccctcgccg gcgaccccca ccccgccctc gacgcggccc agtggccccc
 9601 ggcggggccc gaaccgctgc cgtgggacg cactacgac cgcctcgccg
 9651 acggcggtt cggctacggc ccggtcttc agggcctgag cgcgcctgg
 9701 cgcggcgcg acgtcgtcta cgcgaggtc gagctgccg aggcgggccg
 9751 gtccgacgcc gaggcggtc gcctccacc cgcctgctc gacgcgcgc
 9801 tgcacgcgc gcccttcacc ggctcggcg aacgcggccg gggcggcctg
 9851 ccgttctctt gggagggcgt ctccctccac gccggcgcg ccaccacct
 9901 ccgcgtccgc ctgaccccg tcgccgacga cgcgtcgcc ctgaccgtcg
 9951 ccgacggcac cggcgcgccc gtgctgtcc tcgactcgt cgtcctgcgc
 10001 agcgtggcga cccaacagct cgacacggc gccgccgtc cccgtgacgc
 10051 cctcttcgc ctcgactgga ccccggtcca gccgaccgc accgacccg
 10101 ggcggctgc ctcctcggc gccgaccct tcggcctgt caccacgcc

10

20

30

10151 ggattcgcgcg acgccccggc ataccgcggac ctgcgcgcgc tcgcgcgcgc
 10201 ggacggcccg gtcccgacca cgcctcgtgt gtccctcgcg ggcaaccggg
 10251 acgacgcggc cgaccgggcc cggctcgcac accgctgcgc cgcggaggcc
 10301 ctgcgcgcgc tacagacctg gctcgaccac catgagcgct tcgcgcgcgc
 10351 ccgcctggtc ttctgtaccc gtgggtgcgc ggtcgggcgt gatgttgctg
 10401 ctgctgcggg gtggggctct gtgcgttcgg cgcagtcgga gaatccggg
 10451 tgttttgctc tggtcgatct ggatccggat ggtgcggctg gtgcgcgcgc
 10501 gctcgtcgtc gcgttggtca gtgggtgagc gcagcttcgc gtgcgcgcgc
 10551 atgtgttcgc ggtcgcgcgt ctggtgcgcg ggccgcctac cgaggctcgt
 10601 gcgggtgctg atggcaccgg ggatggcgct ggggatggct ctggtgtgtc
 10651 gttctcgggt gagggtgcgc tcctggctac tgggtgtacg ggtggtctgg
 10701 gtgcgggtgt ggcgctcat ctggtggcgc agtatggggg gcgggatctg
 10751 ctgttggtca gtgcagtggt tgaacgtgcc gtgggtgctg gggagttggt
 10801 ggcgagctt gcgggtgtgg gtgcgcgggt gcgggtggtt gcgtgtgatg
 10851 tgaccgatcg tgccgcgggt gtggagttgg ttggcgggca tgcggtgtcc
 10901 gcgggtggtc atgcgcgtgg tgtgctggat gacggcatgg tgggtgcgtt
 10951 gaccggggag cggttgtccg cgggtgctgc gccgaagggt gatgctgtct
 11001 ggcatctaca tgaggcgacc cgcggcctgg acctggacgc gttcgtcgtc
 11051 ttctcctccc tcgcgggggt cttcggcagt cccggccagg ccaactacgc
 11101 ggccgcgaac gccttcctgg acgcctgat gacgcggcgc cgggcggagg
 11151 gactgcccgg cctgtcactc gcattgggac cgtgggagca gtcgggcgga
 11201 atgacgggca ccctgacgga cgtcgacgcc gaacggctgg ccgcctccgg
 11251 tgtccgcgcg ctctccgtgg cgcaggccct ggccctcttc gacgctgccc
 11301 tggccgggac cgacgccacc tgcgttcggg tccgctgga cctcccgcgc
 11351 ctccgcgcac ggggtgaagt gccgcgcgtg ctgaggtcgt tgatccgcgt
 11401 ccgggcgcgc cgagccgcgc tcgcgggctc cgcaccgcgc ggcaacctcg
 11451 cccagcgcct gcgcgcctg gacgaggacg gccgcgacga gatggtcctg
 11501 gacctcgtcc gcggtcagggt cgcctcgtc ctcgccacgc cgaccggtgg
 11551 cgacgtcgac gccggccgtg ccttcgcgga cctcggtctc gactcgtgga
 11601 ccgcctcgca actgcgcaac cgcctcaaca ccgtcacggg cctgcgcctg
 11651 cccgccaccc tggctcttga ctaccgacc gtccggcacc tcgccacgta
 11701 cgtcctggac gatttgttgg gcacggatgc cgagggtggc accgtgcagc
 11751 cggccgcggg tcgggtggcg gacgatccga tcgtcatcgt gggcatggcc
 11801 tgccgctacc ccggtggcgt cagctcccc gaggacctgt ggcgcgtgct
 11851 caccgaaggc accgacgcgc tctcgggctt cccgaccaac cgtggttggg
 11901 acgtcgaatc cctctatcac ccggacctg accaccggg tacctcctac
 11951 acgcgctcgc gtgggttcct gcattgaggcg ggggagttcg atccggggtt

10

20

30

12001 cttcgggatg agtccgcggg aggcgttggc gaccgattcc cagcagcggg
 12051 tgttggttga gtcgtcgtgg gaggcgatcg agcgggcccg tattgatccg
 12101 gtgagtttgc ggggtagtcg gacgggtgtg ttcgcggggg tgatgtacag
 12151 cgattacagc gcgatgttgg caggtccgga gttcgagggg ttccagggca
 12201 gtgggagttc gccgagtttg gcgtcgggtc ggggtggccta cacgttgggg
 12251 ttggaaggcc cggcggtgac ggtggatacg gcgtgttcgt cgtcgttggg
 12301 ggcgatgcac tgggcgatgc aggcgttgcg tagtggtgag tgtgggttgg
 12351 cgttggcccg tgggtgtgac gtgatgtcga cgcctgcggg gtttgtggac
 12401 tttgctcggc agcgggggtt gtcgcgggat ggccggtgca aggcgtttgc
 12451 ggatgcggcc gatggtgtgg gctggtccga gggcgtcggc gtgttgggtc
 12501 tggagcggca gtcggacgcg gtgcgcaatg gtcacgagat tttggctgtg
 12551 gtgcgggggt cggcgggtcaa ccaggatggg gcgtccaatg gtttgacggc
 12601 gcccaatggt ccgtcgcagc agcgggtgat ccggcaggcg ttggccagtg
 12651 gtggtctgac ggccggtgac gtggatgtgg tggaggcgca tggtagggg
 12701 acgacgctcg gtgatccgat cgaggcgagc gcgttgttgg cgacgtatgg
 12751 gcgggatcgt gagcctgagc ggccgttgtt gttgggttcg gtgaagtcca
 12801 atctggggca tacgcaggct gctgcgggtg tggcgggctg tatcaagatg
 12851 gtgttggcga tgcggcatgg tgtggtgccg cggacgttgc atgtggatgc
 12901 gccttcttcg catgtggact ggtccgaggg tgcggtggag ctgctcagtg
 12951 agcaggcggc ctggccggag acgggtcggg tgcggcgggc ggggtgtctcc
 13001 tccttcggca tcagcgggtac caatgtgcat gtcacgtcgc agcaggcgcc
 13051 gggcgccaag gcgatgcgcg cggccgggtg ggcgcggcgc acgcggggtg
 13101 ccgtgccggg gctgctctcg gggcgtggcc ggagtgcctt gcggggccag
 13151 gccgcccgcc tgcctggaca cctccaggcc cgacccgacg ccgaactcgt
 13201 cgatgtcgca ctgtcgttgg cgaccacccg gtcccgcctc gagcagcggg
 13251 ccgcctcgtt ggcgcaggac cgcgaccagc tgatcgctc gctgggggcg
 13301 ctggcccgcc accgccccga ccccgccgtc gtcgagggcg aggcgcggc
 13351 acgcggcccg accgcgggtg tgttcaaccg acagggcagc cagcgggccc
 13401 ccatggggcg tgaactccac gaggtgcagc cggagttcgc cgcggcgttc
 13451 gacgcgggtg gtgccgtttt cgacccgctg ttggaccggc cgtgcgcga
 13501 ggtggtgttc gccgaggacg gcagcgacga ggccgcactg ctggacgaga
 13551 ccggttggac gcagccggct ctgttcgccg tcgaggtggc gctgttccgg
 13601 ctggtggaga gttggggtgt ccgtccggac ttcgtggccg gccattccat
 13651 cgggtgagatc gcggcggcgc acgtcgccgg ggtgctgacg ttggaggacg
 13701 cctgccgtct ggtggccgcg cgggcgacgc tgatgcaggc gctgccgacc
 13751 ggccggcgca tgatcgcat ccaggccacc gaggacgaga tcgcggcgca
 13801 cctcgacgac acggtggcga tcgccgccgt caacggggcg cagtccttgg

10

20

30

13851 tgatctccgg tgacgaggag gccgccgaaa cgatcgccgc cacgttcgcc
 13901 gaacgcgggc gcaagaccaa gcggctgcgg gtgagccatg ccttccactc
 13951 gccgcggatg gacgggatgc tggacgcctt ccggatcgtc gccgaggggc
 14001 tgacctaccg ggccgcgcgc atcccgctcg tctccgacct caccggccgg
 14051 cgccgcgacg atgcggaggt gtgcaccgcg gactactggg tgcggcacgt
 14101 ccgagaggcc gtgcgggttcg ccgactgcgt gcggacgctg cgcgacgccg
 14151 gggccaccac cttcctggaa ctgggctccg acggcctgct gaccgcgatg
 14201 gccgaggaca ccctcgggtga cgaccacgac gccgaactgg tgcgatgct
 14251 gcgcgcgggg cgccgcgagg aactggccgc ggccaccgcc ctggcccgcc
 14301 tccagggtgc cgccgtggac gtggactggg cggcgtaact cgcgggcacc
 14351 ggcgcccgcg gcaccgacct gccgacctac gccttcacgc acgcgtacta
 14401 ctggccgcag ctgccgaccc cggccgcgcg cctcgccgcc gccgatcccg
 14451 ccgaccagca gctgtggggc gctgtggagc gcggcgacgc ccgcgaactc
 14501 gccgacatcc tcggcctggg cgaacaggac ctacgcgcgc tggactccct
 14551 gctgcccgcc ctacactcgt ggccgcgcgg caaccaggag aagcacctcc
 14601 tggacacctc gcgtaccgc gtggagtggg cagactgag caagccgacc
 14651 gccccggtcc tcgacggcac ctggctgctg gtgcctccg accccaccgc
 14701 ggccgaccag ccagccctcc tcgacggcct ggccgacgcc ctcggtcgc
 14751 acggcgcgcg ggtgcgtcgc ctgcttctgg acgactcctg cgcggaccgc
 14801 gcggtgctcg ccgaacgact ggccgcggacc gccgacgtgg accccgcgac
 14851 ccagggtgctg tccgtgctgc cgctcgacga gcgggacgcc gacgactgcc
 14901 cgccgctcac ccgcggactg gcgctgaccg tcgcgctcgt ccaggccctc
 14951 gccgacaccg gcgcccaggg ccggctgtgg accgccacc gccgcgcgct
 15001 ctccaccaac cccgccgacc cggtcacca ccccgctccg gccgctgcct
 15051 ggggcctggg ccggggcgtc gccctggagc acccaccgct gtggggcggc
 15101 ctggctgacc tgccgcaggt cttcgacgag cgggccggac agcggtcgc
 15151 cgggatcttc gccgtcaagg acgcaccgga cggcgaggac cagggtggcg
 15201 tcggggccac cggagtctcc ggccgccggc tcgtccgcca caccgtcgaa
 15251 gcgctgcca cggccgcgga gttcacgcc accggcactg tcctgatcac
 15301 tgggtggcacc ggtggcctgg gcgccgaggt cgcgcggtgg ctggcccgcg
 15351 ccggcgccca gcacctcgtc ctgaccagcc gcccgggccc ggacgcgcgc
 15401 ggccgcggcc aactccgggc cgaactggag ggctacgggc cgtcggtgct
 15451 cgtcgtcgcc tcgacgctc ccgaccggga cgcgctcgcc gccgtcctca
 15501 ccgactgccc cgaggaaactg ccgctgaccg gtgtcgtgca caccgcaggc
 15551 gtcggccact accgcccgt ggacacctg agcaccgcc agttcgccgc
 15601 cctcacgcc gccaaactcg ccggcgccgc ccacctcgac gccctgctcg
 15651 ccgaccgcga actggacttc ttgctcctct tcggctccat cgcgggtgct

10

20

30

15701 tggggcagtg gcaaccagag cgcctacggc gccgccaacg cctacctcga
 15751 cgcgctcgcc ctgcaaccgc gcgcccggcg cctcgccggc acctccgtcg
 15801 cctggggccc gtgggcccag gccggcatgg ccgcccagca tgccgtttcc
 15851 gagacctgc gccgccaggg cctcgccctg ctcgaccggg ccccggccat
 15901 gaccgagctg cgcgcggcgg tgcctcgga ggaagtcacc gtcaccgtcg
 15951 ccgacgtgga ctggcagcgc tacgcaccgc tgttcacctc cgcgcggccc
 16001 agcgcctga tcgcccgcct gcccgaggtc cgcgcctcg ccgcccagca
 16051 gcgcaccgag caggacgcca ccggcgccct cgaggtcgtc acccgcgctc
 16101 gcgcccctgc cgaaccagag caactgcgc tgctgaccga cctcgtccgc
 16151 accgagtcgg ccacctgctt cggccacagc tccgcccagc ccgtgcccga
 16201 gggccgccc ttccgagag tcggcttcga ctcgctgacc gcggtcgagc
 16251 tccgcaagcg cctgggccc gcgaccgggc tgcctctgcc cagcaccatg
 16301 gtcttcgact acccgacacc gctggaactc gccagtcacc tgcgggaggga
 16351 gatcctcgcc gcggtgctgg aagtcgccc cccggtcgcc accggcgggc
 16401 ccgacgacga gccgatcgcc atcatcgga tggcctgcgc ctccccggc
 16451 ggcgtcagct ccccggaaca gctgtgggac ctggtcgctt ccggcaccga
 16501 cgcgatcagc gagttcccc tcaaccgcgg ctggcagacc gggaacctct
 16551 tcgaccggga ccccgaccgg ccgggacca cctactccac ccaggggggc
 16601 ttctccacg aggcgacga gttcgacccc accttcttcg gcatctcgcc
 16651 ccgagggcg ctggtcatgg acccgagca gcggtcctg ctggagacca
 16701 cctgggagtc cttegagcgc gccgggatcc gcccggaac cctccgatcc
 16751 accctgaccg gcaccttcgt cggctccagc taccaggagt acggcctggg
 16801 cgcggggcgc ggcaccgagg gccacatggt caccggcagc agccccagtg
 16851 tgctctccgg ccgactgtcg tacgtcttcg gtctggaagg cccggcggtc
 16901 accgtcgaca ccgctgctc gtctctcgct gtggcgctgc acctggcctg
 16951 ccagtcgctg cgcgaaggcg agagcaacct ggccgtcgcc ggcgggccca
 17001 cgatcatgac gacgccaac ccgttcacg cgttcagccg gcagcgcgcc
 17051 ctgcgaagg acggccgctg caaggcgttc tccgacgagc cggacggcat
 17101 gacgctcgcc gagggcgctg gcgtcgctct cgtcgagcgg ctctccgagc
 17151 cgcagcgcaa cggccaccgg gtcctggcgg tcctccggcg ctccgccatc
 17201 aaccaggagc gcgcctccaa cggcctgacc gcgcccacg gcccgctcca
 17251 gcagagggtc atccgccagg ccctcgccaa cgcgcgcctc gcgcccggg
 17301 acatcgagc cctggaggcg caccgcaccg gcacaccgct cggcgacccc
 17351 atcgaggccc aggcactgtt cgcacacct gcgcgcgacc gcgacccga
 17401 gagcgcgctg ctgctcggtt cgggtgaagt caacatcggc cacaccagt
 17451 ccgcccgggg catcgccagc gtcacaaaga tggtcatggc gctgcgccac
 17501 tccgaactgc cgcgaccct gcacgcccgc gcgcgctct cgcacgtgga

10

20

30

17551 ctgggtcgccc gggacgggtcc ggctgctgac ccaggcgcgc gcctggccgg
17601 agaccgggtcg cccgcgcggg gccgcgggtg cctcggttcgg catcagcggg
17651 accaacgccc atgtcctgct ggagcaggcg cccgtcgcgg acaccccggc
17701 cgaggagcgg cccgcggtgg cgcgggtccc gatcgccggc ggctcgtcc
17751 cgtgggtggt caccgcccgc agcgcgcgg cctgcgcgg ccaggccgag
17801 cgcctcctcg cgcacgcga aaccgtcgg accgcctgc cggcgcggg
17851 accgctcgac atcggcctgt cgtggtctc cgcgcgcgc cgtttcgagc
17901 accgtgccc cgtcgtccc cccgcgggca ccgaccgct ggccgcgctg
17951 cgcgcgctcg cgaaggacgg gccctcgccc gtggtcgcgc gtggcgtggc
18001 ggacgtcgag ggtcggacgg tgttcgtgtt ccccggtcag ggttcgcagt
18051 gggtagggat ggggtcccaa ctcttgatg agtcggcggg gtgcgcgag
18101 cggattgccg agtgtgcgg gccactcgcc gagttcaccc actggtcgtc
18151 ggtcgatgtg ctgcgggggt tggtaggtgc gccgtcgttg gagcgggtcg
18201 atgtggtgca gccggcgctc ttcgcgggtga tgggtcgtt ggctgcgttg
18251 tggcgttccc gtggtgtgtt gccggtgctg gtggtggggc attcgcaggg
18301 tgagatcgct gctgcgggtg tgcgggtgc gctgtcgttg cgggacgggg
18351 cgcgggtggt ggctcgcgg agtcaggcca ttggtcgtgc gttggcgggg
18401 cggggaggga tgatgtccgt cgcgctgctg gtggacgtgc tcgaaccgcg
18451 gttggtcgag ttcgaggggc ggggtgcggg gccgcgcgtc aacggcccgc
18501 gctccgtcgt ggtcgcggc gagcccgagg cgtggacgc gctgcacgc
18551 cggctgaccg ccgacgacat cccggcccgc cggatcgcg tggactacgc
18601 ctgcactcg caccaggctc aggcctgca cagggaactg ctggaggtgc
18651 tggcggagct ggccgcgcgc acgtcggagg tgccgttctt ctgaccgtg
18701 accggcgact ggctggacac cgcgcggatg gacgcggct actggttccg
18751 caacctgcgc ggacgggtgc ggttcgcgga cgcggtggcg gacctgctgg
18801 cggcggagta ccgcgcatc gtcgaggtca gctcgcaacc ggtgctgtcg
18851 atggcgggtgc aggaggcgat cgacgaggcc ggctgcccgg ccgtcgcgc
18901 cggcaccctg cccgcgcacc agggcggcac cgaccgttc ctgctgtcgg
18951 ccgccgaggt cttcgtgcgc ggtgtggacg tggactgggc ggggctgttc
19001 gaggggaccg gtgcgtccc gatcgacctg cccacctacg ccttccaa
19051 cgaacacctg tgggccgtcc cgcgcgcgc ggaggcgcgc gccgcgcgcg
19101 acccggaaga cgcggccttc tggaccgcgg tcgaggacgg tgacgtctcc
19151 gcgctgaccg ccgcgtcgc caccgaagag gactccgtcg ccgcgtgct
19201 gccgcctcg acctcctggc gccggggccc cgcgcaccgc tcacacgtgg
19251 acgcctggcg ctaccgcgtc gcctggaaac cctcggcg caccctgccg
19301 caccgcctcc tgaccggcac ctggctgctg gtcaccgcgc acggcatcga
19351 cgacaccgat gtggcagggg cgttgagac ctacggcgcc gaggtgcgc

10

20

30

19401 ggctggctcct ggacgaggag tgcgtcgacc gcgccgtcct gcgggagcgg
 19451 ctggccggcg cggaggacgt gaccggcatc gtctccgtcc tcgccggcgc
 19501 cgagcggacc gacgcggtac cgggcacctc cctggtgctc ggcaccgccc
 19551 tgaccgtggc actgateccag gccctcggcg acgccgaaat cgacgctccc
 19601 gtatgggcgt tgaccgcgcg cgcggtctcc accggccggg ccgacgagct
 19651 gaccgcgccc gtccaggcac aggtcaccgg catcggctgg accgcggcgc
 19701 tggagcaccg gcagcgtggt ggcggcaccc tcgacctgcc cgcgcacctc
 19751 gacgcccggg ccgccagcgc gctcgcggcc gtgctgtccg gcgccctcgg
 19801 cagcgacgac cagctggcca tccggccctc cggggtcttc accgcgcga
 19851 tcgtgcgggc cgaggccacc gccgggcggc ccgccggcac ctggacgccg
 19901 cgcggcacca cactggtcac cggcggtccc ggcaccctcg ccccgcacct
 19951 cgcgcgctgg ctggcccaac gcggcgccga gcacctggtc ctgatcagcc
 20001 ggcgcgccac ggccgccccg ggcgccggcg aactcgtcgc ggaactggcc
 20051 gagtcgggca ccgaggcgac cgtcgcggcc tgcgacatca ccgaccgga
 20101 cgcggtcgcc gcgtgctggt ccgacctcaa ggccgacggg cgcaccgtcc
 20151 gcacctgctg gcacaccgac gccaccatcg agctgcacac cctggacgcc
 20201 accaccctgg cggacttcca cgggtgctg caccgcaagg tcaccggcgc
 20251 ccagggtctc gccgaactgc tcgacgacga agagctggac gacttcgtcc
 20301 tgtactctc caccgcccgc atgtggggca gcggcgccca cgcgcctac
 20351 gtcgcgggca acgctacct cgcgcgctc gccgagcacc gccggggcaa
 20401 cggactgccc gccctgtcgc tgtcctgggg catctggggc gacgacctca
 20451 aactggggcg ggtcgatccc cagatgatcc ggcgacggc cctggagtcc
 20501 atggaccgac agctggccct gagcggcctg cagcgggcgc tggacgacaa
 20551 cgagaacgtg ctgcgggtcg ccgacgtgga ctgggagacc taccaccccg
 20601 tctacacctc cggccgaccc accccgctct tcgacgaggt gccggaggtc
 20651 cgcgggtcca ccgcggccgc cgagcagagc gccgggaccg tcgccgaggg
 20701 cgagtctgcc gccgcgctgc gcgccctgtc cgacgccgag cagcagcgca
 20751 ccctgctgga gaccgtccgc accgagggcg cgtccgtcct cgggctgtcc
 20801 tccgcgagg acctcaccga ccagcgggcc ttccgcgacg tcggcttcca
 20851 ctcgctgacc gccgtcggcc tgcgcaaccg gctcgctccc gtcaccggcc
 20901 tgacgctgcc ctcgacgatg gtcttcgact accccaaccc ggccgcgctc
 20951 gccgcctatc tgcacggcga gctggccggc gcccggtccg ccgccgccgg
 21001 cgcgcggccc gtcccgaccg gcgccccga cgcgacgac ccgatcgcca
 21051 tcgtcggcat gagctgccgc taccgccggc gggtcggctc cgcgaggac
 21101 ctgtggcgga tcgccctgga cgaggtcgac gcgatctccg gcttccccgc
 21151 cgaccgcggc tgggacggcg agggcctcta cgaccggac cccgaccggc
 21201 ccggccgcac ctactcgtc caggcgggat tcctgcgga cgtcggcgag

10

20

30

21251 ttcgacccgg gcttcttcgg gatctcgccg cgcgaggcgc tgtcgatgga
 21301 cccgcagcag cggctcctgc tggagaccgc ctgggaggcg ttcgagcacg
 21351 ccggcatcga cccggtcggc cagcgcggca gccgcaccgg caccttcgtc
 21401 ggcgccagct accaggacta cgcctccggc gtgcccaca gcgagggctc
 21451 cgaaggccac atgatcaccg gcacgctctc cagtgtgctg tccggccggg
 21501 tgtcctacct cttcggcttc gagggccccg ccgtcacgct cgacaccgcc
 21551 tgctcctcct ccctggtcgc catgcacctg gcctgccagt ccctgcgcaa
 21601 cggggagagc tcgctggccc tggccggcgg cgtcagcatc atgtccaccc
 21651 cgatgtcgtt cgtcggcttc agccggcagc gcgccctcgc cgaggacggc
 21701 cgtgcaagg cgtacgcgga cggcgccgac gggatgacct tcgccgaggg
 21751 cgtcggcctg gtgctgctgg agcggctgtc cgacgcccgc gccaacgggc
 21801 accagggtgt cgcgctgac cgcggctccg ccgtcaacca ggacggcgcc
 21851 tccaacggcc tgaccgcacc caacggcccc tcccagcagc gcgtcatccg
 21901 ccaggcgctg gccaaactcc cgggtggcgc cggcgacatc gacgtcctgg
 21951 agggccacgg caccggctacc gccctcggcg accccatcga ggcgcaggcc
 22001 ctgctcgcca cctacggcca ggaccgcgcc ccggaacggc cgtgctgct
 22051 cggctcggtg aagtccaaca tcggccacac ccagatggca tccggcgctg
 22101 ccagcgtcat caagctcgtc cgcgccctcc aggaaggcgt ggtgcccag
 22151 tccctgcaca tcgaccggcc ctccaccac gtcgactggt cctcgggcgc
 22201 catcgggctg ctacccgaac gcaccccgct gcccgagacc ggccggccgc
 22251 gccgcgccgc cgtctcctcc ttcggcatca gcggcaccaa cgtccacacc
 22301 atcctcgaac agggccccgc ggacgaggcg ccacgccccg ccgaccgcc
 22351 gcgggacggc ctggtgccgg tcctgctctc cggccgcggc gaggccgcgc
 22401 tgccgcacca ggccgccgc ctgctcgctc tcgtcgagga gcggcccgag
 22451 gccacacctc ccgacctcgc ccactccctc gccacctcgc gcgcgcgct
 22501 ggaacgccgc gccgccgtca tcgccgccga ccgcgacacc ctgaccgcg
 22551 gcctgcgcgc cctgtccgac ggccggcccc accccggcct ggtccagggc
 22601 accgcgggac gcggccggac cgccttcctg ttcaccggac agggcagcca
 22651 gcgccccggc atgggcgcg aactccacga ccgctaccgc gtgttcgccg
 22701 acgcgctgga cgagggtgct gcccggtcgc acgacggacc ggaccggccg
 22751 ctgcgcgagg tgcgttctgc cgcgcccgac tccgccgagg ccgcgctcct
 22801 ggaccggacc ggctacgcc agcccgcgct gtccgccgct gaggtcgcgc
 22851 tgttcgcctc gctgacgtcc tggggcctga ccccgacta cctggccggc
 22901 cactccgctc gcgaactcgc cgcgcgcac gtcgccggcg tgcgtcgcct
 22951 ggacgacgcc tgcaactctg tcgccgccgc cggccggctc atgcaggcgc
 23001 tgcccgaggg cggcgcgatg gtcgccctgg aggcgcgga ggacgaggtc
 23051 ctgcgctcc tggaggggct caccgaccgc gtgtccgctc ccgccgtcaa

10

20

30

23101 cgggccgogg tccgtgggtcg tcgccggcgt cgaggaggac gtgctcctcc
 23151 tcgccgacct cttcgccgcc gacgggcgcc ggaccaagcg gctgcgggtg
 23201 agccatgcct tccactcgcc gctgatggac gccatgctcg acgacttcgc
 23251 cgcgctcgcc cgcggggtga cctaccaccc gccgacgac ccgttcgtgt
 23301 cgaacgtcag cggcgccctg gccaccgcc aacaggtccg cacgcccagc
 23351 tactgggtcg ggcacgtccg cgcgcgggtc cgcttcgccg acggcatcga
 23401 ctggctcgcc acccagggcg acgtccacac cttcctggag ctcgggcccg
 23451 acggcgtgct cagcgccatg gcccgggaga gcctcaccga cccgtcccgc
 23501 acggcactgc tgccgaccct gcgcggcgac cggcccagg aacctgccct
 23551 ggtcacccgc gtcgcgcggg ccacgcgcga cggcgcccgc gtcgactgga
 23601 gcgggtactt cgccgaccac ggcgcgccg ggaccacgct gccgacctac
 23651 gcgttcacaac gcgagcggtg ctggcccgac accaccgcc ccacgagcgc
 23701 ccacacgccc ggatccgccc tcgacgccga gttctgggcc gccgtcgagc
 23751 gggacgacgt cgcgcacctc gccgcctccc tggacctgga cgacgccacc
 23801 gtcaccgoga tgggtcccgc gtcaccgcc tggcgccggc gccgcggcga
 23851 gcagaccgaa ctggactcct ggcgctaccg cgtaacctgg aagccgcgcg
 23901 gcggcgccac cgcaccgcc gccctcaccg gccgctggct cgtgctcgtc
 23951 ccgcacgacc accaggaccg tcaggacgac gcgaccggg cctgggcagc
 24001 cgacgtcgag accgcctcgg gcaccaccac cgtccggctg acggtcacca
 24051 ccaccgaccg cgcgcgcgtg gccgcccga tcaccgaagc cgcggcgac
 24101 cagggcccggt tcagcggtgt gctgtccctg ctgcgcctcg ccaccggcga
 24151 cgcgggccac cccggtgcgc ccgcgcctc caccctcacc accaccgccg
 24201 tccaggccct cggcgacgcc ggcatcgacg cgcgcgtgtg gaacgtcacc
 24251 cgcggggccg tggcgctcgg ccgcgcgaa caggtcaccg cgcgcgaaca
 24301 ggcgcgcgtc tggggcctgg gccgcgcgt cgcctggaa ctgcggccc
 24351 ggttcggcgg caccctcgac ctgcgcgcca cctggacgg ccaggccgcc
 24401 cgcgggttgc gcgcggtgct cgcgctacc gacggcgagg acgcggtggc
 24451 cctgcggccc tcggcgctc tcctccgcc cctggcccac gcccgggcg
 24501 gcccgcacac cgcgcgacc gccttcgacc cggcgcccg caccgtcctg
 24551 atcacccggc gcaccggcg catcgccggc cagtcgccc gccgcctggc
 24601 ccgcgacggc gccaccacc tgctgctcac cagccgcgc ggcccggcg
 24651 cccccggcgc cgacgcgctc cgcgcggaac tggaggaaact gggcgcccgg
 24701 gtcaccctcg ccgcctgcga cgcgcgcgac cgcgacgcgc tggccgcgct
 24751 cctgcgcgaa ctgcccgcg acgccccgt gtgcgcggtg ttccacaccg
 24801 ccggcgctcgt cgaggaccac gtcgtggacg cgctcacacc ggagaacttc
 24851 gccgcggtgc tgcgcgccaa gacgctcgc gccaccacc tgacagagct
 24901 gaccgcgac ctggacctcg ccgcttctgt gctgttctcc tccacggccg

10

20

30

24951 gcgtectcgg cgccgccgga cagggaact acgccgcggc caacgcccac
25001 ctggacgccc tcgcgaaca ccgccgtcc caggcctga ccgcgtgtc
25051 cgtcgcctgg ggcctgtgg ccggtccgg catggtcgc gacgcccg
25101 aactcaccga ccgggtacgg ccggtcggt tcgaaccgt ccccccgaa
25151 ccggccgtgc gcgcctgct gcgcgccatc gagaacgac acaccaccgt
25201 ccgcgtcgc gacatcgact gggagcgctt ccagcgcgcc ttgcgcgg
25251 tcgcctcgct gccgttcgt gccgacctc ccgagaccgg ccgggccacc
25301 ccgcgcaccg ccaccggcgc gccaccggc ctgcccagc aactcgccga
25351 actgcccggg cagcagcgcc ccgagcggt cctggacctg ctgcgtacc
25401 aggtcgccgc cgtcctcggc caccgcgacc ccgcaccgt cagggacgac
25451 caccgcttcc ggcacctggg cttcgactcg ctgacctcc tggaaactgc
25501 caacgcccctc aacgcgcga ccggcctgag cctgcccgc accctggtct
25551 acgacctgcc caccgcgcg gagatggcg acttcctgct cgcgaactc
25601 ctccgacccc tgcccaccga caccgcgcg accgtcgcca gcacggcctc
25651 cccaagctc tcagcttcgt tcgagcagg ccgtaccccc ttgcgacacc
25701 cgatcgccgt catcgccatc ggctgccgt tcccggcg cgtcaccacc
25751 ccggaggagc tctggcagct cctcgacgag ggcgcgacg gcatcagccg
25801 ctccccgac gaccgcggct gggacctgc ccgcgtggg gccggcgct
25851 ccgacacct ggaggcgcc ttctgaccg gcgtcgccga cttcgacgc
25901 ccgttcttcg gcatctcgc ccgcaggcg ctggccatgg acccccagca
25951 gcggctgctg ctggagacca cctgggagg gctggagcgg gccggcatcg
26001 acccgaccac gtcgcgggc tccaccacc gcgtcttcgt ccgcaccaac
26051 ggccaggact acccgacgt gttgcgcgc tccgcctcg acgtggccgg
26101 ctacgtcgc accggcaaca ccgccagct gatgtccgg ccctgtcct
26151 acgcgtcgg cctgaaggc ccggcgtca ccatcgacac ccctgtctc
26201 tcctcgctcg tcgcctgca ctgggcggc cgggcgctgc gcgccggga
26251 gtgcgacctc gtggtggcg ggcggtctc ggtcatggc agccggact
26301 ccttcgtcga gttctccacc caggcgggc tggcaccga ccggcgctgc
26351 aaggcgttct ccgacgcgc ccagggcacc gcctggtcc aaggcgtcg
26401 catcctcgtc ctggaacgc tctccgcgc ccgcgcgaac ggccaccagg
26451 tcctcgccct gatccgggc accgcgtca accaggacg ccggtccaac
26501 ggcctgaccg ccccaacgg cctctcccag cagcgctca tcgccaggc
26551 actcgccgac gccgcctgc gcccgcgga catcgacgc atcgaggcg
26601 accgcaccg caccacctc ggcgaccca tcgaggccg ccctgatc
26651 accgcctac gccgggacc ggacgcgaa cggcgctgc tgcggggac
26701 cgtcaagtcc aacatcggg acaccaggc ccgcgcggg gccgcggcg
26751 tcataagat gctgatggg atgcgccac gcacctgcc caggacgtg

10

20

30

26801 cacgtgggca ccccggtccag ccacgtcgac tggagcggcg gcaccgtcgc
 26851 gctcctcgac gacgcgcggc cctggccacg gaccgggacg ccgcggcgcg
 26901 ccggcgtctc cgccttcggc gtcagcggca ccaacgcca cgctcgtcgc
 26951 gagcaggccc cggaaaaccga agcccccgcc gccccggccg ccgagccggc
 27001 gccggaggcc acgcccaccg tcgtcccctg ggtcgtctcc ggacgcagcc
 27051 ggggaagcgt ccaggcgcag ctggaccggc tcaccgcgca caccgccggc
 27101 caccgccgcg gctcggcggc ggacgtcggc cgctcgtcgg ccaccgaccg
 27151 caccgtcttc ccgcaccgcg ccgtgctgct cgcgggcccg gacgggggtc
 27201 gcgaggccgc ccgcgcgcgc gcgccccgca ccccgggccg caccgcgttc
 27251 ctgttctccg gacaggggcg ccagcacgcc ctgatgggcc acgacctgta
 27301 ccagcgcttc ccggtctacg ccgacgcact ggacaccgtc ctgcccagt
 27351 tcgacaccgt gctggacgtc ccgtgcgcg ccgcgctgtt cgcgcgcgcg
 27401 ggcacccccg aggcgcgcgt cctggaccag accggcttca cccagcccg
 27451 gctgttcgcc gtcgaggtcg cactgttcgc gctgcgcgag tcctggcggc
 27501 tgacgcggga ctctcgtcgc gccactcca tcggcgagat c

10

SEQ ID NO: 3

NysD2 (incomplete - probably C-terminally truncated).

20

1 msftypvsmv wlqgreldyv teavgggwis sqgpyvrrfe eafaayndvp
 51 fgvacssgtt altlalralg vpggdevivp eftmiasawa vtytgatpvf
 101 vdcgddlnid vsrieekitp rtkvimpvhi ygrqcdmdav lnlayeynlr
 151 vvedsaeahg vrprgdiacf slfankiisa geggvclthd phlaeqmahl
 201 ramafktdhs flhkklaynf rmtnmqaava laqteqltdi lalrrdiekr
 251 ydealrdipg itlmpprdv1 wmydlraerr delcaylage gietrvffkp
 301 msrgpgyfsa dwpalnaarl sadgfylpth tgltaeqef itgri

SEQ ID NO: 4

30

NysD1

1 vtlpsgntrl gwrrrrmhsp gdragrvrga rarrpatfrg vlsmganrrp
 51 ilfvsyaesg llnpllvlag elsrrdvadl wfatdekard evaavvdgsp
 101 vrfaslgtv sqmsavtwd atyaevtqrs rfkahaavir hsfapesrma
 151 kyrrleeive evepalmvie smcqfgyela itkgipfvlv vpfvpsnvl
 201 shvpfaksyt psqfpvphsg lpaamslaqr ienqlfrlrt lgmfltsdvr

251 kvveednrvr telgiapgar qmmaridhae qvlycysvrel dypfpmhpk1
 301 rlvgtmvppl pqapdddgl dwlsaqksvv ymgfgtitrl treqvaslve
 351 varrldgrgh qvlwklprgq qellppaael pdnlriegwv psqldvlahp
 401 nvkaiffthag gngyheglyf gkplvvrplw vdcddgaairg qdfgvs1tld
 451 rpetvdtedv ldkitrldq psfteraehf agllrdaggr aaaad1llgl
 501 palatd

SEQ ID NO: 5

NysA

10

1 mtigadedpv vvvgmacryp ggvagpedlw elvrtgrdat tafpddrgwd
 51 laalagdpgp rsatreggfl tgaadfdAAF fgmspreavs tdpqqr1vle
 101 tawealerag idphslrgsr tgvfvgasgq dyaavthasp ddldghaltg
 151 lapgvasgrl ayvlglegpa vtvdttssss lvalhwavra lragecstal
 201 aggvvmstp aafvghtrgg glapdgrckp fsddadgtaw aegvgivvle
 251 hlstaraagn pvlavlrsga vnqdgasdgl tapsgpaqer viraaladar
 301 lapadidlve ahgtgtrlgd pvearallaa yggdrdpdrp lrlgslkstl
 351 ghaqaaagig gviktvtlrl hglmprrhl atptrqvdws qgavapltdh
 401 tpwppadrpr ragvssfgis gtnahvilee appadvptr pgtlrpstvp
 451 wpvsaatpea ldaqrlrlra hlrrhsldlp ldvgyslatg raalrhra1v
 501 lppadgtaad avehargaah qrrtavlfsg qgsqrpgmgr elaarfpvfa
 551 dalddalral drhldgpvre vmwgtdaall drtgwtgpal favevalhrl
 601 vaslgvtpdf vggshvgeia aahvagvls1 edacrlvaar atlmqalp1ag
 651 gamaaleate devapllgah lalaavngpt avvvagaeda vrqltarfad
 701 rgrtrslrav shafhsplme pmldafrdvv srltfhqpsi plvsnl1tgel
 751 agseitsaey wvrhvrtdvr fadgitalak agadvlielg pggvlsamar
 801 dtlcpdsttd vvpalskgrp eetafagalgl rlhtlgvpvd wpafyagtga
 851 rrvclptyaf qhvrhwptpp rpngagpgal ghpllgsave ladggg1tcs
 901 galslrthpw ladhtvagr vlpatallel avragdeagc dvlhelhltt
 951 ppalpddaal hvqv hvgp ad ttgrravtvh trpdhhpagd wtrcatgtlg
 1001 stppsaaeee tggtpaawpp adaepldlad hyerladrgf dygptfrglr
 1051 aawrrgaeif advecppgta ddapdhglhp alldaarhaa mavdgtvpva
 1101 whgvrlhavg atalrvrirp tttgtltlta vdvhgapvvt vealtarpl1t
 1151 deeraaprt1p rqargetpad arparpaaar pgpageplpd ttgshptag1h
 1201 laalppaare rqlldlvrtq aaavlghpgp eavgtrsvfk elgfdslagv
 1251 eladrltart glrlpatlvf nfptperaah rlgellaata pldpgaygee

20

30

1301 ltrfeaiivtn lpqdgperra vadrldaivs alrqnspacev pssdedidtv
 1351 svdrllidiid eefett

SEQ ID NO: 6

NysB

1 mqepqggqpd qqekivdylr rvtsdlrrar rrigesksd nepiaivgm
 51 crlpggvnsp eslwdlvrsq gdaisgfpvd rgwdletltg ngdgssathe
 101 ggflydaaef daaffgispr eatamdpqqr lllavaweal eragiaptal
 151 rgrsrgvfvq syhwgapsad aatelhghal tgtaasvlsg rlaytlgleg
 201 pavtvdtacs sslvalhlaa qslrvgessl aviggvtilt epsvfvefsa
 251 qgglapdgrc kafsdaadgt gwaegvgvlv aerlsdaqrn ghpvlavlrq
 301 savnqdgasn gltapngpsq erviqqalar tgltpadida veahgtgtrl
 351 gdpieaqall atyggqhtpd qplwlsgslks nightqaaag vagvikmvma
 401 lrhghlpptl hadapsshvd wsagsvrllt egqqwpetgr prraavssfg
 451 isgtinahall eqaphpadta dagddaapte pagapaalpww ivsghspgal
 501 rdqaaalaar vetdpalrpq dightlhtar allerravvv apdraellaa
 551 thelaagrea navvegladv egrtvfvfpg qgsqwvgmga qlldesavfa
 601 eriaecaaal aeftdwsldv vlrgvvgaps lervdvvqpa sfavmvslaa
 651 lwsgrgvlpd avvgshqgei aaavvgals lrdgarvval rsqaignala
 701 grggmmsval svdvleprlv efegrsvvaa vngprsvvva gepealdalh
 751 arltaddira rriavdyash shqvedlhee llevlaelap rtsevpffst
 801 vtgdwldtar mdagywfrnl rgrvrfadav adllaaeyra fvevsshpvl
 851 tmavldliee agvtavatgt lrrdgggagr flsaaevfv rgvdvdwaga
 901 fegtgaarvd lptyafqrer ywntrtaadr tpadapmdae fwaaveqadv
 951 saltaalgtd edsvaailpg ltswrrarsq rttldswryr vtwtplaqvp
 1001 ratltgtwll vttgdiddtd vagalesyga evrrlvldde ctdravlrer
 1051 lagaedvtgi vsvlaaaedd aarhpgltrg laltvslvqa lgdaeatapl
 1101 wfltrgafat gpsdpvtrpl qsqiagvgwt talehpqrwg gtvdlpdtld
 1151 araaqrlaaa lsgalgaedq lavraagvla rrivraghra grpartwapr
 1201 gttlitggsg tlapqlarwl aergaehvvl vsrrgadapg apeliaeaae
 1251 sgtevtvaac ditdrdavaa lladltadgr tlrtvihaaa aielsaladt
 1301 tvaefadvvh akvtgarild ellddaeldd fvlysstagm wsggvhaayv
 1351 agnaylsala eqrrarglrt tsihwgkwpd drareladph rirrsogleyl
 1401 dpelaltalg hvldddetvi glmdidwdty hdvftagrpa hlfdqipevr
 1451 rrlldqasvpd pagpaadgla arlhglaaa qdrllltlvr teaaavlgla

10

20

30

1501 saesfperra frdlgfdsvt avdlrnlrva gtglrlpstm vfdhpncaal
 1551 aafikktalg vpgaapqgha atgt padddp iavigmscry pggaatpeel
 1601 lrlaldgadv isefpadrgw darglydpdp drpghtysvq ggflheaagf
 1651 dpgffgispr eavamdpqr lletsweaf eragidpasl rgsaagttffg
 1701 asyqdyssstv qngtgeseah mvtgtaasvl sgrvsyllgl egpavtvdt
 1751 cssslvalhl acqslrdges slalaggaav matphafvgf srqralakdg
 1801 rckpfsdtad gmtlaegvgv vllrlshar anghrvlavi rgsavnqdg
 1851 snglitapngp sqqrvirqal anagltgadv daveahgtgt klgdpieaga
 1901 llatyggdrd aerplllgsv ksnightqaa agvagvikmv lamdagelpg
 1951 tlhldapssh vdwtagavel lrgtrpwpes grprragvss fgisgtnahl
 2001 ileqapatep padpdlrldt atdtvvpwpl aakspaalra qaarllatve
 2051 hdpdlppapv ghalattraa lehravvge rredflrgla alstgastag
 2101 lvsgiaagdp egavfvfpgg gsqwwmgre llatsevfirt aiddcatala
 2151 pyvdwslhdv lagegd pall ervdvvpal fammvglsl wrshgvvpaa
 2201 vvghsqgeia aacvagalsl adaarvvalr sqalpqslgr ggmmvsapv
 2251 ervtallapw qealsvaavn gpssvvsgd tdaldalhta cgegvrark
 2301 vsdyashgr hveavrdela rvlavdpdra pevfyfystvt gdrvddaafd
 2351 gaywytnlrq tvrmeeatra llaaghrvfi evsphpvlaa piqetqeava
 2401 eatggsavvl gslrrdegpp rrfltslaea hthgapvdt ttarsayqp
 2451 vdlptypfqr qdfwpearpa tpaagadasd aafwqlvenq dlaaladal
 2501 vpaddehtal gtvlpalsaw rakagartri delryhvqwt rvaepaaapt
 2551 tgrllvavpp dhada pwaa aldalgt dtv rfeakgt dra gwaaqiaqlv
 2601 edgeeftgvv sllaaedlh pdfgsvplgl gqtlvlvqal gdagltapl
 2651 cltrgavatg rddaldsptq galwglgrvv alehpdrwgg lidlpatlda
 2701 raaarl tgl adpagedqla vratgvlarr mvhaapsapr tgrwrgrgt
 2751 clitgggtggi ggrvarwmae hgaahlvlt rrgpdapgaa alraelealg
 2801 arvtlaacd v adrdalaall adlpadqplt svfhsagvad gdaraadltl
 2851 dqldallrak ltaahhihel tapldldafv lfssgaavwg sggqpgyaaa
 2901 nayldalaah rrsldlpgas vawgtwgevg matvpevher lhrqgvrame
 2951 pdhaigalqg mledddttla vtlmdweafa psftatrpsa lfstvpeavr
 3001 avtgdpgtta gddvdsatpp lrrhleelsa aergralvea vraeasatlg
 3051 hdt pdaipag rafrdvgfde vtave lnrnl rtalglplpa alvfdhptpt
 3101 alaghl gall fgtapedagt grpddpdari realatvpig rlrkaglldm
 3151 vlkladgdat dapapeadap sesldmdae allrlatens an

10

20

30

SEO ID NO: 7

NysC

1 mstnpdkyve alrsslkeie rlrrqneqlv aaavepvavv gigcrfpggv
 51 tspedlwelv aegrdivgpf pqdrgwdlek lagggeggsl aqvvgfveda
 101 agfdpgffgi spreavamp qqrilleitw ealeragidp stlrgtptgv
 151 fvgttgqdyg evikasaedv evysttggha svisgrlsyt lgaegpavtv
 201 dtgcssslva lhwavgalrg gecsmalagg asimatpgpf vaftaqscla
 251 adgrckpfsd radgtgwgeg agmlvlmrls daqregrpvl avlrgsaing
 301 dgasngltap ngpsqqrvir aaldsahlta adidaveahg tgttlgdpie
 351 aqallatygg drprplwlgs vksnightqa asgaagvikm imalqrgvlp
 401 rslhateptt dvdwtagsvd lldetvawpe tgrarragvs sfgisgtnah
 451 vileqaptap eeptteptvr pavvpwalsa rtaaaldaqr arltghladt
 501 pdadpldvgy aladgratfe hravllpdgt elahgtageg pcavlfsggg
 551 sqrpqmgrel harfpvfaaa fdeitalldt hldrplrevv wgtdadllnd
 601 tgwaqpalpa vevalyrlva slgvtpdfvg ghsigelaaa hvagvlsled
 651 actlvaarar lmqalprgga mlairatede vtphltddvs iaavngptsv
 701 vvagteeava aigarftaqd rkttrlrsh afhsplmdpm laefravaag
 751 ltyhepripv lsnlgtvae vadlcsadyw vrhvreaervf adgvtaltlr
 801 gvttlvelgp dgvlsamaqe slpdgaaavp llrkdrpeel savtglarah
 851 vrgvtvrwag lfdgtgarra dlptypfghq rfwpataaraa qdvtaaglg
 901 adhpllgatv eladgagylf tsrlsvrthp wladhgvqgr allpgtafve
 951 lavragdeag cdrveeltla aplvlpergg vqlqvrvgap daagrtrtgi
 1001 fsrvedgfdl pwsqhatgvl tagagapdpt fdatvwppsg aepvdltagy
 1051 erlaalgfgy gpafqglraa wrrdtevyae valpdgadtd paafglhpal
 1101 ldaaqhaaay adlgaisrgg lpfawegvsl aaagattvra riapagedtv
 1151 tiavydaagg tvlsvdsivs revpadapga agtvhrdslf hvewtqlqgr
 1201 pgpapatvav lgpdpdalad tlratgirtt aprdlaalad aegvpvdlvv
 1251 ttltttpgap vpdaabatta avlalaqqwl addrfadarl vlvtrgatdg
 1301 tdpaaaaagg lirtartenp grfalldlap dtgrpdpetl atalaashde
 1351 pdlavrgtdv haarlarvpl atepttwnpd gtvlitgggt glgavlarhl
 1401 vathgvrhll lasxrgpaad gaddltaelt glgatvhiaa cdvadpaala
 1451 dllgtvpagh pltvvvhtag vvddgvlgs ltpqrldtvlr pkadaawhlh
 1501 eatrhldlda fvlfssvaat lgspgganya agnafldala arraatglpa
 1551 tslawgpwtq svgmtsslsd ldveriarsg mppltleqgt alfaalaag
 1601 paalapvrl dlpvlrtqgd apllrglirt pvrtaaqvs qtadglaql

