



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0035346
(43) 공개일자 2017년03월30일

- (51) 국제특허분류(Int. Cl.)
C12N 15/74 (2006.01) C12N 9/00 (2006.01)
C12P 17/16 (2006.01) C12R 1/465 (2006.01)
- (52) CPC특허분류
C12N 15/74 (2013.01)
C12N 9/93 (2013.01)
- (21) 출원번호 10-2016-0121760
(22) 출원일자 2016년09월22일
심사청구일자 2016년09월22일
- (30) 우선권주장
1020150133892 2015년09월22일 대한민국(KR)

- (71) 출원인
주식회사 팜한농
서울특별시 영등포구 여의대로 24(여의도동, 전경련회관)
- 이화여자대학교 산학협력단
서울특별시 서대문구 이화여대길 52 (대현동, 이화여자대학교)
- (72) 발명자
남상집
경기도 광명시 하안로 320, 1012동 607호(하안동, 고층주공아파트)
- 윤여준
서울특별시 강남구 압구정로 151, 126동 202호(압구정동, 현대아파트)
- (74) 대리인
유미특허법인

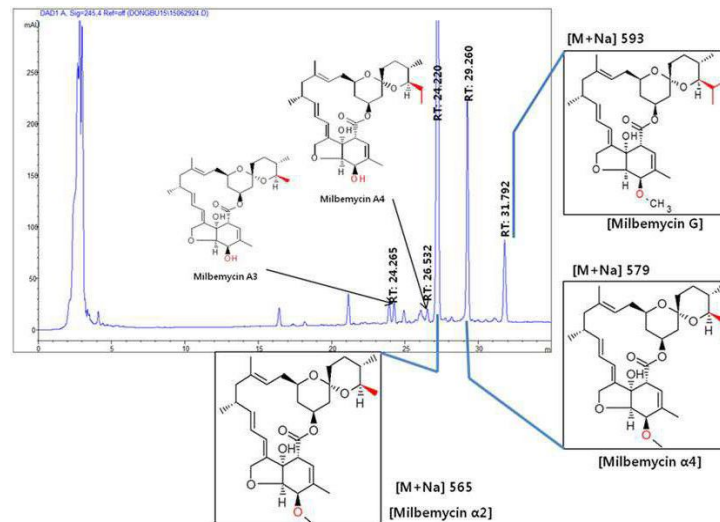
전체 청구항 수 : 총 19 항

(54) 발명의 명칭 밀베마이신을 생산하는 재조합 미생물 및 이를 이용한 밀베마이신 생산 방법

(57) 요약

밀베마이신을 생산하는 재조합 스트렙토마이세스 아베르미틸리스(*Streptomyces avermitilis*) 균주 및 이를 이용한 밀베마이신 생산 방법이 제공된다.

대표도



(52) CPC특허분류

C12P 17/162 (2013.01)

C12R 1/465 (2013.01)

(72) 발명자

유영지

서울특별시 서대문구 독립문로 10, 105동 306호(영천동, 삼호아파트)

김면수

서울특별시 서대문구 이화여대길 50-12, 106동 203호(대현동, 럭키아파트)

오기훈

대전광역시 서구 도안동로 77, 1803동 3304호(가수원동, 도안18단지 린 풀하우스)

정봉진

경기도 과천시 관문로 143, 1102동 804호(중앙동, 래미안에코팰리스아파트)

조완제

대전광역시 서구 도안동로 77, 1803동 2102호(가수원동, 도안18단지 린 풀하우스)

송남숙

대전광역시 서구 청사로 282, 16동 1005호(둔산동, 수정타운아파트)

소은수

대전광역시 서구 도안동로 77, 1805동 2904호(가수원동, 도안18단지 린 풀하우스)

박성환

대전광역시 유성구 유성대로 1741, 107동 501호(전민동, 세종아파트)

김가은

대전광역시 서구 도안동로 77, 1803동 2102호(가수원동, 도안18단지 린 풀하우스)

전희선

대전광역시 서구 도안동로 77, 1808동 2601호(가수원동, 도안18단지 린 풀하우스)

이 발명을 지원한 국가연구개발사업

과제고유번호 114048-3

부처명 농림축산식품부

연구관리전문기관 농림수산식품 기술기획평가원

연구사업명 첨단생산기술사업

연구과제명 친환경 살충제 개발을 위한 방선균 이차대사산물 고생산균주 및 대량생산 시스템 개발

기 여 율 1/1

주관기관 동부팜한농(주)

연구기간 2014.09.25 ~ 2017.09.24

명세서

청구범위

청구항 1

스트렙토마이세스 아베르미틸리스(*Streptomyces avermitilis*) 균주에서,

아베멕틴 생합성을 수행하는 폴리케타이드 합성효소의 1번 유전자(aveA1)의 밀베마이신 생산 균주의 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소의 1번 유전자(milA1)로의 치환; 및

아베멕틴 생합성을 수행하는 폴리케타이드 합성효소의 3번 유전자(aveA3) 또는 이의 일부가 밀베마이신 생산 균주에서 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소의 3번 유전자(milA3) 또는 이의 일부로 치환된,

재조합 스트렙토마이세스 아베르미틸리스 균주.

청구항 2

제1항에 있어서, 상기 aveA3 유전자의 일부가 밀베마이신 생산 균주의 milA3 유전자의 일부로 치환된 것은, aveA3 유전자 중 적어도 모듈 7 코딩 유전자 또는 모듈 7의 DH (dehydratase) 도메인 코딩 유전자가 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 코딩 유전자 또는 모듈 7의 DH (dehydratase) 도메인 및 ER (Enoyl reductase) 도메인 코딩 유전자로 치환된 것을 포함하는, 재조합 스트렙토마이세스 아베르미틸리스 균주.

청구항 3

제2항에 있어서, 상기 aveA3 유전자의 일부가 밀베마이신 생산 균주의 milA3 유전자의 일부로 치환된 것은,

(i) aveA3 유전자의 모듈 7 코딩 유전자의 KS 도메인 코딩 유전자, AT 도메인 코딩 유전자, KR 도메인 코딩 유전자 및 ACP 도메인 코딩 유전자로 이루어진 군에서 선택된 하나 이상의 유전자가 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 코딩 유전자의 KS 도메인 코딩 유전자, AT 도메인 코딩 유전자, KR 도메인 코딩 유전자 및 ACP 도메인 코딩 유전자로 이루어진 군에서 선택된 하나 이상으로 치환된 것;

(ii) aveA3 유전자의 모듈 8 및 모듈 9의 도메인들 중 하나 이상의 코딩 유전자가 밀베마이신 생산 균주의 milA3 유전자의 모듈 8 및 모듈 9의 도메인들 중 하나 이상의 코딩 유전자로의 치환된 것; 또는

(iii) 치환 (i) 및 (ii)의 조합

을 추가로 포함하는 것인,

재조합 스트렙토마이세스 아베르미틸리스 균주.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 밀베마이신 생산 균주는 스트렙토마이세스 밀베마이시니쿠스(*Streptomyces milbemycinicus*), 스트렙토마이세스 시아네오그리세우스 아속 논시아노게누스(*Streptomyces cyaneogriseus* subsp. *Noncyanogenus*), 스트렙토마이세스 난찬젠시스(*Streptomyces nanchangensis*), 또는 스트렙토마이세스 빙첸젠시스(*Streptomyces bingchenggensis*)인, 재조합 스트렙토마이세스 아베르미틸리스 균주.

청구항 5

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 재조합 균주는 밀베마이신 A3, 밀베마이신 A4, 밀베마이신 α2, 밀베마이신 α4, 밀베마이신 G, 및 밀베마이신 D로 이루어진 군에서 선택되는 1종 이상의 밀베마이신을 생산하는 것인, 재조합 스트렙토마이세스 아베르미틸리스 균주.

청구항 6

제1항에 있어서, 상기 재조합 스트렙토마이세스 아베르미틸리스 균주는 스트렙토마이세스 아베르미틸리스 DBM-03-A 균주(수탁번호 KCTC12890BP)인, 재조합 스트렙토마이세스 아베르미틸리스 균주.

청구항 7

제1항에 있어서, 상기 재조합 스트렙토마이세스 아베르미틸리스 균주는 스트렙토마이세스 아베르미틸리스 DBM-03-B 균주(수탁번호 KCTC12891BP)인, 재조합 스트렙토마이세스 아베르미틸리스 균주.

청구항 8

제1항에 있어서, 상기 재조합 스트렙토마이세스 아베르미틸리스 균주는 스트렙토마이세스 아베르미틸리스 DBM-03-C 균주(수탁번호 KCTC13083BP)인, 재조합 스트렙토마이세스 아베르미틸리스 균주.

청구항 9

스트렙토마이세스 아베르미틸리스 균주에서, aveA1 유전자가 밀베마이신 생산 균주의 milA1로 치환되고,

22,23-디하이드로-25-이소프로필-아베멕틴(22,23-dihydro-25-isopropyl-avermectin), 5-메톡시-22,23-디하이드로-25-메틸-아베멕틴(5-methoxy-22,23-dihydro-25-methyl-avermectin), 5-메톡시-22,23-디하이드로-25-에틸-아베멕틴(5-methoxy-22,23-dihydro-25-ethyl-avermectin) 및 5-메톡시-22,23-디하이드로-25-이소프로필-아베멕틴(5-methoxy-22,23-dihydro-25-isopropyl-avermectin)으로 이루어진 군에서 선택되는 1종 이상의 아베멕틴 유도체를 생산하는,

재조합 스트렙토마이세스 아베르미틸리스 균주.

청구항 10

제9항에 있어서, 상기 재조합 스트렙토마이세스 아베르미틸리스 균주는, 스트렙토마이세스 아베르미틸리스 DBM-01 균주(수탁번호 KCTC12889BP)인, 재조합 스트렙토마이세스 아베르미틸리스 균주.

청구항 11

(1) 밀베마이신 생산 균주의 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소의 1번 유전자(milA1); 및
(2) 밀베마이신 생산 균주의 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소의 3번 유전자(milA3)의 전부 또는 일부를 포함하는 재조합 벡터.

청구항 12

제11항에 있어서, 상기 milA3의 일부는 적어도 milA3 유전자 중 모듈 7 코딩 유전자 또는 모듈 7의 DH 도메인 코딩 유전자 및 ER 도메인 코딩 유전자를 포함하는 것인, 재조합 벡터.

청구항 13

제12항에 있어서, 다음으로 이루어진 군에서 선택된 하나 이상을 추가로 포함하는 재조합 벡터:

밀베마이신 생산 균주의 milA3의 모듈 7의 KS 도메인, AT 도메인, KR 도메인, 및 ACP 도메인 중에서 선택된 하나 이상의 코딩 유전자;

밀베마이신 생산 균주의 milA3의 모듈 8의 코딩 유전자 또는 모듈 8의 도메인들 중 하나 이상의 코딩 유전자; 및

밀베마이신 생산 균주의 milA3의 모듈 9의 코딩 유전자 또는 모듈 9의 도메인들 중 하나 이상의 코딩 유전자.

청구항 14

제11항에 있어서,

밀베마이신 생산 균주의 milA1 유전자; 및

밀베마이신 생산 균주의 milA3 유전자 중의 모듈 7의 DH 도메인부터 모듈 9의 KR 도메인까지의 영역의 코딩 유전자를 포함하는 milA3 유전자의 일부

를 포함하는, 재조합 벡터.

청구항 15

제11항 내지 제14항 중 어느 한 항에 있어서, 상기 밀베마이신 생산 균주는 스트렙토마이세스 밀베마이시니쿠스(*Streptomyces milbemycinicus*), 스트렙토마이세스 시아네오그리세우스 아속 논시아노게누스(*Streptomyces cyaneogriseus* subsp. *noncyanogenus*), 스트렙토마이세스 난찬젠시스(*Streptomyces nanchangensis*) 또는 스트렙토마이세스 빙첸젠시스(*Streptomyces bingchenggensis*)인, 제조합 벡터.

청구항 16

제1항 내지 제3항 및 제6항 내지 제8항 중 어느 한 항의 제조합 스트렙토마이세스 아베르미틸리스 균주를 포함하는, 밀베마이신 생산용 조성물.

청구항 17

제1항 내지 제3항 및 제6항 내지 제8항 중 어느 한 항의 제조합 스트렙토마이세스 아베르미틸리스 균주를 배양하는 단계, 및

상기 배양된 균주 또는 균주의 배양물로부터 밀베마이신을 수득하는 단계를 포함하는, 밀베마이신을 생산하는 방법.

청구항 18

하기에서 선택되는 아베멕틴 유도체 또는 이의 염:

5-메톡시-22,23-디하이드로-25-메틸-아베멕틴(5-methoxy-22,23-dihydro-25-methyl-avermectin), 및
5-메톡시-22,23-디하이드로-25-에틸-아베멕틴(5-methoxy-22,23-dihydro-25-ethyl-avermectin).

청구항 19

제9항 또는 제10항 중 어느 한 항의 제조합 스트렙토마이세스 아베르미틸리스 균주 또는 상기 균주의 배양물; 또는

제18항의 아베멕틴 유도체 또는 이의 염
를 포함하는, 해충 방제용 조성물.

발명의 설명

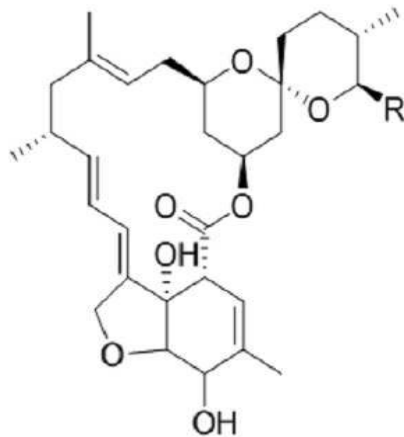
기술 분야

[0001] 밀베마이신을 생산하는 제조합 스트렙토마이세스 아베르미틸리스(*Streptomyces avermitilis*) 균주 및 이를 이용한 밀베마이신 생산 방법이 제공된다.

배경 기술

[0003] 밀베마이신(Milbemycin)은 스트렙토마이세스 하이그로스코피쿠스 아속 아우레오라크리모수스(*Streptomyces hygroscopicus* subsp. *aureolacrimosus*; 스트렙토마이세스 밀베마이시니쿠스(*Streptomyces milbemycinicus*) (NRRL5739 또는 DSM 41911)라고도 칭함), 스트렙토마이세스 시아네오그리세우스 아속 논시아노게누스(*Streptomyces cyaneogriseus* subsp. *noncyanogenus*), 스트렙토마이세스 난찬젠시스(*Streptomyces nanchangensis*) 또는 스트렙토마이세스 빙첸젠시스(*Streptomyces bingchenggensis*)등으로부터 생산되는 폴리케타이드 계열의 16-원환 매크로라이드 화합물로서 다양한 형태의 생물학적, 화학적 유도체로서 농업 및 동물의약품 용도로 상품화되어 있다.

[0004] 밀베마이신의 대표적인 예인 밀베멕틴(milbemectin)은 아래 그림과 같은 구조로서, 밀베마이신 A₃ 와 밀베마이신 A₄가 약 30:70 비율로 혼합되어 있는 형태이다.



Compound	R
Milbemycin A ₃	CH ₃
Milbemycin A ₄	CH ₂ CH ₃

Milbemectin

[0005]

[0006]

이 외에도, 밀베마이신 계열로서 밀베멕틴의 화학적 유도체인 레피멕틴(Lepimectin)이 살충제로서 이용되고 있으며, 동물의약품으로서 밀베마이신 옥심(Milbemycin Oxime)과 네마멕틴(Nemadectin)의 화학적 유도체인 목시텍틴(Moxidectin)이 상업화되어 있다.

[0007]

한편, 아베멕틴(Avermectin)은 밀베마이신과 같이 맥틴계 살충제로 분류되며, 스트렙토마이세스 아베르미틸리스(*Streptomyces avermitilis*) 등으로부터 생산되는 폴리케타이드 계열의 16-원환 매크로라이드 화합물이다. 이들 중 아베멕틴 B1a와 아베멕틴 B1b의 혼합물은 아바멕틴(Abamectin)으로서 응애 등에 대한 살충제로, 아베멕틴의 화학적 유도체인 에마멕틴 벤조에이트(Emamectin benzoate)는 나방 등에 대한 살충제 성분으로 사용되고 있으며, 동물의약품으로서 아베멕틴의 화학적, 생물학적 유도체인 이버멕틴(Ivermectin), 도라멕틴(Doramectin), 셀라멕틴(Selamectin), 에프리노멕틴(Eprinomectin) 등이 상업화되어 있다.

[0008]

스트렙토마이세스 아베르미틸리스의 아베멕틴 생산성을 향상시키기 위하여 무작위적인 돌연변이법, 트랜스포존(transposon) 돌연변이법 등이 적용되었으며, 이를 효율적으로 스크리닝하기 위한 고속스크리닝법이 개발되었다 [Ikeda et al. 1993; Weaden and Dyson 1998; Gao¹ et al. 2010; Gao² et al. 2010; Wang et al. 2010]. 또한, 생산균주의 주요 유전자를 제거하거나 과발현하여 생산성을 증가시키는 연구가 성공적으로 진행되었다 [Duong et al. 2009; Li et al. 2010; Zhuo et al. 2010; Qiu et al. 2011; He et al. 2014; Liu et al. 2015]. 이케다 등은 스트렙토마이세스 아베르미틸리스가 외래의 물질합성 유전자를 효과적으로 발현하여 물질을 생산할 수 있음을 밝히고 지놈엔지니어링을 통해 보다 효과적인 발현 균주로 개발하고 있다 [Ikeda et al. 2014]. 또한, 아베멕틴의 경우 현재 대규모의 상업적 생산이 이뤄지고 있어 아베멕틴 생산균주인 스트렙토마이세스 아베르미틸리스가 산업균주로서 보다 유리할 것으로 여겨지고 있다.

[0009]

밀베마이신의 생산은 스트렙토마이세스 속에 속하는 세균을 배양하고, 균체를 배양액으로부터 분리하여 생산된 밀베마이신을 균체로부터 유기 용매로 추출한 후, 얻어진 목적 화합물을 포함하는 물질을 정제하는 등의 방법에 의해 행해지고 있다. 종래 밀베마이신의 생산에 관한 기술들을 살펴보면, 일본의 Sankyo 사에서 1980년대 스트렙토마이세스 밀베마이시니쿠스(*Streptomyces milbemycinicus*)가 밀베마이신을 생산하는 것을 발견한 후, 무작위적인 돌연변이 개발 기술을 통해 상업적 수준의 밀베마이신의 생산을 시작하였다. 그러나, 이와 같이 개발된 균주를 이용하여 밀베마이신을 대량생산 할 때 발효조 내의 용존 산소량에 의한 생산성 변화가 상당히 크다는 것이 보고되었고, 개발된 균주의 생산성을 유지시키는 것이 상당히 어렵다는 것이 보고된 바 있다 [Okada and Iwamatu. 1997; Ide et al. 1993]. 또한 상기 균주와 밀베마이신을 생산하는 스트렙토마이세스 빙첸젠시스에서 조절 유전자인 nsdA의 삭제를 통하여 밀베마이신 A4의 생산성을 향상시켰다는 보고가 있고 [Wang et al. 2009], 스트렙토마이세스 빙첸젠시스에서 miID를 삭제하여 밀베마이신 A3/A4의 생산성을 향상시켰다는 보고가 있다 [Zhang et al. 2013].

[0010]

그러나, 종래 기술은 원 균주인 스트렙토마이세스 밀베마이시니쿠스(*Streptomyces milbemycinicus*)를 이용하여 상업적 수준으로 생산하는 경우에 있어서도 대량생산의 어려움 때문에 생산성 향상을 이루기 어려우며, 신규의 균주를 이용하는 경우에 있어서도 산업적 수준의 생산성을 얻기 까지 장기간이 걸린다는 문제가 있다.

선행기술문헌

비특허문헌

- [0012] (비특허문헌 0001) Ikeda et al., J. Bacteriol., 175(7): 2077-2078 (1993년 4월)
- (비특허문헌 0002) Weaden and Dyson, Microbiology, 144:1963-1970 (1998년 7월)
- (비특허문헌 0003) Gao et al., Bioresource Technology, 100:4012-4016 (2009년 4월)
- (비특허문헌 0004) Gao1 et al., Appl. Environ. Microbiol., 76(13):4583-4586 (2010년 7월)
- (비특허문헌 0005) Gao2 et al., Appl. Microbiol. Biotechnol., 85(4):1219-1225 (2010년 1월)
- (비특허문헌 0006) Wang et al., J. Appl. Microbiol., 108(3):851-858 (2010년 3월)
- (비특허문헌 0007) Duong et al., J. Microbiol. Biotechnol., 19(2): 136-139 (2009년 2월)
- (비특허문헌 0008) Li et al., Bioresour. Technol., 101(23): 9228-9235 (2010년 12월)
- (비특허문헌 0009) Zhou et al., Proc. Natl. Acad. Sci. (107(25):11250-11254 (2010년 6월)
- (비특허문헌 0010) Qiu et al., Appl. Microbiol. Biotechnol., 92(2):337-345 (2011년 10월)
- (비특허문헌 0011) He et al., Appl. Microbiol. Biotechnol., 98(1):399-409 (2014년 1월)
- (비특허문헌 0012) Liu et al., Appl. Environ. Microbiol., 81(15):5157-5173 (2015년 8월)
- (비특허문헌 0013) Wang et al., Bioorg. Med. Chem. Lett., 21(11):3320-3333 (2011년 6월)
- (비특허문헌 0014) Ikeda et al., J. Ind. Microbiol. Biotechnol., 41:233-250 (2014년 2월)
- (비특허문헌 0015) Wang et al., The Journal of Antibiotics, 62:309-313(2009년 5월)
- (비특허문헌 0016) Zhang et al., Appl Microbiol Biotechnol., 97(23):10091-10101(2013년 12월)
- (비특허문헌 0017) Okada and Iwamatu., J Chem Technol Biotechnol 70(2):179-187(1997년 10월)
- (비특허문헌 0018) Ide et al., Annu. Rep. Sankyo Res. Lab. 45:1-98 (1993년)

발명의 내용

해결하려는 과제

- [0013] 본 발명은 밀베마이신의 산업용 생산 균주로 사용할 정도의 높은 생산성을 가지는 재조합 미생물 및 이를 이용한 밀베마이신 생산 방법과 관련 있다.
- [0014] 일 예는 아베멕틴의 고생산성 균주에서 아베멕틴 생산 유전자의 일부를 밀베마이신 생산 유전자로 치환함으로써 밀베마이신을 고생산하는 재조합 미생물을 제공한다.
- [0015] 다른 예는 밀베마이신 생산 유전자를 포함하는 밀베마이신을 고생산하는 재조합 미생물 제조에 사용 가능한 재조합 벡터를 제공한다.
- [0016] 다른 예는 상기 재조합 미생물을 이용한 밀베마이신 생산 방법을 제공한다.
- [0017] 다른 예는 아베멕틴의 고생산성 균주에서 아베멕틴 생산 유전자의 일부를 밀베마이신 생산 유전자로 치환함으로써 아베멕틴 유도체를 생산하는 재조합 스트렙토마이세스 아베르미틸리스 균주를 제공한다.
- [0018] 다른 예는 신규한 아베멕틴 유도체 또는 이의 염을 제공한다.
- [0019] 다른 예는 상기 신규한 아베멕틴 유도체 또는 이의 염을 포함하는 작물 보호제 및/또는 동물의약품 조성물을 제공한다.

과제의 해결 수단

- [0021] 상기 문제를 해결하기 위한 하나의 방안으로서, 본 명세서에서는, 아베멕틴(ivermectin)과 밀베마이신(milbemycin)의 분자구조가 매우 유사하고 이들을 생산하는 폴리케타이드 합성효소의 모듈(module) 및 도메인(domain)의 구성이 유사함에 근거하여 아베멕틴 고생산성 균주의 아베멕틴 합성 폴리케타이드 합성효소를 밀베마이신 합성 폴리케타이드 합성효소로 전환함으로써 고생산성의 밀베마이신 생산균주를 제작하는 기술이 제공된다 (도 18a-18d 참조).
- [0022] 하나의 양태로서, 본 발명은 아베멕틴의 고생산성 균주에서 아베멕틴 생산 유전자의 일부를 밀베마이신 생산 유전자로 치환하여 얻어진 재조합 미생물에 관한 것이다. 상기 재조합 미생물은 밀베마이신을 고효율로 생산하는 특징으로 갖는 것일 수 있다.
- [0023] 도 18a-18c는 아베멕틴과 밀베마이신의 분자구조 차이 및 이를 유발하는 유전자의 구성을 보여주는 그림이고, 도 18d는 밀베마이신 폴리케타이드 합성효소 유전자군 (gene cluster)과 아베멕틴 폴리케타이드 합성효소 유전자군을 비교하여 보여주는 그림이다. 도 18a-18d에서, 각 원은 각각 도메인을 나타내며, 각 도메인으로부터 코딩되는 단백질은 각각 다음의 활성을 갖는다 (AT: 아실트랜스퍼라제 (Acyltransferase), KS: 케토-아실기 합성효소 (ketosynthase), KR: 케토-아실기 환원효소 (ketoreductase), DH: 탈수효소 (dehydratase), ER: 엔올 환원효소 (enoyl reductase), ACP: 아실기 운반단백질 (Acyl Carrier Protein), TE: 티오에스테라제 (Thioesterase)). 각각의 합성효소는 모듈로 구성되고 각 모듈은 도메인들로 구성되며, 합성효소 유전자군은 모듈 코딩 유전자들로 구성되며, 각 모듈 코딩 유전자는 각 도메인 코딩 유전자들로 구성된다.
- [0024] 각 모듈 내의 각 도메인은 KS 도메인부터 시작하여 ACP 도메인으로 끝나는 순서로 배열될 수 있다 [예컨대, N-말단에서 C-말단 방향으로, (KS)-(AT)-(DH 및/또는 KR; in any order)-(ACP) 순서로 배열, milA3의 모듈 7의 경우 (KS)-(AT)-(DH)-(ER)-(KR)-(ACP) 순서로 배열]. 밀베마이신 및 아베멕틴 폴리케타이드 합성효소의 모듈 10의 KR 도메인은 폴리케타이드 합성에 관여하지 않으며, 밀베마이신 폴리케타이드 합성효소 모듈 11의 KR 도메인과 아베멕틴 폴리케타이드 합성효소 모듈 7의 DH 도메인은 활성이 없다.
- [0025] 도 18a-18d에서 알 수 있듯이, 아베멕틴과 밀베마이신의 분자 구조가 매우 유사하고, 아베멕틴 폴리케타이드 합성효소 유전자군과 베마이신 폴리케타이드 합성효소 유전자군은 그 구성이 매우 유사하다.
- [0026] 두 화합물(아베멕틴과 밀베마이신)의 합성에 있어 차이는 다음에 기인한다:
- [0027] 1) 폴리케타이드 합성효소의 loading 모듈의 AT의 기질특이성 차이로 밀베마이신 합성효소의 loading 모듈의 경우 acetyl-CoA와 propionyl-CoA를 주요 기질로 사용하고, 아베멕틴 합성효소의 loading 모듈의 경우 isobutyryl-CoA와 2-methylbutanoyl-CoA를 주요기질로 사용하고 있어서, 두 화합물의 각 탄소위치 25번에 연결되어 있는 탄소 사슬의 길이가 달라진다.
- [0028] 2) 두 화합물의 탄소위치 22-23의 탄소결합은 각 합성효소의 모듈 2의 도메인 구성에 의해 결정되어, DH-KR 도메인을 갖는 아베멕틴 폴리케타이드 합성효소 모듈 2의 경우 이중 결합 또는 -OH 결합을 생성하는 반면, DH-ER-KR 도메인을 갖는 밀베마이신 폴리케타이드 합성효소 모듈 2는 단일결합을 생성한다.
- [0029] 3) 두 화합물의 탄소위치 13의 -OH기의 생성은 각 합성효소의 모듈 7의 도메인 구성에 의한다. 아베멕틴 폴리케타이드 합성효소의 경우 KR 도메인만 작용하여 -OH가 생성되지만, 밀베마이신 폴리케타이드 합성효소의 경우 DH-ER-KR 도메인이 있어 탄소 단일결합의 형태가 생성된다.
- [0030] 상기 내용에 기초하여, 아베멕틴 생성 균주를 밀베마이신 생성 균주로 전환시키기 위하여 필수적으로 필요한 최소한의 도메인만을 선택적으로 치환할 수 있다. 본 명세서에서는 도메인 치환시의 활성 유지 정도를 고려하여, 적절한 모듈 또는 도메인 단위를 선택하여 치환을 시도하였다.
- [0031] 보다 구체적으로, aveA3와 milA3의 기능적 차이는 모듈 7의 DH (dehydratase) 와 ER (Enoyl reductase) 활성 도메인의 포함 여부에 의존할 수 있다. aveA3의 모듈은 DH 도메인을 갖고 있지만 활성이 없는 반면, milA3는 모듈 7에 DH와 ER 활성을 갖는 도메인을 포함한다. 이외에, aveA3의 모듈 8과 모듈 9는 milA3의 모듈8과 모듈 9와 동일 내지 유사한 기능을 갖는다. 따라서, 기능적인 측면을 고려하면, 최소한 aveA3의 모듈 7 코딩 유전자 중의 DH 도메인 코딩 유전자를 milA3의 모듈7 코딩 유전자 중의 DH-ER 도메인 코딩 유전자로 치환시킴으로써, 아베멕틴 생성균주에 밀베마이신 생산 활성을 부여할 수 있다. 임의로, 폴리케타이드 합성효소와 같은 거대 효소에 있

어 단백질 상호작용에 의해 합성이 조절되는 것이 매우 중요하기 때문에, 단순한 도메인 치환에 의하여 생길 수 있는 활성 감소를 방지하기 위하여, 모듈 단위 (예컨대, 모듈 7)로 치환하거나, 또는 앞서 설명한 도메인 이외의 적절한 도메인 또는 모듈을 추가로 치환할 수 있다.

- [0032] 일 예에서, 스트렙토마이세스 아베르미틸리스(*Streptomyces avermitilis*) 균주에서,
- [0033] 아베멕틴 생합성을 수행하는 폴리케타이드 합성효소 유전자군의 1번 유전자(aveA1; 예컨대, *Streptomyces avermitilis* MA-4680의 aveA1 유전자 (GeneBank Accession number AB032367.1의 101번째부터 12019번째까지의 폴리뉴클레오타이드 부위(DNA 서열: 서열번호 66; 아미노산 서열 (BAA84474.1): 서열번호 67)) 등)가 밀베마이신 생산 균주의 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소 유전자군의 1번 유전자(milA1)로 치환되고,
- [0034] 아베멕틴 생합성을 수행하는 폴리케타이드 합성효소 유전자군의 3번 유전자 (aveA3)의 전부 또는 일부가 밀베마이신 생산 균주에서 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소 유전자군의 3번 유전자 (milA3)의 전부 또는 일부로 치환된,
- [0035] 재조합 스트렙토마이세스 아베르미틸리스 균주가 제공된다.
- [0036] 상기 aveA3 일부가 milA3의 일부로 치환된 것은, aveA3 (예컨대, *Streptomyces avermitilis* MA-4680의 aveA3 유전자의 적어도 모듈 7 코딩 유전자의 전부 또는 일부 (적어도 DH 도메인 코딩 유전자를 포함)가 밀베마이신 생산 균주에서 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소 유전자군의 3번 유전자 (milA3)의 모듈 7 코딩 유전자의 전부 또는 일부 (적어도 DH (dehydratase) 도메인 코딩 유전자 및 ER (enoyl reductase) 도메인 코딩 유전자를 포함)로 치환된 것을 의미할 수 있다.
- [0037] 예컨대, 상기 aveA3는 GeneBank Accession number AB032367.1의 33436번째부터 50034번째까지의 폴리뉴클레오타이드 부위(DNA 서열: 서열번호 71 (polyketide synthase gene cluster (AB032367.1)의 33436번째부터 50034번째까지의 폴리뉴클레오타이드 부위); 아미노산 서열 (BAA84478.1): 서열번호 72 (Module 7: 38번째부터 1841까지 폴리펩티드; Module 8: 1862 번째부터 3641까지 폴리펩티드; Module 9: 3660번째부터 5438까지 폴리펩티드; module 7의 DH domain: 976번째부터 1148번째까지 폴리펩티드))일 수 있으며, 상기 aveA3의 일부는 적어도 모듈 7 (BAA84478.1 (서열번호 72)의 35번째부터 1841번째 또는 38번째부터 1841번째까지의 부위)의 코딩 유전자의 전부 또는 일부 (적어도 DH (dehydratase) 도메인 (BAA84478.1 (서열번호 72)의 976번째부터 1148번째까지의 부위)의 코딩 유전자를 포함)을 의미할 수 있다.
- [0038] aveA3 유전자의 각 모듈, 도메인 및 이의 코딩 유전자는 웹사이트 (<http://www.ncbi.nlm.nih.gov/protein/5902891>)를 통하여 명확하게 특정할 수 있다.
- [0039] 상기 재조합 스트렙토마이세스 아베르미틸리스 균주는 밀베마이신의 고효율 생산 활성을 갖는 것일 수 있다.
- [0040] 폴리케타이드 합성효소 상에서 도메인, 모듈, 및/또는 단백질의 변이(예컨대, 치환) 시에, 각 영역을 암호화하는 유전자를 연결하는 docking 부분간의 상호작용이 해당 폴리케타이드 합성효소의 정상적 발현에 매우 중요하다. 특히 KS (ketosynthase) 도메인, AT (acyltransferase) 도메인, DH (dehydratase), ER (enoylreductase) 도메인, KR (ketoreductase) 도메인, ACP (acyl carrier protein) 도메인 등을 포함하는 복합효소를 포함하는 폴리케타이드 합성효소 (상기한 도메인들은 N-말단에서 C-말단 방향으로 상기 기재 순서대로 위치하는 것일 수 있다)에서, 각 모듈의 KS와 AT는 upstream의 ACP와 상호작용한다. 이를 고려할 때, 아베멕틴 폴리케타이드 합성효소의 정상적 발현을 위해서, aveA3 유전자의 일부가 milA3 유전자의 일부로 치환 시 치환이 일어나는 염기 서열의 위치가 중요하게 고려되어야 한다.
- [0041] 상기 milA3 유전자의 일부로 치환되는 aveA3 유전자의 일부는 적어도 모듈 7의 전부 또는 일부 (적어도 DH 도메인을 포함)의 코딩 유전자를 포함할 수 있고, 이를 치환하는 milA3 유전자의 일부는 적어도 모듈 7의 전부 또는 일부 (적어도 DH 도메인 및 ER 도메인을 포함)의 코딩 유전자를 포함할 수 있다.
- [0042] aveA3 유전자의 일부를 밀베마이신 생산 균주의 milA3 유전자의 일부로 치환 시, aveA3 유전자의 N-말단 코딩 부위 및 C-말단 코딩 부위의 docking 영역에 더하여, 모듈 7의 KS (KS7), 모듈 7의 AT 도메인 (AT7), 및/또는 모듈 9의 ACP 도메인 (ACP9)의 코딩 유전자 부위들이 아베멕틴 폴리케타이드 합성효소(aveA3 유전자)의 상기 도메인의 코딩 유전자 부위들로 보존되도록 유전자 치환이 일어날 수 있다.
- [0043] 예컨대, aveA3 유전자의 일부 치환은

- [0044] (a) aveA3 유전자 중 적어도 모듈 7 코딩 유전자 또는 모듈 7의 DH (dehydratase) 도메인 코딩 유전자가 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 코딩 유전자 또는 모듈 7의 DH (dehydratase) 도메인 및 ER (enoyl reductase) 도메인 코딩 유전자로 치환되거나; 또는
- [0045] (b) 상기 (a) 치환에 더하여, 다음을 추가로 포함하는 치환일 수 있다:
- [0046] (i) aveA3 유전자의 모듈 7 코딩 유전자 중의 DH 도메인 코딩 유전자를 제외한 나머지 도메인 중 하나 이상의 코딩 유전자 (예컨대, KS 도메인 코딩 유전자, AT 도메인 코딩 유전자, KR 도메인 코딩 유전자 및 ACP 도메인 코딩 유전자로 이루어진 군에서 선택된 하나 이상의 유전자)의 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 코딩 유전자 중의 DH 도메인 및 ER 도메인을 제외한 나머지 도메인 중 하나 이상의 코딩 유전자 (예컨대, KS 도메인 코딩 유전자, AT 도메인 코딩 유전자, KR 도메인 코딩 유전자 및 ACP 도메인 코딩 유전자로 이루어진 군에서 선택된 하나 이상)로의 치환;
- [0047] (ii) aveA3 유전자의 모듈 8 및 모듈 9의 도메인들 중 하나 이상의 코딩 유전자의 밀베마이신 생산 균주의 milA3 유전자의 모듈 8 및 모듈 9의 도메인들 중 하나 이상의 코딩 유전자로의 치환; 또는
- [0048] (iii) 치환 (i) 및 (ii)의 조합.
- [0049] 앞서 설명한 aveA1 유전자의 전부가 milA1 유전자의 전부로 치환 및/또는 aveA3 유전자의 전부 또는 일부가 milA3 유전자의 전부 또는 일부로 치환된 균주는, aveA1 및/또는 aveA3 유전자의 유전자 치환 부위의 N-말단 및/또는 C-말단의 docking 영역 코딩 부위가 보존되어 있는 것일 수 있다. 즉, aveA1 유전자의 전부가 milA1 유전자의 전부로 치환 및/또는 aveA3 유전자의 전부 또는 일부가 milA3 유전자의 전부 또는 일부로 치환된 균주는, aveA1 및/또는 aveA3 유전자의 유전자 치환 부위의 upstream 및/또는 downstream 영역과 상동성을 가지는 부위를 상기 치환(도입)된 milA1 유전자 및/또는 milA3 유전자의 전부 또는 일부의 어느 한 말단 또는 양쪽 말단에 연결된 형태로 포함할 수 있다. 이러한 상동성을 가지는 부위의 도입은 상동 재조합시에 유리할 수 있다. 상기 상동성은 원래의 유전자 서열과 90% 이상, 95% 이상, 또는 98% 이상 동일성을 가짐을 의미한다.
- [0050] 상기 사용된 용어 "Docking 영역"은 폴리케타이드 합성효소를 코딩하고 있는 유전자 양 끝 부위에 의하여 코딩되는 부위로, 폴리케타이드 합성시 다음 단계의 합성효소와의 단백질-단백질 상호작용을 원활하게 하는 역할을 하는 부위이다. 통상적인 상동성 염기서열 분석을 범위의 특징이 가능하다.
- [0051] 상기 밀베마이신 생산 균주는 스트렙토마이세스 밀베마이시니쿠스 (*Streptomyces milbemycinicus*) (milA1 유전자: 서열번호 70; milA3 유전자: 서열번호 73; milA3 단백질: 서열번호 74 (BAA84478.1) (Module 7: 34번째부터 2139 번째까지 폴리펩티드; Module 8: 2163번째부터 3927번째까지 폴리펩티드; Module 9: 3951번째부터 5731 번째까지 폴리펩티드: DH-ER domain in module 7: 953번째부터 1775 번째까지 폴리펩티드 (DH domain in module 7: 953번째부터 1129 번째까지 폴리펩티드; ER domain in module 7: 1497 번째부터 1775 번째까지 폴리펩티드))), 스트렙토마이세스 시아네오그리세우스 아속 논시아노게누스(*Streptomyces cyaneogriseus subsp. Noncyanogenus*; 예컨대, Accession no. CP010849), 스트렙토마이세스 난창젠시스(*Streptomyces nanchangensis*; 예컨대, Accession no. FJ952082; milA1 유전자 (Meilingmycin이 milbemycin과 구조가 유사하고 기능이 동일하여, 본 명세서에서는 meiA1와 milA1을 동일한 의미로 사용함): 서열번호 69; milA3 유전자(본 명세서에서는 meiA3와 milA3을 동일한 의미로 사용함): 서열번호 75 (GenBank: FJ952082 Sequence 영역 78606번째부터 96074 번째까지 폴리뉴클레오티드); milA3 단백질: 서열번호 76 (Module 7: 39번째부터 2143 번째까지 폴리펩티드; Module 8: 2166번째부터 3931 번째까지 폴리펩티드; Module 9: 3952번째부터 5734 번째까지 폴리펩티드; Module 7의 DH-ER didomain 957번째부터 2143 번째까지 폴리펩티드 (Module 7의 DH domain: 957번째부터 1133 번째까지 폴리펩티드; 모듈 7의 ER domain: 1501번째부터 1779 번째까지 폴리펩티드))) 또는 스트렙토마이세스 빙첸젠시스(*Streptomyces bingchenggensis*; 예컨대, Accession no. CP002047.1; milA1 유전자: 서열번호 68; milA3 유전자: 서열번호 77 (Accession no. CP002047.1 중 1063754::1081234 (유전자가 3'→5' 이어서 reverse complement seq를 기재함)); 단백질: 서열번호 78 (Genbank Accession No. ADI03854; Module 7: 35::2150, Module 8: 2173::3938, Module 9: 3990::5738, DH-ER didomain in module 7: 950::1772 (DH domain: 950::1126; ER domain: 1494::1772)))일 수 있다.
- [0052] 다른 구체예로, 상기 재조합 균주는 밀베마이신 A3, 밀베마이신 A4, 밀베마이신 α2, 밀베마이신 α4 및 밀베마이신 G 로 이루어진 군에서 선택되는 1종 이상의 밀베마이신을 생산할 수 있다.
- [0053] 다른 구체예로, 상기 재조합 균주는 스트렙토마이세스 아베르미틸리스 DBM-03-A 균주(수탁번호 KCTC12890BP), 스트렙토마이세스 아베르미틸리스 DBM-03-B 균주(수탁번호 KCTC12891BP), 및 스트렙토마이세스 아베르미틸리스

DBM-03-C 균주(수탁번호 KCTC13083BP)로 이루어진 균에서 선택된 1종 이상일 수 있다.

- [0054] 다른 양태로서, 앞서 설명한 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소 유전자 milA1 및 milA3의 전부 또는 일부 (적어도 모듈 7 코딩 유전자 또는 모듈 7의 DH-ER 도메인 코딩 유전자를 포함)를 포함하는 재조합 벡터를 적절한 숙주세포 (예컨대, 스트렙토마이세스 아베르미틸리스와 같은 아베멕틴 생성 균주)에 도입하는 단계를 포함하는 재조합 미생물 (예컨대, 재조합 스트렙토마이세스 아베르미틸리스 균주)의 제조 방법을 제공한다.
- [0055] 상기와 같이 제조된 재조합 미생물은 상기 재조합 벡터가 도입되지 않은 미생물과 비교하여 밀베마이신을 생산 효율이 현저히 증가한 특징을 갖는다.
- [0056] 상기 도입시에, 숙주세포 (예컨대, 스트렙토마이세스 아베르미틸리스와 같은 아베멕틴 생성 균주)의 폴리케타이드 합성효소 유전자 aveA1 및 aveA3의 전부 또는 일부 (적어도 모듈 7 코딩 유전자 또는 모듈 7의 DH 도메인 코딩 유전자를 포함)가 milA1 및 milA3의 전부 또는 일부 (적어도 모듈 7 코딩 유전자 또는 모듈 7의 DH-ER 도메인 코딩 유전자를 포함)로 치환이 일어날 수 있다. 또한, aveA1 및 aveA3의 활성이 제거된 상태에서 milA1 및 milA3의 전부 또는 일부 (적어도 모듈 7 코딩 유전자 또는 모듈 7의 DH-ER 도메인 코딩 유전자를 포함)를 포함하는 재조합 벡터가 도입되는 경우 보다 상승된 밀베마이신 생산 효과를 얻을 수 있다.
- [0057] 또 하나의 양태로서, 상기 재조합 미생물 (즉, 상기 재조합 스트렙토마이세스 아베르미틸리스 균주) 또는 상기 재조합 벡터를 포함하는 밀베마이신 제조용 조성물이 제공된다.
- [0058] 또 하나의 양태로서, 상기 재조합 미생물을 이용한 밀베마이신 생산 방법이 제공된다.
- [0059] 구체예로, 본 발명은 상기 재조합 스트렙토마이세스 아베르미틸리스 균주를 배양하는 단계, 및 상기 배양된 균주 또는 균주의 배양물로부터 밀베마이신을 수득하는 단계를 포함하는, 밀베마이신을 생산하는 방법에 관한 것이다.
- [0060] 또 하나의 양태로서, 본 발명은 스트렙토마이세스 아베르미틸리스(*Streptomyces avermitilis*) 균주에서 아베멕틴 생합성을 수행하는 폴리케타이드 합성효소의 1번 유전자(aveA1)가 밀베마이신 생산 균주의 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소의 1번 유전자(milA1)로 치환되며, 22,23-디하이드로-25-이소프로필-아베멕틴(22,23-dihydro-25-isopropyl-avermectin), 5-메톡시-22,23-디하이드로-25-메틸-아베멕틴(5-methoxy-22,23-dihydro-25-methyl-avermectin), 5-메톡시-22,23-디하이드로-25-에틸-아베멕틴(5-methoxy-22,23-dihydro-25-ethyl-avermectin) 및 5-메톡시-22,23-디하이드로-25-이소프로필-아베멕틴(5-methoxy-22,23-dihydro-25-isopropyl-avermectin)으로 이루어진 균에서 선택되는 1종 이상의 아베멕틴 유도체를 생산하는 재조합 스트렙토마이세스 아베르미틸리스 균주에 관한 것이다. 예컨대, 상기 재조합 스트렙토마이세스 아베르미틸리스 균주는 앞서 설명한 milA1을 포함하는 재조합 벡터를 사용하여 제작된 것일 수 있다.
- [0061] 구체 예로, 상기 재조합 스트렙토마이세스 아베르미틸리스 균주는, 스트렙토마이세스 아베르미틸리스 DBM-01 균주(수탁번호 KCTC12889BP)일 수 있다.
- [0062] 또 하나의 양태로서, 본 발명은 하기에서 선택되는, 아베멕틴 유도체 또는 이의 염에 관한 것이다: 5-메톡시-22,23-디하이드로-25-메틸-아베멕틴(5-methoxy-22,23-dihydro-25-methyl-avermectin) 또는 이의 염, 및 5-메톡시-22,23-디하이드로-25-에틸-아베멕틴(5-methoxy-22,23-dihydro-25-ethyl-avermectin) 또는 이의 염.
- [0063] 또 하나의 양태로서, 본 발명은 상기 아베멕틴 유도체 또는 이의 염의 작물 보호제 및/또는 동물의약품으로서의 용도에 관한 것이다.
- [0064] 이하, 본 발명을 보다 상세하게 설명한다.
- [0065] 스트렙토마이세스 아베르미틸리스는 아베멕틴을 생산하는 균주로서, 그 예로는, *S. avermitilis* SA-01 균주, *S. avermitilis* MA-4680 균주(NCBI 등록번호: NC_003155.4), *S. avermitilis* 76-02-e 균주(He et al. 2014), *S. avermitilis* 14-12A 균주(Gao et al. 2009), *S. avermitilis* 3-115 균주(Gao et al. 2010) 등을 들 수 있으나, 이에 제한되는 것은 아니다.
- [0066] 스트렙토마이세스 아베르미틸리스는 아베멕틴을 생산하기 위한 PKS 클러스터(Polyketide synthase gene cluster; 폴리케타이드 합성효소 유전자 클러스터)라는 거대 유전자군을 포함하고 있다. 상기 아베멕틴 PKS 유전자군은 aveA1, aveA2, aveA3 및 aveA4 유전자를 포함하고 있으며, aveA1 유전자는 로딩 모듈과 모듈 1번과 모듈 2번, aveA2 유전자는 모듈 3번 내지 모듈 6번, aveA3 유전자는 모듈 7 내지 모듈 9, aveA4 유전자는 모듈 10

번 내지 모듈 12번을 포함하고 있고, 각 모듈은 하위 도메인들로 구성되어 있다.

[0067] 밀베마이신 생산 균주는 스트렙토마이세스 밀베마이시니쿠스 (*Streptomyces milbemycinicus*), 스트렙토마이세스 시아네오그리세우스 아속 논시아노게누스(*Streptomyces cyaneogriseus* sp. *Noncyanogenus*), 스트렙토마이세스 난창젠시스(*Streptomyces nanchangensis*) 또는 스트렙토마이세스 빙첸젠시스(*Streptomyces bingchenggensis*)를 포함하나, 이에 제한되는 것은 아니다. 스트렙토마이세스 하이그로스코피쿠스 아속 아우레오라크리모수스로는, 스트렙토마이세스 밀베마이시니쿠스 (*Streptomyces milbemycinicus*) NRRL 5739 균주를 사용할 수 있으나, 이에 제한되는 것은 아니다.

[0068] 밀베마이신 생산 균주는 밀베마이신을 생산하기 위한 PKS 클러스터 유전자군을 포함하고 있다. 상기 밀베마이신 PKS 유전자군의 구성은 아베멕틴 PKS 유전자군의 구성과 유사하며, milA1, milA2, milA3 및 milA4 유전자를 포함하고 있으며, 각 유전자는 모듈 및 하위 도메인들로 구성되어 있다.

[0069] 본 발명은 아베멕틴 생산 균주인 스트렙토마이세스 아베르미틸리스의 aveA1 유전자, 이에 상응하는 밀베마이신 생산 균주의 milA1 유전자로 치환시킴으로써 스트렙토마이세스 아베르미틸리스 내에 하이브리드 PKS 유전자가 포함되도록 하였고, 상기 하이브리드 PKS 를 포함하는 재조합 스트렙토마이세스 아베르미틸리스 균주가 다양한 아베멕틴 유도체들을 생산할 수 있음을 확인하였다.

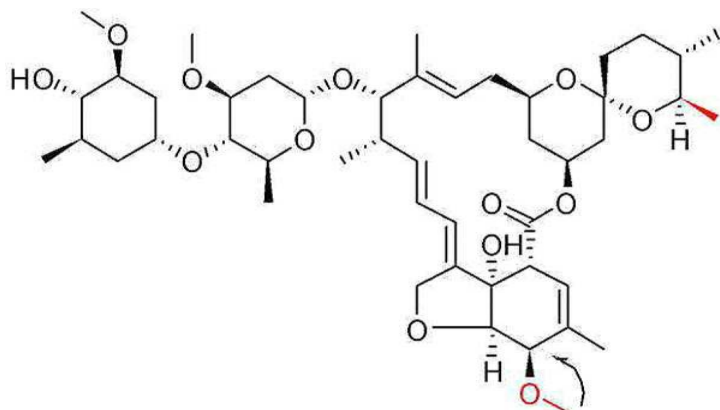
[0070] 상기 재조합 스트렙토마이세스 아베르미틸리스 균주가 생산하는 아베멕틴 유도체는, 22,23-디하이드로-25-이소프로필-아베멕틴(22,23-dihydro-25-isopropyl-avermectin), 5-메톡시-22,23-디하이드로-25-메틸-아베멕틴(5-methoxy-22,23-dihydro-25-methyl-avermectin), 5-메톡시-22,23-디하이드로-25-에틸-아베멕틴(5-methoxy-22,23-dihydro-25-ethyl-avermectin) 및 5-메톡시-22,23-디하이드로-25-이소프로필-아베멕틴(5-methoxy-22,23-dihydro-25-isopropyl-avermectin)으로 이루어진 군에서 선택되는 1종 이상을 포함한다.

[0071] 또한, 이 중에서 5-메톡시-22,23-디하이드로-25-메틸-아베멕틴(5-methoxy-22,23-dihydro-25-methyl-avermectin), 5-메톡시-22,23-디하이드로-25-에틸-아베멕틴(5-methoxy-22,23-dihydro-25-ethyl-avermectin)은 종래 구축된 화합물 데이터베이스에서 검색되지 않는 신규한 화합물임이 확인되었다. 이들 신규 아베멕틴 유도체들은 작물 보호제 및/또는 동물의약품 용도로 사용될 수 있으며, 종래 개발된 아베멕틴 유도체들을 대체할 수 있다.

[0072] 다른 예로, 스트렙토마이세스 아베르미틸리스(*Streptomyces avermitilis*)의 aveA1 유전자가 밀베마이신 생산 균주의 milA1 유전자의 전부 또는 일부로 치환되며, 아베멕틴 유도체들을 생산하는 재조합 스트렙토마이세스 아베르미틸리스 균주가 제공된다.

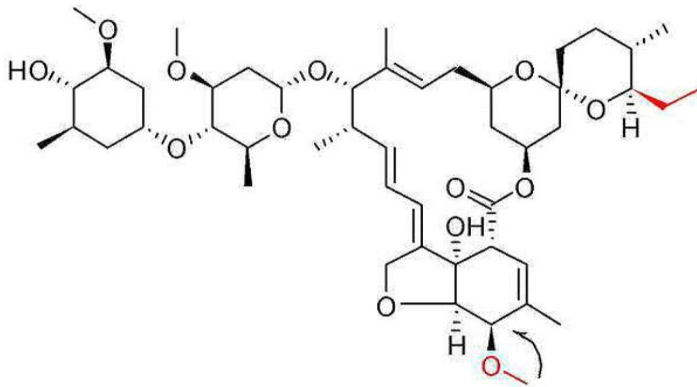
[0073] 다른 예로, 5-메톡시-22,23-디하이드로-25-메틸-아베멕틴(5-methoxy-22,23-dihydro-25-methyl-avermectin) 또는 이의 염, 5-메톡시-22,23-디하이드로-25-에틸-아베멕틴(5-methoxy-22,23-dihydro-25-ethyl-avermectin) 또는 이의 염, 및 이들 화합물들의 작물 보호제 및/또는 동물의약품으로서의 용도가 제공된다.

[0074] 상기 5-메톡시-22,23-디하이드로-25-메틸-아베멕틴은 다음의 화학식으로 나타낼 수 있다:



[0075]

[0076] 또한, 상기 5-메톡시-22,23-디하이드로-25-에틸-아베멕틴은 다음의 화학식으로 나타낼 수 있다:



[0077]

[0078] 또한, 앞서 설명한 바와 같이, 아베멕틴 생산 균주인 스트렙토마이세스 아베르미틸리스의 aveA1 유전자 및 aveA3 유전자의 전체 또는 일부 (적어도 aveA3의 모듈 7의 DH 도메인 코딩 유전자를 포함)를, 이에 상응하는 밀베마이신 생산 균주의 milA1 및 milA3 유전자의 전체 또는 일부 (적어도 milA3의 모듈 7의 DH-ER 도메인 코딩 유전자를 포함)로 치환시킴으로써 스트렙토마이세스 아베르미틸리스 내에 하이브리드 PKS 유전자가 포함되도록 하였고, 상기 하이브리드 PKS 유전자를 포함하는 재조합 균주가 밀베마이신을 생산할 수 있음이 확인되었다.

[0079] 따라서, 다른 예로, 스트렙토마이세스 아베르미틸리스(*Streptomyces avermitilis*)의 aveA1 유전자의 전부 또는 일부가 밀베마이신 생산 균주의 milA1 유전자의 전부 또는 일부로 치환되고, aveA3 유전자의 전부 또는 일부(적어도 aveA3의 모듈 7의 DH 도메인 코딩 유전자를 포함)가 밀베마이신 생산 균주의 milA3 유전자의 전부 또는 일부(적어도 milA3의 모듈 7의 DH-ER 도메인 코딩 유전자를 포함)로 치환된, 밀베마이신을 생산하는 재조합 스트렙토마이세스 아베르미틸리스 균주가 제공된다.

[0080] 상기 aveA3 유전자는 전체가 milA3 유전자로 치환될 수도 있으나, milA3 유전자 중에서 숙주 균주에 밀베마이신을 생산능을 부여할 수 있는 최소한의 부위만 aveA3 유전자의 상응하는 부위에 치환(삽입)되는 것이 가능하다. 예를 들어, 상기 milA3 유전자 중에서 숙주 균주에 밀베마이신을 생산능을 부여할 수 있는 최소한의 부위는 milA3 유전자 중 모듈 7 전부 또는 모듈 7의 DH-ER 도메인의 코딩 유전자를 포함할 수 있으며, aveA3 유전자 중 모듈 7 전부 또는 적어도 모듈 7의 DH 도메인을 포함하는 모듈 7의 일부의 코딩 유전자가, 밀베마이신 생산 균주의 milA3 유전자 중 모듈 7 전부 또는 모듈 7의 DH-ER 도메인의 코딩 유전자로 치환될 수 있다.

[0081] 일 구현예로, 스트렙토마이세스 아베르미틸리스 균주에서 aveA1 유전자가 밀베마이신 생산 균주의 milA1 유전자로 치환된 균주는, 스트렙토마이세스 아베르미틸리스 균주는, 스트렙토마이세스 아베르미틸리스 DBM-01 균주(수탁번호 KCTC12889BP)일 수 있다.

[0082] 다른 예로, 스트렙토마이세스 아베르미틸리스 균주에서 aveA1 유전자가 밀베마이신 생산 균주의 milA1 유전자로 치환되고, aveA3 유전자의 모듈 7의 코딩 유전자 부위가 밀베마이신 생산 균주의 milA3 유전자의 모듈 7의 코딩 유전자 부위로 치환된 균주는, 스트렙토마이세스 아베르미틸리스 DBM-03-A 균주(수탁번호 KCTC12890BP)일 수 있다.

[0083] 다른 예로, 스트렙토마이세스 아베르미틸리스 균주에서 aveA1 유전자가 밀베마이신 생산 균주의 milA1 유전자로 치환되고, aveA3 유전자의 모듈 7 중 KS 도메인을 제외한 모듈 7의 도메인들의 코딩 유전자 부위가 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 중 KS 도메인을 제외한 모듈 7의 도메인들의 코딩 유전자 부위로 치환된 균주는, 스트렙토마이세스 아베르미틸리스 DBM-03-B 균주(수탁번호 KCTC12891BP)일 수 있다.

[0084] 다른 예로, 스트렙토마이세스 아베르미틸리스 균주에서 aveA1 유전자가 밀베마이신 생산 균주의 milA1 유전자로 치환되고, aveA3 유전자의 모듈 7 중 KS 도메인 및 AT를 제외한 모듈 7의 도메인들의 코딩 유전자 부위가 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 중 KS 도메인 및 AT 도메인을 제외한 모듈 7의 도메인들의 코딩 유전자 부위로 치환되고, aveA3 유전자의 모듈 9 중 ACP 도메인을 제외한 모듈 9의 도메인들의 코딩 유전자 부위가 밀베마이신 생산 균주의 milA3 유전자의 모듈 9 중 ACP 도메인을 제외한 모듈 9의 도메인들의 코딩 유전자 부위로 치환된 균주가 제공된다. 상기 균주는 aveA3의 모듈 7의 DH 도메인부터 모듈 9의 KR 도메인까지의 영역(aveA3의 모듈 7의 DH 도메인, KR 도메인, 및 ACP 도메인, 모듈 8, 모듈 9의 KS 도메인, AT 도메인, DH 도메인, 및 KR 도메인을 N-말단에서 C-말단 방향으로 포함하는 영역)의 코딩 유전자 부위가 밀베마이신 생산 균주의 milA3의 모듈 7의 DH 도메인부터 모듈 9의 KR 도메인까지의 영역(milA3의 모듈 7의 DH 도메인, ER 도메인,

KR 도메인, 및 ACP 도메인, 모듈 8, 모듈 9의 KS 도메인, AT 도메인, DH 도메인, ER 도메인, 및 KR 도메인을 N-말단에서 C-말단 방향으로 포함하는 영역)의 코딩 유전자 부위로 치환된 것일 수 있으며, 예컨대, 스트렙토마이세스 아베르미틸리스 DBM-03-C 균주(수탁번호 KCTC13083BP)일 수 있다.

- [0085] 그러나, 상기 기탁 균주들은 본 발명의 대표적인 구현예에 불과하며, 본 발명의 범위가 이에 제한되는 것은 아니다.
- [0086] 스트렙토마이세스 아베르미틸리스의 aveA1 및/또는 aveA3 유전자의 치환은, 당업계에 알려진 공지의 기술들, 예를 들어 상동 재조합(homologous recombination)에 의하여 수행될 수 있다.
- [0087] 일 구현예로, 밀베마이신 생산 균주에서 분리한 milA1 및/또는 milA3 유전자의 전체 또는 일부가 상동 재조합을 통하여 숙주 균주 계놈에 통합될 수 있도록 하기 위한 유전자 치환용 벡터를 제조할 수 있다. 상기 벡터는 숙주 계놈의 특정 유전자 위치로 목적하는 유전자를 제거 또는 삽입할 수 있는 벡터로, 상동 재조합이 일어나도록 타겟팅하고자 하는 특정 유전자 부위에 상동인 염기 서열을 포함할 수 있다.
- [0088] 본 발명의 일 양태로서, 상기 스트렙토마이세스 아베르미틸리스 재조합 균주 생산에 사용되는 재조합 벡터가 제공된다.
- [0089] 상기 재조합 벡터는
- [0090] (1) 밀베마이신 생산 균주의 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소의 1번 유전자(milA1); 및
- [0091] (2) 밀베마이신 생산 균주의 밀베마이신 생합성을 수행하는 폴리케타이드 합성효소의 3번 유전자(milA3)의 전부 또는 일부 (적어도 milA3 유전자 중 모듈 7 코딩 유전자 또는 모듈 7의 DH 도메인 코딩 유전자 및 ER 도메인 코딩 유전자를 포함)
- [0092] 을 포함하는 것일 수 있다. 이 경우 milA1 유전자와 milA3의 전부 또는 일부는 하나의 벡터에 함께 포함되거나 각각 별개의 벡터에 포함될 수 있다.
- [0093] 일 구체예에서, 상기 재조합 벡터는 밀베마이신 생산 균주의 milA1 유전자; 및 밀베마이신 생산 균주의 milA3 유전자 또는 밀베마이신 생산 균주의 milA3의 모듈 7의 전부 또는 일부의 코딩 유전자에 더하여, 다음으로 이루어진 군에서 선택된 하나 이상을 추가로 포함할 수 있다:
- [0094] 밀베마이신 생산 균주의 milA3의 모듈 7의 DH-ER 도메인을 제외한 도메인들 (예컨대, KS, AT, KR, 및 ACP 도메인) 중에서 선택된 하나 이상의 코딩 유전자;
- [0095] 밀베마이신 생산 균주의 milA3의 모듈 8의 코딩 유전자 또는 모듈 8의 도메인들 중 하나 이상의 코딩 유전자; 및
- [0096] 밀베마이신 생산 균주의 milA3의 모듈 9의 코딩 유전자 또는 모듈 9의 도메인들 중 하나 이상의 코딩 유전자.
- [0097] 일 구체예에서, 상기 재조합 벡터는
- [0098] 밀베마이신 생산 균주의 milA1 유전자; 및 다음에서 선택된 milA3 유전자 일부 또는 이의 변이체를 포함할 수 있다:
- [0099] 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 중의 KS 도메인의 코딩 유전자 부위가 결실된, 예컨대, milA3 유전자의 모듈 7의 AT 도메인, DH 도메인, ER 도메인, KR 도메인 및 ACP 도메인의 코딩 유전자를 포함하거나 이들 로 구성된 milA3 유전자의 일부;
- [0100] 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 중의 KS 도메인의 코딩 유전자가 aveA3 유전자의 모듈 7의 KS 도메인의 코딩 유전자로 치환된 milA3의 모듈 7코딩 유전자 변이체;
- [0101] 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 중의 AT 도메인의 코딩 유전자 부위가 결실된, 예컨대, milA3 유전자의 모듈 7의 KS 도메인, DH 도메인, ER 도메인, KR 도메인 및 ACP 도메인의 코딩 유전자를 포함하는 milA3의 모듈 7코딩 유전자 변이체;
- [0102] 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 중의 AT 도메인의 코딩 유전자가 aveA3 유전자의 모듈 7의 AT 도메인의 코딩 유전자로 치환된 milA3의 모듈 7의 코딩 유전자 변이체;
- [0103] 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 중의 KS 도메인 및 AT 도메인의 코딩 유전자가 결실된, 예컨대, milA3 유전자의 모듈 7의 DH 도메인, ER 도메인, KR 도메인 및 ACP 도메인의 코딩 유전자를 포함하는 milA3의

모듈 7의 코딩 유전자 변이체; 및

- [0104] 밀베마이신 생산 균주의 milA3 유전자의 모듈 7 중의 KS 도메인 및 AT 도메인의 코딩 유전자가 각각 aveA3 유전자의 모듈 7의 KS 도메인 및 AT 도메인의 코딩 유전자로 치환된 milA3의 모듈 7의 코딩 유전자 부위 변이체.
- [0105] 상기 재조합 벡터는 다음으로 이루어진 군에서 선택된 하나 이상을 추가로 포함할 수 있다:
- [0106] 밀베마이신 생산 균주의 milA3 유전자의 모듈 9 중의 ACP 도메인의 코딩 유전자가 결실된, 예컨대, milA3 유전자의 모듈 9의 KS 도메인, AT 도메인, DH 도메인, 및 KR 도메인의 코딩 유전자를 포함하는 milA3의 모듈 9의 코딩 유전자 변이체; 및
- [0107] 밀베마이신 생산 균주의 milA3 유전자의 모듈 9 중의 ACP 도메인의 코딩 유전자가 aveA3 유전자의 모듈 9의 ACP 도메인의 코딩 유전자로 치환된 milA3의 모듈 9의 코딩 유전자 변이체.
- [0108] 일 예에서, 상기 재조합 벡터에 포함되는 milA3 유전자 변이체는
- [0109] 밀베마이신 생산 균주의 milA3 유전자 중의 모듈 7의 DH 도메인부터 모듈 9의 KR 도메인까지의 영역(예컨대, milA3의 모듈 7의 DH 도메인, ER 도메인, KR 도메인, 및 ACP 도메인, 모듈 8, 모듈 9의 KS 도메인, AT 도메인, DH 도메인, ER 도메인, 및 KR 도메인을 N-말단에서 C-말단 방향으로 포함하는 영역)의 코딩 유전자를 포함하는 milA3 유전자 변이체 (일부)일 수 있으나 이에 제한되는 것은 아니다.
- [0110] 본원의 일 실시예에서, milA1을 치환하기 위한 재조합 벡터는 milA1 유전자 전체를 포함하고, 상동 재조합을 위하여 aveA1 유전자의 upstream 및 downstream 지역과 상동성을 가지는 부위를 milA1의 어느 한쪽 말단 또는 양쪽 말단에 연결된 형태로 구성될 수 있다. 또한, 다른 실시예에서는, milA3을 치환하기 위한 재조합 벡터는, milA3 유전자의 7번 모듈의 코딩 유전자의 전체 또는 일부를 포함하고, 상동 재조합을 위하여 aveA3 및/또는 aveA4 유전자 지역과 상동성을 가지는 부위를 milA3 유전자의 7번 모듈의 코딩 유전자의 전체 또는 일부의 어느 한쪽 말단 또는 양쪽 말단에 연결된 형태로 구성될 수 있다. 상기 벡터는 본원의 대표적인 실시예일뿐이며, 본원의 범위가 이에 제한되는 것은 아니다.
- [0111] 상기 벡터는, 상동 재조합을 위하여 숙주 유전자 부위와 상동성을 가지는 부위를 포함하는데, 여기에서 상동이란 숙주 유전자 부위의 염기 서열과의 동일성 정도를 나타내는 것으로, 예를 들어, 숙주 유전자의 염기서열과 90% 이상, 95% 이상 또는 98% 이상 동일한 것일 수 있다.
- [0112] 스트렙토마이세스에서 사용하기 위한 다양한 여러 벡터들, 예를 들어 과지, 고 카피수(high copy number) 플라스미드, 저 카피수 플라스미드, 및 대장균(E. Coli)-스트렙토마이세스 서플 벡터들이 개발되어 왔으며, 이들 벡터들을 본 발명을 구현하기 위하여 사용할 수 있다. 예를 들어, pCR-Blunt, pCR2.1(Invitrogen), pGEM3Zf(Promega) 및 서플 벡터 pWHM3, pKC1139 등을 예시할 수 있으나, 이에 제한되는 것은 아니다.
- [0113] 상기 벡터는, 형질전환된 세포를 선별하기 위한 목적으로 선별마커(selection marker)를 더욱 포함할 수 있으며, 예를 들어, 약물 내성, 영양 요구성, 세포 독성제에 대한 내성 또는 표면 단백질의 발현과 같은 선택가능 표현형을 부여하는 마커들이 사용될 수 있고, 포지티브 선별마커와 네가티브 선별마커를 예시할 수 있다. 포지티브 선별마커란 선택제(selective agent)가 처리된 환경에서 선택 마커를 발현하는 세포만 생존하도록 하여 포지티브 선택을 가능하게 하는 마커로, 아프라마이신, 네오마이신, 하이그로마이신, 히스티딘올 디하이드로게나제(histidinol dehydrogenase gene: hisD) 또는 구아닌 포스포리보실트랜스퍼라제(guanine phosphoribosyltransferase: Gpt) 등이 있으나, 이에 제한되지 않는다. 네가티브 선별마커란 무작위적 삽입(random insertion)이 일어난 세포를 선별하여 제거하는 네가티브 선택을 가능하게 하는 마커로, 허피스 심플렉스 바이러스-싸이미딘 키나제(Herpes simplex virus-thymidine kinase: HSV-tk), 하이포잔틴 포스포리보실 트랜스퍼라제(hypoxanthine phosphoribosyl transferase: Hprt), 싸이토신 디아미네즈(cytosine deaminase), 디프테리아 독신(Diphtheria toxin) 등이 있으나, 이에 제한되지 않는다.
- [0114] 상기 벡터 제작은 당해 기술분야에서 잘 알려진 유전자 재조합 기술을 이용하여 제조할 수 있으며, 부위-특이적 DNA 절단 및 연결은 당해 기술 분야에서 일반적으로 알려진 제한효소 등을 사용할 수 있다.
- [0115] 본 발명의 벡터는 스트렙토마이세스 세포에서 작용하지만, 예를 들어 클로닝 또는 발현 목적을 위해 다른 박테리아 또는 진핵 세포로 또한 형질전환될 수 있다. 예를 들어, 아메리칸 타입 컬처 컬렉션(American Type Culture Collection; ATCC)로부터 구입할 수 있거나, 시판 중인 DH5α 균주와 같은 에스케리키아 콜라이 균주가 전형적으로 사용될 수 있다. 바람직한 진핵 숙주 세포로는 포유동물 세포, 곤충 세포 또는 효모 세포가 또한 효

과적으로 사용될 수 있다.

- [0116] 숙주 균주 내로 본 발명의 벡터를 도입 방법은 핵산을 세포 내로 도입하는 어떠한 방법도 포함되며, 당 분야에서 공지된 바와 같이 적합한 표준 기술을 선택하여 수행할 수 있다. 예를 들어, 원형질체(protoplast) 형질전환, 전기천공법(electroporation), 전기주입법(electroinjection), 미세주입법 (microinjection), 인산 칼슘공동-침전법(calcium phosphate co-precipitation), 염화칼슘/염화루비듐법, 레트로바이러스 감염(retroviral infection), DEAE-덱스트란(DEAE-dextran), 양이온 리포솜(cationic liposome)법, 폴리에틸렌글리콜 침전법(polyethylene glycol-mediated uptake), 유전자총(gene gun) 등을 이용할 수 있으나, 이에 제한되는 것은 아니다. 이 때 원형의 벡터를 적절한 제한효소로 절단하여 선형의 벡터 형태 또는 플라스미드를 제거한 선형의 벡터 형태로 도입할 수 있다. 형질전환체는 선별마커, 예를 들어 전술한 바와 같이 재조합 벡터와 연관된 항생물질 저항성을 발현시키는 세포를 선택하는 방법과 같은 표준 절차에 따라 선택할 수 있다.
- [0117] 이와 같이 제조된 재조합 균주를 배양하여 밀베마이신을 생산할 수 있으며, 예를 들어, 밀베마이신 A3, 밀베마이신 A4, 밀베마이신 $\alpha 2$, 밀베마이신 $\alpha 4$ 및 밀베마이신 G 로 이루어진 군에서 선택되는 1종 이상의 밀베마이신을 생산할 수 있다.
- [0118] 재조합 균주의 배양은 균주의 생육과 밀베마이신의 대량 생산에 적합하도록 온도, 배지의 pH 및 배양시간 등의 조건들을 적절하게 조절할 수 있다. 상기 배양 방법의 예에는, 회분식, 연속식 및 유가식 배양이 포함되나, 이에 제한되는 것은 아니다.
- [0119] 배양에 사용되는 배지는 특정한 균주의 요구조건을 적절하게 만족시켜야 한다. 상기 배지는 다양한 탄소원, 질소원, 인원 및 미량원소 성분을 포함할 수 있다. 발현 벡터가 유도성 촉진자를 포함하는 경우에는, 온도 변화, 영양원의 고갈, 무상 유도 물질(예를 들어, 이소프로필- β -D-티오갈락토피라노시드(IPTG)와 같은 탄수화물의 유사 물질)의 첨가, 과잉 대사 부산물의 축적 등과 같은 적절한 유도 조건은 발현을 유도하는데 필요에 따라 적용될 수 있다.
- [0120] 배지 내 탄소원으로는 글루코즈, 사카로즈, 락토즈, 프락토즈, 말토즈, 전분, 셀룰로스와 같은 당 및 탄수화물, 대두유, 해바라기유, 피마자유, 코코넛유 등과 같은 오일 및 지방, 팔미트산, 스테아린산, 리놀레산과 같은 지방산, 글리세롤, 에탄올과 같은 알코올, 아세트산과 같은 유기산을 예시할 수 있으나, 이에 제한되는 것은 아니다. 이들 물질은 개별적으로 또는 혼합물로서 사용될 수 있다. 배지 내 질소원으로는 펩톤, 효모 추출물, 육즙, 맥아 추출물, 옥수수 침지액, 대두밀 및 요소 또는 무기 화합물, 예를 들면 황산암모늄, 염화암모늄, 인산암모늄, 탄산암모늄 및 질산암모늄을 예시할 수 있으나, 이에 제한되는 것은 아니다. 질소원 또한 개별적으로 또는 혼합물로서 사용할 수 있다. 배지 내 인원으로는 인산이수소칼륨 또는 인산수소이칼륨 또는 상응하는 나트륨-함유 염을 예시할 수 있으나, 이에 제한되는 것은 아니다. 또한, 배양 배지는 성장에 필요한 황산마그네슘 또는 황산철과 같은 금속염을 포함하거나, 아미노산 및 비타민과 같은 필수 성장 물질을 포함할 수 있으나, 이에 제한되는 것은 아니다. 상기된 원료들은 배양 과정에서 배양물에 적절한 방식에 의해 회분식으로 또는 연속식으로 첨가될 수 있다.
- [0121] 또한, 필요에 따라, 수산화나트륨, 수산화칼륨, 암모니아와 같은 기초 화합물 또는 인산 또는 황산과 같은 산 화합물을 적절한 방식으로 사용하여 배양물의 pH를 조절할 수 있다. 또한, 지방산 폴리글리콜 에스테르와 같은 소포체를 사용하여 기포 생성을 억제할 수 있다. 호기 상태를 유지하기 위해 배양물 내로 산소 또는 산소-함유 기체(예, 공기)를 주입할 수 있으며, 배양물의 온도는 보통 20 내지 45, 바람직하게는 25 내지 40 일 수 있다. 배양은 원하는 밀베마이신의 생산량이 최대로 얻어질 때까지 계속될 수 있다.
- [0122] 재조합 미생물로부터 생산된 밀베마이신은, 당업계에 널리 알려져 있는 방법으로 세포, 세포 용해물 또는 배양 배지로부터 분리되거나 실질적으로 정제될 수 있다. 밀베마이신의 회수 방법의 예로서, 유기용매 추출법, 원심 분리, 초음파파쇄, 여과, 결정법(crystallization), 이온교환 크로마토그래피, 친화성 크로마토그래피, 고성능 액체 크로마토그래피(high performance liquid chromatography: HPLC), 가스 크로마토그래피(gas chromatography: GC) 등의 방법이 있으나, 이들 예에 한정되는 것은 아니다. 구체적으로, 균주 배양물로부터 유기용매를 이용하여 생성 물질들을 추출한 후, 실리카겔, 알루미나, 덱스트란겔, 이온교환 수지, 합성흡착제, 분자체, C_8H_{17} , $C_{18}H_{37}$, C_6H_5 등의 화학 결합형 실리카겔 등의 담체를 사용한 크로마토그래피에 부여하여, 얻어진 목적 화합물을 함유하는 분획을 농축 건조시키는 방법으로 회수할 수 있다.
- [0123] 본 발명의 다른 예는 앞서 설명한 재조합 스트렙토마이세스 아베르미틸리스 균주 또는 상기 균주의 배양물; 재조합 벡터; 및 아베멕틴 유도체 또는 이의 염으로 이루어진 군에서 선택된 1종 이상을 유효성분으로 포함하는

해충 방제용 조성물을 제공한다.

- [0124] 다른 예는 앞서 설명한 재조합 스트렙토마이세스 아베르미틸리스 균주 또는 상기 균주의 배양물; 재조합 벡터; 및 아베멕틴 유도체 또는 이의 염으로 이루어진 군에서 선택된 1종 이상을 작물에 적용하는 단계를 포함하는 해충 방제 방법을 제공한다.
- [0125] 상기 해충 방제는 해충에 대한 살충, 기피 등의 저해 효과를 포괄적으로 의미한다.
- [0126] 상기 해충은 동물 및/또는 식물에 위해를 가하는 모든 절지 동물의 성충, 유충 및 알을 총칭하는 것으로, 밀베마이신 또는 이의 유도체, 및/또는 아베멕틴 또는 이의 유도체가 방제 활성을 갖는 모든 해충의 성충 및/또는 유충일 수 있다. 예컨대, 상기 해충은 다음으로 이루어진 군에서 선택된 1종 이상일 수 있다:
- [0127] - 반시류 (Hemiptera) (예컨대, 다음을 포함함:
- [0128] 멸구류 (Delphacidae), 예컨대, 애멸구 (*Laodelphax striatellus*), 벼멸구 (*Nilaparvata lugens*), 흰등멸구 (*Sogatella furcifera*) 등; 매미충류 (Deltocephalidae), 예컨대, 끝동매미충 (*Nephotettix cincticeps*), 두점 끝동매미충 (*Nephotettix virescens*) 등; 진딧물류 (Aphididae), 예컨대, 목화진딧물 (*Aphis gossypii*), 복숭아혹진딧물 (*Myzus persicae*), 양배추가루진딧물 (*Brevicoryne brassicae*), 감자수염진딧물 (*Macrosiphum euphorbiae*), 찌리수염진딧물 (*Aulacorthum solani*), 기장태두리진딧물 (*Rhopalosiphum padi*), 굴소리진딧물 (*Toxoptera citricidus*) 등; 노린재류 (Pentatomidae), 예컨대, 풀색노린재 (*Nezara antennata*), 톱다리개미허리노린재 (*Riptortus clavatus*), 호리허리노린재 (*Leptocoris chinensis*), 가시점등글노린재 (*Eysarcoris parvus*), 썩덩나무노린재 (*Halyomorpha mista*), 장님노린재 (*Lysus lineolaris*) 등; 가루이류 (Aleyrodidae), 예컨대, 온실가루이 (*Trialeurodes vaporariorum*), 담배가루이 (*Bemisia tabaci*), 은빛잎가루이 (*Bemisia argentifolii*) 등; 깍지벌레류 (Coccidae), 예컨대, 캘리포니아붉은깍지벌레 (*Aonidiella aurantii*), 설희제깍지벌레 (*Comstockaspis perniciosus*), 화살깍지벌레 (*Unaspis citri*), 루비깍지벌레 (*Ceroplastes rubens*), 이세리아깍지벌레 (*Icerya purchasi*) 등; 방패벌레류 (Tingidae); 나무이류 (Psyllidae) 등);
- [0129] - 인시류 (Lepidoptera) (예컨대, 다음을 포함함: 명나방류 (Pyralidae), 예컨대, 이화명나방 (*Chilo suppressalis*), 옐로우 라이스 나무좀 (*Tryporyza incertulas*), 흑명나방 (*Cnaphalocrocis medinalis*), 목화명나방 (*Notarcha derogata*), 화랑곡나방 (*Plodia interpunctella*), 조명나방 (*Ostrinia furnacalis*), 유럽조명나방 (*Ostrinia nubilalis*), 배추순나방 (*Hellula undalis*), 잔디포충나방 (*Pediasia teterrellus*) 등; 밤나방류 (Noctuidae), 예컨대, 담배거세미나방 (*Spodoptera litura*), 과밤나방 (*Spodoptera exigua*), 멸강나방 (*Pseudaletia separata*), 도둑나방 (*Mamestra brassicae*), 검거세미나방 (*Agrotis ipsilon*), 가두배추금날개밤나비 (*Plusia nigrisigna*), 토리코플루시아류 (*Thoricoplusia* spp.), 담배밤나방류 (*Heliothis* spp.), 담배나방류 (*Helicoverpa* spp.) 등; 흰나비류 (Pieridae), 예컨대, 배추흰나비 (*Pieris rapae*) 등; 잎말이나방류 (Tortricidae), 예컨대, 애모무늬잎말이나방류 (*Adoxophyes* spp.), 복숭아순나방 (*Grapholita molesta*), 콩나방 (*Leguminivora glycinivorella*), 팔나방 (*Matsumuraesia azukivora*), 사과애모무늬잎말이나방 (*Adoxophyes orana fasciata*), 차애모무늬잎말이나방 (*Adoxophyes honmai*), 차잎말이나방 (*Homona magnanima*), 검모무늬잎말이나방 (*Archips fuscocupreanus*), 코드린나방 (*Cydia pomonella*) 등; 가는나방류 (Gracillariidae), 예컨대, 동백가는나방 (*Caloptilia theivora*), 사과굴나방 (*Phyllonorycter ringoniella*) 등; 심식나방류 (Carposinidae), 예컨대, 복숭아심식나방 (*Carposina niponensis*) 등; 굴나방류 (Lyonetiidae), 예컨대, 은무늬굴나방류 (*Lyonetia* spp.) 등; 독나방류 (Lymantriidae), 예컨대, 매미나방류 (*Lymantria* spp.), 독나방류 (*Euproctis* spp.) 등; 집나방류 (Yponomeutidae), 예컨대, 배추좀나방 (*Plutella xylostella*) 등; 뿔나방류 (Gelechiidae), 예컨대, 목화다래나방 (*Pectinophora gossypiella*), 감자뿔나방 (*Phthorimaea operculella*) 등; 불나방류 (Arctiidae), 예컨대, 미국흰불나방 (*Hyphantria cunea*) 등; 곡식좀나방류 (Tineidae), 예컨대, 옷좀나방 (*Tinea translucens*) 등);
- [0130] - 총채벌레목 (Thysanoptera), 예컨대, 다음을 포함함: 꽃노랑총채벌레 (*Frankliniella occidentalis*), 오이총채벌레 (*Thrips palmi*), 볼록총채벌레 (*Scirtothrips dorsalis*), 과총채벌레 (*Thrips tabaci*), 대만총채벌레 (*Frankliniella intonsa*), 담배총채벌레 (*Frankliniella fusca*) 등;
- [0131] - 쌍시류 (Diptera) (예컨대, 다음을 포함함:
- [0132] 굴파리류 (Agromyzidae), 예컨대, 고자리파리 (*Hylemya antiqua*), 씨고자리파리 (*Hylemya platura*), 벼잎굴파리 (*Agromyza oryzae*), 벼애잎굴파리 (*Hydrellia griseola*), 벼노랑굴파리 (*Chlorops oryzae*), 아메리카잎굴파리 (*Liriomyza trifolii*) 등; 뿔파리 (*Dacus cucurbitae*), 지중해과실파리 (*Ceratitidis capitata*) 등);

- [0133] - 초시류 (Coleoptera), 예컨대, 점박이 무당벌레 (*Epilachna vigintioctopunctata*), 오이잎벌레 (*Aulacophora femoralis*), 벼룩잎벌레 (*Phyllotreta striolata*), 벼잎벌레 (*Oulema oryzae*), 벼뿌리바구미 (*Echinocnemus squameus*), 벼물바구미 (*Lissorhoptus oryzophilus*), 목화바구미 (*Anthonomus grandis*), 팔바구미 (*Callosobruchus chinensis*), 헌팅 바구미 (*Sphenophorus venatus*), 왜콩풍뎡이 (*Popillia japonica*), 구리풍뎡이 (*Anomala cuprea*), 옥수수잎벌레류 (*Diabrotica* spp.), 콜로라도감자잎벌레 (*Leptinotarsa decemlineata*), 방아벌레류 (*Agriotes* spp.), 쉼벌레 (*Lasioderma serricorne*) 등;
- [0134] - 메뚜기류 (Orthoptera), 예컨대, 땅강아지 (*Gryllotalpa africana*), 벼메뚜기 (*Oxya yezoensis*), 벼메뚜기 (*Oxya japonica*) 등;
- [0135] - 막시류 (Hymenoptera), 예컨대, 무잎벌 (*Athalia rosae*), 가위개미(일개미)류 (*Acromyrmex* spp.), 불개미류 (*Solenopsis* spp.) 등;
- [0136] - 이목 (Phthiraptera), 예컨대, 소 이 (*Damalinea bovis*, *Haematopinus eurysternus*, *Linognathus vituli*), 돼지 이 (*Haematopinus suis*) 등;
- [0137] - 진드기류 (Acarina) (예컨대, 다음을 포함함:
- [0138] 응애류 (Tetranychidae), 예컨대, 점박이응애 (*Tetranychus urticae*), 굴응애 (*Panonychus citri*), 나무응애류 (*Oligonychus* spp.) 등; 흑응애류 (*Eriophyidae*), 예컨대, 굴녹응애 (*Aculops pelekassi*); 먼지응애류 (*Tarsonemidae*), 예컨대, 차면지응애 (*Polyphagotarsonemus latus*); 지응애류 (*Tenuipalpidae*); 치레응애류 (*tuckerellidae*); 가루진드기류 (*Acaridae*), 예컨대, 긴털가루진드기 (*Tyrophagus putrescentiae*) 등; 먼지진드기류 (*Pyroglyphidae*), 예컨대, 큰다리먼지진드기 (*Dermatophagoides farinae*), 세로무늬먼지진드기 (*Dermatophagoides pteronyssinus*) 등; 발톱진드기류 (*Cheyletidae*), 예컨대, 짧은빗살발톱진드기 (*Cheyletus eruditus*), 발톱진드기 (*Cheyletus malaccensis*), 케일레투스 무레이 (*Cheyletus moorei*) 등; 움진드기류 (*scabies mites*), 예컨대, 움진드기 (*Sarcoptes scabiei*) 등;
- [0139] 선충류 (Nematodes), 예컨대, 소나무재선충 (*Bursaphelenchus xylophilus*), 뿌리혹선충 (*Meloidogyne* spp.) 벼 잎선충 (*Aphelenchoides besseyi*), 딸기아선충 (*Nothotylenchus acris*), 개구충 (*Ancylostoma caninum*), 모두충 (*Capillaria* spp.), 개사상충 (*Dirofilaria immitis*), 가축에 기생하는 선충류 (roundworm) 등.
- [0140] 상기 작물은 해충에 의하여 피해를 입을 수 있는 모든 식물일 수 있다.
- [0141] 상기 작물에 적용하는 단계는 침지, 분사, 도포 등의 통상적인 모든 방법에 의하여 수행될 수 있다.
- [0142] 본 명세서에 기재된 유전자 서열 및 아미노산 서열은 기능성 동등성을 유지하는 한, 60% 이상, 65% 이상, 70% 이상, 75% 이상, 80% 이상, 85% 이상, 90% 이상, 92% 이상, 94% 이상, 96% 이상, 98% 이상, 또는 99% 이상의 서열 상동성을 갖는 경우를 포함하는 것으로 해석될 수 있다.

발명의 효과

- [0144] 본 발명에서 제공된 재조합 균주를 이용하여 밀베마이신을 대량으로 생산할 수 있으며, 이에 의해 생산된 밀베마이신은 농업 및 동물의약품 분야에 폭넓게 활용될 수 있다.

도면의 간단한 설명

- [0146] 도 1은 일 실시예에 따른 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739)의 milA1 유전자의 구성을 나타낸 것이다.
- 도 2는 일 실시예에 따른 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739)의 milA1 유전자를 확보하는 과정을 나타낸 것이다.
- 도 3은 pGEMB2-MA-DOWN-fu 벡터를 구성하는 염기서열을 나타낸 것이다.
- 도 4a 내지 도 4c는 일 실시예에 따른 pCR2.1-milA1-AC2 벡터를 제조하는 과정을 나타낸 것이다.
- 도 5는 일 실시예에 따른 pCR2.1-Adu-M1Age 벡터를 제조하는 과정을 나타낸 것이다.

도 6은 일 실시예에 따른 pCR2.1-Adu-milA1-Ac2 벡터를 제조하는 과정을 나타낸 것이다.

도 7a는 일 실시예에 따른 milA1 integration 벡터가 도입된 균주에서 교차를 통해 milA1 유전자가 삽입되는 과정을 나타낸 모식도이다.

도 7b는 일 실시예에 따른 milA1 integration 벡터가 도입된 균주에서 milA1이 성공적으로 삽입되었는지를 확인한 결과이다.

도 8은 일 실시예에 따른 *S. avermitilis* DBM-01 균주의 배양액으로부터 생성된 물질을 분석한 HPLC 결과를 나타낸다.

도 9a 및 도 9b는 일 실시예에 따른 pCR2.1-A3Mmo7A 및 pCR2.1-A3Mmo7B 벡터를 제조하는 과정을 나타낸 것이다.

도 10a는 일 실시예에 따른 모듈7 교체 벡터가 도입된 균주에서 교차를 통해 모듈7 유전자가 삽입되는 과정을 나타낸 모식도이다.

도 10b는 일 실시예에 따른 모듈7 교체 벡터가 도입된 균주에서 모듈 7 이 성공적으로 삽입되었는지를 확인한 결과이다.

도 11a 내지 도 11c는 실시예에 따른 DBM-03-A와 DBM-03-B가 생성한 물질을 분석한 HPLC 분석 결과를 보여준다. 형성된 각 peak 중 retention time이 표기된 peak는 Milbemycin 표준품과 동일한 UV-흡광패턴 보여주는 것이다. 도 11a는 밀베마이신 표준품의 HPLC 분석결과, 도 11b는 *S. avermitilis* DBM-03-A가 생성한 물질의 HPLC 분석결과, 도 11c는 *S. avermitilis* DBM-03-B가 생성한 물질의 HPLC 분석결과를 각각 나타낸다.

도 12는 일 실시예에 따른 모듈7 치환 균주의 배양액으로부터 생성된 물질을 분석한 HPLC 결과 및 각 피크에서 생성된 물질을 나타낸다.

도 13은 aveA3 유전자와 milA3 유전자의 AT-DH 모듈간 영역의 코딩 유전자 부위 및 KR-ACP 모듈간 영역의 코딩 유전자 부위의 염기서열 분석 결과를 보여준다.

도 14는 일 실시예에 따른 milA3의 모듈 7의 DH 도메인부터 모듈 9의 KR 도메인까지의 영역을 코딩하는 유전자 부위를 포함하는 pD7K9milA3 벡터의 제작 과정을 모식적으로 보여준다.

도 15는 일 실시예에 따른 milA3 삽입 벡터 pD7K9milA3가 삽입된 스트렙토마이세스 아베미틸리스 균주의 제작 과정을 모식적으로 보여준다.

도 16은 일 실시예에 따라서 제작된 *S. avermitilis* DBM-03-C 균주의 생성 물질의 HPLC 분석 결과를 보여주는 그래프이다.

도 17a 내지 17f는 일 실시예에 따라서 제작된 *S. avermitilis* DBM-03-C 균주의 생성 물질의 UPLC-qTOF-HR-MS 분석 결과를 보여주는 그래프이다 (도 17a: 밀베마이신 A3; 도 17b: 밀베마이신 A4; 도 17c: 밀베마이신 D; 도 17d: 밀베마이신 alpha2; 도 17e: 밀베마이신 alpha4; 도 17f: 밀베마이신 G).

도 18a-18c는 아베멕틴과 밀베마이신의 분자구조 차이 및 이를 유발하는 유전자의 구성을 보여주는 그림이고, 도 18d는 밀베마이신 폴리케타이드 합성효소 유전자군 (gene cluster)과 아베멕틴 폴리케타이드 합성효소 유전자군을 비교하여 보여주는 그림이다.