10

20

30

1651 agldaaarre allelvrtqi aqvlghadat evetgrqfqd lgfdsltave
 1701 lrnalntatg lrlpatmvfd yptphaladh lrdellgtea esttavpvpt
 1751 rtagtddpiv ivgmacrypg giaspedlwr lvsqgadatg pfptnrgwdl
 1801 dnlydpdpdr pgrthvragg flhdagsfda dffgmsprea matdsqgrll
 1851 lelsweaver agidpaslrd sgtgvfagvm yndygttltg deyeafgrng
 1901 sapsvasgrv sytlglegpa vtvdtaacss lvalhwaaga lragecslal
 1951 aggvttmstp stfvefsrqr glapdgrska faeaadgvaw segvgmlvle
 2001 rqsdaavrng h eilavvrgsa vnqdgasngl tapngpsqqr virgalasgg
 2051 lstadvdave ahgtgttlgd pieagallat ygrdrdpenp lllgsiksni
 2101 ghtqaaagva gvikvmamr hgvlpqtlhv dapsshvdws vgavellteq
 2151 tawpetgrar ragvssfgis gtnahvvieq sptavpatpa sadrsveepp
 2201 avpwalsgkt pdalrdqaar llahveahpa lrpvdisysl iatrtafdhr
 2251 avvlgtdrae alraltalaa getdpaaltg tvrtgrtafl fsgqgsqrlg
 2301 mgrvlyerfp afaealtdvl taldaelghp lrdiiwgeda qlvdrtygtq
 2351 palfaieval frlleawgit pdfvaghsig eiaaahvagv lslgdacrly
 2401 varavlmqsl peggamiavq atedevlpil tddvsiaavn sptsvvvsgy
 2451 enatlavarh fadqgrtrtr lrvshafhsp lmapmlddfr avvesltfta
 2501 pttpvvsnl gelapaealc sadywvrhvr eavrfadgir tladrgvttf
 2551 velgpdsvls amagesapeg agtipllrrd rpeeqavlaa lchlqvlge
 2601 adwsatfrgl dpvrldlpty afghrwwpa arparpddvr aaglgaaehp
 2651 llgaavqlpd ddgaiftgrl slrthpwlad htvlgtilp gtalvelavr
 2701 agdetgsghl eeltlaaplt lpedgatllq vrvgsaddtg rrtvtvharp
 2751 dtdadrtwtl hatgvlattp paaaafdtv wppadaeplt tddcyahftt
 2801 hrfaygpafg glraawragd vlyaevalpe satdeaaafg lhpalldagl
 2851 haalladdrd tglpfswegv tlhasgatal rvrlapngpn glsvtaadpa
 2901 gnpvatvtrl larpldaeql tihsaltrda lfhlwdtpvp lpdansapp
 2951 allgpdtavl adalgdpava rhatlddlla gdtppatvl vplgapldgd
 3001 taghahalt r saltlvqqwl atdrladsrl vfvthgavat ddapptdlaa
 3051 aavwglirsa qtenpgtftl lldltepdst talsraltld epqlllragr
 3101 araarltrtp aptttthtpw sadgtvlvtg gtgglgglva rhlvrscgvr
 3151 hllltsrsgv gaagaaglvv eleslgarvv vaacdvgdgs avaelvagvs
 3201 esyplsavvh aagvlddgvv gsltperlaa vlrpkvdgaw nlheatrgld
 3251 ldafvvfssv agvfggagqa nyaagnafld almvhrvagg lpgvslawga
 3301 wdqvgvmtag lterdvrraa esgmplltvd ggvalfdaal atgsaalvpv
 3351 rldlaalrtr gdiaplrlgl vraplrrtaa tglatgadtg lvqrlgrldh
 3401 aqrhealldm vrssaalvlq hadgnaidae rafrdlgfds ltavelnrnl
 3451 rtatghlrsa tmvfdhptls alaehirdel fgavesevrv pvqalpptad

10

20

30

3501 dpivvvgmac rfpggvtspe dlwrlvddgt daittfptnr gwldldnlydp
 3551 dpehfgtsyt rsggflheag efdpaffgms prealatdsq qrllesswe
 3601 aieragidpl tlrgsatgvf agvmysdygs ilggkefegf qgggsagsva
 3651 sgrvsyalgf egpavtvdta cssslvalhw aaqalragec slalaggvtv
 3701 mstpstfvef srqrglapdg rskafaeaad gvgwsegvgi lvlerqsdav
 3751 rngheilavi rgsavnqdg sngltapngp sqgrvirgal asggilstadv
 3801 daveahgtgt tlgdpieaqa llatygrdrd penplwlgsi ksnightqaa
 3851 agvagvikmv mamrhgvlpq tlhvdapssh vdwsvgavel lteqtawpet
 3901 grvrragvss fgisgtnahv iveqpaves paaepsgrep gvvplplsge
 3951 spealrdqaa rllaglaerp alrpdlgys lattrsafdh ravvlatdra
 4001 davraltala aadadlsavv gdtrtgrhav lfsgqgsqrl gmgrelyerf
 4051 pvfaealdva idhladaalpa qaslrevmwg ddvelldetg wtqpalfave
 4101 valfrlvesw gvrpfdvagh sigeiaaahv vgvfsledac rlvaaratlm
 4151 qalptggami aiqaaedevt qhltdvsvia avngptsvvv sgaesaartv
 4201 adrlaengrk ttrlrshaf hsplmdpmla efravaegls yatptlpvvs
 4251 nltgrlatad dlcsaeywar hvreavrfad gvstlenegv ttfllelpgdg
 4301 vlsamaqqsl tgdaatvpal rkdrdeetsa ltalahlhta glrvdwaaff
 4351 agsgatravl ptyafghaty wptgtlptah aaavgltaae hpllngsvel
 4401 aegegvltfg rlsqshpwl adhavgqvl lpgtallela fragdeagcd
 4451 rveeltlaap lvlpergavq tqvrvgvadd tgrrtvtvhs rpehatdvs
 4501 tqhatgtltm gsapadtgfd atawppadae platddcyar fttlgfaygp
 4551 vfgglraawr agdvlyaeva laestgdeat afglhpalld aalhaslvah
 4601 egeesngglp fswegatlya tgatalrvrl tptgtmgrsv aiavadtagr
 4651 pvaaidnlvs rrvsgdqltg aaglardalf tldwnpvpen lvpnpvpen
 4701 tggghaqdqd grpaaatval vgadgtaiaa dltaagihtt lhpdlttlat
 4751 tdadvpktvl ipltgtgtgt gtgtestdgi gtgaaesdas apspaevaht
 4801 lstaalalvq ewtaqerfag srlafvttga taaggtvmd vaaaavwglv
 4851 rsagseapdt fvlidrdpgp agthdrtaaa ergqlllral htdepqlalr
 4901 dggvlaarla rfdtaaaltp padrawrlds takgsingla ltpypaalap
 4951 ltghevrvv raaglnfrdv lnalgmypygd dvgsfgseae gvvvevgpev
 5001 tglapgdgvm gmitgsfgsl avddarrrlar lpedswetg asvplvflta
 5051 yyalkelggl ragekvlvha gagvgmaai qiarhvgaev fatasegkwd
 5101 vlrlsgvadd hiasrtldf eaafaevagd rgldvvlsl agdfvdasmr
 5151 llgdggrfle mgktldiraad svpdglisyqs fdlawvvpel igtmlaelmd
 5201 lfrtgalarpl pvrtwvvrha kdafrfmsma khigkivltl prswkpegtv
 5251 lvtggtgglg glvarhlvrs cgvrhlills rsgvgaagaa givaeleslg
 5301 arvvaacdvd gdgsavaelv agvsesypls avvhaagvld dgvgsltpe

10

20

30

5351 rlaavlrpkv dgawnlheat rgldldafvv fssvagvfgg agqanyaagn
 5401 aflaldalmvhr vaggldpgvsl awgawdqgvg mtaglterdv rraaesgmpl
 5451 ltvdqgvalf daalatgsaa lvpvrldlaa ltrrgdiapl lrglvkapir
 5501 raaattpgdt glaeqltrfq raerrdtlla lvrdaqamvl ghtsgdgvdp
 5551 srafrldlgfd sltavelrnr igaatglrlp atavfdypta dalaahlte
 5601 llgpdaesdp depgdptagp tddpiviigm scrfgpdigs pedlwrlldg
 5651 gadvvtdfpt nrgwdldnly dpdpahagts yartggfild aadfdadffg
 5701 mspreamatd sqqrllless weaieragid pltlrdsrtg vfagvmysgy
 5751 gtrldgaefe gfqgqgsals vasgrvsytf gfeqpamtvd tacssslval
 5801 hlaaqalrgg ectlalaggv tvmsipdtfi efsrqrglap dgrskpfses
 5851 adgvwsegv gmlillerqsd avrnghqila vvrsgavnqd gasngltapn
 5901 gpsqqrvirq alasgglsta dvdaveahgt gttlgdpiea gallatygrd
 5951 rdpenplllg sikslnightq aaagvagvik mvamrhgvl prslniteps
 6001 shvdsagav ellteqtawp etgrarragi ssfgisgtna hvilegpeaa
 6051 rhsapeeadt aaaaakapat ahlpvmpwal sgktpealra qaarllahlq
 6101 qrpelapadi alsatqrsq fthravvlst drdeatrlas alattaasdp
 6151 saltgtvtmg rcavlfsggg sqrlgmgrl yerfpvfaea ldvvidhlda
 6201 alpagaqlre vmwgddvell netgwtqpal faievalfrl veswgvrpdf
 6251 vaghsigeia aahvvgvfl edacrlvaar atlmqalpam gamiavgate
 6301 deviphltde vaiaavngpt svvisgaea tqtvahfad qgrrttalrv
 6351 shafhsplmm laefravaeg lsyatptlpv vsnlgtgvat adelcsaeyw
 6401 vrhvreaervf adgvtaleae gvrtflelpg dgvlaamare tvaddtvtvp
 6451 vlrrnmpeer tlitalgrlh ttgtpidwaa llaptgarpv dlptyafqhr
 6501 pfwpsgprdt adaaavgia ashpllngiv eladeegllf tgrlsqshp
 6551 wladhavmgq vllpgtalle lalragdevg cdhveeltla aplvlperga
 6601 vqtqvrvgva dttgrrtvti hsrparattt dsdthtgdtd pwtqhatgvl
 6651 vaglpatatv pfdatvwppa haepvldadf yasragegfg ygpafqglra
 6701 awrrdgevfa dvalpeagrt eaeayglhpa lldaglhaw lvapdgeptr
 6751 tgsvpfswrg vflaasgass vrvrlgrdsd gtlsalaiadt tgapvasvga
 6801 lsmrtvsvta lsataglard alfrldwasa pepacqpddt vtvipavavv
 6851 gtetseltse ltaalraaga dvdvrttlst depapalial plvasdqtgt
 6901 aeaapvpaav hdltrralal vqtrlqeqhf adtkfvfvr gatvgrdvaa
 6951 aavwglvrse qsenpgcfal vldldpgavg aaalvaalvs gepqlavrgd
 7001 vlrvarlvrr pltevgagad gtgdgvgggs gvsfsggav lvtgggtgglg
 7051 avlarhlvae ygvrldllvs rsgervagag elvaelagvg arvrvvacdv
 7101 tdraavvelv gghavsavvh aagvlddgmvt galtgerlsa vlrpkvdavw
 7151 hlheatrgld ldafvvfssl agvfgspgqa nyaaanafld almrtrraeg

10

20

30

7201 lpglslawgp wsltdgtsgm ladaeadrlt rsgvppltae qglalfdaal
 7251 atgdatcvpv rldlsalraq gevppllrsl irgrsrraaa aesatatglr
 7301 erlvglnpve rqevlldlvr gqvalvlgha daddvhpara frelgfdsit
 7351 svelnrnlnt vtglrlpatm vfdyptvevl vsyvldellg tdaevatvqp
 7401 aavavaddpi vivgmacryp ggvaspddlw rlvtdgvdav spfptnrgwd
 7451 veslyhpdpd hlgtstyrsg gflheagefd pgffgmspre alatdsqqr1
 7501 llessweaie ragidpvslr gsrtgvfagv mysdysamla spefegfqgs
 7551 gsspslasgr vaytlglegp avtvdtacss slvamhwamq alrsgcgla
 7601 laggvvmst pavfvdfarq rglsdpgrck afadaadvgv wsegvgvlvl
 7651 erqsdavrng heilavvrgs avnqdgasng ltapngpsqq rvirgalasg
 7701 gltagdvdiv eahgtgttlg dpieaqalla tygrdreper plllgsvksn
 7751 lghtqaaagv agvikmvlam rhgvvprrth vdapsshvbw segavellse
 7801 qaawpetgrv rragvssfgi sgtnahvile rpeaarprvm etntvepstv
 7851 pwlsgktpe alraqaakll ssieerpelr lvdvgmslv grstfehrav
 7901 vlaadradaa ralsaiade adaaaatgrv gagrhavlf sggagrlgm
 7951 relyerfpvf aealdvvvdh ldaalpaqag lrevmwgdda ellnetgwtq
 8001 palfaieval frlveswgv pdfvaghsig eiaaahvagv fsledacrlv
 8051 aaratlmqai paggamiaqv atedevtphl tddvaiaain gpnalvvsgv
 8101 edaaveigar faaegrtrtr lhwshafhsp lmdpmlaefr vvaeglsyaa
 8151 pslpvvsnl1 gqvatalelc saeywvrhvr eavrfadgvt aleaegvrtf
 8201 lelpgdgvla amagasltes slavpllrkd rpeepaala laqlhiagar
 8251 vdwpvlfagv gagrvelpty afqrgwfwpv grvgvggdvg avglgsaghp
 8301 llgaavelaa gagvltgrl slsshgwlad havmgrvfv gtallemvmr
 8351 agdevgcgrv eeltlaapl1 lperggvrvg vavdapdaag rrgvgvyscp
 8401 dgvgqavwsq havgvlavsv adqvggfgdg gwppqgavs vdaegcyelf
 8451 adagfgygpv fgglravwrr geelfaeval sdevaesadt atgfglhpal
 8501 ldaslhasll sslegqsadg gpalfaweg vslfasgata lrvrlapage
 8551 havsvtavdp tgapvisida lrtrrltlde vnashtqlsd alfgvqwtv
 8601 pstpaadhps vaiigtthlg laealssssa gattttttaa yesldaliaa
 8651 gpevsvpdvt liglttedai agyvndhdat vagggtigag aaavdaarrl
 8701 taealrtiqa wladerlaar rlvfvtrgaa dgqdvaaaav qglvrsaqte
 8751 npgtfglldl dgteastavl gealtsdepq lllrdghlha arltrlaspa
 8801 dtavptewna dgtvltitgt gglgaqfarh lvdrygvznl llvsrrgpda
 8851 pgttelvael tahgaevavq acdvadgdav aalvagvpde hplravvhta
 8901 gvlddgvig1 lteerlatvl rpkadaawhl heatrgldld afvvfssvag
 8951 vfggaggany aaanafldal maqrraaglp glslawgpwd qtggmtgmls
 9001 daeadrlars gipplsaeqg lalfdaalal agtstpdraa gsaaastsgt

10

20

30

9051 gdtiaipaaa lvapvrldla alaaggevpa ilrglvtrtr rrrtaaggsvt
 9101 vaglvnrlsg ltaderrqel lelvrtaaal vlghadpasv dstaqfrdlg
 9151 fdsltaveir nrlstatglr ltatlvfdyp ntdalaehlr delfgavese
 9201 vrvpvqalpp taddpivvvg macrfpggvt spedlwrlvd agtdaittfp
 9251 tnrgwdlesl ydpdpahlgt sytrsggflh eagefdpaff gmsprealat
 9301 dsqqrllles sweaieragi dpltlrgsat gvfvagmysd ygsilggkef
 9351 egfggqgsag svasgrvsyt lgfegpavtv dtacssslva lhlaaqalra
 9401 gectlalagg vtmstpgtf vefsrqrgla pdgrskafae aadgvwgseg
 9451 vgilvlerqs davnrgheil avirgsavnq dgasngltap ngpsqqrvir
 9501 qalasggilst advdaveahg tgttlgdpie aqallatygr drdpenplll
 9551 gsiksnlght qaaagvagvi kmvmamrhgv lpqtlhvdap sshvdwsvga
 9601 vellteqtvw petgrvrrag vssfgisgtn ahvileqpea vqrlapgaae
 9651 tvepvai kps aepslvpwal sgkspealra qaarlrdfia erpeprsidi
 9701 ghslavtrsq fdhraivlvd dakapadsla alaalasgva dpavvsdavs
 9751 tggsavlfth qgaqrlgmgr elygrfpvfa ealdivvdhl daalpaqagl
 9801 revmwgddve llnetgwtqp alfavevalf rlvrgwvvp dfvagsige
 9851 iaahvagvf sledacrlva aratlmqalp tggamiavqa tedevtphlt
 9901 devaiaavng ptsvvisgae eatqtvaqhf adqgrrttal rvshafhspl
 9951 mdpmlaefra vaeglsyatp slpvvsnlgt wlatadelcs aeywvrhvre
 10001 avrfadgitt leaegvrtfl elgpdgilsa laqqslagea vtvplrkdr
 10051 geestaltar ahlhtrglie dwqdfagvg agrvelptya fqrqfwfwpvg
 10101 rvgvggdvga vglgsaghpl lgaavelaag agvvltrls lsshgwladh
 10151 avmgrvfapg tallemvra gdevgcgrve eltlaaplvl perggvrvqv
 10201 avdapdaagr rgvgvyscpd gvgqavwsqh avgvasgaa dqvggfgdgg
 10251 vwppqgavsv daegcyelfa dagfgygpvf qglravwrrg eelfaevals
 10301 devaesadta tgfglhpall daslhaslls sleggsadgg palpfawegv
 10351 slfasgatal rvrlapageh avsvtavdpt gapvisidal rtrrltldev
 10401 nashtqlsda lfgvqwtvtp stpaadhpsv aiigtdpfgl adglldalpl
 10451 veergdlaal aasehpvpdl vlvpvagtrr tgvpadaegh tdagtsdmlr
 10501 svreataqvl eqiqqwladd rfeaarlvfv trgavsvgeg giadlaasav
 10551 wglvrseqse npgcflldl dldlaldsl apevdierdr drdpvggtvq
 10601 palaaalhat adepqlalrg gtvqaarltr ipapqtdrae tdpactdrpe
 10651 idtrrrpgtvl itggtgglgg llarhlvaer gvrsylvlasr sglaaegaek
 10701 lvadlealga vvavqtcvda dgdavaalva gvsdeyplta vvhtagvldd
 10751 gvigslteer latvlrpkad aawhlheatr dldldafvfv sslagvlgga
 10801 gganyaaant fldalmaqrr aaglpgvsla wgpwdraggm tgtlsdaead
 10851 rlarsgvpipi saeggalyd aatagerplv vpvrlldlaal rglgdvpall

10

20

30

10901 rglvrtpparr taaagaapsa dvltrqlagl ggaeeevll rlvrgqaavv
 10951 lghadgsaig agrqfqlglf dsltavefrn rlnaatglrl patllfdypt
 11001 padvvghlrg rltggevsga gsvlaaldnl eaviaglsld dagehqlvag
 11051 rlevlrakwa dmrsaegavd ggadvdieea sdddmfalld delgln

SEQ ID NO. 8

Nyse

1 mttsteeslw arcfhpapaa pvrlfcfpha ggsasfyfpv
 41 saqlssvaev faiqypgrqd rrkeagvsdl atladvqyda
 81 lrpllkerps tffghsmgat lafevarrfe addgdlvrlf
 121 asgrapsrv reevhrrsd dgiveelkll agtntallgd
 161 eeilrmilpa irsdygaiet yrcppdvtvr apltvltdgr
 201 dpktsldeae awrghttgdf dlkvlpqghf fvsseapali
 241 dllrahlagh g

10

SEQ ID NO. 9

NysR1

1 mrkqsgssgl lttlvgrdde lrtlarhaaa ardgraglvl lhgpagmgkt
 51 sllrsftasd vcrgmtvlyg tcgetvagag yggvrellgg lglsggdarr
 101 splleglaar alpaltadpa gpdaatgayp vlhglywlaa rlmagrplvl
 151 vladdvhwcde rslawidflr rraedlpllv vlawrseaep vapavladia
 201 aqxrptvlg hplgpddige mvrsvftrta apsfvsrvaa vsggnplala
 251 rlldelraeg vrpdaagerr aaevgshvla rsvrcllerr ppwvrgvara
 301 iavlgpecte llaalagvpa atvdeallvl rragilaadr vdfvhdvvr
 351 avlddvappt laelrtnaal llsdagrpse elagqlmlp vldqpwwaav
 401 lrdaaaqaes rgapeagvrc lyrvleved nvavriqmar alaeinppa
 451 mrllkealsl agdvrtraqv avqygftcla vqespsgvrml ledalaelta
 501 elgpepgpvd relrtlivesv llivgadekv tigavrdraa rltmppgdtp
 551 aqrqmlamt vltamdgrda rsavdqarra lrapgvelep wslsasfal
 601 sladevadaq yaldlmlqyg qdnaavwtyv lalstrallh hgvgaapeal
 651 adaqtaveil geerwadgav lprvalatal vdrgeperae hvldgitrpr
 701 lerfvieyhw ylqarayarw vrgdfqgald lllacgrsle esrfsnpafv
 751 pwwadgavll atldrhdaqr elaaygsela erwgtarglg lafmaaggvaa

20

30

801 pgragidhlt eavslldasp aramearael llghahlkrd dlraarehlt
 851 aaadlaqrcg avklgvdark llvtaggrvr rmtaspldml tgmertvadl
 901 avtgasnrai aealfvtvrt iethltsvyx klvgvgrael savletrtat
 951 sgrqppawvs qargra

SEQ ID NO. 10

NysR2

1 vprskarnqp ttctpgcapd ahgdptmille cgreqrllgd llhrlgggrp
 51 svlsitgrpg haqnalvrwg acrarhdglr vlraqatpae relrygavlq
 101 llavldgphg stldaaairhd gppplpvpgi eevlrrtgta ptlvvvedvq
 151 wldpasltwl qillrhlgpd tplavlassc gdttafdtdp kapavpgppd
 201 tvpvarfvvp altdrgvaat vravcgtpgd eefiaaltsa tagnpailrd
 251 alrafvdhgl padadhlpel haltagvvgd htvraldglp aevnavlral
 301 avcgdlldfh rvralagahs lsedrirtll asvgltvsvg dkvhirfpas
 351 karviedmpa aeradlyvra aelthscgvn dedvahlllr ssplgapwv
 401 pllrrgfaaa lrredhhrac aclsrallqp ldprersllt lelaaaava
 451 rpeagdrllg elvrstvadit dptssgegvg vraidlgfar gnsewvrrta
 501 gealpyagpa dreelvalfw laavrdddap mipvvprlpd rpvppagaga
 551 rawqlatage dadkarklar ialtggvnes lmpkllaaca alfatddnde
 601 avhgltdmtl aarsahlrsm aarifnlrar ihlcaarlea aerddsaer
 651 alpptswhpr alpnliatri lvsmetgrp d rarrlaeapv paggeegvw
 701 palllararv aaddgdweea lrslrecgrw lfrhwanpa mlswrplaae
 751 aclklgdvte arrlrdeelf fadrwgtasa rgiairtttr lfddgdgrav
 801 rrireaaaall rdsparlayl wsrlsqagae tahgdtaaaa rswqavarnt
 851 aahpasrlat aartltvpsv pvatapptav vppgwrdis aekdtvllaa
 901 rghgnrqiae qlavsrrtve lrslsnayrkl riggrkelyl llealegpva
 951 das

10

20

30

SEQ ID NO. 11

NysR3

1 mllerenela riraaldae agdsslllin gplsgsrsal lrripelagd
 51 gtrvlrasaa wrerdfpfgi arqlfdhlls gaggagpaer tagaehfsrl
 101 mdtgdrptgt gpalevsqav lqgaqallad asaerrllil vddlqwadgp

151 slrwlahltr rihglrallv ctladgdhrg ryplvrevag aahtvlrlap
 201 lsrdatravl agpggrppqd alvravyeas rgnplfita rsalratgrp
 251 pggdhfgavr elsptvlrdr laghlriqpq pvrevavava algdhsdpvl
 301 laqlagvdei gfagarralv dagllargrd vrfvhgvvrd avdsliltde
 351 rershdadaad llyrcgrpae qvaghlavv hpgrpwseav lrsaahnalr
 401 agrpadaary lrrallhbrrt qdgcrarilv dlataerald pdacvrhvsq
 451 avalldtsrd raaavlrpp sllaapspsa velvrqaaag ldepgqrdee
 501 gadelalrle awlrhsgghen pvelassvar lrrmgarppv dsvaerelva
 551 vllsagalrg rlsaaeiadt gnrilerepa taahahtplp lvmlslfva
 601 svqgvaswla seqhtrrrya tgaddvllta erafvltvgg rpaaarehve
 651 ralvmdagdw sepavmmfaa vafelrpal serilerird rrpagialta
 701 tgqmlqaavd vhfgrgrdal dtllacgrrl etvgwrnsal lpwrpyaigl
 751 hqrlgetdaa lqlaedelrw arewgattnl gralrlkgwl lqdegldllr
 801 esveilrass yatelartlv vlgrrlpggp eaeavlreaa giaaacgvpw
 851 laeraelglg saivppvatl tpserrvasl vsrgltnqai atelgvssra
 901 vekhltsayr klgvsgrrrel vnalpgr

10

SEQ ID NO. 12**NysR4**

20

1 visaqtapag esvgpglmas ldrdltikha nqefrrridd sagdvgrsf
 51 rdlmhpsvqq plmrqfsrli egkrhrfash vvavgaqdaa fagtltasav
 101 tgktptdiagi lvlmdssgaa daadagvvtg qkkflteida rilegiaagl
 151 stiplasrly lsrqgveyhv tglrlklrvp nraalvsray smgilnvgtw
 201 ppkvvddfik

SEQ ID NO. 13**NysR5**

30

1 vdaegrrrdm lelirrgsa dvvrlaeefa vsketvrrdl nvlegghlir
 51 rrhggaypmv rpgseavfvs rtaqpipees riataaaell seaetvfide
 101 gftpqliada lprdrplativ taslpvvsaf atspqanvll lggrvrrgtt
 151 atvdhwavhm lsgfvidlaf lgaegisrry glttdpava evkaqairva
 201 rrpvlagvht kfgtasfcrrf gevgdletiv tgaglpvaea hryhlmgpkv
 251 lrv

SEQ ID NO. 14**ORF2**

1 vaqdsqgtpr sldhvdqalv halqitpras wtrigsvlgl dautvarrwn
 51 rlvetaaawi schpavlaa sgqgclafve idcapgrlld varalaaaph
 101 vvalshvtgd rdlqlnvmar dpamlsrwtv hdlalldgvr aarthlagpv
 151 htegrswrlr algrhqvarl aadasrhrtv tpaflvldeld qqlvtalsvd
 201 gratyralae qcgagpdtvr rrvqrlfaad mlharcevar plsewptvs
 251 fwgqvpaarl revtrrvtgm revrlcasvi srhnlhlvaw vrslddaqrf
 301 evrlaeraad ltvteraval whmkhghll deegyrvgvt plalwreptd
 351 arrg

10

SEQ ID NO. 15**ORF1**

1 mqpelerlrr slhrepelgl alprtqekvl aaldglplei tpgtgltsvt
 51 avlrggrpgg avllrgdmda lpvteesgvp yaseipgrmh acghdlhtag
 101 lvgaarllae rrerlhgdvv fmfgpggeegh ngagamieeg vldaagkpld
 151 aayalhasn qvpggvitr sgtitsasdl ltvtrvgegg hgstpatak
 201 pvpaacemvt avqnwvtraf difdpvvvtv gtfhagtkas vipdnrrvpg
 251 hgaqlfrirp gparrrrelpr lvraiaaahg leaevdylrh ypvtnvdtde
 301 tafavatard lfapgevves piptngsedf ayvlrrvpga mllvgaapeg
 351 sdwqhaphmh spravfddrv lyrqaallte laarilaata paepavag

20

Translation products of SEQ ID No. 2 (Nys 2):SEQ ID NO. 16

30

NysF

1 vielilpatv ateaayddrp rpgdrllsse revlaraves rgrefttvrh
 51 larralrrlg hpdraillpr rgapqwpvgi vgsnthcagy raaavspael
 101 saavsideap ngplpagvln aialpserph lvalaahrpv vhwdrillfa
 151 kesvfkawyp ltqrelidfse aeividptqg aftarllvpg pllggrrvtv
 201 fpgrwhstpa llttavhlpa pprrdrehr thltvnsplp rptfg

SEQ ID NO. 17**NysG**

1 maspddleee rtaprrparrl vgllrphrrs valavsmgv givlnafgpl
 51 llgrvtdlia dgvlggvpqp apgidfaaig rlllvllaly vvaslfmlaq
 101 grlvasavwr tihelrrdar ekltrlplrh fdrqpagell srttndidnl
 151 gqtlqqtlae litsifslit mlvlmvisp slavvmlsv pvsaliaari
 201 skraqphyaa qwsangtlna hveevctgha likgfdrraa aeerfdacnd
 251 avyraaakaq fasgamepvm mfvanlgyvl vavigawkvi ngltltgdivq
 301 afilyarqfs qpiveiasva grlgsgiasa qrvftlldap eqapdplrpg
 351 tparaegrve ftdvsfrysp dtplienlsl tvepgstvai vgpptgagkt
 401 lgnllmrbye pdsgrilldg tdtatmrdd lrsrfglvlq dtwlfggia
 451 eniaygapga cradieear atcadrfirt lpqgydtvld desgtvsage
 501 kqlltvaraf larpavlvld eatssvdrtr evliqramns lragrtsfvi
 551 ahrlstirda dlivvmdagr iveqgthdql lcaqglyarl haarthtpta
 601 gaaag

10

SEQ ID NO. 18**NysH**

1 vllrllraql rpyawataal valqlvqilg tlllptlgaa lidqgvvrgd
 51 ggritelgvv mgvvalvqia aalgaaalaa rtatamgrdl rsalfrrild
 101 fsareigrfg tpslltrsvn dvqqvgnlaq tgfgivvcap lmcigsvlla
 151 lrqdvplall lvalvlvvav cfglllarmg tlyarmqltl drlgrllrea
 201 itgvrvvrsf vrddherarf aqtnaflvv srrvgrliat mlpvvlillm
 251 gftvallwtg shridagrmv igslsallsy lslilmsvmm lafvflsvpr
 301 arvcagriae vldtgssvap paapqpvrpg agrielcaag yrypgaeepv
 351 lrdvdlvtep geriavlgst gsgkttllnl vlrladateg avrvvgtdvr
 401 eltaatlaaa vgfvpqrppl fsgtvasnir fgrpdatdee lwealrvaga
 451 adfvampdg ldaeitqggg nvsggqrqrl slarallrrp eiylfddcfs
 501 aldgatdaal rtalvpytag atvitvaqri sagrdadriv vldrgrvvaq
 551 gthdvllrts ptyreialsq lteeeahgl agrp

20

30

SEQ ID NO. 19**NysD3**

1 mskralitgi tgqdgssylae hllsqgyqvw glirgqanpr ksrvsrlase
 51 ldfidgdldmd qgslvsavdt vqpdevynlg aisfvpmswg qaelvtevng
 101 mgvlrmleai rmvsglstsr tvsprgqirf yqasssemfg kaaetpgret
 151 tlfhprspyg aakayghyit rnyresfgmy avsgmlfnhe sprrgqefvt
 201 rkislavari kgglqdklal gnldavrdwg yagdyvramh lmlqqdagdd
 251 yvigtgqmhs vrdavriafe hvglnwedyv vidpdlvrpa evevlcadsa
 301 kaqdrilgwkp dvdfptlmmr mvdsdlaqvs renqygdvll aanw

10

SEQ ID NO. 20**NysI aa sequence (partial)**

1 mdneqklrdy lklatadlrr trrrvhkles aaqepvaiig mtcrypggvr
 51 spedlwrme agehgvtpfp tdrqwdleal aaaptasggf lhdapdfdad
 101 ffgispreav amdpqgrvvl esaweaferra gidptsvkgs rtgvfigama
 151 qdyrvgpapg aegfqltgnt gsvlsgrisy tfgtvgpavt vdtacssslv
 201 avhlatqalr agectlalag gvtimgpggt fiemgrqggf sadgrcrsfg
 251 dtadgtgwae gvgilvlerl sdavrnghei lavvrgtavn qdgasnglta
 301 pngpsqqqvi qqalvnarla agdidvveah gtgttlgdpv eaqallatyg
 351 qnrpadrp1l lgsyksenlsh tqaaagvagv ikmvmmamrhg tlprtlhaee
 401 pthhvdwsqg avrlltdttd wpatgaprra avssfgisgt nahtiieqap
 451 epqpadaata qddaagstpa tapvvpgvvp vllsgrtpda lrgqaaalra
 501 aldtgrrpdl ldlahtlatt ragfehravl latdhpaltl gltaladadd
 551 paaapawitg ttraetrlav lftgqgaqrl gagrelaarf pafataldaa
 601 ldaftphldr plrevlwgtl aalldraya qpalfaveva lyrliesfgv
 651 rpdhlaghsv geivaahlag vlsladaatl vaargrlmqv lpdggamiav
 701 gaseadvapl laghedqvai aavngpsavv lsgaeatvta laeqaadgr
 751 ktrrlrvsha fhsp1mepml dafravvedl tlqppllpvv snltgkpatv
 801 aqltsadywv dhvrhavrfa dgidwlarhd ttafielgpd gvlsamaqdc
 851 ldaadadavt lpalragrpe ehtlttalag lnhvgatldw tgcfaagtgar
 901 rtdlptyafq rrrywpkalq sgtadlrsvg lgaahhp1ls aavsladagg
 951 tlltgrlsrq thpwladhtv rgttllpgta flelavragd evgcdrveel

20

30

1001 tlaaplillpe qggvqvqlwi gnpdvsgrrt vnvharpdtg ddtptwahat
 1051 gvlttadasr qlpasseggg tplagdphpa ldaaqwppag aeplpldghy
 1101 drladggfgy gpvfqglraa wrggdvvyae velpeagrsd aeafglhpal
 1151 ldaalhaapf tglgergrgg lpfswegvsl haggattlr v rltpvaddal
 1201 altvadgtga pvlsvdsllv rsvatqqltd aaavardalf rldwtpvqpt
 1251 atdpgpvall gadpfgllth agfadapayp dlaalaaadg pvpttvvlsl
 1301 agtgddaadp arsahrcaae alaavgtwld hherfaaarl vfvtrgatvg
 1351 rdvaavaavwg lvrsaqsenp gcfaivldlp dgavgaalv aalvsgepql
 1401 avrgdvlrva rlvrplte v gagadgtgdg vgdsgsvsfs gegavltvgg
 1451 tgglgavlar hlvaeygvrd lllvsrger avgagelvae lagvgarvr
 1501 vacdvtdraa vvelvgghav savvhaagvl ddgmvgaltg erlsavlrpk
 1551 vdavwhlhea trglldlafv vfsslagvfg spgganyaaa nafldalmtr
 1601 rraeglppls lawgpweqsg gmtgtltdvd aerlarsgvp plsvaaglal
 1651 fdaavagtda tcvpvrldlp vlrangevpp llrslirvra rraavagsat
 1701 agnlaqrlrr ldedgrdemv ldlvrgqval vlghatggdv dagrafrdlg
 1751 fdsltavelr nrIntvtglr lpatlvfdyp tvrhlatyvl dellgtdae v
 1801 atvqpaavav addpivivgm acrypggvss pedlwrulte gtdavsgfpt
 1851 nrgwdvesly hpdphpgts ytrsggfibe agefdpgffg msprealatd
 1901 sqqrlless weaieragid pvslrgsrtg vfagvmysdy samlaspefe
 1951 gfgsgssps lasgrvaytl glegpavtvd tacssslvam hwamqalrsg
 2001 ecglalaggv tvmtspavfv dfarqrglsp dgrckafada adgvwgsegv
 2051 gvlvlerqsd avrnghela vvrsgavnqd gasngltapn gpsqqrvirq
 2101 alasggltag dvdvveahgt gttlgdpiea gallatygrd reperplllg
 2151 svksnlghtq aaagvagvik mvlamrhgvv prtllhdaps shvdwsegav
 2201 ellseqaawp etgrvrragv ssfgisgtv hviveqapga kaiaaagaar
 2251 rtpgavpvll sgrgrsalrg qaarllghlq arpdaelv dv alslatrsr
 2301 feqraavvaq drdqliaslg alaadrpdpa vvegeaagrg rtavliftggg
 2351 sqraamgrel hevqpefaaa fdavcavfdp lldrplrevv faedgsdeaa
 2401 lldetgwtqp alfavevalf rlveswgvrp dfvaghsige iaahvagvl
 2451 tledacrlva aratlmqalp tggamiaiaq tedeiaahld dtvaiaavng
 2501 pqsvvisgde eaaetiaatf aergrktrkl rvshafhspr mdgmldafri
 2551 vaegltyrp riplvsdlgt rredaevct aeywvrhvre avrfadcvrt
 2601 lrdagattfl elgsdglita maedtlgddh daelvpmra graeelaaat
 2651 alarlqvrvg dvdwaaylag tgarrrdlt yafqbayywp qlptpaaala
 2701 aadpadqglw aavergdare ladilglgeq dltpldsllp altswrrng
 2751 ekhlldtlry rvewtrlskp tapvldgtwl lvasdataad qpallidglad
 2801 algshgarvr rlllddscad ravlaerlar tadvdaatqv lsvlplderd

10

20

30

2851 addcppltrg laltvalvqa ladtgaggrl wtatrgavst npadpvthpv
 2901 qaaawglgrg valehprlwg glvdlpqvfd eragqrlagi lavkdapdge
 2951 dqvalratgv sgrrlvrhtv ealptaaeft atgtvilitgg tgglgaevar
 3001 wlaragaqhl vltstrrgpda pgaaelrael egygpsvsyv acdvadrdal
 3051 aavltalpee lpltgvvhta gvghygpldt lstaefaglt aaklagaahl
 3101 dalladreld fflvfgsiag vwgsgnqsay gaanayldal alhrrargla
 3151 atsvawgpwa eagmaaddav setlrrqglg lldpapamte lrravvrqdv
 3201 tvtvadvdwq ryaplftsar psaliaglpe vralaadert eqdatgasev
 3251 vtrvralaep eqlrlldlv rtesatvlgh ssadavpegr afrdvvgfddl
 3301 tavelrkrig aatglslpst mvfdyptple laqylraeil gavlevagpv
 3351 atggaddepi aigmacrfp ggvsspeqlw dlvasgtdai sefpvnrgwq
 3401 tghlfdpdpd rpgttystgg gflheadefd ptffgispre alvmdpqxrl
 3451 llettwesfe ragirpetlr stltgtfvgg syqeyglgag dgteghmvtg
 3501 sspsvlsgrl syvfglegpa vtvdtaacss lvalhlacqs lrrgesnlav
 3551 aggatimttp npfiafsrqr alakdgrcka fsddadgmtl aegvgvvlve
 3601 rlsdaqrngh pvlavlrqsa inqdgasngl tapngpsqqr virqalanar
 3651 lapgdidale ahgtgtplgd pieaqalfat ygrdrdpesa lllgsvksni
 3701 ghtqsaagia svikvmalr hselptilha dapsshvds agtvrlltqa
 3751 rawpetgrpr raavssfgis gtnahvilleq apvadtpae rpavavvpia
 3801 agvvpwvta rsaaalrgqa erllahaetv gtalpaagpl diglslvsar
 3851 arfehravvv ppagtdplaa lravatdgpv pvvargvadv egrtvfvfpg
 3901 ggsqwvgmgs qlldesavfa eriaecaaal aeftdwsldv vlrgvvgaps
 3951 lervdvvqpa sfavmvslaa lwsrgvlpd avvghsqgei aaavvsgals
 4001 lrdgarvval rsqaigrala grggmmsval svdvleprlv efegrsvvaa
 4051 vngprsvvva gepealdalh arltaddira rriavdyash shqvedlhee
 4101 llevlaelap rtsevpffst vtgdwldtar mdagywfrnl rgrvrfadav
 4151 adllaaeyra fvevsshpvl smavqeaide agvpavaagt lrrdgggtdr
 4201 flisaaevfv rgvdvdwagl fegtgasrid lptyafqheh lwavppapea
 4251 vaaadpddaa fwtavedgdv saltaalgtd edsvaavlpv ltswrrarrd
 4301 rstvdawryr vawkplggtl phpsltgtwl lvtadgiddt dvagaletyg
 4351 aevrrlvldv ecvdravlr rlagaedvtg ivsvlaaaer tdavpgtslv
 4401 lgtaltvali qalgdaeida pvwaltrgav stgradelta pvgaqvtgig
 4451 wtaalehpqr wggtdlplaa ldaraaqrla avlsgalgsd dqlairpsgv
 4501 ftrrivraea tagrpagtwr prgttlvtgg sgtlaphlar wlaqrgaehl
 4551 vlisrrgtaa pgaaelvael aesgteatva acditdrdav aalladlkad
 4601 grtvrtvvht aatielhtld attladfdrv lhakvtgaqv laellddeel
 4651 ddfvlyssva gmwsggahaa yvagnaylaa laehrrangl palslswgiw

10

20

30

4701 addlklgrvd pqmirrsgle fmdpqalalsg lqralddnen vlavadvdwe
 4751 tyhpvytsgr ptplfdevpe vrrltaaaeq sagtvaegef aaalralsda
 4801 eqqrtlletv rteaasvlg1 ssaedltdqr afrdvfgdsl tavglrnrla
 4851 svtgltlpst mvfdypnpaa laaylhgela garsaaagaa avptgapdad
 4901 dpiaivgmsc rypggvggae dlwrialdev daisgfpadr gwdaeglydp
 4951 dpdrpgrtys vqggflrdva efdpgffgis prealsmdpq qrilletawe
 5001 afehagidpv gqrgsrtgtf vgasyqdyas gvpnsegseg hmitgtlssv
 5051 lsgrvsylfg fegpavtldt acssslvamh lacqslrnge sslalaggvs
 5101 imstpmfsvg fsrqralaed grckayadga dgmtlaegvg lvillerlsda
 5151 ranghqvlav irgsavnqdg asngltapng psqqrvirqa lansavapg1
 5201 idvleghgtg talgdpieaq allatyggdr aperplllgs vksnightqm
 5251 asgvasvikl vralqegvvp kslhidrpst hvdwssgaig lltertpwpe
 5301 tgrprraavs sfgisgtnvh tileqapade aptpadpprd glvpvllsgr
 5351 geaalraqaa rllafveerp eahltdlahs latsraaler raaviaadrd
 5401 tltrglrals dgrdpplvg gtagrgrtaf lftgqgsgrp gmgrelhdry
 5451 pvfadaldev larlddgpdr plrevlfaap dsaeaalldr tgyaqpalfa
 5501 vevalfrllt swgltpdyla ghsvgelaaa hvagvlslld actlvaargr
 5551 lmqalpegga mvaleaaede vlpillegltd rvsvaavngp rsvvvagvee
 5601 dvllladlfa adgrtrklr vshafhsp1m damlddfaav argltyhppt
 5651 ipfvsnvsgg lataeqvrt1p dywvghvraa vrfadgidwl atqgdvhtfl
 5701 elgpdgvlsa maresltdps rtallptlrg drpeepalvt avaaahahga
 5751 rvdwsgyfad hgarrttlpt yafqrerywp dttaatsaht pgsaldaefw
 5801 aaverddvaa laasldldda tvtamvpalt awrrrrrgeqt eldswryrvt
 5851 wkprggatap aaltgrwlv1 vphdhqdrqd dataawaadv etalgttt1v
 5901 lvtvttdraa laariteaag dqgpfsgvls llplatgdag hpgapaalt1
 5951 tttavqalgd agidaplwnv trgavavgra eqvtapeqaa vwg1lgraval
 6001 elparfggtl dlpatl1dgga arrlrvl1aa tdgedavalr psgvflrrla
 6051 hapagpdtar tafdpagtv litgg1tg1g ghvarrlard gathllltsr
 6101 rgpaapgada lraeleelga rvtlaacdaa drdalaalla elpddaplca
 6151 vfhtagvved hvvdalt1pen faavl1raktv aahlhlhel1ta dldlaafvlf
 6201 sstagvl1gaa gqgn1yaaana hldalaehrr shgltalsva wgpwagsgmv
 6251 adaaelt1drv rrggfeplap epavrallra ienddt1tval adidwerfqr
 6301 afaavrplpf vadlpetgra tpatatgaat glrqqlaelp eherpaavld
 6351 llrtqvaav1 ghadprtved dhafrdlgfd sltilelrna lnaatgls1p
 6401 atlvydl1pt remadfl1lae llgtlptdta atvastaspk lsasfeqggt
 6451 pfddpiavig igcrfpggvt tpeelwqild egrdgisr1fp ddr1gwdlaal
 6501 gagasdtleg gfltgvadfd arffg1spre alamdpqqr1 llettweale

10

20

30

6551 ragidpttlr gttgvfvgt ngqdyptllr rsasdvagv atgntasvms
6601 grlsyalgle gpavtidtac ssslvalhwa gralragecd lvvaggvsvm
6651 aspdsvfevs tqgglapdgr ckafsdaadg tawsevgil vlerlsaarr
6701 nghqvlglir gtavnqdgas ngltapngls qqrviaqala darlrpadid
6751 aieahgtggt lgdpiaral itaygrdrda explllgtvk snightqaaa
6801 gaagvikmlm amrhgtlprt lhvgtpsshv dwsggtvall ddarpwprtq
6851 qprragvsaf gvsgtnahvv veqapeteap aapaaepape atptvvpwv
6901 sgrsrealqa qldrltahta ahparsaadv grslatdrtl fphravllag
6951 pdgvreaara aaprtpgrta flfsgggaqh almghdlyqr fpvyadaltd
7001 vlaqfdtvld vplraalfaa pgtpeaalld qtgftqpalf avevalfrla
7051 eswrltpdfv aghsigei

10

SEQ ID NO: 21

Glu(Asp)Leu Gly Phe(Leu, Val) Asp Ser Leu

SEQ ID NO. 22

GAG/C CTG/C GGC/G T/CTG/C GAC TCC/G CTG/C

20

SEQ ID NO. 23

Val Asp Thr Ala Cys Ser Ser

SEQ ID NO. 24

G/CGA G/CGA G/ACA/ G/CGC C/GGT GTC G/CAC

SEQ ID NO. 25

30

GTTGGTACCCCACTCCCGGTCCGCAC

SEQ ID NO. 26

CCAGCCGCATGCACCACC

SEO ID NO. 27

CCG CGT CGG ATC CGC CGA C

SEO ID NO. 28

AGC CTT CGA ATT CGG CGC C

SEO ID NO: 29

10

DNA sequence of ERD48.seq

1 gatccgcccga cgacaccggc cgccgcaccg tcaccgtcca cgcccgcgcc
 51 gacgacaccg ccgaccgcac ctggacgctg cagcccaccg gtgtgctcgc
 101 caccacgcca ccggccgcgc cggcgttcga caccacggtc tggccgcgcc
 151 ccgacgcccga accctcacc accgacgact gctacgcaca cttcaccacc
 201 caccgcttcg cctacggccc cgccttcag ggctgcggg ccgcctggcg
 251 cgccggcgac gtgctgtacg ccgaggtcgc cctgccggag tccgccaccg
 301 acgaagcggc cgccttcggc ctgcaccggc cgctcctgga cgccggcctg
 351 cagccgcgcg tcctcgcgca cgaccgcgac accggactcc cgttctctg
 401 ggaaggcgtc actctgcacg cctccggcgc caccgcgcta cgcgtccggc
 451 tcgccccgaa cggccccaac ggctgttcg tcaccgcgcg cgaccgggcc
 501 ggcaaccccg tcgccaccgt caccgcctg ctgcgccgc ccctggacgc
 551 cgagcagttg accatccaca gcgcctgac ccgacgcgc cttctccacc
 601 tggactggac cccggtcccg cttcccgaca ccgccaactc cgcgcgcgcg
 651 gccctcctcg gcccggaac cgcgtgtc gccgacgccc tcggcgaccc
 701 ggccgtcgca cgccacgcaa ccctcgacga cctcctggcc ggggacacca
 751 ccccgccgc caggtctc gtccccctcg gcgccccact cgacggcgac
 801 accgcgcagc acgcgcacgc cctcaccgc agcgcgctga cctcgtcca
 851 gcagtggctc gccaccgacc gcctcgccga ctcgcgctg gtcttcgtca
 901 cccacggagc cgtcgccacc gacgacgcgc ccccaccga cctggccgc
 951 gccgcggctc ggggcctgat ccgctccgcg cagaccgaga accccggcac
 1001 cttcaccctc ctgcacctg acaccgagcc cgactcgacc accgcgctca
 1051 gccgcgcctt gaccctcgac gaaccacagc tcctctctcg cgccggccgc
 1101 gccgcgcgc cccgcctcac ccgcacccc gcccccacca ccaccacca
 1151 cagccgctgg tcgcggacg gaacggtgtt ggtgacgggt ggtacgggtg
 1201 gtctgggtgg gttggtggcc cggcatctgg tgcggtcgtg tggggtgcg