발명을 실시하기 위한 구체적인 내용

[0147] 이하, 본 발명을 실시예에 의해 상세히 설명한다. 단, 하기 실시예는 본 발명을 예시하는 것일 뿐, 본 발명이 하기 실시예에 의해 한정되는 것은 아니다.

[0149] 실시예 1. milA1 삽입 벡터(milA1 integration vector) 제작

[0150] 밀베마이신 생산 균주인 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739)의 milA1 유전자를 *S. avermitilis* SA-01의 aveA1에 치환하기 위한 벡터를 제조하였다. 상기 벡터는 milA1 유전자 전체를 포함하고, 통합을 위한 상동성을 가지는 aveA1 유전자의 upstream 및 downstream 지역이 milA1의 양쪽 말단에 붙여진 형태로 구성하였다. 이하, 벡터 제작 과정을 상세히 설명한다.

[0152] **1-1. milA1 유전자 확보**

[0153] 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739) genomic DNA로부터 milA1 유전자를 확보하였다. 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739)의 경우 밀베마이신 생합성 유전자들의 DNA 염기서열이 알려져 있지 않기 때문에, NCBI에 등록된 스트렙토마이세스 난창젠시스(*Streptomyces nanchangensis*; Accession no. FJ952082) 또는 스트렙토마이세스 빙청젠시스(*Streptomyces bingchenggensis*; Accession no. CP002047.1)의 밀베마이신 생합성 유전자군의 염기서열을 기반으로 프라이머를 디자인하여 유전자를 증폭하여 이용하였다. 또한, 밀베마이신 생합성에 관여하는 폴리케타이드 생합성효소의 유전자는 매우 크고 높은 GC비율을 갖기 때문에 한 번에 증폭하는 것은 매우 어려운 일로 이를 극복하기 위해 유전자를 임의의 두 구역으로 나눠 증폭하였다 (도 1).

[0154] 우선, milA1 유전자의 N-terminus쪽의 단편을 확보하기 위하여, 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739)의 genomic DNA를 A1NT2-F 프라이머(GAACCGTGTGCGGAAGAA: 서열번호 1) 및 A1CT1-R 프라이머(ATCAGGTCGGTGAGACGGT: 서열번호 2)를 사용하여 PCR(95도에서 5분; 98도에서 20초, 60도에서 20초 및 68도에서 10분으로 30 사이클; 및 68도에서 10분)을 통해 증폭한 후 pCR2.1 벡터(TA 클로닝 벡터)에 클로닝하였다. 이 벡터를 pCR2.1-NM으로 명명하였다. 다음으로, milA1 유전자의 C-terminus쪽의 단편을 확보하기 위하여, 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739)의 genomic DNA를 A1CT1-F 프라이머(GCAAGGTCGTCTTTGTGTGC: 서열번호 3) 및 A1CT2-R 프라이머(CCACCAACTCGATATGGGC: 서열번호 4)를 사용하여 PCR(95도에서 5분; 98도에서 20초, 60도에서 20초 및 68도에서 10분으로 30 사이클; 및 68도에서 10분)을 통해 증폭한 후 pCR2.1 벡터(TA 클로닝 벡터)에 클로닝하였다. 이 벡터를 pCR2.1-CT으로 명명하였다. 상기 제조한 pCR-NM과 pCR-CT 벡터를 각각 NcoI 제한효소로 처리하여 단편(NM fragment, CT fragment)을 확보한 후(도 1), 이들 단편을 연결(ligation)하여 하나의 벡터를 제조하였으며, 이를 pCR2.1-milA1으로 명명하였다(도 2; milA1 gene 서열 (서열번호: 70) 포함). 이 때, 연결된 단편의 ORF의 염기서열을 확인한 결과, 스트렙토마이세스 빙청젠시스(*Streptomyces bingchenggensis*)의 염기서열과 100% 일치하는 것을 확인하였다.

[0156] **1-2. C-terminus쪽 arm 제작**

[0157] milA1의 모듈 2의 위치에 해당하는 일부 DNA sequence와, aveA1 docking domain으로 추정되는 부위, 및 aveA2 sequence를 포함하는 DNA를 인공적으로 합성하였고, 합성된 DNA는 pGEM vector에 도입하여 pGEM2-MA_DOWN_fu를 제작하였다(도 3; 서열번호 35 내지 서열번호 37). 상기 제작된 pGEM2-MA_DOWN_fu를 BamHI/HindIII 제한효소로 잘라 insert를 분리한 후, BamHI/HindIII로 자른 pCR2.1-CT에 연결하여 milA1의 C-terminal 부위가 aveA1의 C-terminal 부위로 교체된 pCR2.1-CT-Ac 벡터를 제작하였다.

[0158] pCR2.1-CT-Ac 벡터의 결합부위를 연장하기 위하여, *S. avermitilis* SA-01균주의 genomic DNA로부터 A2-404F 프라이머(GGCGCTGGCAATGACC: 서열번호 5) 및 A2PH-1500R 프라이머(GATTACGCCAAGCTTGGCTGATTCTCCCGGACG: 서열번호 6)를 이용하여 aveA2의 N-terminus region을 PCR 증폭하고, 이를 infusion ligation을 통해 pCR2.1-CT-AC의 3' 말단에 결합하여, pCR2.1-CT-Ac2 벡터를 제작하였다.

[0159] pCR2.1-CT-Ac2 벡터와 pCR2.1-NM 벡터를 각각 NdeI/XbaI 제한효소로 처리한 후 결합하여, milA1 유전자의 C-terminal부분이 aveA1의 것으로 치환된 milA1 유전자를 갖는 pCR2.1-milA1-Ac2 벡터를 제작하였다(도 4a 내지 도 4c 참조).

[0161] **1-3. N-terminus쪽 arm 제작**

[0162] *S. avermitilis* SA-01 균주의 genomic DNA로부터 AveD2-F 프라이머(GATGACGGAATCCCTGG: 서열번호 7), AveD2-R 프라이머(AGAATGAAAGGAGCGCGA: 서열번호 8), MA1NU-F 프라이머(GAGGACACAGCTCAGTTGCCCAAGCCGAGAAC: 서열번호 9), M1Age-R 프라이머(TGATCGAAGAGCCGGGAGT: 서열번호 10), PHR-M1Ag 프라이머(CCGGCTCTTCGATCAAGCTTGGCGTAATCATGG: 서열번호 11) 및 PXF-AD2 프라이머(GGGATTTCGTCATCAAGCTTGGCGTAATCATGG: 서열번호 12)를 이용하여 N-terminus arm에 해당하는 3개의 단편을 각각 PCR 증폭하고, 이를 infusion ligation을 통해 결합하여 pCR2.1-Adu-M1Age 벡터를 제작하였다(도 5).

[0164] 1-4. milA1 유전자에 양쪽 arm이 결합된 형태의 vector 제작

[0165] pCR2.1-milA1-Ac2 벡터와 pCR2.1-Adu-M1Age 벡터를 XmnI 제한효소로 처리한 후 이를 infusion ligation을 통해 결합하여, pCR2.1-Adu-milA1-Ac2 벡터 (서열번호 79)를 제작하였다(도 6). 이를 아베택틴 생산균주로 도입하기 위해 스트랩토마이시스-대장균 서플렉터인 pKC1139 벡터(M. Bierman et al., Gene, 116:43-49)와 pCR2.1-Adu-milA1-Ac2 벡터를 각각 XbaI, HindIII 제한효소로 처리한 후 연결(ligation)하여 pKC1139에 Adu-milA1-Ac2 구조가 들어가 있는 형태인, pKC-Adu-milA1-Ac2 벡터를 제작하였다.

[0167] 실시예 2. milA1 치환 균주 제작 및 생성물질 확인

[0168] milA1 integration 벡터를 S. avertmils SA-01 균주 내로 도입시키고, 상동재조합이 일어나도록 유도하여 S. avertmils SA-01 균주의 aveA1 유전자가 milA1 유전자로 치환된 형태의 균주를 확보하고, 배양을 통해 생성되는 신규물질을 확인하였다.

[0170] 2-1. 접합법을 통한 형질전환

[0171] pKC-Adu-milA1-Ac2 벡터를 포함하고 있는 E. coli 균주(ET12567/pUZ8002 strain)를 chloramphenicol, kanamycin, apramycin을 넣은 LB 액체배지 3ml에 접종한 후 Shaking incubator에서 37, 200rpm 조건으로 18~24 시간 배양하였다. 멸균된 250ml flask에 25ml LB broth를 위의 항생제와 함께 넣고, 배양액 250ul을 접종한 뒤 동일 조건으로 3시간 배양하였다. OD가 0.4~0.5에 이르렀을 때, 배양을 종료하고 5000rpm 에서 5분간 원심분리하였고(4 조건), Pellet만 취하여 냉각된 LB broth 1ml에 풀어주고 세척하여 균체 현탁액을 확보하였다. 또한, S. avertmils SA-01 균주를 germination 하기 위하여, ISP4 agar plate에 host로 사용할 방선균을 7일 이상 배양하고, Spore를 loop로 긁어서 2X TY 배지에 풀어주고, 50 heating block에 넣고 10분간 activation을 실시하였다. 이와 같이 준비된 E. coli 균주 500ul과 germination된 host 균주 S. avertmils SA-01 500ul을 microtube에 넣고 inverting으로 섞고, 8000rpm 에서 2분간 원심분리하여 상층액을 제거하는 방법으로 접합(conjugation)을 유도하여, pKC-Adu-milA1-Ac2 벡터를 S. avertmils SA-01 균주 내로 도입시켰다.

[0173] 2-2. 교차에 의한 유전자 치환

[0174] 상기와 같이 pKC-Adu-milA1-Ac2(Et) 벡터가 도입된 균주에서 1차 교차를 유도하여 벡터 전체가 균주의 염색체에 삽입되도록 하였다. 이를 위하여, apramycin이 25ug/ml 농도로 첨가된 ISP-4 고체배지를 제조한 후, pKC-Adu-milA1-Ac2 벡터 도입에 의해 생성된 colony를 취한 후, 준비된 고체 배지에 streaking 하였다. 이후 고온에 민감한 pKC1139 벡터의 성질을 이용하여 벡터가 복제 되는 것을 막기 위하여 37에서 7일간 배양하였다.

[0176] 2-3.2차 교차

[0177] 상기 1차 교차가 일어난 균주에서 aveA1 유전자를 milA1 유전자로 치환하기 위해 2차 교차를 유도하였다(도 7). 이를 위하여, apramycin이 25ug/ml 농도로 첨가된 ISP-4 고체배지와 항생제가 첨가 되지 않은 배지를 각각 제조하고, 1차 교차를 유도한 균주를 ISP-4 고체 배지에 도말하여 2차 교차를 유도한 후, apramycin 저항성 유무로 후보 균주를 선발하였다. 스크리닝을 통해 확보한 후보 균주의 genomic DNA를 추출하여, aveA1, milA1 일부지역을 확인할 수 있는 프라이머(표 1)를 사용하여 유전자의 교체가 정확하게 이루어졌는지 확인하였다. 이와 같이 최종 균주를 선발하고 S. avermitilis DBM-01로 명명하였으며, KCTC(Korean Collection for Type Cultures, 한국생명공학연구원 미생물자원센터)에 기탁하여 2015년 8월 31일자로 기탁번호 KCTC12889BP 를 부여받았다.

표 1

Primer Name	Sequence	용도
aveFi-F	GGCAACACCAACCGTGAGAT(서열번호 13)	N-terminus fusion 확인
milA1i-R	GATGACCGAGGAGTGAACG(서열번호 14)	
milA1i-F	GTACTGGACCTGTTGAGCG(서열번호 15)	C-terminus fusion 확인

aveA2i-R2	ATGAATGCGGTCTTTCCTGC(서열번호 16)	aveA1 일부지역 확인
aveA1C-F	GGACTCGCTGATTCGGACAA(서열번호 17)	
aveA1C-R	GGAGCAGATGAATGCCCTCT(서열번호 18)	
milA1C-F	CCTGTGGGAGTTCATCACCG(서열번호 19)	milA1 일부지역 확인
milA1C-R	CGAGATTGTGGTGGACGAGG(서열번호 20)	

[0180] 2-4.S. avermitilis DBM-01의 생성물질 확인

[0181] S. avermitilis DBM-01 균주를 배양하기 위한 종균 배양용 배지와 생산 배양용 배지를 각각 제조하였다. 종균 배지를 제조하기 위하여, soluble starch 30g/L(Junsei, Japan), yeast extract 15g/L(Duchefa, 네덜란드), KH₂PO₄ 0.4g/L(Junsei)의 농도가 되도록 적정량의 DW에 혼합한 후, pH를 7.2로 맞추고 121, 15분간 고압살균을 수행하였다. 충분히 식은 후, 멸살한 Corn steep liquor(Sigma)를 5g/L의 농도가 되도록 첨가하였다. 또한, 생산 배지를 제조하기 위하여, soluble starch 80g/L, soybean meal 10g/L(Sigma), skim milk 15g/L(Difco), KH₂PO₄ 0.5 g/L의 농도가 되도록 적정량의 DW에 혼합한 후, pH를 7.2로 맞추고, 121, 15분간 고압살균을 실시하였다. 생산 배양에 접종할 균주의 양을 충분히 만들기 위해, 상기 제조한 종균 배지 25ml을 멸균된 250ml baffled flask에 넣고, 균주 mycelium을 agar plate로부터 1loop씩 떼서 종균 배지에 접종하였다. 230 rpm 및 28도 조건하에서 48시간 동안 배양하였다. 또한, 생산배지 25ml을 멸균된 250ml non-baffled flask 넣고 종균 배양액 1.25ml을 접종한 후, 230 rpm 및 28도 조건하에서 10일 동안 배양하였다. 생산 배양액으로부터 생성된 물질을 추출하기 위하여, 유기용매(acetonitrile : MeOH = 1 : 1 혼합 용액) 1.5ml과 배양액 0.5ml을 섞고 충분히 vortexing하였다. 5분간 sonication 한 후, 1시간 동안 교반하였다. 12000rpm에서 5분간 원심분리한 후 상등액만 취하여 분석하였다.

[0182] 추출한 시료를 아래의 조건 6에서 HPLC/UVD 분석을 통해 분석하였다.

[0183] - Mobile phase : Acetonitrile/Water(v/v)= 40/60, v/v(5min), 90/10, v/v(30min), 90/10, v/v(10min)

[0184] - Flow rate : 1 ml/min

[0185] - wavelength : 245 nm

[0186] - Run time : 45 min

[0187] 또한, DBM-01을 배양하여 나온 생성 물질 중 신규한 물질 4종(도 8에서 IV1, IV2, IV3, IV5로 표시함)에 대해 NMR 분석을 실시하여 구조를 분석한 결과는 다음과 같다:

[0188] 600 MHz in CD₃OD(the internal references residual solvent protons and carbons; δ_H 4.78, δ_C 49.15)

[0189] IV1: δ_H : 3.16(m, H-2), 5.35(d, J = 1.4, H-3), 1.75(s, 3H, H-4a), 4.16(m, H-5), 3.70(m, H-6), 4.55(m, H-8a), 5.74(dd, J = 1.9 & 11.3, H-9), 5.81(ddd, J = 3.4 & 11.4 & 14.2, H-10), 5.65(ddd, J = 1.9 & 9.7 & 14.8, H-11), 2.59(m, H-12), 1.09(m, H-12a), 3.93(m, H-13), 1.48(s, 3H, H-14a), 5.15(m, H-15), 2.24(m, H-16), 3.63(m, H-17), 0.75(m, H-18), 1.86(m, H-18), 4.98(m, H-19), 1.14(m, H-20), 2.11(m, H-20), 1.24(m, H-22), 1.44(m, H-23), 1.52(m, H-23), 1.18(m, H-24), 0.77(d, J = 6.6, 3H, H-24a), 3.25(m, H-25), 1.08(m, H-26), 4.76(m, H-1'), 2.23(m, H-2'), 3.46(m, H-3'), 3.34(s, 3H, 3'-OMe), 3.12(m, H-4'), 3.76(m, H-5'), 1.17(m, H-6'), 5.26(d, J = 3.2, H-1''), 1.57(m, H-2''), 3.36(H-3''), 3.40(s, 3H, 3''-OMe), 3.64(m, H-4''), 2.97(m, H-5''), 1.15(m, H-6'') δ_C : 173.6(C-1), 47.1(C-2), 120.3(C-3), 137.2(C-4), 19.9(C-4a), 69.1(C-5), 82.3(C-6), 82.2(C-7), 141.9(C-8), 68.7(C-8a), 121.8(C-9), 126.7(C-10), 138.4(C-11), 41.2(C-12), 21.0(C-12a), 82.7(C-13), 136.3(C-14), 15.4(C-14a), 120.1(C-15), 35.2(C-16), 69.1(C-17), 38.1(C-18), 69.9(C-19), 42.7(C-20), 99.1(C-21), 30.7(C-22), 36.9(C-23), 38.0(C-24), 18.2(C-24a), 72.7(C-25), 19.8(C-26), 96.3(C-1'), 36.2(C-2'), 79.7(C-3'), 57.0(3'-OMe), 82.3(C-4'), 69.9(C-5'), 19.1(C-6'), 99.9(C-1''), 37.3(C-2''), 79.5(C-3''), 57.6(3''-OMe), 69.2(C-4''), 77.1(C-5''), 18.3(C-6'').
HRESQTOF m/z 855.4507 [M+Na]⁺(calcd for C₄₅H₆₈O₁₄Na)

[0191] IV2: δ_{H} : 3.16(m, H-2), 5.36(d, $J = 1.5$, H-3), 1.75(s, 3H, H-4a), 4.17(m, H-5), 3.71(m, H-6), 4.55(m, H-8a), 5.75(dt, $J = 2.2$ & 11.3, H-9), 5.82(dd, $J = 11.2$ & 14.9, H-10), 5.65(dd, $J = 9.8$ & 14.9, H-11), 2.59(m, H-12), 1.10(d, $J = 7.0$, H-12a), 3.92(m, H-13), 1.48(s, 3H, H-14a), 5.13(m, H-15), 2.44(m, H-16), 3.62(m, H-17), 0.77(m, H-18), 1.85(m, H-18), 5.00(m, H-19), 1.16(m, H-20), 2.13(m, H-20), 1.26(m, H-22), 1.66(m, H-22), 1.15(m, H-23), 1.43(m, H-23), 1.24(m, H-24), 0.76(d, $J = 6.5$, 3H, H-24a), 3.05(m, H-25), 1.29(m, H-26), 0.93(t, $J = 7.3$, H-27), 4.75(m, H-1'), 2.24(m, H-2'), 3.56(m, H-3'), 3.32(s, 3H, 3'-OMe), 3.12(m, H-4'), 3.65(m, H-5'), 1.18(d, $J = 6.2$, H-6'), 5.26(d, $J = 3.2$, H-1''), 1.56(m, H-2''), 3.35(H-3''), 3.35(s, 3H, 3''-OMe), 3.64(m, H-4''), 2.94(m, H-5''), 1.15(d, $J = 6.2$, H-6'') δ_{C} : 173.6(C-1), 47.1(C-2), 120.1(C-3), 137.2(C-4), 19.9(C-4a), 69.1(C-5), 82.2(C-6), 82.0(C-7), 142.0(C-8), 68.6(C-8a), 121.8(C-9), 126.7(C-10), 138.4(C-11), 41.2(C-12), 20.9(C-12a), 83.2(C-13), 136.5(C-14), 15.3(C-14a), 120.3(C-15), 35.1(C-16), 69.1(C-17), 38.1(C-18), 70.1(C-19), 42.9(C-20), 99.0(C-21), 30.9(C-22), 29.2(C-23), 35.6(C-24), 18.2(C-24a), 77.5(C-25), 26.9(C-26), 10.6(C-27), 96.4(C-1'), 36.2(C-2'), 80.9(C-3'), 56.9(3'-OMe), 82.2(C-4'), 70.1(C-5'), 19.1(C-6'), 99.9(C-1''), 36.8(C-2''), 79.4(C-3''), 57.6(3''-OMe), 70.2(C-4''), 77.5(C-5''), 18.2(C-6''). HRESQTOF m/z 869.4663 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{46}\text{H}_{70}\text{O}_{14}\text{Na}$)

[0193] IV3: δ_{H} : 3.16(m, H-2), 5.42(d, $J = 1.4$, H-3), 1.74(s, 3H, H-4a), 3.92(m, H-5), 3.41(s, 5-OMe), 3.90(m, H-6), 4.58(m, H-8a), 5.77(dd, $J = 1.9$ & 11.3, H-9), 5.85(ddd, $J = 3.4$ & 11.4 & 14.2, H-10), 5.68(ddd, $J = 1.9$ & 9.7 & 14.8, H-11), 2.59(m, H-12), 1.09(m, H-12a), 3.90(m, H-13), 1.48(s, 3H, H-14a), 5.17(m, H-15), 2.24(m, H-16), 3.63(m, H-17), 0.75(m, H-18), 1.86(m, H-18), 5.00(m, H-19), 1.14(m, H-20), 2.11(m, H-20), 1.24(m, H-22), 1.44(m, H-23), 1.52(m, H-23), 1.18(m, H-24), 0.77(d, $J = 6.6$, 3H, H-24a), 3.25(m, H-25), 1.08(m, H-26), 4.76(m, H-1'), 2.23(m, H-2'), 3.46(m, H-3'), 3.34(s, 3H, 3'-OMe), 3.13(m, H-4'), 3.81(m, H-5'), 1.17(m, H-6'), 5.28(d, $J = 3.2$, H-1''), 1.57(m, H-2''), 3.36(H-3''), 3.40(s, 3H, 3''-OMe), 3.64(m, H-4''), 2.97(m, H-5''), 1.15(m, H-6'') δ_{C} : 173.4(C-1), 47.0(C-2), 121.4(C-3), 135.4(C-4), 20.0(C-4a), 78.56(C-5), 57.8(5-OMe), 80.9(C-6), 82.2(C-7), 141.5(C-8), 68.7(C-8a), 121.9(C-9), 126.7(C-10), 138.5(C-11), 41.2(C-12), 21.0(C-12a), 82.7(C-13), 136.3(C-14), 15.4(C-14a), 120.1(C-15), 35.2(C-16), 69.1(C-17), 38.1(C-18), 69.9(C-19), 42.7(C-20), 99.1(C-21), 30.7(C-22), 36.9(C-23), 38.0(C-24), 18.2(C-24a), 72.7(C-25), 19.8(C-26), 96.3(C-1'), 36.2(C-2'), 79.7(C-3'), 57.0(3'-OMe), 82.3(C-4'), 69.9(C-5'), 19.1(C-6'), 99.9(C-1''), 37.3(C-2''), 79.5(C-3''), 57.6(3''-OMe), 69.2(C-4''), 77.1(C-5''), 18.3(C-6''). HRESQTOF m/z 869.4662 $[\text{M}+\text{Na}]^+$ (calcd for $\text{C}_{46}\text{H}_{70}\text{O}_{14}\text{Na}$)

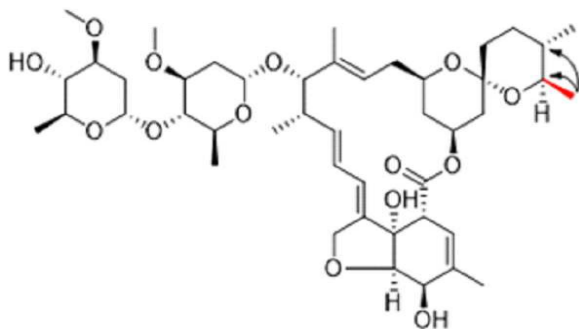
[0195] IV5: δ_{H} : 3.16(m, H-2), 5.41(d, $J = 1.5$, H-3), 1.74(s, 3H, H-4a), 3.93(m, H-5), 3.41(s, 5-OMe), 3.91(m, H-6), 4.57(m, H-8a), 5.76(dt, $J = 2.2$ & 11.3, H-9), 5.82(dd, $J = 11.2$ & 14.9, H-10), 5.66(dd, $J = 9.8$ & 14.9, H-11), 2.59(m, H-12), 1.11(d, $J = 7.0$, H-12a), 3.92(m, H-13), 1.47(s, 3H, H-14a), 5.13(m, H-15), 2.43(m, H-16), 3.62(m, H-17), 0.78(m, H-18), 1.85(m, H-18), 5.00(m, H-19), 1.15(m, H-20), 2.13(m, H-20), 1.25(m, H-22), 1.66(m, H-22), 1.14(m, H-23), 1.43(m, H-23), 1.24(m, H-24), 0.76(d, $J = 6.5$, 3H, H-24a), 3.04(m, H-25), 1.29(m, H-26), 0.93(t, $J = 7.3$, H-27), 4.75(m, H-1'), 2.24(m, H-2'), 3.56(m, H-3'), 3.32(s, 3H, 3'-OMe), 3.12(m, H-4'), 3.65(m, H-5'), 1.18(d, $J = 6.2$, H-6'), 5.26(d, $J = 3.2$, H-1''), 1.56(m, H-2''), 3.35(H-3''), 3.35(s, 3H, 3''-OMe), 3.64(m, H-4''), 2.94(m, H-5''), 1.15(d, $J = 6.2$, H-6'') δ_{C} : 173.5(C-1), 47.1(C-2), 121.5(C-3), 135.3(C-4), 20.2(C-4a), 78.4(C-5), 57.7(5-OMe), 80.8(C-6), 82.1(C-7), 141.4(C-8), 68.8(C-8a), 121.9(C-9), 126.8(C-10), 138.3(C-11), 41.2(C-12), 21.0(C-12a), 83.3(C-13), 136.4(C-14), 15.3(C-14a), 120.4(C-15), 35.2(C-16), 69.1(C-17), 38.0(C-18), 70.2(C-19), 42.7(C-20), 99.1(C-21), 30.7(C-22), 29.2(C-23), 35.6(C-24), 18.2(C-24a), 77.5(C-25), 26.8(C-26), 10.6(C-27), 96.4(C-1'), 36.2(C-2'), 80.9(C-3'), 56.9(3'-OMe), 82.2(C-4'), 70.1(C-5'), 19.1(C-6'), 99.9(C-1''), 36.8(C-2''), 79.4(C-3''), 57.6(3''-OMe), 70.2(C-4''), 77.5(C-5''), 18.2(C-6'').

HRESQTOF m/z 883.4823 $[M+Na]^+$ (calcd for $C_{47}H_{72}O_{14}Na$)

[0197] Ivermectin 유도체 1(IV1)의 구조는 화학식 1과 같이 결정하였으며, HMBC correlation을 화살표로 표시하였다.

[0198] [화학식 1]

[0199] 22,23-dihydro-25-methyl-ivermectin

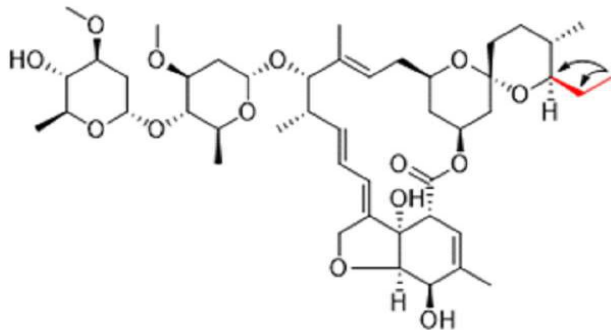


[0200]

[0201] Ivermectin 유도체 2(IV2)의 구조는 화학식 2와 같이 결정하였으며, HMBC correlation을 화살표로 표시하였다.

[0202] [화학식 2]

[0203] 22,23-dihydro-25-ethyl-ivermectin

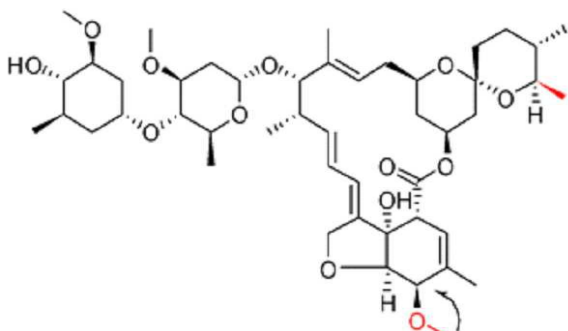


[0204]

[0205] Ivermectin 유도체 3(IV3)의 구조는 화학식 3과 같이 결정하였으며, HMBC correlation을 화살표로 표시하였다.

[0206] [화학식 3]

[0207] 5-methoxy-22,23-dihydro-25-methyl-ivermectin

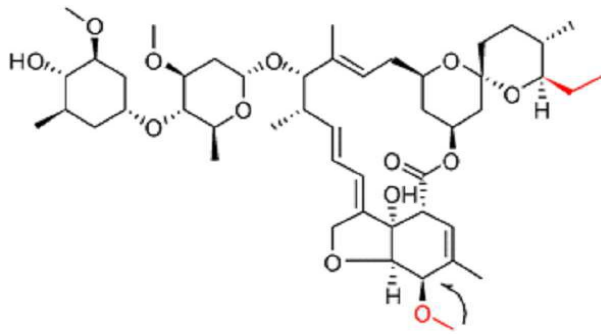


[0208]

[0209] Ivermectin 유도체 5(IV5)의 구조는 화학식 4와 같이 결정하였으며, HMBC correlation을 화살표로 표시하였다.

[0210] [화학식 4]

[0211] 5-methoxy-22,23-dihydro-25-ethyl-avermectin



[0212]

[0213] 결과적으로, 상기 물질의 분석을 통해 avermectin 생산 균주인 *S. avermitilis* SA-01의 avermectin 생산 유전자군의 첫번째 유전자인 *aveA1*을 *S. hygroscopicus* subsp. *aureolacrimosus*의 밀베마이신 생산 유전자군의 첫번째 유전자인 *milA1*으로 교체하였을 때, 숙주 균주와 다른 형태의 물질 생산 패턴을 나타냄을 확인할 수 있었으며, *milA1*이 삽입된 균주는 각각 22,23-dihydro-25-methyl-avermectin (CAS No. 137330-81-5; Ref: DE4031039), 22,23-dihydro-25-ethyl-avermectin (CAS No. 71972-13-9; Ref: FR2387231), 22,23-dihydro-25-isopropyl-avermectin(즉, Ivermectin B1b, CAS NO. 70209-81-3; Ref. US4199569), 5-methoxy-22,23-dihydro-25-methyl-avermectin, 5-methoxy-22,23-dihydro-25-ethyl-avermectin, 및 5-methoxy-22,23-dihydro-25-isopropyl-avermectin (즉, Ivermectin A1b; Ref. US4199569)을 생성함을 알 수 있었다.

[0214] 이들 중, 5-methoxy-22,23-dihydro-25-methyl-avermectin (화학식 3; IV3) 및 5-methoxy-22,23-dihydro-25-ethyl-avermectin (화학식 4; IV5)의 경우, 물질 구조를 이용한 SciFinder 검색에서 일치하는 구조를 찾지 못하였는 바, 신규 화합물인 것으로 판단되었다.

[0216] 2-5. *S. avermitilis* DBM-01 생성물질인 IV3와 IV5의 생물학적 활성

[0217] 상기 실시예 2-4에서 신규 화합물로 확인된 IV3 (화학식 3) 및 IV4 (화학식 4)의 생물학적 활성을 평가하여, 상기 화합물류가 해충의 유충 및 성충, 특히 점박이응애와 배추좀나방의 해충 구제 (살충)에 활성효과가 있음을 확인하였다.

[0218] 상기 화합물류의 살충 활성을 측정하기 위하여 다음과 같은 실험을 수행하였다. 상기 시험 화합물 IV3 및 IV4을 각각 용매인 acetone에 용해시키고 농도가 100ppm이 되도록 용해시켜 stock를 만든 후, 각 시험에서의 처리 농도 (표 2 및 표 3 참조)가 되도록 희석하여 사용하였다. 점박이 응애의 경우, 성충을 대상으로 하였으며, 상기 준비된 희석액을 배추 (Φ5cm/disk)에 1ml을 spray 처리하고, 1 내지 2일 후 생충수를 측정하였다. 이 경우, 대조약제는 아바멕틴(Sigma)을 사용하였다. 배추좀나방의 경우, 2령충을 대상으로 하였으며, 배추 (Φ5cm/disk) 절편에 침지 처리한 후 시험충을 절편당 10마리 접종하여 1 내지 2일 후 생충수를 측정 하였다. 이 경우, 대조약제는 에바멕틴(Sigma)를 사용하였다.

[0219] 상기 얻어진 결과를 아래의 표 그 결과를 표 2 및 표 3에 각각 나타내었다.

표 2

[0220] 점박이 응애 성충에 대한 IV3 및 IV5의 생물학적 활성 (살충률)

물질	처리 농도	처리기간	
		1일차	2일차
IV3	0.6 ppm	100.0	100.0
IV5		76.7	86.7
abamectin		100.0	100.0

표 3

[0221]

배추좀나방 유충에 대한 IV3, IV5의 생물학적 활성 (살충률)

물질	처리 농도	처리기간 (2일차)
IV3	0.06 ppm	100
IV5		100
Abamectin		96.7

[0222]

상기 표 3 및 표 4에 나타난 바와 같이, 화합물 IV3 및 IV4는 대조약물과 비교하여 동등 이상의 해충의 유충 및 성충에 대한 방제 활성을 나타냄을 확인할 수 있다.

[0224]

실시예 3. 모듈 7 교체 벡터 제작

[0225]

스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739)의 milA3 유전자에 존재하는 7번 모듈을 *S. avermitilis* DBM-01의 물질 생산 유전자군의 7번 모듈과 치환하기 위한 벡터를 제조하였다. 상기 벡터는 7번 모듈 전체 또는 일부를 포함하고, 통합을 위한 상동성을 가지는 aveA3, aveA4 유전자의 지역이 양쪽 말단에 붙여진 형태로 구성하였다. 이하, 벡터 제작 과정을 상세히 설명한다.

[0227]

3-1. 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739) genomic DNA로부터 모듈7 유전자 확보

[0228]

모듈 7 전체와 모듈 7 중 KS(Ketosynthase) 도메인만 숙주 균주인 *S. avermitilis* DBM-01의 것을 사용한 두 가지 형태로 모듈의 치환을 계획하고 이에 따른 유전자를 확보하고자 하였다. 앞서 milA의 클로닝 경우와 마찬가지로 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739)의 모듈 7의 염기서열이 알려져 있지 않기 때문에, NCBI에 등록된 스트렙토마이세스 난찬젠시스(*Streptomyces nanchangensis*) 또는 스트렙토마이세스 빙청젠시스(*Streptomyces bingchenggensis*)의 milA3의 염기서열을 활용하여 프라이머를 디자인하였다.

[0229]

스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739) genomic DNA로부터 모듈7에 해당하는 부위를 두 가지 형태로(A 및 B)로 클로닝하기 위하여, MA3-6503R 프라이머(TGCCAATGATCGCGATGGCCTCATC: 서열번호 21), MA3-104Fin(A) 프라이머(GGGTCAGTGGCAGCAACGCCATGTGATCCTCG: 서열번호 22) 및 MA3-1368Ffu(B) 프라이머(GGAGGCGTCCAGCAACCCATCGCCATCATC: 서열번호 23)를 사용하여 PCR(95도에서 5분; 98도에서 20초, 60도에서 20초 및 68도에서 6분으로 30 사이클; 및 68도에서 10분)을 통해 증폭한 후 pCR2.1 벡터(TA 클로닝 벡터)에 클로닝하였다. 이 벡터를 pCR2.1-Mo7A, pCR2.1-Mo7B로 명명하였다 (도 9a 및 도 9b).

[0231]

3-2. 양쪽 arm 제작

[0232]

S. avermitilis DBM-01 균주의 모듈7 지역에 교차를 일으키기 위한 결합부위를 제작하였다. *S. avermitilis* DBM-01 균주의 genomic DNA를 아래 프라이머를 사용하여 증폭하여 양쪽 arm에 해당하는 부위의 fragment를 확보한 후, pCR2.1 벡터(TA 클로닝 벡터)에 각각 클로닝하였다.

[0233]

AA3-8600RH : aaaaagcttAACGGTGTGTGGAGCGTCAG (서열번호 24)

[0234]

AA3-5597FSp : aaaactagTCGCGATCATTTGGCATG (서열번호 25)

[0235]

AA3-111RSp : aaaactagtGCTGGAGCCTCCATGGC (서열번호 26)

[0236]

AA3-1369RSp : aaaactagTCGTGCCACTGACCCGAAC (서열번호 27)

[0237]

AA4-2300FNs : aaatctagaGCTCCGTGCAGACCGAGAAC (서열번호 28)

[0238]

벡터를 회수하여 NsiI/SpeI, SpeI/HindIII 제한효소로 각각 처리 한 후 ligation 하여 두 fragment를 붙인 형태로 양쪽 arm이 벡터 하나에 존재하도록 제작하였다. 이 벡터를 pCR2.1-armA, pCR2.1-armB 로 명명하였다(도 9).

[0239]

확립한 pCR2.1-armA, pCR2.1-armB 벡터의 양쪽 arm 사이를 SpeI 제한효소로 자르고, pCR2.1-mo7A, pCR2.1-mo7B 벡터를 EcoRI 으로 잘라서 insert를 수득하였다. 이 두 단편을 각각 infusion ligation을 통해 결합하여 7번 모듈 교체 벡터의 구조를 완성하였다. 이 벡터를 각각 pCR2.1-A3Mmo7A, pCR2.1-A3Mmo7B로 명명하였다(도9)

[0240] pKC1139 벡터와 pCR2.1-A3Mmo7A, pCR2.1-A3Mmo7B 벡터를 각각 XbaI, HindIII 제한효소로 처리하고, pKC1139 벡터 fragment와 pCR2.1-A3Mmo7A, pCR2.1-A3Mmo7B 벡터의 insert fragment를 접합하여 pKC1139에 A3Mmo7A 또는 A3Mmo7B 구조가 들어가 있는 형태를 제작하였다. 이를 각각 pKC-A3Mmo7A, pKC-A3Mmo7B 벡터로 명명하였다 (도 9).

[0242] **실시예 4. 모듈 7 교체 균주 제작 및 생성물질 확인**

[0243] 모듈7 교체 벡터를 DBM-01 균주 내로 도입시키고, 상동 재조합이 일어나도록 유도하여 *S. avermitilis* DBM-01 균주의 모듈7이 milA3의 모듈7로 치환된 형태의 균주를 확보하고, 배양을 통해 생성되는 신규물질을 확인하였다.

[0245] **4-1.PEG-mediated protoplast transformation을 통한 벡터 도입**

[0246] Lysozyme을 처리하여 protoplast를 제조하고, pKC-A3Mmo7A(110), pKC-A3Mmo7B(110) 벡터를 섞고 PEG1000를 처리하여, PEG-mediated protoplast transformation을 통해 상기 벡터를 DBM-01 균주 내로 도입하였다. 24시간 배양한 후 항생제(apramycin)를 처리하였고, 위 벡터의 apramycin 저항성 유전자인 aac 유전자를 확인 할 수 있는 aac3iv-F 프라이머(GGTTCATGTGCAGCTCCATC: 서열번호 29) 및 aac3iv-R 프라이머(CGTCGCGGTGAGTTCAGG: 서열번호 30)를 사용하여 도입이 이루어졌음을 확인하였다.

[0248] **4-2.*S. avermitilis* DBM-01 균주로 모듈 7의 삽입**

[0249] 상기와 같이 모듈7 교체 벡터의 도입이 이루어진 균주에 1차 교차를 유도하여 벡터 전체가 *S. avermitilis* DBM-01 균주의 genomic DNA에 삽입되도록 하였다(도 10). 이를 위하여, apramycin이 25 µg/ml 농도로 첨가된 ISP-4 고체배지를 제조한 후, 벡터 도입에 의해 생성된 colony를 취한 후, 준비된 고체 배지에 streaking 하였다. 이후 고온에 민감한 pKC1139 벡터의 성질을 이용하여 벡터가 복제 되는 것을 막기 위하여 37에서 7일간 배양하였다.

[0250] 상기 1차 교차가 일어난 균주에서 aveA3의 모듈7을 milA3 모듈7로 치환하기 위해 2차 교차를 유도하였다(도 10). 이를 위하여, apramycin이 25ug/ml 농도로 첨가된 ISP-4 고체배지와 항생제가 첨가 되지 않은 배지를 각각 제조하고, 1차 교차를 유도한 균주를 ISP-4 고체 배지에 도말하여 2차 교차를 유도한 후, apramycin 저항성 유무로 후보 균주를 선별하였다. 스크리닝을 통해 확보한 후보 균주의 genomic DNA를 추출하여, aveA3 내의 모듈 7, 및 milA3 내의 모듈7 일부지역을 확인할 수 있는 프라이머(표 4)를 사용하여 유전자의 교체가 정확하게 이루어졌는지 확인하였다 (도 10). 이를 통해 확보된 균주 중 A3Mmo7A가 삽입된 균주를 *S. avermitilis* DBM-03-A, A3Mmo7B가 삽입된 균주를 *S. avermitilis* DBM-03-B로 명명하였다. *S. avermitilis* DBM-03-A 균주는 2015년 8월 31일자로 기탁번호 KCTC12890BP 를 부여받았고, *S. avermitilis* DBM-03-B 균주는 2015년 8월 31일자로 기탁번호 KCTC12891BP 를 부여받았다.

표 4

Primer Name	Sequence	용도
Amo7i-F	GTGTGCGGTGTGCTCCG (서열번호 31)	aveA3의 모듈7 확인
Amo7i-R	CACATCACACGCAACAATCC (서열번호 32)	
Mmo7i-F	ACATCAGCGCCTCTACGA (서열번호 33)	milA3의 모듈7 확인
Mmo7i-R	CCAGGTCATCACTCCCTC (서열번호 34)	

[0253] **4-4.모듈7 치환 균주의 생성물질 확인**

[0254] 모듈7 치환 균주를 배양하기 위한 종균 배양용 배지와 생산 배양용 배지를 각각 제조하였다. 종균 배지를 제조하기 위하여, soluble starch 30g/L, yeast extract 15g/L, KH₂PO₄ 0.4g/L의 농도가 되도록 적정량의 DW에 혼합하여 교반기로 충분히 섞어주고, 충분히 식은 후, 멸살한 Corn steep liquor 을 5g/L의 농도가 되도록 첨가하였

다. 또한, 생산 배지를 제조하기 위하여, soluble starch 80g/L, soybean meal 10g/L, skim milk 15g/L, KH₂PO₄ 0.5g/L의 농도가 되도록 적정량의 DW에 혼합하여 교반기로 충분히 섞어주고, pH를 7.2로 맞춘 후, 121, 15분간 고압살균을 실시하였다. 생산 배양에 접종할 균주의 양을 충분히 만들기 위해, 상기 제조한 종균 배지 25ml을 멸균된 250ml baffled flask에 넣고, 균주 mycelium을 agar plate로부터 1loop씩 떼서 종균 배지에 접종하였다. 230 rpm 및 28도 조건하에서 48시간 동안 배양하였다. 또한, 생산배지 25ml을 멸균된 250ml non-baffled flask 넣고 종균 배양액 1.25ml을 접종한 후, 230 rpm 및 28도 조건하에서 10일 동안 배양하였다. 생산 배양액으로부터 생성된 물질을 추출하기 위하여, 유기용매(acetonitrile : MeOH = 1 : 1 혼합 용액) 1.5ml과 배양액 0.5ml을 섞고 충분히 vortexing하였다. 5분간 sonication 한 후, 1시간 동안 교반하였다. 12000rpm에서 5분간 원심분리한 후 상등액만 취하여 분석하였다.

[0255] 추출한 시료를 아래의 조건에서 HPLC/UVD 분석을 통해 분석하였다.

[0256] - Mobile phase : Acetonitrile/Water(v/v)= 40/60 v/v, 90/10 v/v, 90/10 v/v - Flow rate : 0.7 ml/min

[0257] - wavelength : 245 nm

[0258] - Run time : 40 min

[0259] 그 결과, 밀베마이신 표준품의 HPLC 결과에서 나온 milbemycin A3와 A4의 peak (도 11a 참조)가 S. avermitilis DBM-03-A (도 11b 참조), S. avermitilis DBM-03-B 배양액 (도 11c 참조)의 HPLC분석 결과에서도 형성됨을 확인하였다(도 11a 내지 도 11c 참조).

[0260] 또한, 모듈 7 치환 균주를 배양하여 수득한 생성물 중, HPLC로 확인 한 milbemycin A3, milbemycin A4 이외의 생성물질에 대한 분석을 실시하였다. 그 결과, HPLC 상에 보인 major peak에 대한 mass 분석을 실시하여 분자량을 근거로 예측했을 때, milbemycin a2, milbemycin a4, milbemycin G 가 생성 되었음을 확인할 수 있었다 (도 12).

[0262] 실시예 5. aveA3의 변이를 위한 milA3 삽입 벡터(milA3 integration vector) 제작

[0263] 아베멕틴과 밀베마이신의 AT-DH 모듈간 영역 및 KR-ACP 모듈간 영역의 염기서열 분석을 통해 공통적인 amino acid를 코딩하는 지점을 치환이 일어나는 fusion site로 결정하고, amino acid의 서열에 변화를 주지 않는 범위 내에서 제한효소 site를 도입하였다. 상기 결과를 도 13에 나타내었으며, 하기의 유전자 변형을 위한 재조합 벡터의 제조의 제작시에 이를 참조하였다.

[0264] 상기 실시예 2-3에서 제작된 스트렙토마이세스 아베미틸리스 DBM-01의 aveA3 유전자의 모듈 7의 DH 도메인부터 모듈 9의 KR 도메인을 포함하는 영역의 코딩 유전자 부위를 스트렙토마이세스 밀베마이시니쿠스 (S. milbemycinicus NRRL5739)의 물질 생산 유전자군 milA3 유전자의 모듈 7의 DH 도메인부터 모듈 9의 KR 도메인 포함 영역의 코딩 유전자 부위로 치환하기 위한 벡터를 제작하였다. 이 벡터는 milA3의 일부를 포함하고 균주 내 재조합을 위한 상동성을 지니는 aveE 및 aveA3 유전자의 영역이 양쪽 말단에 연결된 형태로 구성하였다.

[0266] 5-1. MilA3 의 모듈 7의 DH 도메인부터 모듈 9의 KR 도메인까지의 코딩 유전자 부위의 확보

[0267] 밀베마이신 생산 균주인 스트렙토마이세스 밀베마이시니쿠스 (S. milbemycinicus NRRL5739)로부터의 milA3 모듈 7 DH~모듈 9 KR 영역까지의 코딩 유전자의 확보를 위해 NCBI에 등록된 또 다른 밀베마이신 생산 균주 스트렙토마이세스 빙첸젠시스 (Streptomyces bingchenggensis)의 밀베마이신 폴리케타이드 합성효소의 염기서열을 활용하여, 하기의 표 5와 같은 염기서열로 프라이머를 디자인하였다 (제작하고자 하는 DNA fragment가 매우 거대하여, 3개의 구역(fragments)로 나누어 PCR을 수행한 후, 하기의 실시예 5-3의 과정을 통해 합성함).

표 5

Primer	Sequence (5' → 3')	용도
DK1F	GCGTCCGGGGTCGACATGGGATGCAGGCCTGAGCGGTGATGCGGGCCCG (서열번호 38)	Fragment 1'
DK1R	ATGGGACTGCGGCTGCTTAAGACCTCACCCGTC(서열번호 39)	합성
DK2F	GCTAGCGGGGTGTGGTCG(서열번호 40)	Fragment 3
DK2R	TTCGAACACGCGGGAATCGAC(서열번호 41)	

DK3F	GCATGCCGTGTCCACGCGCCC(서열번호 42)	Fragment 5-6'
DK3R	ACACCCACCCCCACAACCACCACCTAGATCTGCCACCTACGCCTTCCA(서열번호 43)	합성

[0269] 상기 프라이머를 사용하여 얻어진 각 fragment를 PCR (1. 97 °C 3분, 2. 72°C 1분, 3. 98 °C 10초, 4. 55 °C 15초, 5. 68 °C 1분 30초, 6. 스텝 3~ 스텝 5 에 대해 31 사이클 반복 수행, 7. 72 °C 10분, 8. 4 °C, 9. end) 을 통해 증폭하였다. 각 fragment는 pGEMTeasy 벡터 (Promega)에 클로닝되어 DNA sequencing을 통해 염기서열을 확인하였다. 이와 같이 얻어진 벡터를 pYJ1950, pYJ1951, pYJ1952로 명명하였다. 이들 각 플라스미드의 insert는 도14의 fragment 1', 3과 5-6'에 해당하며, 각 fragment는 클로닝의 용이성을 위해 milA3 모듈 7의 DH 도메인부터 모듈 9의 KR 도메인 영역에 존재하는 제한효소 위치를 참조하여 디자인 하였다.

[0271] 5-2. Left arm 및 Right arm의 제작

[0272] 스트렙토마이세스 아베미틸리스 DBM-01 균주의 모듈 7 DH~모듈 9 KR 영역에 상동재조합을 통한 교차를 일으키기 위해 각 벡터를 제작하였다. 스트렙토마이세스 아베미틸리스 DBM-01 균주의 genomic DNA를 주형으로 사용하고, DKLF/DKLR 및 DKRF/DKRR 프라이머를 이용하여 (아래의 표 6 참조), milA3 유전자의 모듈 7 DH~모듈 9 KR 영역에 연결될 left 및 right flanking 영역 fragment를 PCR로 확보한 뒤, 각각 pGEMTeasy 벡터에 클로닝하였다. 이때 각 fragment는 연결 될 milA3 영역과 상동성을 갖도록 디자인 되었고 이를 각각 pYJ1948 및 pYJ1949로 명명하였다. 각각 플라스미드의 insert는 도 14의 LA와 RA에 해당한다.

표 6

Primer	Sequence (5'→3')	용도
DKLF	AACCCGAACGACCGAGCGCAGCGCGGCCGGAATTCATACGGCTACGAGGAT(서열번호 44)	Left arm (L A)의 제작
DKLR	CTCGACCGCACCGACCGGGCCGCGATCAACCGCTCAGGCCTGCATCCCATGTCGACCCC(서열번호 45)	
DKRF	GCCAGTAGCGGGCGCGCTGGAAGGCGTAGGTGGGCAGATCTAGGTGGTGGTTGTGGGGGTG(서열번호 46)	Right arm (RA)의 제작
DKRR	CTTAGGACCGTTATAGTTACGCGCTTCCTCGCTCACTGACTTTAATTAAGCTTACGTTTCGCTTGCCGATG(서열번호 47)	

[0275] 5-3. milA3 유전자에 양쪽 arm이 결합된 형태의 vector 제작

[0276] 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739)로부터의 PCR을 통해 확보된 milA3영역을 포함하는 pM5-2벡터 및 pM6-4벡터를 제작하고, pM5-2벡터 에 제한효소 AfIII/BmtI를 처리하여 도14의 fragment 2를 얻었으며, pM6-4벡터에 제한효소 BstBI/SphI를 처리하여 도14의 fragment 4를 확보하였다.

[0277] 상기 과정을 보다 구체적으로 기재하면 다음과 같다:

[0278] 스트렙토마이세스 밀베마이시니쿠스 (*S. milbemycinicus* NRRL5739)의 milbemycin 생산유전자군을 클로닝하기 위하여, 스트렙토마이세스 빙첸젠시스 (*S. bingchenggensis* BCW-1) 의 genome 염기서열 정보 (Accession no. CP002047.1)를 기반으로 milA2, milA3, milA4 지역을 포함하는 61kb의 영역을 각각 10kb 씩 6개 영역으로 나누어 프라이머를 디자인 하였다(표 7 및 도 19 참조)

표 7

Fragment 명칭	Primer Name	Sequence	용도
M1	M1171-1NF	aaaatgcatCTACCGCGCTTCAACAGC(서열번호 48)	Fragment 1 증폭
	M10810-1R	aaaactagtGGAGAGGAAGAAGGAGAGCCG(서열번호 49)	
M2	M10762-2NF	aaaatgcatCACCCACACAGGACAGGGA(서열번호 50)	Fragment 2 증폭
	MC20300-2R	aaaactagtATGACAGGCCGGGTACATCA(서열번호 51)	

M3	M20255-3NF	aaaatgcatCGCGTACCTGATGTCCCTGA(서열번호 52)	Fragment 3 증폭
	M30434-3R	aaaactagtGGCCACACGGTCTTCATCG(서열번호 53)	
M4	M30377-NF2	aaaatgcatCCTGCACACCGACCGAGAC(서열번호 54)	Fragment 4 증폭
	M40160-4R	aaaactagtACCTTCGTACCCGTCTCGC(서열번호 55)	
M5	M40091-5NF	aaaatgcatGTGCGGACGATCGTGAGAAG(서열번호 56)	Fragment 5 증폭
	M50273-5R	aaaactagtAACACGCGGGAATCGACC(서열번호 57)	
M6	M50196-6NF	aaaatgcatGTAGTCCGTCGGCATCACAC(서열번호 58)	Fragment 6 증폭
	M60390-6R	aaaactagtTCGAGACCGTCCCTATGTCC(서열번호 59)	
M4N	M30290-F	ACGTCAGGAAGCGTTCA(서열번호 60)	Fragment M4N 증폭
	M34887-R	CGTCCGAACGTTTCTGGA(서열번호 61)	
M4CN	M4M4502F	GAGCTCCAGAAACGTTTCGG(서열번호 62)	Fragment M4CN 증폭
	M4F7910R	GGATCGTCAACATCTCCTCC(서열번호 63)	
M4CC	M4C-ScF	CGGCCACCAGAACACCTTC(서열번호 64)	Fragment M4CC 증폭
	M4C25R	ACACGCTTCTCAGATCGTC(서열번호 65)	

[0281] 다음의 표 8의 조건으로 PCR 용 mixture를 제조하였다:

표 8

[0282] PCR 용 mixture의 조성

Component	1 reaction 당 (uL)	
5X GXL buffer	5	Primer를 제외한 component를 모두 혼합 후 21 ul / tube 첨가
2.5 mM dNTPs	2	
S. milbemycinicus Genomic DNA (1/50 dil.)	2	
DW	16.5	
PrimeSTAR Pol.	0.5	
Primer for. (N), 5 pmol/ul	2	
Primer rev. (C), 5 pmol/ul	2	

[0284] 상기 mixture에 대하여 MyCycler (Bio-Rad 사)를 사용하여 다음의 표 9의 조건에서 PCR을 수행하였다

표 9

Step	Temperature	Time	Cycles
First denaturation	95 °C	5 min	1
Denaturation	98 °C	20 sec.	30
Extension	68 °C	10 min	
Final extension	68 °C	10 min	1

[0286] 증폭된 DNA fragment 중 milA3에 해당되는 fragment인 M5와 M6를 pCR2.1 TOPO cloning 시스템 (Invitrogen, 미국)을 이용하여 클로닝하고 *E. coli* DH5a를 형질전환하여 각각 pM5-2, pM6-4를 확보하였다.

[0287] 상기 얻어진 fragments와, 상기 표 5의 프라이머 DK1F/DK1R, DK2F/DK2R, DK3F/DK3R 를 사용한 PCR 결과물 및 left, right flanking 영역과 단계적으로 overlap PCR 방법과 통상적인 ligation반응을 통하여 pLitmus28 (NEB, 미국) 에서 한 fragment로 연결하였다. 이와 같이 얻어진 재조합 벡터를 pYJ1957로 명명 하고, 이로부터 EcoRI/HindIII 제한효소 처리를 통해 얻은 left flanking region-milA3 모듈 7 DH~모듈 9 KR-right flanking region을 pKC1139로 옮겨 milA3 모듈 7 DH~모듈 9 KR 영역의 발현을 위한 재조합 벡터 (milA3 삽입 벡터; milA3 integration vector)를 제작하였다(도 14 참조; LA: left arm; RA: right arm; 1', 2, 3, 4, 5, 6: fragments). 이를 pD7K9milA3으로 명명하였다 (서열번호 80; LA: 6074-7261 부위; 1': 7256-10476 부위; 2: 10471-15043 부위; 3: 15038-17193 부위; 4: 17188-20776 부위; 5 및 6': 20771-21226 부위; RA: 21221-22436

부위).

[0289] **실시예 6. milA3 삽입 벡터(milA3 integration vector)가 삽입된 균주 제작 및 생성물질 확인**

[0290] 상기 실시예 5-3에서 제작된 milA3 integration 벡터 (pD7K9milA3) 를 *S. avermtilis* DBM-01 균주 내로 도입시키고, 상동재조합이 일어나도록 유도하여 *S. avermtilis* DBM-01 균주의 aveA3 유전자가 milA3 유전자로 치환된 형태의 균주를 확보하고, 배양을 통해 생성되는 신규물질을 확인하였다.

[0292] **6-1. 접합법을 통한 벡터 도입 및 유전자 치환**

[0293] *E. coli* ET12567/pUZ8002 (The John Innes Foundation, 영국) 를 치환 벡터(pD7K9milA3)로 형질전환시킨 형질 전환체를 아프라마이신 (apramycin), 클로람페니콜 (chloramphenicol), 카나마이신 (kanamycin)이 포함된 2xTY (trptone 16 g/L, yeast extract 10 g/L, NaCl 5 g/L) 또는 LB (peptone 10g/L, yeast extract 5g/L, sodium chloride 10g/L) 액체 배지를 이용하여 37 °C에서 배양하였다. 12 시간 후, 1 ul의 배양액을 25 mL의 2xTY에 옮기고, 37°C에서 광학 밀도 (optical density)가 0.6이 될 때까지 배양하였다. 배양액의 세포를 동일한 액체 배지로 두 번 세척한 뒤, 500 uL의 2xTY로 현탁하여 얼음에 보관하였다. 스트렙토마이세스 아베미틸리스 DBM-01의 포자액 30-50 uL를 원심분리 (12,000 rpm으로 10분간)하고, 500 uL의 2xTY 에 현탁한 후, 50°C에서 10분 동안 열충격을 가하여 균사액을 준비하였다. 얼음에 보관해둔 *E. coli* 현탁액 500 uL과 상기 준비된 균사액 500 uL을 서로 혼합한 후 MS (agar 20g/L, D-mannitol 20g/L, soya flour 20g/L) 배지에 도포하여 28°C에서 24 시간 동안 배양한 뒤 날리딕신산 (nalidixic acid)과 아프라마이신 (apramycin) 각 1mg을 배지위에 도말하였다. 7~10일 뒤 아프라마이신 저항성을 지닌 콜로니를 선별하고 37 °C에서의 배양을 통해 항생제 저항성 여부를 재확인 후, 삼차 스트리킹 (streaking)을 통해 얻은 콜로니에 대해 PCR 및 PCR 결과물에 대한 sequencing을 수행하여 일차교차 균주를 선별하였다.

[0294] 상기 선별된 콜로니로부터의 계대배양을 통해 이차교차를 유도하고, 이를 아프라마이신이 첨가된 ISP4 고체배지와 아프라마이신이 첨가되지 않은 ISP4 고체배지에 각각 도말하여 아프라마이신 저항성을 잃은 콜로니를 선별하였다.

[0295] 상기 사용된 ISP4 고체배지의 조성을 아래의 표 10에 나타내었다:

표 10

Reagents	g/L
Soluble starch	10.0
KH ₂ PO ₄	1.0
MgSO ₄ ·7H ₂ O	1.0
NaCl	1.0
(NH ₄) ₂ SO ₄	2.0
CaCO ₃	2.0
Trace salt solution	1 mL
Agar	20.0

[0296]

[0297] PCR 분석을 통해 최종 milA3가 치환된 이차교차 균주가 확보되었다. 상기 균주 제작 과정을 도 15에 모식적으로 나타내었다.

[0298] 상기와 같이 확보된 균주를 *S. avermitilis* DBM-03-C로 명명하였다. *S. avermitilis* DBM-03-C 균주는 2016년 8월 24일자로 대한민국 전라북도 정읍시에 위치하는 한국생명공학연구원(Korean collection for type cultur

e)에 기탁하여 수탁번호 KCTC13083BP를 부여받았다.

6-2. milA3 치환균주의 배양

Milbemyci의 생산확인을 위한 종균배양 및 생산배양을 수행하였으며, 이 때 상용된 배지의 조성을 다음의 표 11에 나타내었다:

표 11

Seed culture	g/L	Main culture	g/L
Soluble starch	30.0	Soluble starch	80.0
Yeast extract	15.0	Soybean flour	10.0
Corn steep liquor	5.0	Skim milk	15.0
KH ₂ PO ₄	0.4	KH ₂ PO ₄	0.5

각 분량의 soluble starch, yeast extract, corn steep liquor, 및 KH₂PO₄를 DW에 녹여서 혼합한 후, NaOH를 사용하여 pH 7.2로 맞춘 후, 121℃에서 20 분간 고압살균하였다. ISP4 고체배지 (상기 표 5 참조)에 배양한 milA3 발현 균주 (S. avermitilis DBM-03-C)의 포자 혹은 mycelium을 긁어 종균배지에 접종한 뒤 28 °C, 230rpm에서 40~48 시간 배양하여 종균배양액을 준비하였다. Soluble starch, soybean flour, skim milk, 및 KH₂PO₄를 DW에 녹이고 pH 7.2에 맞춰 제조한 생산배양 배지를 종균배양 배지와 같은 조건하에서 고압 멸균한 뒤 생산배양 배지 총 부피의 5 % 비율로 상기 준비된 종균배양액을 접종하여 230 rpm 및 28 °C 조건에서 10~15 일간 배양하였다.

6-3. 생성물질의 추출 및 분석

분석을 위한 시료의 전처리를 위하여, 상기 얻어진 배양액 0.5 mL을 동일 부피의 acetonitrile 및 methanol 혼합 용액 (acetonitrile:methanol=1:1 (v:v)) 또는 동일 부피의 ethyl acetate 와 혼합하여 2 분간 vortex 처리한 뒤 10 분간 초음파 추출을 수행하였다. 5분간 원심분리하여 고형물을 제외하고 상층액만을 취하였다. 동일한 방법으로 추출을 3회 반복하여 분석시료를 확보하였다. 상기 확보된 시료는 HPLC 및 UPLC-qTOF-HR-MS를 통해 분석하였다.

HPLC는 PDA detector와 gradient pump로 구성된 HPLC system (YL Instrument Co. Ltd)하에서 Phenomenex Luna (4.6x100 mm, 5 mm) column을 이용하여 수행하였고, 물 기반의 0.05% (v/v) trifluoroacetic acid (TFA)를 용매 A로 하고 acetonitrile을 용매 B로 하고 표 12와 같은 용매 기울기를 사용하여, 238~245 nm 에서 1 mL/min 의 flow rate로 총 60 분간 시료를 분석하였다.

표 12

Time (min)	0.05 % TFA in water (%(v/v))	Acetonitrile (%(v/v))
0	60	40
5	60	40
35	10	90
55	10	90
60	60	40

Milbemycin A3/A4의 30%/70% (w/w) 표준품인 milbemectin (Bioaustralis, Australia) 및 milbemycin D (Bioaustralis, Australia)를 사용하여 본 HPLC 분석 조건에서의 해당 milbemycin의 retention time (RT) 및

245 nm에서의 absorbance를 확인하였다. 상기 확보된 시료 (S. avermitilis DBM-03-C 균주)의 생성 물질을 분석한 결과를 도 16에 나타내었다. 도 16에서, (a)는 표준품 milbemectin (Bioaustralis, Australia) 및 milbemycin D (Bioaustralis, Australia)의 분석 결과를, (b)는 S. S. avermitilis DBM-03-C 균주의 생성물질 분석 결과를 각각 보여준다. 도 16에 나타난 바와 같이, 상기 S. avermitilis DBM-03-C 균주는 milbemycin A3 (CAS NO. 51596-10-2), A4 (CAS NO. 51596-11-3), alpha2 (CAS No. 51596-12-4), D (CAS NO. 77855-81-3), alpha4 (CAS No. 51596-13-5), 및 G (CAS No. 83471-31-2)로 추정되는 피크를 확인할 수 있고, 이 중에서, milbemycin A3, A4, 및 D는 표준품 milbemycin과 동일한 RT값을 지님을 확인할 수 있다. 특히 milbemycin A3, A4 및 D로 추정되는 peak는 표준품 milbemycin 과 동일한 RT 값을 지님을 알 수 있다. 도 16의 HPLC 분석 결과에 나타난 각 물질의 RT는 다음과 같다:

[0310] Milbemycin A3 RT: 29.0 min

[0311] Milbemycin A4 RT: 32.1 min

[0312] Milbemycin D RT: 35.4 min

[0313] Milbemycin alpha 2 RT: 33.7 min

[0314] Milbemycin alpha 4 RT: 36.3 min

[0315] Milbemycin G RT: 39.4 min.

[0316] HPLC 상에서 milbemycin으로 추정되는 물질의 분자량을 비교하기 위하여, UPLC-qTOF-HR-MS를 다음과 같이 수행하였다. 상기 HPLC 상에서 분석된 여섯 종류의 milbemycin의 해당 분자량 값을 Xselect® CSH column XP (2.1x100 mm, 2.5 um; Waters) 을 통해 Waters XEVO® G2S Q-TOF mass spectrometer; Waters) 상에서 확인하였다. 이 때 사용된 용매 기울기는 아래의 표 13과 같이 하였다.

표 13

Time (min)	Water (vol%)	Acetonitrile (vol%)
0	50	50
1	50	50
7	25	75
15	10	90
18	10	90
20	50	50

[0317]

[0318] 상기 얻어진 질량 스펙트럼 결과를 도 17a 내지 17f에 나타내었다 (도 17a: 밀베마이신 A3; 도 17b: 밀베마이신 A4; 도 17c: 밀베마이신 D; 도 17d: 밀베마이신 alpha2; 도 17e: 밀베마이신 alpha4; 도 17f: 밀베마이신 G). 도 17a 내지 17f에 나타난 바와 같이, milbemycin A3에 해당하는 m/z=527.3014, milbemycin A4 및 a2에 상응하는 m/z=541.3171, milbemycin D및 a4를 나타내는 m/z=555.3327, milbemycin G의 값인 m/z=569.3484을 negative ion mode에서 확인할 수 있었으며, 각각 분자의 알려진 분자량과 잘 일치하는 것으로 보아 제작된 재조합 스트렙토마이세스 아베르미틸리스 DBM-03-C 가 밀베마이신을 생산함을 알 수 있었다.

수탁번호

[0320]