20

30

1251 catattgttgt tgaccagtcg ttctggtgtg ggtgctgcgg gtgcggccgg
 1301 gttggtcgcg gagttggagt cgttgggcgc gcgggttgtg gttgcggcgt
 1351 gtgatgtggg tgatggtcgc gctgttgccg agttggttgc cgggtgtgtc
 1401 gagtcgtatc cgttgctctgc ggtggtgcat gcggctggtg tgttggtatga
 1451 cgggtgtggtg ggttcgttga cgcggagcgc gttggctgcg gtgttgctgc
 1501 cgaagggtgga tgggtgcgtg aacctgcatg aggcgacgcg tggctctggat
 1551 ctggacgcgt ttgttgtctt ctgctctgtt gcgggtgtgt tcgggggtgc
 1601 gggtcaggcc aactatgcgg cgggtaatgc gtttttgac gcgttgatgg
 1651 ttcatcgggt ggctggtggg ttgectggtg tgcgttggc gtggggtgct
 1701 tgggatcagg gtgtggggat gacggcgggg ctgacggagc gggatgtccg
 1751 tcgtgctgct gagtcgggta tgccgttgtt gacggttgat cagggtgtgg
 1801 cgttgttcga tgcggcgttg gcgacgggga gtgccgcgtt ggtgccggtc
 1851 cgtctggacc tggccgcact gcgcaccccg ggcgacatcg caccgctcct
 1901 ccgcggcctg gtcaaggcgc ccatccgcgc gcgagccgcc accacacccg
 1951 gcgacaccgc actcgcgcag cagctacccc ggctccagcg cgcgagcga
 2001 cgggacaccc tcctcgcgtc cgtccgcgac caggccgcga tggctcctcg
 2051 ccacacctcg ggcgacggcg tcgacccgtc ccgcgccttc cgcgacctcg
 2101 gcttcgactc gctcaccgcg gtcgaactcc gcaaccgcat cggcgcgggc
 2151 accggcctgc ggctaccggc caccggcgtc ttcgactacc ccacgcgcga
 2201 tgccctcgcc gcacacctgc tcaccgaact gctcggcccc gacgcgcagt
 2251 cggacccccg cgagcccggc gacccaccg cgggaaccgac cgacgacccc
 2301 atcgtcatca tcggcatgag ctgccgcttc cccggcgaca tcggctcgcc
 2351 ggaggacctg tggcgctcgc tcggcgacgg cgcgacgtc gtcaccgact
 2401 tcccgaacaa ccgcggctgg gacctggaca acctctacga ccccgacccc
 2451 gcgcacgcgc gcacctcgta cgcgcgacc ggcggtttcc tgcacgacgc
 2501 cgccgacttc gacgcgcact tcttcggcat gagccccgc gaggcctatg
 2551 ccacggactc ccagcagcgc ctgctgttgg agtcctctg ggaggcgatc
 2601 gagcggggcg gcatcgaccc gctgacctg cgcgacagcc gcaccggcgt
 2651 cttcgccggc gtcattgtaca gcggctacgg caccgcctc gacggcgccg

10

20

30

SEQ ID NO: 30

Translation product of SEQ ID NO. 29 (ERD48.seq)

1 saddtgrrtv tvharpddta drtwlhatg vlattpaaa afdtvwp
 51 daeplttddc yahftthrfa ygpafqglra awragdvlya evalpesatd
 101 eaaafglhpa lldaglhlaal laddrdtglp fswegvtlha sgatalrvrl

151 apngpnnglsv taadpagnpv atvtrllarp ldaeqltihs altrdalfhl
 201 dwtpvplpdt ansappallg pdtavladal gdpavarhat lddllagdt
 251 ppatvlpplg apldgdtagh ahaltrsalt lvqqwlatdr ladsrlvft
 301 hgavatddap ptdlaaaavw glirsagten pgtftlldld tepdsttals
 351 raltldepql llragraraa rltrtpaptt tthtpwsadg tvlvtggtgg
 401 lgglvarhlv rscgvrhlld tsrsgvgaag aaglvaeles lgarvvvaac
 451 dvgdgsavae lvagvsesyp lsavvhaagv lddgvvgsit perlaavlrp
 501 kvdgawnlhe atrgldldaf vvfssvagvf ggagqanyaa gnafldalmv
 551 hrvagglpgv slawgawdgg vgmtaglter dvrraesgm plltvdqgva
 601 lfdaalatgs aalvpvrldl aalrtrgdia pllrglvkap irraaattpg
 651 dtglaeqltr lgraerrdtl lalvrdqaam vlightsgdgv dpsrafrdlg
 701 fdsltavelr nrigaatglr lpatavfdyp tadalaahl tellgpdaes
 751 dpdepdgpta gptddpivii gmscrfgpdi gspedlwrll gdgadvvtdf
 801 ptnrgwldn lydpdpahag tsyartggfl hdaadfdad fgmspreama
 851 tdsqqrllle ssweaierag idpltlrdsr tgvfagvmys gygtrldga

10

SEO ID NO. 31

CGCCGCATGCTGTTCTCACCCCACGT

20

SEO ID NO. 32

GGCGCGACCCGGTTCGGCCT

SEO ID NO. 33

GCGAGCGGCCGCTTCACCCCGCAACTCA

SEO ID NO. 34

CGCGAAGCTTGGCCGACTGCTCGACGTC

30

SEO ID NO. 35DNA sequence - 125401 bps (entire cluster):

1 gaattcggat cgatgctgct cgtggtgtgt gccgcgggag gtgccgacgg

51 tgcggtgcgt gaatcggggc gtggtcgccg tcgcgatcgc ggtggtgtcg
 101 gtgctgctcg tctgaccgt ggtggtggcg acggcctact tgcgtagac
 151 ccggtccttg cgcgcgacc tgggcagcgg atgcacctct tcggtcgacg
 201 ccaaggggcg gtggcgagg tcgagatagg cgcggacgga gacccgggc
 251 tgccccggcg tctgccaggg cgcggagggg ccgctggaac gccacagacg
 301 gcccgctcg gcccgggcaa ggtggcggt ggcttcgtcg gcggcggtga
 351 cgtcgatgac gaccatgccg ggttcgtga catgttctc gtggatatgc
 401 acagtctgt atgacctatt tcgccccgtg gcgtaaggag tagccggcag
 451 gtttcatccg aagggtgggc gcgggagcgg cgagttgacc gtgaggtgcg
 501 tcggtgttc ccggtcgcg cggggggtgg gggccggcag gtggaccgcc
 551 gtggtcagca gggcggggt gctgtgccag cggccgggga agaccgtgac
 601 ccggcgaccg ccagcaggg gggccgggac caggagtcgg gcggtgaagg
 651 cgccctgtgt cgggtcgatg acgatctcgg cctcggagaa gtccagtctg
 701 cgtgggtca ggggatacca cgcctgaag acgctctct tggcgctgaa
 751 gagcagccgg tcccagtga cgtccggacg atgggcccgc agggccacga
 801 ggtgcggcgg ttcgaggggc agggcgatgg cgttcaggac gccggccggc
 851 agcggaccgt tcggttcggc gtcgatgctc accgcggccg acagctccgc
 901 gggggagacg gcggcgccgc ggtagccggc gcagtgcgtc atgctgccga
 951 cgatgcccg cggccactgc ggggcgccgc gccgattggg cagtatggcc
 1001 cggtcgggt gggcgagccg gcgcagggcc cggcgggcga gatggcgac
 1051 ggtggtgaac tcgcgctgcc gggactccac gggccgggcg atgacctgc
 1101 gtcccgaggg gagcggccgg tcgccccggc gcggccggtc gtcgtacgcc
 1151 gcctcggtgg cgaccgtggc gggcaggatc agttcgatca ccgcatacct
 1201 ccggcgaaac gtaagaacag ggggttgctt cggggcacgg ttccgtcctt
 1251 gacggggcgc cgtggggcgc cgggtcagcc ggcgcggcg ccgcgggtgg
 1301 gagtgtgggt gcgggctgcg tcgagccggg cgtacaggcc ctgcgcgcac
 1351 aggagttgat cgtgggtgcc ctgctcgacg atgcggccgg cgtccatcac
 1401 gacgatcagg tccgcgtcgc ggatcgtgga caggcggtgc gcgatcacga
 1451 agctggtcgg gcccgcccgc agggagttca tggcgcttg gatcaggacc
 1501 tcggtccggg tgtccacgga gctggtggcc tcgtccagca cgaggacggc
 1551 cggctctggc aggaaggccc gggccacggt cagcagctgc ttctcgccg
 1601 cgtgacggg gcccgactcg tcgtccagca cgggtctgta gccctgcggc
 1651 agggtgcgga tgaagcggtc ggcacaggtc gcgcgggccc cctcctcgat
 1701 gtccgcacgg caggcacccg gtgcgccgta cgcgatgttc tcccgatgg
 1751 tgccgccgaa cagccaggtg tcctggagca ccagcccga gcgggaccgc
 1801 aggtcgtcgc ggtcatcgt cgcggtgtcg gtgccgtcca ggaggatgcg
 1851 gccggagtcg ggttcgtaga agcgcacag gaggttgccg aggggtggtt

10

20

30

1901 tgccggcgcc ggtggggccg acgatcgcca cgggtgctgcc cggttccacg
 1951 gtcagcgaga ggttctcgat gaggggcggtg tcgggggagt agcgaagga
 2001 cacgtcgtg aactcgacgc ggccctcggc gcgggcgggc gtgccgggcc
 2051 ggagcgggtc cggggcctgc tcggggcggt cgaacagggt gaagacgcgc
 2101 tgggcggagg cgaacccgga ctggagccgg cccgccaccg aggcgatctc
 2151 cacgatcggc tggctgaact ggcgggcgta gaggatgaac gcctgcacgt
 2201 caccgagggt cagggtgccc tttatgacct tcaggcgcc gatgacggcc
 2251 accagcacat agccgaggtt ggcgaagaa atcatgacc gttccatggc
 2301 accggaggcg aactgcgcct tggccgcagc ccggtagacc gcgtcgttgc
 2351 aggcgtcgaa gcgctcctcg gccgcgcgc gccggtcgaa gcccttgatc
 2401 agcgcacgac cgggtgcacac ctctccaca tgggcgttga gggtgccgtt
 2451 cgcggaccac tgcggggcgt agtggggctg cgcgcgctt ctgatccggg
 2501 ccgcgatcaa cgcgagacc ggcacgctga gcagcatcac caccgccagc
 2551 gacggcgaga tcaccagcat cagcaccagc atcgtcaaca gcgagaagat
 2601 cgaggatgac agctcggcga gggctcgtg gagggtctgt tggaggttgt
 2651 cgatgtcgtt ggtggtgccc ctgagcagct caccggccgg ctgccggtcg
 2701 aagtggcgca cgcgcagccg ggtcagcttc tcccgggcgt cgcggcgag
 2751 ttcgtggatg gtgcgccaca ccgcggacgc caccagccgg cctgcgcca
 2801 gcatgaacag cgcgcaccg acgtagagcg ccagcaggac cagcagcagc
 2851 cggccgatcg cgcgaagtc gatcccgggc gccggggccc ggacgccgcc
 2901 gacacgcgc tcggcgatca gatcggtagc ccggccgagc agcagcgggc
 2951 cgaacgcgtt gacacgac ccgcgcagc ccacgacac ggcaggtgcc
 3001 acggagcggc ggtgcggagc cagcaggccg acgagccgac gtgccggccg
 3051 gggcgcggtc cgtcctcct caaggctcgt cggcgaggcc atggcgggct
 3101 tcctcctcgg tcagctgcga gacgcgac tcgcggtagg tcgggtcgtt
 3151 gcgcagcagc acgtcgtggg tgccctgggc gaccacgcgc ccgcgggtca
 3201 gcaccacgat ccggtccgcg tcgcggccgg cggagatccg ctgcgccacg
 3251 gtgatcaccg tggcgccgc ggtgtacggc accagcgcg tcgcgagcg
 3301 cgcacgggtg gcctggtcga gcgcgagaa acagtcgtcg aagagataga
 3351 tctccggccg gcgcagcaac gcccgggcga gggacaggcg ttggcgctgg
 3401 ccgcgggaga cattgccgcc gccctgggtg atctccggt cgaggccgtc
 3451 cggcatccgc gccacgaagt cggccgctg ggcgacccgc agcgccctcc
 3501 acagctcctc tgcggtggcg tcgggcccgc cgaagcgag attgctcgc
 3551 acggtgccgc agaacaggt cggccgctg ggcacgaacc cgcggcggc
 3601 ggcgagcgtg gccgcggtca gtcgcggac gtcggtgccc ccgacccgca
 3651 ccgcgcctc ggtggcgctg gccagccgca gcaccaagtt caacagggtc
 3701 gtcttgccgc tgcgggtgct gccgagcacg gcgatccgt cggccggctc

10

20

30

3751 gacgggtcagg tcgacgtccc gcagcacggg ctctctggcg cccgggtagc
 3801 ggtaccgggc cgcgcacagt tcgatccggc cggcggggccc gcgcaccggc
 3851 tgccggcgcg cccggggcgc cagctcgcac ccggtgtcca ggacctcgc
 3901 gatccggcgc gcacagaccc gggcccgcgc caccgacagg aacacgaagg
 3951 cgagcatcac gacggacatc aggatcagcg agagatagct caggagggcg
 4001 ctgagcgagc cgtacggcat ccggcccgcg tcgatccggt gggagccggt
 4051 ccacagcagg gctacgggtg aaccgttcat cagcagcagc acgaccggca
 4101 gcatcgtcgc gatcagccga cccaccggc gcgacaccac gaggaacgcg
 4151 tcgtttggtct gcgcgaaccg cgcgcgctcg tggctcgtcg ggacgaaggga
 4201 ccggaccacc cgcaccccgc tgatcgctc gcgcagcagc cggcccagcc
 4251 ggtccagggt cagctgcac cgcgcgtaca gggtgcccat cggggccagc
 4301 agcaggccga agcagaccgc caccaccagc accagcgcca ccagcagcag
 4351 tgccagcggg acgtcctggc gcagcgccag cagcacgctg cccaggcaca
 4401 tcagcggcgc gcagacgacg atgccgaagc cggctcgggc gaggttctgc
 4451 acctgctgca cgtcgttcac cgaccgggtc agcagggagg gggtgccgaa
 4501 ccggccgcat tcgcggggcg agaagtccag gatgcggcgc aagagcgcgc
 4551 accgcagatc gcggcccacg gccgtcgcgc tccggggcgc cagcgcggcc
 4601 gcaccgagcg cggcggcgat ctgcaccagc gccaccacgc ccatcaccac
 4651 acccagctcg gtgatgcgcc cgcgctcgc gcgcaccacc cctgtgtcga
 4701 taagtgcggc gcccaagctc ggcagcagca aagtgcacag gatctggacg
 4751 agttgaaggc cgacgagagc ggcgggtggc caggcgtagg ggcgcagctg
 4801 tgcccgcaga agtctcaaca gcacggagga acaccccggt tgacggcggc
 4851 gtgcggccgc gcgggtggac gagtcggccg gcccgccgac ccggtcattg
 4901 caccgcagtc cgtcccgaaa ttctactagt gttgggggtg gcaacggact
 4951 tttgacggcc cgttactcgt gaattcccta gaaagcccgc gatgcgttga
 5001 cagcatcttc cgcggcttgc gagcgtgcgc gtgtctatcg tccggactgc
 5051 gtattcgacg ccaggctcgt gccgagctca agcgtttgga cggctctctt
 5101 cgtgttcgag aagggtttgc catgtccaaa cgagcgctga tcaccggaat
 5151 caccggccag gacggctcct atctcgcgga gcacctgctg tcccagggt
 5201 accagggtgt gggctctgat cgcggccagg ccaatccccg caagtcccg
 5251 gtcagccgcc tcgctccga actcgacttc atcgacggg acctgatgga
 5301 ccagggcagc ctggtctccg ccgtcgacac cgtgcagccc gacgaggtct
 5351 acaacctcgc cgcctctcgc ttctgtccga tgtcctggca gcaggccgag
 5401 ctggtcaccg aggtcaacgc catgggcgtg ctgcgcctgc tggaaagcat
 5451 ccgcatggtc agcggactgt ccacctccg cccggtcagc ccgcgcggcc
 5501 agatccgctt ctaccagcg tccagctcgc agatgttcgc caaggccgc
 5551 gagacgccgc agcgcgagac cccctcttc caccgcgca gccctacgg

10

20

30

5601 cgcggaag ggcgtacgggc actacatcac ccgcaactac cgcgagtcct
 5651 tcggcatgta cgcggtctcc ggcatgctct tcaaccacga atccccgcgc
 5701 cgcggccagg aattcgtcac ccgcaagatc agcctggcgg tcgcccgcac
 5751 caagcaaggc ctccaggaca agctggcact cggcaacctc gacgcggtgc
 5801 ggcactgggg ctatgcgggc gactacgtcc gcgccatgca cctgatgctc
 5851 cagcaggacg ccggcgacga ctacgtcatc ggcaccgggc agatgcactc
 5901 ggtgcgcgac gcggttcgga tcggttcga acacgtcggc ctgaactggg
 5951 aggactacgt cgtcatcgac cccgaacctg tgcggccgc cgaggtcgag
 6001 gtgctgtgcg ccgacagcgc caaggcccag gaccgcctcg gctggaagcc
 6051 ggacgtcgac tccccaccc tcattgcgcac gatggtcgat tccgacctgg
 6101 cgcaggtttc ccgcgaaaac caatacggcg acgtgctgct cgcgcacaac
 6151 tggtagcagt tctcaagctt tcgaaaacta gtgaattcct gccggaattc
 6201 cgacgacact cgcacatgg attcccccg tgagtggcga atccagggtg
 6251 cgaatccgaa cgtaccgcgc aaacggcgtg gagaagtcgg acgccattca
 6301 cgtgcggccg tcggtgtcg agatgggttg agttgagatg gacaacgaac
 6351 aaaaactccg ggattacctc aagcttgcca cggccgacct tcgacgcacc
 6401 cgcgcggcgc tccacaagct ggagtcggcg gcccggaac cgttgccat
 6451 catcgccatg acctgtcgt acctcgcgcg cgtccgcagc cccgaagacc
 6501 tctggcgcat ggtcgaggcc ggcgagcacg gcgtcaccac gttccccacc
 6551 gaccgcggtt gggacctgga ggcgtgccc gccgcgcga ccgcctccgg
 6601 cggattcctg cagcagcac ccgacttga cgcggacttc ttggcatct
 6651 cgcgcgcga ggcggtcgcc atggaccgc aacagcgcgt cgtcctgga
 6701 tccgcctggg aggcgttcga acgcgcgcgc atcgaccga cgtcctgga
 6751 gggcagccgc accggagtct tcacggcgc gatggcccag gactaccggg
 6801 tcggcccccgc cgcggcgcc gagggcttcc aactcaccgg caacaccggc
 6851 agcgtgctgt ccggccgcat ctctacacc ttcggcacgg tcggccccgc
 6901 cgtcacgctc gacaccgcct gctcctctc cctcgtcgc gtccacctcg
 6951 ccaccagggc gctgcgggcc ggcgagtga cctcgcct cgcggcgcc
 7001 gtcaccatca tgtccggccc cggcaccttc atcgaaatgg gccgccaggg
 7051 cgggctctcc gccgacggcc gctgcgcctc ctccggcgac accgccgacg
 7101 gcaccggctg ggcgaaggc gtcggcatcc tcgtcctgga acggtgtcc
 7151 gacgcgctcc gcaacggcca cgagatctc gccgtcgtcc ggcgcaccgc
 7201 cgtcaaccag gacggcgcc ccaacggcct gaccgcccc aacggcccc
 7251 cccagcagca ggtcatccag caggccctgg tcaacgcccg actcgccgc
 7301 ggggacatcg acgtcgtcga ggcgcacggc accggacca cctcggcga
 7351 ccccgctcag gcccgagccc tgcctgcac ctacgggcag aaccgcccgg
 7401 cggaccggcc gctgctgtg ggctcggtca agtccacct cagccacacc

10

20

30

7451 caggccgccc cggcgctcgc cggcgtgatc aagatggtca tggcgtatgcg
7501 gcacgggcacc ctgccgcgca cctgcacgc cgaggagccc acccaccacg
7551 tcgactggtc gcaggggcgc gtgcggctgc tgaccgacac caccgaactgg
7601 cccgccaccg gggcgccgcg ccgcgcgcgc gtctctctct tcggcatcag
7651 cggcaccaaac gcccaaccca tcctcgagca ggccccgaa ccgcagcccg
7701 aggaagccgc gaccgcgcag gacgacgccg ccggcagcac gccggccacc
7751 gcccccgtag tgccccggt cgtaccggtc ctgctctccg gccgcacccc
7801 ggacgcctc cgcggccag ccgcggccct gcgcgcgcgc ctcgacaccg
7851 gccggcgccc cgacctgctc gacctcgac actccctcgc caccaccgcg
7901 gccgggttcg agcaccgcgc cgtctctctc gccaccgacc acccgcctct
7951 gaccgacggc ctacccgcc tcgcgcagc cgacgaccgc gccgcgcgc
8001 ccgcctgat caccggcacc acccgggcgc agaccgcgt cgcgtctctg
8051 ttcaccggc agggcgccca acgcctcgc gcgggacggg aactcgccgc
8101 ccgtttcccg gcgttcgcca ccgccctga cgcggcgctc gacgccttca
8151 cccgcacct cgaccgcccc ctgcgcgagc tctgtgggg caccgacgcc
8201 gccttgctcg accgcaccgc atacgccag ccggccctct tcgcgtcga
8251 agtggcgctc taccggctga tcgaatcgtt cggcgctccg cccgaccacc
8301 tcgcgggcca ctccgtcggc gagatcgtc ccgcgcacct cgcggggctc
8351 ctctccctgg ccgaagccgc caccctcgtc gccgccgcgc gtgcctgat
8401 gcaggcgctc cccgacggcg gggcgatgat cgcgtccag gctcgggag
8451 ccgacgtcgc cccgctgctc gccgggcacg aggaccaggc cgcgatcgc
8501 gccgtcaacg gccctcgcgc cgtcgtctc tcggcgccgc aagccaccgt
8551 caccgcgctc gccgaacagc tcgcgcgcga cggccgcaag accgcgcgc
8601 tcgcggtctc gcaagccttc cactcgccgc tcattggagc gatgctcgac
8651 gccttccgcg ccgtcgtcga agacctcacg ctccagccgc cgtctctgc
8701 ggtcgtctcc aacctgaccg gcaagccgc caccgtcgc caactacct
8751 ccgccgacta ctgggtcgac cactcgcgc acgcgtccg ctctgcgcac
8801 ggcactgact ggtcgcgcgc gcacgacacc accgccttc tcgaactcgc
8851 cccgcagcgc gtgctgtccg ccattggcca ggaactgctg gacgcgcgc
8901 acgcagacgc cgtcacctc ccgcctcgc gcgcggggcg ccccgaggag
8951 cacacccca ccaagccct cgcgggtctg cactccacg gcgccacct
9001 ggaactggacc ggtcgtctc ccggcaccg cgcgcgcgc accgacctgc
9051 cgacctacgc ctccagcgc cgcgctact ggcccaaggc cctccagac
9101 ggcaccgcgc acctgcgctc ggtcggcctc ggtgcgcgc accaccgct
9151 gctctccgc gccgtctccc tcgcgcgcgc aggcggcacc ctgctcaccg
9201 gccgcctctc ccggcagacc caccctggc tcgcgcacca caccgtcgc
9251 ggcaccaccc tgcgtcccg taccgcctc ctggaactc ccgtcgcgc

10

20

30

9301 cgccgacgag gtcggctgcg accgcgtcga ggaactcacc ctccgccac
 9351 cgctectgct gcccgacag ggcggcgtcc aggtccagtt gtggatcggc
 9401 aaccccgacg tgtccggtcg ccgcaccgtc aacgtccacg cccgccccga
 9451 caccggcgac gacacccctt ggaccgccca cgcaccggc gtcctacca
 9501 ccgcggacgc ctcccgccag ctcccggtt cgtccgagca gggcggcacc
 9551 cccctcgccg gcgaacccca ccccgccctc gacggggccc agtggccccc
 9601 ggccggcgcc gaaccgctgc cgtcggacgg ccactacgac cgcctcgccg
 9651 acggcggtt cggtacggc ccgtcttcc agggcctcg cgcgccttg
 9701 cgcggcgcg acgtcgtcta cgcggaggtc gagctgcccg aggcggcgcg
 9751 gtccgacgcc gaggcgttcg gctccacccc cgcctgctc gacgcggccc
 9801 tgcacgccgc gcccttcacc ggctcggcg aacgcggccg gggcggtcg
 9851 ccgttctcct gggagggcgt ctccctccac gccggcgcg ccaccacct
 9901 ccgcgtccgc ctgaccccg tgcggacga cgcgtccgc ctgaccgtcg
 9951 ccgacggcac cggcgccccc gtgctgtccg tcgactcgct cgtcctgcgc
 10001 agcgtggcga cccaacagct cgacacggcc gccgcgctcg cccgtgacgc
 10051 cctcttcgc ctgcactgga ccccgctca gccgacgcc accgaccccg
 10101 ggcggctgc cctcctcgcc gccgacccct tggcctgct caccacgcc
 10151 ggattcgccg acgccccggc ataccggac ctgcggccc tgcggcgcc
 10201 ggacggcccg gtcccgacca ccgtcgtgct gtccctcgc gccacgggg
 10251 acgacggcg cgaccggcc cgtccgcac accgctgcgc cgcggaggcc
 10301 ctgcggcccg tacagacctg gctcgaccac catgagcgt tgcggccgc
 10351 ccgcctggtc ttcgtgaccc gtggtgcgac ggtcggcggt gatgttgctg
 10401 ctgctgcggt gtggggtctg gtgcgttcgg cgcagtcgga gaatccggg
 10451 tgttttgctc tggctgatct ggatccgat ggtgcggtg gtgcggctgc
 10501 gctcgtcgt cgttgggtca gtggtgagcc gcagcttcg gtgcggggtg
 10551 atgtgttgcg ggtcgcgcgt ctggtgcggc gccgcctcac cgaggtcggt
 10601 gcgggtgctg atggcaccgg ggatggcgtc ggggtggt ctggtgtgtc
 10651 gttctcggtg gagggtgcg tccgtggtac tgggtgtaag ggtggtctg
 10701 gtgcgggtt gccgcgtcat ctggtggccg agtatgggt gccggatctg
 10751 ctgttggtca tgcgcagtgg tgaacgtgcc gtgggtgctg gggagttggt
 10801 ggcggagctt gcgggtgtgg gtgcggggt gccgggtgtt gcgtgtgatg
 10851 tgaccgatcg tgcccggtg gtggagttg ttggcgggca tgcggtgtcc
 10901 gcgggtggtt atgcggctg tgtgctggat gacggcatgg tgggtgcgtt
 10951 gaccggggag cggttgtccg cgtgctgcg gccgaagggt gatgctgtct
 11001 ggcattctaca tgaggcgacc cgcggcctg acctggacgc gttcgtcgtc
 11051 ttctctccc tcgcgggggt cttcggcagt cccggccagg ccaactacgc
 11101 ggcccggaac gccttcttg accgcgtgat gacgcggcg cggcgagg

10

20

30

11151 gactgccccg cctgtcactc gcatggggac cgtgggagca gtcgggaggga
 11201 atgacgggca ccctgacgga cgtcgacgcc gaacggctgg ccgctccgg
 11251 tgtcccgccg ctctccgtgg cgcagggcct ggccctcttc gacgctgccg
 11301 tggccgggac cgaagccacc tgcgttccgg tccgcttggg cctccccgtc
 11351 ctccgcgcac ggggtgaagt gccgcgctg ctgaggctgt tgatccgct
 11401 ccgggcgcgc cgagccgccc tcgcccggct cgcaccgcg ggcaacctcg
 11451 cccagcgccct gcgcgcctg gacgaggacg gccgcgacga gatggtcctg
 11501 gacctcgtcc gcggtcaggt cgcctcgtc ctcgccaccg cgaccggtgg
 11551 cgacgtcgac gccggccgtg ccttccgcga cctcggcttc gactcgtgga
 11601 ccgccgtcga actgcgcaac cgcctcaaca cgtcaccgg cctgcgcctg
 11651 cccgccaccc tggctcttga ctaccgcacc gtccggcacc tcgccacgta
 11701 cgtcctggac gatttgttgg gcacggatgc cgaggtggcg accgtgcagc
 11751 cggccgcggt tcggttggcg gacgatccga tcgtcatcgt gggcatggcc
 11801 tgccgctacc ccggtggcgt cagctccccc gaggacctgt ggcgctgct
 11851 caccgaaggc accgacgcg tctcgggctt cccgaccaac cgtggttggg
 11901 acgtcgaatc cctctatcac ccggacctg accaccggg tacctcctac
 11951 acgcgctcgg gtgggttctt gcatgaggcg ggggagttcg atccggggtt
 12001 ctccgggatg agtccgcggg aggcgttggc gaccgattcc cagcagcgtt
 12051 tgttgttggg gtcgtcgtgg gaggcgatcg agcgggcccg tattgatccg
 12101 gtgagtttgc ggggttagtc gacgggtgtg ttcgcggggg tgatgtacag
 12151 cgattacagc gcgatgttgg cgaagtcgga gttcgagggt tccagggca
 12201 gtgggagttc gccgagttt gctcgggtc ggggtggccta cacgttgggg
 12251 ttggaaggcc cggcggtgac ggtggatacg gcgtgttcgt cgtcgttgg
 12301 ggcgatgcac tgggcgatgc aggcgttgcg tagtggtagg tgtgggttgg
 12351 cgttggcccg tgggtgtgac gtgatgtcga cgcctgcggt gtttgtggac
 12401 tttgctcggc agcggggtt gtcgcccgat ggccggtgca aggcgtttgc
 12451 ggatgcggcc gatggtgtgg gctggtccga gggcgtcggc gtgttgggtc
 12501 tggagcggca gtcggacgcg gtgcgcaatg gtcacgagat tttggctgtg
 12551 gtgcgggggt cggcggtcaa ccaggatggt gcgtccaatg gtttgacgc
 12601 gcccaatggt ccgtcgcagc agcgggtgat ccggcaggcg ttggccagtg
 12651 gtggtctgac gcccggtgac gtggatgtgg tggaggcgca tggtaagggt
 12701 acgacgctcg gtgatccgat cgaggcgag gcgttgttgg cgacgtatgg
 12751 gcgggatcgt gagcctgagc gccggttgtt gttgggttcg gtgaagtcga
 12801 atctggggca tacgcaggct gctgcgggtg tggcgggct tatcaagatg
 12851 gtgttggcga tcggcatgg tgtggtgccg cggacgttgc atgtggatgc
 12901 gcccttcttc catgtggact ggtccgaggg tgcggtggag ctgctcagtg
 12951 agcaggcggc ctggccggag acgggtcggg tcggcggggc ggggtgtctc

10

20

30

13001 tccttcggca tcagcggtac caatgtgcat gtcctcgtcg agcaggcgcc
 13051 gggcgccaag gcgategccg cggccggtgc ggcggcgccg acgcccggtg
 13101 ccgtgccggt gctgctctcg gggcggtggc ggagtgcctt gcggggccag
 13151 gcgcgccgce tgctcggaca cctccaggcc cgacccgacg ccgaactcgt
 13201 cgatgtcgca ctgtcgttgg cgaccacccg gtcccgttc gagcagcggg
 13251 ccgccgtcgt ggcgcaggac cgcgaccagc tgatcgctc gctggggcg
 13301 ctggccgccc accgcccga ccccgccgtc gtcgagggcg aggcggccgg
 13351 acgcgccgg accgcggtgc tgttcaccgg acagggcagc cagcggccgg
 13401 ccatggggcg tgaactccac gaggtgcagc cggagttcgc cgcggcgctc
 13451 gacgcggtgt gtgccgtttt cgacccgctg ttggaccggc cgtgcgcga
 13501 ggtggtgttc gccgaggacg gcagcgacga ggccgactg ctggacgaga
 13551 ccggttgac gcagccggt ctgttcgccc tcgaggtggc gctgttcgg
 13601 ctggtggaga gttgggtgtt ccgtccggac ttctggccc gccattccat
 13651 cggtgagatc gcggcgccgc acgtcgccgg ggtgctgacg ttggaggacg
 13701 cctgccgtct ggtggccgcg cgggcgacgc tgatgcaggc gctgccgacc
 13751 ggcggcgcca tgatcgcat ccaggccacc gaggacgaga tcgcggcgca
 13801 cctcgacgac acggtggcga tcgccccgt caacgggccg cagtccgtgg
 13851 tgatctccgg tgacgaggag gccgcgaaa cgatcgccc cacttcgcc
 13901 gaacgcgggc gcaagaccaa gcggctgccc gtgagccatg ccttcactc
 13951 gccgcggatg gacgggatgc tggacgcctt ccggtatgct gccgagggc
 14001 tgacctaccg ggcgcgcgc atcccgtctg totccgacct caccggccgg
 14051 cgcgccgacg atgcggaggt gtgcaccgcg gagtactggg tgcggcacgt
 14101 ccgagaggcc gtgcggttcg ccgactgctg gcggacgctg cgcgacgcc
 14151 gggccaccac ctctctggaa ctgggctccg acggcctgct gaccgcatg
 14201 gccgaggaca ccctcgttga cgaccacgac gccgaactgg tgcgatgct
 14251 gcgcgcggg cgcgcggagg aactggccgc ggccaccgcc ctggcccgc
 14301 tccaggtgcg cggcggtggc gtggactggg cggcgctacct cgcggcacc
 14351 ggcgcccgc gcaccgacct gccgacctac gccttcacg acgcgtacta
 14401 ctggccgcag ctgccgacc cggccgcgc cctcgcgcc gccgatccg
 14451 ccgaccagca gctgtgggccc gctgtggagc gcggcgacgc ccgcgaactc
 14501 gccgacatcc tcggcctggg cgaacaggac ctacgcgcg tggactccct
 14551 gctgcccgcc ctacactcgt ggcggcgccg caaccaggag aagcacctc
 14601 tggacaccct gcgtaccgc gtggagtggc caogactgag caagccgacc
 14651 gcccgggtcc tcgacggcac ctggctgctg gtgcctccg acgccaccgc
 14701 ggccgaccag ccagccctcc tcgacggcct ggccgacgcc ctcggtctgc
 14751 acggcgccg ggtgcgtcgc ctgcttctgg acgactcctg cgcggaccgc
 14801 gcggtgctcg ccgaacgact ggcgcggacc gccgacgtgg acgcgcgac

10

20

30

14851 ccaggtgctg tccgtgctgc cgctcgacga gcgggacgcc gacgactgcc
 14901 cgccgctcac ccgcggaactg gcgctgaccg tcgcgctcgt ccaggccctc
 14951 gccgacaccc gcgcccaggc ccggctgtgg accgccaccc gcggcgccgt
 15001 ctccaccaac ccgcccgaac cggtcaccca ccccgtccag gccgctgcct
 15051 ggggctctgg ccggggcgct gccctggagc acccacggct gtggggcggc
 15101 ctggctgacc tgccgcaggc cttcgacgag cgggccggac agcggctcgc
 15151 cgggatccct gccgtcaagg acgcaccgga cggcgaggac cagggtggcg
 15201 tgccggccac cggagtctcc ggccgcccgc tcgtccgcca caecgtcgaa
 15251 gcgctgcccc cggcccgcca gtccaccgcc accggcactg tccgtatcac
 15301 tggtggcacc ggtggcctgg gcgcccaggc cggccggtgg ctggcccgcg
 15351 ccggcgcccc gacccctcgc ctgaccagcc gccgcccgcc ggacgcgcgc
 15401 ggccgcccgc aactccgggc cgaactggag ggctacgggc cgtcgggtgc
 15451 cgtcgtcgcc tgcgacgtcg ccgaccggga cgcgctcgcc gccgtcctca
 15501 ccgcaactgc cgaggaactg ccgctgaccg gtgtcgtgca caccgcaggc
 15551 gtccggccact accgcccgcct ggacaccctg agcacccgcc agttcccgcg
 15601 cctcacccgc gccaaagctcg ccggcgccgc ccacctcgac gccctgctcg
 15651 ccgaccgcga actggacttc ttcgctctct tcggctccat cgcgggtgct
 15701 tggggcagtg gcaaccagag cgcctacggc gccgccaacg cctacctcga
 15751 cgcgctcgcc ctgcaccgcc gcgcccgcgg cctcgccgcg acctccgtcg
 15801 cctggggccc gtgggcccga gccggcatgg ccgcccagca tgccgtttcc
 15851 gagaccctgc gccgcccagg cctcggcctg ctcgaccggc ccccgcccat
 15901 gaccgagctg cgcgcgcgcg tcgtccggca ggacgtcacc gtcaccgtcg
 15951 ccgacgtgga ctggcagcgc tacgcaccgc tgttcacctc cgcgcggccc
 16001 agcgccctga tcgcccgcct gcccgaggtc cgcgccctcg ccgcccagca
 16051 gcgcaccgag caggacgcca ccggcgccct cagggtcgtc acccgctcc
 16101 gcgcccctgg cgaaccggag caactgcgcc tgctgaccga cctcgtccgc
 16151 accgagtcgc ccaecgtcct cggccacagc tcgcccagcg ccgtgcccga
 16201 gggccgcgcc ttccgcgacg tcggcttcga ctcgctgacc gcggtcgagc
 16251 tcgccaagcg cctgggcgcc gcgaccgggc tgtccctgcc cagcaccatg
 16301 gtcttcgact acccgacacc gctggaactc gccagttacc tgcgggcgga
 16351 gatcctcgcc gcggtgctgg aagtcgcccg cccggtcgcc accggcgccg
 16401 ccgacgacga gccgatcgcc atcctcgcca tggcctgcc ctcccccggc
 16451 ggcgtcagct ccccggaaca gctgtgggac ctggtcgcc ctcccccggc
 16501 cgcgatcagc gagttccccc tcaaccgcgg ctggcagacc gggcacctct
 16551 tcgaccggga ccccgaccgg ccgggacca cctactccac ccaggcgccg
 16601 ttectccacg aggcgcgacg gtccgacccc accttcttcg gcctctcgcc
 16651 ccgcgaggcg ctggtcatgg acccgacgca cgggctcctg ctggagacca

10

20

30

16701 cctgggagtc ctctgagcgc gccgggatcc gcccggaac cctccgatcc
 16751 accctgaccg gcaccttcgt cggctccagc taccaggagt acggcctggg
 16801 cgcgggcgac ggcaccgagg gccacatggt caccggcagc agccccagt
 16851 tgctctccgg ccgactgtcg tacgtcttcg gtctggaagg cccggcggtc
 16901 accgtcgaca ccgctgctc gtctctgctc gtggcgctgc acctggcctg
 16951 ccagtcgctg cgcaacggcg agagcaacct ggccgtcgcc ggcggcgcca
 17001 cgatcatgac gacgccaac ccgttcacgc cgttcagccg gcagcgcgcc
 17051 ctcgccaagg acggccgctg caaggcgctc tccgacgacg cggacggcat
 17101 gacgctcgcc gaggcgctcg gcgtcgctct cgtcgagcgg ctctccgacg
 17151 cgcagcgcaa cggccaccgc gtcttgcccg tctcccgagg ctccgccatc
 17201 aaccaggacg gcgctccaa cggcctgacc gcgccaacg gcccgctcca
 17251 gcagagggtc atccgccagg ccttcgcca cgcgcgcctc gcgcccggg
 17301 acatcgacgc cctggaggcg caccggcacc gcacaccgct cggcgacccc
 17351 atcgaggccc aggcactgtt cgccacctac ggccgcgacc gcgacccga
 17401 gagcgcgctg ctgctcggtc cgggtgaagc caacatcggc cacaccagt
 17451 ccgcccgggg catcgccagc gtcataaga tggatcatgc gctgcgccac
 17501 tccgaactgc cgcgacccct gcacgcccgc gcgcgctct cgcacgtgga
 17551 ctggctcgcc gggacggctc ggtgctgac ccaggcggcg gcctggccgg
 17601 agaccggtcg ccgcgcggg gccgcggtgt cctcgctcgg catcagcggt
 17651 accaacgccc atgtctgtgt ggagcaggcg cccgtcgagg acaccccggc
 17701 cgaggagcgg ccgcgcgtgg cgcgcgtccc gatcgccgcc ggcgctgctc
 17751 cgtgggtggt caccgcccgc agcgccgcgc cctgcgcgg ccaggccgag
 17801 cgcctctctc cgcacgcga aaccgtcgga accgcccgtc cggccgcgg
 17851 accgctcgac atcggcctgt cgtggtctc cgcgcgcgc cgtttcgagc
 17901 accgtgccgt cgtcgctccc cccgcgggca ccgacccgct ggccgcgtc
 17951 cgcgcgctcg cgacggacgg gccctcgccc gtggtcgccc gtggcggtgc
 18001 ggacgtcgag ggtcggaagg tgttcgtgtt ccccggtcag ggttcgcagt
 18051 ggggtgggat ggggtcccaa ctctttgatg agtcggcggt gttcgcgag
 18101 cggattgccc agtgtcggc ggcactcgcc gagttcaccc actggtcgct
 18151 ggtcgatgtg ctgcggggtg tgggtgggtg gccgtcgctg gagcggtcg
 18201 atgtggtgca gccgcgctcg ttcgcggtga tgggtgctgt ggctgcgttg
 18251 tggcgctccc gtggtgtgtt gccggatgag gtggtggggc attcgaggg
 18301 tgagatcgct gctgcggtgg tgtcggtgct gctgtcgctg cgggacgggg
 18351 cgcgggtggt ggcgctcgcg agtcaggcca ttggtcgctg gttggcggg
 18401 cggggaggga tgatgtccgt cgcgctgtcg gtggacgtgc tcgaaccgag
 18451 gttggtcgag ttcgaggggc ggggtgtcgt ggccgcgctc aacggcccgc
 18501 gctccgctgt ggtcgccggc gagcccgagg cgtggacgc gctgcacgcc

10

20

30

18551 cggtgaccg ccgacgacat cggggcccgc cggatcgcgg tggactacgc
 18601 .ctcgactcg caccaggctcg aggacctgca cgaggaactg ctggagggtgc
 18651 tggcggagct ggccgcgcgc acgtcggagg tgcggttctt ctcgaccgtg
 18701 accggcgact ggtcggacac cgcgcggatg gacgcgggt actggttcgc
 18751 caacctgcgc ggacgggtgc ggttcgcgga cgcggtggcg gacctgctgc
 18801 cggcggagta ccgcgcattc gtcgaggcca gtcgcaccc ggtgctgtcg
 18851 atggcgggtgc aggaggcgat cgacgaggcc ggctgcccgc ccgtcgcgc
 18901 cggcacccctg cgcgcgcacc agggcggcac cgaccgcttc ctgctgtcgc
 18951 ccgcggaggt ctctgtgcgc ggtgtggacg tggactgggc ggggtgttc
 19001 gaggggaccg gtgcgtcccg gatcgacctg cccacctacg ccttccaaca
 19051 cgaacacctg tgggcccgtc cgcgcgcgc ggaggccgtc gccgcgcgc
 19101 acccggacga cgcggccttc tggaccgcgg tcgaggacgg tgacgtctcc
 19151 gcgtgaccg ccgcgctcg caccgacgag gactcgtcg ccgcggtgt
 19201 gccgcgcctg acctcctggc gccgggcccgc ccgcgaccgc tccaccgtgg
 19251 acgctggcg ctaccgcgtc gcctggaac cctcggcg gcacctgcgc
 19301 caccgctccc tgaccggcac ctggtgtctg gtcaccgccc acggcatcga
 19351 cgacaccgat gtggcagggg cgttgagac ctacggcgcc gaggtgcgc
 19401 ggctggtcct ggacgaggag tgcgtcgacc gcgcgctct ccgggagcgg
 19451 ctggccggcg cggaggacgt gaccggcacc gtctccgtcc tcgcgcgcgc
 19501 cgagcggacc gacgcggtac cgggcacctc cctggtgttc ggacccgcgc
 19551 tgaccgtggc actgatccag gcctcggcg acgcgaaat cgacgctccc
 19601 gtatgggctg tgaccgcgcg cgcggtctcc accggccggg ccgacgagct
 19651 gaccgcgcgc gtccaggcac aggtcaccgg catcggtgg accgcggcg
 19701 tggagcacc gcagcgtcg ggcggcacc tcgacctgcc cgcgcctc
 19751 gacgcgcggg ccgcaccagc gctgcgcgc gtgtgtctcg gcgcctcgc
 19801 cagcgacgac cagctggcca tccggcctc cggggtcttc accgcgcga
 19851 tcgtgcgggc cgaggccacc gccgggcggc ccgcgcgac ctggacgcgc
 19901 cgcggcacca cactggtcac cggcggtcc gccacctcg cccgcacct
 19951 cgcgcgctg ctggcccaac gcggcgccga gcacctggtc ctgatcagcc
 20001 ggcgcggcac ggccgccccg ggcgcgcgc aactcgtcg ggaactggcc
 20051 gagtcgggca ccgaggcgac cgtgcgcgc tgcgacatca ccgaccgga
 20101 cgcggtcgcc gcgtgctgg ccgacctca ggcgcaggg cgcacctcc
 20151 gcacctgctg gcacaccgcc gccaccatcg agctgcacac cctggacgcc
 20201 accacctgg cggacttcga ccgggtgtcg cagccaagg tcaccggcg
 20251 ccaggctctc gccgaactgc tcgacgacga agagctggac gacttcgtcc
 20301 tgtactctc caccgcggc atgtgggga gcggcgcca cgcgcctac
 20351 gtcgcggga acgcctacct cgcgcgctc gcgagcacc gccgggcca

10

20

30

20401 cggactgccc gccctgtcgc tgtcctgggg catctgggce gacgacctca
 20451 aactggggccg ggtcgatccc cagatgatcc ggcgagcgcg cctggagttc
 20501 atggaccgcgc agctggccct gagcggcctg cagcgggcgc tggacgacaa
 20551 cgagaaactg ctcggggtcg ccgacgtgga ctgggagacc taccaccccg
 20601 tctacacctc cggccgaccc accccgctct tcgacgaggt gccggaggtc
 20651 cgccggctca ccggggccgc cgagcagagc gccgggaccg tcgccgaggg
 20701 cgagttcgcc gccgcgtgc gcgccctgtc cgacgccgag cagcagcgca
 20751 cctgtctgga gaccgtccgc accgagggcg cgtccgtcct cgggctgtcc
 20801 tccgccgagg acctcaccca ccagcgggcc ttccgcgacg tcggcttcga
 20851 ctcgctgacc gccgtcgcc tcgcgaaccg gctcgctcc gtcaccggcc
 20901 tgacgttgc ctcgacgatg gtcttcgact accccaaccc gcccgcgctc
 20951 gccgcctatc tgcacggcga gctggccggc gcccggtccg ccgccgcccg
 21001 cgccgcgcgc gtcccgaccg gcgcgcccca cgccgacgac ccgatcgcca
 21051 tcgtcgccat gagctgccc taccgccggc gggtcggtc cgccgagga
 21101 ctgtggcgga tcgccctgga cgaggtcgac gcgatctcc gcttccccgc
 21151 cgaccgcggc tgggacggcg agggcctcta cgaccggac ccgacccggc
 21201 ccggccgcac ctactccgtc cagggcggtt tctgcgcga cgtcgccgag
 21251 ttgaccccg gcttcttcgg gatctcgcc cgcgagggcg tgtcgatgga
 21301 cccgcagcag cggctcctgc tggagaccgc ctgggaggcg ttcgagcacg
 21351 ccggcatcga ccgggtcggc cagcgcgcca gccgcaccg cacttcgtc
 21401 ggcgcagct accaggacta cgcctccggc gtgcccaaca gcgagggtc
 21451 cgaagggcac atgatcacc gcacgctctc cagtgtgctg tcgggcccgg
 21501 tgtcctacct cttcggcttc gagggccccc ccgtcacgt cgacaccgcc
 21551 tgcctcctc cctgtgtgc catgcacct gccgtccagt cctgcgcaa
 21601 cggggagagc tcgctggccc tggccggcg cgtcagcat atgtccacc
 21651 cgatgtcgtt cgtcggttc agccggcagc gccccctcgc cgaggacggc
 21701 cgctgcaagg cgtacgcgga cggcgccgac gggatgaccc tcgccgagg
 21751 cgtcgccctg gtgtgtgtg agcggtgtc cgacggccc gcccaacggc
 21801 accaggtgct cgcgtgac cgcggctcc cgtcaacca ggacggcgcc
 21851 tccaacggcc tgaccgcacc caacggccc tcccagcagc gcgtcatcc
 21901 ccaggcgctg gccaaactcc cgggtggcgc cggcgacatc gacgtcctgg
 21951 agggccacgg caccggtacc gccctcggcg accccatcga ggcgcaggcc
 22001 ctgctcgcca cctacggcca ggaccgcgc cccgaacggc cgctgtgtct
 22051 cggctcggtg aagtccaaca tcggccacac ccagatggca tccggcgctg
 22101 ccagcgtcat caagctcgtc cgcgccctcc aggaaggcgt ggtgcccaag
 22151 tcctgcaca tcgaccggcc ctccaccac gtcgactggt cctcgggcgc
 22201 catcgggctg ctaccgaac gcaccccgct gcccgagacc ggccggccgc

10

20

30

22251 gccgcgcgc cgtctctctc ttccggcatca gcggcaccac cgtccacacc
 22301 atcctcgaa aggcctccgc ggacgaggcg cccacgccc cgcaccgcgc
 22351 gcgggacggc ctgggtgccg tctgtctctc cggccgcggc gaggcgcgc
 22401 tgcgcgccca ggcgcgcgc ctgctgcctc tgcgcaggga gcggcccgag
 22451 gccacactca ccgacctcgc ccactccctc gccacctcgc gcgcgcgcct
 22501 ggaaacgcgc gcgcgcgtca tcgcgcgcga ccgcgcacac ctgaccgcgc
 22551 gcctgcgcgc cctgtccgac ggccggccc accccggcct ggtccagggc
 22601 accgcgggac gcgcgcggac cgccttctctg ttaccggac aggcagacca
 22651 gcgccccgyc atgggcgcgc aactccacga ccgtaccgc gtgttcgcgc
 22701 acgcgctgga cggggtgctg gcccggtcgc acgacggacc ggaccggcgc
 22751 ctgcgcgagg tgctgttcgc cgcgcgcgc tccgcgcagg ccgcgcctct
 22801 ggaccggacc ggctacgccc agccgcgcct gttcgcgcgc gaggtcgcgc
 22851 tggtccgcct gctgacgtcc tggggcctga ccccgacta cctggccgc
 22901 cactccgtcg gcgaactcgc cgcgcgcgc gtgcgcgcgc tgctgtcgtc
 22951 ggacgacgcc tgcactctgg tcgcgcgcgc cggccgcctc atgcaggcgc
 23001 tgcccgaggg cggcgcgatg gtgcgcctgg aggcgcgga ggacgaggtc
 23051 ctgcgcctcc tggaggggct caccgaccgc gtgtccgtcg ccgcgcctaa
 23101 cgggcgcgcgc tccgtggtcg tcgcgcgcgc cggaggagac gtgtcctcc
 23151 tcgcgcacct ctgcgcgcgc gacgggcgcgc ggaccaagcg gctgcgggtg
 23201 agccatgcct tccactcgc gctgatggac gccatgctcg accacttcgc
 23251 cgcgcgcgc cgcgggctga cctaccaccc gccgacgatc ccgttcgtgt
 23301 cgaacgtcag cggcggcctg gccaccgcgc aacaggctcg caccgccgc
 23351 tactgggtcg ggacgcctcg cgcgcgcgc cgttcgcgc accgcacga
 23401 ctggctcgc acccaggcg acgtccacac ctctctggag ctgcgcgcgc
 23451 acggcgtgct cagcgcctatg gcccgggaga gcctcaccga cccgtccgc
 23501 acggcactgc tgcgcaccct gcgcgcgcgc cggcccgagg aacctgcct
 23551 ggtcaccgcgc gtgcgcgcgc cccacgcgc cggcgccgc gtgcactgga
 23601 gcgggtactt cgcgcaccac ggcgcgcgc ggaccacgt gccgacctac
 23651 gcgttccaac gcgagcggtg ctggccgcgc accaccgcgc ccacgagcgc
 23701 ccacacgcgc gcatccgcgc tcgacgcgc gttctgggc gccgtcgagc
 23751 gggaacgacgt cgcgcctcgc gccgcctccc tggacctgga cgaacccacc
 23801 gtcaccgcga tggctccgc gtcaccgc tggcgccgc gccgcggcga
 23851 gcagaccgaa ctggactcct ggcgtaccg cgtcacctg aagccgcgc
 23901 gcggcgccac cgcaccgc gccctaccg gccgctggct cgtgtcgtc
 23951 ccgcacgacc accaggaccg tcaggacgac gcgaccgcgc cctgggcagc
 24001 cgacgtcgag accgcctgg gcaccaccac cgtccgctg accgtacca
 24051 ccaccgacc cgcgcgcctg gccgcgcgc tcaccgaagc cgcgcgcgc