기탁기관명 : 한국생명공학연구원

수탁번호 : KCTC12889

수탁일자 : 20150831

기탁기관명 : 한국생명공학연구원

수탁번호 : KCTC12890

수탁일자 : 20150831

기탁기관명 : 한국생명공학연구원

수탁번호 : KCTC12891

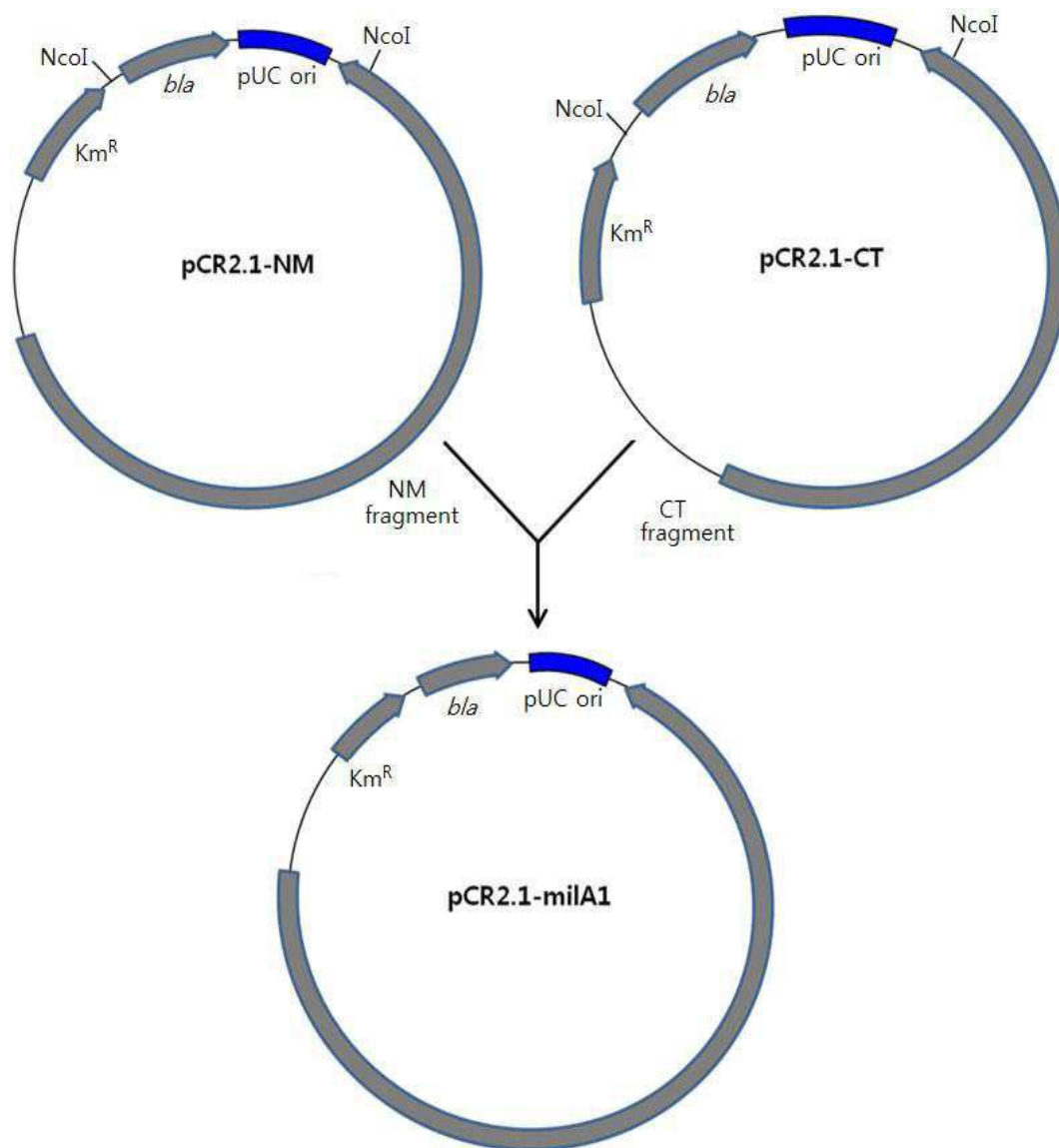
수탁일자 : 20150831

기탁기관명 : 한국생명공학연구원

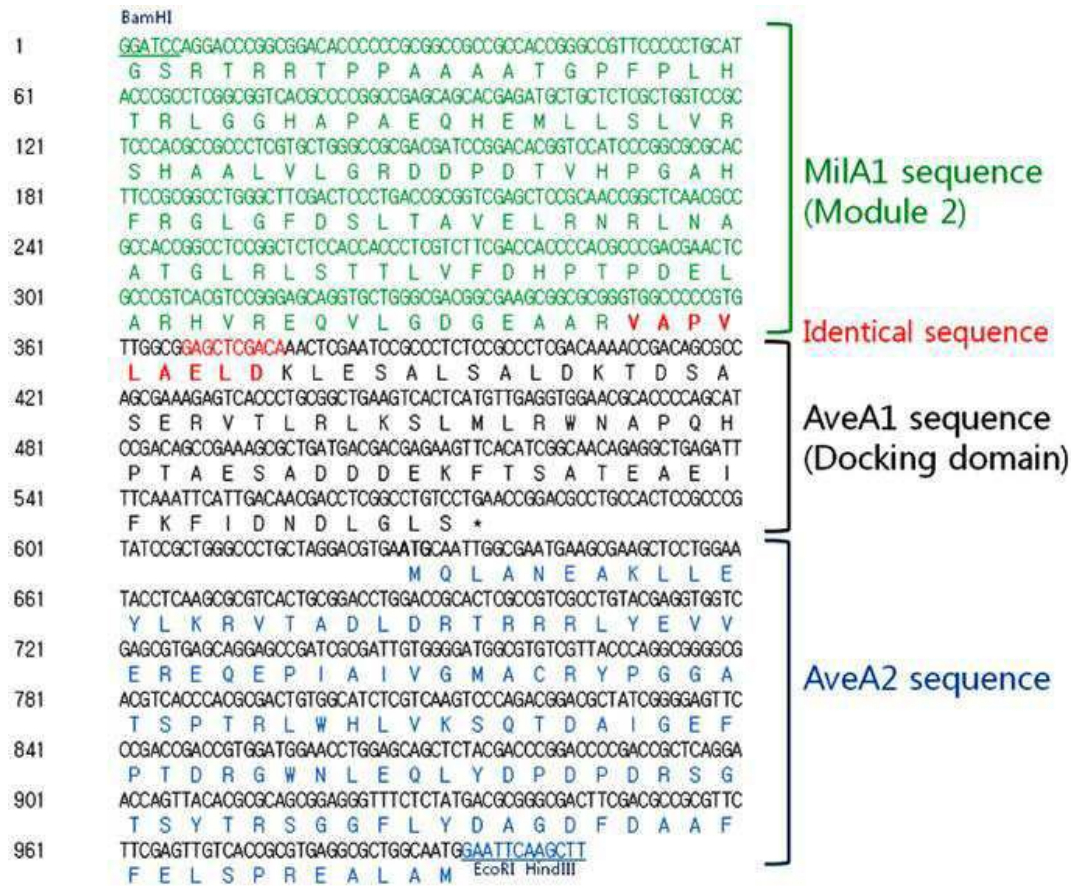
수탁번호 : KCTC13083

수탁일자 : 20160824

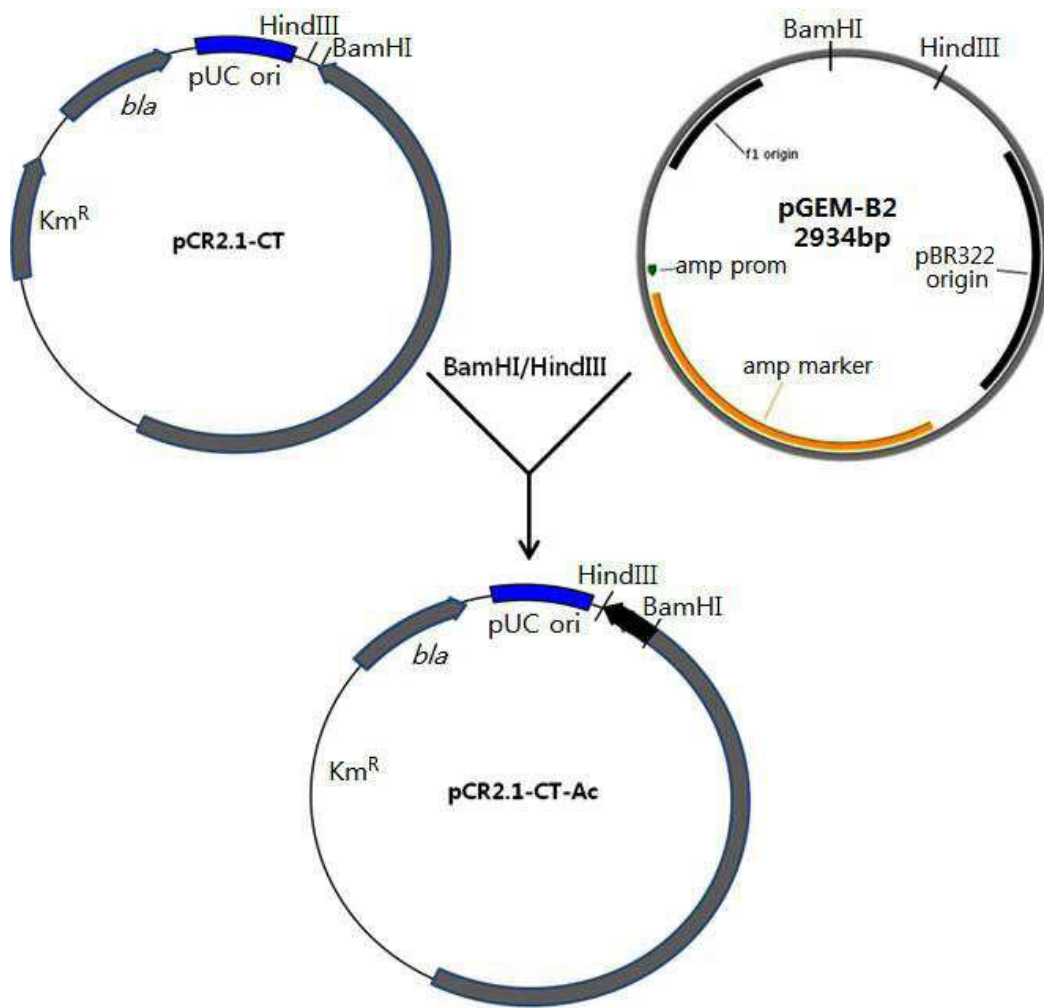
도면2



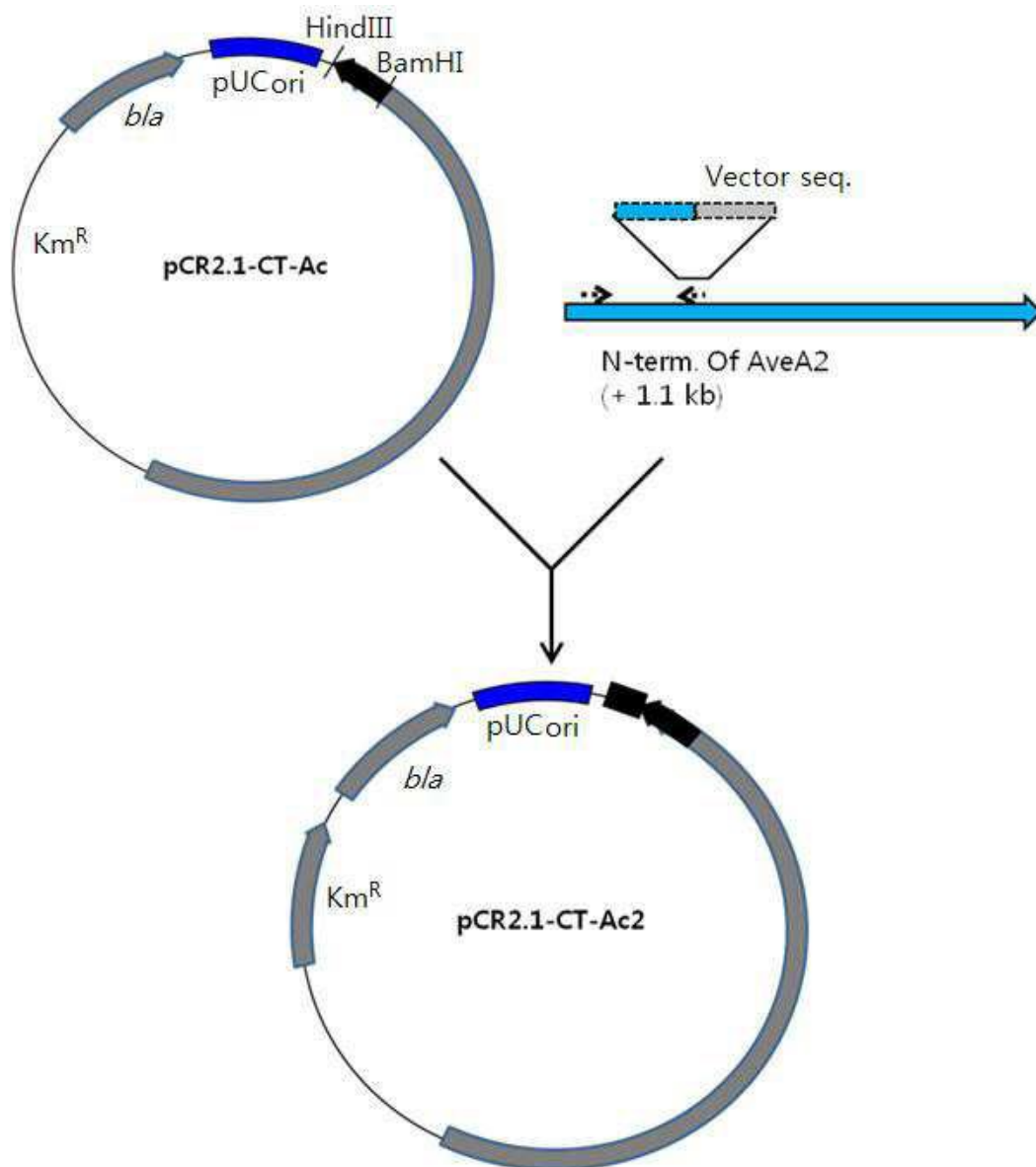
도면3



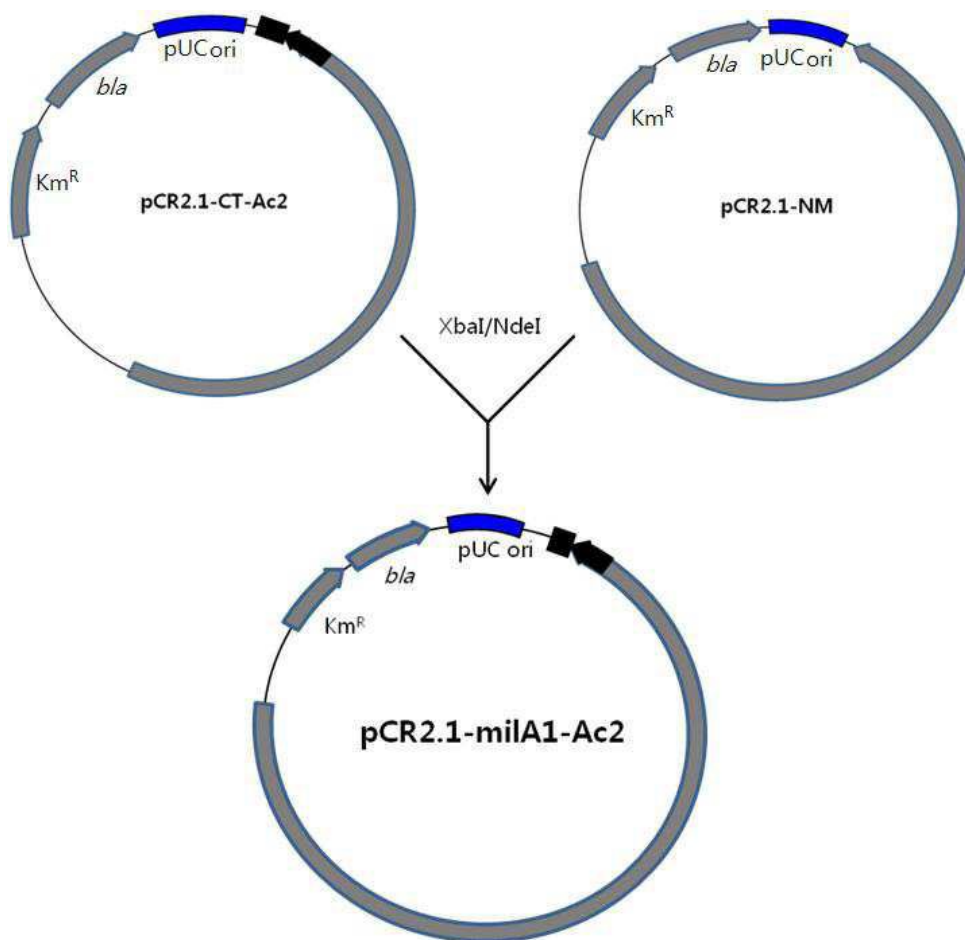
도면4a



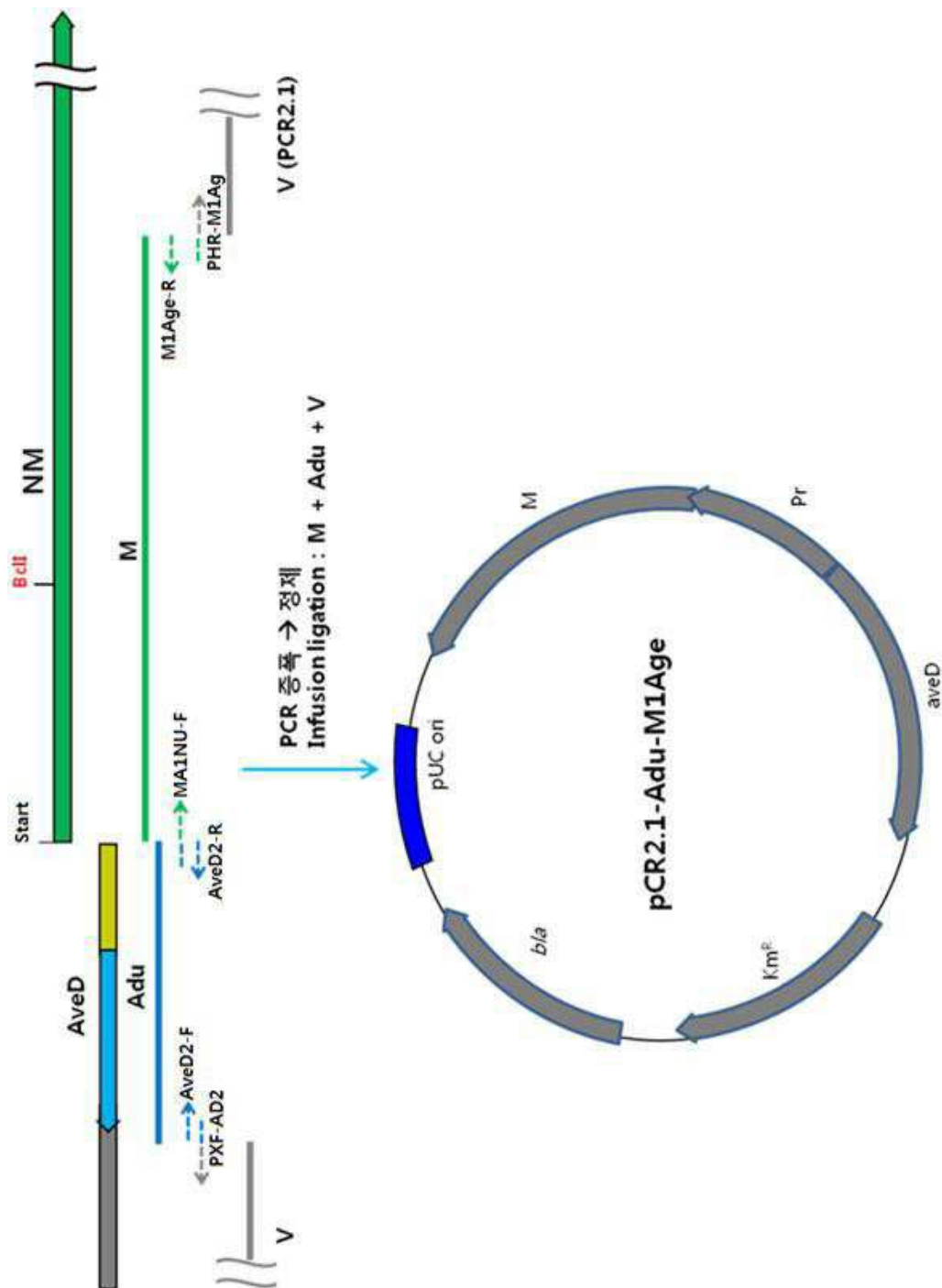
도면4b



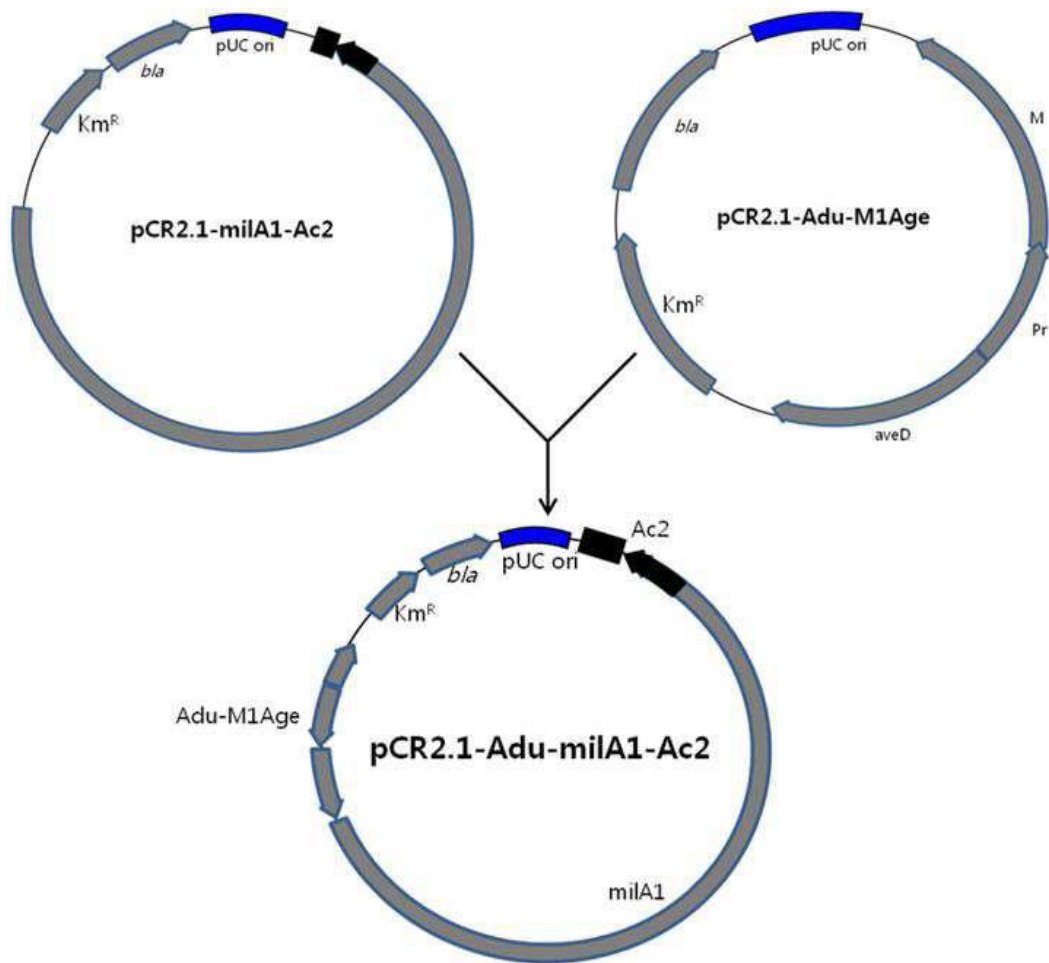
도면4c



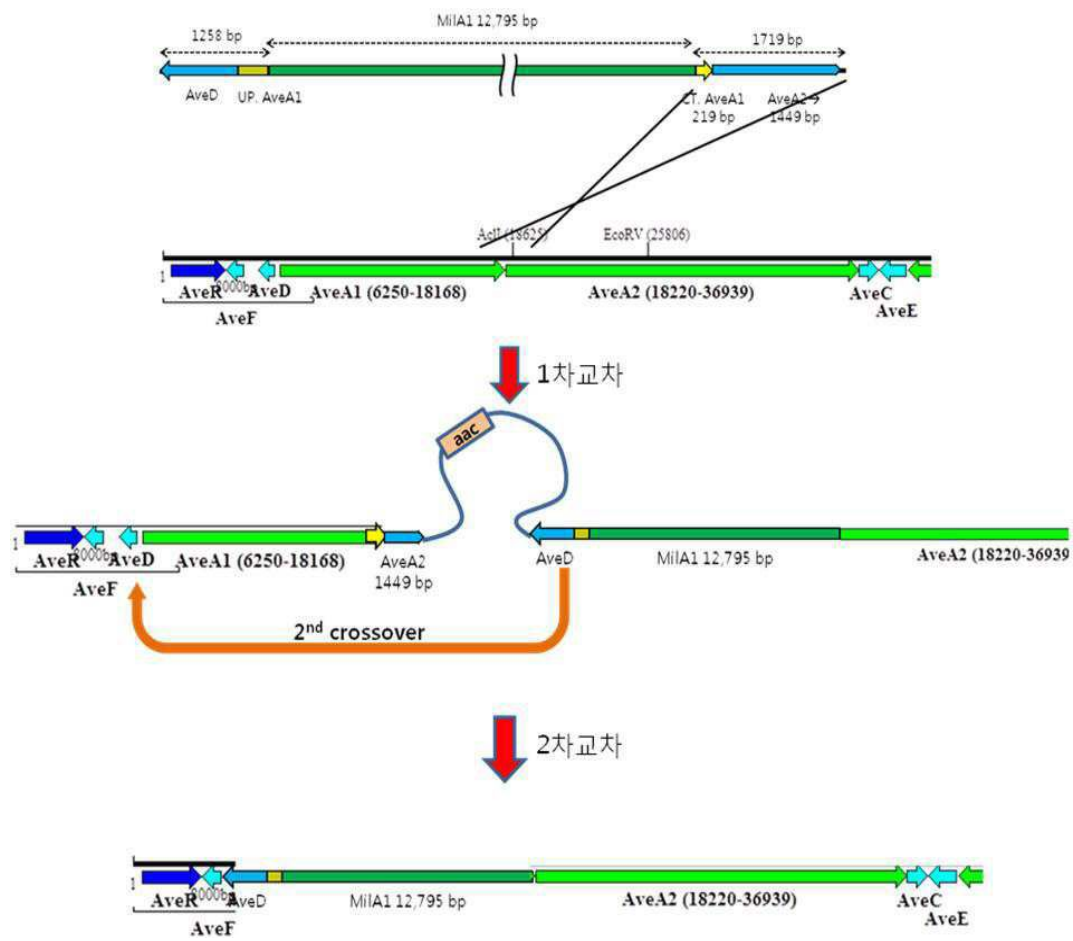
도면5



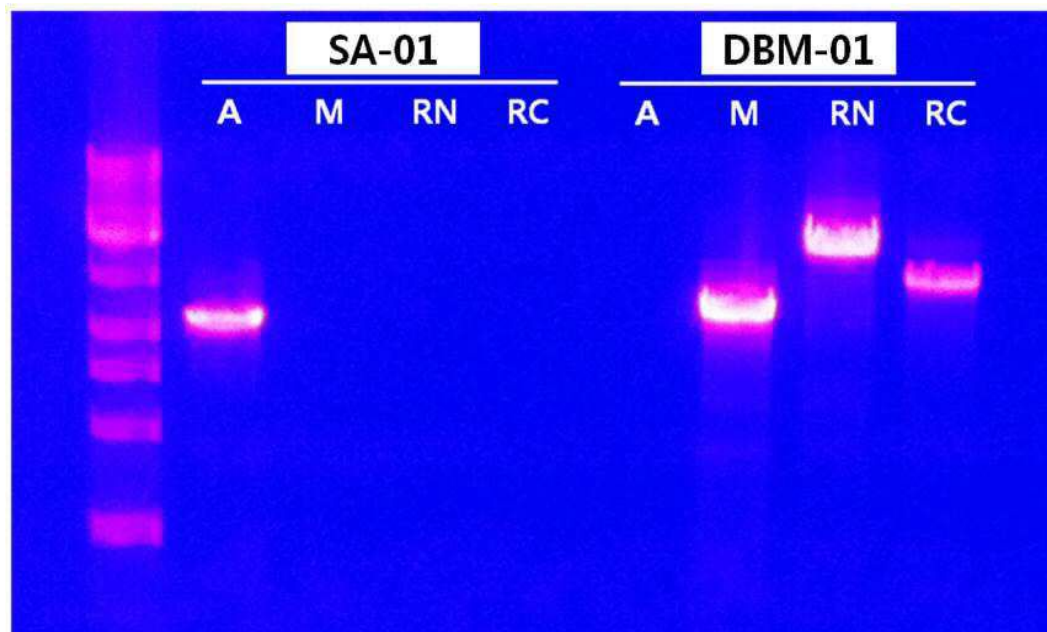
도면6



도면7a

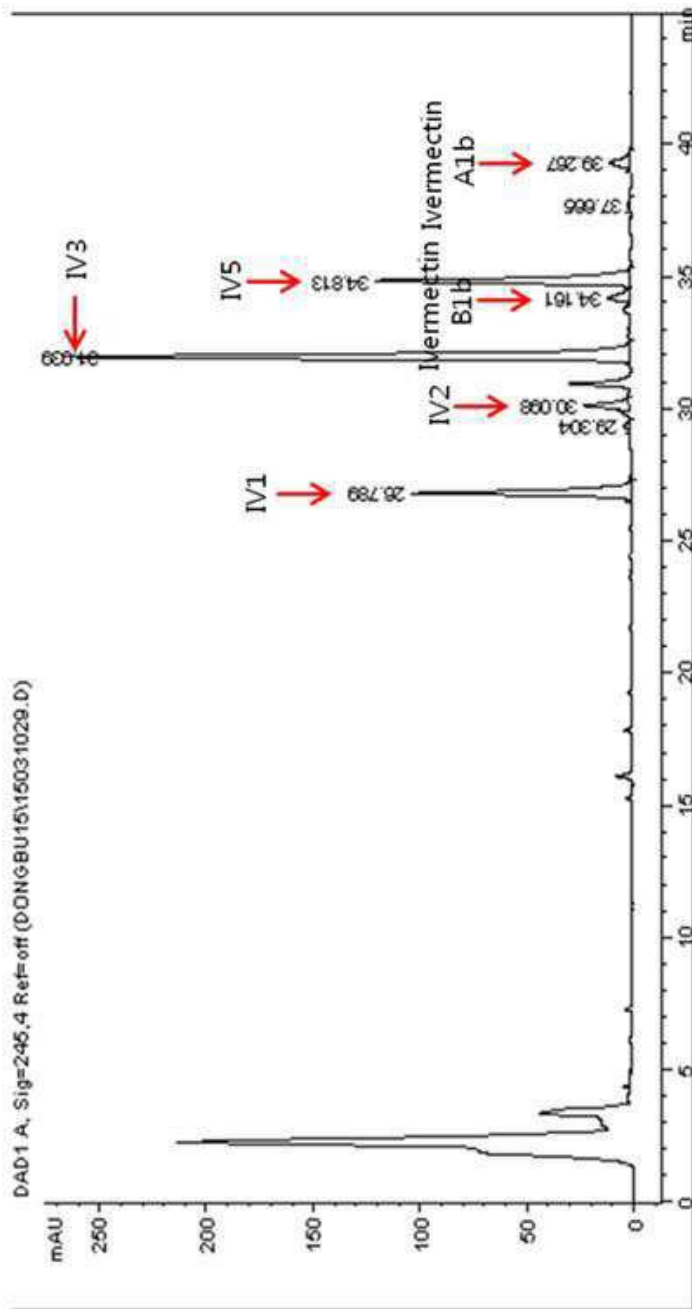


도면7b

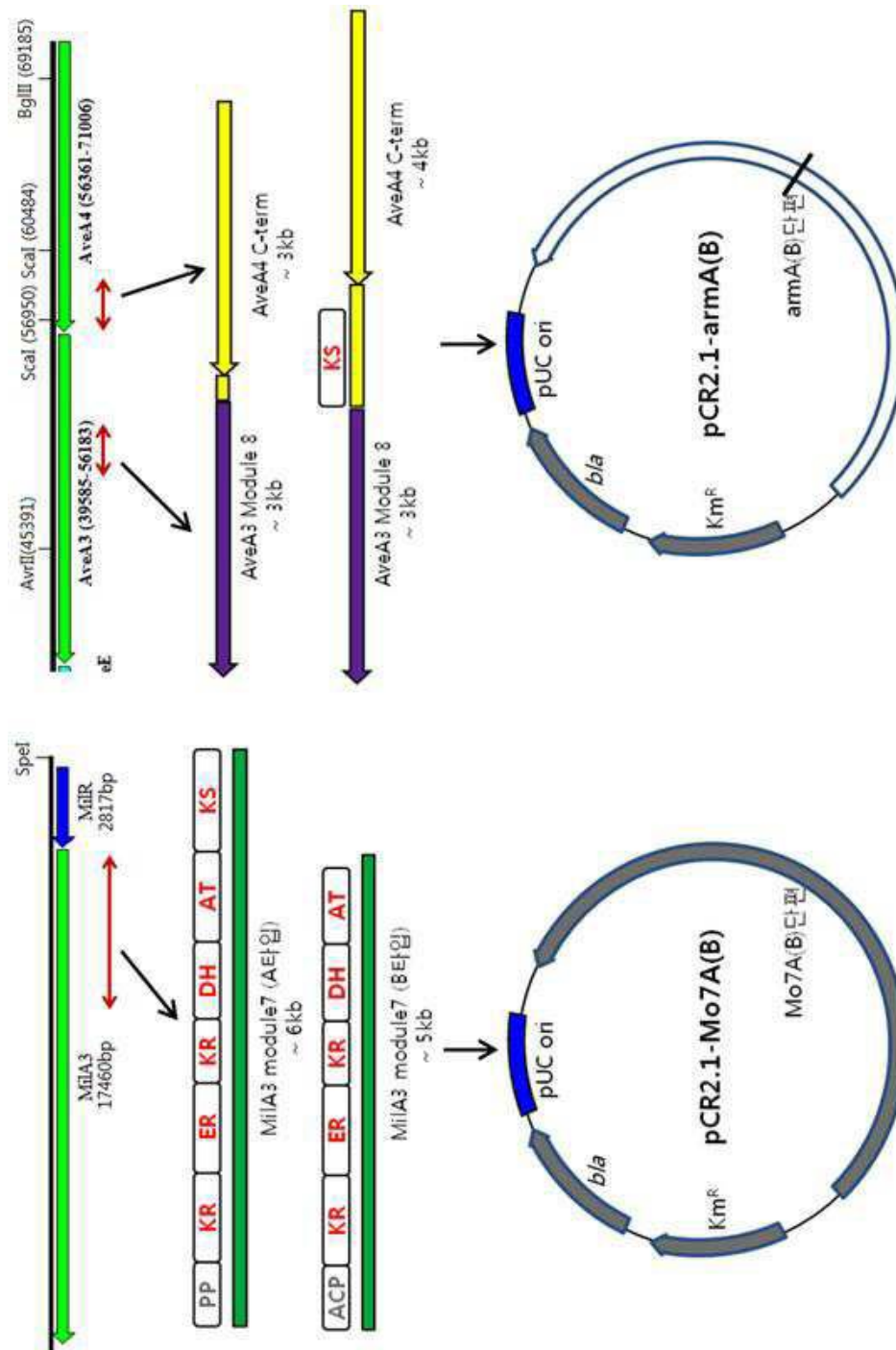


A : aveA1 지역 (2.2K)
M : milA1 지역 (2.4K)
RN : N-fusion (4.0K)
RC : C-fusion (3.3K)

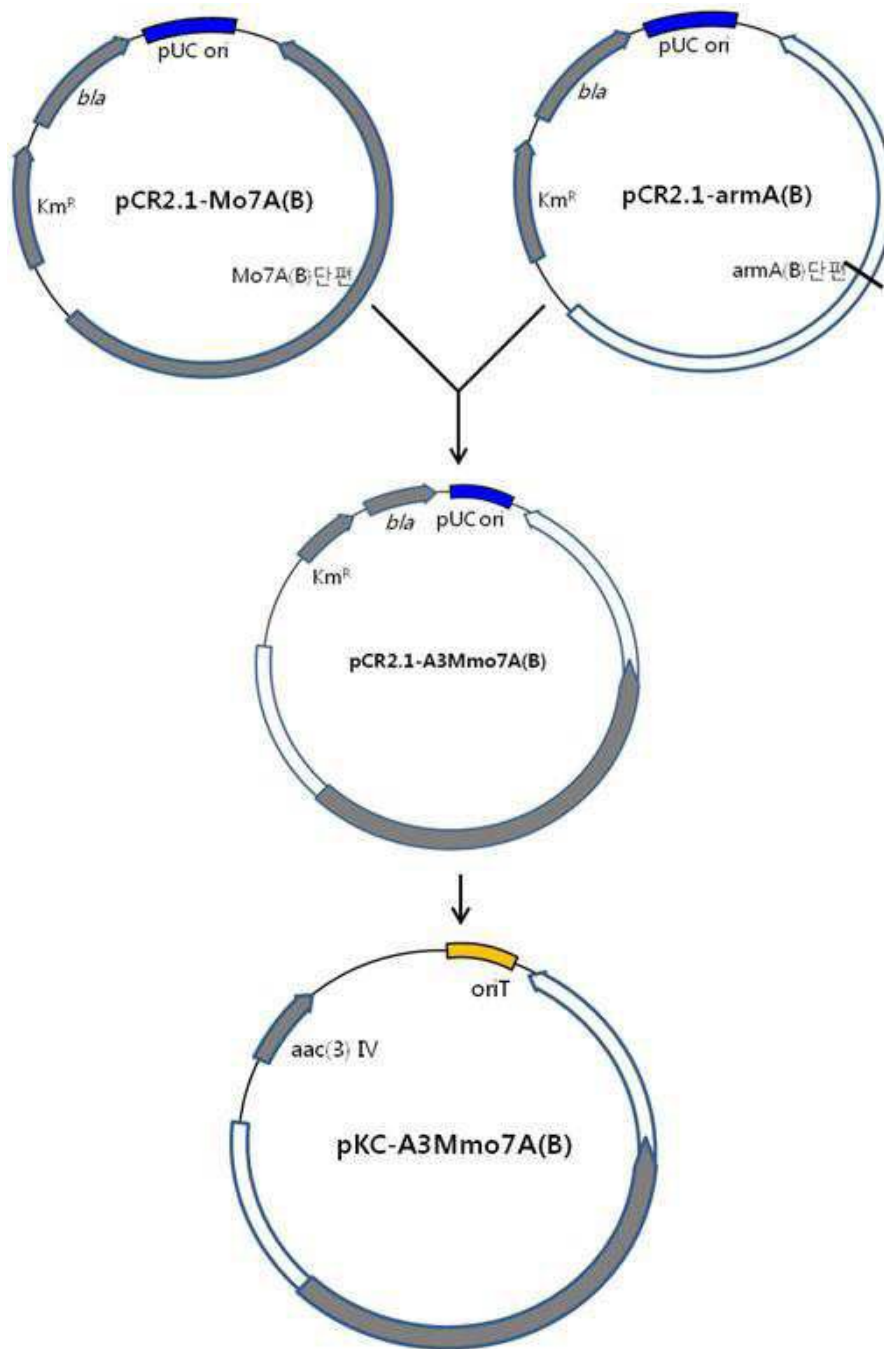
도면8



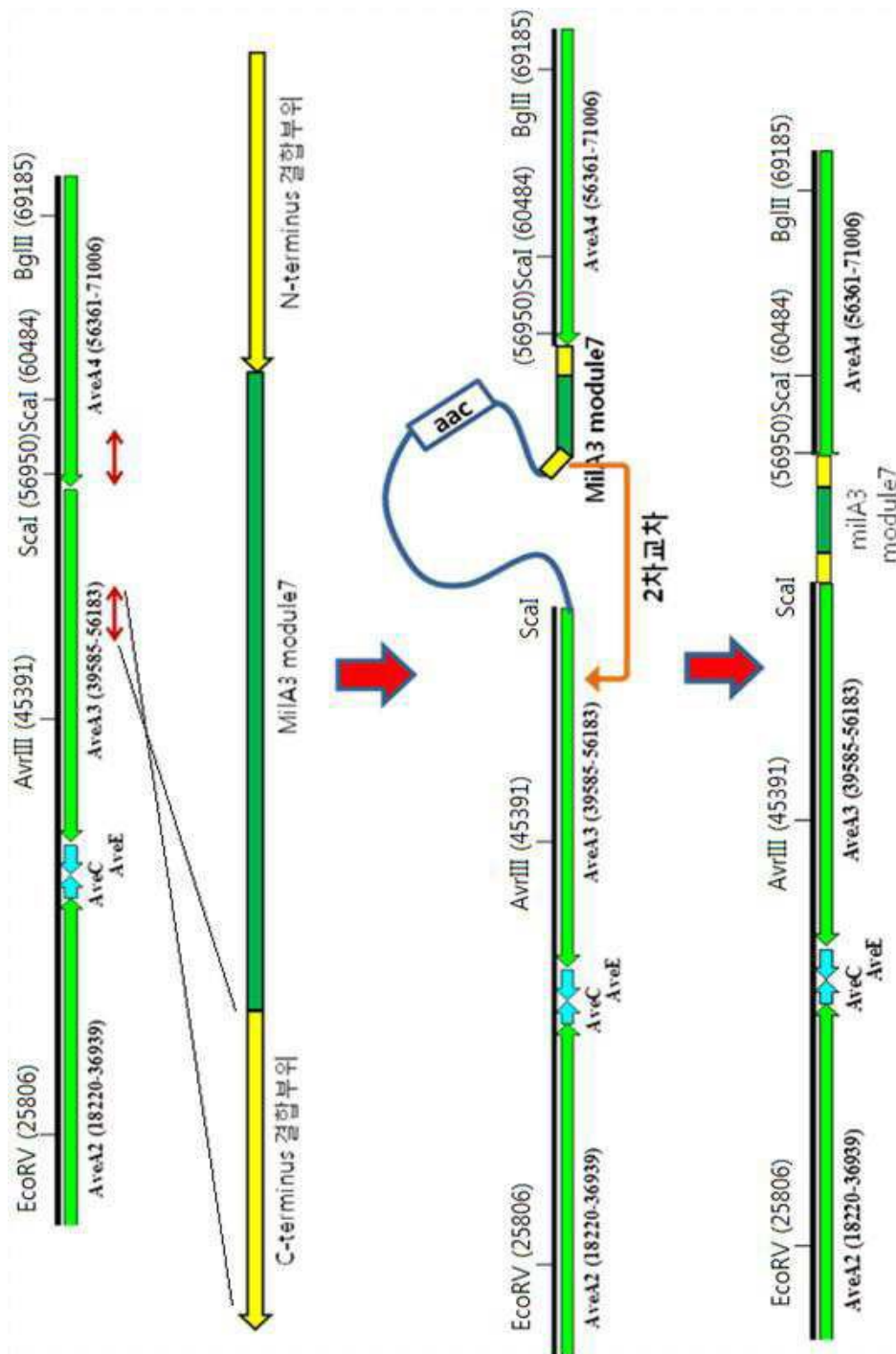
도면9a



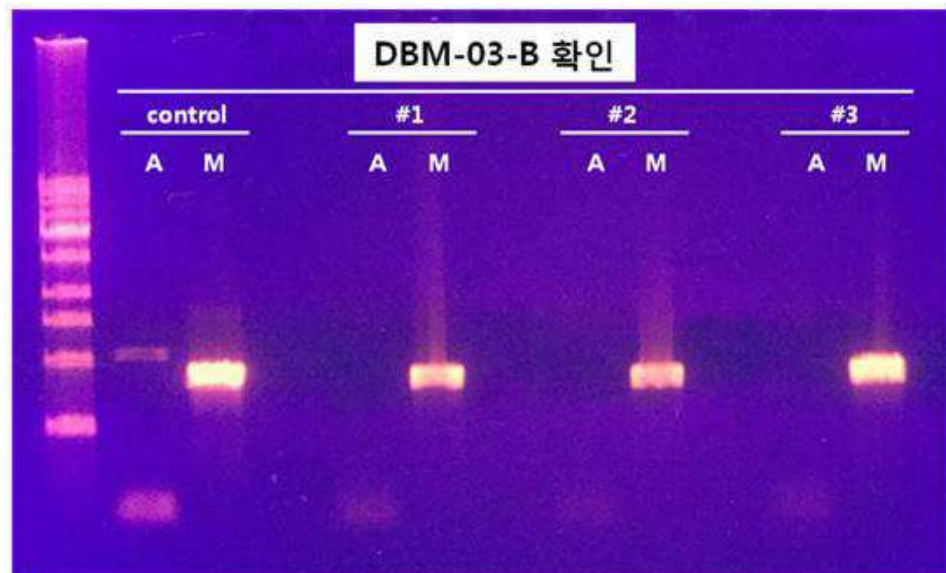
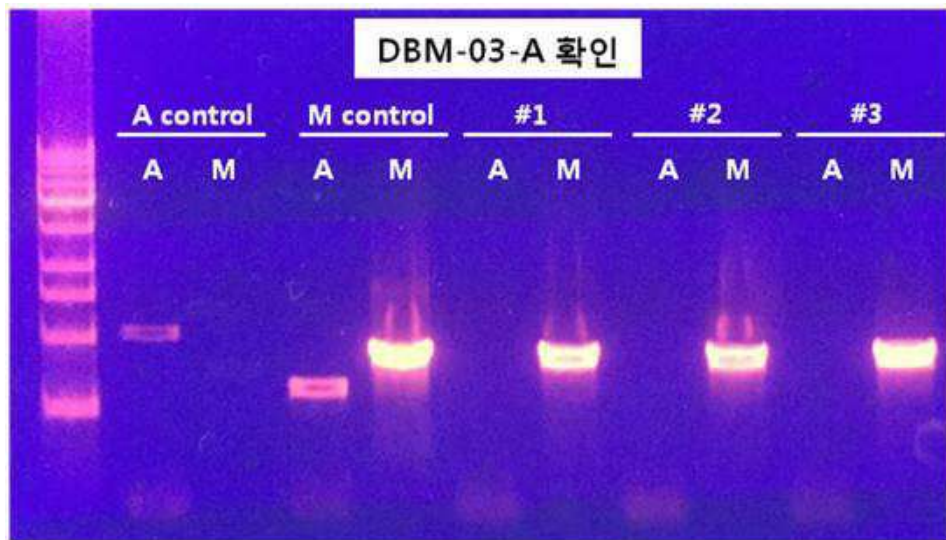
도면9b



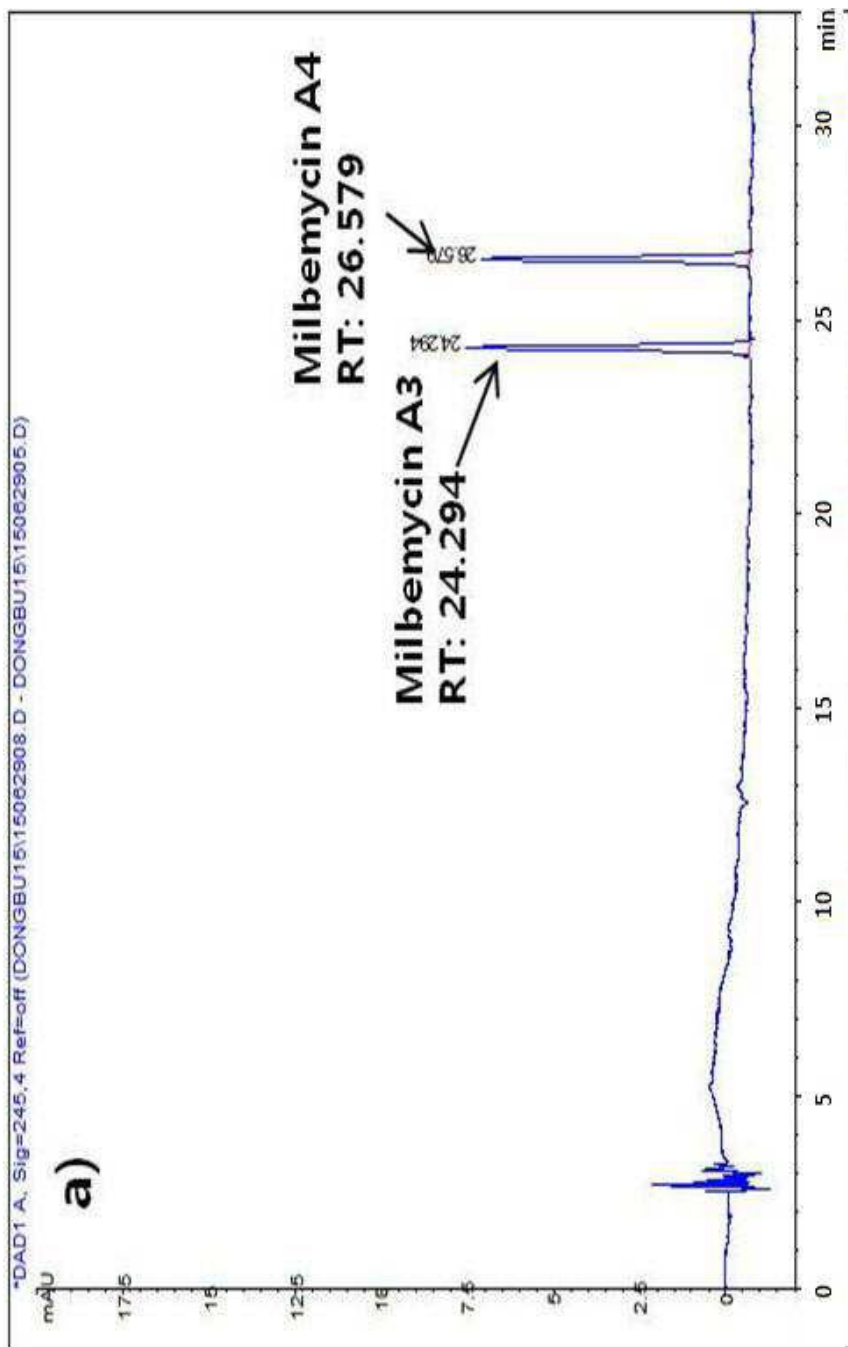
도면10a



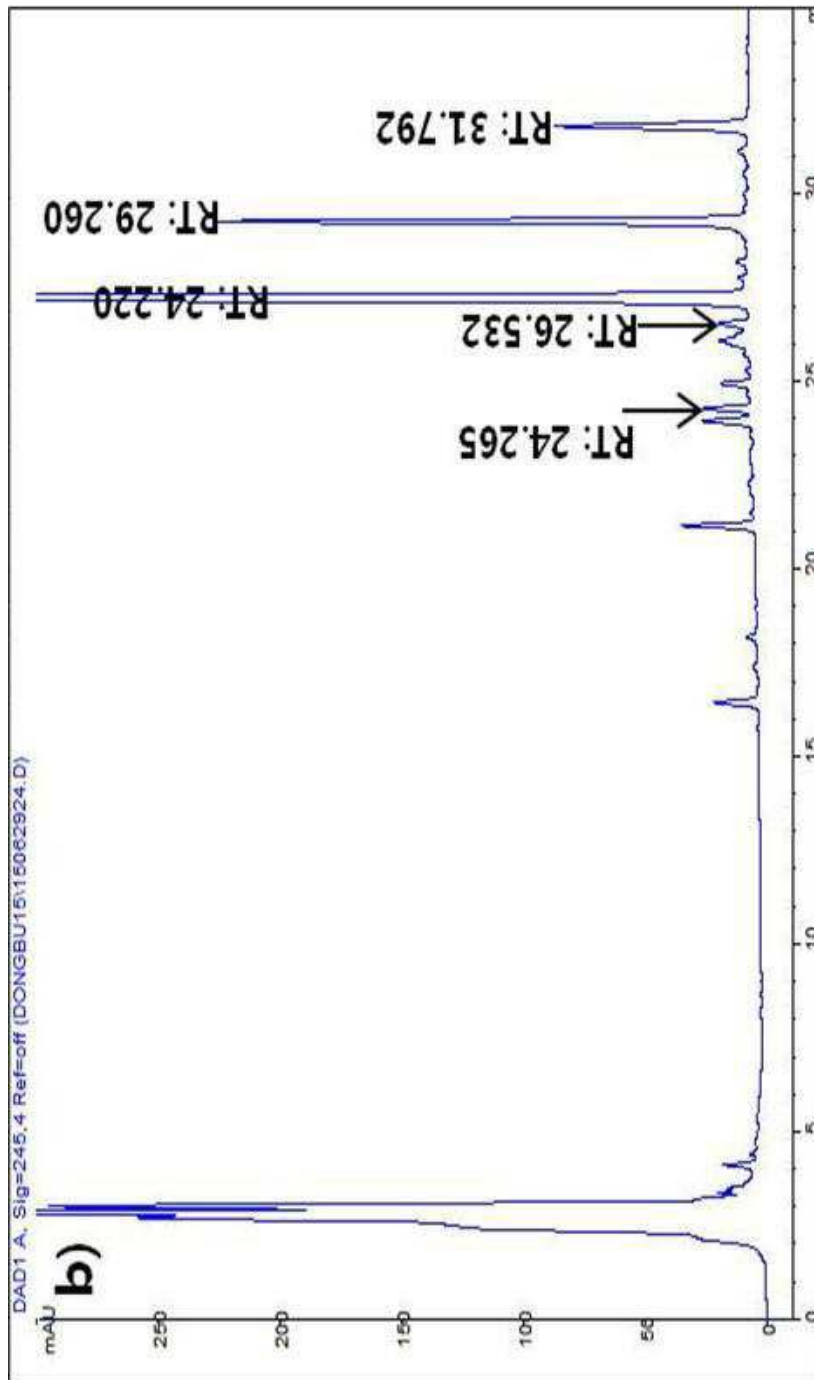
도면10b



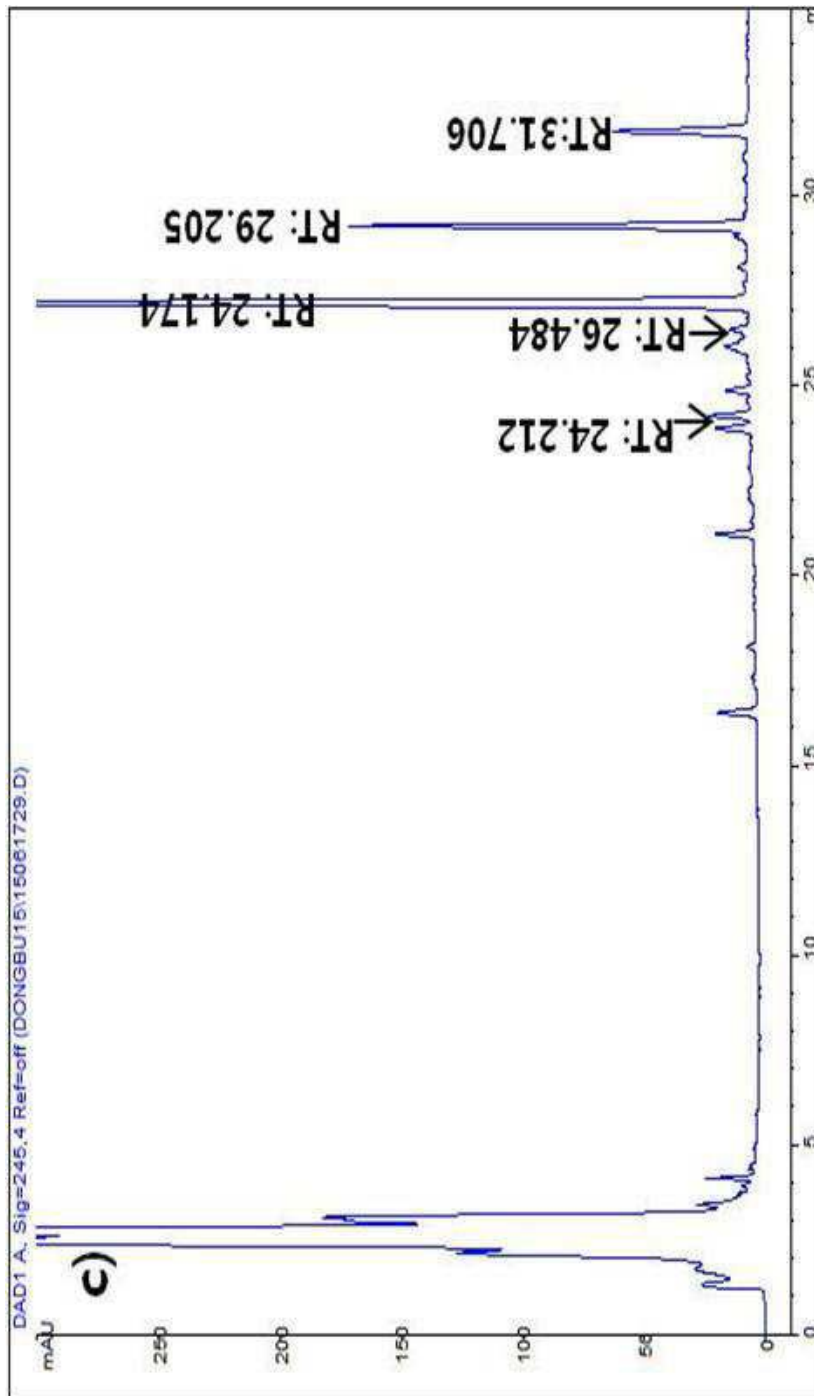
도면11a



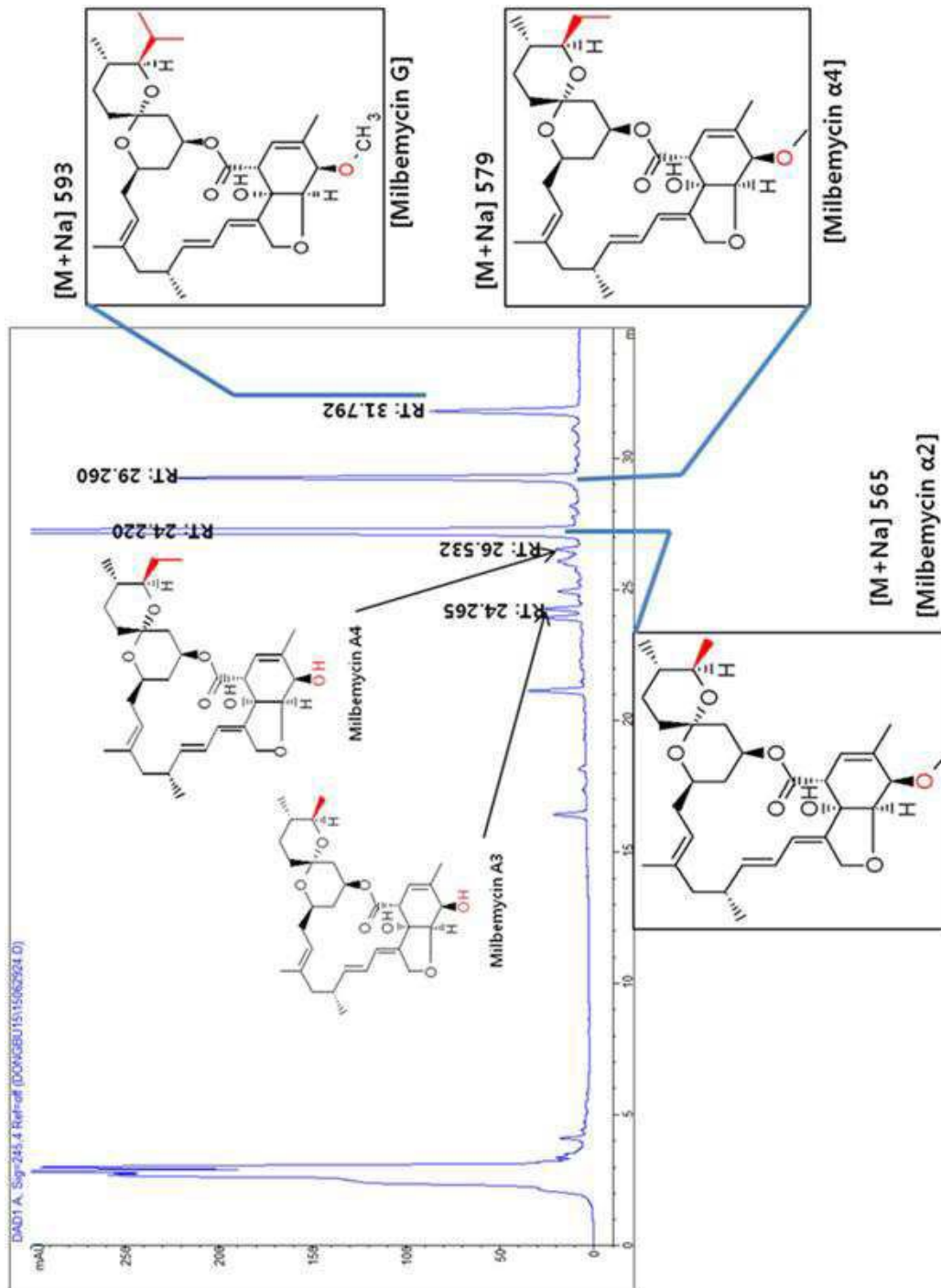
도면11b



도면11c



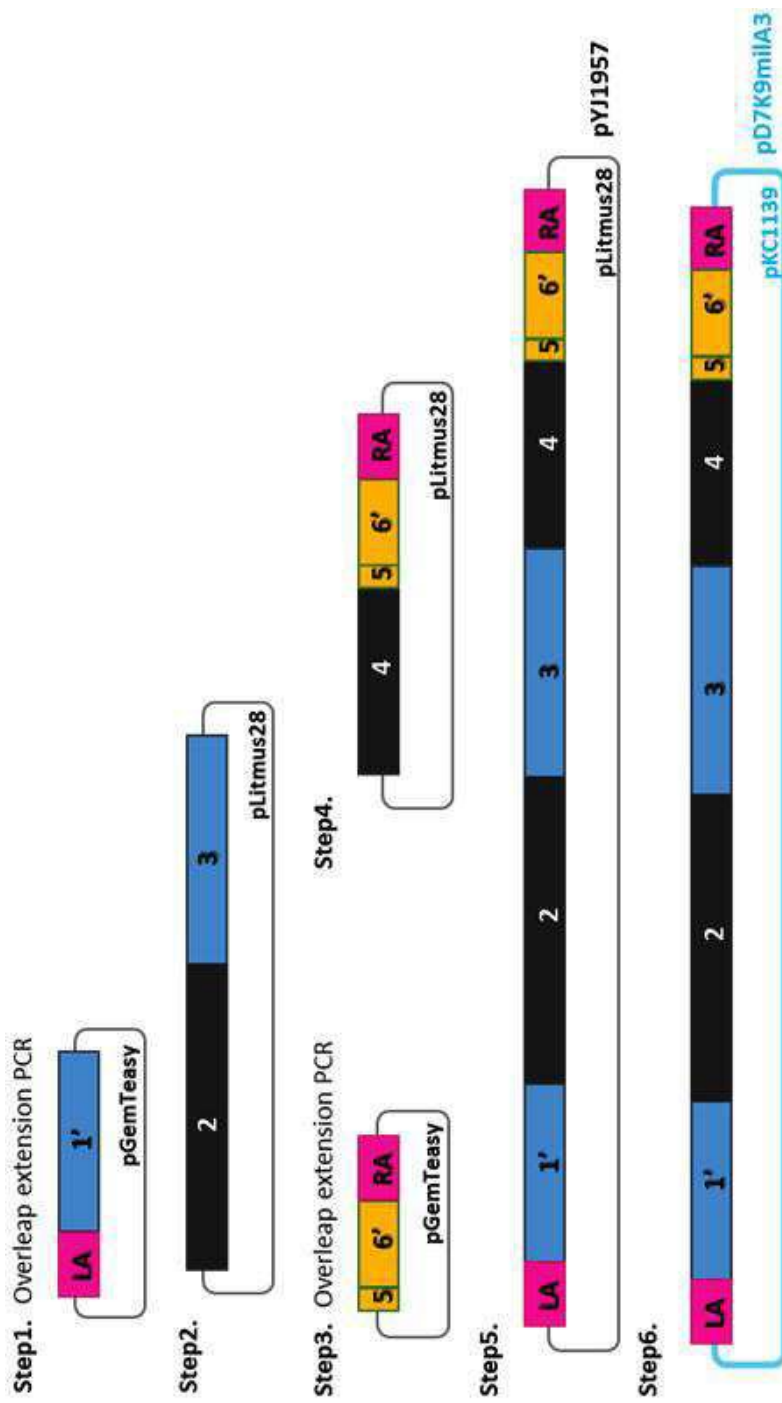
도면12



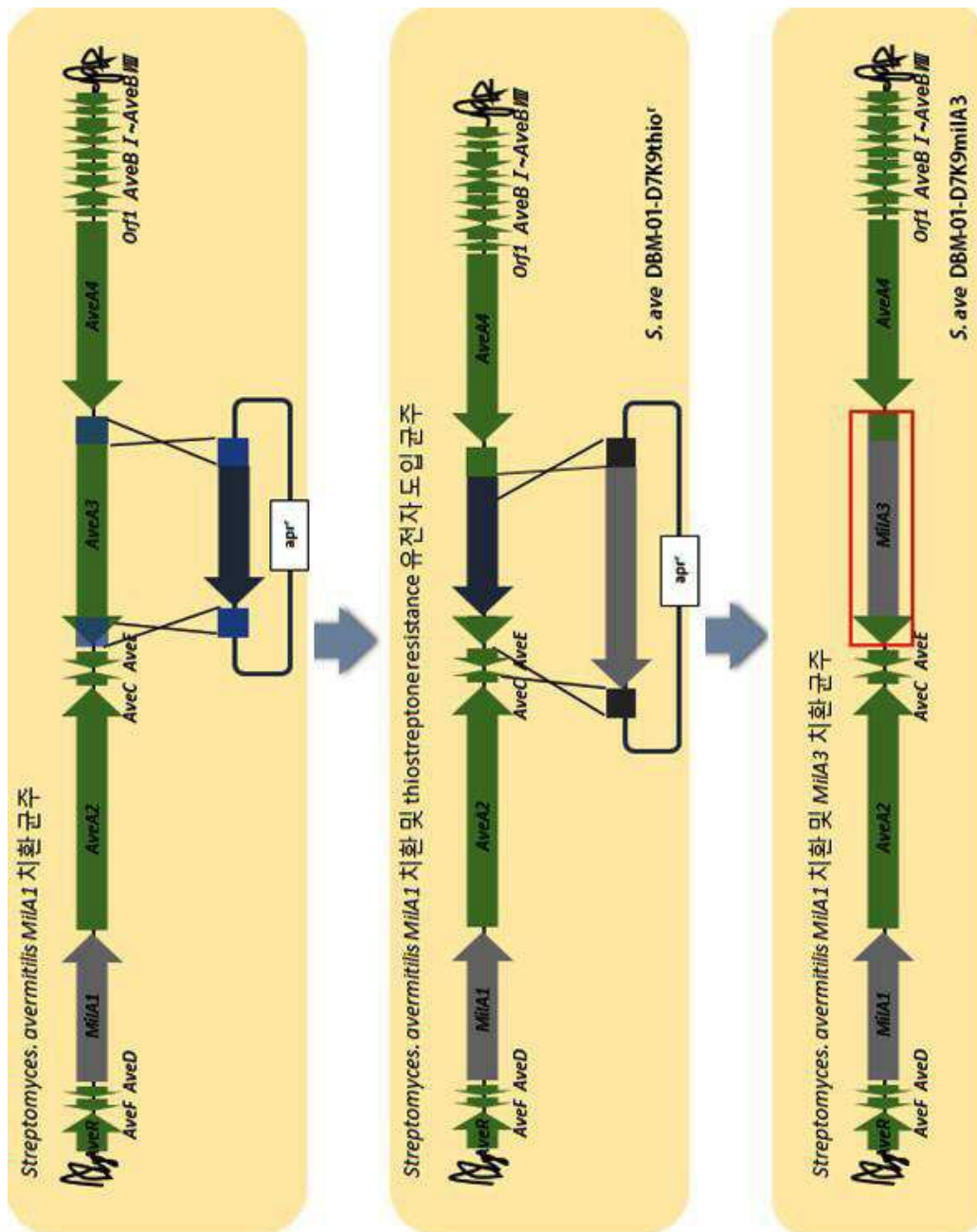
도면13

aveA3	890	AEDVTAIGSLRRGENDTRREFALAMTHHTTIGTFTTNHHYTHHHHTPHNHIDLEPTYP T +GSLRRG+D R LA+LA H G+ P W HY P +R+DLPTV	[] : A17-DH7 intermodular region
	868	GVAATVVGSLRRGEDDAHRILSLARAHISL--PVANDHY--PALAPEANHVDLPYIA	
	950	FORQYWLDAFTGAGDYAAAGLEPAEHPLLAATVOLADTDGCLLTGRLSLRSHFWLGDYE DH7	
	924	FOR+ YMLDAP GDV AAGL P HPLL A V LA+ DG LLTGRL+ +HFWL D+ FORRYWLDAFTGDTAAGLAPVGHPLLGAAVGLAEGGYLLTGRLAPHTHFWLTDHA	
	1010	VGGAVLLSGSAFVELAVOVGERVGTTRIEQLTVHAPLVVPGGVGGVSVOTGVAAADGEGRR V G VLL G+A+VELAV VG +GC R+E+LT+HAPLV+P GGV++QV V A D GRR	
miIA3	984	VAGTVLLPGTAYVELAVHVGHLGCPALELTTHAPLVLPDTGGVALQVAVGAPDETGRR	[] : fusion site / restriction enzyme
aveA3	5221	AAGILGSAGQGNYYAANAALDALAYRRRAAGLPGVSLANGLWEASGMTGHLAGTDHRI KR9	[] : KR9-ACP9 intermodular region
	5520	AAG+LG+AGQGNYYAANA LDALA RPA GLP +SLANGLWE+ SGMTHGL TD RI AAGVILGAAGQGNYYAANAFLDALAEHRRPAQGLFALSANGLWEQSGMTGHLDRTRARI Stu I	
	5281	IRSGLHPNSTPDALALFDALALDRPVLLPA--DLR---PAPPLPQLLDLLPATRRRTT RSGL P++T DALALFDALAL DRP L+PA DLR A PL + PA RTT	
	5580	NRSGLAPLATEDALALFDALAGDRPFLVPAFLDLRGSAAETPAPLFSRIAPA---RTT	
	5336	RTTTGGADNGAQLHGRLAGQTHEQQHTLLALVRSNIATVLGHTTPTDTPPDPAFRDYG R + GA+ A L RLA Q +Q TLI +VR+H A VLGH T + PD AFR+IG	
miIA3	5637	RGRSP-GAEGNADLRTRLAQAQDAAEQRDTLLIVRTHTAAVLGHDTAAAVRPDGAFRELG ACP9	[] : KR9-ACP9 intermodular region