10

20

30

24101 cagggcccggt tcagcgggtgt gctgtccctg ctgccgctcg ccaccggcga
 24151 cgcgggcac ccgggtgcgc ccgccgccct caccctcacc accaccgccg
 24201 tccagggcct cggcgacgcc ggcatcgacg cgcgcgtgtg gaacgtcacc
 24251 cgcggggcgg tggcgtcgg ccgcgcgaa caggtcaccg cgcgcgaaca
 24301 ggcgcgcgtc tggggcctgg gccgcgcgt cgccttgaa ctgccggccc
 24351 ggttcggcgg caccctcgac ctgccgccca ccttggaagg ccaggccgcc
 24401 cgcgggttgc gcgcggtgct cgcggctacc gacggcgagg acgcggtggc
 24451 cctgcggccc tcggcgtct tcctccgcc cctggccacc gccccggccg
 24501 gccccgacac cgcgcgcacc gccttcgacc cggcggcgg caccgtcctg
 24551 atcacggcgc gcaccggcgc catcgcggc cagtcgccc gccgcctggc
 24601 ccgcgacggc gccaccacc tgcgtctcac cagcgcgcg ggcccggccg
 24651 cccccggcgc cgcgcgcctc ccgcgcgaa tggaggaact gggcgcccg
 24701 gtcaccctcg ccgcctgcga cgcgcgcgac cgcgacgcg tggcgcgct
 24751 cctgcgcgaa ctgcccgacg acccccgct gtgcgcggtg ttccacaccg
 24801 ccggcgtcgt cagggaccac gtcgtggacg cgtcacacc ggagaacttc
 24851 gccgcggtgc tgcgcgcaa gaccgtcgcc gccaccacc tgcacgagct
 24901 gaccgcgac ctggacctcg ccgtttcgt gctgttctcc tcacggccg
 24951 gcgtcctcgg cgcgcgcgga cagggaact accgcgcgc caacgcccac
 25001 ctggacgccc tcgcgaaca ccgcgcctcc cagggcctga ccgcgctgtc
 25051 cgtgcctcgg gccgcgtgg ccgcctcgg catggtcgcc gacgcgcg
 25101 aactcacga cgggttacg cgcgcgcgct tcgaaccgt cgcgcgcgaa
 25151 ccggcgtg gcgcctgct gcgcgcctc gagaacgacg acaccaccgt
 25201 cgcgctcgcc gacatcgact gggagcgtt ccagcgcgc ttgcgcg
 25251 tcgcgcgct gccgttcgt gcgcacctcc ccgagaccg ccgggccacc
 25301 ccgcgacgc ccaccggcgc cgcaccggc ctgcgcgac aactgcgga
 25351 actgccggag cagcgcgc ccgcgcggt cctggacctg ctgcgtaccc
 25401 aggtcgccgc cgtcctcggc cagcgcgacc cgcgcaccgt cagggacgac
 25451 cagccttcc gcgacctggg cttegactcg ctgacctcc tggaaactgc
 25501 caacgcctc aacgcgcga ccggcctgag cctgcgcgc accctggtct
 25551 acgacctgcc caccgcgcg gagatggcg acttctgct cgcggaactc
 25601 ctgcgcacc tgcaccga caccgcgcg accgtcgcca gcacggcctc
 25651 ccccaagctc tcagcttcgt tcgagcaggg cgttaccctc ttgcgacacc
 25701 cgatcgccgt catcgccatc ggtgcgct tcgccggcg cgtcaccacc
 25751 ccggaggagc tctggcagct cctcgacgag gccgcgcgac gcatcagccg
 25801 ctccccgac gaccgcggt gggacctcgc cgcgctggg gccggcgct
 25851 ccgacacct ggagggcggc ttctgaccg gcgtgcgga cttegacgcc
 25901 cggttcttcg gcatctcgcc ccgcgaggc ctggccatgg accccagca

10

20

30

25951 gcggtctgtg ctggagacca cctgggaggg gctggagcgg gccggcatcg
 26001 acccgaccac gctgcgcggc tccaccaccg gcgtcttcgt cggcaccaac
 26051 ggccaggact acccgacgct gttgcgcgcg tccgcctcgg acgtggccgg
 26101 ctacgtcgcc accggcaaca ccgccagcgt gatgtccggc cgcctgtcct
 26151 acgcgctcgg cctcgaaggc ccggccgtca ccategacac cgcctgtctc
 26201 tcctcgctcg tcgccttgca ctgggcgggc cgggcgtgc gcgccggcga
 26251 gtgcgacctc gtggtggcgg gcggcgtctc ggicattggc agcccgact
 26301 ccttcgtcga gttctccacc caggggcgcc tggcaccgga cgggcgtgc
 26351 aaggcggtct ccgacgcgcg cgacggcacc gcctggtcgg aaggcgtcgg
 26401 catcctcgtc ctggaacgcc tctccgcgcg ccgccgaac ggccaccagg
 26451 tcctcggcct gatccgcggc accgccgtca accaggacgg cgcgtccaac
 26501 ggccctgacc cgcaccaagg cctctcccag cagcgcgtca tcgccaggc
 26551 actgcgcgac gccgcctgc gccccggcga catcgacgcg atcgaggcgc
 26601 acggcacccg caccacctc ggcgaccgga tcgaggcccg cgcctgac
 26651 accgcctacg gccgggaccg ggacgcccga cggccgtgc tgcggggcac
 26701 cgtcaagtc aacatcgggc aaccccaggc cggcgccggt gccgcggcg
 26751 tcatcaagat gctgatggcg atgcgccacg gcacctgac caggacgctg
 26801 cacgtgggca cccgctccag ccacgtcgac tggagggcg gcaccgtcg
 26851 gtcctcgac gacgcgcggc cctggccacg gaccgggag ccgcggcgcg
 26901 ccggcgtctc cgccttcggc gtcagcgga ccaacgcca cgtcgtcgtc
 26951 gagcaggccc cggaaaccga agccccgcc gccccggcg ccgagccggc
 27001 gccggaggcc acgcccaccg tcgtccctg ggctgtctcc ggacgagcc
 27051 gggaagcgt ccaggcgag ctggaccggc tcaccgcga caccgccgc
 27101 caccgcgcg gctcggcggc ggacgtcggc cgtcgttg ccaccgaccg
 27151 cagctcttc ccgcaccgcg ccgtgtgct cgcgggccg gacggggtgc
 27201 gcgaggccgc ccgcgcgcg gcgccccgca ccccgcccg caccggttc
 27251 ctgttctccg gacaggcgcc ccagcacgcc ctgatgggc acgacctga
 27301 ccagcgttc ccggtctacg ccgacgact ggacaccgtc ctgcgccagt
 27351 tcgacaccgt gctggacgtc ccgtgcgcg ccgcgtgtt cgcgcgcg
 27401 ggcaccccg aggcgcgct cctggaccag accggctca cccagccgc
 27451 gctgttcgcc gtcgaggtcg cactgttcg gctcgcgag tcctggcggc
 27501 tgacgcggga ctctgtcgc ggccactcca tcggcgagat cgcgcgcg
 27551 cagctcgcg ggggtgtctc cctggaggac gcctgcacg tggtcgcgc
 27601 ccgcgcctc ctcatgcag aactgccgcg ggacggtgc atggtggcc
 27651 tgggaagcac cgaggacgag gtcgcgcgcg tgcaccca cgcgtcgca
 27701 ctgcgcgcg tcaacggccc ccgtcgggtg gtcgtcgcg gcgccagga
 27751 cgcgtcgcg gcggtcgcc accggctcgc cgcgacggc cgcgcaccc

10

20

30

27801 gccggctgac ggtcagccac gccttccact cggcgctgat ggacccgatg
 27851 ctcaccgact tcgcccggtt cgcggagggc ctgacctacc acgagccgcg
 27901 cateccctc gtctccaccc tcctcggcgc cccggccggc gccgaactgc
 27951 gcaccccca ctactgggtg cggcacgtcc gcgagaccgt gcggttcgcc
 28001 gacggcgtgc gcgcctgca cgcgcggcgc gccggcacct tcgtggagat
 28051 cggcccgga cgcgtgctca ccgcctgac ccagcagacc ctcgacaccg
 28101 tcgaggccgg cgcgcggcgc gtcgtcgtgc cgtccagcg ccgcgaccgc
 28151 gccggcgacc tcgccctcct ggaggggcctg gccaccctgc acaccacgg
 28201 caccggcccg tcctggcccg cctacttcga ggccaccggc ggccaccgga
 28251 ccgatctgcc caccacgco ttccagcggg agcggtagtg gccgaactc
 28301 ggcgaccccg tcgccaccgc ccgcaggac ccggcgccct ggcgctacca
 28351 cgagacctgg gcccgcgtgc cggcaccca ggcggccgcg ccggccggcc
 28401 gcgcctcgtt cctcgtcccg gccgggaacc gcgacaccgc gtggatgacg
 28451 gccgtcgcg acgcgtcgtg cgcgcgacc gtcacggccg aaccgcgacg
 28501 cctggccgag cagttgacgg ccgcggcgca cacaccctgg cgcgtcgtgg
 28551 tgctcgtgct cgcgcgcgcg tcgagggggc tgcccgcgga cggcgccctg
 28601 ccgcgcgcc tcctcggcac cctggacgag gccggcgtgc acgcgccct
 28651 gtggtcgtgc accgcggcg ccgtcgcgtt cgcgggggag gcccgaccg
 28701 ccgtcggcca ggccgcctt tggggcctgg gccgggtcgc cgcgttgac
 28751 caccgggacc gcttcggcgg cctggccgac ctgcccgccg acaccgacgc
 28801 gcacgcgcgc gggtcgtcgt cgcgcgacct ggccgcgcgc ggccaccgag
 28851 ccgagatcgc ggtccgcgc accggcgctc acgccgtcg cctggtccgt
 28901 acgcgcggcg ccgcgcgag tgccacctgg ctgccgaccg gcaccgtcct
 28951 ggttgctcgc ggccaccgag gcaccggcac catgggcggc cgggcgcgcc
 29001 gctggctggt ccgcgagggc gcccgccacc tcgtcctgac cgcgcccgac
 29051 ggccaccgga ccgcgcgga caccgaggcc ctgacggccg aactggccgc
 29101 gctcggcgcc cggatcacgc tcgtggacca cgacccacc gccccggacg
 29151 gcttcgcgc gctcctcgac ggaactgccg acgacacccc gctcacgcgc
 29201 gtcgtgtacg cgcggagggc cgaacgcgcc ccggcacccg cggcgagct
 29251 gtccgcgcga ctgcgcccg tcaccgcctt aggcgcgcgc ctacccggcc
 29301 ggccgctgga cgccttcgtc ctcttcggct ccacgcgcg gctctggggc
 29351 gtgcgcggcc ggccgcgcga gcccgcgctc ggccgctacc tcgacgcctt
 29401 cgcgcgcgcc tgccgcgacc gcggcacccc ggcaactggc gtcgcctggg
 29451 gcgcctgggc cgacctggtc ggcccgctcc tcgcgcgca cctgcggatg
 29501 aacggcctgc cgggtgatga cgcggacacc gcaactgacc cctcagccg
 29551 ggccgctcgc gacgggtccg ccgcgaggc ggtcgcgcgc gtcgcgtggg
 29601 agaccttcgc gccctccac cagaggccc gccgcaccgc cctgttcgac

10

20

30

29651 gccctgcccg aggcccgccg cgcgctcgcg gaggccgccc gggaccgcg
 29701 cgaccggaag accgccgccg gcgactacgg ccggtggctc gccgagcagc
 29751 ccgccgcgga ccacgacgcc atcctgctgg cactggtcac cgagaaggcc
 29801 gcgaccgtcc tcggccacgc cgaccacgac ctgctcgaac ccgacctgcc
 29851 ctcccgcgac ctgggcttcg actcgctgac cgcggctgac ctgcgcaacc
 29901 agctcaccgc ggaaaccggc ctacccctgc ccgccaccct cgtcttcgac
 29951 caecccaacc cggccgcccct cgcgcgccac ctgcgcgccc aactcctcgy
 30001 cgaggcgagc gactccgccg caccggtggc cgcgcccgtc gccctcggtg
 30051 ccgacgacga cgcgatcgtc atcgctggca tggcctgccg ctaccccggc
 30101 ggggtcacct cgcgcgagga cctgtggcag ctggtcggcg acgaggtcga
 30151 cgcggtcggc gacttccgga ccgaccgagg ctgggacctg gccgcgctcg
 30201 ccggcgacgg accggggcgc agtgccaccg ccgaggcgcg attcctctac
 30251 gacgccaccg acttcgacct cggcctgttc ggcattctcg cgcgcgagge
 30301 cctggtgatg gaccgcgagc agcggtacct gctcgaaacg tcctgggagg
 30351 ccctggagcg ggccggcacc gaccggcgga cgtgcgcggy cagcggcacc
 30401 accggcgctc tcgtcggcgg cggctccggc gactaccggy cgcgcgagga
 30451 ggccggggcag tggcagaccg ccagtcgcgc cagcctgtct tcgggtcgcc
 30501 tcgcctacac cttcggcacc caggggccca ccgtgtcggc cgacaccgcc
 30551 tgcctcctgt cgtcgtcgcg gctgcacctg gccgcgcagg ccctgcgcgc
 30601 cggcggaatgc tcgatcgcgc tggccggcgg cgtcaccgtg atggccaccc
 30651 cgggtgggctt cgtcgagttc agcggccagg gcgcctgtct gccggacggc
 30701 cgtgcgcggy ccttctccga cgacgccaac ggaccggctt ggtccgaagg
 30751 cgtgggcatg ctcgctcgcg aacggctctc cgacgcccgc cgcaacggcc
 30801 accgcgtcct cgcggtgctc cgcggctccg ccataacca ggacggcgcg
 30851 tccaacggcc tgaccgcccc cagcgccccg gccacgacgc gcgtcatccg
 30901 ccaggccctc gccaacgccc gaetgcgccc cgcgcgacac gacgcgctcg
 30951 aggcccaagg caaccggcac aggctcggcg accccatcga ggcccaggcc
 31001 ctgctcgcca cctacggcca ggaccgcgag cggcccgtgc tgcctcgctc
 31051 gctcaagtcc aacatcgccc acaccaggc cgcctccggc gtcggcgcg
 31101 tcatcaagat ggtcctcgcc atgcagcacg gcgaactgcc gcgtccctg
 31151 tacgccgaga acccctcgtc gcacgtggac tggaccgccc gccgcgcccc
 31201 cctgctcacc gccaggaccc cgtggcccca ctccggctcg ccgcgcgcg
 31251 ccgcgctctc ctccctcggc gccagcgcca ccaacgcccc cgcctcctg
 31301 gagcagcgcc cgcgcgagga actccccgcy cgcgccgcyg acgacggcgc
 31351 cccgctgcgg ttcttgctct ccggccgctc gcagaacgcc ctgcgcgccc
 31401 aggcccgccg actcctggcc cgcctcaccg cccaccccca caaccgggc
 31451 gccgacctgg cgtactccct ggccgaccac cgggcccgtc tcgagcaccg

10

20

30

31501 ggccgcgata accgccaccg accacgaagg cctccgcacc ggcctgaccg
 31551 ccgtcgccga ggccaccacc gccccgcaca ccgccgaaca ccacctccag
 31601 ggcaccggaa agcgcgccgt gctcttctcc ggccagggct cccagcgcc
 31651 gggcatgggc cgcgaactgc acgagcgcca cccggtgttc gccgagggct
 31701 tcgactcgtt actggccccc ctcgacgacc ggctcgacac cccctgcgg
 31751 gacgtcgtct ggggcaccga cgaggaggcg ctgcacgcca ccgggaacac
 31801 ccagcccggc ctgttcggcg tcgaagtcgc gctctaccgc ctgacgaat
 31851 cctggggcgt ggggcccgcac ttctgtggcg gccactccgt cggcgagctc
 31901 gccgcggccc acgtcgcggg ggtgctctcc ctggacgacg cctgccgcct
 31951 ggtcgcggcc cgcgcggccc tcatgcagcg cctcccgccc ggccggcgcca
 32001 tgatcgccgt cgaggccacc gaggacgagg tcaccccgct cctcaccgac
 32051 ggcgtgtccc tcgcgcgggt caacggaccg accgcctggg tctctccgg
 32101 cgcggggcgc gccgtgaccg ccctggggcca ggcgctggcc gaacggggcc
 32151 accgcaccac ccggctgcgg gtcagccacg ccttccactc gcacctcatg
 32201 gacccgatgc tggcggaact ccgcaccgtc gccgagggcc tggaaatcca
 32251 cccgcgcgcg atcccgtgg tctccaacct caccggggac gtcgccgacg
 32301 cggccgacct gtgtccgcgc gactactggg tgcgccacgt ccgcggcacc
 32351 gtacggttcg ccgacggcgt gcgcaccatg gccgaccgcg gcgtgcacct
 32401 ctccctcgaa ctggcccggg acgcccgtgt gtcggccatg gcccgccagt
 32451 gcgcaccgga cgcgctcgtc gtcccggccc tgcgcggcaa ccgcgacgag
 32501 gacgagacgc tggtcggcgc cgtcgcgcga ctgcacgtcc acggcgcggg
 32551 tccgcgctgg gacgcgtact tcgcggcggc cggcgccag tggttgacc
 32601 ttccgacgta ccccttccag cgcggccgct tctggccgga gtcccttccg
 32651 ggcgcgcgat cggccgcccc ggcagccgga cagccggcgg agaccgacgc
 32701 ggccttcttg gacgcgcgtc cacaggagga cttaccgca ttggaatccg
 32751 tactcgacgt cgagagcgac gcactgtcca aggtgctgcc ggccctgatg
 32801 gactggcgca gccgccaggc cgacgagtc caactggcag gctggcgcca
 32851 ccgcacgtc tggaaagcgg tcaccggcgc cgccttgga caccgcaagg
 32901 cgtcagcgg cacctggctc gcggtggtcc ccgagggctt cgcgcgacg
 32951 ccctgggtga ccaccacct ggacggcctc ggtaccacc tctgtcatct
 33001 ggaggtcgcg gaggccgacc gggccgcgt ggccgacgag atcgcgccc
 33051 gcaccgcca cggcaccgc ttcggcggg taatctccct gctggccctg
 33101 cgcgaggagc tcaccggcgc ggtgcccgag gggaccgccc tgaccaccac
 33151 cctcctccag gccctcggcg acgcccgcgt cgacgcaccg ctgtggtgcg
 33201 tcaccgcgag cgcgctctcc gccggccgca ccgaccggcc gcaaccgacg
 33251 ctccaaggcg ccgtctgggg cctgggccc gtcgcggccc ttgagtacc
 33301 gcagcgtcgg ggcggcctgg tggacctgcc ggaggagccc gacgagcgg

10

20

30

33351 ccgcggccgg cctcgccgcc gtccctggccg gtctggacgg cgaggaccag
 33401 gtcgccgtgc ggggcaccgc ggtgctcgcc cgcgcctgg tgccggctcc
 33451 cggccgcaag ccgtcccgcc cctggcacc gtccggcacc gtccctggtca
 33501 ccggcgccac cggcgccctc ggcgcgcacg tcgcccgcgg cctggccaag
 33551 gacggcgccc agcacctcgt cctgctcagc cgcgcgggcc cggacgctcc
 33601 cgggtcgggc gaactgcggc cggaaactga cgcgttgggc accgacgtca
 33651 cggtegcgcg ctgcgacgtc gccgaccgcg accagctgac ggccgtcctg
 33701 gacgcgctgc ccgcgcaccg gccgctgacc ggtgtggtgc acaccgcggg
 33751 cgtccctcag cagcgcgtae tggaccggct cacccccag cggttccagg
 33801 aggtgttccg cgccaaggte acctcgcccc tgctgctgga cgagctgacc
 33851 cgcgaccgcg agctggccgc gttcgtcctc ttctcctccg cctccgcgcg
 33901 ggtcggcaac ccgggcccag ccaactacgc cgtcgccaac gccgtcctgg
 33951 acgcgctcgc cgaacagcgc cgggtgctcg gcctgccgcg caccctcggtc
 34001 tcctggggtg cctggggagg cggcgccatg gccgacgcgg acggcgcgga
 34051 cgaggccgcc cggcgccgcc gcgtcggcgc catggaccgc caccctgcgcg
 34101 tggaagccct gctgcgcctg gtcgcccaga aggagccgac cgcggtggtc
 34151 gccgaggtgg ccctggaccg gttcgcgggc gccttcggcg gcagccgacc
 34201 cagcgccctg ctgcgggagt tccccggcta ccgcgaggcg ctgcgcgcc
 34251 aggcggagca ggccgcggac ggcggcgggc tggccgcccg actggccgcg
 34301 ctgcgcgccg ccgcgcgcct ggacaccgtt gtggacctgg tgcgcacccg
 34351 cgcgcgcgag gtgctcggtt accccgacac cgaagcggtc gccgcggaac
 34401 ggtccttccg cgacctgggt gtcgactcgc tcggcgccgt cgagctgcgc
 34451 aaccaactga gcgcggccac cggcctgaac ctgcggcgga cgtggtggtt
 34501 cgaccacccg acccccctgg tcctggggga gcacatcctc ggccggctct
 34551 tcccggaaga gcccgccggg tccgacgacg agacggagat ccgggccctg
 34601 ctggcctccg tcccgctega ccaactgcgg gagatcgggg tcctggagcc
 34651 cctgctccag ctgcgccgac gcggcgcccg ggccgcggac ggcgacgacg
 34701 gcgagtcctg cgactcgatg acagtggcag acctggtgcg ggccgcgctc
 34751 aacggccagt ccgacctgta gcgcgattga tggagcagac gatgaacgcg
 34801 cccgagaacc ccgagacccc cgagaacaac gtagtcgccg cactccgcgc
 34851 cgcggtcaag gagaccgacc ggctccggcg gcagaaccgg atgctggtcg
 34901 cggcgcccaa ggaaccgacg gccgtggctg gcatggcctg ccgcttcccc
 34951 ggcgccgtcg actccccgga agcgtgtggt gagatggctg ccaccggcac
 35001 cgacgtgatc tcggattcc ccgacgaccg cggctgggac ctggagggcg
 35051 tgcgcaacag cggcacgcgac gcccgcgaca ccgacgtcag ccagcgcggc
 35101 ggattcctgg actgcatcgc cgacttcgac cccggcttct tcgggatctc
 35151 accgcgcgag gcggtcacca tggaccgcga acagcggtc ctgctgacca

10

20

30

35201 ccgcctggga ggccgtcgag cgggccggca tcgacgccac cacgtgcgc
 35251 gccacccgca ccggcgcggt catcggcacc aacggccagg actacgccta
 35301 cctgctcgtc cgtccctgg acgacgccac cggcgacgtc ggcacccgca
 35351 tcgcggccag cgcgcctcc ggtcggtct cctacacct cggcctcgaa
 35401 ggcgccgcgc tcaccgtcga caccgcctgc tcctcgtcgc tggtcgcct
 35451 gcaactggcc gtgcaggcgc tcgcgaacgg cgagtgcggc atggcgctgg
 35501 ccggcgcggt caacgtgatg gccacaccgg gctcgctggt cgagttcagc
 35551 cgccaggggg ggctggcccg ggacggccgc tcgaaggcgt tcgcggacgc
 35601 cgccgacggc accggctggt ccgagggcgc cggcgtgctg ctgctggaac
 35651 ggctctccga cgcgcagcgc aacggccacc cgtgctcgc cgtggtccgc
 35701 ggctccggcg tcaaccagga cggcgccctc aacggcttca ccgccccaa
 35751 cggccctcc cagcagcgcg tcctccgcca ggcctcgc aacgccggcc
 35801 tggccaccgg cgacatcgac gcggtcgagg cgcacgggac cggcaccgcc
 35851 ctggcgaccc ccacgagggc gcagagcctc ctgccacct acggccagga
 35901 ccgcgcccac ccggtgctgc tcggtcgtat caagtgcac atgggccaca
 35951 cccaggccgc gtccggcgtc gccggcgta tcaagatgat catggcgatg
 36001 cggcacggcg tcctcccgcg gacctgcac gtcgacggc cctccacca
 36051 cgtcgactgg accaccggca gcgtcgaact cctaccgac gccacccgt
 36101 ggcgcgagac tggcaggccg cgcgcacccg gcctctctc cttcggcgtc
 36151 agcggcacca acgcccacgt catcgtcgaa caggcccccg acacccccgc
 36201 cgaggcggtt gacgacactc cgcgccgcac ccgcgggacc ctgcggtggc
 36251 tgctctccgc ccgcaccggc gccgcctgc gcgaccaggc caccgcgctg
 36301 ctgcaccacc tcgaccgccc cgacggcgac cgcgggccc aacccctgga
 36351 caccgcgttc tcctcgcga ccaccgcgc cgcctgga aaccggctcg
 36401 ccgtcgtcac cggcaccgac ggcaccgcgc gacgggacgc cctgaccgcc
 36451 tggctggcgc acggcacccg ccccgacgcc caagaaggac acgcgcggg
 36501 acgcaccgcg tcgcgggccc tcttctccg ccaggcgccc cagcgctgg
 36551 gcatggggcg cgaactccac gcccgtttc cgggtgttcg acgggcccct
 36601 gacaccgcgc tggacctgct cgacggcgaa ctggcgggca cctgcggga
 36651 ggtgatctgg ggcaccgacg acgcgcgct caacgagacc ggttcaccc
 36701 agcccgccct gttcgccgtc gaggtcgccc tctaccgct gatcgaatcc
 36751 tggggcgctg ccccggaact cgtcgccggc cactccatcg gcgagatcg
 36801 cgccgcgcac gtgcggggg tgttctccct ggaggacgcc tcgacgctgg
 36851 tggcgcccg cgcgggctg atgcaggcgc tcgcgcggc cggggcgatg
 36901 gtgcgctcg aggccaccga ggaagggc agcccgctgc tcaccgacgg
 36951 cgtcgcgatc gccgcgatca acggccccc ctcgctcgtc gtctccggcg
 37001 acgagaccgc caccctcgcc gtcgcggccc gactcgccga acaggggcgc

10

20

30

38901 cgagtcgcc gcccaactcc cgttgctgcc ggccctcctg gacgccgacg
 38951 agccgcaggc cgtgggtccgc gagggcaccg tccgggtcgg ccggctcgc
 39001 cgcctggact ccggccgcgg cctcgtcccg ccgcccggca cccctcggcg
 39051 cctgggcagc cgggccaagg gcagcctcga cggcctcgc ctgctgccc
 39101 accccgaggc ccgacgaccc ctcaccggcc acgaggtccg cgtcggcatc
 39151 cgcgcgcggc gctgaactt ccgtgacgtg ctcaacgcgt tgggatgta
 39201 tcccggggat gggggctgt tgggttcgga ggccggcggc gtgggtcgtg
 39251 aggtcggacc ggaggtcacc ggcctggcac ccggcgaccg ggtcatgggc
 39301 atgctcttcg gcggcttcgg accgctcggc atcgccgacg cccggctgct
 39351 ccccccgtc ccggccgact ggtcctggga gacgggtgcg tcgggtcgcgt
 39401 tgggtgtcct caccgcgtac tacgccctga aggagttggg tggctcgcgg
 39451 gcgggggaga aggtgctggt gcatgccggt gccgggtggt tcgggtatggc
 39501 ggcgatccag atcgcccgcc atgtcgggtc cagaggtgtc gccacggcca
 39551 gtgagggcaa gtgggacgtg ctgcgctccc tgggcgtggc cgacgaccac
 39601 atcgctcct cccgcaccct cgacttcgag gcggccttcg ccgaagtcgc
 39651 cggcgaccgc ggcctggacg tcgtactgaa cgcgtgttcc ggcgagttcg
 39701 tcgacgcctc gatgcggctg ctcggcgacg gcggccgggt cctggagatg
 39751 ggcaagaccg acatccgcgc cgcggactcc gttcccgacg gctctccta
 39801 ccactccttc gacctcggca tggtcgatcc ggaacacatc cagcgatgc
 39851 tgetcgacct cgtcgagctg ttcgaccgcg gcgcgctggc cgcgttgccg
 39901 gtccgcagct gggacgtgcg ccgcgcgggc gaggcgttcc gcttcatgag
 39951 cctggcccag cacatcgga agatcgtgct caccgtgccc caaccctcg
 40001 accccgacgg caccgtgctc ctcaccggcg gcaccggcgg cctggccggc
 40051 ctgctcgcgc gccacctggt caccgagcac ggcccccgc acctgctgct
 40101 ggccggcccg cgcggcccg accgcgcccg cgcgcgcga ctccacgcg
 40151 aactgaccgc cctgggcgcg gaggtcaccg tcgcgcctg cgacgtcgc
 40201 gaccgcaccg cgtcgcgcg gctgctcgc accgtgccc ccgaacacc
 40251 cctcaccgcg gtcgtgcaca ccgcggcgt cctggacgac ggcaccctca
 40301 ccgccctgaa ccccgaccgc ctgccaccg tctacggcc caaggtggac
 40351 gccgcctggc acctgcacga cctcaccgc cactcgacc tggcgcggtt
 40401 cgtgctctac tctccaccg ccggcgtcat gggcggaacc ggccaggcca
 40451 actacgcggc cggcaacacc ttctcgcagc cgtcgcgcg ccaccgacac
 40501 gccctcggcc tgcccgccac ctgcctggcc tggggcgcc gggagcaggg
 40551 cgccggcatg accggcgac tgaccgacca cgacctgcgc cgggtcagcg
 40601 acgcggcgcg ccaaccgctg ctcaccgcg aacgcggcct cgcctctac
 40651 gacgcgcga ccgcgcga cgaaccctg atcgtccgc tcggcctcac
 40701 cggcggtgcg ctgcccgcg gggtcggcgt ccccgccgtg ctgcgcggc

10

20

30

40751 tgggtccgcac cgcggggccgc cggggccaggg ccggcaccgc cggcgtctcc
 40801 cgcgcgcggcc tcgcggaacg cctcgcgcgc ctgcccagg aggagcgcac
 40851 ccccttcctc gtgcagctgg tgcgcaccga ggcgcgcacc gtccctcggcc
 40901 acggtctccac cgaccgcgtg gacgcgcgc gcgagttccg ccaactcggc
 40951 ttgcactcgc tgaccgccat cgaactgcgc aaccgactcg gcaaggccac
 41001 cggcctcacc ctgcccgcga cctcctctt cgactaccg acccccgacc
 41051 gcctcgcgt ccacctccac gacgaactcc tcggcgcgga cgcgcgcgtg
 41101 accgtcaccg ccgcgcgaca ggcgcgcgac ccggagcacg acccggctcgt
 41151 catcgtcggc atgagctgcc gcttccccg cggcgtcagc tccccgagg
 41201 agctgtggga cctgggtggca tcgggcaccg acgcgatcac cggcttcccc
 41251 gccgaccgcg catgggaccg ccaccgcgag ctgcgcggcg cccccggcgc
 41301 ccgcaccggc caggggcgat tcctcgcgga catcgcgcgac ttcgacgcgc
 41351 ccttcttcgg catctcgcgc cgcgaggccc tggccatgga cccgcagcag
 41401 cgcactctcc tcgaagtgcg ctgggaggcc gccgagcgcg ccggcatcga
 41451 cccgcagacc ctgcgcggca gcgacaccgg cgtgttcag ggcgtcagcg
 41501 gccaggacta cgcgcgcctc gtgatgcgt cccgcgacga catcgcgcgc
 41551 cagccacca ccggcctcgc cgtcagcgtc gtctcgcgc gcctcgccta
 41601 cgcgctcggc ctggagggcc cggccctgic cgtggacacc gcctgctcct
 41651 cctccctggg gtgcgtgcac ctggccgccc aggcgctgcg cgcgggggag
 41701 tgcaccatgg ccctggccgg cgcgcgcacc gtcgatgacca ccgccgcaa
 41751 ctccaccggc ttctcccgga tgggcggcct cgcgcggac ggcgcgtgca
 41801 aggcgttctc cgactccgcg gacggcaccg gctggtccga gggcgcgcgc
 41851 gtccctggtc tggaacgcct ctccgacgc cgcgcgcgcg gccaccgcgt
 41901 actggccgtg gtgcgcggct cggcgggtcaa ccaggacggt gcgtccaacg
 41951 gtctgacggc gcccaacggt ccgcgccagc agcgcgtcat ccggcaggcc
 42001 ctggccaacg cgggcctgac ccccgctgac gtggacgcgc tcgaggcgca
 42051 cggcaccggc acccgcctcg gcgaaccat cgaggcccg gccctgatcg
 42101 ccgcctacgg caccgaccgc gacccgaa acccgcgtgt gctcggctcg
 42151 gtgaagtcca acatcgcca caccagtcg gcgcgcggcg cggccgggct
 42201 ggtcaagatg gtcattggca tgcgcacag catcctgccg cagaccctgc
 42251 acctcaccga accgtcctcg cacgtggact ggtcggcggg cacggtgcg
 42301 ctgctcaccg agcggaccgc ctggccgcgg acggatcgtc cgcgtcgggc
 42351 cggggtctcc tcgttcggca tcagcggcac caacgcccac gtcctcctgg
 42401 aacagccgc cgcgcgagcc accccgcgc ccgacccggg ccggcccgca
 42451 cccaccgtgg tggcctggcc cgtctccgcg cagaccccg ccgcctcga
 42501 cgcccaactg gaccggttgc gcacgcgcgc cgcctggcg ccgctcgaca
 42551 ccgcccaac cctcgcacc ggcgcctcgc tcttcgaaca ccgcgcgtc

10

20

30

42601 ctgctcgcca ccgtcggcga cccggcgacc ggcccccgcg acctgcccga
 42651 ggtcgccagg ggagcggcga cgcgcaccg caccgcgttc ctcttctcgg
 42701 ggcagggtgc tcagcggtcg gggatggggc gtgaactgca tgctgctttc
 42751 ccggtgttcg cggcggcgtt cgacgaggtg gtggctgtgt tggatgcgga
 42801 gctgggttct gatgctgatg ggggtgtgtc gctgcgggag gtgatgtggg
 42851 gcggggggtc ggagtgttg gatcgaacgc gtttcacgca gccggcgttg
 42901 ttccgggtgg aggtggcgtt gttccgtttg gtggcctcgt ggggggtggg
 42951 gcctgagttc gtggcggggc attcgggtgg tgagattgcg gccgcgcag
 43001 tggccggggt gttctcgttg gtggatgcgt gtcgttttgt ggtggcgcg
 43051 gcttcgttga tggatgcgtt gccggtgggt ggcgtgatgg ttgcggtgga
 43101 ggcggccgag gccgaggtgg tgcgcctgtt ggtcgaatgg gtggcgatcg
 43151 ccgcgggtcaa cggcccggtc tcggtggtgg tctccggtgt ggaggcggcc
 43201 gttgggcagg tcgtggatca gttggtggag cggggccggc gggtcctcgt
 43251 gttggcgtc agtcagtctt tccactcgc gttgatggat ccgatgttgg
 43301 atgccttcg gccgctcgc gagggcctgg agtaccacca gccgcgcac
 43351 cccgtggtgt ccaacgtgac gggcgagggt gccgcggcgg aggagctgtg
 43401 cgcggccgac tactgggtgc ggcaagtcgc ggcaaccgtg cggttcgcgc
 43451 acggcgtccg caccctggcc gagcgcggcg ccaccgcctt cctggagatc
 43501 ggccccgacg gcgtactgtc cgcgctcgc cgcggcgtcc tgcccccgga
 43551 ggcgctcgtg acgcccaccc tcgcgaagga ccgcgacgag gagagcgccc
 43601 tgctcgccgg actggcccgg ctgcacgtcg ccggcgtgac cgtcgactgg
 43651 agcgcgcgcc tgaccggcac cgcgcgcgcg ggcaaccgac tgccgacctg
 43701 cgccttccaa cgcgagcggg actggccgga gttggccgc gaaccgcgcg
 43751 gcggcgcgcg ggtgacgcg gacgcggagt tctgggccgc ggtggagcgc
 43801 gcggacgcca ccgcgctcgc cgcacacctg gacatcgacg gcgaccagct
 43851 cggcgcgtg ctgcccgcac tgtccgcctg gcgcacccgg cgcgcacca
 43901 catcgccac caacgcctg cggcaaccgg agagttggga accgctgtcg
 43951 ctgcgcgga cgcgcacac cggcggcgtc ctggtgctgg tgcccgcgc
 44001 cgcgaccacc gaccctggg tcgcgacgt cgtgcgcgc ctgcggcccg
 44051 acgcccgcg ggtggacgtc ccggccgacg gcaccgaccg ggccgcgtc
 44101 gccgcctgc tcaccgaagc ggccgacgac accgcccga ccgcgtggt
 44151 ctccctgctc gcgctcgacg agaccagcgg cgcgacgcg gtaccggccg
 44201 gcaccaccgc caccgcgcg ctgctccagg ccctcgccga caccggcgc
 44251 ccggccccgc tgtgggccct gaccgggggc gcggtcgcg cgtccccga
 44301 cgagcagccg accgccccg ccagggcgc cgtctggggc ctgcggccga
 44351 tcgcccctc cgaactccg cgcactggg gcggactggg cgacctgcc
 44401 gccgacctc acgagcgcac cgcacgcga ctgcccgcg cactggccga

10

20

30

44451 cgccggcgac gaggaccagc tcgcgctgcg ggccaccggc gcctacggcc
 44501 gccggatcac cccggcgccc gcccccgacg acgccccggg caccgggtgg
 44551 cagccgaccg gcaccgtcct gatcacggcg ggccaccggc cgctcggccg
 44601 gcacaccgcc cgtgggtcgt cggcccacgg cggcgagcac ctgctgctgc
 44651 tcagccgcag cggccccgac gcgcccggcg cggccgaact caccaccgaa
 44701 ctacccggcc tcggcgcccg cgtcaccctc gtggcctgcg acgccgcgga
 44751 ccgggagcag ctgaccaggg tcctcgccga ggtaccggcg gactgcccgc
 44801 tgaccggcgt ggtgcacacc gccggagtgc tcgacgacgg cgtgctcacc
 44851 ggccctcacc cggaccggtt cggcacggtc ttccggcca aggtggcctc
 44901 cgccgtgctc ctgacgagc tgaccggga caccgacctg gccgtcttcg
 44951 cgctgttctc gtccgtcgcg ggcgctgtcg gcaaccccg ccaggccggc
 45001 tacgcccggc ccaacggcgt cctcgacgcc ctcgccgcc gccgcccggc
 45051 ccagggcctg gccggcacct cgatcgccgt ggggtgcttg gccggcgacg
 45101 gcatggcggc cgtcacacc cggcccggcg ccgaaccgt cggcctgctc
 45151 gaccccgaac tcgccgtacc ggccctggcc cgcgcggtga cggagcccca
 45201 gcccaccctc gtctcgccg acctccagca gcccgggctg ctggaatccc
 45251 tgctggcgt gcgcccagc ccgtcctga gccggctgcc cgcgcgccg
 45301 accgcccgc gcgcggtcca ggaaggcgac cgcgcgcgag ccggcgccgc
 45351 ggccgacctg cgcgaccaac tcgcccgcac cgcacccgcc gaccgccacg
 45401 ccgtctcct ccgcctggtg cggaccacgg ccgcccgggt cctcgccac
 45451 accggtgcg acgccatcg ggcgcacaag cccttcgcg acctcggtt
 45501 cgactcgtc acccggttg aactgagcag cgtctcgcc gccgccaccg
 45551 gcctggccct ccgcccagc ctgctcttcg accaccctc ccgcccggcg
 45601 ctgcccgaac acctgcgcg cgaactcacc ggcgaccggc cggaatccgc
 45651 ccccgcgcc ccgcccgcac cggteccgc cgcggacgac gatccgatc
 45701 tcgtcgtcg catggcctgc cgttccccg ggcggctcac cacccccag
 45751 gagttctgg agctgctgc cgaaggccgg gacggcatcg acgcgttccc
 45801 caccgaccgc ggctgggacc tcgaegtgt cggccggcga cggccagggc
 45851 cacagcgc cccgaggtcg gcggttcc cttacgagc ggccgccttc
 45901 gaccccggct tcttcgacat ctccccgcg gaggcgctcg ccatggaccc
 45951 gcagcagcgg ctgctgtgg agaccgctg ggaggccgtc gaacgcaccg
 46001 gcaccgaccc gaccggctg cgcggcagcc gcaccggcgt gttcgtcggc
 46051 accaacggcc aggaactacg cggcctcgtc ctgcccgcc aggaggacgt
 46101 cgagggggcac gccggcaccg gactggccgc cagcgtgatc tccggccgcc
 46151 tcgcctacgc cttcggcttc gagggcccg ccgtcacgt cgacaccgcc
 46201 tgctcctcct cgtcgtcgc cctgcaactg gccgtccagg cgtgcgcgc
 46251 gggggagtgc tccctggccc tggccggcg cgtcaccgtc atgaagacct

10

20

30

46301 cgacgagctt cgccggcttc acccggcagg gcggcctggc gccggagcggg
 46351 cactgcaagg cgttctccga ctccgccgac ggcaccggct ggtccgaggg
 46401 tgtgggcgtc ctctcgtcg aacgccgctc cgacgcgctc cgcaacggcc
 46451 atgagatcct ggccgtggcg ccgggctcgg cggtaacca ggaagggtcg
 46501 tccaacggtc tgacggcgcc caacgggtccc gccagcagc gcgtcatccg
 46551 gcaggccctg gccaacgcgg gcttggcccc cggcgacgtg gacgcggtcg
 46601 agggccacgg caccggcacc gtctctggcg accccatcga gggccaggcg
 46651 ctgctcggca cctacggcca ggaacggccc gccgaacggc cgttgtggct
 46701 cggctcggcg aagtccaaca tcggccacac ccaggccgcc gccggcgccg
 46751 ccggcctgat gaagatggcg ctggccctcc aacacggcac gctgcgcgcg
 46801 accctgcacg tcaccgagcc ctcgaccggg gtcgactggg cggccggcgc
 46851 ggtgcggctg ctccaccgag ggaacgtctg gccgcggacg gatcgtccgc
 46901 gtccggcccg ggtctcctcg ttccggcatc gcggcaccaa cggccacgtc
 46951 atcctggaac agccgcgccg cgagcccaac cccacggccc ctgccgaccg
 47001 cccaccccgg acgcccgcgg tctctccatg ggtcgtctcg gcccgatcgg
 47051 ccaccgcgct cgacgcgcag ctcgcgcgac tcggggcggt cggccgcgag
 47101 cggccggacc tgccgcccg cgaacgtcgc cactcgctcg tcaccagccg
 47151 cgccaccttc gaacaccggg cggctctgct ggccgcgccg gacggcatca
 47201 ccggcgccgc ccggcgcgag gcccggaac gcagcacccg gtctctcttc
 47251 tcggggcagg gtgtcagcg gtccgggatg gggcgtaac tgcattgtgc
 47301 tttcccggtg ttccggcgcg cgttcgacga ggtggtyggc gtgttgatg
 47351 cggagtggc gacgggttc cgtgggggtg tctcgctcg ggaggtgatg
 47401 tggggcgggg ggtcggagtt gttggatcgg acgcgtttca cgcagccggc
 47451 gttgttcggc gtggaggtag cgttgttcgg tttggtaggc tcgtgggggg
 47501 tggggcctga gttcgtggcg gggcattcgg tgggtgagat tcggcgcgcg
 47551 tatgtggccg ggggtgtctc gttgggtgat gcgtgtcgtt tgggtgtggc
 47601 gcgggcttcg ttgatggatg cgttgccggg ggggtggcgt atggttcggg
 47651 tggaggcggc cgaggcgag gtggtgccgc tgttggtcga tggggtggcg
 47701 atcgcccgcg tcaacgggcc ggtttcggcg gtggtctccg gtgtggaggg
 47751 ggccgttggg caggctcgtg atcagttggt ggagcggggc cggcggttc
 47801 gtcggttggc ggtcagtcac gctttccact cggcgttgat ggatccgatg
 47851 ttggatgcct tccgggccgt cggcgagggc ctggagtacc accagcccg
 47901 catcccgctg gtgtccaacg tgacgggcga ggtggcccg gccgaggagc
 47951 tgtgcgcggc cgactactgg gtgcggcacg tccggcgac cgtgcggctc
 48001 gccgacggcg tccgcacct ggccgagcgc ggcgccaccg ccttctgga
 48051 gatcgcccc gacggcgtag tgtccgccct ggcccgggc tcctgttcg
 48101 acacggacgc cgaagtggcg cccgcgctgc gcaagggcg ccccgaggag

10

20

30

48151 cacaccgccc tcaccgccgc cgcccaactc cacgtggccg gcgtggacat
 48201 cgactggacc gcggtcctgg ccggcaccgg cggcgggcgg atcgccctgc
 48251 ccacctatgc ctccagcgc gagcgtact ggcctcgct cgcgcacag
 48301 gccccggcg acgcggcg cgctcgccctg gaagccgggc ggcaccgct
 48351 gctcggggccc gcgaccaccg tcgccggatc cgcggagatc ctgctcaccg
 48401 gccgectgtc caccaccgcc cagccgtggc tcgcggtcta cgaggcgga
 48451 ggccgcaccg tcctgccggc cgcggtcctc gccgaactcg ccgtccgcgc
 48501 cggcgaccag gccgactgcc cgaccgtggc ggaactgacc gtccgcgcac
 48551 cgctcgctcct caccggcgcg gcggcccagc gcctccaggt ccgggtggcc
 48601 gcccccgagc acaccggccg gcgcgcgctg tcggtgcacg cccgaccgga
 48651 cgactccccc gacagccctt ggacgctgca cgcaccgcg gtctcacc
 48701 acgacacccc gcgcccccg gcgcgggaca ccggtggcc gccggagcgc
 48751 gccgtgcccc tcgacgcccct gccaccgcc accggcccgg cccggatcgc
 48801 ggccggcctgg cagtggggcg acgaactctg cgcgagatc gaactcccgc
 48851 aaccgggccc ggcgagcgg gcattcgccc tgcaaccggc gctgctggac
 48901 accggggtcc gcgcggcg cctgctggac ggcgacgcca cctggacgc
 48951 cctcggtgg cggggcccctg cctgcacgc cgcgtccgcc accgcccctg
 49001 ggggtccgct caccgggac ggcacggaca cctgggtctt ggaggccacc
 49051 gaccgcagcg gcgctccggt cgtctccgtc accgggctca cctggggac
 49101 gccacccgtc gaccggtcgg gggccggggc ggccgatgac ggcgcgacc
 49151 tgctcgacct ggagtgggtg cccgcgccgc aggcgcgcc caccggcggc
 49201 gaccacctcc cgtacgccgt gctcgcgat caactcgcg agctggacgg
 49251 gcagttgagg atcgcgggc acgggcccgc gcgctcgca tcgctggccg
 49301 cgctgctgga cggcggtgc ccgctgccc ggctcgctc cgcgcgggtg
 49351 ctggcgctgc cgaccgggga aggcgacctg cccgcgcgg tgcgcggcac
 49401 caccacggcg gtgctggagc tgctgcagcg ctggaccgcc gacgcccga
 49451 ccgcccagag ccacctggtg atcgtcacc gcggcgccgt cgcgcggg
 49501 gcggaggacg tgcacgacct ggcggcggcc ccggtctggg gcctggtcg
 49551 ctcggcacag tcggaacacc ccggcagctt cctgctgctc gacctcgacc
 49601 ccgcgctacc cgcgggagcc tcccgcgcgc cgcgcgggc caccctggcg
 49651 gccctgctcg acgcyggcga gaccagggc gcggtgcgc ccgacacgt
 49701 caccgtcgcc cggctgacct gggccgcga cggaccgag gccaccgcc
 49751 gacacccggt gcgggactgg gaccgcgac gcaccgtcct gatcaccggc
 49801 ggcaccggcg gcctggcgcg cctcctggcc cgcacctgg tcaccggaca
 49851 cggcatcaag cactgctgc tcgcccggcg ccggggccc gacgccccg
 49901 gcgcgcgggc cctgcgcgac gaactggcg cctcggcgc cgaggtgacc
 49951 gtcccgccct gcgacgtggc cgaccgtgcc gcaactggac gactcctcgc