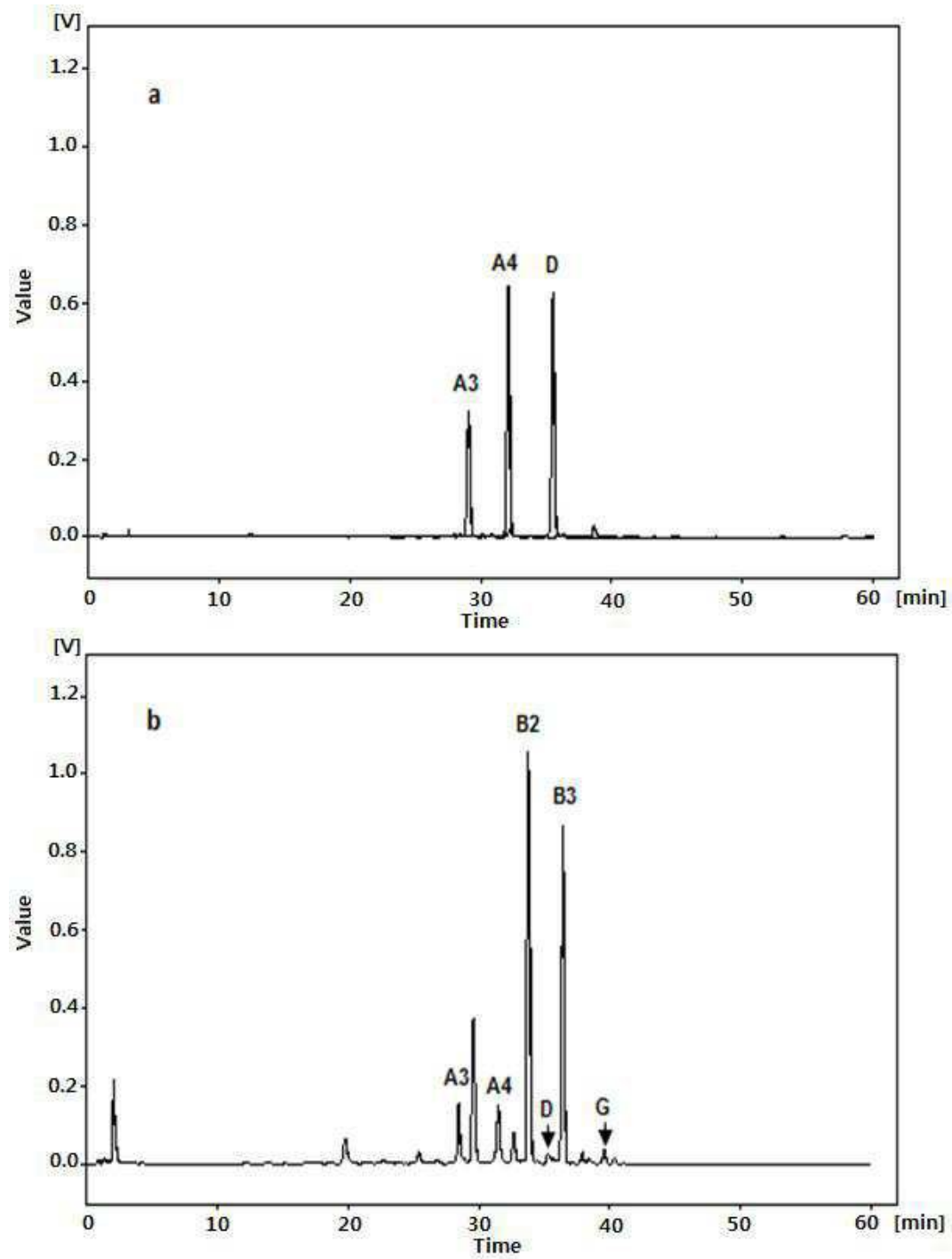
도면14



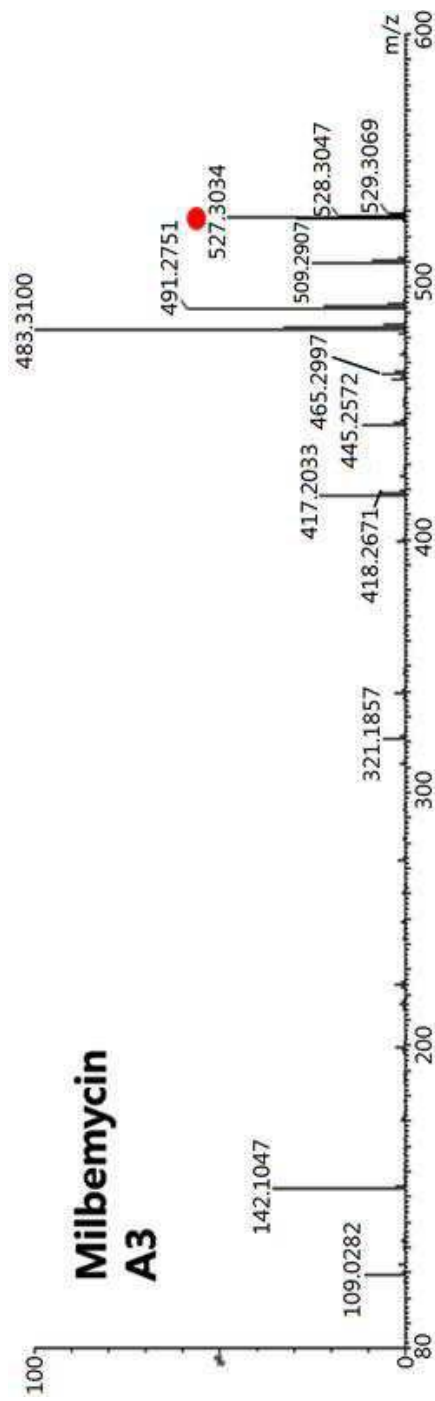
도면15



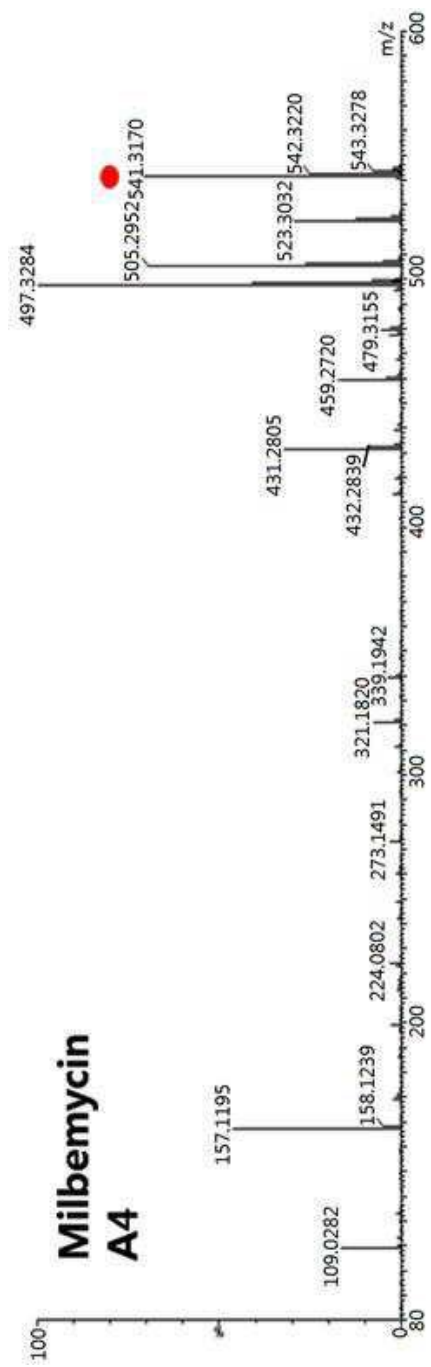
도면16



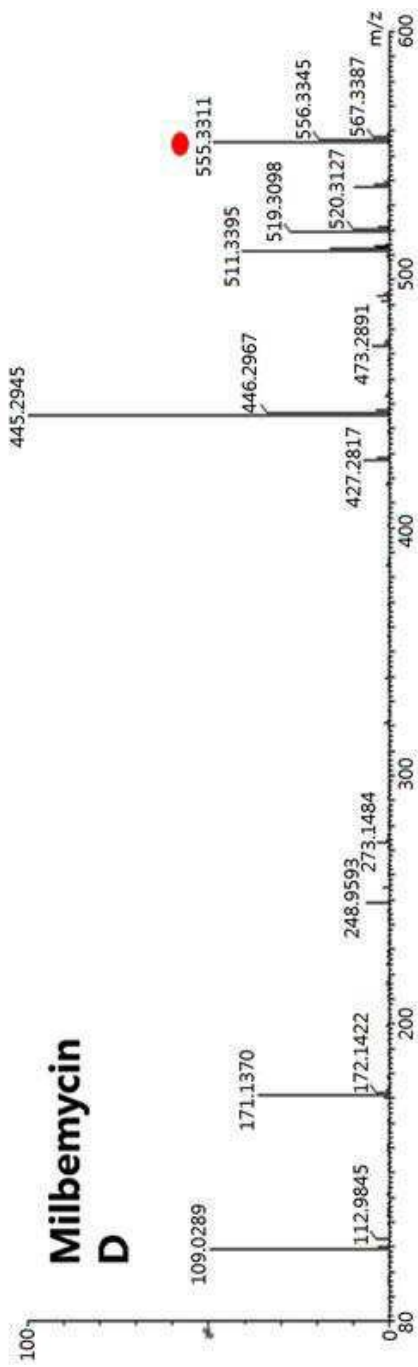
도면17a



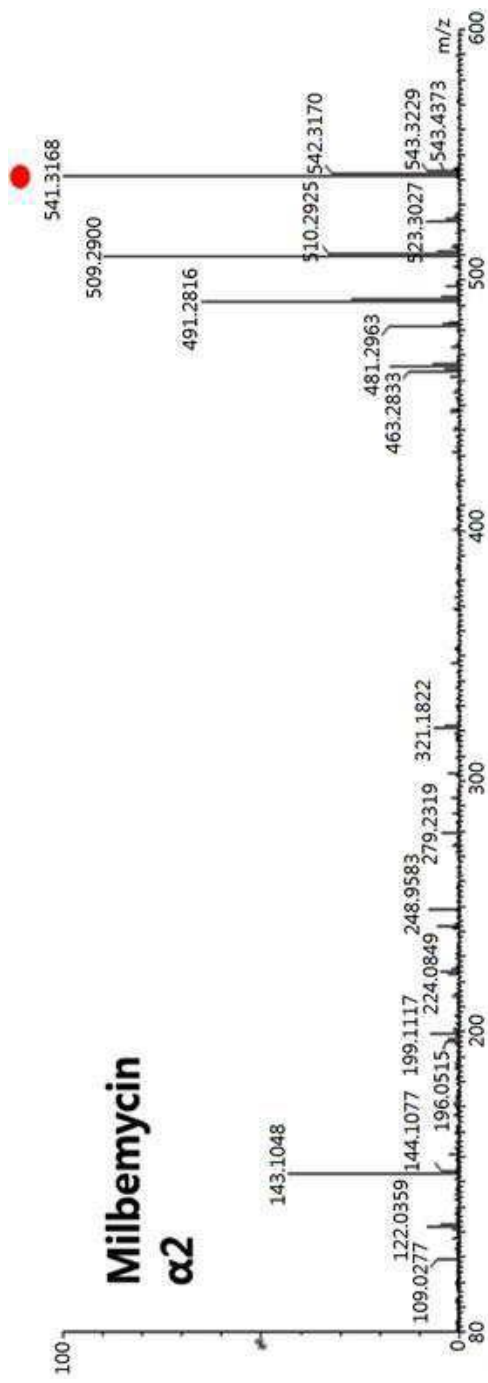
도면17b



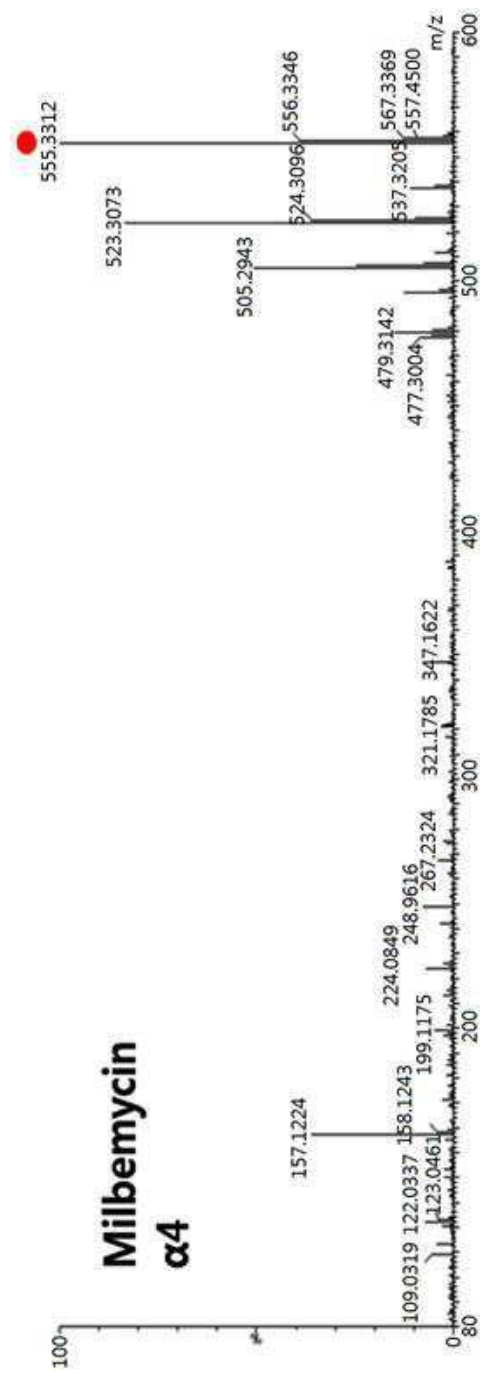
도면17c



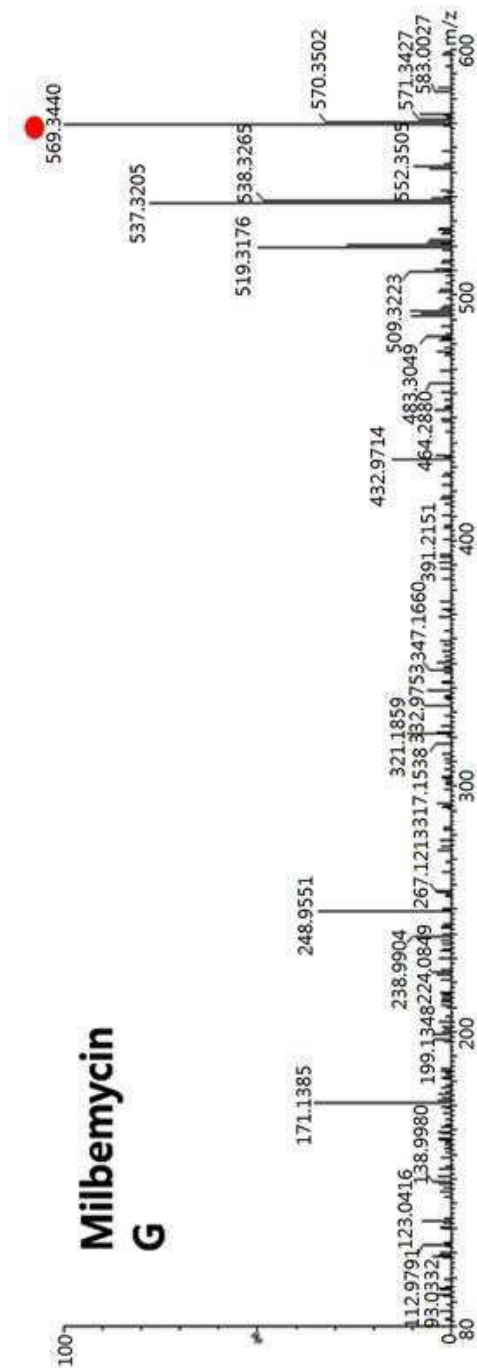
도면17d



도면17e

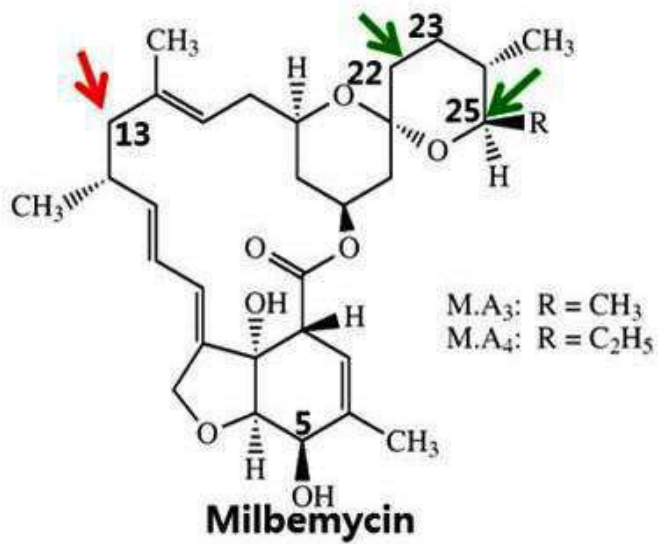
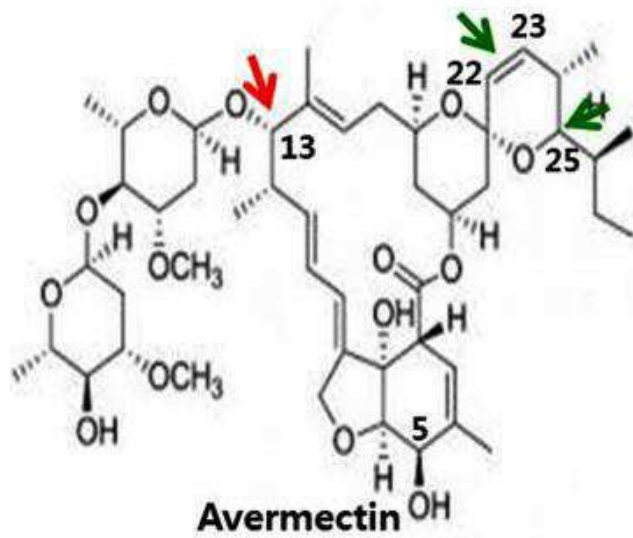


도면17f



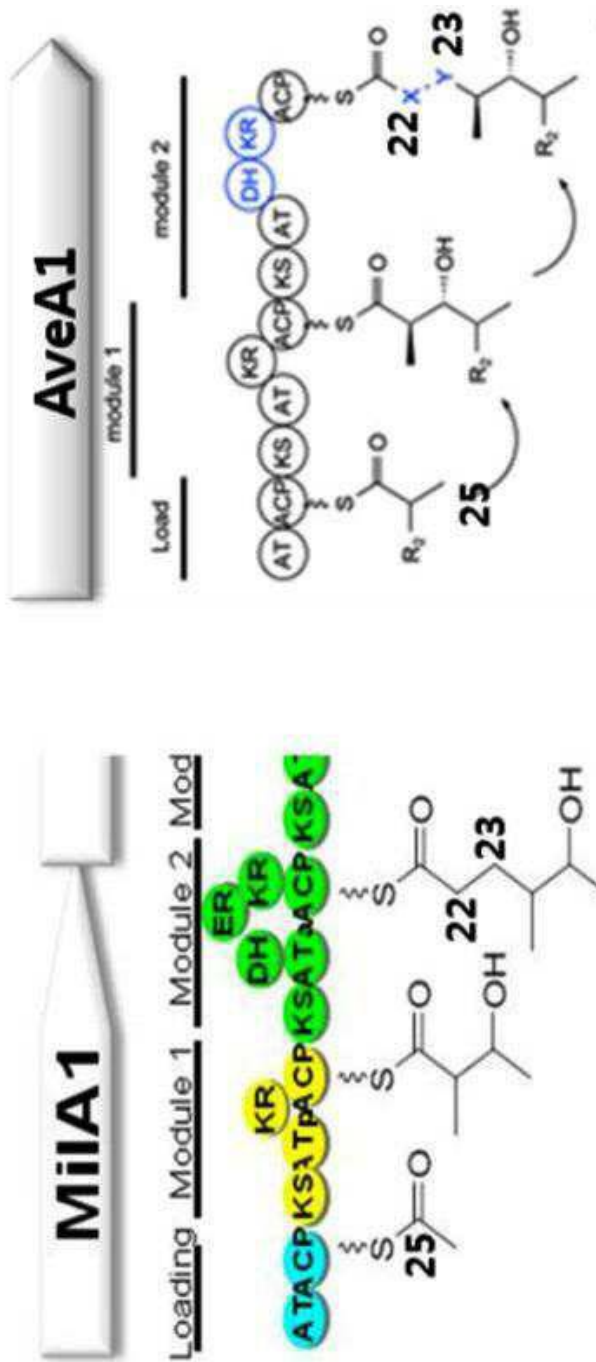
도면18a

Milbemycin 과 Avermectin



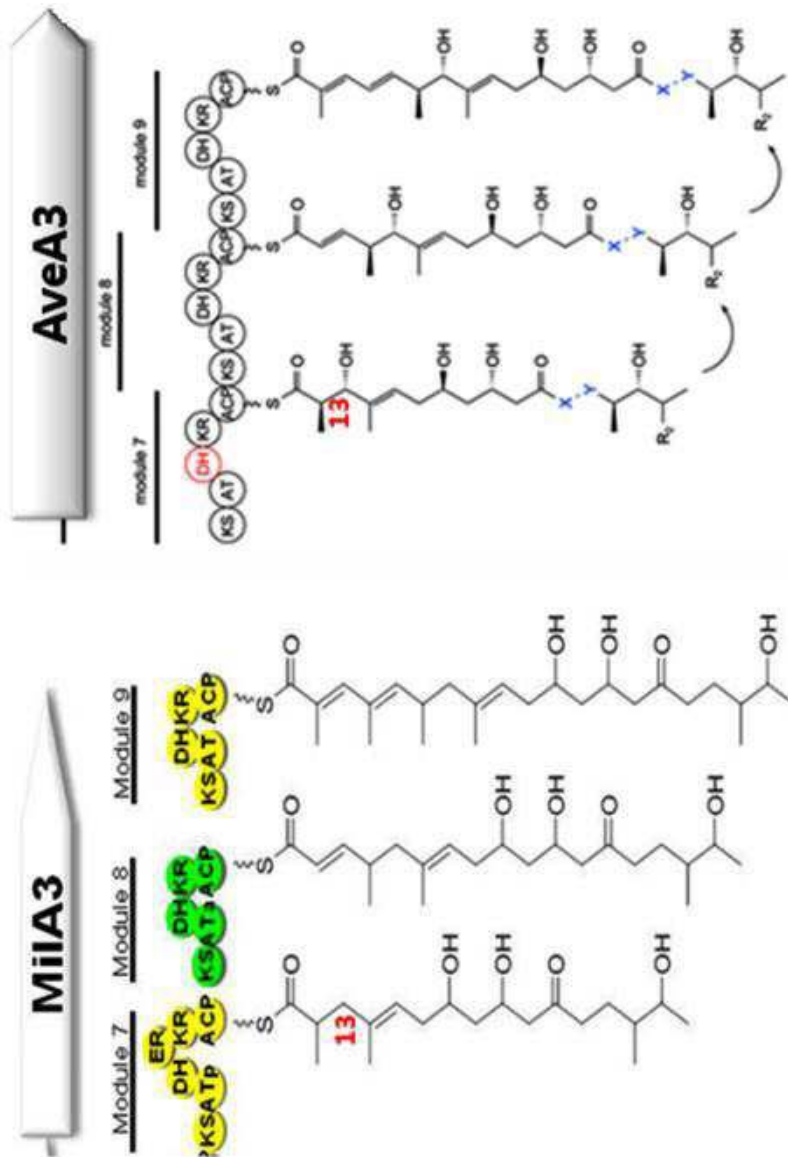
도면18b

MilA1과 AveA1의 차이

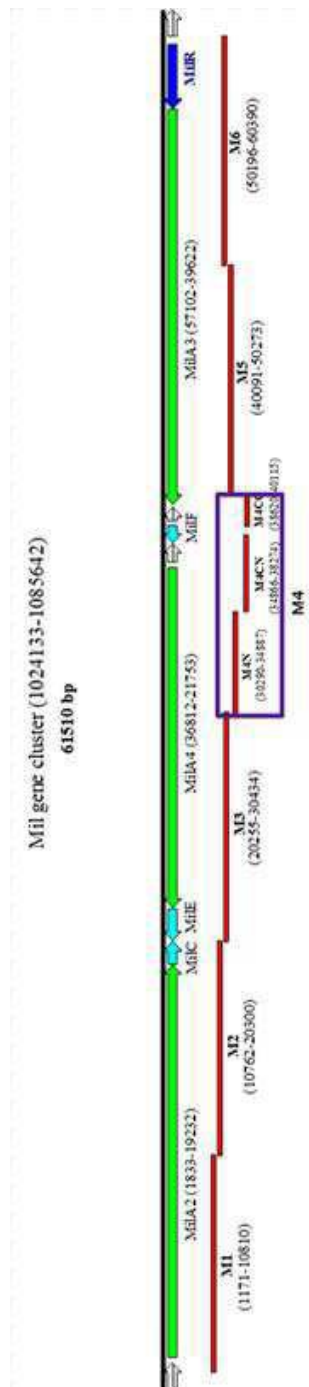


도면18c

MiIA3과 AveA3의 차이



도면19



서열 목록

- <110> Dongbu Farm Hannong Co., Ltd.
Ewha University - Industry Collaboration Foundation
- <120> Recombinant microorganism producing milbemycin and method for
producing milbemycin using the same
- <130> DPP20163328KR
- <150> KR10-2015-0133892
- <151> 2015-09-22

<160> 80
 <170> Kopatent In 1.71
 <210> 1
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220><223> A1NT2-F primer
 <400> 1
 gaaccgtgtg cggaagaa 18

<210> 2
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> A1CT1-R primer
 <400> 2
 atcaggtcgg tgagacggt 19

<210> 3
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> A1CT1-F primer
 <400> 3
 gcaaggtcgt ctttgtgtgc 20

<210> 4
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> A1CT2-R primer
 <400> 4
 ccaccaactc gatatgggc 19

<210> 5
 <211> 17
 <212> DNA

<213>	Artificial Sequence	
<220><223>	A2-404F primer	
<400>	5	
	ggcgctggca atggacc	17
<210>	6	
<211>	34	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	A2PH-1500R primer	
<400>	6	
	gattacgcca agcttggtg attctcccgg gacg	34
<210>	7	
<211>	19	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	AveD2-F primer	
<400>	7	
	gatgacggaa atcccctgg	19
<210>	8	
<211>	18	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	AveD2-R primer	
<400>	8	
	agaatgaaag gagcgca	18
<210>	9	
<211>	33	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	MA1NU-F primer	
<400>	9	
	gaggacacag ctcagttgcc caaagcccag aac	33
<210>	10	
<211>	19	

<212> DNA
 <213> Artificial Sequence
 <220><223> M1Age-R primer
 <400> 10
 tgatcgaaga gccgggagt 19

<210> 11
 <211> 33
 <212> DNA
 <213> Artificial Sequence
 <220><223> PHR-M1Ag primer
 <400> 11
 ccggctcttc gatcaagctt ggcgtaatca tgg 33

<210> 12
 <211> 34
 <212> DNA
 <213> Artificial Sequence
 <220><223> PXF-AD2 primer
 <400> 12
 gggatttcg tcatcaagct tggcgtaatc atgg 34

<210> 13
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> aveFi-F primer
 <400> 13
 ggcaacacca accgtgagat 20

<210> 14
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> milAli-R primer
 <400> 14
 gatgaccgag gagtgaacg 20

<210> 15
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> milA1i-F primer
 <400> 15
 gtactggacc tgttcgagcg 20
 <210> 16
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> aveA2i-R2 primer
 <400> 16
 atgaatgcgg tctttcctgc 20
 <210> 17
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> aveA1C-F primer
 <400> 17
 ggactcgctg attcggacaa 20
 <210> 18
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> aveA1C-R primer
 <400> 18
 ggagcagatg aatgccctct 20
 <210> 19
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> milA1C-F primer

<400>	19	
cctgtgggag ttcacacccg		20
<210>	20	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	milA1C-R primer	
<400>	20	
cgagattgtg gtggacgagg		20
<210>	21	
<211>	25	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	MA3-6503R primer	
<400>	21	
tgccaatgat cgcgatggcc tcatc		25
<210>	22	
<211>	33	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	MA3-104Fin(A) primer	
<400>	22	
gggtcagtg cacaacgcc catgtgatcc tcg		33
<210>	23	
<211>	31	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	MA3-1368Ffu(B) primer	
<400>	23	
ggaggcgtcc agcgaaccca tcgcatcat c		31
<210>	24	
<211>	29	
<212>	DNA	

<213>	Artificial Sequence	
<220><223>	AA3-8600RH primer	
<400>	24	
	aaaaagctta acggtgtgtg gagcgtcag	29
<210>	25	
<211>	25	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	AA3-5597FSp primer	
<400>	25	
	aaaactagtc gcgatcattg gcatg	25
<210>	26	
<211>	27	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	AA3-111RSp primer	
<400>	26	
	aaaactagtg ctggacgcct ccatggc	27
<210>	27	
<211>	28	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	AA3-1369RSp primer	
<400>	27	
	aaaactagtc gtgccactga cccggaac	28
<210>	28	
<211>	29	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	AA4-2300FNs primer	
<400>	28	
	aaatctagag ctccgtgcag accgagaac	29
<210>	29	
<211>	20	

<212> DNA
 <213> Artificial Sequence
 <220><223> aac3iv-F primer
 <400>
 > 29
 ggttcatgtg cagctccatc 20
 <210> 30
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <220><223> aac3iv-R primer
 <400> 30
 cgtcgcggtg agttcagg 18
 <210> 31
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> Amo7i-F primer
 <400> 31
 gttgtcgggtg ttgcttccg 19
 <210> 32
 <211> 20
 <212> DNA
 <213> Artificial Sequence
 <220><223> Amo7i-R primer
 <400> 32
 cacatcacac gcaacaatcc 20
 <210> 33
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> Mmo7i-F primer
 <400> 33
 acatcagcgg cctctacga 19

<210> 34
 <211> 19
 <212> DNA
 <213> Artificial Sequence
 <220><223> Mmo7i-R primer
 <400> 34
 ccaggtccat cactccctc 19

<210> 35
 <211> 1005
 <212> DNA
 <213> Artificial Sequence
 <220><223> pGEMB2-MA-DOWN-fu (Fig. 3)
 <400> 35
 ggatccagga cccggcggac acccccccgcg gccgccgcca ccgggcggtt cccctgcat 60
 accgcctcg gcggtcacgc cccggccgag cagcacgaga tgctgctctc gctggtcgcg 120
 tcccacgcc cctcgtgct gggccgcgac gatccggaca cgggccatcc cggcgcgcac 180
 ttccgcggcc tgggcttcga ctccctgacc gcggtcgagc tccgcaaccg gctcaacgcc 240
 gccaccggcc tccggctctc caccacctc gtcttcgacc accccacgcc cgacgaactc 300
 gcccgtcacg tccgggagca ggtgctgggc gacggcgaag cggcgcgggt ggccccctg 360
 ttggcggagc tcgacaaact cgaatccgcc ctctccgccc tcgacaaaac cgacagcgcc 420
 agcgaagag tcaccctgcg gctgaagtca ctcatgtga ggtggaacgc accccagcat 480
 ccgacagccg aaagcgctga tgacgacgag aagttcacat cggcaacaga ggctgagatt 540
 ttcaaattca ttgacaacga cctcggcctg tctgaaccg gacgcctgcc actccgcccg 600
 tatccgtgg gccctgctag gacgtgaatg caattggcga atgaagcgaa gtccttgaa 660
 tacctcaagc gcgtcactgc ggacctggac cgcaactgcc gtcgcctgta cgaggtggtc 720
 gagcgtgagc aggagccgat cgcgatttg gggatggcgt gtcgttacc aggcggggcg 780
 acgtcaccca cgcgactgtg gcatctctc aagtcacga cggacgtat cggggagtgc 840
 ccgaccgacc gtgatggaa cctggagcag ctctacgacc cggacccga ccgctcagga 900
 accagttaca cgcgcagcgg agggtttctc tatgacgcgg gcgacttcga cgccgcgttc 960
 ttcgagttgt caccgcgtga ggcgtggca atggaattca agctt 1005

<210> 36
 <211> 191

<212> PRT

<213> Artificial Sequence

<220><223> MilA1 sequence (module 2) and AveA1 sequence (Docking domain) in

pGEMB2-MA-DOWN-fu (Fig. 3)

<400> 36

Gly Ser Arg Thr Arg Arg Thr Pro Pro Ala Ala Ala Thr Gly Pro

1 5 10 15

Phe Pro Leu His Thr Arg Leu Gly Gly His Ala Pro Ala Glu Gln His

20 25 30

Glu Met Leu Leu Ser Leu Val Arg Ser His Ala Ala Leu Val Leu Gly

35 40 45

Arg Asp Asp Pro Asp Thr Val His Pro Gly Ala His Phe Arg Gly Leu

50 55 60

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

65 70 75 80

Ala Thr Gly Leu Arg Leu Ser Thr Thr Leu Val Phe Asp His Pro Thr

85 90 95

Pro Asp Glu Leu Ala Arg His Val Arg Glu Gln Val Leu Gly Asp Gly

100 105 110

Glu Ala Ala Arg Val Ala Pro Val Leu Ala Glu Leu Asp Lys Leu Glu

115 120 125

Ser Ala Leu Ser Ala Leu Asp Lys Thr Asp Ser Ala Ser Glu Arg Val

130 135 140

Thr Leu Arg Leu Lys Ser Leu Met Leu Arg Trp Asn Ala Pro Gln His

145 150 155 160

Pro Thr Ala Glu Ser Ala Asp Asp Asp Glu Lys Phe Thr Ser Ala Thr

165 170 175

Glu Ala Glu Ile Phe Lys Phe Ile Asp Asn Asp Leu Gly Leu Ser

180 185 190

<210> 37

<211> 122

<212> PRT

<213> Artificial Sequence

<220><223>

> AveA2 sequence in pGEMB2-MA-DOWN-fu (Fig.3)

<400> 37

Met Gln Leu Ala Asn Glu Ala Lys Leu Leu Glu Tyr Leu Lys Arg Val

1 5 10 15

Thr Ala Asp Leu Asp Arg Thr Arg Arg Arg Leu Tyr Glu Val Val Glu

20 25 30

Arg Glu Gln Glu Pro Ile Ala Ile Val Gly Met Ala Cys Arg Tyr Pro

35 40 45

Gly Gly Ala Thr Ser Pro Thr Arg Leu Trp His Leu Val Lys Ser Gln

50 55 60

Thr Asp Ala Ile Gly Glu Phe Pro Thr Asp Arg Gly Trp Asn Leu Glu

65 70 75 80

Gln Leu Tyr Asp Pro Asp Pro Asp Arg Ser Gly Thr Ser Tyr Thr Arg

85 90 95

Ser Gly Gly Phe Leu Tyr Asp Ala Gly Asp Phe Asp Ala Ala Phe Phe

100 105 110

Glu Leu Ser Pro Arg Glu Ala Leu Ala Met

115 120

<210> 38

<211> 50

<212> DNA

<213>

> Artificial Sequence

<220><223> DK1F primer

<400> 38

gcgtccgggg tcgacatggg atgcaggcct gagcggttga tgcgggcccg

50

<210> 39

<211> 33

<212> DNA

<213> Artificial Sequence

<220><223> DK1R primer

<400> 39

atgggactgc ggctgcttaa gacctcacc gtc	33
<210> 40	
<211> 18	
<212> DNA	
<213> Artificial Sequence	
<220><223> DK2F primer	
<400> 40	
gctagcgggg tgtggtcg	18
<210> 41	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> DK2R primer	
<400> 41	
ttcgaacacg cgggaatcga c	21
<210> 42	
<211> 21	
<212> DNA	
<213> Artificial Sequence	
<220><223> DK3F primer	
<400> 42	
gcatgccgtg tccacgcgc c	21
<210> 43	
<211> 49	
<212> DNA	
<213> Artificial Sequence	
<220><223> DK3R primer	
<400> 43	
acaccaccc ccacaaccac cacctagatc tgcccaccta cgccttcca	49
<210> 44	
<211> 53	
<212> DNA	
<213> Artificial Sequence	

<220><223>	DKLF primer	
<400>	44	
aacccgaacg accgagcgca gcggcggccg cgaattcata cggctacgag gat		53
<210>	45	
<211>	59	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	DKLR primer	
<400>	45	
ctcgaccgca ccgaccgggc ccgcatcaac cgctcaggcc tgcatcccat gtcgacccc		59
<210>	46	
<211>	61	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	DKRF primer	
<400>	46	
gccagtagcg gcggcgctgg aaggcgtagg tgggcagatc taggtggtgg ttgtgggggt		60
g		61
<210>	47	
<211>	73	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	DKRR primer	
<400>	47	
cttaggaccg ttatagttac gcgcttcctc gctcactgac tttaattaaa agcttacgtt		60
tcgcttgccg atg		73
<210>	48	
<211>	28	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M1171-1NF primer	
<400>	48	
aaaatgcatc taccgccgct tcaacagc		28

<210> 49
 <211> 30
 <212> DNA
 <213> Artificial Sequence
 <220><223> M10810-1R primer
 <400> 49
 aaaactagtg gagaggaaga aggagagccg 30
 <210> 50
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 <220><223> M10762-2NF primer
 <400> 50
 aaaatgcatc acccacacag gacaggga 28
 <210> 51
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 <220><223> MC20300-2R primer
 <400>
 > 51
 aaaactagta tgacaggccg ggtacatca 29
 <210> 52
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 <220><223> M20255-3NF primer
 <400> 52
 aaaatgcatc gcgtacctga tgtccctga 29
 <210> 53
 <211> 28
 <212> DNA
 <213> Artificial Sequence
 <220><223> M30434-3R primer
 <400> 53

aaaactagtg gccacacggt cttcatcg	28
<210> 54	
<211> 28	
<212> DNA	
<213> Artificial Sequence	
<220><223> M30377-NF2 primer	
<400> 54	
aaaatgcatc ctgcacaccg accgagac	28
<210> 55	
<211> 28	
<212> DNA	
<213> Artificial Sequence	
<220><223> M40160-4R primer	
<400> 55	
aaaactagta ccttcgtacc cgtctcgc	28
<210> 56	
<211> 29	
<212> DNA	
<213> Artificial Sequence	
<220><223> M40091-5NF primer	
<400> 56	
aaaatgcatg tgccgacgat cgtgagaag	29
<210> 57	
<211> 27	
<212> DNA	
<213> Artificial Sequence	
<220><223> M50273-5R primer	
<400> 57	
aaaactagta acacgcggga atcgacc	27
<210> 58	
<211> 29	
<212> DNA	
<213> Artificial Sequence	

<220><223>	M50196-6NF primer	
<400>	58	
	aaaatgcatg tagtccgtcg gcatcacac	29
<210>	59	
<211>	28	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M60390-6R primer	
<400>		
	59	
	aaaactagtc gagaccgtcc ctatgtcc	28
<210>	60	
<211>	18	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M30290-F primer	
<400>	60	
	acgtcaggaa gcggttca	18
<210>	61	
<211>	18	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M34887-R primer	
<400>	61	
	cgtccgaacg tttctgga	18
<210>	62	
<211>	19	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M4M4502F primer	
<400>	62	
	gagtcccaga aacgttcgg	19
<210>	63	
<211>	20	

<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M4F7910R primer	
<400>	63	
ggatcgtcaa catctcctcc		20
<210>	64	
<211>	19	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M4C-ScF primer	
<400>	64	
cggccaccag aacaccttc		19
<210>	65	
<211>	20	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	M4C25R primer	
<400>	65	
acacgtttct cagatcgtc		20
<210>	66	
<211>	11919	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	aveA1 gene of Streptomyces avermitilis MA-4680	
<400>	66	
gtgcagagga tggacggcgg ggaagaacc cgccctgcgg caggggaggt cctcggagtg		60
gccgacgagg cggacggcgg cgtcgtcttc gtttttcccg ggcagggccc gcaatggccg		120
ggcatgggaa gggaacttct cgacgttcc gacgtcttc gggagagcgt ccgcgcctgc		180
gaagccgcgt tcgcgcccta cgtcgactgg tcggtggagc aggtgttgcg ggactcgccg		240
gacgtcccc ggctggaccg ggtggacgtc gtccagccga ccctgttcgc cgtcatgac		300
tccctggccg cctcttggcg ctcgcaagg gtcgagccgt gcgcggtgt gggacacagc		360
ctgggcgaga tcgcggcagc ccacgtctcg ggaggcctgt ccctggccga cgccgcacgc		420
gtggtgacgc tttggagcca ggcacagacc acccttgccg ggaccggcgc gtcgtctcc		480

gtcgccgcca cgccggatga gctcctgccc cgaatcgctc cgtggaccga ggacaacccg	540
gcgcggctcg ccgtcgagc cgtcaacgga ccccgagca cagtcgtttc cggtgccgc	600
gaggcgtcg cggacctggt ggccgacctc accgccgcgc aggtgcgac gcgcatgac	660
ccggtggacg ttcccgcca ctccccctg atgtacgcca tcgaggaacg ggtcgtcagc	720
ggcctgctgc ccatcacccc acgcccctcc cgcacccct tccactctc ggtgaccgac	780
ggccgcctcg acacccgca gctagacgcg gcgtactggt accgcaacat gtcgagcacg	840
gtccggttcg agcccgccgc ccggtgctt ctgcagcagg ggccaagac gttcgtcag	900
atgagccgc acccggtgct gaccatgggc ctccaggagc tcgcccggga cctgggcgac	960
accacggca ccgccgacac cgtgatcatg ggcacgtgc gccgcggcca gggcaccctg	1020
gaccacttcc tgacgtctct cgcccaacta cgggggcatg gtgagacgtc ggcgaccacc	1080
gtctctcgg cagcctgac gcgctgtcc ccacgcagc agcagtcgt gtcctggac	1140
ctggtgcgcg ccacaccat ggcggtgctg aacgacgacg gaaacgagc caccgcgtcg	1200
gatgccggcc catcggcgag ttctgcccac ctcggttcg actccgtcat ggggtgtcga	1260
ctgcgaacc gcctcagca ggccacgggc ctgcggttc ccgtgacgt catcttcgac	1320
cacaccagc cgcccgcggt cgccgcgcg ctccggaccg cggcgctcgg ccacctgac	1380
gaggacaccg cgccgtacc ggactcacc agcgccacg gaggcacgc agcggcggac	1440
gaccgatcg ccatcatcg catggcatgc cgtttcccg gcggagtccg gtcccgaag	1500
gacctgtgg agctggccgc ctcgggcgga gacccatcg ggccgttccc caccgaccg	1560
ggatggcca cggaacagc tcaccccag gacccacgc agcccgac gtctatccg	1620
cagggaggcg ggttcctca cgacgcggcg cacttcgac ccggttctt cggaatcagt	1680
ccactgagg cactggcat ggatccgag cagcggtgc tgcggagac gtctgggag	1740
gcgttcgagc gggcggaat cgatccgtg tcggtacgc ggtcccgtac gggcgtcttc	1800
gcgggcgcc tctccttca ctacggccc cgtatggaca ccgctcgtc ggaggcgcc	1860
gcggacgtgg agggccacat cctcaccgt accacgggca gcgtcctgtc gggccgtatc	1920
gcctacagct tcgggtgga agggccggcg atcacgtgg acacgggtg ctcgcatcg	1980
ctcgtgacg tgcattggc gtgccagtgc ctgcggtcgg gtgagtgcac gtcgcgctg	2040
gccggcgcg tctcggtcat gtccacctc ggcatgttca tcgagttctc ccggcagcg	2100
gggctgtcgg tggacggcag gtgcaaggcg tactcggtg cagccgacgg caccggtgg	2160
ggcgaggcg tcgggatgct gttggtggag cggttgtcgg atcggtgctc gctggggcat	2220

cggtgtctgg cggtggtacg cggcagtcgc gtcaaccagg acgtgtcgtc gaatgggctg	2280
acggcgccga acgttccggc tcaggagcgg gtgatccggc aggcgttggc gaacgcgggg	2340
ttgtccgtgg cggatgtgga tgtggtggag gggcacggga cgggcacgac gctgggtgat	2400
ccgatcgagg cacaggcgtt gctcgccacg tacgggcagc gggccggtga caggccgctg	2460
tggctggggt ctctgaagtc caacatcggg cacacatgg ctgccgcggg tgtgggtggg	2520
gtcatcaaga tggatgatgc gttgcgggag ggggtgttgc cgcggacgtt gcatgtggat	2580
aagccgtcgc cgcaggtgga ctggtccgcg ggggcggtgc ggctgctgac ggaggcgggtg	2640
ccgtggccgg gggacgcggc agggcggttg cggcgggcgg gagtgtcgtc gttcgggatc	2700
ggcggcacga atgcgcatgt gatattggag gaggcgccgg cggcgggggg ctgtgttgc	2760
gggggtgggg tgttggaggg tgctccgggt ctgccattt cggtagctga gtcggtggcc	2820
gctccagtgg ctgtgtctgc gccggtggct gagtccgtgc cggtagcggg gccggtgccg	2880
gttcctgtgc cgggtgtcgc taggtctgag gctgggttgc gggcgaggc ggaggcgttg	2940
cgtcagtagc tggcagtcgc gccggacgtt tcgcttgccg atgtgggtgc gggctctggc	3000
tgtggcggg ctgtgttga gcatcgtgcg gtcgtcttgg ccgcggaccg tgaggagctg	3060
gtgcaagggt tgggggcgct ggccgcgggt gagccggatc ggcgggtgac cacgggtcat	3120
gcgccgggtg gtgaccgggg cgggtgtcgtc ttcgtgtttc ccggacaggg tgggcagtgg	3180
gccgggatgg gtgtgcgtct gctgcctcc tctccggtgt tcgccggcg gatgcaggcg	3240
tgcgaggagg ctctggcgcc gtgggtggac tggctcttgg tggacatcct gcgccgggac	3300
gcgggggatg cgggtgtgga gccggccgat gtggtccagc ctgtgtgtt cagcgtcatg	3360
gtgtctttgg ctgtctgtg gcgttcctac ggtatcgaac ccgacgcggg ccttggccat	3420
tcccagggcg agatcgcgcc cgcgcagtgt tgtggggcgc tgagcctgaa ggacgcggcg	3480
aagactgttg cgtgtgcgag ccgggcgctg gccgtgtgc ggggccgggg cggcatggcc	3540
tcagtgccgc tgctgtccca ggaggtggag cagctcattg gtgagcgggt ggcggggcgg	3600
ttgtgggtgg cggcgggtcaa cgccccgcg tccaccgcg tctcggggga tgccgaggcg	3660
gtggacgagg tgctggcgta ctgtccggc accgggttgc gggcccgcg gatcccgtc	3720
gactatgcct cgcactgccc ccatgtgcag ccctgcggg aggagtgtct ggagctgctg	3780
ggggacatca gcccgcagc gtccggcgtg ccgttcttct ccacggtgga gggcaccttg	3840
ctggacacca caacctgga cggccctac tggtagcga acctgcacca gccgggtccgt	3900
ttcagcatg ccgtccaggc cctggcggat gacggacacc gcgtcttctg cgaagtcagc	3960