10

20

30

50001 gcaactgccg ccggagcacc cgtcgaccgc cgtcgtgcac accgccggcg
 50051 tcctcgacga cgccaccgtc ggcaccctga cgcccgagcg gctggacacc
 50101 gtcctgcgcg ccaaggcggg cggcgcttg caccctgcac acgccacccg
 50151 cgaccgcgac ctggcagggt tcgtgctgta ctctcggtc gccggtgtca
 50201 ccggcgggcc cgcccgaggc aactacgcgc ccggcaaac gtctctcgac
 50251 gcgctcgccg cgcaccgcgc cgcccgaggc ctgcccgga tgtcgctggc
 50301 ctggggaccg tgggggcagg acgccggcat gaccggcacc ctcggcgcgc
 50351 ccgacctggc ccgcttgagg cgtccggca tgcgcgcgt caccocggaa
 50401 caggggcctg ccctgttcga cgcgcgcgc gcccgcgcg accgggttcgc
 50451 ggtggcggtg cggctcgccc gtggcgccgc cgcaccgggc gccgacgagg
 50501 tcccgcggt gctgcgtgcc ctgggtgcgc gccggcgccg caccggcgcc
 50551 gcggccgggc acccgcggtg actggccgc cggtggccg ccctggacgc
 50601 cgagcagcgc catcaggcgc tgcctgacct ggtccgcacc gagacggccg
 50651 cgggtgctcg ccaactccgg gccgacgcgc tcccggccga gcgggaactc
 50701 aaccggctgg gcttcgactc gctgatggcg gtcgaaactg ggacgcggct
 50751 ggccaccgcc acccgagccc ggtgccgcg caccgtcgtc ttcgaccacc
 50801 cgacgcggga cgcggtcgcc cggcacctcg cgtcgacgct gcccggtggg
 50851 acccgggccg gtccggaccg tcccccgctg gccgaactcg accggatcgc
 50901 cgccgagttg tcgcccgaag gcgcggacga cgccaccga caggcgctcg
 50951 tcgggcggct gcggcacctg ctggcgcagt gggacggcac ccgacaggac
 51001 ggcggtggga cgaccgtcga cgaccgcac gaagcgcgga gcgccgaaga
 51051 ggtctctgcc ttcactgacc acgagctcgg ccggcaggcg gactcctgac
 51101 ccgccccact ccgctcgctc gcgcgcacca catctgaggg aggtttcacg
 51151 gaccatgccg gacgaaaaga agctcgtcga ctatctgaag tgggtcacga
 51201 aggacctcca ccagaccgc cagcgcttc aggaggtgga ggcggggcgc
 51251 caccgaaccg tggcgatcgt cggcatggcc tgcgccttc ccggcggtgt
 51301 gcgctccccg gaggaacctg gggagctgct gtccgcgggc cgggacggca
 51351 tcggggccgt ccccgccgac cgcggctggg acctggcggc gctggccggc
 51401 gacgggcccg gtccgagcgc caccaggaa ggcgggttcc tgcgccagc
 51451 ggccgccttc gaccggggt tcttcgacat ctcccgccg gaggcgctcg
 51501 ccatggaccg gcagcagcgc ctgctgctgg agaccgcctg ggaaggccgtc
 51551 gaacgctccg gcctcgaccc ggccgggctg cgcggcagcc gcaccggcgt
 51601 tttcgtcggc accaacggcc aggactaegc gcacctggtc ctgcgcgcgc
 51651 aggacgacat gggcggtac gcgggcaac gcctggccgc cagcgtgctc
 51701 tccggccgac tggccttcgc gctcgccctg gaaggcccg ccgtcacctt
 51751 cgacaccgcc tgcctcctcg cactggtgac cctgcacctg gccgcacagg
 51801 ccgtgcgcgc cggcggaatg ggcctcgccc tggccgggtg cgtcacggtc

10

20

30

51851 atgacgacct cgtcgagctt cgcggtcttc agcctccagg gcggcctggc
 51901 gccggagcgc cgtcgcaagg cgttcgccga ggccggccgac ggcaccggct
 51951 ggtccgaggg catcggcctg ctctctctcg agcggctctc cgacgcgcag
 52001 cgcaacggcc acccggtgct cgcggtgctg cgcggctccg ccgtcaacca
 52051 ggacggcgcg tccaacggcc tcagcgcgcc caacggtccg tcccagcagc
 52101 gggtcacccg ccaggcgctg gccggcgccg gactcgctcc cggcgacgtg
 52151 gacggcgctg agggcgacgg caccggcacc cggctcgcgg accccatcga
 52201 ggccgggtcg ctgctcgcca cctacggcca ggaccggccc gccgaccggc
 52251 cgttgtggct cggctcggtg aagtccaacc tcggccacac ccaggccgcc
 52301 gcgggcgctg ccggcgctcat caagatggtg ctggccctcg ggcattggcgt
 52351 cctcccgag accctgcacg tggacgcgc ctcctcgcac gtcgactggg
 52401 agagcggcgc ggtgcggtg ctcaccgcac ccgtcgctcg gtcgagggc
 52451 gacgaccggg tgcgcggggc cggcgctctg tcgttcggca tcagcggcac
 52501 caacgcccac gtcctcctcg aacaagcccc cgatcagccg gaaccgaccg
 52551 cggaagagac ggctgccgcg gcgcccggcg gcaccgccga ggagcggggc
 52601 gccgctcccg tcgcccgcg cgcggtgccc tggccggtcg cggcacgcac
 52651 cgcgcggccc ctgcagccc aactggtccg ggtccgcgcg ctgaccaccg
 52701 cgcgcggccc caccgcgcgc gacgtcggtc acgcgctggc caccgcccgt
 52751 acccccttcg agcaccgggc gctgctggtc cagcagggcg gcgccgtcac
 52801 cgaggtggcg cgcggcgccg tcccaccgg tgaccggggc gggctggccg
 52851 tgctgttctc cggacagggc tcccaacggc cgggcatggg gcggaactc
 52901 cagcccgct acccggtctt cgcgcgcgc ctcgacgaga ccgtcgccct
 52951 gctcgacgcc cggctcgcca cgtcgctcg cgacatcgtc tgggaccagg
 53001 accgcacccg gctcgacgac acccgccaca cccagcccgc gctgttcgcc
 53051 gtcgaggtcg cgtgttaccc cctgctggcc tcctggggca tccggcccga
 53101 ccacgtcacc ggacactcca tcggcgagat caccgcggcg cactcgccg
 53151 gtgtgctgac cctcgcgac gcctgcacc tggtggcgc ccgcgccacc
 53201 gccatgagcg aactgccgcc cggcggcgc atggtggcgc tggaggccac
 53251 cgaggacgag gtgcgtccgc tgctcaccga cgacctcgcg atcgccgcg
 53301 tcaacgcccc ccggtccgtg gtcgtcgccg gcgcccagga cgcgcctc
 53351 gccgtccgcc ggcacttcga cgacctgggc cgcggacca cccggctccc
 53401 ggtcagccac gccttccact cgcgcctcat ggaccgatg ctcgacgcct
 53451 tcggagcggc ctcgccccg ctgaccttcg ccgagccgga gatcccggtc
 53501 gtctccaacc tcaccggcct cccggccacc gccgaggaac tcgccacccc
 53551 gcaactactg gtgtgccacg tccggcaggc cgtccgcttc ggcgacggcg
 53601 tgccgcccct cgcgcaccgc ggcgtgcgga ccttcctcga actcggccc
 53651 gacggcgctg tgtccgccct ggtccgggag aacctcccc agccgggcct

10

20

30

53701 gggtcgccgtg cccgtgctgc gcaaggagcg gcccgaggag accaccgtgc
 53751 tggccgccct gggaaccctg tgggcgcacg gcgcggacgt ggactgggac
 53801 gcggtgttcg ccggcacccg caccocgcag gccgaccocg tcgagctgcc
 53851 gacgtacgcc ttccaacggg cccgctactg gcccaacctc ggccgccgcc
 53901 acggcgaccc ggccgacctc gggcagaccg ccgcgcacca cccgtgctg
 53951 ggccgccgcc tcacctcgc cgacgccgac gagaccgtgc tcaccggccg
 54001 cctcgcgctg cctccccc cctggctcgg cgaccaccgc agcgacggcc
 54051 ggatcaccgt ccccgccgct gccttcgcgc aactcgccgt ccgcgccggc
 54101 gacctgagcg gcaccccgca cctggcgccg ctcgacctgc cggcgccgct
 54151 caccctcggc gacggcgaca ccgtcaccct ccagggtccg gtcggcgccc
 54201 ccgacccgcg ggccacccgg ccgctgaccg tcacgcgcgc cctcgacgcc
 54251 accgaggacg cccctgggac cactgcgcg accggtctgc tcgcccggga
 54301 cgcccccgaa gcgcgcgcgg atccgatcgg cccggccgac gccgggtggc
 54351 cgccgcggga cgcccgcccg gtgcccgctc cgacacctga cgcggccgcc
 54401 accgcgcgag gccgcacta cggcccccac ttccagggcc tgaccgggct
 54451 ctggcgccgc gacggcgagg tcttcgcgga ggtggccctg cccaccgcca
 54501 ccgcgcgcga ccgcgccttc ggcattccac ccgcgctgct ggccaccgcg
 54551 ctccgcgcca ccgcgcact ggacgacgac cacaccgcgc gccacacccc
 54601 cgaaccgacc ggcatcaccg gactcgccct gcacgccacc ggggccaccg
 54651 caetgccggt ccggtcacc gcgaccgggc ccgacaccgt ggccctcgc
 54701 gccgcggacg ccacggcgcg cgcggctcct accgcgcgac ccgtcaccct
 54751 cggctccccc caggaccgc ccgctccgc accggccggc cacaccgggc
 54801 agggcgccct gttccacctc gactgggtgc cggtcgacc cggcagccga
 54851 gccaccggca cccgtgggc cgtcgtcggc gacgacgaac tcgacctcgg
 54901 ctacgcctc caccgcgcgc acgagacggt cagtgcctac gcggcgctgc
 54951 tgggcggagc catcggcgac agcggtctgg cgcgcgacgt ctccctcgtc
 55001 cccgtcgtcg gcggcccgga cgcgggccc gacgcggtgc accgcgtcac
 55051 cgcgcgcgc ctggggctgc tccaggagtg gctgaacgag ccgcggttgg
 55101 ccggcgcccg cctggtcttc gtcacccgcg gcgcgcgcgc ggtgccgggc
 55151 gagaccgtca ccgaccggc cggcgccgc gcttggggcc tgcctgcctc
 55201 cgcaccagac gagaaccgg gcagctctgt gctggtcgac ctcgacgacg
 55251 cgttcgggtc cgcgggatg ctgccgacg tccctaccct cgacgaacag
 55301 cagctcgtcg tccgcgacca cgcggtccgc gccgcgcgc tggcccggt
 55351 gccggagccg gccgcggca ccgcgcggc ccgcgcctgg gaccggacg
 55401 gcaccgtcct gatcaccggc ggcaccggcg gcctggcgcg cgcgctcgc
 55451 cgcacacctg tcaccgtccg cggcgccgc cactgctgc tcgcccgcg
 55501 ccgcggcccc gaggcgcgg gcgcggcgga actggtggcg gagctgaccg

10

20

30

55551 cacagggcgc ggacgtgcgg gtggccgcct gcgacgtcgg cgaccgcacc
 55601 gccctcgacg cgtcctcgc caggtcccc ggggcgcacc cgtgaccgc
 55651 cgtcgtgcac accgcggcg tctggacga cgcctgac ggctcgtca
 55701 cccccgacca actggccacc gtgctacggc ccaaggccga cgcgcctgg
 55751 catctgcacg acgccaccgc cggcctcgac ctggccggct tcgtcctgta
 55801 ctctcggtc tccgggtcc tgggcagccc cggccagggc aactacgccg
 55851 ccgccaacgc ctacctcgac gcgtcgcgc ggcaccgcgc cgaccagggc
 55901 ctcccggcgc tctccctcgc ctggggcccc tggggtcggg gcagcggcat
 55951 gaccgcgtcg gtcagcgacg ccgacctgga gcggtggcg cgcggcggcc
 56001 tgccgcgcgt gaccgtcgag gacggcctgg cctgttcga cgcgcgcgtc
 56051 ggccgccccg agccggccct ggtgcccagc cgcataacg tcgccggcct
 56101 gcgggaccag caggcactgc cggcactctg gcgcgacctg gtaccgcggg
 56151 ccgcgcgcac cgcggccacc gccgaccgct ccccggtcac ggtgcgcgag
 56201 cgcctcgcgc acctcgacga gaccggccag gacgagctgc tcctcgacct
 56251 cgtcgtcggc tacaccgcgc gctgctcgg ccaccccgac cccaccgcgc
 56301 tcgataccga acggggcttc ctggagctgg gcttcgactc cctggtctcg
 56351 gtcggcctgc gcaaccagct cgcgagatc ctggcctgc gctgcctc
 56401 gtccatcgtc ttcgacagca agtcgcgggt gaagctggcg cgttggctgc
 56451 accaggaaact cgcacaacgc ccccgccgg gcgccaccgg ccccgccgc
 56501 gcggacgcgc gtcccgcggt gcgtccgac gacacctgg agggcctgtt
 56551 ctacaacgcg gtgcgcggcg gcaagctcgt cgaggcgatg cggatgctca
 56601 aggcgctcgc caacacccgc ccgatgttcg acaccccgcc cgagctggag
 56651 gagctctccg agccggtgac gctgcgcgac ggcccgggcc ggcccggtct
 56701 gatcttcgtc agcgcgccgc gcgccaccgc cggcgtccac cagtacgcgc
 56751 gcatcgccgc gcacttcgc gccagccgc atgtctccgc gctgcccctg
 56801 atgggcttcg ccccgggcga gctcctccgc gccaccagcg aggcgcgggc
 56851 ccgtatcgtc gccgagagcg tctgatggc cagcgagggc gaaccgttcg
 56901 tcatggtcgg caactccacc ggcggtcgc tggcctacct cgcgcgggc
 56951 gtcttgagg acactggga cgtccggccc gaagcggtgg tctcctcga
 57001 caccgcgtcc atccgctaca accccggcga gggcaacgac ctggaccgca
 57051 ccacgaggtt ctacctggcc gacatcgact cgcctcgggt gacgtcaac
 57101 agcgcgccga gtccgccat ggcccactgg ttcattggcg tgaccgacat
 57151 ccaggcgcgc gcaccgaccg cccccacct cctcgtgcgc gccgcccggg
 57201 cctcgcagcg ctccggctc gacacctcgt ccgtccccgc cgacgaggtc
 57251 cgggacatcg acgcgacca cctctccctc gccaaaggag actcggcact
 57301 gaccgcgcag gccatcgagg gatggctcgc ggaactgcgc gaccccgcg
 57351 cctgatctcc ggcccggcgc gcccccga ca ggccgggccc atccccccac

10

20

30

57401 cggggaggcg ccaccaggt gcgccccggc gcaccgggc ccgcgcaggc
57451 agccggagcg cccatcccc cgcctggcac caccctccag aggagtcctt
57501 ccatgagcac accgaccgca ccgcccctcc tgaaagcgga ggtgccgccc
57551 gtccctgcgc tgagcccgcg gctgcgcgaa ctccagtcgc gcgcccccg
57601 ctgcaaggtc cgcacccccg ccggcgacga gggctggctg gtgacccggc
57651 acaccgaact caagcagctg ctgcacgacg accggctggc ccgcgcccac
57701 gccgacccgg ccaacgcccc gcgttatgtg cacaacccgt tcctggacct
57751 gctcgtcgtc gacgacttcg acctggcccg cagctgcac gccgagatgc
57801 gtccttctgt caccocgcag ttctcggccc gcgcgctcat ggacctgacg
57851 ccgagggtgg aagccctcgc cgagggggta ctggccact tcgtcgcaca
57901 gggacgccc gccgacctgc acaacgactt ctgctgccg ttctccctgt
57951 cgggtgctgt cgcgctcatc ggcgtcccg ccgaggaaca ggggaagctg
58001 atcgccgccc tcacaaact gggcgaactc gacgacccgg cagcgtcca
58051 ggaaggccag gacgagctgt tcggcctgct gtccggcctg gcacgcgcga
58101 agcgcctcac acccgaggac gacgtcatct cccggctctg cctgaagggtg
58151 ccctccgacg agcgcctcgg ccgctcgc tcgggtctgc tcttcgcccg
58201 cctggacagc gtcgccagcc acatcgacct gggcacggtg ctgttcaccc
58251 agcaccggga ccagctcgc gcggccctgg ccgacgagaa gctgatgcgc
58301 ggcgcgctg aggagatcct gcggctccgc aaggccggcg gttcgggtgt
58351 cccgcggtac gcgaccgccc atgtaccgat cggcgacgtg accatcaggg
58401 ccggcgacct ggtgctgctg gacttcaccc tgggtgaactt cgaccgcag
58451 gtcttcgacg agecggagct cttcgacatc cggcgcgccc ccaacccgca
58501 cctgacgttc ggccacggca tgtggcactg catcggcgcg ccgctggccc
58551 ggggtcaatct gcgcaccgcc tacaccctgc tgttcacccg cctgcccggc
58601 ctgcgctggt tgcccccggg cgaggaactg cgggtgctgt cggggcagtt
58651 gtcggccggc ctgacggagc tgcccgtcac ctggtgacgt gatgtcggac
58701 ccgcccggcc cggttccggg caggcggaag tgagggcccc cgcccgcctc
58751 ggcggggggc cctcacgcac gggggagcgg gctcctcagt ccgcccgcac
58801 gctgatcgcc ccggacgggc agagcgccgc cgcgtcgcgt acgtcgccgg
58851 gatccgcggc gtcgcgcgc ccggccagga cggtcaccag accgtcgtcg
58901 tcctggtcga acaggtccgg tcgggtcagg acgactggc cgccccgac
58951 acagcgcccg ggggtccagg tgatgcgcac gatcgttcct ccgagggtcg
59001 gttcgtcgc ggggtcgggc cgcgcgggg cggcccgacc gggctcactt
59051 ccaggtgacg ggcagttcgt gcaggccgaa cagcaccocg tcgtacttca
59101 acggcagttc ctccaccggc acggcgagcg tgagggaggg gatccgggcg
59151 aagagcgtgc ggtaggcgat gtccatctcc tcccgacca ggttctggcc
59201 caggcactgg tggacaccgt agccgaacgc cagctgcgag cgcgcgagc

10

20

30

59251 gctcgggggtc gaactccgag ggcgccgcga aggccgtggc gtcgtggttg
 59301 gcggcgccga tcagcggcac gatgccctcg ccggcccgga tcagctgacc
 59351 gccgatctcc acgtccctga cggccacccg gaatgccacc agatcggcca
 59401 ccgagtagta gcgcagcagt tctcgcgga tccggtcgtc gccgatccac
 59451 cgcgggttgg tgagcaactg caccacgccc aggccgatgt tgttcgccgt
 59501 cgtctcgtgc ccggcgatca gcagcagcat cggcaccocg gacagctcct
 59551 gccgggacag tgtgcccgcg gcgacagcc ggcgtgatcag gtcgtcacccg
 59601 ggtcgcgcgt gcttgatctg gatgagtcgg ccagatagc gcagcagcgc
 59651 ctgggtcgcc ttgtcgcgt cctcgtcggc cgaactgagc cggaccagca
 59701 cccgggtccg gtcctcgaag aagtcccggt cgaccttggg caccocgagc
 59751 agactggaga tcaccaggga cggcaccggc agcgcgaagg actcgacgag
 59801 gtccgcgcgag gtgcccgcg ccagcatcgc gtcgatccgc tcgtccacga
 59851 tctgctggat caccggccgc agttcccgca ccttgccac ggtgaactcg
 59901 gggatcaagg tcttgccgaa ccggccgtgc tccggtgggt ccatcgccac
 59951 gaaccacccc ggcacctcgt actgcgacgg ggcgccaccg gtcgcaccgg
 60001 ccttggggaa accgggtttg gaggggtcgg cgctgatccg ggggtcggtc
 60051 agcaccgccc ggcgcctcct gtgccgggtc accaaccaga ccgggcccgt
 60101 ggggagttcg gagcggaaca ggcccgcccc gccgcgtag gtcgcgtact
 60151 ccggcggggg gaaggcccg cccggcctgc gcagcggaa ggccaccggg
 60201 cactgcggcg cggcggtccg cgcgtcggct tcggtgctca tgccactccg
 60251 tagaacgcgc ggatccggcc ggtgatgaac tctgctcct gtgcggtcag
 60301 tccggtgtgc gtcggcagat agaagccgtc ggcgctgagc cgggcccgt
 60351 tcaacgcggg ccagtcggcg gagaagtaac cgggctgacg gtcctcggc
 60401 ttgaagaaga cccgggtctc gatgccctcg ccggcgaggt acgcgcacag
 60451 ctcgtcccg cgtcgcggcc gcaggtcgtc catccacagc acgtcgcgog
 60501 gcggcatcag cgtgatgccg gggatgtcgc gcagcgcctc gtcgtagcgc
 60551 ttctcgatgt cgcggcgca ggcgaggatg gtgtccagct gtcggtctg
 60601 cgcgagcgcc acccgggcct gcatgttgg catccggaag ttgtaggcca
 60651 gcttcttgtg caggaaagctg tggctccttg tgaacgccat cgcgccaga
 60701 tgggccatct gctcgccag gtgcgggtcg tgggtcaggc agacgccgc
 60751 ctgcgcggcc gagatgatct tgttgccgaa gagcgagaaa caggcgatgt
 60801 cgcgcccgcg ccgcaccccg tgcgcctcg cggagtcctc caccacccgc
 60851 aggttgtaact cgtacgccag gttcagcacg gcgtccatgt cgcactgccg
 60901 gccgtagatg tgcaccggca tgatacttt ggtgcgcggg gtgatcttct
 60951 cctcgatgag ggacacgtcg atgttcaggt cgtcgcgcga gtccacgaac
 61001 accggcgtgg caccggtgta ggtcacgcc caggcggacg cgatcatcgt
 61051 gaactccggg acgatcacct cgtcacggg gccgacgcc agcgcgcgca

10

20

30

61101 gcgcccagcgt cagcgccgtg gtgccggagg agcaggcgac gccgaacggc
 61151 acgtcgttgt acgcgggcgaa cgccctctcg aaccgcctga cgtacggccc
 61201 ctgcgaagag atccagccgc cgccgacggc ctccgtcaca tagtcgagct
 61251 cgcggccctg gagccacggc atggacaccg gatacgtaaa ggacatgggt
 61301 ttctagatctt cctcggtcag tcgggttgcca gggcgggaa ggcgagcagc
 61351 aggtccgcgc cgcgggccc gcccgcggcg tcccgcagca gcccggcgaa
 61401 gtgctccgcg cgtcgggtga aggagggttg gtcgagcacg cgggtgatct
 61451 tgtccaggac gtccctcggg tccacggtct ccggccggtc cagggtcagg
 61501 ctcaccccca agtcctggcc ccggatcgcc tggctcgtcg agtccacca
 61551 caacggcccg accaccagcg gctttccgaa gtacaggccc tcgtggtagc
 61601 cglttgccacc ggcatgggtg aagaacgcct tcacgttcgg atgggccagc
 61651 acgtccagct gcgacggcac ccagccctcg atccgcagg tgcctggcag
 61701 ctggcgggcc ggcggcagca actcctgttg gccgcgggg agtttccaca
 61751 acacctgggtg gcccggcccg tccagtcgcc gggcgacctc caccagcgac
 61801 gccacctgct caccgggtcag ccgggtgatc gtgccgaagc ccatgtacac
 61851 caeggacttc tgcgcccaga gccagtcgga caggccgtcg tcgtccggtg
 61901 cctggggcag cggcggcacc atcgtgccca ccagcccgag cttcggatgc
 61951 atcgggaacg ggtagtccaa ctcccttacg gagtagcaca agacctgctc
 62001 cgcattggtc atccgcgcca tcattctgcg cgccctggggc gcgatgccca
 62051 gctcgggtgc gaccgggttg tcctctctga cgaccttgcg gacgtccgac
 62101 gtcaggaaac tcccagcgtt ccgcagccgg aacagctggt tctcgatccg
 62151 ctgagccagg gacatcgccg ccggcagccc cgagtgcggc accgggaaac
 62201 ccgacggggg gtaggacttg gcgaacggga cgtgcgaggt gaggacgttg
 62251 ctgggcacga accggacccc gagcacgaac ggaatgcctc tggatgatcg
 62301 cagctcgtac ccgaactggc acatgctctc gatcaccatc agcgccggct
 62351 cgacctctc gacgatctcc tccaggcggc ggtacttcgc catccgcgac
 62401 tccggcgcca acgaatgcg aatcccgcc gcgtgcgcct tgaaccgcca
 62451 ccgtgcgtc acctccgcat acgtcgcgtc gtcccattgt accgccgaca
 62501 tctgcgagac ggtgtcgcg agcgacgcca accgaaccgg gctgccgtcc
 62551 accacggccg ccacctcgtc gcgcgcttcc tcgtcgggtg cgaaccacag
 62601 gtccgccacg tcgcgcgggg acaattcccc ggccagcacg agcagcgat
 62651 tgagcaggcc gctttcgcca taactgacga acaggatcgg ccgccgattc
 62701 gcgcccattg acaacacccc tcggaatgtg gcgggcccgc gggcccgcgc
 62751 gccacgcacc cgcggggccc ggtcgcgggg tgagtgcatt cgccgacgcc
 62801 gccacccgag gcgcgtgttg ccggaaggaa gggtcaccgg ccggcaccgc
 62851 gaacgcgcgc cgtggaaaac gggtcgggta cttggtctca tgccacggac
 62901 cgggggaatca ctagtcttcg gcgcgcgacg gccctttccg ggccgtgtgg

10

20

30

62951 ccaatgcccg tccccggcgc ccgtcattcc ttagggaaaa gtacagcgtt
 63001 tgcgaacgta cgatccggca cgcagaggcg acctgaggcc aacttttccg
 63051 caggggtgag caaggcatga cgatcggagc cgacgaggac ccggtggtgg
 63101 tcgtcggaat ggcttccgtt tatccgggtg gggtcgccgg ccgaggagac
 63151 ctgtgggaac tgggtccgac cggccgcgac gcgaccaccg ctttcccgga
 63201 cgaccgcggc tgggacctgg ccgcactggc cggcgacgga ccggccgca
 63251 gcgcgacccg cgaggcgga ttcctcaccg gcgcgcgga cttcgacgac
 63301 gccttcttcg gcatgtcgcc ccgcgaggcc gtctccaccg acccgcaaca
 63351 gcgcctcgtc ctggagaccg cctgggaagc cctggagcgc gccggcatcg
 63401 acccgcactc cctgcgcggc agccgcaccg gggctcttct cggcgccagc
 63451 ggccaggact acgcgcctt caccacgccc tcgcccagc accctggacgg
 63501 acacgccttc accggcctgg ccccgccgtt cgcctccggt cgcctggcgt
 63551 acgtcctggg cctcgaaggc cccgcgctca ccgtcgacac cagtccttc
 63601 tcgtcgtcgg tcgcgtgca ctgggcggtc cgcgcctcgc gcgcggggga
 63651 gtgcagcacc gccctggcgc gcggcgctac ggtgatgtcc acccggccgc
 63701 ccttcgtcgg ccacacccga cagggcggcc tcgcgcgcga cggccgctgc
 63751 aagccgttct ccgacgagc cgacggcacc gcctgggcgg agggcgctcg
 63801 catcgtcgtc ctggagcacc tgctccaccg ccgcgcgcgc ggcaaccgcg
 63851 tcttcgcgtt gctgcgcggc tcggccgtca accaggacgg cgcctccgac
 63901 ggcctcaccg caccagcgg tcccgcccag gaacgcgtca tccgcgcgc
 63951 cctcgcgcgc gcccgactcg ccccgccgga catcgatctc gtcgaggcgc
 64001 acggcaccgg caccgcgtc ggccgacccg tcgaggccgc ggcgtgctc
 64051 gccgcctacg gccaggaccg ggaccgggac cgaccgctgc gcctcggttc
 64101 cctgaagtcc accctcggcc acgcacaggc cgcgcgcgcg atcggcggag
 64151 tgatcaagac cgtcctgacc ctgcgcgacg gcctgatgac gcgcacccg
 64201 caccctggcca ccccccaccg ccaagtcgac tggteccagg gcgcgtggc
 64251 cccctcaccg gaccacacgc cctggccacc ggccgaccga ccgcgcgcg
 64301 ccggcgtctc ctcttcggc atcagcggca ccaacgccc tgtgatctc
 64351 gaagaggcgc cgcgcgcga cgtcccgctc acccgcccg gcaccctccg
 64401 ccccgaccgc gtccctggc cgggtctcgc cgcacgccc gaagccctcg
 64451 acgcccact cgcgcgctc cgcgcgcacc tgcgaccca ctcggacctg
 64501 gaccgcgtgg acgtcggcta ctccctggc accggccgcg ccgcgtccg
 64551 ccaccgggcg gtcctcctgc cgcgcgcga cggcaccgac gcggacgccc
 64601 tcgagcagc ccgcgtgcg gccaccagc gccgcaccgc cgtcctcttc
 64651 tccggccagg gcagccagc cccggcatg ggccgcgaac tcgcccgcg
 64701 ctcccggtg ttcgcgcag cactggacga cgcgtcgcg gccctggacc
 64751 ggcacctgga cggcccggtg cgcgaggtga tgtggggcac cgaecgcgcg

10

20

30

64801 ctctctggacc ggaccggctg gaccacagccc gccctgttcg ccgtcgaggt
 64851 cgcctctccac cgcctggctg cgtccctcgg cgtcaccccc gacttcgtcg
 64901 gcggccactc cgtcggcgag atcgccgcgg cccacgtcgc cggcgtcctg
 64951 agcctggagg acgcctgccg cctgggtggc gcccgcgcca cgtgatgca
 65001 ggcgtctccg gcggggcgg cgtggccgc gctggaggcc accgaggacg
 65051 aagtggcccc gctgctcggc gcacacctcg cgtggccgc cgtcaacggc
 65101 cccaccgcgg tcgtcgtcgc cggagccgag gacgccgtgc ggcaactgac
 65151 cgcctccttc gccgaccgg gccggcgcac cagccggctg gccgtctcgc
 65201 acgccttcca ctgcctcgtg atggagccca tgctcgacgc ctccggggac
 65251 gtcgtgagcc gactgacctt ccaccagccg tcgatccgc tggctctcaa
 65301 cctcacccgt gaactcgccg gcagtgaat caccagcgc gagtactggg
 65351 tccggcacgt ccgcgacacc gtccgtctcg ccgacggcat caccgcactg
 65401 gccaaaggcc gcgcgcagct cctgatcgaa ctccggcccc gcggcgtgct
 65451 gtccgcgatg gcccgcgaca cctcggccc cgacagcacc accgacgtcg
 65501 tccccgcct gagcaaggga cggcccgagg agaccgcctt cgcggcgcc
 65551 ctccggccgc tgcacacct cggcgtcccc gtcgactggc ccgccttcta
 65601 cgcggcacc gcgcgccgc gcgtcgaact gccacctac gccttcacgc
 65651 acgtgcgcca ctggccacc ccgcgccgc cgaacggcgc cgggcccggc
 65701 gccctcggcc accccctgct cggctccgc gtcgaactcg ccgacggcgg
 65751 cggcacccgtc tgctccggcg cctctccct ccgcacccac ccttggtcgtg
 65801 ccgaccacac cgtcgcggcg cgggtcgtgc tgcggccac cgcgtgctg
 65851 gaactcgccg tgcgcgcgg cgacgaggcg ggtgcgacg tctgcacga
 65901 actccacctc accaccgcgc cggccctgcc cgacgacgc gccctgcacg
 65951 tccagtgca cgtcggcccc gccgacacca cggggcgccg cgcctcacc
 66001 gtccacacc gccccgacca ccccccggc gcgactgga ccgatgcgc
 66051 caccggcacc ctccggcagc ccccgccgtc cgcagccgaa gccgccacgg
 66101 gcggcaccgc ggccgcctgg ccgcggcgc acgcggaacc cctcgacctc
 66151 gccgaccact acgaacggct cgcgaccgc ggcttcgact acggcccgac
 66201 ctcccgcggc ctgcggggcg cctggcgacg cggcgcgagg atcttcggcg
 66251 acgtggaatg ccgcgccggc accgcgacg acgccccga ccacggactg
 66301 caccccgcgc tgctcgacgc ggcccggcac gccgcatgg cggtggaagg
 66351 caccgtgccc gtgcctggc acggcgtccg gctgcacgc gtcggcgcca
 66401 ccgcgtcgtg ggtccgcac cgcgccacca cgaccggcac gctgacctc
 66451 accgcggctg acgtgcacgg cgcgcggctc gtcacgtcg aggcctcac
 66501 cgcggcgccg ctgaccgacg aggaacgcgc cgcggcgcg acgcgcggc
 66551 agggccgcgg cgagacgccc gccgacgccc gcccgggccc gcccgcgcg
 66601 gcccgccccg gcccgggcgg cgaacccctc ccggacacca ccgggtcca

10

20

30

66651 cccacacgcc ggccacctcg ccgcgctgcc gccggccgcc cgggagcgcc
 66701 agctgctgga cctgggtgcgc acccaggccg ccgccgtcct gggccacccc
 66751 ggccccgagg ccgtcggcac ccgcagcgtc ttcaaggagc tgggcttcga
 66801 ctggttggcc gggtcgaaac tcgccgaccg gctcaccgcc cgcaccggac
 66851 tgccgctgcc ggccaccctc gtcttcaact tccccacccc cgaacgtgct
 66901 gccacacgcc tcggggaaact cctcgccgca accgcccccc tcgaccccg
 66951 gggtacgga gaggaactca ccaggttcga ggcatcgtg acgaacctgc
 67001 cgcaggacgg ccccgaaacgc cgggccgtcg cggaccggtt ggacgccatc
 67051 gtctccgcac tcgccagaa ctcgctgca gaggtgccct cctcggacga
 67101 ggacatcgac acggtgtcgg tcgacagact gctcgacatc atcgatgaag
 67151 agttcgaaac cacatagaga aattgttgct ttcgttcgg acccgatgac
 67201 gaggacggac cgtgacgga accccagcaa ggccagccgg accagcagga
 67251 gaaaatcgtc gactatctcc ggccgggtcac ttcagatctt cgcggtgcc
 67301 gccgccgcat tggcgaaactg gaatccaagg acaacgagcc catcgccatc
 67351 gtgggaatgg gctgccgact tcccgccggc gtcaattcgc cggaatccct
 67401 gtgggacctg gtggttccg gcggcgacgc catttccgga tccccgtcg
 67451 accgcgctg ggacctggaa acctcaccg gaaacggcga cggcagcgc
 67501 gccaccacgc aaggcggtt cctctacgac gcccggaat tcgacgcgc
 67551 ctctctcggc atctcgcgc gcgagcgac tgcctatggc cccagcagc
 67601 gcctcctcct cgaagtgcgc tgggagcgcc tggagcgcc cggcatcgcc
 67651 cccacagccc tcgcggcgag ccggtccggc gtgttcgtcg gtcctacca
 67701 ctggggcgcg ccctcgccg acgccgccac cgaactgcac ggccacgcc
 67751 tgaccggcac cgcgccagc gtgctgtccg gccgcctggc ctacaccctc
 67801 ggccctcgaag gcccgccgt caccgtcgac accgcctgct cctcctccct
 67851 cgtcgcctcg caccgtggcg cccagtcctt gcgcgtcggc gaatcctcgc
 67901 tcgccgtgat cggggcggtc acgacctca ccgagccgtc cgtcttcgtc
 67951 gagttcagcg ccagggcgcg cctggcaccg gacggccgtt gcaaggcggt
 68001 ctccgacgcc gccgacggca ccggttgggc cgaggcggtc ggtgtcctcg
 68051 tcgccgagcg gctctccgac gcgcagcgca accgccatcc ggtcctcgcc
 68101 gtgctgcgcg gctccgctt caaccaggac ggccctcca accgctgac
 68151 cgcacccaac ggccctccc aggaacgggt catccagcag gccctcgcc
 68201 ggaccggcct gacccccgc gacatcgac ccgtcgaggc gcacggcacc
 68251 ggacccggc tcggcgccc catcgaggcc caggccctgc tcgccacctc
 68301 cggccagggc cacaccccc accagccgt gtggtcggc tcctgaagt
 68351 ccaacatcgg gcacaccag gcggccggc gcgtcgccg tgcctcaag
 68401 atggctatgg cgtgcgcga cggccacctg ccgcgaccc tgcacgccga
 68451 cgcgcctcc tcgacgtgg actggtccgc cggatcggtc cgcctgctga

10

20

30

68501 ccgagggcca gcagtggccg gagaccggac gtccgcgccg ggccgcgggtg
 68551 tccctcgttcg gcatcagcgg caccaacgcg cacgccctgc tggaacaggc
 68601 accccacccc gcggacaccg cggacgccgg cgacgacgcc gcgccaccg
 68651 aaccggccgg cgcgcccgcc gcgctgccct ggatcgtctc cggacactcc
 68701 ccgcaggcgc tcgcgcacca ggccgcgcc ctagccgcca gggtcgagac
 68751 cgaccccgcg ctccgcccc aggacatcgg gcacaccctg cacaccgcc
 68801 gcgccctgct cgaacgacgc gccgtcgtcg tcgccccga ccgcgccga
 68851 ctcttcgcgg ctacccacga gttggccgcc ggccgggtccg cgaacgccgt
 68901 cgtcgagggc ctccgggacg tcgagggtcg gacgggtgtc gtgttcccg
 68951 gtcagggttc gcagtgggtg gggatggggg cccaactcct cgtatagtcg
 69001 gcgggtgtcg cggagcggat tgcgagtggt gcggcgccac tcgccgagtt
 69051 caccgactgg tcgctgggtc atgtgctgcg ggggtgtgtg ggtgcgccgt
 69101 cgttgaggcg ggtcgatgtg gtgcagccgg cgtcgttcgc ggtgatgggt
 69151 tcgttggtcg cgttgtgggg ttcccggtgt gtgttgccgg atgcggtggt
 69201 ggggcattcg cagggtgaga tcgctgccgc ggtggtgtcg ggtgcgctgt
 69251 cgttgccgga cggggcgccg gtggtggcgc tcgaggatca ggcattggt
 69301 cgtgcgttgg cggggcgggg cgggatgatg tcgctgcgc tgcggtgga
 69351 cgtgctcga cgcgggttgg tcgagttcga gggcggggtg tcggtggccg
 69401 ccgtcaacgg ccgcgcctcc gtcgtggtcg ccggcgagcc cgaggcgtg
 69451 gacgcgtgc acgcccggt gaccgcgac gacatccggg ccgcccgat
 69501 cgcggtggac tacgcctcgc actcgacca ggtcgaggac ctgcacgagg
 69551 aactgctgga ggtgctggcg gagctggcgc cgcgcacgtc ggagggtgccg
 69601 ttcttctcga ccgtgaccgg cgaactggctg gacaccgcgc ggatggacgc
 69651 cggctactgg ttccgcgaac tgcgcggacg ggtgcggttc gcggacgcgg
 69701 tggcggacct gctggcgcg gagtaccgcg cgttcgtcga ggtcagctcg
 69751 caccgggtgc tgacgatggc ggtcttggac ctgatcgagg aggcgggggt
 69801 caccggcgtc gcgaccggca cctgcgcgg tgaccagggt ggcgcgggcc
 69851 gcttctctgt gtcggccgcc gaggtcttcg tcgcgggtgt ggacgtggac
 69901 tgggcggggg cgttcgaggg gaccgggtcg gcccggtcg acctgccac
 69951 ctacgccttc cagcgcgagc ggtactgga caccgcacc gccgccgacc
 70001 gcaccccgcc cgaaccccg atggacgcc aattctgggc cgcgctcga
 70051 caggcggacg tctccgcgt gaccgcgcg ctgggcaccg acgaggactc
 70101 cgtcgcgcc atcctgcccg gcctcacctc ctggcgccgg gcccgctccc
 70151 agcgcaccac cctcgactcc tggcgctacc gcgtcacctg gacgccctc
 70201 gccaggtgc ccgcgccac cctgaccggc acctggctgc tggtcaccac
 70251 cgacggcatc gacgacacc atgtggcagg ggcgttgag agctacggcg
 70301 ccgaggtgcg ccggtgtgtc ctggacgagg agtgcaccga ccgcgccgtc

10

20

30

70351 ctgcgggagc ggctggccgg cgcggaggac gtgaccggca tcgtctccgt
70401 cctcgccgcc gccgaggacg acgcgcacg ccaccccgcc ctcaccggg
70451 gactcgcgct caccgtctcc ctggtccagg ccctgggcga cgcgaggcg
70501 accgcgcgcg tgtgggttcc gaccgcgcgc gccttcgcca ccggcccgtc
70551 cgaccccgtc acccgggccc tgcagagcca gatcgcgggc gtcggctgga
70601 ccaccgcgct ggagcaccgc cagcgctggg gcggcaccgt ggacctgcc
70651 gacacctctg acgcccgggc cggccagcgg ctgcgcgcgc cgtgtccgg
70701 cgccctcggc gccgaggacc agctcgccgt ccgcgcgcgc ggggtactgg
70751 cccgcgcgat cgtgcgtgcc ggacaccgcg ccggacgacc ggcacggacc
70801 tgggcgcgcg gcggcaccac cctgatcacc ggcggtccgc gacacctcgc
70851 cccgcagctc gcccgctggc tggccgaacg cggcgccgag cactgtgtgc
70901 tggtcagccg gcgcggtgcc gacgccccg gagcgccga actcatcgcg
70951 gaggcagccg agtcgggcac cgaggtgacc gtcgcgcct gcgacatcac
71001 cgaccgcgac gcggtcgccg cgtgtctggc cgacctcag gccgacggcc
71051 gcacctcgcg caccgtctc caccgcgcgc ccgccatcga gctgtccgcg
71101 ctgcgcgaca ccaccgtggc ggagttcgcc gacgtcgtgc acgccaaggc
71151 caccggcgca cggatcctcg acgaactgct cgacgacgcg gaactggacg
71201 acttcgtctc gtactcctcc accgcgggca tgtggggcag cggcgtgcac
71251 gccgcctacg tcgcgcgcaa cgcctatctg tccgcgctcg ccgagcagcg
71301 ccgcgcgcgc ggactgcgca ccacctccat ccactggggc aagtggcccg
71351 acgaccgggc acgcgagctg gccgaccgc accggatccg ccgcagcggc
71401 ctggagtacc tcgacccgga gctggcgctc accgcgctcc agcacgtct
71451 ggacgacgac gagaccgtca tcggcctcat ggacatcgac tgggacacct
71501 accacgacgt gttcaccgcg ggccggcccg cgcacctctt cgaccagatc
71551 cccgaggtgc ggcgcgcct cgcaccaggca tccgtcccg accccgcggg
71601 cccggccgcc gaaggcctgg ccgcgcgcct gcacggcctc gccgcgcgcg
71651 aacaggaccg gctgctgctc accctggtcc gcaccgaggc cgcgcgcgtc
71701 ctgggccacg cctcggccga gtcttcccc gagcgccgc ccttcctgga
71751 cctcggtctc gactcggtca ccgcctgga cctgcgcaac cggctcgtgg
71801 ccggcaccgg actgcggtg ccctcgacga tggctcttga ccaccccaac
71851 tgcgcggcgc tcgcgcggt cctgaagacg acggcgctcg gcgtcccg
71901 cgccgcaccg cagcagcag ccgctaccgg caccgcggcc gacgacgaac
71951 cgatcgccgt gatcgcatg agctgcgct accccggcg cgcgcgcc
72001 cccgaggaac tgctgcggct cgccctcgac ggcgccgacg tcctctcgga
72051 gttcccgcg gaccgcggct gggacgcccg gggcctgtac gaccggacc
72101 ccgaccgcc cgccacacc tactcgtcc agggcggtt cctccacgag
72151 gccgcggct tcgatcccg cttcttcggg atctccccgc gcgaggcgt

10

20

30

72201 cgccatggac ccgcagcagc ggctcctgct ggagacctcc tgggaggcgt
 72251 tcgaacgcgc cggatctgac ccgcgtcac tgcgcggcag cgcgcgcgc
 72301 accttcttcg gcgcagcta ccaggactac tctccaccg tgcagaacgg
 72351 caggggggag tccgagcgc acatggtgac cggcaccgcg gccagtgtcc
 72401 tgctcggcgc ggtctcctac ctgctcggcc tggagggccc cgcggtcacc
 72451 gtggacaccg cctgctcctc ctcaactggtc gccctgcacc tggcctgcc
 72501 gtccctgcgc gaaggcgaga gctccctcgc gctggccggc ggtgcggcgc
 72551 tgatggccac ccgcacgcg ttcgtcggct tcagccgca cgtgacctg
 72601 gccaaaggac gccgctgcaa gccgttctcc gacaccgcg acggcatgac
 72651 gctcgcgcg ggcgtcggc tctcctgct ggagcgctg tccacgccc
 72701 gcgccaacgg gcaccgggta ctggcctga tccgcggtc cgcgctaac
 72751 caggacggcg cctccaacgg cctgaccgcg cccaacggcc cgtcccagca
 72801 gcgcgtcatc cgccaggcgc tcgccaacgc cggcctgacc ggcgcgcgac
 72851 tcgacgcggt cgaggcgac gccaccgca ccaagctggg cgacccatc
 72901 gagggccagg cctgctcgc cactacggc caggaccgcg acgccaacg
 72951 gccgctgctg ctgggctcgg tgaagtccaa catcgccac acccaggccg
 73001 cggcgcgcgt cgccgcgtc atcaagatg tgctggccat ggacgcggc
 73051 gaactgcccg gcacctgca cctcgacgcg cctccagcc acgtcgactg
 73101 gaaccgcggc gccgtggaac tgctgcgcg gcgcacccc tggcccgaga
 73151 gcgggcgcgc ccgcgcggc ggtgtctct cgttcggcat cagcggcacc
 73201 aacgccacc tgatcctcga acaggcccg gccaccgagc cgcagccga
 73251 ccccgaccgc ctccgggaca ccgccaccga caccgtcgtc cctggccgc
 73301 tcgcgcgcaa gtcccggcc gccctgcgcg ccagggcgc cggctctct
 73351 gccaccgtcg agcacgacc cgacctccg ccgcgcccg tgggccacgc
 73401 cctggccacc acccgcgccg cctcgaaca ccgcgcgtc gtcgtcggc
 73451 agcgcgcgga ggacttctg cgcggcctg ccgcctgtc caccggcgc
 73501 tcgacggccg gcctggtcag cggcatcgcc ggccccgacc ccgagggagc
 73551 ggtctctgct tccccggcc agggatccca gtggtggga atgggcccgc
 73601 aactcctcgc cactccgag gtgttcgca ccgcgatcga tgactgcgc
 73651 acggccctcg cccgtacgt cgaactggtc ctgcacgac tctggccgc
 73701 cgagggcgac ccgcctcgc tggagcgggt ggacgtggtc cagccgcgc
 73751 tgcttcgcat gatggtcgg ctgtccgcgc tctggcctc ccacggcgtc
 73801 gtcccggcgc cgtggtcgc ccaactcgag ggcgagatcg ccgcgcctg
 73851 cgtcgcgga gccctcagcc tggccgacgc cgcgcgcgtg gtggcgtgc
 73901 gcagccaggc actgcgcaa ctgtccggac gcggcgcat gatgtcgtc
 73951 tccgccccg tagagcgggt caccgacctc ctgccccgt ggcaggaggc
 74001 gctgtcgtc gccgcggtca acggccctc gtcgtggtc gtctcggcgc

10

20

30

74051 acaccgacgc gctcgacgcc ctgcacaccg cctgccagga acagggcggtg
 74101 cgggcccgcga aggtgtccgt ggactacgcc tcgcacgggc ggcacgtcga
 74151 ggcggtccgc gacgaactcg cccgcgtcct cgcgcgggtc gaccgcgcgc
 74201 cccccgaggt gccgtttctac tcgacggtea ccggcgaccg cgtggacgac
 74251 gccgccttcg acggcgcccta ctggtacacc aacctccgcc agaccgtccg
 74301 catggaggag gccacccgcg ccctcctcgc cgcgggacac cgcgtcttca
 74351 tcgaggtcag cccgcacccg gtgctcgcgc ccccgatcca ggagacgcag
 74401 gaggccgtag cggaggccac cggcggtcc cgggtggtcc tcggctcgtc
 74451 ccgcgcgcgc gaggcgggcc cgcggcgctt cctgacgtcg ctgcgcgagg
 74501 cccacaccca cggcgccccg gtcgactgga ccaccacctt cgcgcgtcc
 74551 gcctaccagc cgggtggacct gccgacctac cccttccaac gacaggactt
 74601 ctggcccagc gcccgccccg ccaccccggc cgcggcgccc gacgcgtccg
 74651 acgcgcgctt ctggcaactg gtgcagaacc aggacctcgc cgcgtcgcgc
 74701 gacgcgctcg gcgtccccgc cgacgacgag caccacgcgc tcggcacctg
 74751 gctgccggcc ctgtccgcct ggcgcgccaa ggcccaggcc cgcacccgga
 74801 tcgacgaact ccgctaccac gtccagtgga cccgggtcgc cgagcccgcg
 74851 gcggccccca ccacccggcc gctgctggtc gccgtccgc cggaccacgc
 74901 cgacgcccc tggttcgcgc cggcgctcga cgcctgggc accgacaccg
 74951 tcgccttcga ggccaaggcc accgacccgc cgggatgggc cgcacagatc
 75001 gcccaactcg tcgaggacgc cgaggagtcc accggcggtg tgcgtcgtc
 75051 ggccgcgcgc gaggatctcc acccgactt cggctcggtc ccgtcggggc
 75101 tggggcagac cctcgtcctc gtccaggccc tcggcgacgc cgcctgacc
 75151 gcgcacctgt ggtgcctgac ccgcggcgcc gtgccaccg gccgcgacga
 75201 cgccttcgac agcccgaccc agggcgccct gtggggcctc ggccgggtcg
 75251 tggccctgga acaccccgac cgtggggcg gcctgatcga cctgcccgcc
 75301 accctcgacg cccgcgcgcg ggcccgcctc accggcctgc tcgccgaccc
 75351 cgcgggtgag gaccaactcg ccgtccgcgc caccggcgta ctgcccgcc
 75401 gcatggtgca cgcgcgcgcg tcgcgcgcc gcaccggggc ccgtcggcgc
 75451 ggccgcggca cctgcctgat caccggcggc accggcgga tcggcggcgc
 75501 ggtcgcgcgc tggttggcgc agcagggcgc cgcacacctg gtcctgacca
 75551 gccggcgcg cccggacgca cccggcgcc cgcactccg ggccgaactg
 75601 gaggccctgg gcgcgcgcgt caccctcgc gcctgcgacg tcgccgaccg
 75651 cgacgcctcg gccgcctcgc tggccgacct cccgcgcgac cagccgctca
 75701 cctcgtctt cactccgcc ggctgggcgc acggggacgc ccgggcagcc
 75751 gacctgaccc tagatcagct cgacgcgctg ctgcgcgcca aactgaccgc
 75801 cgcaccacac ctgcacgagc tgaccgcccc cctcgacctc gacgcgttcg
 75851 tgcctcttc ctccggcgcc gcggtctggg gcagcgcggc ccagcccggc

10

20

30

75901 taogccgccc ccaacgccta cctcgacgcc ctgcccgcc accgcaggtc
 75951 cctcgacctg cccggcgctt ccgtcgcttg ggccacctgg ggcgaggtcg
 76001 gcatggccac cgtccccgag gtccacgagc gactgcaccg ccaaggggtc
 76051 cgcgccatgg aaccggacca cgcgatcggc gcgtccagc agatgctgga
 76101 ggacgacgac accaccctcg ccgtgaccct catggactgg gaggcggttcg
 76151 caccgagctt caccgcgacc cgaccacgcg ccctgttcag caccgtgccc
 76201 gaagccgttc gcgcggtgac cggcgacccg ggcaccacgg ccggcgacga
 76251 cgtggacttc gcgaccccg cgtcccgccg ccacctggag gagctgtccg
 76301 ccgcccagcg cggccggggc ctggtcgagg cggtcggcgc cgaggcgctc
 76351 gcgacctctg gccacgacac ccccgacgcc atcccccgcc gccgtgcctt
 76401 ccgcgacgtc ggcttcgact cggtcaccgc cgtcgaactg cgcaaccggc
 76451 tgcgcaccgc cctcgccctg ccgtgcggc ccgcgctcgt cttcgaccac
 76501 cccaccccca cggccctcgc cggccacctc ggcgcgctgc tcttcggcac
 76551 cgcgcccgag gacgcgggca ccggccggcc cgacgacccc gacgcccgca
 76601 tccgcgaggc gctcgccacc gtcccatcgc gacggctgcg caaggcgggc
 76651 ctcttcgaca tgggtgctgaa actcgccgac ggagacgaga ccgacgcccc
 76701 cgcgcccgag gccgacgccc cctcggaatc cctcgacgac atggacgccc
 76751 aagccctgct cgggctggcc accgagaact cggcgaactg aaagagagct
 76801 ggagcccacc atgagcagga accccgacaa gtacgtcgag gcactccggt
 76851 cgtccctgaa ggagatcgag cggctgcgcc ggcagaaaga acagctggtg
 76901 gcggcgggcg tcgagccctt cgcggctcgc ggcacggct gccgcttccc
 76951 cggcgggcgc acctccccc aggaacctgt ggagctggtc gccgaggggc
 77001 gcgacgtcat cgggcccgtc ccgacgacc ggcgctggga cctggagaag
 77051 ttggccggcg cggcgagggg cggcagcctc gcgaggtcg gcggcttcgt
 77101 cgaggacgcc gccggttcg accccggtt ctccggcctc tccccgcgcg
 77151 aagcggtcgc catggaccgg cagcagcgca tcctgctgga gatcacctgg
 77201 gaagccctgg aacgcgcggc catcgaccgg tccacctgc gcggtacccc
 77251 caccggcgtc ttcgtcgga ccaccggcca ggactacggc gaggctcatca
 77301 aggcgtccgc cgaggacgtc gaggcttact cgaccaccgg ccacgcccgc
 77351 agcgtcatct ccggccggct ctctacacc ctggcgccg agggcccggc
 77401 cgtcacctgc gacaccggct gctcctcgt cctggtcgcc ctgcaactgg
 77451 ccgtccaggc gctgcggggc ggcgagtgt ccatggccct ggccggcggc
 77501 gcgtccatca tggccacccc gggcccgttc gtcgcttca ccgcgcagag
 77551 cggcctggcc gccgacggcc gctgcaagcc cttctccgac cgggcccagc
 77601 gcaccggctg gggcgagggc gccggcatgc tggctctgat gcggtcttc
 77651 gacgcccagc gcgagggcgc cccggtcctc gccgtgctgc gcggctccgc
 77701 catcaaccag gacggcgctt ccaacggcct gaccgctccc aacggccct

10

20

30

77751 cccagcagcg cgtcatccgc gccgcgctgg acagcgccca cctcaccgcc
 77801 gccgacatcg acgcgctcga ggcccacggc accggcacca cctcggcgga
 77851 cccgatcgag gcccaggcgc tcttgccgac ctacggacag gaccggccgc
 77901 ggccctctgt gctcggctcg gtgaagtcca acatcggccca caccagggcc
 77951 gcctccggtg ccgcgcggct gatcaaatg atcatggcgt tgcagcgcg
 78001 cgtgctgcgg cgcagcctgc acgccaccga acccaccacg gacgtcgact
 78051 ggaccgcccgg ctccgtcgac ctctctgacg agacggtcgc ctggcccag
 78101 accggacggc ccgcgcggc cggcgtctcc tcttcggca tcagcggcac
 78151 caacgcccac gtcatcctcg aacaggcccc caccgcccc gaagagccca
 78201 ccaccgaacc caccgtccgc ccgcgcgctg tcccgtaggc gctctccgc
 78251 cgcaccgccc ccgcctcga cgcacagcgc gcccgctca ccggccacct
 78301 cgcgcacacc ccgcagcggc acccctcga cgtcggctac ggcctcgcc
 78351 acggacgcgc caccctcga caccgcgcgc tctgtctccc cgacggcacc
 78401 gaactcgccc acggaaccgc cggcgaaggc cctgcggccg tctcttctc
 78451 cggccagggc tcccagcgc cggcatggg acgcgaactc caccgccgt
 78501 tcccggtgtt cgcgcgggc ttccgacgaga tcacagcgt cctcgacacc
 78551 caccctgacc gccgcgtcg cgaggctcgc tggggaccg acccgacct
 78601 gctgaacgac accggctggg cccaaccgc cctgttcgcc gtcgaggtcg
 78651 cctctacgg cctggctcg tccctcgcg tcaccccgga ctctcgtggc
 78701 ggccactcca tcggcgagct cgcgcgcgc cactcgccg gggctctctc
 78751 cctcgaagac gcctgcacc tcgtcgccgc ccgcgcccgc ctcatgcagg
 78801 cctgcgcgc cggcggcgc atgctcgca tccgcgccac cgaggacgag
 78851 gtcaccccc acctcaccga cgacgtctcg atcgccgcg tcaacgggcc
 78901 caccctccgtc gtcgtcgcc gcaccgagga agcgtcgcc gcgatcgggg
 78951 cgcgcttcac cgcacaggac cgaagacca ccgggtcgg ggtcagccac
 79001 gccttcact cgcgcctcat ggaccgatg ctggcggaat tccgcgcgt
 79051 cgcgcgggc ctgacctacc acgagccgc catcccggtc ctctccaacc
 79101 tcaccggcac cgtcgccgc gtcgcgcacc tgtgtctcgc cgactactgg
 79151 gtccgccacg tccgcgagc ggtccgcttc gccgacggc tcaccgccct
 79201 caccgaccgc ggcgtgacca cgtcgtcga actcggccc gacggcgtgc
 79251 tgtccgccat ggcccaggaa tccctgccg acggcgccgc cgcgtgccc
 79301 ctgctcgca aggaccgccc cgaggagctc tccgcgcta ccggcctggc
 79351 ccgcgccac gtcgcggcg tcacggctcg ctgggccggc ctcttcgacg
 79401 gcaccggcgc gcgcgcgc gacctgcca cctaccctt ccagcaccag
 79451 cggttctggc cgaccgcgc ccgcgcgc caggacgta ccgccggg
 79501 actgggcgc gccgaccacc cgtgctcgg cgcaccgtc gaactcgcc
 79551 acggggcgg ctacttgttc accagccgc tctcgtcgc gaccacccc

10

20

30

79601 tggctcgccg accacggggt ccaggggccgg gccctgctgc ccggcaccgc
 79651 ctctgctgaa ctggccgtcc gcgcgggcga cgaggccggc tgcgaccgcg
 79701 tcgaggaaact gaccctggcc gccccctgg tgcctgccga gcgcggcggc
 79751 gtccaaactcc aggtccgcgt cggcgccccc gacgcgcgcg gccgcgcac
 79801 cctcggcacg ttctcccgcg tcgaggacgg ctctgacctg ccttggctgc
 79851 aacacgccac cggcgctctg accgcggcg ccggcgcccc cgaccccacc
 79901 ttgacgcca ccgtctggcc cccagcggc gccgaacccg tcgacctcac
 79951 cggcgcgta cggcgctgg ccgcactcgg ctccagta cggcccgcc
 80001 tccagggcct gcgcgcgcg tggcgcccg acaccgaggt ctacgcgaa
 80051 gtggccctgc ccgacggcg ggacaccgac ccgcgcgcct tcggactgca
 80101 cccggccctg ctggacggcg cacaacacgc cgcgcctac gccgacctg
 80151 gcgccatcag ccgcggcggc ctgcgcttcg cctgggaagg cgtctcgctc
 80201 gccgcgcgcg gccccaccac cgtccgcgc ccgatcgccc cggccggcga
 80251 ggacaccgtc accatcgccg tctacgacgc cgcggcggc accgtgctgt
 80301 ccgtcgactc cctggctcc cgcgaggtcc ccgcgacgc acccgcgcc
 80351 gccggcaccg tccaccgga ctccctcttc cacgtcgagt ggaccccgct
 80401 ccaggggcgc ccgggcccc caccggccac cgtcgccgtc ctgggccccg
 80451 acccgga cctcgccgac accctccgc ccaccggcat ccggaaccac
 80501 gcccccgcg acctggcgc cctcgccgac gccgaagggc ccgtcccca
 80551 cctggctgct accacctca ccaccacccc ggcgcccccc gtccccgacg
 80601 ccgcgcacgc caccaccgcc gccgtcctcg cctcgccca acagtggctc
 80651 gccgacgacc gcttcgcga cgcgcgcctg gtctctgta ccgcggcgcc
 80701 caccgacggc accgaccccg ccgcgcggc cgcggcggc ctgatccga
 80751 ccgcgcgac cgaqaacccc ggccgtttcg cctcctcga cctcgcccc
 80801 gacaccggcc ggcccgacc cgagacctg gccaccgcc tggccgccag
 80851 ccacgacgag cccgacctcg ccgtccgcgg caccgacgtg caccgcgcc
 80901 gcctggcccg tgteccgctc gccaccgaac ccaccaactg gaaccggac
 80951 ggcaccgtcc tgatcaccgg cggcaccggc ggccctggcg cggctcctgc
 81001 ccgccacctg gtccgccccc acggcgctcg ccacctgctg ctgcgcagcc
 81051 gccgcggccc ggccgccgac ggccgcgacg acctgacggc cgaactcacc
 81101 gggctcgggc ccaccgtcca catcgccgc tgcgacgtcg ccgacccggc
 81151 cgcctcgtcc gacctgctcg gcaaccgtcc ggccgggac ccgctcaccg
 81201 tcgtcgta caccgccggc gtcgtcgac aggcgtcct cggctccctc
 81251 accccgcagc gcctggacac cgtcctgagg cccaaggccg acgcgcctg
 81301 gcacctgac gagggcacc gccacctga cctggacgcc ttctcctct
 81351 tctcgtccgt cgcgcgccc ctcggcagcc ccggacaggc caactacgc
 81401 gccgggaaag cctcctgga gccctcgcg gccggcgcg ccgccaccgg

10

20

30

81451 cctgccccgc acctccctcg cctggggccc gtggaccag agcgtcggca
 81501 tgacaagcag cctgtccgac ctgacgtcg agcgcacgc ccgctccggc
 81551 atgccccgc tgacctgga acagggcacc gccctcttcg acgcgccct
 81601 ggccgcccgg ccgcccgcgc tcgcccgggt ccgctctgac ctgcccgtcc
 81651 tgcgcaccca gggcgacatc gccccgctgc tgcggggcct gatccgcacc
 81701 cccgtgcggc gcaccgcccgc ccaggtctcg cagaccgccc acggcctcgc
 81751 ccagcggctc gccggcctcg acgcgcgcgc ccgcccggaa gccctcctgy
 81801 aactcgtccg caccagatc gccaggtcc tcggccacgc ggacgccacc
 81851 gaggtggaga ccggccgcca gtccagac ctcggttcg actccctcac
 81901 cgccgtcgaa ctccgaacg cctgaacac cgcacccggc ctgcggtcgc
 81951 ccgccaccat ggtgttcgac taccgcacac cacacgccct cgcgcaccac
 82001 ctgcgcgacg aactcctggg caccgaggcc gactcgacca ccgccgtccc
 82051 cgtgccgacc cgtaccgccc gcaccgacga ccgacgcgc atcgtcggca
 82101 tggcctgccg ctaccccggc ggcatcgctt caccgagga cctctggcgc
 82151 ctggtcagcc agggcgccga cgcactggc ccgttccca ccaaccggg
 82201 ctgggacctg gacaacctt acgaccccga cccgaccgc ccgggcccga
 82251 cccacgtccg cgcggcgccc ttctgcacg acgcccgtc ctctgacgcc
 82301 gacttcttcg ggatgagccc gcgcgaggcg atggccaccg actccagca
 82351 gcgcctgctg ctcgaaactt cctgggaagc cgtcgaaacg gccggcatcg
 82401 accccgcctc actgcgcgac tcgggcaccg gcgtcttcgc cggcgtcatg
 82451 tacaacgact acggcaccac cctgaccggc gacgagtacg aggcgttcg
 82501 cggcaacggc agcgcgccga gcgtcgctc cggccgcgtc tctacaccc
 82551 tcggcctgga agggccggcc gtcacggtgg acaccgcctg ctcttcctct
 82601 ctggtcgcgc tgcactgggc ggcgagggcg ttgcggggcg gggagtgtc
 82651 gttggcgttg gccgggtggt tgacggtgat gtcgacgccg agcacgttcg
 82701 tggagtcttc gcggcagcgg ggtctggcgc ctgatggtcg ttcgaaggcg
 82751 ttgcgcgagg ccgcgacggc cgtggcctgg tccgaggcg tcggcatgct
 82801 ggtcctggag cggcagtcgg acgcgggtcg caacggtcac gagatcctgg
 82851 ccgtggtgcg cggtcggcg gtcacaccag acggtgcgtc caacggtctg
 82901 accgcgccc acggcccgtc ccagcagcgg gtgatccgtc aggcgttggc
 82951 cagtggcggc ctgtccacgg ccgacgtgga cgcggttgag gcgcacggca
 83001 cgggtacgac gctcggtgac ccgacgagg ccagggcgt cctggccacc
 83051 tacggtcgcg accgcgaccc cgagaaccgg ctgctgctcg gtcgatcaa
 83101 gtccaacatc ggtcacaccc aggcagcggc cgggtgcgcc ggtgtcatca
 83151 agatgggtcat ggcatgagg cacggcgtgc tgcgcagac cctgcatgtc
 83201 gacgcgccgt cctcgacgt cgattggagc gtcgggcggc tcgaactgct
 83251 caccgagcag accgcctggc cggagaccgg ccgggcccgt cgcgcgggtg

10

20

30

83301 tctcctcctt cggeatcagc ggcaccaacg cccacgtcgt catcgagcag
 83351 tccccgaccg ccgtccccc cagcccgcg tccgocgacc ggtccgtcga
 83401 ggaaccgcg gcggtccctt gggccctgtc cggcaagacc cccgacgccc
 83451 tccgcgacca ggcgcgcgc ctctctgcgc acgtcagagg ccccccgcga
 83501 ctgcccccg tcgacatcag ctactccctg atcgccacc gcaccgcctt
 83551 cgaccaccgc gccgtcgtcc tcggcaccca ccgcgcgag gccctgcgcg
 83601 cctcacgcgc cctgcgcgc ggcgagacgc acccggcgc cctcacgcgc
 83651 accgtccgca ccggccgcac cgccttctc ttctccggcc agggctccca
 83701 acggctcggc atgggcgcg tcctctacga cgggttccc gccttcgcgc
 83751 aagccctcga caccgtctc accgccctcg acgcggaact cggccacccc
 83801 ctccgcgaca tcctctgggg cggagacgct caactcgtcg accggaccgg
 83851 ctacacccaa cccgcctctg tcgccatcga ggtggcactc ttccgcctcc
 83901 tgggaagcctg gggcatcaca ccggaactcg tggccggcca ctccatcggc
 83951 gagatcgcg ccgcacacgt cgcggcgctg ctctccctcg gcgacgcctg
 84001 ccgcctcgtc gtggccgcg ccgtgctgat gcagtcgtg cccgaaggcg
 84051 gcgcgatgat cgcggtccag gccaccgagg acgaggtcct gccctctctc
 84101 accgacgacg tctcgatcgc cgcgctcaac agcccgacct ccgtcgtcgt
 84151 ctccggttac gagaacgcca cctcgcggt cgcgcgcac ttccgcgacc
 84201 agggccgcg caccacgcg ctgcgcgta gccacgcctt ccactcgcg
 84251 ctgatggcgc cgtatgctga cgaattccgc gccgtcgtg agagcctcac
 84301 ctccaccgc cccacgacc ccgtcgtctc caactgacc ggcgaactgg
 84351 ccccgccga ggcgctctgc tcggccgact actgggtccg gcacgtccgc
 84401 gaggcggtcc gcttcgcga cggcatccgc accctcgcg accgcggcgt
 84451 caccaccttc gtcgaaactg gcccgcagag cgtgctgtcc gccatggccc
 84501 aggagtcgc cccgaaggc gccggacca tcccgctcct gcgcgcgac
 84551 cggcccgagg aacaggccgt cctggccgc ctctgccacc tccaggtgct
 84601 cggcgtcag gccgaactgt ccgccacctt ccgcggcctc gaccccgctc
 84651 gcgtcgacct gccgacctac gccttcacgc accgctggtt ctggcccgcc
 84701 gcccgacccg ccgcgccga cgaactcgc gccgcggcc tgggcgcgcg
 84751 cgaacacccc ctctcggcg ccgcggtgca actcccgac gacgacggcg
 84801 cactcttcac cggccgcctc tccctgcgca cccaccgtg gctggccgac
 84851 cacaccgtcc tgggcaccgt cctgctcccg gccaccgcac tgggtggaact
 84901 cgcggtccgc gcgggcgag agaccggcag cggccacctc gaagaactca
 84951 cctcgcgcgc gccctgacc ctccccgagg acggcgccac cctcctccag
 85001 gtccgcgtcg gatccgcga cgaacccgc cgcgcaccg tcaccgtcca
 85051 cgcgcgcgc gacgacacc ccgaccgcac ctggacgctg caccgcccg
 85101 gtgtgctcgc caccacgcca ccggccgcg cggcgttcga caccacggtc

10

20

30

85151 tggccgccc cgcacgcga acccctcacc accgacgact gctacgcaca
 85201 cttcaccacc caccgcttcg cctacggccc cgccctccag ggccctgcggg
 85251 ccgcctggcg cgcggcgac gtgctgtacg ccgaggtcgc cctgccggag
 85301 tcggccaccg acgaagcggc cgcccttcggc ctgcacccgg cgctcctgga
 85351 cgccggcctg caccgcccgc tccctgcgga cgaccgcgac accggactcc
 85401 cgttctcctg ggaaggcgtc actctgcacg cctccggcgc caccgcgcta
 85451 cgcgtccggc tcgcccga cggcccacac ggccctgtccg tcaccgcgc
 85501 cgaccggcc ggaaccccg tcgccaccgt caccgcgctg ctgcccgcc
 85551 ccctggacgc cgagcagttg accatccaca gcgcctgac ccgcgacgcg
 85601 ctcttcacc tggactggac cccggtcccg ctcccgaca ccgccaactc
 85651 cgcgcgcgcg gccctcctcg gcccgacac cgccgtgctc gccgacgcc
 85701 tcggcgacc gcccgctgca cgccacgcaa ccctcgacga cctcctggcc
 85751 ggggacacca ccccgcccgc caccgtctc gtcccccctg gcgcccact
 85801 cgacggcgac accgcgcagc acgcgcacgc cctcaccgc agcgcgctga
 85851 ccctcgtcca gcagtggctc gccaccgacc gcctcgccga ctcccgccg
 85901 gtcttcgtca cccacggagc cgtcgccacc gacgacgcgc cccccaccga
 85951 cctggcgccc gcccggtct ggggcctgat ccgctccggc cagaccgaga
 86001 acccgggcac cttcaccctc ctgcacctg acaccgagcc cgactcgacc
 86051 accgcgctca gccgcgcct gaacctcgac gaaccacagc tccctcctcg
 86101 cgccggcgc gcccgcgcc cccgcctcac ccgcacccc gcccccacca
 86151 ccaccacca caccgcgtg tccgcggacg gaacgggtgt ggtgacgggt
 86201 ggtacgggtg gtctgggtgg gtgggtggcc cggcatctgg tcgggtcgtg
 86251 tggggtgcgg ctttgtgtg tgaccagtgc ttctggtgtg ggtgctgcgg
 86301 gtgcggccgg gttggtcgg gagttggagt cgttgggcgc gcgggttgtg
 86351 gttgcggcgt gtgatgtgg tgatggctcg gctgttcgg agttggtgc
 86401 cgggtgtgct gagtcgtatc cgttgtctgc ggtgggtcat gcggtcgtg
 86451 tgttgatga cgggtgtgtg ggttcgtga cgccggagcg gttggctgcg
 86501 gtgttgctc cgaagggtga tgggtcgtgg aacctgcat aggcgacgcg
 86551 tggctcggat ctggacgcgt ttgttgtct ctcgtctgtt gcgggtgtgt
 86601 tcgggggtgc gggtcaggcc aactatgcgg cgggtaatgc gtttttgga
 86651 gcgttgatgg tcatcgggt ggctgggtgg ttgcctggtg tgcgttggc
 86701 gtgggtgct tggatcagg gtgtgggat gacggcggg ctgacggagc
 86751 gggatgtccg tcgtgctgct gagtcgggta tgcggtgtt gacggtgat
 86801 cagggtgtgg cgttgttca tgcggcgtt gcgacgggga gtgccgctt
 86851 ggtgccggtc cgtctggacc tggccgact gcgcacccgg ggcgacatc
 86901 caccgctcct ccgcggcctc gtccgcgac cgtgcgcgc caccgcggcc
 86951 accggcctcg ccaccggcgc ggacaccggc ctgctccaac ggctcggccg

10

20

30

87001 actcgaccac gcccaacgcc acgaggcact gctcgacatg gtccgcagca
 87051 gcgcgcgcgt cgtcctcggc cagcccgacg gcaacgccat cgacgccgaa
 87101 cgcgccttcc gcgacctcgg ctctgactcg ctaccgcgg tcgaactccg
 87151 caaccgtctg cgcaccgcca ccggcctgca cctgtcggcc accatggtct
 87201 tcgaccaccc caccctgtcc gccctcggcg agcacctgcg ggacgagttg
 87251 ttccggcggg tcgagagcga ggtgcgggtg ccggtccagg cactgcgcgc
 87301 gaccgcgcac gatcccatcg tggtggtggg catggcctgc cgtttccccc
 87351 gtggtgtgac ctgcgccgag gacctgtggc gccgtgtcga cgacggcacc
 87401 gacgccatca ccaccttccc gaccaaccgc ggctgggacc tggacaacct
 87451 ctacgacccg gaccccgagc acttcggcac glcgtacacc cgctccggtg
 87501 gcttcctgca cgaggcgggg gagttcgacc cggcgttctt cggaatgagc
 87551 ccgcgtgagg cgctggcaac cgactcccaa cagcgtctcc tgctggaatc
 87601 ctctcgggag gcgatcagc gggccggcat cgaccgcctg accctgcgcg
 87651 gcagcgcac ccggcgtcttc gccggcgtga tgtacagcga ctacgggagc
 87701 atctcggcg gcaaggagtt cgaggccttc caaggccagg gaagtgcggg
 87751 cagcgtggcc tcgggcccgc tctcctacgc cctcggcttc gagggcccgc
 87801 ccgtcacggg ggacacggct tgccttctct ctctggtcgc cctgcactgg
 87851 gcggcgcagg cgttgccggc gggggagtg tcgttgccgt tggccggtg
 87901 tgtgacggtg atgtcgacgc cgagcacggt cgtggagttc tcgcggcagc
 87951 ggggtctggc gccgtatggt cgttccaagg cgttcgccga ggccgcggac
 88001 ggcgtcggct ggtccgaggg cgtcggcatc ctctcctcgg agcggcagtc
 88051 ggacgcggtg cgcaacggcc acgagatcct cgccgtgate cgccgctcgg
 88101 cgggtcaacca ggacggtgcg tccaacggcc tgaccgcgc caacggcccg
 88151 tcccagcagc gcgtcatccg tcaggcgttg gccagtggcg gccgtccac
 88201 ggccgacgtg gacgcgctcg aggcgcacgg caccggtagc acgctcgtg
 88251 acccgatcga ggcccaggcg ctctcggcca cctacggccg tgaccgcgac
 88301 cccgagaacc cctgtgggtt gggctccctg aagtccaaca tcgggcacac
 88351 ccaggcagcg gccggtgtcg ccggtgtcat caagatggtc atggcgatgc
 88401 ggcacggcgt gctgccgcag accctgcatg tcgacgcgc gtctctgcac
 88451 gtcgattgga gcgtcggcgc cgtcgaactg ctaccgagc agaccgcctg
 88501 gccggagacc ggccgggtcc gtccgcggcg tgtctctcc ttccgcatca
 88551 gcggcaccac cgcccacgtc atcgtggaac agccggcgct cgtcgaaggc
 88601 ccggcccgcg agccgagcgg acgcgaaccc ggctcgttc cgtgcggct
 88651 gtccggaaag tccccgagg ccctgcgcga ccaggccgca cgctcgtgg
 88701 ccgggttgcc ggagcggccc gcgtgcgcgc cgctcgacct cggctactcg
 88751 ctggcgacga cccgttcggc gttcgaccac cgggcgggtg tgctcgccac
 88801 cgaccgcgc gatgcggtcc gcgcgctgac ggcgctcgc gccgcgacg

10

20

30

88851 cggatctctc cggcgtctgc ggcgacaccc gcacgggtcg tcacgcggtg
 88901 ctgtttctcg gtcagggctc gcaacgcctg ggcattggggc gtgagttgta
 88951 cgagcgtttc cgggtcttcg ccgaggtctc cgatgtcgcg atcgaccacc
 89001 tggacgcgcg ctgcccgcg caggccagtc tgcgtgaggt gatgtggggc
 89051 gacgatgtcg agctgctgga cgagacgggt tggacgcagc cggctctggt
 89101 cggcgtcgag gtggccctgt tccggctggt ggagagttgg ggtgtccgtc
 89151 cggacttcgt ggcgggtcat tccatcggtg agatcgcggc ggcgcattgc
 89201 gtgggggtgt tctcgtcgtg ggaagcctgc cgtctggtgg ccgcccgtgc
 89251 gacgctgatg caggcgtcgc cgaccggcgg cgcgatgacg gcgatccagg
 89301 ccgcccagga cgaagtcacc cagcacctga ctgacgacgt ctcgatcgcc
 89351 gccgtcaacg gccgcacctc cgtggtcgtc tccggggccg agagcgtgc
 89401 ccgcacggtg gccgaccggc tcgcggagaa cggccgcaag acgaccggc
 89451 tgggggtctc gcatgcgttc cactcgcctg tgatggatcc gatgctggcg
 89501 gagttccgtg cggtaggcga gggcctgtcc tacgccacc cgaacctccc
 89551 cgtcgtctcg aacctgacgg gccggctggc cagggccgat gacctctgct
 89601 cggccgagta ctgggcgcgc cagtcgcgg aggggltccg ctctgcgcac
 89651 ggcgtcagca ccctggagaa cgagggcgtc accacgttcc tggaaactggg
 89701 accggacggc gtgtctgtcc ccatggccca gcagtcgctc accgacgacg
 89751 ccgccaccgt cccggccctc cgcaaggacc gcgacgagga gacgtccgcg
 89801 ctacacgccc tcgccacct ccacacggca ggtctccgcg tcgactgggc
 89851 ggcgttcttc gccggcagcg gcgccaccgg cgtggacctg ccgacctacg
 89901 ccttcagca cggcacctac tggcccaccg gcaacctgac caccgcgcac
 89951 gccgcggccg tcggcctcac cgcgcgcgag caccgcgtgc tgaacggttc
 90001 ggtcgaactc gccgaaggcg aagggtgtgt gttcaccgga cggctgtcac
 90051 tgcagtcaca tccgtggctg gccgaccacg ccgtcatggg acaggtcctg
 90101 ctgcccggca ccgactgct ggaactggcg ttccgggccc gcgacgaggc
 90151 cggttgcgac cgcgtcgagg aactgacgct cgcgcaccg ctgctcctgc
 90201 ccgagcgcgg tcgggtacag acccaggtcc gggtcggcgt cgcgcacgac
 90251 accggccgcc gtaccgtcac cgtccactcc cggcccagag acgcgaccga
 90301 cgtgtcgtgg acccagcacg cgaccggcac cctgaccatg ggtccgccc
 90351 cggccgacac cggtttcgac gccactgcct ggccgccgc cgcgcggaa
 90401 cccctcgcca ccgacgactg ctacgcgcgc ttacacgacg tcggcttgcg
 90451 ctacggggcg gtcttcagg gcctgcgggc cgcctggcgc gccggtgacg
 90501 tgctgtacgc cgaggtggcc ctggcggagt ccaccggcg cagggcgacc
 90551 gccttcggtc tgcaaccgc actgctcgac gccgcctgc acgcctccct
 90601 cgtcggccac gagggcgagg agagcaaccg cggactgccg ttctcctggg
 90651 agggcgcgac cctctacgcg accggcgcca ccgcgtgcg cgtccgctg

10

20

30

90701 accccgacgg gcaccgacgg ccgttcggtg gccatcgccg tggccgacac
 90751 cgccgggtcgt ccggtcgccg ccacgcacaa cctcgtctcg cgccgggtct
 90801 ccggcgacca gttgaccggc gccgcgggac tggcccgca cgccctgttc
 90851 accctggact ggaaccccggt accggagaac ctctaccgg agaaccctgt
 90901 accggagaac accggcgggg gccacgcca ggaccaggac ggccggcccg
 90951 ccgcggccac cgtcgcgctg gtcggcgcg accggcccg gatcgccgc
 91001 gacctgaccg ccgcggcat ccacaccacc ctccacccg acctaccac
 91051 cctcgccacg accgacggc acgttcgaa gacggctctc atccccctca
 91101 ccggaaccgg aaccggaacc ggcaccggga ctgagtcgac ggacggaac
 91151 gggacggggg ccgcggatc ggacgcgtcc gccccctccc cgccggaggt
 91201 cgccacacc ctgtccaccg ccgcactcgc cctcgtccag gagtggaccg
 91251 cacaggagcg cttcggcgcc tcccgctgg cgttcgtcac gaccggggcg
 91301 accggcgccg gcggtaccga cgtcatggac gtggcgccg ccgcggctcg
 91351 gggcctggtc cgttcggccc agtcggaag ccgggacac ttcgtcctga
 91401 tcgaccgtga ccccgggccc gcgggcacgc acgaccgac agccggccgc
 91451 gaacggggcc aactgctcct accgggcactg cacaccgacg aaccgcagct
 91501 cgccctgctg gacggcgggc tgcctcgccg ccgcctggcc cgcttcgaca
 91551 ccgcggccgc gctcaccgg ccggccgacc gggcctggcg gctcgacgc
 91601 accggcaagg gcagcctcaa cggcctcgcc ctgacccgt atccggcgcc
 91651 actggcgccg ctaccggcc acgaggtgcg ggtcgagggt cgtgcccgcg
 91701 gcctgaactt ccgtgacgtg ctcaacgcgt tggggatgta tccgggtgat
 91751 gatgtcggt cgttcggctc ggaggcgcc ggtgtggtcg tcgaggtcgg
 91801 accggaggtg accggcctgg ccccgcgca ccaggtcatg ggcattgaca
 91851 ccggcagctt cggctcgctc gccgtggacg acgcgcggg cctcgcccg
 91901 ctgcccagg actggtcctg ggagacgggt gcgtcggtgc cgttggtgtt
 91951 cctcaccgg taccacggcc tgaaggagtt ggggtggtcg cgggcggggg
 92001 agaagggtct ggtgcatgcc ggtgccggtg gtgtcggtat ggccggcgc
 92051 cagatcgccc ggcattgctg tgcgaggtg ttcgccaacg ccagtgaggg
 92101 caagtgggac gtgttgctc cgttcggcgt ggccgacgac cacatcgct
 92151 cctcccgac cctcgacttc gagggcgcc tgcggaagt cgccggcgac
 92201 cgcgccctgg acgtcgtcct caactccctc gccggtgact tcgtcgacgc
 92251 ctcgatcgcg ttgctcggg accggcgccg gttcctggag atgggcaaga
 92301 ccgacatccg cgccgggac tccgttcctc accgctctc ctaccagtcc
 92351 ttcgacctcg cctgggtggt gccggaaacc atcggcacca tgctggccga
 92401 gctgatggac ctcttcgca ccggcgact gcggccact ccggtccgca
 92451 cctgggacgt ccggcacgcc aaggacgcgt tccgctcat gagcatggcc
 92501 aagcacatcg gcaagatcgt gctcaccctg ccccgctcct ggaagccga

10

20

30

92551 gggaacggtg ttggtgacgg gtggtacggg tggctcgggt gggttggtgg
 92601 cccggcatct ggtgoggtcg tgtggggtgc ggcatttgtt gttgaccagt
 92651 cgttctggtg tgggtgctgc ggggtgcggc gggttggtcg cggagttgga
 92701 gtcgttgggc gcgcgggttg tggttgcggc gtgtgatgtg ggtgatggct
 92751 cggctgttgc ggagttggtt gccggtgtgt cggagtcgta tccgttgtct
 92801 gcggtggtgc atgcggctgg tgtgttggaat gacggtgtgg tgggttcgtt
 92851 gacgccggag cggttggtcg cgggtgttgc tccgaaggta gatggtgcgt
 92901 ggaacctgca tgaggcgacg cgtggtctgg atctggacgc gtttgttgc
 92951 ttctcgtctg ttgcgggtgt gtccgggggt gcgggtcagg ccaactatgc
 93001 ggccggtaat gcgttttttg acgcgttgat ggttcctcgg gtggctggtg
 93051 ggttgccctg tgtgtcgttg gcgtggggtg ctctggatca ggggtggtgg
 93101 atgacggcgg ggctgacgga gcgggatgtc cgtcgtgctg ctgagtcggg
 93151 tatgccgttg ttgacggttg atcagggtgt ggcgttgttc gatgcggcgt
 93201 tggcgacggg gagtgcgcgg ttggtgcggg tccgtctgga cctggccgca
 93251 ctgcgcaccc ggggcgacat cgcaccgctc ctccgcggcc tggtaaggc
 93301 gcccatccgc cgcgcagccg ccaccacacc cggcgacacc ggaactcggc
 93351 agcagctcac ccggtccag cgcgcggagc gacgggacac cctcctcgcg
 93401 ctgctccggg accaggccgc gatggtcttc ggccacacct cgggcgacgg
 93451 cgtcgacccg tccgcgcctt tccgcgacct cggttcgac tgcctcaccc
 93501 cggctgaact ccgcaaccgc atcggcgcgg ccaccggcct gcggctaccg
 93551 gccacggccg tcttcgacta cccaccgccc gatgccctcg ccgcacacct
 93601 gctcaccgaa ctgctcggcc ccgacgcca gtcggacccc gacgagcccg
 93651 gcgaccccac cgcgggaccc accgacgacc ccatcgtcat catcggcag
 93701 agctgcgcgt tccccggcga catcggtcgg ccggaggacc tgtggcgctt
 93751 gctcggcgac ggcgccgacg tcgtcaccga ctcccgacc aaccgcggct
 93801 gggacctgga caacctctac gaccccgacc ccgcgcacgc cggcacctcg
 93851 tacgcccgca ccggcggttt cctgcacgac gccgcgact tcgacgcca
 93901 cttcttcggc atgagccccc gcgaggccat ggccacggac tcccagcagc
 93951 gcctgtgtgt ggagtcctcg tgggaggcga tcgagcgggc cggcatcgac
 94001 ccgctgaccc tgcgcgacag ccgcaccggc gtcttcgccc gcgtcatgta
 94051 cagcggctac ggcacccgcc tcgacggcgc cgaattcgaa ggcttcacgg
 94101 ggcagggcag cgcactgagc gtggcctccg gccgggtctc ctacaccttc
 94151 ggcttcgaag gcccgcccat gacggtcgac accgcctgct cctcctcgtt
 94201 ggtcgccctg caacctcgcc caccggcact ccgcggcggt gagtgcaccc
 94251 tcgcccctgc cgggtgtgtc accgtgatgt ccatcccgga caacctcatc
 94301 gagtctctcc ggcagcgcgg actggccccc gacggccgct ccaagccgtt
 94351 ctccgagtcg gccgacggcg tcggctggtc cgaggcggtc ggaatgctgc

10

20

30

94401 tccctggagcg ccagtcggac gccgtgcgca acggccacca gatcctggcc
 94451 gtggtgcgcg gctcggcggt caaccaggac ggtgcgtcca acggcctgac
 94501 cgcgcccaac ggcccgtccc agcagcgggt gatccgtcag gcgttgacca
 94551 gcggcgccct gtccacggcc gacgtggacg ccgtcgaggc gcacggcacg
 94601 ggcaccacgc tcggtgaccc gatcgaggcc caggccctcc tggccacctc
 94651 cggccgcgac cgcgacccc agaacccgct gctgctcggt tcgatcaagt
 94701 ccaacctcgg ccacacccag gcagcccgcc gtgtcgccgg cgtcatcaag
 94751 atggtcatgg cgaatgcgca cggcgtgctg ccccgagcc tgaacatcac
 94801 cgagccgtcc tcgcacgtcg attggagcgc cggcgccgtc gaactgctca
 94851 ccgagcagac cgcctggccc gagaccggcc gggcccgctc cgcgggtatc
 94901 tccctccttcg gcatcagcgg caccaacgcc cactcatcc tggagcagcc
 94951 ggaggcccg cggcactcgg cgccggaaga agccgacacg gcggaggcag
 95001 ccgccaaagg gccggccacc gcgcacctgc ccgtaatgcc gtgggcactg
 95051 tccggcaaga cgcgggaggc cctgcgtgcc caggccgcac gcctcctcgc
 95101 ccacctccag cagcgcccc aactcgcacc cgcgcacatc gccctgtccc
 95151 tcgccacca gcgctcccag ttcacccacc gggcagtcgt cctgagcacc
 95201 gaccgtgacg aggcgacccg cgcgtgttcc gccctcgcca ccaccgccgc
 95251 gtcgacccc tcggccctca ccggcacggc caccatggga cgttgccggg
 95301 tgctgttctc gggtcagggc tcgcaacgtc tgggcatggg gcgtgagttg
 95351 tacgagcggt tcccggctct cgcgaggct ctcgatgtcg tgatcgatca
 95401 cctggacgcc gccttgcgcc cccaggccgg tttgcgtgag gtgatgtggg
 95451 gcgacgatgt cgaattgctg aacgagacgg gttggacca gcccgcgctc
 95501 ttcgccatcg aggtggcgct gtttcggctg gtggagagtt ggggtgtccg
 95551 tccggacttc gtggccggtc attccatcgg tgagatcgcg gcggcgcatg
 95601 tcgtcggggg gttctcgttg gaggacgct gccgtctggt ggcgcgcgg
 95651 gcgacgctga tgcaggcgtt gccggccggc ggcgcgatga tcgcggtcca
 95701 ggcgaccgag gacgaagtca tcccgacct gaccgacgag gtggcgatcg
 95751 cggccgtcaa cggcccgaac tccgtggtga tctcgggcgc agaagaggcc
 95801 acgcagaccg tggcacaaca cttcgccgac caggggcgcc ggacgaccgc
 95851 gctgcgggtc tcgcatcggt tccactcgcc gctgatgatg ctggcgaggt
 95901 tccgtgcggt ggccgagggc ctgtcctacg ccaccccgac cctcccgtc
 95951 gtctcgaatc tgacgggcca ggtggccacg gccgacgaac tctgctcggc
 96001 cgagtactgg gtgcgccacg tccgtgaggc ggtccgcttc gccgacggtg
 96051 tgacggccct cgaagccgag ggcgtgcgga ccttcctgga actcggcccg
 96101 gacggcgctc tcgcccctat ggccaggga accgtcgccg acgacacggt
 96151 caccgtcccc gtctccgca ggaacatgcc cgaggaaacg accctgctca
 96201 ccgcactcgg ccggctccac accaccgga ccccgatcga ctgggccgcc

10

20

30

96251 ctccctggccc cgaccggcgc ccgcccgggtg gacctgccga catacgcgtt
 96301 ccaacaccgt cccttctggc cctccggccc ccgcgacacc gcggatgccg
 96351 ccgccgtcgg catcgccggc gcgagccacc cgctcctcaa cggcatcgtc
 96401 gaactcgccg acgaagaggg cctgttggtt accggacggc tgtcactgca
 96451 gtgcgcatcc tggctggccc accacgccgt catgggacag gtccctgctgc
 96501 ccggcaccgc actgctggaa ctgcctctgc gcgcggcgga cgaggtcggc
 96551 tgtgaccacg tcgaggaaat gacgctcgcc gcaccgctcg tccctgccga
 96601 gcggcgccgc gtacagaccc aggtccgggt cggcgtcgcc gacaccaccg
 96651 ggccgcgcac cgtcacgato cactcgctgc ccgcacgcgc caccaccacc
 96701 gacagtgaca cccacaccgg caccgacacc ccgtggaccc aacacgccac
 96751 cggcgtcctc gtcccgggcc tgcggcgac ggcaaccgtc ccgttcgatg
 96801 ccaccgtgtg gccgcggcg cagccgaac ccgttgacct ggccgacttc
 96851 tacgctcccc gggccggcga aggattcggc tacgggcccg ctttcagggg
 96901 cctgcgagcg gcctggcgcc gcgacggcga ggtgttcgac gatgtcgac
 96951 tgccggaggc cggccgtacc gaagccgagg cgtacgggct gcctccggca
 97001 ctgctcgacg ccggactgca cgcagcctgg ctgctcgccc cggacgggga
 97051 gcccacacgg acgggcagcg tgccgttctc gtggcgcgcc gtcttccttg
 97101 ccgcttcgcg tgccctcctc gtccgcgtcc gactcggccg cgactccgac
 97151 ggaacgctga gcctggccat cgcgacacc accggtgcac cggtcgctc
 97201 cgtacaggcc ctctccatgc gcaccgtctc ggtgacggcc ctgagcgcca
 97251 ccgcgggccc cgcgcgcgac gcgctgttcc gcctggactg ggccctcgcc
 97301 ccggagccgg cgtgccagcc ggacgacacg gtgaccgtga tcccgcggt
 97351 ccgggtcgtc ggtacggaaa cctccgaact caccctcgag ctcaaccggg
 97401 ccctgcgtgc ggcgggcgcc gacgtcgacg tccgcacgac cctgtcgacc
 97451 gatgaacccg cgcgcgcgct gatcgcgctg ccgctcgctg cctccgacca
 97501 gaccggcacc gcagaagcgg caccgctccc ggccggccgtg caccacctca
 97551 cccgccgagc cctggccctc gtacagaccc gcctgcaaga gcagcacttc
 97601 gcggacacga agttcgtctt cgtgaccggt ggtgcgacgg tcgggctgga
 97651 tgtggctgct gctgcggtgt ggggtctggt gcgttcggcg cagtcggaga
 97701 atccgggttg ttttgcctcg gtccgatctg atccggatgg tcgggtgggt
 97751 gcggctgcgc tcgtcgctgc gttggtcagt ggtgagccgc agcttgcggt
 97801 gcgcggtgat gtgttgccgg tcgcgcgtct ggtgcggcgg ccgctcaccg
 97851 aggtcggctg cgggtgctgat ggcaccgggg atggcgtcgg ggggtggctct
 97901 ggtgtgtcgt tctcgggtga ggggtcggtc ctgggtcactg gtggtacggg
 97951 tgggtctgggt gcgggtgttg cgcgtcatct ggtggccgag tatggggtgc
 98001 gggatctgct gttggtcagt cgcagtgggt aacgtgccgt ggggtgctggg
 98051 gagttggtgg cggagcttgc ggggtgtgggt gcgcgggtgc ggggtggtgc

10

20

30

98101 gtgtgatgtg accgatcgtg ccgcggtggt ggagttggtt ggcgggcatg
 98151 cgggtgtccgc ggtggttcat gcggctggtg tgctggatga cggcatggtg
 98201 ggtgcgttga ccggggagcg gttgtccgcg gtgctgcggc cgaagggtga
 98251 tgctgtctgg catctgcatg aggcgacccg cggcctggat ctggacgcgt
 98301 tcgtcgtctt ctctctctc gccgggtct tcggcagtc ccgccaggcc
 98351 aactacgcgg ccgcgaacgc ctctctggac gcgctgatga cgcggcgccg
 98401 ggcggaggga ctgcccgcc tgtcactcgc atggggaccg tggtcgctga
 98451 ccgatggcac atcgggcatg ctgcgggacg ccgaggccga tcgcctgacc
 98501 cgttcgggag tgccaccgct gaccgcggag caaggactgg cactgttcga
 98551 cgcggccctg gcgaccggtg acgccacctg cgtcccggtc cgcctggacc
 98601 tgtcggcact gcgtgcccg ggtgaggtgc cgccttgcct gcggtccctg
 98651 atccgaggcc gctcgcgccg cgcgcgcccc gcggaatcgg caaccgccac
 98701 cgggctgcgg gaacggctcg tcggactgaa cccggctcgg cgacaggaag
 98751 tcctcttga cctcgtgcgc gccaggtgg ccctggtcct cggccacgcc
 98801 gacgccgacg acgttcaccc ggcacgtgcc ttcaggaggt tgggcttcga
 98851 ctccctcacc tcggtggaac tgcgcaaccg cctcaacacc gtcaccggtc
 98901 tgcggctccc cgcgaccatg gtgttcgact atccgaccgt cgaagtgtcg
 98951 gtctctacg tcctggacga gttgctgggc accgatgccg aggtggcgac
 99001 cgtgcagccg gccgcggttg cggtgccgga cgatccgac gtcacgtgg
 99051 gcatggcctg ccgctacccc ggcggggctg cctctccgga cgacctgtgg
 99101 aggcctcgtc cggacggcgt cgacgcggtg tcccccttc cgaccaaccg
 99151 tggttgggac gtogaatccc tctatcacc ggaccccgac catctcggta
 99201 cctcctacac gcgctcgggt gggttcctgc atgaggcggg ggagttcgat
 99251 ccgggggtct tcgggatgag tcgcgggag gcgttgccga ccgattccca
 99301 gcagcggttg ttgttggagt cgtcgtggga ggcgatcgg cgggccggta
 99351 ttgatccggt gagtttgccg ggtagtcgga cgggtgtgtt cgcgggggtg
 99401 atgtacagcg attacagcgc gatgttggcg agtcggagt tcgagggttt
 99451 ccagggcagt gggagttcgc cgagtttggc gtcgggtcgg gtggcctaca
 99501 cgttgggggtt ggaaggcccc gcggtgacgg tggatacggc gtgttcgtcg
 99551 tcgttggttg cgatgcactg ggcgatgcag gcgttgcgta gtggtgagtg
 99601 tgggttggtg ttggccggtg gtgtgacggt gatgtcgac cctgcggtgt
 99651 ttgtggactt tgctcggcag cggggtttgt cgcgggatgg tcggtgcaag
 99701 gcgtttgcgg atgcggccga tgggtgtggc tggtcggagg gcgtcggcgt
 99751 gttggtcctg gagcggcagt cggacgcggt gcgcaatggt caccagattt
 99801 tggctgtggt gccgggttcg gcggtcaacc aggatggtgc gtccaatggt
 99851 ttgacggcgc ccaatggtcc gtcgcagcag cgggtgatcc ggcaggcgtt
 99901 ggccagtggt ggtctgacgg ccggtgacgt ggatgtggtg gaggcgcatg

10

20

30

99951 gtacgggtac gacgctcggg gatccgatcg aggcgcaggc gttgttgccg
 100001 acgtatgggc gggatcgtga gcctgagcgg ccgttggtgt tgggttcggg
 100051 gaagtccaat ctggggcata cgcaggctgc tgcgggtgtg gcgggcgtta
 100101 tcaagatggg gttggcgtg cggcatgggt tgggcccggg gacgttgcat
 100151 gtggatgcgc ctctctcgca tgtggactgg tccgaggggt cgggtggagct
 100201 gctcagtgag caggcggcct ggccggagac gggtcgggtg cggcgggcgg
 100251 gtgtctcttc ctctggcatc agcggcacca acgcgcacgt gattctggag
 100301 cgtccggagg ccgcgcggcg tccgggtgat gagacgaaca ccgtggagcc
 100351 gtccacgggt ccgtgggtgc ttcccgcaa gacaccggag gccctgcgcg
 100401 cccaggcggc gaaactgttg tcgtcgatcg aggaacgccc ggagcttcgc
 100451 ctctcgacg tgggaatgtc cctggtcacc ggccgctcga cctttgaaca
 100501 ccgtgcgggt gtactcgcg ccgaccgtgc cgacgcgcgc cgcgcattgt
 100551 cggcgattgc tgcgcagag gccgatgctg ccgctgccac cggccgcgtc
 100601 ggccgcgggt gtcacgcggg gctgttctcg ggtcagggtg ctcaacgtct
 100651 gggcatgggg cgtgagttgt acgagcgttt cccggtcttc gccgaggtc
 100701 tcgatgtcgt ggtcgaccac ctggacgcgc ccttgcccgc ccaggccggg
 100751 ttgcgtgagg tgatgtgggg cgcagatgct gagttgctga acgagacggg
 100801 ttggacgcag ccggctctgt tcgccatcga ggtggcgtg ttccggctgg
 100851 tggagagttg ggggtgtcgt ccggacttcg tggccggtca ttccatcggg
 100901 gagatcgcgg ccgcgcgtgt cgcgggggtg ttctcgtctg aggaacgctg
 100951 ccgcctgggt gccgcccggt cgcgcgtgat gcaggcgttg ccggccggcg
 101001 gtgcgatgat cgcggtccag gccaccgagg acgaagtcac ccgcgatctg
 101051 accgacgacg tagcgatcgc cgcctcaac gggccgaacg cactggtcgt
 101101 atcgggtgtg gaggatgccg ccgtcgagat cggggcgcgg ttccggcccg
 101151 aggggcgtcg cagcaccga ctccatgtgt cgcgtgcgtt cactcgcgcg
 101201 ttgatggatc cgatgctggc ggagttccgt gtgggtggcg agggcctgtc
 101251 ctatgctgct ccgtccctcc ccgtcgtctc gaatctgacg ggccaggtgg
 101301 ccacggccga cgaactgtgc tcggccgagt actgggtgcg ccacgtccgc
 101351 gaggcgggtc gcttcgccga cggggtgacg gccctcgaag ccgagggcgt
 101401 gccgaccttc ctggaaactc gcccggaagg cgtcctcgcc gccatggcag
 101451 gagcctcgct caccgaatcc tccctcgcgg taccgctgct ccgtaaggac
 101501 cggccggagg aaccggcggc actcgcgcgc ctggcccagt tgcacatcgc
 101551 cggcgcgcgc gtcgattggc ccgtgctctt cgtgtgtgtg ggtgcggggc
 101601 ggggtggagtt gccgacgtat gcgttcacgc gtgggtgggt ctggccgggt
 101651 ggtcgggttg gtgttggtgg tgatgtgggt gctgtggggc ttgggtctgc
 101701 ggggcacccg ttgttggtg ctgcgggtga gttggctgcg ggtgcggggg
 101751 tgggtttgac gggctcgtctg tcgttgctgt cgcattggtg gttggctgat