ccccaccca cctcgtccc cgccatcgaa gacaccaccg aagacaccgc cgaagacgtc	4020
accgcgatcg gcagcctccg ccgcggcgac aacgacaccc gccgttctct caccgccctc	4080
gcccacaccc ataccaccgg catcggcaca cccaccacct ggcaccacca ctacaccac	4140
caccacaccc accccaccc ccacacgcac ctgcacctgc ccactaccc ctccaacac	4200
cagcactact ggctcgagag ctacacccg ggtgccgat ccggttcggg tgccggtgcc	4260
ggttcgggtg ccggttccgg gcgggcaggg actgcgggcg ggacggcaga ggtggagtcg	4320
cggttctggg acgcggtggc ccgccaggac ctggaaacgg tcgcgaccac actcgccgtg	4380
ccccctccg ccggcctgga caggttggtg cccgcactct ccgctggca ccgccacaa	4440
cacgaccaag ccgcatcaa cactggacc taccaggaaa cctggaaacc ctcaccctc	4500
cccaccacc accaaccaca ccaaacctgg ctcatcgcca tccccgaaac ccagaccac	4560
cacccacaca tcaccaatc ctcaccaac ctccaccacc acggcatcac ccccatccc	4620
ctcaccctca accacacca caccacccc caacacctcc accacacct ccaccacacc	4680
cgacaacaag cccaaaacca caccaccgga gccatcaccc gcctgctctc cctcctgcc	4740
ctcgacgaaa caccacccc ccaccacccc cacacacca ccggcaccct cctcaacctc	4800
accctaccc aaaccacac ccaaacccac ccaccaaccc cctctggta cgccaccacc	4860
aacgccacca ccaccaccc caacgacccc ctcacacacc ccaccaagc ccaaacctgg	4920
ggactcgccc gcaccacct cctcgaacac cccaccaca ccgcgggaat catcgacctc	4980
cccaccacc ccacccccc caccctccag cactcacc aaacctcac ccaacccac	5040
caccaaacc aactcgccat ccgcaccacc ggcaccaca cccgccct caccaccacc	5100
accctaccc ccacacacca accaccacc cccaccccc acggaaccac ctcatacc	5160
ggcggaaccg gcgccctgc caccacctc accaccacc tcaccacca ccaaccacc	5220
caacacctcc tctcaccag ccgaaccggc cccacaccc cccacgcaca acacctacc	5280
acccaactcc aaaaaaagg catccacctc accatacca cctgcgacac cagcaacca	5340
gaccaactcc aacaactct caacaccatc cccccacaac acccctcac caccgtcatc	5400
cacacgcag gcatcctga cgacgccacc ctaccaaac tcacccccc ccaactaac	5460
aacgtctcc gcgcaaagc ccacagcgc cactctctc accaactcac ccaacacacc	5520
ccctcaccg cttcgtctc ctactctcc gccgcgcca ccttcggcgc acccgccaa	5580
gccaaactag ccgcagccaa cgcctacctc gacgccctg cccaccacc ccacaccac	5640
cactccccg ccaccagcat cgcctggggc acctggcaag gaaacggact cgtgattcg	5700
gacaaggccc gcgcatact cgaccgccgc ggttttcgac ccatgtcacc cgagttggcc	5760
acggcagcgg tcacgcaggc gatcgcgac accgaacggc cgtatgtcgt catcgccgac	5820

atcgactgga gcaagatcga acacacctct cagaccagcg acctggtgag cgcggcccgg	5880
gaaagggagc cagctgtcca gcgccccact ccaccggcgg agttgcacaa aacgttgccc	5940
catcagacgt cggccgacca acgggccgca ttgctcgagc tcgtacgaga ccatgtggcg	6000
gcagtgtctc ggacgcgga cccgaaagcc atcgcgccc accagtcgtt ccgtgcactc	6060
ggcttcgatt cactcacggc cgtcgagttc cgaaacctgc tgatcaaggc aacaggactc	6120
cgccttctctg tctcgctggt cttcgaccac ccgaccctg ccaaactcgc cgtacacctg	6180
cagaaccaac tgcggggcac agcagcggag tcggctcctt cagcggcagc cgttaccgcc	6240
gaggcttctg tcaccgagcc gatcgccatc gttggcatgg cctgtcgttt ccccgcgga	6300
gtgacctcgg cggacgactt ctgggatctg atctcctccg agcaggacgc gatcggcgga	6360
ttccccaccg acccgggctg ggacctggac acgctctacg accccgaccc cgaccacccc	6420
ggcacctgct acaccgaaa cggcggattc ctctacgacg caggccactt cgacgccgaa	6480
ttcttcggca tcagcccccg cgaagccctc gccatggacc cccagcaacg actcctctc	6540
gaaaccgcct gggaaacat cgaacacgcc ggcatcaacc cccacaccct ccacggcacc	6600
cccacggag tcttcaccgg caccaacgga caggactacg cacttcgctg gcacaacgcg	6660
ggccagtcaa ccgatggttt cgcactgacc ggaaccgccg gcagcgtcat ctccggtcgt	6720
atctcgtaca cgtttggtt tgagggtcct gcggtgtcgg tggacacggc ttgttcctcg	6780
tcgttggtgg ctttgcatct ggccgtgcag gcgttgctg cgggtgagtg ctcgatggcg	6840
cttgccgggg gtgtgacggt gatgtcgtct ccgggtgcct tcgtggagt ttcgcggcag	6900
cggggtctgg ccgcggacgg gcattgcaag gcgttctcgg cggcggcgga cgggaccggc	6960
tggggtgagg gtgtggggat gctgctggtg gagcggctct ccgacgcca tcgcaacggt	7020
caccgtgtcc tgccctggt gcgtggcagt gcggtcaacc aggacggtgc gagcaacggt	7080
ctgaccgcgc ccaacgggcc gtcccagcag cgtgtcatcc gccaggccct cgccaacgcc	7140
ggcttgtcgg ccggtgatgt cgacgcggtg gaggccacg gcaccggcac cactttgggc	7200
gaccgatcg agggccagcc cctcctcgcg acctacggac aggaccgtgc cggcgagggg	7260
ccgctgtggc tgggctcgtt caagtccaat gtcggtcaca cacaggctgc cgcggcgctc	7320
gccggggtga tcaagatggt gatggcgtg cggcatggtc tgctgccgcg gacgttgcat	7380
gtggatgagc cgtgccgca tgtggactgg tccgcggtg cggtcagct gctgacggag	7440
acggtgccct ggcccggcgg ggagggcgcg ctacggcggg caggagtgtc atcattcggc	7500
gtcagcgga ccaacgcca cgtcatctc gaagaagcac ccgccgacga cgttcgggg	7560
ggaccacccg ccggcgaggg tgacggggc agcgacgatg aggtgtgtgc cggcagtcct	7620
ggggtgtggc cgtggctggt gtcggccaag tcgcagcgg ccctgcgcgc ccaggcccag	7680

gccctgcacg cccacctcac cgaccacccc ggccctcgacc tcgcggatgt cggatacacc	7740
ctcgccccacg cccgcgccgt gttcgaccac cgcgccaccc tcatcgccgc ggaccgcgac	7800
acgttccctgc aagcactcca ggcaactgcc gcaggcgagc cccaccccgc cgtcatccac	7860
agcagcgccc cgggcgggac cgggaccggg gaggccgcag gaaagaccgc attcatctgc	7920
tccggacagg gcacccaacg ccccggcatg gcccacggcc tctaccacac ccaccccgtc	7980
ttcgccgccc cactcaacga catctgcacc cacctcgacc cccacctcga ccacccctc	8040
ctccccctcc tcacccaaaa cgacaacgac aacgaggacg cggccgcact gctccagcag	8100
accgctacg ccagcccgc cctcttcgcc ttccaggtcg cctccaccg cctctcacc	8160
gacggctacc acatcacccc ccactactac gccggacact ccctcggcga aatcacggcc	8220
gcccacctcg ccggcatcct caccctcacc gacgccacca ccctcatcac ccaacgcgcc	8280
acctcatgc aaacatgcc ccccggcacc atgaccaccc tccacaccac cccccaccac	8340
atcacccacc acctcacgc ccacgaaaac gacctcgcca tcgccgcat caacaccccc	8400
acctccctcg tcacagcgg cccccccac accgtccaac acatcaccac cctctgccaa	8460
caacaaggca tcaaaaacca aacctcccc accaaccag ccttccact cccccacacc	8520
aaccccatcc tcaaccaact ccaccagcac acccaaacc tcacctacca cccacccac	8580
acccccctca tcaccgcaa cccccacc gaccaactcc tcacccccca ctactggacc	8640
caacaagccc gcaacaccgt cgactacgcc accaccaccc aaacctcca ccaacacggc	8700
gtcaccact acatcgaact cggaccgcac aacacctca ccacctcac ccaccacaac	8760
ctcccaaac cccccacc caccctcacc ctacccacc cccaccacca ccccaaac	8820
cacctctca ccaacctgc caaaaccacc accacctggc accccacca ctacaccac	8880
cacgacaacc aaccacac ccacaccac ctgacctcc ccactacce cttcaaac	8940
caccactact ggtcgaag cacacagccc ggtgccggca acgtgtcagc agccggactc	9000
gacccaccg aacacccct actcggcgcc acattggaac tggcgactga cggaggagcg	9060
cttcttcgag ggcgcttgct tttaggtcg catccgtggc tggctgacca tgccgtcggc	9120
ggcacggtgc tgctgtcgg cgccaccttc ctggaactcg cccttcatgc gggcacatac	9180
gtgggctgc accgagtgga tgagctgacg ctgcatgcgc cgctggtggt tcctgtggat	9240
gggggtgtga gtgtgcaggt tggggttcg gctgcggatg gggaggggcg gcgtttggtg	9300
agtgtgatg cgcggggtgg gagtgttgt ggtgggggtg gtgcgtcggg tggggtgtgg	9360
acgtgtcatg cctcgggggt gctggttag gctgctgctg gtggtgtggt ggtggatggt	9420

ctggcggggg tgtggccgcc gcggggtgcg gtggcgggtgg atgtcgatgg tgtccgtgac	9480
cgttttgctg gggctggttg tgttttgggg ccggtgtttt cggggctgcg tgcggtgtgg	9540
cgtgatgggg gggatttctt ggctgaggtg tgtctgccgg aggaggcgtg gggatgatgcg	9600
gctggttttg ggctgcatcc ggcgttgctg gatggtgtgg tccagccgtt gtcggtgttg	9660
cttccgggtg ggacgggggt tggggagggg gcggggttcg gggagggtgt tccggtgccg	9720
gctgtgtggg gtggtgtgtc gcttcaccgg gcgggtgtga ccggtgtgcg ggtgcgtgtg	9780
tccgctgtcg ggcggggcgg cgggcgtgag gcggtgtcgg tcgtggtcgg ggatgaggcg	9840
gggtgtccgg tggcgtcggc cgatcgtctt gagttgcggc ctgtggatat gggtcagtgt	9900
cgtgctgtct cggtttcggc ggggcggcgg ggttcgctgt atgcggtgca gtgggctgag	9960
gtgggtcctg tgcgggtgtg tgggcaggcg tgggcgtggc acgaggacgt gggtagagac	10020
gggtgtgggc ctgtgccggg ggtggtgtgt ttgcggtgcc cggatgccgg tgcgggtggc	10080
gggtggcggg gcggtgtgtg cgggtgtgtg ggtgaggttg ttggtggggt gttgggtgtg	10140
gtgcagggtt ggctggggct ggagcgggtt gcgggttcgc ggctggtgtt ggtgaccggg	10200
gggtgcggtg tggccggccc ggaggacggc ccggtggatg tgggtgggtgc gtcggtgttg	10260
gggctgggtg gttcggcgca ggctgagcat ccggaccggc ttgtcctcct cgacctcgac	10320
accgacaccg gcaccgacct cgacaccggc gctggtgtctg gttggggcgt ggatggtggg	10380
cgtgtggcgg cgggtgtgtg gtgtggtgag ccgcagttgg cgggtcgtgg ggagcgggtt	10440
ctggccgcac gcctgaaacg acttgagtca tccggtgatg ttccagccca gcggtccggt	10500
gacacacgag cccggcggtc cgacgtgcct gcccagcgt ccggtggcgt gcctgtcgg	10560
cggctcgggt atgtatcggg tccggagggt ttccgtgtgt tgtcgggtgg gtcggtgttg	10620
gtgacgggtg ggacgggtgt gctgggtgcg gcggtggcgc ggcatctggc tgggtgtgt	10680
ggggtgcggg atctgtgtt ggtgagccgg cgtggtccgg atgtccggg tgcggagggt	10740
ctgcggcgcg agctggccgc gttggggcg gaggtgcgga ttgttgcgtg tgatgtgggg	10800
gagcggcggg aggtgttcg cgtgtgtgag ggtgttcctg ccgggtgtcc gctgacgggt	10860
gtcgtcatg ccgctgtgtt gctggacgat gcgacgatc cctctctcac gcccagcgg	10920
ctgggcacgg ttttcgggc caaggtggat gccgtcttt tgctggatga gctgacggg	10980
ggatatggag ttttcggctt cgtgtgttc tcctcgcccg cggggatcct ggggtcggc	11040
gggcagggca actacgcgc ggccaatgcc gctctggacg cgtggcgta ccggcgcg	11100
gcggcgggtc tgcgggggtt gtcgtgtggt tgggggtgtt gggaagaggc cagcgggatg	11160
accgggcacc tggccggcac cgaccaccgg cgcacatcc gttccggtct gcatcccatg	11220
tcgaccccg acgcactggc cctcttcgat gcggccctgg ctctggaccg gccggtcctg	11280

ctgccgcgcg acctgcgtcc cgccccgccc ctgccgcccc tgctgcagga cctcctgccc 11340
gccacccgcc gccgcaccac ccgcaccacc actaccggtg gtgcggacaa cggcgcccag 11400
ctgcacgccc ggctggccgg ccagacacac gaacaacagc acaccaccct cctcgccctg 11460
gtccgctccc acatgccac cgtcctgggc cacaccacc cgcacacat ccccccgac 11520

cgcgcggttc gcgacctcgg cttcgactcc ctcaccgcg tcgaactacg caaccggctc 11580
tcccgacca cgggactcgg cctccccacc accctcgctc tgaccaccc caaceccacc 11640
accctaccc accacctcca cacacaactc cagccacaac cggacaacgc tgcgcccc 11700
gtgttgccgg agctcgacaa actcgaatcc gccctctcgg cctcgacaa aaccgacagc 11760
gccagcgaaa gattcacct gcggctgaag tcactcatgt tgaggaggaa cgcaccccag 11820
catccgacag ccgaaagcgc tgatgacgac gagaagttca catcggcaac agaggctgag 11880
attttcaaat tcattgacaa cgacctcggc ctgtcctga 11919

<210> 67
<211> 3972
<212> PRT
<213> Artificial Sequence
<220><223> type I polyketide synthase AVES 1 (BAA84474.1)
<400> 67

Met Gln Arg Met Asp Gly Gly Glu Glu Pro Arg Pro Ala Ala Gly Glu
1 5 10 15
Val Leu Gly Val Ala Asp Glu Ala Asp Gly Gly Val Val Phe Val Phe
20 25 30
Pro Gly Gln Gly Pro Gln Trp Pro Gly Met Gly Arg Glu Leu Leu Asp
35 40 45
Ala Ser Asp Val Phe Arg Glu Ser Val Arg Ala Cys Glu Ala Ala Phe
50 55 60
Ala Pro Tyr Val Asp Trp Ser Val Glu Gln Val Leu Arg Asp Ser Pro
65 70 75 80
Asp Ala Pro Gly Leu Asp Arg Val Asp Val Val Gln Pro Thr Leu Phe
85 90 95
Ala Val Met Ile Ser Leu Ala Ala Leu Trp Arg Ser Gln Gly Val Glu
100 105 110

Pro Cys Ala Val Leu Gly His Ser Leu Gly Glu Ile Ala Ala Ala His

115 120 125
Val Ser Gly Gly Leu Ser Leu Ala Asp Ala Ala Arg Val Val Thr Leu

130 135 140
Trp Ser Gln Ala Gln Thr Thr Leu Ala Gly Thr Gly Ala Leu Val Ser
145 150 155 160

Val Ala Ala Thr Pro Asp Glu Leu Leu Pro Arg Ile Ala Pro Trp Thr
165 170 175
Glu Asp Asn Pro Ala Arg Leu Ala Val Ala Ala Val Asn Gly Pro Arg
180 185 190

Ser Thr Val Val Ser Gly Ala Arg Glu Ala Val Ala Asp Leu Val Ala
195 200 205

Asp Leu Thr Ala Ala Gln Val Arg Thr Arg Met Ile Pro Val Asp Val
210 215 220

Pro Ala His Ser Pro Leu Met Tyr Ala Ile Glu Glu Arg Val Val Ser
225 230 235 240

Gly Leu Leu Pro Ile Thr Pro Arg Pro Ser Arg Ile Pro Phe His Ser
245 250 255

Ser Val Thr Gly Gly Arg Leu Asp Thr Arg Glu Leu Asp Ala Ala Tyr

260 265 270
Trp Tyr Arg Asn Met Ser Ser Thr Val Arg Phe Glu Pro Ala Ala Arg

275 280 285
Leu Leu Leu Gln Gln Gly Pro Lys Thr Phe Val Glu Met Ser Pro His

290 295 300

Pro Val Leu Thr Met Gly Leu Gln Glu Leu Ala Pro Asp Leu Gly Asp
305 310 315 320

Thr Thr Gly Thr Ala Asp Thr Val Ile Met Gly Thr Leu Arg Arg Gly
325 330 335

Gln Gly Thr Leu Asp His Phe Leu Thr Ser Leu Ala Gln Leu Arg Gly
340 345 350

His Gly Glu Thr Ser Ala Thr Thr Val Leu Ser Ala Arg Leu Thr Ala

355 360 365
 Leu Ser Pro Thr Gln Gln Gln Ser Leu Leu Leu Asp Leu Val Arg Ala
 370 375 380
 His Thr Met Ala Val Leu Asn Asp Asp Gly Asn Glu Arg Thr Ala Ser
 385 390 395 400
 Asp Ala Gly Pro Ser Ala Ser Phe Ala His Leu Gly Phe Asp Ser Val

 405 410 415
 Met Gly Val Glu Leu Arg Asn Arg Leu Ser Lys Ala Thr Gly Leu Arg
 420 425 430
 Leu Pro Val Thr Leu Ile Phe Asp His Thr Thr Pro Ala Ala Val Ala
 435 440 445
 Ala Arg Leu Arg Thr Ala Ala Leu Gly His Leu Asp Glu Asp Thr Ala
 450 455 460
 Pro Val Pro Asp Ser Pro Ser Gly His Gly Gly Thr Ala Ala Ala Asp
 465 470 475 480

 Asp Pro Ile Ala Ile Ile Gly Met Ala Cys Arg Phe Pro Gly Gly Val
 485 490 495
 Arg Ser Pro Lys Asp Leu Trp Glu Leu Ala Ala Ser Gly Gly Asp Ala
 500 505 510
 Ile Gly Pro Phe Pro Thr Asp Arg Gly Trp Pro Thr Glu Gln Arg His
 515 520 525
 Ala Gln Asp Pro Thr Gln Pro Gly Thr Phe Tyr Pro Gln Gly Gly Gly
 530 535 540
 Phe Leu His Asp Ala Ala His Phe Asp Ala Gly Phe Phe Gly Ile Ser

 545 550 555 560
 Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu
 565 570 575
 Thr Ser Trp Glu Ala Phe Glu Arg Ala Gly Ile Asp Pro Leu Ser Val
 580 585 590
 Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Ala Leu Ser Phe Asp Tyr
 595 600 605
 Gly Pro Arg Met Asp Thr Ala Ser Ser Glu Gly Ala Ala Asp Val Glu

610					615					620									
Gly	His	Ile	Leu	Thr	Gly	Thr	Thr	Gly	Ser	Val	Leu	Ser	Gly	Arg	Ile				
625					630					635					640				
Ala	Tyr	Ser	Phe	Gly	Leu	Glu	Gly	Pro	Ala	Ile	Thr	Val	Asp	Thr	Gly				
645					650					655									
Cys	Ser	Ala	Ser	Leu	Val	Thr	Leu	His	Leu	Ala	Cys	Gln	Ser	Leu	Arg				
660					665					670									
Ser	Gly	Glu	Cys	Thr	Leu	Ala	Leu	Ala	Gly	Gly	Val	Ser	Val	Met	Ser				
675					680					685									
Thr	Leu	Gly	Met	Phe	Ile	Glu	Phe	Ser	Arg	Gln	Arg	Gly	Leu	Ser	Val				
690					695					700									
Asp	Gly	Arg	Cys	Lys	Ala	Tyr	Ser	Ala	Ala	Ala	Asp	Gly	Thr	Gly	Trp				
705					710					715					720				
Gly	Glu	Gly	Val	Gly	Met	Leu	Leu	Val	Glu	Arg	Leu	Ser	Asp	Ala	Val				
725					730					735									
Arg	Leu	Gly	His	Arg	Val	Leu	Ala	Val	Val	Arg	Gly	Ser	Ala	Val	Asn				
740					745					750									
Gln	Asp	Gly	Ala	Ser	Asn	Gly	Leu	Thr	Ala	Pro	Asn	Gly	Pro	Ala	Gln				
755					760					765									
Glu	Arg	Val	Ile	Arg	Gln	Ala	Leu	Ala	Asn	Ala	Gly	Leu	Ser	Val	Ala				
770					775					780									
Asp	Val	Asp	Val	Val	Glu	Gly	His	Gly	Thr	Gly	Thr	Thr	Leu	Gly	Asp				
785					790					795					800				
Pro	Ile	Glu	Ala	Gln	Ala	Leu	Leu	Ala	Thr	Tyr	Gly	Gln	Arg	Ala	Gly				
805					810					815									
Asp	Arg	Pro	Leu	Trp	Leu	Gly	Ser	Leu	Lys	Ser	Asn	Ile	Gly	His	Thr				
820					825					830									
Met	Ala	Ala	Ala	Gly	Val	Gly	Gly	Val	Ile	Lys	Met	Val	Met	Ala	Leu				
835					840					845									
Arg	Glu	Gly	Val	Leu	Pro	Arg	Thr	Leu	His	Val	Asp	Lys	Pro	Ser	Pro				
850					855					860									

Gln Val Asp Trp Ser Ala Gly Ala Val Arg Leu Leu Thr Glu Ala Val
865 870 875 880
Pro Trp Pro Gly Asp Ala Ala Gly Arg Leu Arg Arg Ala Gly Val Ser
885 890 895
Ser Phe Gly Ile Gly Gly Thr Asn Ala His Val Ile Leu Glu Glu Ala
900 905 910

Pro Ala Ala Gly Gly Cys Val Ala Gly Gly Gly Val Leu Glu Gly Ala
915 920 925
Pro Gly Leu Ala Ile Ser Val Ala Glu Ser Val Ala Ala Pro Val Ala
930 935 940
Val Ser Ala Pro Val Ala Glu Ser Val Pro Val Pro Val Pro Val Pro
945 950 955 960
Val Pro Val Pro Val Ser Ala Arg Ser Glu Ala Gly Leu Arg Ala Gln
965 970 975
Ala Glu Ala Leu Arg Gln Tyr Val Ala Val Arg Pro Asp Val Ser Leu
980 985 990
Ala Asp Val Gly Ala Gly Leu Ala Cys Gly Arg Ala Val Leu Glu His
995 1000 1005
Arg Ala Val Val Leu Ala Ala Asp Arg Glu Glu Leu Val Gln Gly Leu
1010 1015 1020
Gly Ala Leu Ala Ala Gly Glu Pro Asp Arg Arg Val Thr Thr Gly His
1025 1030 1035 1040
Ala Pro Gly Gly Asp Arg Gly Gly Val Val Phe Val Phe Pro Gly Gln
1045 1050 1055

Gly Gly Gln Trp Ala Gly Met Gly Val Arg Leu Leu Ala Ser Ser Pro
1060 1065 1070
Val Phe Ala Arg Arg Met Gln Ala Cys Glu Glu Ala Leu Ala Pro Trp
1075 1080 1085
Val Asp Trp Ser Val Val Asp Ile Leu Arg Arg Asp Ala Gly Asp Ala
1090 1095 1100
Val Trp Glu Arg Ala Asp Val Val Gln Pro Val Leu Phe Ser Val Met
1105 1110 1115 1120

Val Ser Leu Ala Ala Leu Trp Arg Ser Tyr Gly Ile Glu Pro Asp Ala

1125 1130 1135

Val Leu Gly His Ser Gln Gly Glu Ile Ala Ala Ala His Val Cys Gly

1140 1145 1150

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

1155 1160 1165

Ala Leu Ala Ala Val Arg Gly Arg Gly Gly Met Ala Ser Val Pro Leu

1170 1175 1180

Pro Ala Gln Glu Val Glu Gln Leu Ile Gly Glu Arg Trp Ala Gly Arg

1185 1190 1195 1200

Leu Trp Val Ala Ala Val Asn Gly Pro Arg Ser Thr Ala Val Ser Gly

1205 1210 1215

Asp Ala Glu Ala Val Asp Glu Val Leu Ala Tyr Cys Ala Gly Thr Gly

1220 1225 1230

Val Arg Ala Arg Arg Ile Pro Val Asp Tyr Ala Ser His Cys Pro His

1235 1240 1245

Val Gln Pro Leu Arg Glu Glu Leu Leu Glu Leu Leu Gly Asp Ile Ser

1250 1255 1260

Pro Gln Pro Ser Gly Val Pro Phe Phe Ser Thr Val Glu Gly Thr Trp

1265 1270 1275 1280

Leu Asp Thr Thr Thr Leu Asp Ala Ala Tyr Trp Tyr Arg Asn Leu His

1285 1290 1295

Gln Pro Val Arg Phe Ser Asp Ala Val Gln Ala Leu Ala Asp Asp Gly

1300 1305 1310

His Arg Val Phe Val Glu Val Ser Pro His Pro Thr Leu Val Pro Ala

1315 1320 1325

Ile Glu Asp Thr Thr Glu Asp Thr Ala Glu Asp Val Thr Ala Ile Gly

1330 1335 1340

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

1345 1350 1355 1360

Ala His Thr His Thr Thr Gly Ile Gly Thr Pro Thr Thr Trp His His

1365	1370	1375	
His Tyr Thr His His His Thr His Pro His Pro His Thr His Leu Asp			
1380	1385	1390	
Leu Pro Thr Tyr Pro Phe Gln His Gln His Tyr Trp Leu Glu Ser Ser			
1395	1400	1405	
Gln Pro Gly Ala Gly Ser Gly Ser Gly Ala Gly Ala Gly Ser Gly Ala			
1410	1415	1420	
Gly Ser Gly Arg Ala Gly Thr Ala Gly Gly Thr Ala Glu Val Glu Ser			
1425	1430	1435	1440
Arg Phe Trp Asp Ala Val Ala Arg Gln Asp Leu Glu Thr Val Ala Thr			
1445	1450	1455	
Thr Leu Ala Val Pro Pro Ser Ala Gly Leu Asp Thr Val Val Pro Ala			
1460	1465	1470	
Leu Ser Ala Trp His Arg His Gln His Asp Gln Ala Arg Ile Asn Thr			
1475	1480	1485	
Trp Thr Tyr Gln Glu Thr Trp Lys Pro Leu Thr Leu Pro Thr Thr His			
1490	1495	1500	
Gln Pro His Gln Thr Trp Leu Ile Ala Ile Pro Glu Thr Gln Thr His			
1505	1510	1515	1520
His Pro His Ile Thr Asn Ile Leu Thr Asn Leu His His His Gly Ile			
1525	1530	1535	
Thr Pro Ile Pro Leu Thr Leu Asn His Thr His Thr Asn Pro Gln His			
1540	1545	1550	
Leu His His Thr Leu His His Thr Arg Gln Gln Ala Gln Asn His Thr			
1555	1560	1565	
Thr Gly Ala Ile Thr Gly Leu Leu Ser Leu Leu Ala Leu Asp Glu Thr			
1570	1575	1580	
Pro His Pro His His Pro His Thr Pro Thr Gly Thr Leu Leu Asn Leu			
1585	1590	1595	1600
Thr Leu Thr Gln Thr His Thr Gln Thr His Pro Pro Thr Pro Leu Trp			
1605	1610	1615	
Tyr Ala Thr Thr Asn Ala Thr Thr Thr His Pro Asn Asp Pro Leu Thr			

1620	1625	1630	
His Pro Thr Gln Ala Gln Thr Trp Gly Leu Ala Arg Thr Thr Leu Leu			
1635	1640	1645	
Glu His Pro Thr His Thr Ala Gly Ile Ile Asp Leu Pro Thr Thr Pro			
1650	1655	1660	
Thr Pro His Thr Leu Gln His Leu Thr Gln Thr Leu Thr Gln Pro His			
1665	1670	1675	1680
His Gln Thr Gln Leu Ala Ile Arg Thr Thr Gly Thr His Thr Arg Arg			
1685	1690	1695	
Leu Thr Pro Thr Thr Leu Thr Pro Thr His Gln Pro Pro Thr Pro Thr			
1700	1705	1710	
Pro His Gly Thr Thr Leu Ile Thr Gly Gly Thr Gly Ala Leu Ala Thr			
1715	1720	1725	
His Leu Thr His His Leu Thr Thr His Gln Pro Thr Gln His Leu Leu			
1730	1735	1740	
Leu Thr Ser Arg Thr Gly Pro His Thr Pro His Ala Gln His Leu Thr			
1745	1750	1755	1760
Thr Gln Leu Gln Gln Lys Gly Ile His Leu Thr Ile Thr Thr Cys Asp			
1765	1770	1775	
Thr Ser Asn Pro Asp Gln Leu Gln Gln Leu Leu Asn Thr Ile Pro Pro			
1780	1785	1790	
Gln His Pro Leu Thr Thr Val Ile His Thr Ala Gly Ile Leu Asp Asp			
1795	1800	1805	
Ala Thr Leu Thr Asn Leu Thr Pro Thr Gln Leu Asn Asn Val Leu Arg			
1810	1815	1820	
Ala Lys Ala His Ser Ala His Leu Leu His Gln Leu Thr Gln His Thr			
1825	1830	1835	1840
Pro Leu Thr Ala Phe Val Leu Tyr Ser Ser Ala Ala Ala Thr Phe Gly			
1845	1850	1855	
Ala Pro Gly Gln Ala Asn Tyr Ala Ala Ala Asn Ala Tyr Leu Asp Ala			
1860	1865	1870	

Leu Ala His His Arg His Thr His His Leu Pro Ala Thr Ser Ile Ala
 1875 1880 1885
 Trp Gly Thr Trp Gln Gly Asn Gly Leu Ala Asp Ser Asp Lys Ala Arg
 1890 1895 1900
 Ala Tyr Leu Asp Arg Arg Gly Phe Arg Pro Met Ser Pro Glu Leu Ala
 1905 1910 1915 1920

 Thr Ala Ala Val Thr Gln Ala Ile Ala Asp Thr Glu Arg Pro Tyr Val
 1925 1930 1935
 Val Ile Ala Asp Ile Asp Trp Ser Lys Ile Glu His Thr Ser Gln Thr
 1940 1945 1950
 Ser Asp Leu Val Ser Ala Ala Arg Glu Arg Glu Pro Ala Val Gln Arg
 1955 1960 1965
 Pro Thr Pro Pro Ala Glu Leu His Lys Thr Leu Ala His Gln Thr Ser
 1970 1975 1980
 Ala Asp Gln Arg Ala Ala Leu Leu Glu Leu Val Arg Asp His Val Ala

 1985 1990 1995 2000
 Ala Val Leu Arg His Ala Asp Pro Lys Ala Ile Ala Pro Asp Gln Ser
 2005 2010 2015
 Phe Arg Ala Leu Gly Phe Asp Ser Leu Thr Ala Val Glu Phe Arg Asn
 2020 2025 2030
 Leu Leu Ile Lys Ala Thr Gly Leu Arg Leu Pro Val Ser Leu Val Phe
 2035 2040 2045
 Asp His Pro Thr Pro Ala Lys Leu Ala Val His Leu Gln Asn Gln Leu
 2050 2055 2060

 Arg Gly Thr Ala Ala Glu Ser Ala Pro Ser Ala Ala Ala Val Thr Ala
 2065 2070 2075 2080
 Glu Ala Ser Val Thr Glu Pro Ile Ala Ile Val Gly Met Ala Cys Arg
 2085 2090 2095
 Phe Pro Gly Gly Val Thr Ser Ala Asp Asp Phe Trp Asp Leu Ile Ser
 2100 2105 2110
 Ser Glu Gln Asp Ala Ile Gly Gly Phe Pro Thr Asp Arg Gly Trp Asp
 2115 2120 2125

Leu Asp Thr Leu Tyr Asp Pro Asp Pro Asp His Pro Gly Thr Cys Tyr

2130 2135 2140
 Thr Arg Asn Gly Gly Phe Leu Tyr Asp Ala Gly His Phe Asp Ala Glu
 2145 2150 2155 2160
 Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln
 2165 2170 2175
 Arg Leu Leu Leu Glu Thr Ala Trp Glu Thr Ile Glu His Ala Gly Ile
 2180 2185 2190
 Asn Pro His Thr Leu His Gly Thr Pro Thr Gly Val Phe Thr Gly Thr
 2195 2200 2205

Asn Gly Gln Asp Tyr Ala Leu Arg Val His Asn Ala Gly Gln Ser Thr
 2210 2215 2220
 Asp Gly Phe Ala Leu Thr Gly Thr Ala Gly Ser Val Ile Ser Gly Arg
 2225 2230 2235 2240
 Ile Ser Tyr Thr Phe Gly Phe Glu Gly Pro Ala Val Ser Val Asp Thr
 2245 2250 2255
 Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Cys Gln Ala Leu
 2260 2265 2270
 Arg Ala Gly Glu Cys Ser Met Ala Leu Ala Gly Gly Val Thr Val Met

2275 2280 2285
 Ser Ser Pro Gly Ala Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala
 2290 2295 2300
 Ala Asp Gly His Cys Lys Ala Phe Ser Ala Ala Asp Gly Thr Gly
 2305 2310 2315 2320
 Trp Gly Glu Gly Val Gly Met Leu Leu Val Glu Arg Leu Ser Asp Ala
 2325 2330 2335
 His Arg Asn Gly His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val
 2340 2345 2350

Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser
 2355 2360 2365
 Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Asn Ala Gly Leu Ser Ala

2370 2375 2380
 Gly Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr Thr Leu Gly
 2385 2390 2395 2400
 Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Asp Arg
 2405 2410 2415
 Ala Gly Glu Gly Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Val Gly

 2420 2425 2430
 His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met
 2435 2440 2445
 Ala Leu Arg His Gly Leu Leu Pro Arg Thr Leu His Val Asp Glu Pro
 2450 2455 2460
 Ser Pro His Val Asp Trp Ser Ala Gly Ala Val Gln Leu Leu Thr Glu
 2465 2470 2475 2480
 Thr Val Pro Trp Pro Gly Gly Glu Gly Arg Leu Arg Arg Ala Gly Val
 2485 2490 2495

 Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Val Ile Leu Glu Glu
 2500 2505 2510
 Ala Pro Ala Asp Asp Val Pro Gly Gly Pro Pro Ala Gly Glu Gly Asp
 2515 2520 2525
 Ala Gly Ser Asp Asp Glu Ala Ala Ala Gly Ser Pro Gly Val Trp Pro
 2530 2535 2540
 Trp Leu Val Ser Ala Lys Ser Gln Pro Ala Leu Arg Ala Gln Ala Gln
 2545 2550 2555 2560
 Ala Leu His Ala His Leu Thr Asp His Pro Gly Leu Asp Leu Ala Asp

 2565 2570 2575
 Val Gly Tyr Thr Leu Ala His Ala Arg Ala Val Phe Asp His Arg Ala
 2580 2585 2590
 Thr Leu Ile Ala Ala Asp Arg Asp Thr Phe Leu Gln Ala Leu Gln Ala
 2595 2600 2605
 Leu Ala Ala Gly Glu Pro His Pro Ala Val Ile His Ser Ser Ala Pro
 2610 2615 2620
 Gly Gly Thr Gly Thr Gly Glu Ala Ala Gly Lys Thr Ala Phe Ile Cys

2625 2630 2635 2640
 Ser Gly Gln Gly Thr Gln Arg Pro Gly Met Ala His Gly Leu Tyr His
 2645 2650 2655
 Thr His Pro Val Phe Ala Ala Ala Leu Asn Asp Ile Cys Thr His Leu
 2660 2665 2670
 Asp Pro His Leu Asp His Pro Leu Leu Pro Leu Leu Thr Gln Asn Asp
 2675 2680 2685
 Asn Asp Asn Glu Asp Ala Ala Ala Leu Leu Gln Gln Thr Arg Tyr Ala
 2690 2695 2700
 Gln Pro Ala Leu Phe Ala Phe Gln Val Ala Leu His Arg Leu Leu Thr

 2705 2710 2715 2720
 Asp Gly Tyr His Ile Thr Pro His Tyr Tyr Ala Gly His Ser Leu Gly
 2725 2730 2735
 Glu Ile Thr Ala Ala His Leu Ala Gly Ile Leu Thr Leu Thr Asp Ala
 2740 2745 2750
 Thr Thr Leu Ile Thr Gln Arg Ala Thr Leu Met Gln Thr Met Pro Pro
 2755 2760 2765
 Gly Thr Met Thr Thr Leu His Thr Thr Pro His His Ile Thr His His
 2770 2775 2780

 Leu Thr Ala His Glu Asn Asp Leu Ala Ile Ala Ala Ile Asn Thr Pro
 2785 2790 2795 2800
 Thr Ser Leu Val Ile Ser Gly Thr Pro His Thr Val Gln His Ile Thr
 2805 2810 2815
 Thr Leu Cys Gln Gln Gln Gly Ile Lys Thr Lys Thr Leu Pro Thr Asn
 2820 2825 2830
 His Ala Phe His Ser Pro His Thr Asn Pro Ile Leu Asn Gln Leu His
 2835 2840 2845
 Gln His Thr Gln Thr Leu Thr Tyr His Pro Pro His Thr Pro Leu Ile

 2850 2855 2860
 Thr Ala Asn Thr Pro Pro Asp Gln Leu Leu Thr Pro His Tyr Trp Thr
 2865 2870 2875 2880

Gln Gln Ala Arg Asn Thr Val Asp Tyr Ala Thr Thr Thr Gln Thr Leu
2885 2890 2895

His Gln His Gly Val Thr Thr Tyr Ile Glu Leu Gly Pro Asp Asn Thr
2900 2905 2910

Leu Thr Thr Leu Thr His His Asn Leu Pro Asn Pro Pro Thr Thr Thr
2915 2920 2925

Leu Thr Leu Thr His Pro His His His Pro Gln Thr His Leu Leu Thr
2930 2935 2940

Asn Leu Ala Lys Thr Thr Thr Thr Trp His Pro His His Tyr Thr His
2945 2950 2955 2960

His Asp Asn Gln Pro His Thr His Thr His Leu Asp Leu Pro Thr Tyr
2965 2970 2975

Pro Phe Gln His His His Tyr Trp Leu Glu Ser Thr Gln Pro Gly Ala
2980 2985 2990

Gly Asn Val Ser Ala Ala Gly Leu Asp Pro Thr Glu His Pro Leu Leu
2995 3000 3005

Gly Ala Thr Leu Glu Leu Ala Thr Asp Gly Gly Ala Leu Leu Ala Gly
3010 3015 3020

Arg Leu Ser Leu Arg Ser His Pro Trp Leu Ala Asp His Ala Val Gly
3025 3030 3035 3040

Gly Thr Val Leu Leu Ser Gly Ala Thr Phe Leu Glu Leu Ala Leu His
3045 3050 3055

Ala Gly Thr Tyr Val Gly Cys Asp Arg Val Asp Glu Leu Thr Leu His
3060 3065 3070

Ala Pro Leu Val Val Pro Val Asp Gly Gly Val Ser Val Gln Val Gly
3075 3080 3085

Val Ala Ala Ala Asp Gly Glu Gly Arg Arg Leu Val Ser Val Tyr Ala
3090 3095 3100

Arg Gly Gly Ser Ala Cys Gly Gly Gly Gly Ala Ser Gly Gly Val Trp
3105 3110 3115 3120

Thr Cys His Ala Ser Gly Val Leu Val Glu Ala Ala Ala Gly Gly Val
3125 3130 3135

Val Val Asp Gly Leu Ala Gly Val Trp Pro Pro Arg Gly Ala Val Ala

3140 3145 3150

Val Asp Val Asp Gly Val Arg Asp Arg Leu Ala Gly Ala Gly Cys Val

3155 3160 3165

Leu Gly Pro Val Phe Ser Gly Leu Arg Ala Val Trp Arg Asp Gly Gly

3170 3175 3180

Asp Leu Leu Ala Glu Val Cys Leu Pro Glu Glu Ala Trp Gly Asp Ala

3185 3190 3195 3200

Ala Gly Phe Gly Leu His Pro Ala Leu Leu Asp Gly Val Val Gln Pro

3205 3210 3215

Leu Ser Val Leu Leu Pro Gly Gly Thr Gly Phe Gly Glu Gly Ala Gly

3220 3225 3230

Phe Gly Glu Gly Val Arg Val Pro Ala Val Trp Gly Gly Val Ser Leu

3235 3240 3245

His Arg Ala Gly Val Thr Gly Val Arg Val Arg Val Ser Ala Val Gly

3250 3255 3260

Arg Gly Gly Gly Arg Glu Ala Val Ser Val Val Val Gly Asp Glu Ala

3265 3270 3275 3280

Gly Val Pro Val Ala Ser Val Asp Arg Leu Glu Leu Arg Pro Val Asp

3285 3290 3295

Met Gly Gln Leu Arg Ala Val Ser Val Ser Ala Gly Arg Arg Gly Ser

3300 3305 3310

Leu Tyr Ala Val Gln Trp Ala Glu Val Gly Pro Val Pro Val Cys Gly

3315 3320 3325

Gln Ala Trp Ala Trp His Glu Asp Val Gly Glu Ser Gly Gly Gly Pro

3330 3335 3340

Val Pro Gly Val Val Val Leu Arg Cys Pro Asp Ala Gly Ala Gly Gly

3345 3350 3355 3360

Gly Gly Gly Gly Gly Gly Gly Gly Gly Val Gly Glu Val Val Gly Gly

3365 3370 3375

Val Leu Gly Val Val Gln Gly Trp Leu Gly Leu Glu Arg Phe Ala Gly

3380 3385 3390
 Ser Arg Leu Val Val Val Thr Arg Gly Ala Val Val Ala Gly Pro Glu
 3395 3400 3405
 Asp Gly Pro Val Asp Val Val Gly Ala Ser Val Trp Gly Leu Val Arg
 3410 3415 3420
 Ser Ala Gln Ala Glu His Pro Asp Arg Phe Val Leu Leu Asp Leu Asp

 3425 3430 3435 3440
 Thr Asp Thr Gly Thr Asp Leu Asp Thr Gly Ala Gly Ala Gly Trp Gly
 3445 3450 3455
 Val Asp Gly Gly Arg Val Ala Ala Val Val Ala Cys Gly Glu Pro Gln
 3460 3465 3470
 Leu Ala Val Arg Gly Glu Arg Leu Leu Ala Ala Arg Leu Lys Arg Leu
 3475 3480 3485
 Glu Ser Ser Gly Asp Val Pro Ala Gln Arg Ser Gly Asp Thr Arg Ala
 3490 3495 3500

 Arg Arg Ser Asp Val Pro Ala Gln Arg Ser Gly Gly Val Pro Ala Arg
 3505 3510 3515 3520
 Arg Ser Val Asp Val Ser Gly Arg Glu Val Leu Pro Trp Leu Ser Gly
 3525 3530 3535
 Gly Ser Val Leu Val Thr Gly Gly Thr Gly Val Leu Gly Ala Ala Val
 3540 3545 3550
 Ala Arg His Leu Ala Gly Val Cys Gly Val Arg Asp Leu Leu Leu Val
 3555 3560 3565
 Ser Arg Arg Gly Pro Asp Ala Pro Gly Ala Glu Gly Leu Arg Ala Glu

 3570 3575 3580
 Leu Ala Ala Leu Gly Ala Glu Val Arg Ile Val Ala Cys Asp Val Gly
 3585 3590 3595 3600
 Glu Arg Arg Glu Val Val Arg Leu Leu Glu Gly Val Pro Ala Gly Cys
 3605 3610 3615
 Pro Leu Thr Gly Val Val His Ala Ala Gly Val Leu Asp Asp Ala Thr
 3620 3625 3630
 Ile Ala Ser Leu Thr Pro Glu Arg Leu Gly Thr Val Phe Ala Ala Lys

3635 3640 3645
 Val Asp Ala Ala Leu Leu Leu Asp Glu Leu Thr Arg Gly Met Glu Leu
 3650 3655 3660
 Ser Ala Phe Val Leu Phe Ser Ser Ala Ala Gly Ile Leu Gly Ser Ala
 3665 3670 3675 3680
 Gly Gln Gly Asn Tyr Ala Ala Ala Asn Ala Ala Leu Asp Ala Leu Ala
 3685 3690 3695
 Tyr Arg Arg Arg Ala Ala Gly Leu Pro Gly Val Ser Leu Ala Trp Gly
 3700 3705 3710
 Leu Trp Glu Glu Ala Ser Gly Met Thr Gly His Leu Ala Gly Thr Asp