10

20

30

101801 catgcggtga tggggcgggt gtttgttctt ggtacggcgt tgctggagat
101851 ggtgatgcgt gctggtgatg aggtggggtg tggtcgtgtt gaggagctga
101901 cgttggcggc gccgttgggt ttgcctgagc gtggtggggt gcgggttcag
101951 gttgctgtgg atgctcctga tgctgcgggt cgtcgtggtg tgggggtgta
102001 ttctgtctct gatggtgtgg gtcaggcgggt gtggtcgcag catgctgtcg
102051 gtgtgttggc ctctggtgtg gctgaccagg tcgggtgggt cggtgacgggt
102101 ggtgtgtggc cgcgcagggt tgcggtgtcg gtggatgctg agggctgcta
102151 cgagctgttt gcggatgctg ggttcggtta tggcccggtg ttccagggggt
102201 tgctgtcgggt gtggcgtcgc ggcgaggaac tcttcgcaga ggtgcacctg
102251 tcggacgagg ttgctgagag cgtctgatac gcgaccgggt tcgggttgca
102301 cccggcgttg ctggatgcct cgtccatgc ctgcttctc tctctccttg
102351 aaggtaaatc cgcgatggtt gggcctgcgt tgcggttcgc gtgggaggggt
102401 gtttccctct tcgcctcggg tgcgacggct ttgctcgtgc ggttggcgc
102451 ggcgggtgag catgcggtgt cgggtgaccg ggtggatccg accggtgcgc
102501 cgggtatttc catcgacgcg ctctgtacct gtcgctcac cctcgatgag
102551 gtcaacgcct ctacacacca gctgagcgat gcgtctctcg gcgtccaatg
102601 gaccacggtc ccgagcacc cggccgcgga ccaccgctcg gtagccatca
102651 tcggaaccga ccacctgggt ctgcgcgaag cgtcagcag ttctctgtct
102701 ggcgcgacca cgaccacgac cgcggccgcg tacgagagcc ttgacgcgt
102751 gatcgcggcg gggccgaag tgtccgtccc tgacgtcaca ctcatcggtc
102801 tcaccaccga ggacgccatc gctcagtaag tgaacgacca cgacgccacg
102851 gtggccggcc aaggcaccat cggagccggt gcggcggccg tagacgcggc
102901 tcgcccctc accgccagg ccctgcgcac gatccaggca tggttggccg
102951 acgagcgtct ggcggcaagg cgcctggtct tcgtgacctg cggcgcggcc
103001 gacgggcagg atgtcgcgc ggcggccgtc cagggccttg tgctctcgc
103051 ccagaccgag aaccgggca ccttcggcct cctcgacctc gacgggaccg
103101 aggcgtcgac cgcggtctc ggcgaggctc tcacctcga cgaaccgcaa
103151 ctgcttctgc gcgacggaca cctgcaagcc gccaggtga cccgcctggc
103201 gtcgcccgcg gacacggcgt tgcacacgga gtggaacgcg gacggcacgg
103251 tgctgatcac cgggtgtacc ggcgggctcg gcgcgcagtt cgcacggcac
103301 ctctctgaca ggtacggcgt ccgcaatctc ctgctcgtca gccggcgcgg
103351 cccgatgccc ccgggaacca cggagttggt cgcgcgagctg acggcgcacg
103401 gtgccgaggt ggcctgtcag gcatgtgacg tggccgatgg cgtatgcggtg
103451 gcggcgttgg tcgcccgcgt gccggatgag caccgcctga ggcgggtcgt
103501 ccacacggcc ggtgtgttgg acgacggagt gatcggctcg ctacccgagg
103551 agcggctcgc caccgtcctg cggcccaagg cggatgccgc ctggcatctg
103601 cacgaagcga cccgcggcct ggacctggac gcgttcgtcg tcttctctc

10

20

30

103651 cgctgcaggg gtcttcggcg gcgccggcca ggccaactac gccgcggcca
103701 acgccttctt ggatgcgctg atggcccagc gccgggcagc gggcctgcc
103751 ggactctcgt tggcctgggg gccgtgggac cagaccggcg gaatgacggg
103801 catgctgtcg gacgccgagg ccgaccgcct cgcgcgctcc ggeatcccgc
103851 cgctctccgc ggagcagggc ctgcacctct tcgacgcggc actcgctctt
103901 gccggaacca gcacgccaga cagggcagcc ggacgcgcg ccgccagcac
103951 gtcgggaacc ggcgacacga tcgccatccc ggccgcggcc ctcgctgcac
104001 cggtcgggt cgcacctggcg gcgctggccg cgcagggtga ggtccccgcg
104051 attctgcgcg gactggtgcg gacccgcacc cgcgctacgg cggccggcg
104101 ttccgtcacc gtggccggac tcgtcaaccg cctgtccggg ctcaccgccg
104151 acgagcggcg ccaggaactt ctgaaactgg tcgcactca ggcagccctc
104201 gtcctcgggc acgccgatcc ggcgctccgtg gactccaccg cacagtccg
104251 tgacctcggc ttcgactcgc tgaccgcctg cgagctgcgc aaccggctga
104301 gcacggccac cggcctgcgc ctgaccgcaa ccctggtctt cgactaccg
104351 aacaccgatg ccctcgcgga gcacctgcgg gacgagttgt tcggcgcggt
104401 cgagagcgag gtgcgggtgc cggctccaggc actgccgcg accgccgacg
104451 atcccatcgt ggtggtgggc atggcctgcc gtttccccg tgggtgtgac
104501 tcgcccgagg acctgtggcg cctggtcgac gccggcaccg acgccatcac
104551 caccttcccg accaaccgcg gctgggacct cgaatcgctc tacgaccgg
104601 accccgcaca cctcggcacc tcctacaccg gctccggtgg ctctctgcac
104651 gaggcggggg agttcgaccc ggcgttcttc ggaatgagcc cgcgtgaggc
104701 gctggcaacc gactcccaac agcgtctcct gctggaatcc tcctgggagg
104751 cgatcgagcg ggccggcatc gacccgctga ccctgcgagg cagcgccacc
104801 ggcgtcttcg ccggcgatga gtacagcgac tacgggagca tcctcggcgg
104851 caaggagttc gagggtcttc aaggccaggg aagtgcgggc agcgtggcct
104901 cgggcccgcg ctctacacc ctcggtctcg aaggctccgc cgtcaccgtg
104951 gataccgcct gctcctctc cctggtcgcc ctgcacctgg cagcccaggc
105001 ccttcggggc ggtgagtga cgctcgcgct cgcgggtgg gtgacggtga
105051 tgtccacgcc aggcacgttc gtggagtctt cgcggcagcg gggctctggc
105101 cctgatggtc gttccaaggc gttcgccgag gccgcggacg gcgtcggtg
105151 gtccgagggc gtcggcatcc tcgtcctgga gcgcagtcg gacgcccgtc
105201 gcaacggcca cgagatcctc gccgtgatcc ggggtcggc ggtcaaccag
105251 gacggtgct ccaacggcct gaccgcgcc aacggcccgt cccagcagcg
105301 cgtcatccgt caggcggttg ccagtggcg cctgtccacg gccgacgtg
105351 acgccgtcga ggcgacggc acgggtacga cgctcgggtga cccgatcga
105401 gccagggcg tcctggccac ctacggcgc gaccgcgacc ccgagaaccc
105451 gctgctgctc ggttcgatca agtccaacct cggccacacc caggcagcg

10

20

30

105501 ccggtgtcgc cgggtgtcatc aagatggtca tggcgatgcg gcacggcgtg
 105551 ctgccgcaga ccctgcatgt cgacgcgccg tcctcgcacg tcgattggag
 105601 cgtcggcgcc gtcgaactgc tcaccgagca gaccgtctgg ccggagaccg
 105651 gccgggtccg tcgcgccggt gtctctctct tcggcatcag cggcaccaac
 105701 gcccacgtca tcctggaaca gccggagggc gtgcagcgcc tggcaccggg
 105751 agcagcagag accgtggagc cggtcgccat caagccgtcg gcggaaccgt
 105801 ccctggtgcc gtggggcgtg tccggcaagt cacccgaggc cctgcgcgcc
 105851 caggccgcac gcctccgtga ctctctggcg gaacggcccg aaccgcgtc
 105901 gatcgacatc ggccactcac tggccgtcac acgctcgag ttcgaccacc
 105951 gcgcgatcgt gctggtcgac gatgcgaagg ccccggcgga cagcctggcc
 106001 gccctcgcgg ccctggcctc cgggtgtggc gateccgccg tcgtctccga
 106051 cgcggtatcg accggcgggt cggcagtgtc gtacacaggt cagggtgtc
 106101 aacgtctggg catggggcgt gagttgtacg gccgtttccc ggtcttcgcc
 106151 gaggtctctg atgtcgtggt cgaccacctg gacgcgcctc tgcccgccca
 106201 ggccggtttg cgtgaggtga tgtggggcga cgatgtcgag ttgctgaacg
 106251 agacgggttg gaccagccc gcgtctctcg ccgtcgaggt ggcgctgttc
 106301 cggctggtgg agcgttgggg tgtccgtccg gacttcgtgg ccggtcattc
 106351 catcggtgag atcgcgccg cgcatgtcgc cggggtgttc tcgtcgagg
 106401 acgcctgccg tctggtggcc gcccgtagca cgtgatgca ggcgctgccg
 106451 accggcggcg cgatgatcgc ggtccaggcc accgaggacg aagtgacccc
 106501 gcacctgacc gacgaggtgg cgatcgccg cgtcaacggc ccgacctccg
 106551 tgggtgatctc gggcgagaa gaggccacgc agaccgtggc acaacacttc
 106601 gccgaccagg ggcgcggac gaccgcgtg cgggtctcgc atgcttcca
 106651 ctgcgcgtg atggatccga tgcggcgga gttccgtgcg gtggcggaag
 106701 gactgtccta cgccacccc tccctcccc tcgtctcgaa tctgacgggc
 106751 tggctggcca cggccgacga actgtgctcg gccgagtaact ggggtgcgca
 106801 cgtccgcgag gcggtccgct tcgccgacgg catcaccacc ctcgaaagccg
 106851 agggcggtcg gaccttccct gaactcggcc cggacggcat cctgtccgcg
 106901 ctggtcagc agtccctcgc cggcgaaagc gtcaccgtgc ccgtctcgcg
 106951 caaggaccgc ggtgaggagt ccacggccct gacggcccga gcgcattctc
 107001 acaccgcgg actgatcga gactggcagg acttcttcgc tgggtgtgggt
 107051 gcggggcggg tggagttgcc gacgtatgcg ttccagcgtg ggtggtctc
 107101 gccggttggc cgggttgggt ttggtggtga tgtgggtgct gtggggcttg
 107151 ggtctgcggg gcatccgttg ttgggtgctg cgggtggagtt ggctgcgggt
 107201 gcgggggtgg tgttgacggg tcgtctgtcg ttgtcgtcgc atggttggtt
 107251 ggctgatcat gcggtgatgg ggcgggtggt tgctctggt accgcttg
 107301 tggagatggt gatgcgtgct ggtgatgagg tggggtgtgg tcgtgtgag

10

20

30

107351 gaactgacgt tggcggcgcc gttggtgttg cctgagcgtg gtggggtgcg
 107401 ggttcaggtt gctgtggatg ctctgatgc tgcgggtcgt cgtggtgtgg
 107451 ggggtgattc gtgtcctgat ggtgtgggtc aggcggtgtg gtcgcagcat
 107501 gctgtcgggt tgttggcctc tgggtcgggt gaccaggtcg gtgggttcgg
 107551 tgacggtggt gtgtggccgc cgcagggtgc ggtgtcgggt gatgctgagg
 107601 gctgctacga gctgtttgcg gatgctgggt tcggttatgg ccggtgttc
 107651 caggggttgc gtgcggtgtg gcgtcgcggc gaggaactct tcgcagaggt
 107701 cgcctgtcgc gacgaggttg ctgagagcgc tgatacggcg accggtttcg
 107751 ggttgacccc ggcgttgtcg gatgcctcgc tccatgcctc gcttctctcc
 107801 tcccttgaag gtcaatccgc cgatggtggg cctgcgttgc cgttcgcgtg
 107851 ggagggtgtt tccctcttcg cctcgggtgc gacggctttg cgcgtgcggt
 107901 tggcgccggc gggtagcat gcggtgtcgg tgaccgcggt ggatccgacc
 107951 ggtgcgcggc tgatttccat cgacgcgctt cgtaccgcgc gcctcacct
 108001 cgatgaggtc aacgcacccc acaccagct gagcgatgcg ctcttcggcg
 108051 tccaatggac caggttcccg agcaccgccg ccgcgcacca cccgtcggta
 108101 gccatcatcg gaaccgaccc cttcggcctc gcagacggcc ttctcgacgc
 108151 ctgcccctg gtgcgaggag cgggtgacct cgcggcgctc gcagcgtcgg
 108201 agcaccgggt accggacctc gtctctgtcc cggtagcggg caccgcgcgc
 108251 acaggcgtac ctgcggacgc cgaaggacac accgacgcgc ggacatccga
 108301 catgtccga tccgtgcgtg aggccaccgc acaggtactg gagcagatcc
 108351 agcagtgtt ggcggacgac cggttcggag cggcgcggtt ggtgttcgtg
 108401 acgcgcgggg cgggtttccgt gggtaggggc ggcacgcgcg acctggcggc
 108451 ctcgccgctc tggggtcttg tgcggtcggc gcagtcggag aatccgggct
 108501 gcttcgggtc tctcgacctc gacctcgacc tcgcccctga ctccgacctt
 108551 gcccccgagg tcgacatcga gcgcgacctt gaccgcgac cggtcggtgg
 108601 gaccgtgcag ccgcgcgtcg ccgcggccct gcacgcgacc gccgacgagc
 108651 cgcagttggc actgcgcggc gggaccgtgc aggcgcgccg actgacccga
 108701 atccccgcgc cgcagaccga ccgtgccgag accgacctg ccgagaccga
 108751 ccgtccggag atcgacaccc ggcggcccg caccgtgctc atcaccggtg
 108801 gtaccggtgg cctcgggtgg ttgctcgccc ggcacctcgt cgcgagcgg
 108851 ggggtacgga gcctggtgct cgcagccgg agcggctctc cggccgaggg
 108901 agcggagaag ctggtcgcgc acctcgaagc gctcggtgcc gtggtggccg
 108951 tgcagacgtg tgatgtggcc gatggcgatg cggtggcggc gttggtcgcc
 109001 ggcgtgtcgg acgagtaccc gctgacggcg gtcgtccaca cggccggtgt
 109051 gttggacgac ggagtgatcg gctcgctcac cgaggagcgg ctgcaccgc
 109101 tcctgcggcc caaggcggat gccgcctggc atctgcacga ggcgacccgc
 109151 gatctggacc tggacgcgtt cgtcgtcttc tcctccctcg ccggcgctct

10

20

30

109201 cgggtggcgcc ggtcaggcca actacgcggc ggcgaacacg ttcttgagcg
 109251 ccttgatggc gcagcgtcgc gccgcgggt tgccgggtgt gtcgctggcg
 109301 tgggggtccgt gggaccgggc cggcgcatg acggggaccc tgtcggacgc
 109351 cgaggccgac cgcctcgccc gctccggtgt tccgccgacg tcggcgagcg
 109401 agggccttgc gctgtacgac gcggcgaccg ccggtgagcg gccgctggtg
 109451 gtgccggtgc ggctggacct cgcgcgctc cgggggctcg gtgatgtccc
 109501 ggcgctgctg cgcggactgg tccggacgcc cgcgcggcgg accgcggcgg
 109551 ccggtgcggc gccgtcggcc gatgtgctca cccggcagtt ggccgggctc
 109601 ggcggggcgg agcaggagga ggtcctgctg aggctggtgc gcggtcaggc
 109651 cgcggtggtg ctcgggcacg ccgacggctc ggcgatcggg gcggggcgac
 109701 agttccagga gttgggcttc gactcgtga ccgcggtgga gttccgcaac
 109751 cgactcaacg cggccaccgg actgcggctg ccggccaccc tgcgtttcga
 109801 ctacccgacg ccggccgacg tcgtcgggca cctgcgcggc cggctcggca
 109851 ccggggaggt gtccgggtgc ggctcgggtc tggcgcgctt ggacaacctt
 109901 gaggcggtga tcgccgggct gtccctcgac gacgcggggg agcaccagtt
 109951 ggtggccggc cggctggagg tctcagggc gaagtggcg gacatgcgaa
 110001 gcgcggaggg agctgtggac ggcggtgcgg acgtcgacat cgaggaggcg
 110051 tcggacgacg acatgttcgc gctgctggac gacgagctgg ggctgaactg
 110101 agccgctcgc catgagcagt tccgcaccag aagttccaca gtggcctggt
 110151 ccacagtgcg ctgggtccaca gtgagctgtt ccacagcgac ctgttcccg
 110201 gcggttccgt agtgagccgt tccgcttagg cgggaaaccg gggcccgggt
 110251 ggtcggacgc ttgattccgg ccaccgggag cgtcaaaccg gcctcttcta
 110301 aagaaaggga aagaactagg gaaacaccca cagcccgatc atgtgcaatg
 110351 aattccgtgg ctcggaatac catcgcggg ccatggagtt accgtgtgat
 110401 cacgtgtccg tacggcagtg acggcgtgca cgggttgcca cggcaggtcc
 110451 ccggccgagg cgtctccccc ctgctgcccg gcggtgcgtg ccgcttggtt
 110501 cgtccctagg aggtccccc atgaccaagt ccaccgagga gagcctgtgg
 110551 gcccggtgct tccatccagc accggccgcc ccggtccggc tcttctgttt
 110601 ccacatgcg ggcggctcgg cctccttcta cttcccggtg tcggcccaac
 110651 tgtcctcggg tgccgaggtg ttccgcatcc agtaccggg gcgccaggac
 110701 cggcgcaagg aagccggtgt cagtgaacct gcgaccttg ccgaccaggt
 110751 ctacgacgcg ctgcgcccc tgcggaagg gcggccgagc acgttcttcg
 110801 ggcacagcat ggcgcgacg ctggccttcg aggtggcccg gcgcttcgag
 110851 gccgacgacg gtgacctggt ccggctgttc gcctccgggc gccgggcccc
 110901 ctcccgcgtg cgtgaagagg ccgtgcaccg gcggtccgac gacggcatcg
 110951 tcgaggagct gaagctgctc gccggcacca acaccgcgct gctcggcgac
 111001 gaggagatcc tgcggtgat cctgcccgcg atccgcagcg actaccaggc

10

20

30

111051 cātcgagacc taccgctgcc cgcgccacgt caccgtccgg gcgccgctga
 111101 ccgtccctcac cggcgaccgc gacccgaaga cctccctgga cgaggccgag
 111151 gcgtggcgcg gccacaccac cggggacttc gacctcaagg tgcttcccg
 111201 tgggcacttc ttggtcagct ccgaggcccc ggcgatcāt c gatctgctcc
 111251 gggcgacact cgcgcgcaac ggctāgcggg cgcactgcgg caggccggcg
 111301 gtgccggtct gccgcacggc ccccgcgccg cctgagacgg caccatgcca
 111351 cccggcgacg cgcgttgctg gcgcctgtgg agcāccgtcc cgcgtgcgtā
 111401 cgcgggcgcg cggaggacgt gctcggccgg ccggatgtcc tggaggagac
 111451 cattcggaāā acgtcaāatg tctgacagct gctgctcatt tactccagcc
 111501 cgcactatat gatcttggtg tggggcātg acggagcgtg cactcgttc
 111551 gctcctggtg cātccatccc gaccgtccgg āttgtgtgc accggacggc
 111601 cgtggtcctg gtacgcgtcg gcgcgaggtc ggctcctgtt ctcācccccā
 111651 gtggaggāac gcgtgatgcg gaagcagtcg ggttcttccg gcttactgac
 111701 cācgttggtg ggācgcgacg ācgaactgcg cāccctcgcc cggcācgccg
 111751 cggccgcccc cgaaggcggg gccggcctgg tctgtctcca cgggcccgc
 111801 ggcatgggca āgāccagtct gctgcggtcc ttcācggga gcgatgtctg
 111851 ccgcgcatg ācgggtgctg ācggcācctg cggcgagacc gtgcggcg
 111901 cggggtācgg cgggtgtgcg gaactcctcg gcgggctcgg cctgācgggc
 111951 ggcgācgccc ggcgctcgcc cctcctggag ggctggcg cccgcgcgt
 112001 gcccgāctc āccgcagacc ccgcgggtcc cgaacgcgcc āgggtgct
 112051 ācccggtgct gcācggcctg tactggctcg cgcgccct catggccāā
 112101 cggccgctgg tctcgtctct cgaacgāgtc cactggtgcg ācgaacgtc
 112151 cctggcctgg ātcgācttcc tgcgtcgacg cgcggaggac ctgcgctgc
 112201 tggctgctct ggcttggcgc āgcgāggccg āāccggtcgc gcccgcggtg
 112251 ctgcgcgāāā tgcgcgcccā gcgcgcgcc āccgtgctcg gctgcāccc
 112301 gctcgccccc gācgācatāg gcgāāatggt gcgtcgctc ttcggāccā
 112351 cggccgcācc gtggttcgtc āgcgggtcg ccgcctgtc cggcggaat
 112401 ccgctggccc tgcgccgct cctcgācgāā ctccgcgcg āggcgctccg
 112451 gccggācgcc gccggggāgc gccgggccc cgaaggtcgg āgtācgtcc
 112501 tgcgccgtc ggtgcgtgc ctggtggāgc gccggccgc ctgggtgcgc
 112551 ggcggtggccc gtgccatgc cgtactcggc cggāgtgca ccgāgtgct
 112601 ggcggcgtc gccggcgtcc cggccgcgāc cgtcgācgā gccctgttg
 112651 tgcgtcgāg ggccgcāt ctcggccgcg āccgcgtgga ctctgtccāt
 112701 gācgtcgtcc gctccgcggt gctcgācgā gtgcgccgc ccāccctggc
 112751 cgaactgcgc āccāacgcg cgtgtgtgt gagcgācgcc ggccgcccct
 112801 ccgāggāgct cgcgggccāg ctcatgctgc tgcgggtgct cgaacgāccg
 112851 tggatggccg cgggtgctgc cgaacgcgc gccāggcgg āgācgccgg

10

20

30

112901 cgccccggag gccggtgtgc gctgcctcta ccgggtgttg gaggtggagc
 112951 cggacaacgt tgccgtccgc atccagatgg ccgcgcgcgt cgcagagatc
 113001 aacccgcccg aggcgatgcg cctcctcaag gaagcgctct cctcgcgcgg
 113051 ggacgtccgc acccgcgcgc aggtcgccgt ccagtacggc ttcacctgcc
 113101 tcgcggtgca ggaatccccg tccggggtgc ggatgctgga ggacgcgtc
 113151 gccgagctga ccgcgcgaact ggccccgaa ccagggcccg tggaccggga
 113201 gttgcggacc ctcgaggagt ccgtgctgct catcgtcggg gccgacgaga
 113251 aggtgacgat cgtgcggtc cgcgaccggg cggcccggt caccatgccg
 113301 cccgcgcaca caccggccca gcggcagatg ctggccatga ccacgtgct
 113351 gaccgcgatg gacggccggg acgccggtc ggccgtcgac caggccgcc
 113401 gcgccctgcg cgcaccggc gtcgagctgg aaccctggtc gctgctgtcc
 113451 gcctccttcg cctctccct ggccgacgag gtcgccgacg cgcagtacgc
 113501 actggacctc atgtccagt acggccagga caacgcggcg gtgtggacgt
 113551 acgtcctggc gctctccacg cgcgcctgc tccaccacgg ggtggcgcc
 113601 tccccgagg ccttgccga cgcacagacc gccgtcgaga tctcggcga
 113651 ggagcgctgg gcggacggcg ccgtgctgcc ccgtgtcgcg ctggccaccg
 113701 cctcgtcga ccgcggcgag ccgagcgcg ccgaacacgt cctcgacggc
 113751 atcaccggc ccgcctcga acgttcgtc atcgaatacc actggtacct
 113801 ccaggccgc gccctacgcc gctgggtccg cggggatttc caaggagccc
 113851 tggacctcct tctcgctgc ggtcggtccc tggaggagtc gcgcttcage
 113901 aacccggcgt tcgtgcctg gtgggccgat ggcgcggtgc tcttgccgac
 113951 cctggaccgc cagcaccag cgcgcgaact cgcgcatac ggaagcgagt
 114001 tggccgagcg ttgggggacg gcgcgcggcc tcggactggc cttcatggcc
 114051 cagggcgtcg ccgcaccgg ccgcgcggc atcgatcac tcaccgaggc
 114101 ggtctcgctg ctgcgggact ccccgcccg ggccatggag gcccgggccg
 114151 aacttcttct cggacaagcc cactgaagc gcgacgacct gcgggcccgc
 114201 cgggaacacc tgcgcgcgc cgcgacctc gccacgcgt gcggcgccgt
 114251 gaagctcggc gtcgacgcca gaaaactgct ggtcaccgg ggtggtcggg
 114301 tacgcaggat gaccgcctcc cactcgaca tgcctgaccg gatggaacgc
 114351 acggtggcgg acctcgcggt gaccggcgcg agcaaccgg ccatcgcgga
 114401 agccctcttc gtgactgtaa ggaccatga aaccatctc acgagcgtct
 114451 accggaagct cggggtcggc gggcggtcgg agctgtccgc cgtcctggag
 114501 accaggaccg ccacctccg tcggcagccg ccggcctggg tctcccaggc
 114551 acgcggacgc gcttgagacg aacaggacga gaggtagcgg tgccccggag
 114601 caagcgcgcg aaccaaccga ccacctgcac gccgcagtgc gcacccgacg
 114651 cccacggcga cccaccatg ctctcgaat gcggccggga acagcggtc
 114701 atcgcgacc tctgcaccg cctcgccag gggcgccat cgtgtctcag

10

20

30

114751 cctgaccggc cggcccgggc acgcccagaa cgccctggtc cgtgggggcg
 114801 cgtgccgggc caggcacgac gggctgcgcg tcctgcgagc ccaggcgacg
 114851 cccgcggaac gggaaactccg ctacggcgcc gttctccaac tgetggccgt
 114901 cctcgacggc ccgcacggca gcacctgga cgcgcgata cgcacgacg
 114951 gtcccccgcg actgcccggtg cccggcatcg aggagggtgt gggcgccacc
 115001 ggcacggcac ccacctgggt cgtggtcgaa gacgtccagt ggttggaacc
 115051 ggcctcgctg acgtgggtgc agatcctgct gggccacctc gggccggaca
 115101 cccgcctcgc cgtcctggcc agcagctgcg gtgacaccac ggccttcgac
 115151 accgaccgga agggccccgc cgtcccgggg ccgcccggaca ccgtgcccg
 115201 cgcgcgcttc gtgggtgccg cgtcaccga ccgcggggtc gccgccaccg
 115251 tcgcgcctgt ctgcggcacc cccggcgacg aggagtcat cgcgcgctc
 115301 acctccgcca ccgcgggcaa ccccgccatc ctgcgggacg cctgcgcgc
 115351 cttcgtcgac caeggcctcc ccgcccagcg cgaccacctc ccggagctgc
 115401 acgcccctac cgtggtgctc gtcggcgacc acaccgtgcy gcacctggac
 115451 ggcctgcccg ccgaagtcaa cgcctcctg cgggccctgg ccgtctcgg
 115501 cgacctgctc gaattccacc gagtccgggc cctcgccggc gcgcactcgc
 115551 tgtccgagga ccgatccgc acctgctgg cgcgcgtcgg cctgaccgtg
 115601 tcctcgcgcg acaaggtgca catccgcttc cccgcctcca aggcacgcgt
 115651 catcgaggac atgcccgcgc cggagcgcg cgatctgtac gtccgcgcg
 115701 ccgaactcac ccacagttgc ggcgtcaacg acgaggacgt cgcctatctg
 115751 ctgctgcgct cgtcgccgct cggcgcaacc tgggtcgtgc cctgctccg
 115801 ccgcggattc gccgcgcgc tgcgcggga ggaccaccac cgggctgtg
 115851 cctgcctctc ccgcgcctg caggaaaccc tcgaaccccg ggaacgcagc
 115901 ctgctgacgt tggaaactgg ccgcggcgaa gccgtcgcc ggccggaggc
 115951 gggggatcga cgcctggggg aactcgctcc cagcacctgc gcggacaccg
 116001 accccacgtc gtccggtgag ggggtggggg tcgcgcctat cgacctgggg
 116051 ttgcgccggg gcaacagcga atgggtccgc cgcaccgcgg gcgaggccct
 116101 gccgtacgcc gggccggccg accgggagga actggtcgcg ctgttctggt
 116151 tggccgcctg gcgggacgac gacgcgcga tgatcccggt ggtgcccgg
 116201 ttgcccgacc ggcgggtgcc gccggccag gccggcgccc gtgctggca
 116251 gctggccacg gcgggggagg acgcggacaa ggccagggaag ctgcccggga
 116301 tcgccctcac cggcggggtg aacgagagcc tgatgatgcc gaaactggcg
 116351 gectgcgcgc cgtgttctgc caccgacgac aacgacgagg cgtgacacg
 116401 cctggacacc atgtcaccg ccgccgcag tgcccacctg cgcagcatgg
 116451 ccgccgcctt tttcaacctc cgggcccga tacacctgtg cgcggcccgg
 116501 ctggaggccg ccgaacgcga tctggacagc gccgagcgcg cctgcccgc
 116551 gacgagttgg ccccccggtg cgtgcccga cctgatcgcc acccgcatcc

10

20

30

116601 tgcgcagcat ggagacgggc cgcgcggacc gggcccgccg actcgccgag
 116651 gccccgggtcc ccgcccggcg cgaggagggt gtgtggtggc ccgcctgct
 116701 gctcgcccg ccccggtgtg ccgcccagca cggtgactgg gaggaggccc
 116751 tgccgctgtc gcgggagtgc gggcgctggc tttttcgccg gcaactgggc
 116801 aaccgggcca tgcctcagttg gcgcccgtg gccgcccagg cgtgtctgaa
 116851 gctcgccgac gtgacggagg cgcgcgggt gccggacgag gagctgttct
 116901 tcgcccagcc ctggggcacc gcgagcgccc ggggatgc ccgcctgacg
 116951 acgcgcgac tcttcgacga cgacggcgac cgggcccgtc ggcggatccg
 117001 cgaggccgcc gccctgctcc gcgactcgcc cgcgcgctg gctacactgt
 117051 ggagccggct gagccaggcc ggtgccgaga cggcccacgg cgacaccgcc
 117101 gcggccgac gctcctggca ggcggtcgcc cggatgaccg ccgcccaccc
 117151 cgcacggcg ctcgccaccg ccgcccgcac cctgaccgtc ccgtccgttc
 117201 cggtcgccc cgcgcgccc accgcgctg tcccaccgg atggcgcgac
 117251 ctgtccgagg cggagaagga caccgtgctg ctcgcgccc gcggccacgg
 117301 caaccgccag atcgccgaac aactcgccgt cagcaggcgc accgtggagc
 117351 tccggtgag caacgcctac cgcaagctga ggatcgggcg acgcaaggag
 117401 ctgtacctgc tcttgaggc gctggaagga ccggtcgcg atgcttcttg
 117451 agcgggagaa cgaactggcc cggatccggg ccgcccgtga cgcgcggaa
 117501 gcggcgact cctcgctcct cctgatcaac ggtcccctcg gcagcggcg
 117551 ttcggcgctg ctgcgcccga taccgagct ggcggcgac ggcaccgcg
 117601 tctcggggc cagcgccgc tggcggaac gcgacttccc cttcgggatc
 117651 gccgcacaac tcttcgacca cctgctgtcc gggcgggcg gcgcaggcc
 117701 ggccgaacgc accgcgggg cagagcactt cagccgactg atggacaccg
 117751 gcgacggccc taccgggacc ggcgcgccc tggaggtctc ccaggcagtg
 117801 ctcaggggcg ccagcgct gctcgccgac gcgtccgcg agcggcgct
 117851 gctgacctg gtcgacgacc tccagtgggc cgacggccc tccctgcgt
 117901 ggctggccca cctcaccgg cggtgcacg gcctcgggc gctgctggtg
 117951 tgcacgctgg ccgacggcga ccaccgggc aggtacccc tggtcggga
 118001 ggtcgccggc gccgcgaca ccgtcctg cctggcgccg ctgtcccggg
 118051 acgccaccgg cgtcctgctc gccgggccc agggccggcc gccgcaggac
 118101 gcaactggtg gcgccgtgta cgagcgctcc aggggcaacc cgtgttctc
 118151 gaccgccttc cggagcgctc tgcgcgccac cgggaaggcc ccggcgggcg
 118201 accacttcgg cgcgctccg gagctgagcc cgacggtgct gcgcatcg
 118251 ctcgcgggcc atctcgggat ccagccgag ccggtgcgag aggtcgcggt
 118301 ggcggtggcc gcgctggcg accacagca tccggtgctg ctcgccagc
 118351 tcgcccgggt cgatgaaac ggtttcgccg gtgcccgcg cgcgctggtg
 118401 gacgcgggcc tgttgggccg gggacgggac gtccgcttcg tccacggcgt

10

20

30

118451 cgtccgcgat gcggtggact ccctgctcac cctcgacgag cgggaacgct
 118501 cgcacgacga cgcgcgcgat ctgctgtacc gctgcgggcg gccggccgag
 118551 caggctgcgc gccatctgct ggccgtggtc cacccgggcc ggccctggtc
 118601 ggaggcggtc ctgcgctccg cggcccacaa cgcgctgcgc gccggccggc
 118651 ccgccgacgc ggcccgggtac ctgcgcgcgc ccctgctgca ccaccgcacc
 118701 caggacggct gcgcgcgcgc catcctggtc gatctggcca ccgccgagcg
 118751 cgcctcgcac ccgatgcct gtgtacgcca cgtcagccag gcggtcgcgc
 118801 tgctggacac ctgcgcggat cgggcgcgcgc ccgtgttgcg catccgcgcg
 118851 tccctgctcg ccgccccag ccctgcgcgc gtcgagttgg tgcggcaggc
 118901 cgcgcgcggg ctcgacgaac cggggcagcg ggacgaggag ggagccgacg
 118951 aactgcctct gcgcctggag gcgtggctgc ggcactccgg ccacgagaac
 119001 ccctgcgagc tggcgctctc ggtggcgcgc ctgcgcgcga tgggggcacg
 119051 gccgcgggtg gacagcgtcg ccgaacgcga gttggctgcc gtgctgttga
 119101 gtgcgggcgc gctcagcgcgc cggctcagcg ccgcggagat cgcgcacacc
 119151 ggcaaccgca tctcgaacg tgagcgggc accgcgcgc atgccacac
 119201 ccctgctccg ctggtgatgc tctcgtgtt cgtcgcgcag tccgtgcagg
 119251 gcgtggcctc ctggctggcc agcgaacagc acaccgcgc ccggtacgcg
 119301 accgcgcgcg acgatgtgct gctgaccgc gagcgggct tgcctctgtt
 119351 gacgcagggc cgcgcgcgc ccgcgcggga gcacgtcgaa cgcgcctcg
 119401 tcatggacgc cggcgactgg tcggaaccgc ccgtcatgat gttcgcgcgcg
 119451 gtcgccttcg aactgcgcga ccgcgccttg agcgaacgca tcttgaacg
 119501 gatccgcgac cgcgcgcgcgc ccggaactggc gctcacgcgc accggtcaga
 119551 tgctccaggc cgcgcgcgc gcgcacttcg gccggggcg ggacgcctg
 119601 gacacgctgc tggcctgtgg ccgacgcctg gagaccgtgg gatggcgcaa
 119651 ttccgcaact ctgccttggc gtccgtatgc aatcgggctg caccagcggc
 119701 tcggcgagac cgatgcgcgc ctgcaactcg ccgaggacga gctgaggtgg
 119751 gcccgggagt ggggtgcgac gacgaacctg ggccgggccc tgcgcctgaa
 119801 gggctggctg ttgcaggacg agggactgga tctgctgct gagagcgtcg
 119851 agatcctccg cgcttcgtcc tacgcgacgc aactggcgcg caccctcgtc
 119901 gtctcgggc gtgcgctgcc gggtagggcg gaggcgagg cgggtgtgcg
 119951 ggaggccgcg gggatgcgcgc ccgcctgtgg tgttcctgg ctgcgcgaac
 120001 gggccgaact cggcttgggc agtgccatcg tgccgcgggt cgcaccctg
 120051 accccagcgc agcgcgcgag ggctgcgctg gtgagccgg gcctgacca
 120101 ccaggccatc gcgaccgaac tcggtgtgag ctcccgggcg gtggagaagc
 120151 acctcaccag cgcctatcgc aagctcggcg tctccggcg ccgcgagttg
 120201 gtgaatgcgc tcccggtcg ttgacgcgc cggcgcctcc gttccggacg
 120251 atcacaatcc tctgtgcctg tcagccctgt gggtagggcg tcttctgctg

10

20

30

120301 acgcccgtcc ggtcatgaca gcaatcctca acaaaagctc aaacggaagc
 120351 gggggcaggt gcaccacccc catcaccaat gtctcgtatg ttacttttgg
 120401 cccaccgtag ttgggactct gcggaccgta cctcgtcaag tacggaccac
 120451 ggatttaggg gtgcgttagat gtgtcgcgtt gacgtgccga gtgtgtgcta
 120501 actacgctcc gtttcacgga aatcgaaact cttcggcgtg aatcgcggcg
 120551 aacttcgata gaattctgga ttccgtttct cagaaggcta aaagacgacg
 120601 ggggttatcc gaaaggatgg cggctcctgt actatcactc acctcactga
 120651 ctcgaaacac cgtacgaccg gcggggtgat atccgctcag accgcaccgg
 120701 cgggggagtc cgtcggcccc gggtgatgg cgtccctcga ccgtgacctc
 120751 accatcaagc acgccaacca ggagtccgc cgcgcttcg acgattccgc
 120801 cggtagcgtc tgcggtcgca gcttcggga cctgatgac ccgagcgtgc
 120851 agcagccgct gatgcgccag ttctcccgcc tgatcgaggg caagcggcac
 120901 cgcttcgct cgcacgtggt cgcggtgggc gccagggacg cggccttcgc
 120951 gggcaccctc accgcctccg cggtcaccgg caagaccccg gacatcgccg
 121001 ggatcctggt gctcatggac tcgtccggcg ccgcccagcg cgcgatgcc
 121051 ggcgtcgtga ccagccagaa gaagtccctg accgagatcg atgcgcgat
 121101 cctcgaaggc atcgcggccg ggctgtcgac cattccgctc gcctcgcggc
 121151 tctacctcag ccgccagggc gtggagtacc acgtcaccgg gctgctcggg
 121201 aagttgcggg tgcccaaccg cgcgcgcctc gtctcgcgcg cctactccat
 121251 gggcaccctg aacgtcggca cctggcccc gaaggtcgtg gacgacttca
 121301 tcaagtgacg cccggccggc ctgtcccgca ccggccgggc accgtcggac
 121351 cgctgtcccg cagcccggtg tcgcgcagcg ccgcctctca cggggcgatg
 121401 cccggcggca cgggcgttcc gtgccctggt cgatgtttat ccattctcac
 121451 gttacgaacc cgcaatgctc atgcgcgcga ttccggccac ttttcgttcc
 121501 gtcgtaacca aactgcccggt catatcgttt ccactgaccg tttctgtggt
 121551 ccgcggtgtg ccacctgggc cccactccga cgcgcccgcg acgggtcggg
 121601 agcggcctcc gaatactcac ggtaaccccc gttcctcctg gtggagtcc
 121651 ccgcgcgtca cattcctccg tgccctcgtg ggcgtggctc gccgccttc
 121701 cccgcgcggc cgacagccat tgttacaacg gcacgacaga gcgggggtgg
 121751 tcgacccgct cgcggtcggg cgcagtcag cggaaacctct gctcgatgtt
 121801 cgtgccaccc cgaactctgt gttcttgggt accgttttca ccaatgcccc
 121851 gccggttcgg tggcgaatcc gccgccggtt cctggccaca acggatcgct
 121901 tcggccgccc accacggccg gcgcagccga ggcttgttgg aatcaccccc
 121951 acccagtcgg ggcgcggtt tggcccgga acgcaggagg aacaccgtgg
 122001 atgctgaggg ccgcgcagg gacatgctgg agctgatccg gcgcagcgga
 122051 tccgccgatg tcgtgcgtct cgcgaggag ttccgcgtca gcaaggaaac
 122101 ggtccgcggg gacctgaacg tcctggaggg ccatgggctg atacggcggc

10

20

30

122151 ggacacggtgg cgcctatccc atggtgcggc cgggctccga ggccgtcttc
 122201 gtctcccgga ccgcacagcc gatcccgag gagtcccgga tcgccaccgc
 122251 ggccgcccga ctccctcagcg aggcgcagac ggtcttcac cgcgagggct
 122301 tcaccccgca actcatcgcc gacgccctgc cgcgcgaccg gccgctgacc
 122351 atcgtcaccc cctcgtctcc ggtggtcagc gccttcgcga cgagcccaca
 122401 ggccaacgtg ctgctcctgg gcggccgggt ccgcgggggc acgacggcca
 122451 ccgtcgacca ctgggcccgc cacatgctgt ccggcttcgt catcgacctg
 122501 gccctcctcg gcgcggaggg gatctcgcgc aggtacggcc tgaccacccc
 122551 cgacccgcg gtgcgcgagg tcaaggccca ggccatccgc gtgcgcgcgc
 122601 gcccggtcct cgcgcgggtg cacaccaagt tcggcacggc gagcttctgc
 122651 cggttcggag aggtgggcga cctggagacg atcgtcaccc gcgcggccct
 122701 gcccgctgcc gaggcccacc gctaccacct catgggcccc aaggttttac
 122751 ggggtgtgac ccgcgggggc gtcccgcgc cgcgccacgg cctggcgcgc
 122801 gccggtgcgg gtcagccgcg tcgcgcgtcc gtccggtccc gccagagtgc
 122851 gagcgggtgc acccccaccc ggtagccctc ctgcgcgagg agatggccgc
 122901 cgtgcttcat gtgccagagg gcgaccgcgc gttcggtgac ggtcaggctc
 122951 gcggcccggt cggcgagtcg gacctcgaag cgctgggcgt cgtcgaggga
 123001 gcgcaccag gccaccagg gtccaggtgtg ccggctgatg acgctcgcgc
 123051 acagcgggac ctgcgcgcat ccggtgaccc ggcggtcac ctgcgcgagc
 123101 ctgcgcgcgc ggacctgcc ccagaagctc accgtgaccg gccactcgga
 123151 cagcgggcgc gcgacctgc accgggcgtg caacatgtcc gcggcgaaga
 123201 gccgttgac gcgacgggc acggtgtcgg ggccggcgcc gcactgctcg
 123251 gccagcgcgc ggtaggtggc ccggccgtcc acggacagcg cggtgaccag
 123301 ctgttggtcg agctcgtcga ggacgaaggc gggggtgtcc gtgcgggtgc
 123351 tggacgcgtc ggcggcgagc cgggcgacct ggtgccggcc gagggcccgc
 123401 aggcgccacc ggctgccctc ggtgtgcacc gggccggcca ggtgcgtccg
 123451 ggcgccccg acgcgcgtcc acgcggcgag gtcgtgggtc acccaacggg
 123501 agagcatggc cggatcgccg gccatgacgt tgagttggag gtcacggctc
 123551 ccggtgacgt gggagagcgc caccacgtgc ggcgcgccg ccaacgcccg
 123601 cgcgacgtcg agcagtcggc cgggagcgca gtcgatctcg acgaacgcca
 123651 ggcacccctg gcccgacgc gccagcaccg gggccggatg gcagctgato
 123701 caggcggcgc cggctcgcac cagccggttc cagcgccctg ccaccgtcac
 123751 ggcgtccagt ccgaggacgg agccgatccg ggtccagctg gcccgcgcg
 123801 tgatctgcag cgcgtgcacc agggcctgat ccacatggtc cagggaaccg
 123851 ggggtctgcc cggaatcctg cgcacgggc cacctccttg cgtgtttccg
 123901 gcggtatcgg gccgcggtc ggctcaacct tcagcctgga ctccggtaac
 123951 gccggaccgt accaggcaac ccccgagca acaggagtgg gattcggcat

10

20

30

124001 gtcggcagtg gaacgcgcca aggaatgca acccgagttg gagcggctgc
 124051 gccggtcgct gcaccgggag ccggagttgg gccttgcgct gccgcgcacc
 124101 caggagaagg tgctcgcggc gctcgacggg ctgccgttgg agatcacccc
 124151 gggcacccgc ctcacctcgg tgaccgcccgt cctgcgcggc ggccgccccg
 124201 gcggcgcagtg gctgctgcgc ggcgacatgg acgcgctgcc ggtcaccgag
 124251 gagagcggtg tgccgtacgc ctggagatc cccggccgga tgcaacgctg
 124301 cggacacgac ctgcacaccg ccggactcgt cggcgcgggc aggtctgttg
 124351 ccgaacgcgg ggaacgcctg caccgagcgg ttgtgttcat gttccagccc
 124401 ggcgagcagg gccacaacgg ggccggcgcg atgatcgagg agggcggtgt
 124451 ggacgcggcg ggcaagccgc tcgacgcggc gtacgcccgt cactggtcgt
 124501 ccaaccaggt gcccgcgggc gtggtcatca cccgctccgg caccatcacc
 124551 tcggcctcgg acctgctgac ggtgacggtg cggggcgagg gcggccacgg
 124601 ctcgacgccc gccacggcca aggaccgggt cccggcgggc tcgagatgg
 124651 tcaccgcgggt gcagaactgg gtgaccgcgg ccttcgacat cttcgacccg
 124701 gtcgtggtca ccgtcggcac cttccacgcc gggacgaagg cgagtgtcat
 124751 ccccgacaac cgccgagttc caggccacgg tgccgagtta ttccgaatcc
 124801 gcccgggacc ggctcgcgga agggaaactgc cccgtctggt gcgggccatc
 124851 gcggccgccc acggcctgga agcggaggtc gactacctcc ggcactatcc
 124901 ggtcaccgtc aacgacacgg acgagaccgc cttcgcggtg gccaccgccc
 124951 gcgacctgtt cgcccccggc gaggtgtggg agtcgccgat acccaaccaac
 125001 ggctccgagg acttcgcgta cgtcctccgg cgcgtccccg gcgcgatgt
 125051 gctggtgggc gccgcccccg aggggagcga ctggcagcac gcgccgatga
 125101 accactcgcc gcgcgcgctc ttcgacgacc gggtgctgta ccggcaggcg
 125151 gcgctgctga ccgaactcgc cgcgcgccgg ctgcgcgcca ccgcgcggc
 125201 cgagcccgcc gtggcgggat gacaccgccg ccgcacgcgt ggggcccgca
 125251 cgcgtgcggc gtacgggccc tggccggtct ccttcccgcc tacttgccg
 125301 cggggcactc cttccacgcg aagtggtaga ggggtgtgat gctgccgtcg
 125351 gtggagtcac ggccatgtag ctggtggcct tggggtcagc cccgctgccg
 125401 t

10

20

30

SEQ ID NO. 36

NysDII

1 msftypvsmplwlggrelidyv teavgggwis sqgpyvrrfe eafaayndvp
 51 fgvacssgittaltlralrg vpggdevivp eftmiasawa vtytgatpvf
 101 vdcgddlnid vsrieekitp rtkvimpvhi ygrqcdmdav lnlayeynlr

151 vvedsaeahg vrprgdiacf slfankiisa geggvclthd phlaeqmahl
 201 ramafthkdhf flhkklaynf rmtnmqaava laqteqltdi lalrrdiekr
 251 ydeairdipg itlmprrdvl wmydraerr delcaylage gietrvffkp
 301 msrqpgyfsa dwpalnaarl sadgfylpth tgltaqeqef itgrirafyg
 351 va

SEO ID NO. 37

NysI

1 mdneqklrdy lklatadlrr trrrvhkles aagepvaiig mtcrypggvr
 51 spedlwrnve agehgvtpfp tdrqwdleal aaaptasggf lhdapdfdad
 101 ffgispreav amdqqrvvl esaweaferra gidptsvkgs rtgvfigama
 151 qdyrvqpadg aegfqltgnt gsvlsgrisy tfgtvqpavt vdtacssslv
 201 avhlatqair agectlalag gvtimsgpgt fiemgrqggf sadgrcrsfg
 251 dtadgtguae gvgilvlerl sdavrnghei lavvrgtavn qdgasnglta
 301 pngpsqqgvi qqalvnarla agdidvveah gtgttlgdpv eaqallatyg
 351 qnrpadrpil lgsvksnlsh tqaaagvagv ikmvmamrhg tlprrlhaee
 401 pthhvdwsqg avrlldttdt wpatgaprra avssfgisgt nahtieeqap
 451 epqpadaata qddaagstpa tapvvpqgvp vllsgrtpda lrgqaaalra
 501 aldtgrrpdl ldlahtlalt ragfehravl latdhpaltd gltaladadd
 551 paaaawitg ttraetrlav lftgqgaqrl gagrelaarf pafataldaa
 601 ldaftphldr plrevlwgt d aalldrtaya qpalfaveva lyrliesfgv
 651 rpdhlaghsv geivaahlag vlsladaatl vaargrlmq lpdqgamiaav
 701 qaseadvapl laghedqvai aavngpsavv lsgaeatvta laeqlaadgr
 751 ktrrlrvsha fhspmlmepml dafravvedl tlqppllpvv snltgkpatv
 801 aqltsadywv dhvrhavrfa dgidwlarhd ttaflelqpd gvlsamaqdc
 851 ldaadadavt lpalragrpe ehtlittalag lhhvgatldw tgcfagtgar
 901 rtdlptyafg rrrywpkalq sgtadlrsvg lgaahhpils aavsladagg
 951 tlltgrlsrq thpwladhtv rgttllpgta flelavragd evqcdrveel
 1001 tlaaplllpe qggvqvqlwi gnpdvsgrrt vnvharptg ddtptwatah
 1051 gvlttadasr qlpasseqgg tplagdphpa ldaaqwppag aeplpidghy
 1101 driadggfgy gpvfqqlraa wrggdvvyae velpeagrsd aeafglhpal
 1151 ldaalhaapf tglgergrgg lpfswegvsl haggattlr rltpvaddal
 1201 altvadgtga pvlsvdsvl rsvatqqltd aaavardalf rldwtpvqpt
 1251 atdpgpvall gadpfgllth agfadapayp dlaalaaadg pvpttvvlsl
 1301 agtgddaadp arsahrcaae alaavqtwd hherfaaarl vfvtrgatvg

10

20

30

1351 rdvaavaavwg lvrsaqsenp gcfalvdldp dgavgaalv aalvsgepql
 1401 avrgdvlrva rlvrrpltev gagadgtgdg vgdsgvsfs gegavlvtgg
 1451 tgglgavlar hlvaeygvrd lllvsrger avgagelvae lagvgarvr
 1501 vacdvtddraa vvelvgghav savvhaagvl ddgmvgaltg erlsavlrpk
 1551 vdavwhlhea trglldldafv vfsslagvfg spgganyaaa nafldalmtr
 1601 rraeglpqls lawgpweqsg gmtgtltdvd aerlarsgvp plsvaqqlal
 1651 fdaavagtda tcvpvrldlp vlrangevpp llrslirvra rraavagsat
 1701 agnlaqlrr ldedgrdemv ldlvrgqval vlghatggdv dagrafrdlg
 1751 fdsltavelr nrlntvtglr lpatlvfdyp tvrhlatyvl deligtдав
 1801 atvqpaavav addpivivgm acrypggvss pedlwrvtte gtdavsgfpt
 1851 nrgwdvesly hpdphpgts ytrsggflhe agefdpgffg msprealatd
 1901 sqqrlilless weaieragid pvsrlgrstg vfagvmysdy samlaspefe
 1951 gfgsgssps lasgrvaytl glegpavtvd tacssslvam hwamqalrsg
 2001 ecglalaggv tvmtspavfv dfarqrglsp dgrckafada adgvqwsegv
 2051 gvvlvlerqsd avringheila vvrsgavnqd gasngltapn gpsqqrvirg
 2101 alasggltag dvdvveahgt gttlgdpiea gallatygrd reperplllg
 2151 svksnlghtq aaagvagvik mvlamrhgvv prtlhvdaps shvdwsegav
 2201 ellsegaawp etgrvrragv ssfgisgtv hviveqapga kaiaaagaar
 2251 rtpgavpvl sgrgrsalrg qaarllghlq arpdaelvdr alslatrrsr
 2301 feqraavvaq drdqliaslg alaadrpdpa vvegeaargr rtavliftgg
 2351 sqraamgrei hevqpefaaa fdavcavfdp lldrplrevv faedgsdeaa
 2401 lldetgwtqp alfavevalf rlveswgvvp dfvaghsige laaahvagvl
 2451 tledacriva aratlmqalp tggamiaiaq tedeiaahld dtvaiaavng
 2501 pqsvisgde eaaetiaatf aergrkrkr lrvshafhspr mdgmldafri
 2551 vaegltvrap riplvsdltg rraddaevct aeywvrhvre avrfadcvrt
 2601 lrdagattfl elgsdglta maedtlgddh daelvpmira graeelaaat
 2651 alarlvrgv dvdwaaylag tgarrtdlpt yafghayywp qlptpaaala
 2701 aadpadqqlw aavergdare ladilglgeq dltpldsllp altswrrgnq
 2751 ekhlldtlry rvewtrlskp tapvldgtwl lvasdataad qpallldglad
 2801 algshgarvr rlllddscad ravlaerlar tadvdaatqv lsvlplderd
 2851 addcppltrg laltvalvqa ladtgaqgrl wtatrgavst npadpvthpv
 2901 qaaawglgrg valehprlw glvdlpqvfd eragqrlagi lavkdapdge
 2951 dqvalratgv sgrrlvrhtv ealptaaeft atgtvllitgg tgglgæevær
 3001 wlaragaqhl vltstrrgpda pgaaelrael egygpsvsvv acdvadrdal
 3051 aavltalpee lpltgvvhta gvghygpldt lstaefaglt aaklagaahl
 3101 dalladreld ffvlfgsiag vwgsqnqsay gaanayldal alhrrargla
 3151 atsvawgpwa eagmaaddav setlrrqglg lldpapante lrravvrqdv

10

20

30

3201 tvtvadvdwq ryaplftsar psaliaglpe vrslaadert eqdatgasev
 3251 vtrvralaep eqlrltldlv rtesatvlgh ssadavpegr afrdvqfddl
 3301 tavelrkrig aatglsipst mvfdyptple laqylraeil gavlevagpv
 3351 atggaddepi aiiqmacrfp ggvsseqlw dlvasgtdai sefpvnrgwq
 3401 tghlfdpdpd rpgttystgg gflheadefd ptffgispre alvmdpqgrl
 3451 llettwesfe ragirpetlr stltgtfvgs sygeyglgag dgteghmvtg
 3501 sspsvlsgrl syvfglegpa vtvdacsss lvaihlaqcs lrngesnlav
 3551 aggatimttp npfiafsrqr alakdgrcka fsddadgmtl aegvgvvive
 3601 rlsdaqrngh pvlavlrqsa inqdgasngl tapngpsqqr virqalanar
 3651 lapgdidale ahgtgtplgd pieaqalfat ygrdrdpesa lllqsvksni
 3701 ghtqsaagia svikmvmalr hselpttliha dapsshvds agtvrltqa
 3751 rawpetgrpr raavssfgis gtnahvileg apvadtpae rpavapvpia
 3801 agvvpwvvtas rsaalrgqa erllahaetv gtalpaagpl diglslvsar
 3851 arfehravvv ppagtdplaa lravatdgps pvvargvad egtrvfvfpg
 3901 qgsqwvgmgs qlldesavfa eriaecaaal aeftdwsldv vlrqvv gaps
 3951 lervdvvpaa sfavmvslaa lwrsrgvlpd avvghsqgei aaavvsgals
 4001 lrdgarvval rsqaigrala grggmmsval svdvleprlv efegrvsvaa
 4051 vngprsvvva gepealdalh arltaddira rriavdyash shqvedlhee
 4101 llevlaelap rtsevpffst vtgdwldtar mdagywfrnl rgrvrfadav
 4151 adllaaeyra fivevsshpvl smavqeaide agvpavaagt lrrdqqgtdr
 4201 flsaaevfv rgvdvdwagl fegtgasrid lptyafqheh lwavppapea
 4251 vaaadpddaa fwtavedgdv saltaalgt dsvaavlp itswrrarrd
 4301 rstvdawryr vawkplggtl phpsltgtwl lvtadgidt dvagaletyg
 4351 aevrrlvde ecvdravlr rlaaedvtg ivsvlaaaer tdavpgtslv
 4401 lgtaltvali qalgdaeida pvwaltrgav stgradelta pvqaqvtgig
 4451 wtaalehpqr wggtdlpaa ldaraaqrla avlsgalgsd dqlairpsgv
 4501 ftrrivraea tagrpagtwt prgttlvtgg sgtlaphlar wlaqrqaehl
 4551 vlisrrgtaa pgaaelvael aesgteatva acditdrdav aalladlkad
 4601 grtvrtvht aatielhtld attladfdrv lhakvtgaqv laellddeel
 4651 ddfvlysta gmwgsgahaa yvagnaylaa laehrrangl palslswgiw
 4701 addklgrvd pqmirrsgle fmdpqalsg lqralddnen vlavadvdwe
 4751 tyhpvytsgr ptplfdevpe vrrltaaaeg sagtvaegef aaalralsda
 4801 eqqrtilletv rteaasvlgl ssaedltdqr afrdvqfddl tavglrnrla
 4851 svtgltlpst mvfdypnpaa laaylhgela garsaaagaa avptgapdad
 4901 dpiaivgmsc rypggvgsae dlwrialdev daisgfpadr gwdaeglydp
 4951 dpdrpgrtys vqggflrdva efdpgffgis prealsmdp qrlilletawe
 5001 afehagidpv gqrgsrtgtf vgasyqdyas gvpnsegseg hmitgtlssv

10

20

30

5051 lsgrvsylfg fegpavtldt acssslvamh lacqslrnge sslalaggvs
 5101 imstpmfsfvgrsrqralaed grckayadga dgmtlaegvg lvllerlsda
 5151 ranghqvlav irgsævnqdg asngltapng psqqrvirqa lansavapgd
 5201 idvleghgtg talgdpieaq allatyggdr aperplllgs vksnightqm
 5251 æsgvasvikl vralqegvvp kslhidrpst hvdwssgaig lltertpwpe
 5301 tgrprraavs sfgisgtnvh tileqapade aptpadpprd glvpvllsgr
 5351 geaalraqaa rllafveerp eahltdlæhs latstraaler raaviaadrd
 5401 tltrglrals dgrpdpglvq gtagrgrtaf lftgqgsgrp gmgrelhdry
 5451 pvfadaldev larlddgpdr plrevlfæap dsaeaalldr tgyaqpalfa
 5501 vevalfrllt swqltpdyla ghsvgelaaa hvagvlsldd actlvaargr
 5551 lmqalpegga mvaleaaede vlpillegltd rvsvaavngp rsvvvagvee
 5601 dvllladlfa adgrtrkrir vshafhsplm damlddfaav argltyhppt
 5651 ipfvsnvsgg lataeqvrtv dywvghvraa vrfadgidwl atqgdvhtfl
 5701 elgpdgvlsa maresltdps rtallptlrg drpeepalvt avaaahahga
 5751 rvdwsgyfad hgarrttlpt yafqrerywp dttaatsaht pgsaldaefw
 5801 aaverddvaa laasldldda tvtamvpalt awrrrrgeqt eldswryrvt
 5851 wkprggatap aaltgrwlv1 vphdhqdrqd dataawæadv etalgttttr
 5901 ltvtttdraa laariteaag dggpfsgvls llplatgdag hpgapaaltl
 5951 ttavqalgd agidaplwnv trgavavgra eqvtapeqaa vwglgraval
 6001 elparfggtl dlpatlbgga arrlravlaa tdgedavalr psgvflrrla
 6051 hapagpdtar tafdpagtv litgggtggig ghvarrlard gathllltsr
 6101 rqpapagada lraeleelga rvtlaacdaa drdalaalla elpddaplca
 6151 vfhtagvved hvvdaltpen faavlraktv aahlheltha dldlaafvlf
 6201 sstagvlgaa gggnyaaana hldalaehrr shgltalsva wgpwagsgmv
 6251 adaaeltdrv rrggfeplap epavrallra ienddtval adidwerfqr
 6301 afaavrplpf vadlpetgra tpatatgaat glrqqlaelp eherpaavld
 6351 llrtqvaavl ghadprtved dhafrdlgfd sltilelrna lnaatglsip
 6401 atlvydlptp remadfilæe llgtlptdta atvastaspk lsasfeqgg
 6451 pfddpiavig igrfpggvt tpeelwqlld egrdgisrpf ddrqwdlaal
 6501 gagasdtleg gfltgvadfd arffgispre alæmdpqqrl llettweale
 6551 ragidpttlr gsttgfvgt ngqdyptllr rsasdvagyv atgntasvms
 6601 grlsyalgle gpavtidtac ssslvalhwa gralragecd lvvaggvsvm
 6651 aspdsfvefs tqgglapdgr ckafsdaadg tawsegvgil vlerlsaarr
 6701 nghqvlglir gtavnqdgas ngltapngls qqrviaqala darlrpadid
 6751 aieahgtggt lqdpiearal itaygrdrda erplllgtvk snightqaaa
 6801 gaagvikmlm amrhgtlprt lhvgtpsshv dwsqgtvall ddarpwpptg
 6851 qprragvsaf gvsqgnahvv vegapeteap aapaaepape atptvvpwv

10

20

30

6901 sgrsrealqa qldrltahta ahparsaadv grslatdrtl fphravllag
 6951 pdgvreaara aaprtpgrrta flfsgqgaqh almghdlyqr fpvyadaldt
 7001 vlaqfdtvid vplraalfaa pgtpeaalld qtgftqpalf avevalfrla
 7051 eswrltpdfv aghsigeiaa ahvagvfsle dactlvaara slmqqlprdg
 7101 amvaleated evaplltdgv alaavngprs vvvagaedav ravadrilaad
 7151 grtrrrltvs hafhsplmdp mitdfarvae gltyheprip lvstllgapa
 7201 gaelrtpdyw vrvretvrf adgvralhda gagtfveigp dgvlalttqq
 7251 tldtveagap avvvplqrrd ragdlalleg latlhthgtg pswpayfeat
 7301 gghrtdlpty afqrerywpe lgapvatapq dpaawryhet waplapeaa
 7351 apagralvly pagnrdtawm tavadalgad tvtaepdala eqltaagdtp
 7401 wrvvvsllaa aseglpadga wpaallatld eagvhaplwc vtrgavavag
 7451 eaptavgqaa lwglgrvaal dhpdrfggla dlpadtdaha agllaahlaa
 7501 pgteaaiavr atgvharrlv rtpaaadgat wlptgtvlv ggtggtgtmg
 7551 graarwlvre garhlvltap dgtttaadte altaelaalg aritvvdhdp
 7601 tapdgfaall dglpddtplt avvyapeada apgtaaela alapvtalga
 7651 altqrpldaf vlfqsiaglw gvrgraaeee sgayldafar acrdrgtpal
 7701 avawgawadl vqpslaahlr mnglpvmdad taltalsrav adgsaaeava
 7751 dvrwetfapl bhearrtal f dalpeargal aeaardradr ktaagdygrw
 7801 laeqpaadhd aillalvtek aatvlghadh dllepdlpfr dlgsdsltav
 7851 dlrnqltaet gltlpatlvf dhpnpaala hiraqligea sdsapvaap
 7901 valgaddai vivgmacryp ggvtspeclw qlvgdevdav gdfptdrgwd
 7951 laalagdpgp rsatqggfl ydatdfdpql fgisprealv mdpqqrille
 8001 tswealerag idpatlrqsg ttgvfvvggs gdyrppeeag qwqtaqsasl
 8051 lsgrlaytfg iggptvsdvt acssslvalh laaqalrage csialaggvt
 8101 vmatpvgfve fsaqqalspd gcrarfsdda ngtgwseqvg mlvverlsda
 8151 rrrghrvlav lrgsainqdg asngltapsg paqqrvirqa lanarlrpad
 8201 idaveahgtg trlgdpieaq allatyggdr erpvllgslk snightqaas
 8251 gvggvikmvl amqhgelpsr lyenpsshv dwtagrahl1 tartpwpdsg
 8301 rprraavssf gasgtnahai leqppreelp arpaddgapl pflsgrsqn
 8351 alraqerrll arltahpdtr aadlayscale traafehraa itatdhgdlr
 8401 tgltavaeat taphtaehhl qgtgkravlf sqqgsqrlgm grelherhpv
 8451 faeafdsvla rlddrldtpl rdvvwgtdee alhatgntqp alfavealy
 8501 rlieswgvvp dfvaghsvge laaahvagvl siddacriva araalmqrlp
 8551 aggamiaeva tedevtplt dgvsllaavng ptavvlsgag davtalqqal
 8601 aerghrttrl rvshafhshl mdpmladfrt vaegleyhpp ripvvsnlgt
 8651 dvadaedlcs adyvvrvhrg tvrfadgvrt madrgvhlfl elgpdavlsa
 8701 marqcapdav vvpalrrnrd edetlvgava rlvvhgagpr wdayfagrga

10

20

30

8751 qwldlptypf qrgrfwpesl pgaasaapaa gqpaetdaaf wdavaqedft
 8801 alesvldves dalskvlpal mdwrsrqade sqlagwrhri vwkrltgaal
 8851 ahrkalsgtw lavvpegfad dpwvtttlôg lgthlvhlev aeadraalad
 8901 aiaartadgt rfggvislla lreeltgavp egtaltttll qalgdagvda
 8951 plwcvtrsav sagrtdrphr plggavwglg rvaaleypqr wgglvdlpee
 9001 pdersaagla avlagldged qvavrgtavl arrlvpapqr kpsrpwhpsg
 9051 tvlvtggtga lgahvarria kdgaqhlvll srrgpdapga aelraeldal
 9101 gtdvtvaacd vadrdqltav ldalpadrpl tgvvhtagvl ddgvidrltp
 9151 erfqevfrak vtsallldel trdrelaafv lfssasaavg npgganyaaa
 9201 navldalaeq rrvlqlpats vswgawgggg madadgadea arragvgamd
 9251 phlaveallr lvaekptav vaevaldrfa gafggsrpsa llrefpgyre
 9301 alaaqaeqaa dggglaarla alpparrldt vvdlvrtara qvlgyptea
 9351 vaaersfrdl gvdsigavel rnqlsaatgl nlpatlrvfdh ptplvlgehi
 9401 lgglfpdapa gsddeteira llasvpldql reigvlepll qlagrgraa
 9451 dgddgesvds mtvadlvraa lngqsd1

10

SEQ ID NO. 38

NysJ

20

1 mnapienpetp ennvvaalra avketdrllr qnrmvaaaak epiavvgmac
 51 rfpqavdspe alwemvatgt dvisgfpddr gwdlealrns gtdardtdvs
 101 qrggfldcia dfdpqffgis preavtmdpq qrlilttawe averagidat
 151 tlratrqtgaf igtngqdyay llvrslddat gdvgtgiaas aasgrlsytl
 201 glegpaltvd tacssslval hlavqalrng ecgmalaggv nvmatpgslv
 251 efsrqgglar dgrckafada adgtgwsega qvlllerlsd agrnghpvla
 301 vvrqsavnqd gasngftapn gpsqqrvirq alanaglatg didaveahgt
 351 gtpigdpiea qsilatyqqd rahpvllgsi ksnmghtqaa sgvagvikmi
 401 mamrhgvlpr tlhvdrpsth vdwtggsvel ltdahwpet grprtgiss
 451 fgvsqtnahv ivegapdtpa eaaddtpprt prtlpwllsa rtgaalrdqa
 501 talldhldrp dgdrqptald tafslattra alehrlavvt gtdgtagrda
 551 ltawlahgta pdaheghaag rtrcaalfsg qgaqrlgmgr elharfpvfa
 601 raldtavdll daelggtlre viwgtddapl netgftqpal favevalyrl
 651 ieswgvapdf vaghsigeia aahvagvfl edactlvaar aglmqalprg
 701 gamvaveate devsplltdg vaiaaingpt slvsgdeta tlavaarlae
 751 qgrtttrlv shafhsplmd pmlaefrava eglsygepqi pvvsnltagv
 801 adgtllgtad ywrvhvreav rfadgiralt dagvgaflel gpdgtlaala

30

851 qqsapdavsv pvlrkdrdee paavaalarl htagvpdwt afyagtghr
 901 tdlptyafqy erylwpkatyr padatglglt aadhpllgaa msvagsdell
 951 ltgtlslath pwladhvvgg mvffpgtgfl elavraadqv gcdrveelml
 1001 aapliipatg tvqmqlavga adddggrdlr ffrpgddpd aawaqhatgr
 1051 itegervlal dtttwpprda epvldidglyd ryrangldyg pvfrglraww
 1101 rrdteiyaeav alpegtadad afglhpald avlhstlfas adgddrsllp
 1151 fawngvslha agadalrvri tscgpdavei tavdpqgrpv vsvesltira
 1201 agpdagtagh radagslfrm dwtprtvhap atpatwavlg tdpiglteal
 1251 taagpdtvtg lrdqvdalge ltagddrppv dvvavplrga tdhgpagahd
 1301 ltrtvlallq ewlaeerfar srlilvtrga vadgergpld laaapvwglv
 1351 rsaqsenpqr lllvldldta esaaqlpllp alldadepqa vvregtvrvg
 1401 rlarldsgrg lvpppgtpwr lgsrakgsid glallphpea rrpltghevr
 1451 vgiraaglnf rdvlnalmy pgdaglfgse aagvvvevgp evtglapqdr
 1501 vmgmifggfg plgiadarll tpvpadwswe tgasvplvfl tayyalkeig
 1551 glragekvlv hagaggvgma aiqlarhvg evfatasegk wdvrlslgva
 1601 ddhiassrtl dfeaafeava gdrqldvvl nalsgefvdas mrlldgggrf
 1651 lengktldira adsvpdglis hsfldgmvd ehiqrmlldl velfdrgala
 1701 alpvrswdvr rageafrfms laqhigkivl tvppldpdg tvlltggtg
 1751 lagllarhlv tehgarhlil agrrgpdapg aaalhaelta lgaevtvaac
 1801 dvadrtalaa llatvpaehp ltavvhtagv lddgtltaln pdrlatvlrp
 1851 kvdaawhlhd ltrhldlaaf vlysstagvm gpgganyaa gntfidalaa
 1901 hrhalglpat slawgawegg agmtgalt dhllrvsdagg qpltaergl
 1951 alydaataad eplivplgt ggalpaggv pavlrglvrt agrraragta
 2001 gvsraglaer laalpeeert pflvelvrte aatvlghqst dpvdarreft
 2051 qlgfdsltai elnrnlgkat gltlpatlif dyptpdlav hlhdellgad
 2101 apvtvtaaaq aadpehdpv ivgmscrfgp gvsspeelwd lvasgtdait
 2151 gfpadrawdr hpqlagapga rtggggflrd iadfdaffg isprealamd
 2201 pqrilleva weaaeragid pgtlrgsdg vfmgvsgqdy aglvmsrdd
 2251 iaghattgla vsvvsgrlay alglegpals vdtacssslv slhlaaqalr
 2301 agectmalag gvtvmttaen ftgfsrmggl aqdgckafs dsadgtgwe
 2351 gaavlvlerl sdarraghrv lavvrgsavn qdgasnglta pngpaqqrvi
 2401 rgalanaglt pvdvdaveah gtgtplgdp eaqaliaayg tdrdpheppl
 2451 lgsvksnigh tgsaagaagl vkmvmamrhg ilpqtllhte psshvdsag
 2501 tvrllterta wprtdrpra gvssfgisgt nahvileqpp aeptpaadpg
 2551 rpaptvvawp vsaqtpaald aqldrirtaa alapldtaht latgrslfeh
 2601 ravllatvgd patgapdlpe vargaatphr tafllsgqga qrsgmgrelh
 2651 aafpvfaaaf devvavldae lgsdadggvs lrevmwgggs elldtrftq

10

20

30

2701 palfaveval frlvaswvgv pefvaghsvq eiaaahvagv fslvdacriv
 2751 varaslmdal pvvgvmvave aaeavvpll vdgvaiaavn gpvsvvsvgv
 2801 eaavggvvdq lvergrrvr lavshafhsp lmdpmdafr avaegeleyhq
 2851 pripvvsnt gevaaaaelc aadyvvrhvr atvrfadgvr tlaergataf
 2901 leigpdgvls alargvlpae alvtptlrkd rdeesallag larlhvagvt
 2951 vdwsaaltgt gargtdlpty afqrerywpe laaepagga daadaefwaa
 3001 veradataala ahldidgdql gavlpalsaw rtrrrtsat nalrhreswe
 3051 plslagtpht ggvlvlpaa attdpwvad vaaigpdarr vdvpadgtdr
 3101 aalaalltea addtaptavv slaldetsg ddavpagtta taalvqalad
 3151 tgapaplwai trgavaalpd eqptapaqaa vwglgriaal elprhwgglv
 3201 dlpadldert arrlpaalad agdedqlair atgaygrrit papapddapg
 3251 tgwgptgtvl itgggtgalgr htarwlaehg aebllllsrs gpdapgaael
 3301 tteltalgar vtlvacdaad regltrviae vprdcpltg vhtagvlldg
 3351 vltgltprdf atvfrakvas avlldeltrd tdlavfalfs svagavgnpg
 3401 qagyaaanav ldalaarrra qglagtsiaw gawagdgmaw rhtrpgaepv
 3451 glldpdlavp alaravtepq ptlvladlqg prllesllal rpsllsrlp
 3501 aartaaravq eadrrragaa adlrqqlagt apadrhavll rlvrttaaav
 3551 lghtgadair adkpfrdlgf dsltavelss alaaatglal ppslvfdhps
 3601 praladhira eltqdrpesa paappapvpa adddpivvvg macrfpggvt
 3651 tpeefwqla egrdgidafp tdrqwdldvl grrrpggqrp prsaassyda
 3701 aafdpqffdi sprealamdp qqrllletaw eavertgtdp trlrgsrtgv
 3751 fvgtnngdya glvlraqedv eghagtglaa svisgrlaya fgfegpavtv
 3801 dtacssslva lhwavqalra gecslalagg vtmmtstsf agftrqggla
 3851 pdghckafsd sadgtgweg vgvlvverrs dalrngheil avvrgsavng
 3901 dgasngltap ngpaqqrvir qalenaglap gdvDaveahg tgtvlqdpie
 3951 aqallatygg drpadrplwl gsvksnight qaaagaaglm kmvlalqhg
 4001 lprtlhvtep strvdwsaga vrlltertvw prtdrprrag vssfgisgtn
 4051 ahvileqppa eptptapadr ptrtpavlpw vvsarsatal daqlarlraf
 4101 aaerpdippa dvahslvtsr atfehavrll aapdgitaaa raeaxersta
 4151 flfsgggagr sgmgrelhaa fpvfaafde vvavldaala tsggggvsir
 4201 evmwgggsel ldrtrftqpa lfavevalfr lvaswvggpe fvaghsvgei
 4251 aaayvagvfs lvdacrivva raslmdalpv ggvmvaveaa eaevvpllv
 4301 gvaiaavngp vsvvvsgvea avgqvvdqlv ergrrvrlla vshafhspim
 4351 dpmdafrav aegleyhqpr ipvvsntge vaaaaelcaa dyvvrhvrat
 4401 vrfadgvrtl aergatafle igpdgvlsal aaacldftda evvpalrkgr
 4451 peehtaltaa aqlhvagvdi dwtavlagtg grrialptya fqrerywpsl
 4501 aaqapgdagg lgleagrhlpl lgaattvags aeilltgrls ttaqpwlavy

10

20

30

4551 eadgrtvlpav avlaelavra gdqadcptva eltvaaplvl tgaaaqrlqv
 4601 rvaapddtgr ralsvharpd dspdspwtlh atavltthdtp qppapdtgwp
 4651 peravpldal ptatgparia aawqwgdelc aeieipepgp aerafalhpa
 4701 lldtavragg lldqdatlga lgwrglalha asatalrvrl tpdqtdtwal
 4751 eatdpqgapv vsvgtltlgt ptvdrsgaga addgatlldl ewvpapqaap
 4801 tggdhlpyav lqdqlaeldg qlriagdgpg rvaslaalld ggaplprlvi
 4851 spvlgvptge gdlpaavrgt ttavlellqr wtadartads hlviivtrgav
 4901 aagaedvhdv aaapvwglvr sagsehpgsf lldldpadp agasraaapa
 4951 tlaalldage tqaavradtl tvarltraad gpeataghpv rdwdrdgtvl
 5001 itggtgglgg llarhlvtgh gikhlllagr rgpdapgara lrdelaalga
 5051 evtvaacdva draaldrlla qlppehpita vvhtagvldd atvglttper
 5101 ldtvlrakad aawhlhdatr drdlagfvly ssvagvtggp gggnyaagnt
 5151 fldalaahra aqglpglsla wgpwgqdagm tgtlgaadla rlersgmpl
 5201 tpeqglalfd aagargdgfa vavrlargaa apgadevpav lralvgrrr
 5251 taaaaghagv larllaalda eqrhqalldl vrtetaavlg hsgadavpae
 5301 rdnrlgfsd lmavelrtrl atatgarlpa tlvdhptpd avarhlastl
 5351 pggtaagpdr splaeldria aelspegadd atrqgvvgrl rhllaqwdgt
 5401 rqdgggttvd drieaasae vlafidhelg rqads

10

SEQ ID NO. 39

20

NysK

1 mpdekkivdy lkwtkdihq trqrlqevea grhepvaivg macrfpggvr
 51 spedlwells agrdgigpfp adrgwdlaal agdgpgrsat qeggflpdaa
 101 afdpgffdis prealamdpq qrlilletawe aversgidpa qlrgsrtgvf
 151 vgtnggdyah lvlaaqddmg gyagnlaas vlsgrlafai qlegpavtld
 201 tacssslvtl hlaaqavrag ecglalaggv tvmttsssf qfslqgglap
 251 dgrckafaea adgtgwsegi glilverlsd aqrnghpvla vlrgsavnqd
 301 gasnglsapn gpsqqrvirq alagaglvpg dvdaveahgt qtrlqdpiea
 351 gallatyggd rpadrplwlg svksnlghtq aaagvagvik mvlalrhgvl
 401 pgtlhvdaps shvdwesgav rlltapvaws egddrvrrag vssfgisgtn
 451 ahvileqapd qpeptaeeta aaapggtaee raaapvapra vpwpvaarta
 501 galdaqlvrv ralttapgrt aadvghalat artpfehral lvheggavte
 551 vargavptgd rgglavlfsg qgsqrpgmgr elharypvfa aafdetvall
 601 darlgtslrd iwdqdrtrl ddtrhtqpal favevalyrl laswgirpdh
 651 vtghsigeit aahvagvltl adactlvaar atamselppg gamvaleate

30

701 devrplltdd laiaavnapr svvagaeda alavrrhfd d lgrtttrlpv
 751 shafhsplmd pmldafrtal apltfaepei pvvsnlgtlp ataeelatph
 801 ywvchvrqav rfgdgvrata drgvrtflel gpdgvlsalv renlpepglv
 851 avpvlrkerp eettvlaalg tlwahgadvd wdvafagtrt pqadpvelpt
 901 ysfqrarywp tlgarhgdp a dlqgtaahp llgaavtlad adetvltgrl
 951 alpshpwlgd hrsdgritvp gvafaelavr agdlsgtphl arldipaplt
 1001 lqgdgtvtlg vrvgapdpag hrpltvharl aatedapwtt catgllapda
 1051 peapadpigg adagwpprda rvpvpadlda aataagrhyg phfqgltglw
 1101 rrdgevfav alptataadr afgihpalla talrataald ddhtaghtpe
 1151 ptgitglalh atgatalrvr ltatgpdttva laaadatgga vltadtvtlg
 1201 spqdrpapap aghtgggglf hldwvpvdp g sratgtrwav vgddeldlgy
 1251 alhradetvs ayaaslggai gdsqglapdvf lvpvvggpd a gpdavhavta
 1301 ralglqewl neprlagarl vfvtrgavav pgetvtdpag aavwgllrsa
 1351 qtenpgslll vlddafrsa gmlphvltld eqqlvvrda vraarlarlp
 1401 epaagtapar awdpdgtvli tgggtgglaa larhlvtvrg arhillagrr
 1451 gpeapgagel vaeltaaggad vrvaacdvgd rtaldallat vpaahpltav
 1501 vhtagvldda ligsltpdql atvlrpkada awhlhdatrg ldlagfvlys
 1551 svsgvlgsqg qgnyaanay ldalarhrad qglpalslaw gpwgrgsgmt
 1601 asvsdadler margglpplt vedglalfda avgrpepalv psrinvaglr
 1651 dqgalpalwr dlvrarrrta atadrspvtv rerlrhldet gqeqlldlv
 1701 vgytagllgh pdptavdper gflgelgfdsl vsvglrnqla eilglrlpss
 1751 ivfdskspvk larwlhqela ngpqp gatgp aaadarpavr sddtleglfy
 1801 navrggkive amrmlkavan trpmfdtpae leelsepvtl adgprprli
 1851 fvsapgatgg vhyqariaah frqsrhvsal plmgfaggel lpatseaaar
 1901 ivaesvlmas egepfvmvgh stggslyla agvledtwdv rpeavvldt
 1951 asirynpgeg ndldrttrfy ladidpsvt lnsarmsama hwmamtdiq
 2001 apaptaptll vraaraldgf rldtssvpad evrddidadh slakehsalt
 2051 aqaiegwlae lpdpa

10

20

30

SEO ID NO. 40**NysL**

1 mstptappsl kaevppvlrl splirelqsr apvckvrtpa gdegwlvtrh
 51 telkqllhdd rlarahadpa napryvhn pf ldilvddfd lartilhaemr
 101 siftpqfsar rvmldtprve alaegvlahf vaqggpadlh ndfslpfsls
 151 vlcaligvpa eeqgkliaal tklgelddpa rvqeggdelf gllsglarrk

201 ritpeddvis rlcikvpsde rigpiasgll fagldsvash idlgtvlfliq
 251 hpdqlaaala deklmrgave eilrsakagg svlpnyatad vpigdvтира
 301 gdlvllldftl vnfdrtvfde pelfdirrap nphltfghgm whcigaplar
 351 vnrltaytil ftrlpglrlv rpveelrvls gqlsagltel pvtw

SEQ ID NO. 41**NysM**

1 vritvdpgrc vgagqcvtla pdlfdqddg lvtvlagaad aadpgdvrd
 51 aalcpsgais vaad

10

SEQ ID NO. 42**NysN**

1 msteadarta apqcpvafpl rrpgrpfppp eyatyrggag lvrsepsgp
 51 vwlvrhedv ravltdpris adpskpgfipk agrtggapsq yevpgwfvam
 101 dppehgrfrk tlipeftvrk vrelrpviqq ivderidaml aagtsadlve
 151 sfalpvpslv issllgvpkv drdffedrtr vlvrilstde erdkatqall
 201 rylgrliqik qrrpgddlis rliaagtlr qelsgvamll liaghetan
 251 niglgvvqll tnprwigddr iveellryys vadlvafva vedveiggql
 301 iragegivpl iaaanhdata faapsefdpe rsarshvafg ygvhqclgqn
 351 lvreemdiay rtlfaripsi tlavpveelp lkydgvlfgl helpvtwk

20

SEQ ID No. 43:

Nys RIV (long) 266 aa

30

VTITHLTDSNHRTTGGVISAQTAPAGESVGPGLMASLDRDLTIKHANQEFRRRFD
 DSAGDVCGRSFRDLMHPSVQQPLMRQFSRLIEGKRHRFASHVVAVGAQDAAFAGT
 LTASAVTGKTPDIAGILVLMDSGAADAADAGVVTSQKKFLTEIDARILEGIAAG
 LSTIPLASRLYLSRQGVEYHVTGLLRKLRVPMRAALVSRAYSMGILNVGTWPPKV
 VDDFIK

SEQ ID NO. 44:

GCC GGC ATG CGA CGA ACA GGA CGA GAG GT

SEQ ID NO. 45:

GCC GTG GTC GAC GAA GGC

40

SEQ ID NO. 46:

CTC AGC ATG CCG AAA GGA TGG CGG

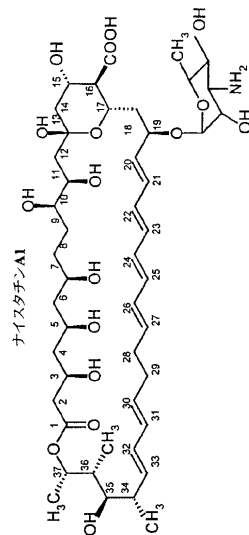
SEQ ID NO. 47:

AGG CAA GCT TCG GCG ACA CGG GCG T

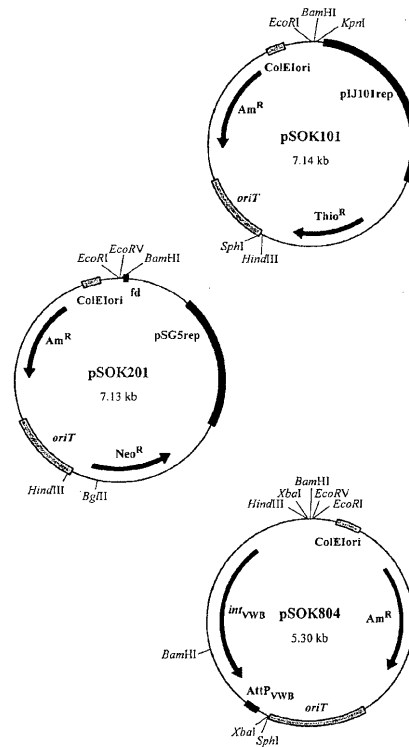
SEQ ID NO. 48:

CTC AGC ATG CGT ACG ACC GGC GGG

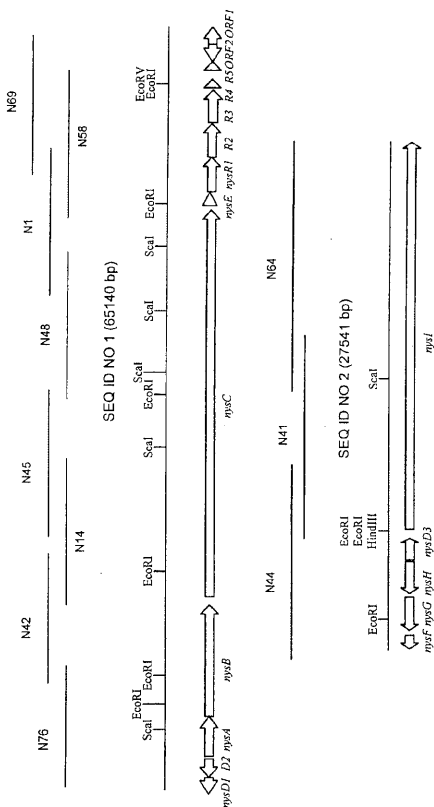
【図 1】



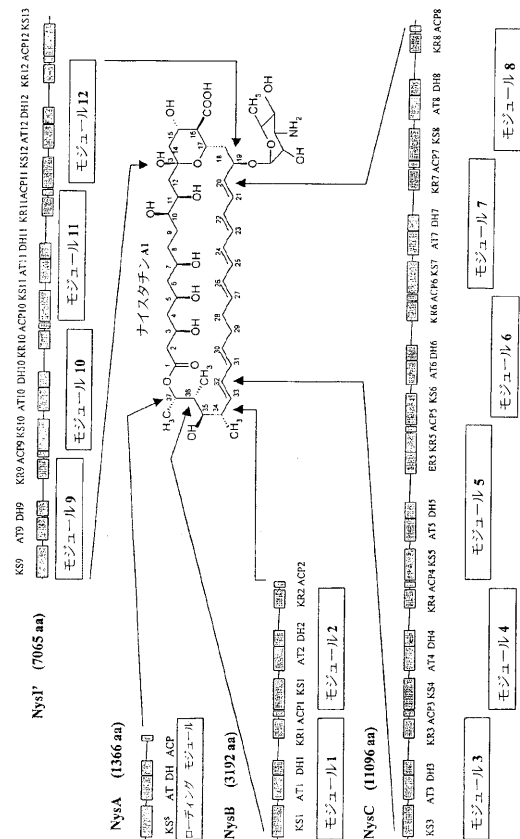
【図 2】



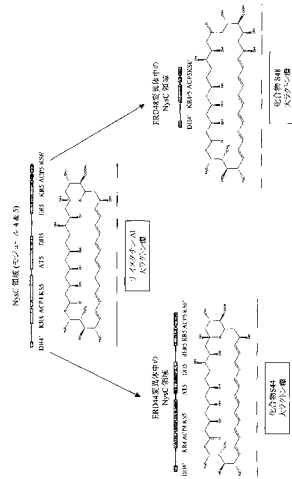
【図 3】



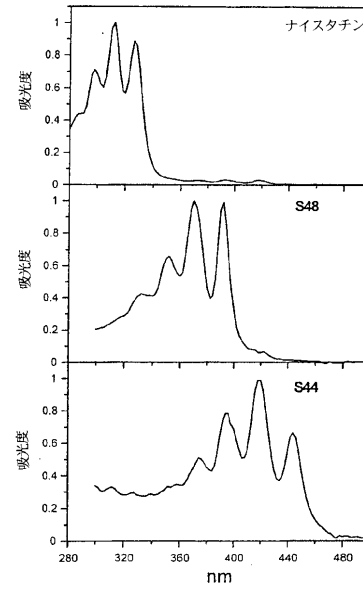
【図 4】



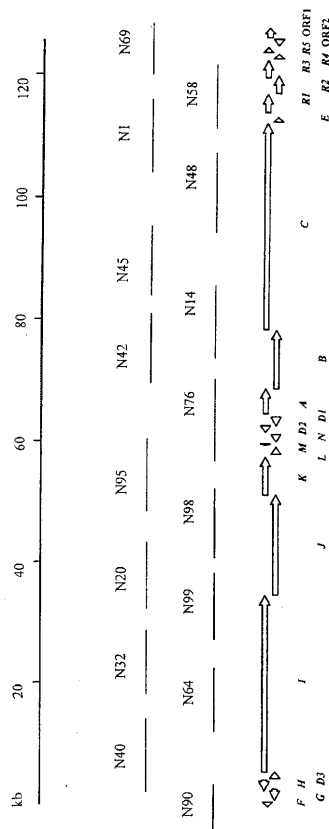
【図 5】



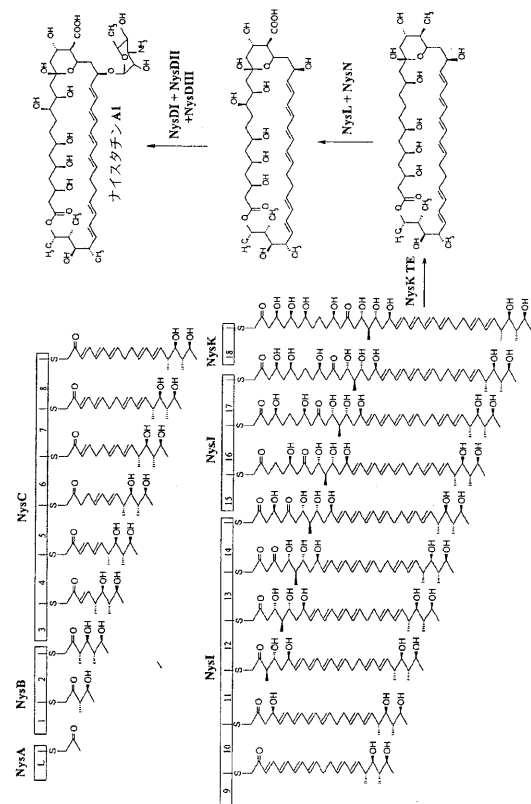
【図 6】



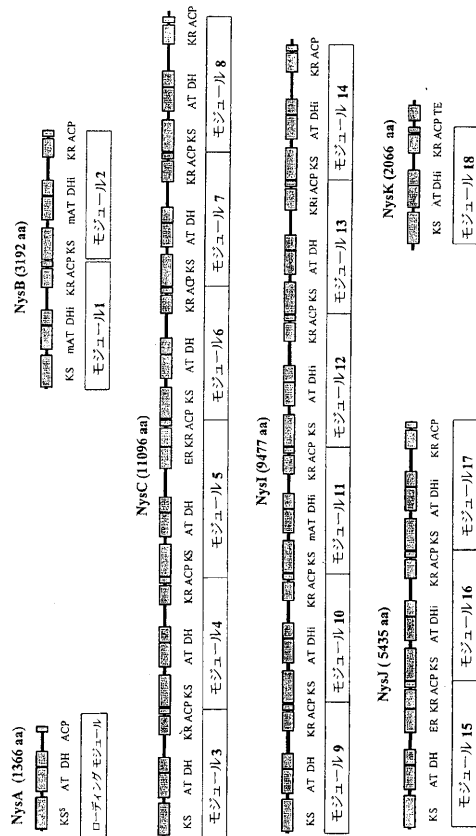
【図 7】



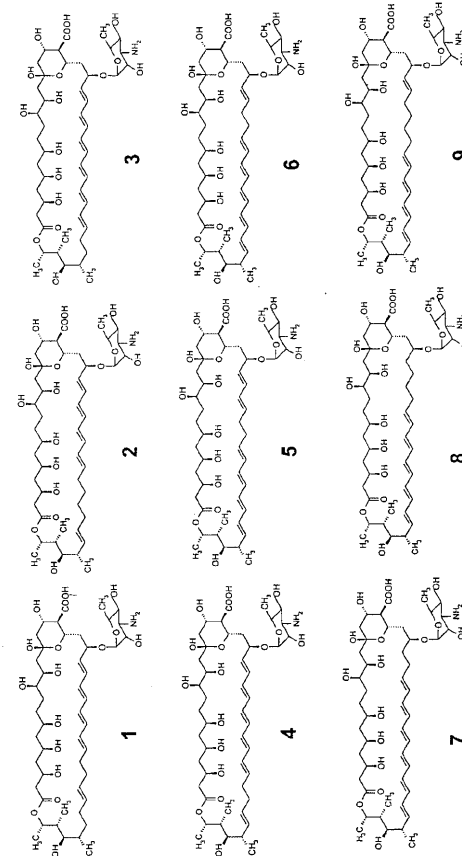
【図 8】



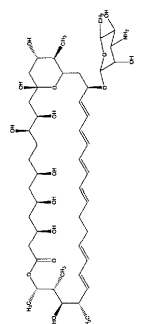
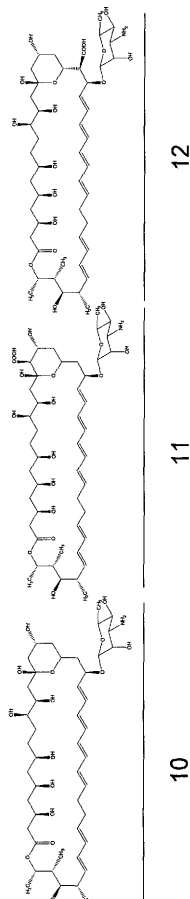
【図 9】



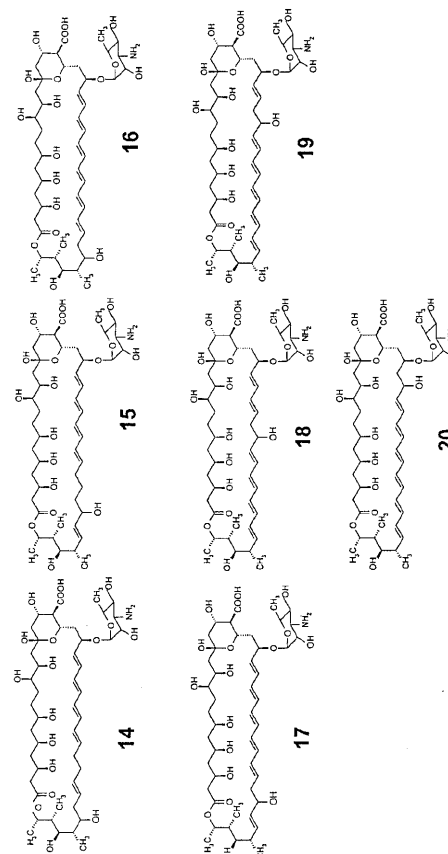
【図 10】



【図 11】



【図 12】



フロントページの続き

(31)優先権主張番号 0009387.2

(32)優先日 平成12年4月14日(2000.4.14)

(33)優先権主張国 英国(GB)

(72)発明者 セクロヴァ, オルガ ニコライブナ

ノールウェー エヌ-7030 トロンハイム エルゲスター ゲート 51

(72)発明者 フヤルヴィク, エスベン

ノールウェー エヌ-7025 トロンハイム リアンファイエン 10

(72)発明者 ブラウタセット, トライヴェ

ノールウェー エヌ-7027 トロンハイム マーティン ストッケンス ファイ 74

(72)発明者 ストレム, アーネ, ライダー

ノールウェー エヌ-7050 トロンハイム コンヴァルヴェゲン 38

(72)発明者 ヴァラ, スヴァイン

ノールウェー エヌ-7560 ヴィクハマー ナウスタンファイエン 2

(72)発明者 エリングセン, トロンド, アーリング

ノールウェー エヌ-7054 ランハイム マーカップラッセン 273

(72)発明者 スレッタ, ホヴァルド

ノールウェー エヌ-7024 トロンハイム ウグラヴァイエン 19

(72)発明者 グリクセン, オレーマーティン

ノールウェー エヌ-0595 オスロ グレブリンゲン. 31 ディー

審査官 幸田 俊希

- (56)参考文献 BRAUTASET, T. et al., Biosynthesis of the polyene antifungal antibiotic nystatin in *Streptomyces noursei* ATCC 11455: analysis of the gene cluster and deduction of the biosynthetic pathway., *Chem. Biol.*, 2000年 5月23日, Vol.7, No.6, pp.395-403
ZOTCHEV, S. et al., Identification of a gene cluster for antibacterial polyketide-derived antibiotic biosynthesis in the nystatin producer *Streptomyces noursei* ATCC 11455., *Microbiology*, 2000年 3月, Vol.146, Pt3, pp.611-9

(58)調査した分野(Int.Cl., DB名)

C12N 15/09

GenBank/EMBL/DBJ/GeneSeq

SwissProt/PIR/GeneSeq

PubMed

CA/BIOSIS/MEDLINE(STN)