 3715 3720 3725
 His Arg Arg Ile Ile Arg Ser Gly Leu His Pro Met Ser Thr Pro Asp
 3730 3735 3740
 Ala Leu Ala Leu Phe Asp Ala Ala Leu Ala Leu Asp Arg Pro Val Leu
 3745 3750 3755 3760
 Leu Pro Ala Asp Leu Arg Pro Ala Pro Pro Leu Pro Pro Leu Leu Gln
 3765 3770 3775
 Asp Leu Leu Pro Ala Thr Arg Arg Arg Thr Thr Arg Thr Thr Thr Thr
 3780 3785 3790

 Gly Gly Ala Asp Asn Gly Ala Gln Leu His Ala Arg Leu Ala Gly Gln
 3795 3800 3805
 Thr His Glu Gln Gln His Thr Thr Leu Leu Ala Leu Val Arg Ser His
 3810 3815 3820
 Ile Ala Thr Val Leu Gly His Thr Thr Pro Asp Thr Ile Pro Pro Asp
 3825 3830 3835 3840
 Arg Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Thr Ala Val Glu Leu
 3845 3850 3855
 Arg Asn Arg Leu Ser Arg Thr Thr Gly Leu Arg Leu Pro Thr Thr Leu

 3860 3865 3870
 Ala Phe Asp His Pro Asn Pro Thr Thr Leu Thr His His Leu His Thr
 3875 3880 3885

Gln Leu Gln Pro Gln Pro Asp Asn Ala Val Ala Pro Val Leu Ala Glu
 3890 3895 3900
 Leu Asp Lys Leu Glu Ser Ala Leu Ser Ala Leu Asp Lys Thr Asp Ser
 3905 3910 3915 3920
 Ala Ser Glu Arg Val Thr Leu Arg Leu Lys Ser Leu Met Leu Arg Trp
 3925 3930 3935

Asn Ala Pro Gln His Pro Thr Ala Glu Ser Ala Asp Asp Asp Glu Lys
 3940 3945 3950
 Phe Thr Ser Ala Thr Glu Ala Glu Ile Phe Lys Phe Ile Asp Asn Asp
 3955 3960 3965

Leu Gly Leu Ser

3970

<210> 68

<211> 13032

<212> DNA

<213> Artificial Sequence

<220><223> milA1 gene of *S. bingchenggensis* BCW-1 (Accession no. CP002047)

<400> 68

ttgccaaag cccagaacga gttcgcagtg gccggtcac cgtggatcct ctccgggcac 60

accggaaccg cgctgcgggc ccaggcacgc cggctccacg accatgtcgc cgaccaccct 120

cggctccgtc cggaagacat cgcccacacg ctggcgagca gcggcccggc gtcacccat 180

cgcgcggcgg tgatcgcggc ggaccgggaa ggacatctcc gggggctcga cgcggtggcc 240

cggggtgagg acacccccgg tgtcgtacgg ggcacggcgg ccgcgggcgg cgacggggtc 300

gcgttcgtct tccccggcca gggcacccag tggcccggta tggccgccga tctgtgacg 360

gtctcccccg cttcagccg ggcggtcgac gcctgcgccg aggcgttcga accgtatgtc 420

tcctggtcac cggaggccgt gctgcggggc gtcggggcgc cgccgccct ggaggggacc 480

gatgtggtgc agccgacgt gttcgcctc atggtggggc tggccgagct gtggcggact 540

cttggggatga gcccgacgtc gatcgtgggc cactgcatc gggagatcgc ggcagcccat 600

ctctcgggcg ccctgtcgt gtccgacgcg gcgcgcgtgg tgatcgagag cagccgggccc 660

caggcgacgc tctccgggtc ggggtgcgtg atcgcggtcg cgcggtccga ggcgcagctg 720

cttccgttgc tgcggcggtg gccgggcagg ctgacgatc ccgcggtcaa cggcccgatg 780

gccacggtcg tctccggcga tcggccggcc gccgacgagc tgttggcgga gttcgcccgt 840

gccgggtgtcc gggcccgcca ggtggcgatc gacatccccg cgcactcgcc gttcatggcc 900

ccctcaggg acggtctgct cgactcgctg tcatcggtca ccgcgggtgc gtcgcggctg 960

ccgttcact cctcggtcat cggggggccg ctggagaccc aagggtcga cgcggcttac 1020

tggtagcga acctcgcca caggtccgc ttgaaagcg tcgtcacggg gctgctgcgg 1080

cagggcacac gctgcttctg ggagctgagc ccgcaccga tgcgtaccat gtgtgtgcag 1140

gccaccgccc aggaggtggt cggcggtag cgcgtcgtga tctgcccac gctgcatcgc 1200

gggcaggccg ccgtcgagtc cgttcgcacc acgtggccg agctgtacgt acggggcgca 1260

ctggatgacc atcgggcggc gttctcggtg ccgggcggcc gcctgatcac cctgcctctc 1320

gagccgcccc cgacacgtc cgtagagctc gccgacccc cggaccggc ggaggcctgc 1380

cggccccct tggtagagcg gcttgcccgg ctctccaccg cggagcggaa gcggcggtg 1440

cgcgagctgg tggcgctga ggcggccaag gtctcgagg acgtcgccgg ggcggacgag 1500

ccgggccacg gcatcgcca gcaggagcac ttctcactt cgggcttga ctccgcggcc 1560

gcggtcgcg tgcgaaccg cctgaacgac gccaccggtt tgcgtgctgc ctccaccctg 1620

gccttcgacc atccgacac cgccgcccgc gccgaccatc tgcactcccg gctcttcgat 1680

caccaggcgc gcgggcagcc gggcgccgac ggccggcccg accccgcccg gcgggcgggt 1740

ccggccaggg ccgacgaca gccgatgcc gtcatcgga tggcgggccc ctccccggg 1800

ggcgcgccga ccccgagga gctgtgggaa ctggtcgccg aaggcaccga cgccctctcg 1860

cccttcccgg agggccgggg ctgggatccg ctgcggctct acgatccga ccccgcccgg 1920

cccggcacgt actaccagc cgaagcggga ttcttcacg acccgacaa gttcgacgcc 1980

gagtctctcg gcatcgccc acgcgaggcc accgcaatgg atccccagca gcggctgctc 2040

ctggagacct cctgggagcg gctcgaacgg gcgcggatcg accgaccgc gctgcgcggc 2100

agccgcaccg ggtgttctg cggcgtggcc ccgtggact acagccccg aatgcaccag 2160

gcgtcgccgg agctggaggg ccatctgctg accggcaaca tcggcgccgc ggcctcgggg 2220

cggatctct acgtactcg gcttgagggg ccgcgggtgt ccgtggacac ggcgtgctcg 2280

tcgtccctgg tcgcccga tctggcggcc caggcgctgc gggccgggga gtgctcgctg 2340

gccctggtcg gcggggcgac ggtcctctcg accccggca tgttcacga gttctcgcg 2400

cagcgcggtc tggtccgga cggccgctgc aaggcgtac cggccgccgc ggacggcacc 2460

ggctggtccg aggtgtggg catgtgtct gtcgagcggc tgtccgacgc gcgacggctc 2520

ggacaccagg tgcttcgggt ggtacggggc tccgccgtca accaggacgg ggcgagcaac 2580

ggcttcacgg cgcccagcgg tccatcacag caacaggtca tccgggcggc cctggccaat	2640
gccgggggtgt cggtccgga ggtcgacgcg gtggaggggc acggcaccgg caccgggttg	2700
ggcgatccga tcgaggcgca ggcgctgctg gccgcctacg ggccggggcg ggccggccgac	2760
cggccgctgt ggctggggtc gatcaagtgc aacatcgac acaccagtg ggccgagggt	2820
gtcatcgggg tcatcaaat ggtgctcgcg ctccagcagc gtgtgctgcc gcgcacgctg	2880
cacgtggaca agccgtcgga ttacgtggac tggtcggccg gggccgtacg gctgttgacg	2940
gagccgggtgc cctggccgga gcggggccac ccgcgccggg cgggggtgtc gtccttcggg	3000
gtgagcggca ccaacgcccc tgtcctctc gagcaggcaa cgccatcgtc cacggtggct	3060
cccggggggc ataccgcca ggccgggcct cccctgccgt ggggtggtctc ggccaagacg	3120
ccccaggcac tgcgcacca ggcccgccgc ctgcacgaac acctcaccgc ccagccacag	3180
ctccaaccgg ccgacgtcgg ccacacctc gccaccggcc gcgccacctt cgaccaccgg	3240
gccgtcctca tcggtccga ccgcaacaa ctctccacg gcctggacgc gctcgccacc	3300
ggccggcccg acccagcggc ccaccagacg tcggaccgtc ccgccaccgc cgacggccgt	3360
atcgtcttcg tcttccccgg acaaggcggc caatgggcgg gcatgggcct acggctgctg	3420
aacgcctcac ccgtcttcac cgagcggatg gccgcctgcg aacaggccct ctccccctac	3480
gtcgactggt cactcacgga catctccac cgcccgcccg acgacgccgt atggcaacgc	3540
gccgacatcg tcagcccgc cctgttctcg atcatggtgt ccttggccgc gctctggcgc	3600
tcttgccgca tcgaaccgga cgccgtctc ggccactccc aaggcgagat cgcccgggcc	3660
cacgtctgcg gcgccctgac gctccacgac gcggccaagg tcatcgccct gcgcagccag	3720
gccctccaag ccgtacgcgg cgccgggggc atggcctccg taccctgcc cgccgaccag	3780
gtcaccgagg atctgcgac ccactggccc gaccggctat gggaggccgc caccaactcc	3840
cccacggcaa ccgtcatctc gggaaacacc gacgcgttg acgaagcgt cgaccactac	3900
cacgcccacg acgtacgggc caagcgcac ccggtcgact acgcctcca ctgccccat	3960
atcgacgcgg tggccgagcg actgcccac ctgctgggcg gcatcgctcc gcgcgccgc	4020
gacatccct tctactccac ggttgacggc cgatgggcg agccgaccga gctcgacgc	4080
gactactggt accgcaacct ccgagcccc gtacgggtcg cccacgccgt ccacgcctc	4140
accgagaccg accaccgac ctttctgaa gtcagccac accccacgt cccccgcc	4200
atcacggcca ccaccgaaac caccgaccgc accaccacg tcctgcctc gctccaccgc	4260
gaccacgacg acaccacca catctcacc aacctgccc agggccacat ccacggccac	4320
accatcgact ggcgacacca ctaccagact ctgcgcccc cccacccca tatcgacctc	4380
cccactacc cttccaaca ccaccactac tggctccacg actccaccga ggacaaggcg	4440

gtgggtacgg acctcgccgc ggcccgttc tgggaggcgg tccacggcga ggacaccaac	4500
gccgtcgccg cgctcctcga cgtcgagccg ggcacctcac tggacgcgct gctgccggcc	4560
ctgtccgcct ggcacggctg cgctcgcgac caggccatca ccgacacctg gtgttaccgc	4620
gacatctgga agccggccga cctcaccgcc gcgcgcccc ggccgtccgg ccgatggctt	4680
gtcgcgatct ccgcaggcgg ggccgatac ctcacgtca gtgccgtcct ggacgtctg	4740
gaacgccagg gtctgccat cgccacctc gtctcgacg acaccacac cgaactcccc	4800
ctgtggagc ggcattcgc acaggcgatc gcgagcgatg ggccggccat cggcggcgtg	4860
ctctcgctgc tcgacctga cgaggggcca catccgcgc acccgaggt gccctcggc	4920
accgccctca cctcagcct gatccaggcg ctcacgcac gcgaggacat ggcccccgg	4980
ctgtggctgg ccaccacga ggccgtcgcc acctcgtccg cggatacgt cgatcacc	5040
ctccaggcga tggctgggg gctgggacgc accgccgac tcgaacacc cgatctgtg	5100
ggcggcctca tcgacctcc ggacacttc accgaacggg tcctccacgg ctcgtcacg	5160
gcgtgacca cctgtcacga cgaggacgaa ctgcgctgc gcgccaccgg cccacgcacc	5220
cggcgctga tcggagcgc gtccaccgcc gcagcggagg acaccgcc gtggacgcc	5280
cgtggcaccg tcctcatcac cggcggcacc ggggccctgg gctcccgt cgcccaccg	5340
atcgccgaac gccacccca ctgccacttg ctgctggtga gccggcgagg gcccaggcc	5400
cccggcgcca ccgcgtccg cgaccagtc atcgaactc gcgccacgt gacctcgcc	5460
gcattgaca ccgccgacc cggcgcgctc gcggatctc tcgccgatgt ccctcggac	5520
cgccccctca ccgggtcgt ccacaccgc ggcgctctgg acgacagcac ctcgccgta	5580
cagaccccg accacctgc cgccgttcg gggcccaagt cccatgccgc acaccatctg	5640
cacgcctcg cccagacca cccctcgac gcgttcgtc tcttcgtc cgtcgcgcg	5700
cccttcggcg ccggggcca ggccaactac gcggcgcca acgcctacct cgacgcctc	5760
gcccagcacc gccgggcca ggggtggcc gccacctca tcgctgggg caactgggac	5820
ggcgacgggc tcgcgagcac acagtccgc cagacgtacc tgcgaaccg cggtttct	5880
ccgatgccg cacacctggc gctggccgc ctggagcgag ccacgtctc gcccacgcc	5940
cagtcgtcg tcgccagct cgactggaag aagctcaagc cggcgccga caccgcgac	6000
atccgggaa gccgccgcc ggccccggcc gccaccgac gcgcggacag gacggccgac	6060
gccaccgca gcctccgtac ccgcctcgcg ggtcagagc cggccgaacg gcaccagac	6120
ctctcgacc tcacagctc tcatacagc gccgtcctc ggacgccac gcccagacg	6180
atccccacgg accgggctt ccgcgacctg ggtttacct cgtgacggc catcgagtc	6240
cgcaaccgc tcgcggcgcc caccgggctc gcctgccga ccacgtcgc cttcgacct	6300

ccgacgccgg acaagctcgc ggccgacctg ctggcgcggt gcgcgccgac aggcccgac	6360
ggcatcgggg tgacccccga cgcgacggcc acgagtggca gttcgcccgg tgcggcgcat	6420
ggcgcgccgg accccgcca gcccatcgcc atcgctgggt tggcctgcgc ctaccccggc	6480
gggatcggct cccccagga cctgtgggag ttcataccg cacaccggga cgccgtcgga	6540
gacttcacca ccgaccgggg ctgggacctg gcgaggctct tcgaccca tccggaccgg	6600
ccgggcacct cgtacagccg acagggcgcc ttcctccgc acgcggcgga cttcgaccg	6660
gagttcttcg ggatcagccc acgggaggcg acggcgacgg accccagca gcgactgctc	6720
ctggaggcgt cctgggaagc cctcgaacga gccgggatca accccacga tctccacggc	6780
agtcgcacgg gcgtcttcac cggcagcaac gcgcaggact tcagcgcgcg gctgcggcag	6840
acgccgtcgg agctggcgga gctgtgcgag ggctatgcgc tgactggcag caacaacagc	6900
gtgcctcgg ggcggtctc gtacgcgtc ggctggaag gcccgcggt cagcatcgac	6960
accgcctgct cgtcctcgt cgtggcgctc catctggcct gccagtcgt ccgggcccgc	7020
gaatgctcgc tggccctggc gggcggcgct acggtcatga tgacccgtt caacttcgtg	7080
gagttctccc ggacggggg cctggcgcg gcagggcggt gcaaggcgtt ctccgccaca	7140
gccgatggca ccggctgggg cgaggcggtg ggcatggtgg tggaggagcg gctgtcgac	7200
gcgcggcgca acggccatcg tgtgtggcc ctgctccgc gcagcgccgt caaccaggac	7260
ggtgccagca atgggtgac tgccccgaac ggccctcgc agcagcggt catccgcgc	7320
gccctggcgc ccgccgggt gcccgcgca gaagtggacg cggctgaggc gcacggcacg	7380
gggacgacgc tcggcgatcc gatcgaggcc caggccctgc tcgccacct cgggcagggg	7440
cggccggcgg accgggcgt gtggctcgt acggtcaagt ccaacatcg acaccccag	7500
tcggcccgcc gtatcgccg ggtcatcaag atggtgctgg ccctcgcgca cgggatgctg	7560
ccgcgtacgc tgcatgtgtc cgagccgtcg ccgatgtgg actggtcggc gggtgcggt	7620
cggctgctga ccgaggacca gccgtggcgc gacaccgggc gccccggcg ggcgggggtg	7680
tcgtccttcg gcgtgagcgg caccaatgcc catgtgatcc tggagcagc ggagccggg	7740
ccggaccggg caccgacggc ctccgcgcc tccctgccc cctggccct ctccgccagg	7800
tcggcgagg ccctgcgggc ccagcccggt aggttgctgg cgtacgtggc cgagcaccg	7860
gatgtcgacc ccgccagct ggggtactcc ctgcgcgcg gacgggccgt gttcgagcac	7920
cgggcgtgc tctcggcac cggccacgac gacttcggc gcgcctgga cgccctggcg	7980
tcgggcgcgc ccgacggcgc ggtcgtccag ggcgcgcg tggggcgga gggcaaggtc	8040

gtcttttgtgt gctcggggca gggcaccag cggcccgca tggcgccgg gctctaccgc	8100
tcgtccacgg cgttcggcg ggcgtggag gaggtgtgcg cccatctgga cccgtatctg	8160
gaacaccctc tgatggaggt gatgttcgcc gatgagaaga gcgatacgtc ggcgtgctg	8220
catctgaccg cctacgcca accggccctc ttccctcc agaccgctc gcaccgcatg	8280
gtaccgagg agttcgggt caccgccgac tatctggcg gccactcct gggcgagctg	8340
accgcgccc atctggcggg catctcagc ctgccgacg ccgcggcgct ggtggcgcc	8400
cgcgcccgcg ccatgcggga cttccagcg accggagcca tggtcgccgt cgaggccacc	8460
gaggcggagc tgcggccccg gctcggcgag ttggcggacc gggtcggcat cgcgcctc	8520
aacgccccg cgtccctggt catcaccggc gaccacgacg ccgtgcacca gatcgcgac	8580
gacttcgcg ggcaggcgag gaaggtcact ccctccagg tcagcggcg cttccactct	8640
ccccatatgg agccctgct gcacgagatc gggcgaccg ccgaaacct cacctaccac	8700
cggccccaca ctccctcgt caccgcgtcg gccgacggcg gcgacgacac gaccgagccg	8760
cgggcccagc acgaccggg caggcccg ttttgccctc tccaggccc gcgcaccgtc	8820
cactacgcgc gggcgtgga gggctgcgc gccgcggcg tcaccagtt cctggaactc	8880
ggccccgact ccacctcac taccctcgtc caccacaatc tcgcgcgca cgatccctg	8940
gccgtctccc tgcctcatc ggagcgggtc gagacgcaca gcgtcctcgg cgactcgc	9000
gcggtccacg cccacagccg cccgtcgac tggacagcc actacaccg acggccggg	9060
ccgacccac accagatcga cgtgccacc tatgccttc ggcaccggcg ctactggctg	9120
cccggccgg cggcggctcg cgatgtgac gccgcggggc tcgacgcgc ggagcaccg	9180
ctgatcggcg ccgcctgtg gctcggcg gagcgaggct gtctgctgac cggcaggatc	9240
tcgcccgta cgcaccgtg gctggccgac catgtcatc ccggcactgt gctccttcg	9300
ggcacgcgt tcgtggagct ggcgtgcgg gccggggcgt acgtgggctg cgaccgtgtg	9360
gaggagctga cctgcacgc gccactccg ctcccccg acggtgaggt ggtgctgacg	9420
gtggcggtag gggccgcca cgagtcggc cgcctgagc tgagcatcca cggccggcg	9480
gcggacgacg gtacatggac acggcacgc atcggcacgc tggcatcggc ccgcggcgtc	9540
ggcctcgacg atggcacggg gcacaatggc cagccccg cgggcgacga gccgttcggg	9600
tcgtgggcca cggcctggc gccgccggt gccgagcct tggacgtcac cggggtctac	9660
gaccggtttg ccgacgcga gttcacgtac ggggaggcat tccaggggct ggtcgcggct	9720
tggcggcacg gcgacgagc gctggcggag gtccgcctc ccgaccagc ggccggtgac	9780
gccctccgt tcgggtgca ccccgctg ctgcacggg cactgcagac catgtggctc	9840
gtggagccc acggcacag gccgagcgt ggctggcg gcccgatcg ggccctgccg	9900

ttcgcctggc aggggggtctc gctgctgacg gcggggcccg cgccctgcg ggtacggctg	9960
cgacggcccg cgccggacac cgtggccgtc gccgtggccg acgcgcccg cgcccggtc	10020
gcgtcggtgg agtcgtgac gctgcgccg gtgccgggg gcgccttgc cggcaccgag	10080
acggcggtgc gcacctcgtt gtacggcctg gactggacgg atgtgccgt gccgacgccg	10140
cagacggccc tgccccggtg tgcgtgatc ggagcggaca cgctcgacct ggtccccgcg	10200
ctcgaggccg cgccgccga ccgcatcacc gacggcgtgg agcgtacgc cgacctggag	10260
gagctggtgc gctccgtggc ggccggcgcc cccggcccg acctcgtcat cgccggtgc	10320
cacgcagccc ctgaagccga cggcgcgagc gaacagccac agcccagac ggtgcgcaca	10380
aggacgggtc aggtgctgga gctgctttag cgttggtcgc gcgcggacgg gctcgccgac	10440
gcacacctgg tgctgttac ctgaggcgcg gtgccaccc ggccggcgga gccggtgcgg	10500
gacctggcgg ggccggcggt ctggggtctg gtgcgtccg gccagtcgga gcatccggag	10560
tgcttcaccg tggtagcat ggacggcgcc caggagtcgc gcgcggcgct gctcgcgcg	10620
ctcgccctcg gcgagccga actggcggtg cgcggcgcc ggccgctggc gccgcgctg	10680
gtgcggccg gtgacccga cgacgacagc ggccggccc tgccgcaggg gccggaaggc	10740
tggcggttgg agtgtcccg cacggcgagc ctggacgggt tgaccacgac cgagtccccg	10800
gcccgggcgg tgccgctcg cccggcgag gtacgggtcg cgtgcgggc cgccgggctg	10860
aacttcgcg atgtgtgat gcgctgggc gtggtgccc ggccgacggc gctgggcagt	10920
gagggggcgg ggatcgtct cgaggtcggg gcggaggtcc gcgatctgc gcccggggac	10980
cgggtggtgg gtaattccc cgaggcgtt ggcccggtgg ccgtggccga gcgggcgacc	11040
ctggcgcggg tccccagcg ctggctgtt gccagggcg cgtcggtccc catcgtgtt	11100
gccaccgctt accacggcct ggtcgatctg gcgcgctgc ggccggggga atcggtgctg	11160
atccatgccg cgccggcggg ggtgggcatg gccgctgc aactggcgcg ccatctgggg	11220
gccgaggtgt acgccacggc cggccccgc aagtggcaca tcctgcgtt ccaaggcatc	11280
gacgacgacc atctggcctc gtgcgcacg ctggagttcg agcagcgtt cgccgcgacc	11340
cgccgggggc gggggatcga tgtcgtcctg gactgtctgg ccatgagtt cgtcgacgc	11400
tcgtcgcgc tggtagcgcg tgacggcggc cgttccttgg agatgggcaa gagcgacatc	11460
cgtgaccgc ggaggtggc gctggacat cggggcgtgc tctaccggc gttcgacctg	11520
ctggaggccg ggccggagcg ggtcggcgag atcctgcga ccgtactgga cctgttcgag	11580
cgcggtgtcc tggcgacct gccgacgacc tgctgggaca tccggcaggc ggagcacgcc	11640
ttccgccatc tgcagcagg ccgtcacatc ggaaagaacg tgctcaccgt cccggccggc	11700
tggaacgccg agggcacct actgatcacc ggcggtatgg gcacctggg cgccgcctc	11760

gccccgtcatc tggcgggtac cgggcgcgcc cgccatctgc tgctggccgg cgcacgcggc 11820

cccgacgccc cgggcgccga ggagctgcga gaggagctga ccgagctggg cgcgcgggtc 11880

accatcgccg catgcgatct cggcgaccgg gcggcggtcg cccggctcct gggggcgatc 11940

cggccgagc ggccgctgac cgctgtcatc cagcggcggt gtgtcgtcga cgatgccacc 12000

ctcgggtccc tcacccccg ccacctggac gccgccctgg ccgccaaggc cgacgccgcc 12060

tggcatctgc acacctcac ccgccagcc gacgtggcgg cgttcgtcct cttctcctcg 12120

gtcgggggtc tgctcggctc gcccgggcag ggcaactacg ccgcgccaa cgccttcttg 12180

gacgcgctcg cccaccaccg gcgcggctct ggcttccgg cgggtgctgct ggcgtggggg 12240

ctgtgggagc agaccagcgg catgaccggg cacctggacc aggcgcaccg cccccgctg 12300

gccccgctcg gcatcagccc gctcacgacc gggcaggcgc tcggcctttt cgacgccgcc 12360

ctcggccacc accgccccgt gctcgtcccc gccgcctcg acgtgccga tccgcacccc 12420

ggctcgtcga ccgtgccgcc cctgtaccgg ggcttggtcg gatccaggac ccggcggaca 12480

ccccccgcgg ccgccgccac cgggcggtt ccctgcata cccgcctcgg cggtcacgcc 12540

ccggccgagc agcacgagat gctgctctcg ctggtccgct cccacgccgc ctcgtgctg 12600

ggccgcgacg atccggacac ggtccatccc ggcgcgact tccgcggcct gggttcgac 12660

tcctgaccg cggtcgagct ccgcaaccgg ctcaacccg ccaccggcct ccggtctcc 12720

accacctcg ttctcgacca cccacgccc gacgaactcg cccgtcacgt ccgggagcag 12780

gtgctgggcg acggcgaagc ggcgcggtg gccccggtgc tggccgagct cgacaggctg 12840

gaagcggcgc tgtcccggt ggacggggac gatgcggtcc gggcgagggt gacggcccg 12900

ttgcaggccc ttctcctgaa gtggaacgag tccgatggtc cggcgacggg cggtgacggt 12960

gcgggcaggc tggcgtccgc cagggccgcc gaggtgctgg atttcacg gaacacctc 13020

ggcctctcct ga 13032

<210> 69

<211> 12993

<212> DNA

<213> Artificial Sequence

<220><223> meilingmycin biosynthetic gene cluster of Streptomyces
nanchangensis strain NS3226

<400> 69

gtggccggac atccgtgat cctctccgga cacaccgga ccgcgctgcg ggcccaggcg 60

cgccggctcc acgacatgt cgccgaccac ccctgctcc gtcggaaga catcgcgac 120

acgctggcga gcggcggccc ggcgctcacc catcgcgagg cggtgatcgc ggccggaccgg	180
gagggatatc tccgggggct cgacgcgggt gcccgagggt aggacgcccc cggtgtccta	240
cggggcacgg cgaccgcggt cggcgacggg gtgcggttcg ttttccccgg ccagggcacc	300
cagtggcccc gtatggccgc ggatctgtcg acggtctccc ctgccttcag ccgggcggtc	360
gacgcctgcg ccgaggcggt cgaaccgtat gtccccgtgt caccggaggc ggtgctgcgg	420
ggcgctccgg gcgcgccgcc cctggagggg accgatgtgg tgcagccgac gctgttcgcc	480
gtcatggtgg ggctggccga gctgtggcgg actcttgggg tgagcccac gacgatcgtg	540
gggcactgca tcggggagat cgcgcgggcc catctctgcg gcgccctgtc gctgtccgac	600
gcggcgcgcg tggatgatga gagcagccgg gcccgaggca cgctctccgg gtcgggtgcg	660
ctgatcgagg tcgcgcggtc cgaggcgag ctgcttccgc tactgcggcg gtggccgggc	720
aggctgacga tcgcccggtt caacggcccc atggccacgg tcgtctccgg cgatcgccgc	780
gccgccgacg agctgttggc ggagttggcc cgtgccggtg tccgggcccg cgaggtgccg	840
atcgacatcc ccgcgcactc ggcggtcatg gccccctca gggacggtct gctcgactcg	900
ctgtcatcgg tcaccgcggg tgcgtcgagg ctgccgttcc actcctcggt catcgggggg	960
ccgtggaga cccaagggtt cgacgcgggt tactggtacc ggaacctcgc cgacacggtc	1020
cgcttcgaaa gcgtggtcac ggggtgtgtg cggcagggca cgcgtgtctt cgtggagctg	1080
agcccgcatc cgatgtgac catgtgtgtg caggccaccg ccgaggaggt ggtcggcggt	1140
gagcgctgcg tgatcctgcc gacgttgcg cgcgggcaag ccgccgtcga gtccgttcgc	1200
accacgtgg ccgagctgta cgtacggggg gcgctggatg accctcgggc ggcggttctcg	1260
gtgccgggcg gccgactgat caccctgccc ctgcagccgc tcgcggacac gtccgtagag	1320
ctcgccgacg ccccgatcc tgcggaggcc tgcggcccc cttgggcgga gcggcttgcc	1380
cggctctcca ccgcggagcg gaagcggcgg ctgtgcgagc tggtagggct cgaggcggcc	1440
aaggtcctcg aggacgtcgc cggggcggac gcgccgcgc acggcatcgc tgagcaggag	1500
cacttcgtcg cttcgggctt cgactccgcg gccgcggtcg cgctgcgcaa ccgctgaac	1560
gacgccaccg gactgtgtgt gcccttcacc ctggccttcg accatccgac acccgccgc	1620
gtcgccgacc atctgcactc ccggtcttct gatcaccggg gcggtgggca gccgggcgcc	1680
gacggctggc ccgacccgcg ggccggcgcc ggtccggcca gggccgacga cgagccgac	1740
gccgtcatcg gcatggcggg ccgcttcccc gggggcgctc gtaccccgga ggagctgtgg	1800
gatctggtcg ccgaaggcac cgacgccctc tcccccttc cggagggccg gggctgggat	1860

ccgctgcggc tctacgatcc ggaccccgcc cggcccggca cgtactacca gcgcgaagcg	1920
ggattcctcc acgacgccga caagtctgac gccgagttct tcggcatcgc gccacgcgag	1980
gccacggcca tggatcccca gcagcggctg ctcttgagga cctcctggga ggcgctcgaa	2040
cgggcgcgga tcgaccgac cgcgctgccc ggagccgca cgggggtgtt cgtcggcgtg	2100
gccccgtgg actacagccc ccgtatgcac caggcgtcgc cggagctgga gggccatctg	2160
ctgaccggca acatcggcgc cgcggcctcg gggcggatct cctacgtact cgggctcgag	2220
gggcccgcgg tgtccgtgga cagggcgtgc tcgtcgtccc tggtagccct gcatctggcg	2280
gctcaggcgc tgcgggcccgg ggagtgtctg ctggccctgg tcggcggggc gacggtcctc	2340
tcgacccccg gcatgttcat cgagtctctg cggcagcgcg gtctggctcc ggacggccgc	2400
tgaaggcgt acgcccgcgc cgcggacggc accggctggg ccgagggcgt gggcatgctg	2460
ctcgtcgagc ggctgtccga cgcacgacgg ctccgacacc aggtgctggc ggtggtagcg	2520
ggctccgccg tcaaccagga cggggcgagc aacggcttca cggcggccag cggccatca	2580
cagcaacagg tcatccgggc ggccctggcc aatgcggggg tgcggctcc ggaggtcgac	2640
gcggtggagg ggcacggcac cggcacccgg ttgggcgac cgtcagaggc gcaggcgtg	2700
ctggcggcgt acgggcaggg cggggcggcc gaccggccgc tgtggctggg ctcgatcaag	2760
tcgaacatcg gacacacca gtggcccgcg ggcgtcatcg gggcatcaa aatggtgctc	2820
gcgtccagc gcggtgtgtt gccgcgacg ctgcacgtgg acaagccgtc ggattacgtg	2880
gactggtcgg cgggggcccgt acggctgttg acggagccgg tgccctggcc ggagaggggc	2940
caccgcgcc gggcgggggt gtctctcttc ggcgtgagcg gcaccaacgc ccatgtcatc	3000
ctcgagcagg caacgccatc gtccacggtg gctcccagg ggcctaccgc cgaggccggg	3060
cctccctgc cgtgggtgat ctccggaag acccccagg cactgcgcga ccaggcccgc	3120
cgctgcacg aacacctcac cgcacgcca cagctccaac cggccgacgt cggccacacc	3180
ctcgccaccg gccgcgccac cttegaccac cgggcgtcc tcatcggtc cgaccgcgaa	3240
caactctcc acggcctgga cgcgtcgcc accggccggc ccgaccagc ggtccaccag	3300
acagcggacc gtccgcacc cgcgacggc cgtatcgtct tcgtcttccc cggacaaggc	3360
ggtcaatggg cgggcatggg tctacggtg ctgaacgct caccgtctt caccgagcgg	3420
atggccgct gcgaacagc cctctcccc tacgtcgaact ggtcactcac ggacatctc	3480
caccggccgg ccgacgacgc cgcatggcaa cgcgccgaca tcgtccagcc cggcctgttc	3540
tcgatcatgg tgtccctggc cgcgtctggt cgtcttgcg gcatcgaacc ggacgcggtc	3600

ctcggccact cccaaggcga gatcgccgcg gcccacgtct gcggcgact gacgtccac	3660
gacgcggcca aggtcatcgc cctgcgcagc caggccctcc aagccgtacg cggcgccggg	3720
ggcatggcct ccgtaccct gtccgcggac caggtcaccg aggatctgca caccactgg	3780
cccgaaccggc tctgggtggc cgccaccaac tccccacgg caaccgtcat ctcggaac	3840
accgacgcac tcgacgaagc gctcgaccac taccacgccc acgacgtacg ggccaaacgc	3900
atcccggctg actacgcctc ccaactgcccc catatcgacg cggtagccga gcgactgccc	3960
gatctgctgg gcggcatcgt cccgcgcgc gccgacatcc ctttctact caccgttgac	4020
ggccgatggg ccgagccgac cgagctcgac gccgactact ggtaccgcaa cctccgcagc	4080
cccgtacggt tcgcccacgc cgtccacgcc ctaccgagg ccgaacacgc caccttcgtc	4140
gaagtacgcc cacacccac gtcaccccc gccatcacgg ccaccgccga aaccaccgac	4200
cgcaccacca ccgtcatcgc ctgcctccac cgcgaccacg aagacgtca ccacatctc	4260
accaacctcg cccaggccca catccacggc cacaccgtcg cctggcgaca ccaactaccg	4320
actctgcgc ccacccgcc ccacatcgac ctccccacct accccttcca acaccagcac	4380
tactggctcc acgactccac cgaggacaag gcggtgggta cggacctgc tcgggccgc	4440
ttctgggagg cagtcgacgg cgaggacacc aacgccgtcg ccgcgtcct cgacgtcgag	4500
ccgggcacct cgttgacgc gctgtgccg gccctgtccg cctggcacgg tcggcgtgc	4560
gaccaggcca tcaccgacac ctgggtttac cgggacatct ggaagccggt cgacctacc	4620
gccgcgcgc cccgaccgtc cagccgatgg ctgtgcgca tctccgagg gcgggccgat	4680
cacctccag tcagtccgt cctggacgt ctggaacgcc aggtctgcc catcgccacc	4740
ctgtctctg acgacacca catcgaact cccctgtgg agcggcatct cgcacaggtg	4800
atcgcgagcg atgggccggc catcgccggc gtgctctgc tgctgcct cgacagggg	4860
ccacatccgc gccaccgga ggtgccgct gccaccgccc tcacctcag cctgatccag	4920
gcgtcatcg cagtgagga catcgcccc cggctctggc tggccacca cgaggccgtc	4980
gccacctgt ccgcggatac gctcgatcac cccctccagg cgatggtctg ggggctggga	5040
cgcaccgcc cctcgaaca cccgatctg tggggcggac tcacgacct tccggacact	5100
ctcaccgaac gggctctccg cggcctctgc acggcgctga ccacctgtca cgacaggac	5160
gagctcgcg tcgcgccac cggccacgc acccgccggc tggtcggac gccgtccac	5220
gccgcggcgg aggacacccc gccgtggacg cccgtggca ccgtcctcat caccggcggc	5280
accggggccc tcggctccg cgtcgccac cgcacgccg aacgtaccc cggtgccac	5340
ttgtctgtg tgagccggcg aggggccaac gccccggcg ccaccgcgt ccgcgaccag	5400
ctcatcgaac tcggcgccac ggtgacctc gccgtatgtg acaccgccga cccggcgcg	5460

ctcgcggtatc tctcgcgca tgtccctcgc gccgcctc tcaccggtg cgtccacacc	5520
gctggcggtcc tggacgacac caccctcgcc gtacagaccc cggaccacct cgcgcggtt	5580
ctggggccca agtcccatgc cgcacacat ctgcacgccc tgcgccagca ccacccctc	5640
gacgcgttcg tctctttctc gtccgtcgcg gcgccttcg gtgccgcggg ccaggccaac	5700
tacgcggccg ccaacgccta cctcgacgcc ctgcgccggc accgccgggc ccaggggctg	5760
gccgccacct ccatcgctg gggcaactgg gacggcgacg ggctcgcgag caccagtc	5820
gcgcagacgt acctgcgcaa ccgcggcttt cctcccatgc cgccacacct ggcgtggcc	5880
gccatggagc gagcggtcgt ctgcgccac gccagctcg tctcgccga cgtcgactgg	5940
aagaagctca agccgacgcc gcacaccgc gacatcccgg aaagccggc cccggccccg	6000
gccgccaccg acggcgagc caggaccgcc gacgccaccg cgagcctccg taccgcctc	6060
gcgggtcaga gccggccga acggcaccag acgtcctcg acctcatcag ctctcataca	6120
gccgcgtcc tggggcagc cagccccag acgatccca cggaccgggc cttccgcgac	6180
ctgggtttca cctcgctgac ggccatcgag ctccgcaacc gcctcgcggc ggccaccggg	6240
ctccgcctgc cgaccacct cgccttcgac cgcgcgacgc cggacaagct cgcggcggac	6300
ctgctggcgc ggtgcgcgcc gacggggccg gacggcatcg gagtacagc cgacgcgacg	6360
gccgcgagcg gcagttcgcc cggtcggcg catggcgcg tggacccgc cgagcccatc	6420
gccatcgctg gctgggctg ccgtacccc ggccgggatcg gctccccga ggacctgtgg	6480
gagttcgta ccgcacaccg ggacgccgtc ggagacttc cgaccgaccg gggtgggac	6540
ctggcgaggc tcttcgaccc cgatccggac cggccgggca cctcgtacag ccgacagggc	6600
gccttctcc acgacgcggg cgaattcgac ccggagtct tgggatcag cccacgggag	6660
gcgacggcga cggaccccca gcagcggtg ctcttgaga cgtcctggga agccctcga	6720
cgagccggga tcaaccgca cgatctccac ggcagtcga cggcgcttt caccgcgac	6780
aacgcgcagg acttcagcg acggctgcgg cagacgccgt cggagctggc ggagctgtgc	6840
gagggtatg cgtgacggg cagcaacaac agcgtcgct cggggcgct ctctacgcg	6900
ctcgccctgg aaggcccggc ggtcagcatc gacaccgct gctcgtctc gctcgtggcg	6960
ctccatctgg cctgccagtc gtcggggcc gcgaatgct cgttgcct gccggcggc	7020
gtcacgtca tgaagaccc gttaacttc gtggagtct cccggcagcg gggcctggcg	7080
gcggacggcc ggtgcaaggc gttctcgcc accgccgatg gcaccggctg gggcgagggc	7140
gtgggcatgg tgggtggga gcggctgtcg gacgcggcg gcaacggcca tctgtgtctg	7200
gccctggtcc gccgcagcg cgtcaaccag gacggtgcca gcaatgggt gactgccccg	7260
aacggccct cgacgagcg ggtcatccg gccgccctgg ccgccggcg ggtcaccgcg	7320

gcagaggtgg acgcggtcga ggcgcacggc acggggacga cgctcggcga tccgatcgag	7380
gcccaggccc tgctcgccac ctatgggcag gggcggccgg cggaccgggc gctgtggctc	7440
ggtacggtca agtccaacat cggacacgcc cagtcggccg ccggtatcgc cggggtcac	7500
aagatggtgc tggccctgcg gcacgggatg ctgcccgta cgctgcatgt gtccgagccg	7560
tcgcccatg tggactggtc ggcgggtgcg gtacggctgc tgaccgagga ccagccgtgg	7620
ccggacaccg ggcgccccg gcgggcgggg gtgtcgtcct tcggcgtgag cggcaccaac	7680
gcccattga tcctggagca ggcggagccg gggccggacc cggaccgggc gccgacggcc	7740
tccgcgact ccgtgctccc ctggccctc tccgccaggt cggcggaggc cctgcgggccc	7800
caggccccga ggttgcgggc gtacgtggcc gagcaccggc atgtcgacc cggcgacgtg	7860
gggtactccc tcgcgcgcgg acgggccacc ttcgagcacc gggccgtgct cctcggcacc	7920
ggccacgacg acttcggcg cggcttggac gccctggtgt cgggcgcgcc cgacggcgcg	7980
gtcgtccagg gcgcggcggg ggggcggcag ggcaaggtcg tctttgtgtg ctcggggcag	8040
ggcaccacg gccccggcat gggccgcggg ctctaccgt cgtccacggc gttcgcggg	8100
gcgttgagg aggtgtgcgc ccatctggac ccgtatctgg aacacctct gatggaggtg	8160
atgttcgccg acgagaagag cgatacgtcg gcgtgtctgc atctgaccgc ctacgcccc	8220
ccggccctct tcgccctcca gaccgcgtg catcgcatgg tcaccgagga gttcgggctc	8280
accccgact atctggccgg ccactccctg ggcgagctga ccgccgccca tctggcgggc	8340
atcctcagcc tggccgacgc cgcgcgctg gttgcggccc gcgcccgcc catgcgggac	8400
cttcggcg cggagccat ggtcgcgtc gaggccaccg aggcgaact gcggcctcgg	8460
ctcgcgagt tggcggagcg ggtcgacatc gccgccgtca acgccccgc gtccctggtc	8520
atcaccggcg accacggcgc cgtgcaccag atcgccgacg acttcgcgc gcagggcagg	8580
aaggtcacct cctccaggt cagcggcgcc ttccactccc cccatatgga gccctgctc	8640
gacgagatcg ggcgcaccgc cgaaacctc acctaccacc ggccccacac tctcctcgtc	8700
accgcatcgg cggacggcgg cgacgacacg atcgagccgc gggccgacga cgaccgggc	8760
acggccgct tctggcctct ccaggcccgg cgcaccgtgc actacgcacg ggccgtggag	8820
cggctgcacg cccgcggcgt caccacgttc ctggaactcg gccccgacgc caccctcacc	8880
gccctcgtcc accacaacct gcgcgcgcac gatccctgg ctgtctccct gctccatccg	8940
gagcgggtgc agacgcacag cgtcctcggc gcgtcgcgc cggtcacgc ccacagccgc	9000
cccgtcgact ggacgcgcca ctacaccgca cggccgcggc cgacgccaca ccagatcgac	9060

gtgcccacct atgccttccg gcaccggcgc tactggctgc ccgccccggc ggcggtcggc	9120
gatgtgacgg ccgcggggct cgacgcggcg gagcaccgc tgatcggcgc cccgtgggg	9180
ctgcggagg gcgacggctg tctgtgacc ggcaggatct cgcccgctac gcaccgtgg	9240
ctggccgacc atgtcatcgt cggcaccgtg ctgcttccgg gcaccgcgtt cgtggagctg	9300
gcgctgcggg ccggggcgta tgtgggctgc ggccgtgtgg aggagctgac cctgcacgcg	9360
ccgctccccg ccgacggtag ggtggtgctc caggtagcgg tggggggcgc cgacgagtc	9420
ggcccgctg agctgagcat tcacggcgcg ccggcggacg acggtacatg gacacggcac	9480
gccatcgga cgtggcacc ggcccacgac gtcgacggg gtcaagatgg ccacggccc	9540
gcgatgacg ggcagttcgg gtcgtgggac acggcctggc cgccggccgg tcgggagccc	9600
ttggacgtca ccggggtcta cggccggttt gccgacggc agttcacgta cggggaggcc	9660
ttccaggggc tggtcggcg ttggcggcac ggcgacgaga cgtggcgga ggtccgcctc	9720
cccaccagc cggccggtag cggccaccgc ttcgggctgc acccgcgct gtcgacgcg	9780
gcactgcaga ccatgtggct cgtggagccc gacggcacac ggccgacggg tggcctgggc	9840
ggcccgatc ggggcctgcc gttcgctgg cagggggtct cgtgcgtac ggcggggccg	9900
tcggccctgc gggtaggct gcgacggcgg gcgccggaca ccgtggccgt cccgtggct	9960
gaccggccg gccgaccgt cgcgctgggt gagtcgctga cgtgcggcc ggtgcccg	10020
ggcgcttgc gggcgccga ggcggcggtg cgacctcgt tgcacggcct ggactggacg	10080
gatgtgccg tgcgacgcc gccccggcc cggccccgt gtgcgtgat cggagcggac	10140
acgtcggcc tgggccccg gtcgaggcc gggcgcccc accgcatcac cgacggcgtg	10200
gagcgtacg ccgacctga ggagctggtg cgtccgtgg cggcggcgcc cccgccccg	10260
gacctgtca tcggacctg ccacacagcc cctgaagccg acggcgcgag cgaacagcca	10320
cagcccgaga cggtagcac aaggacgggt cagggtgctgg agctgcttca gcggtggctc	10380
ggcgcgacg ggctcgcca cgacacctg gtgctgttca cctcaggcgc ggtcgccacc	10440
cggccggcg agctggtgct gacctggcg gggcgggccg tctgggtct ggtgcgtcc	10500
ggccagtgg agcatccga gtgcttcacc gtggtggaca tggacggcgc ccaggagtc	10560
cgcgggcg tgctcgcg gctcgccctc ggcgagcctc aactggcggg gcgcggcg	10620
cgggcgtgg cggcgccct ggtgcggccg ggtgccgag ccgacgacag cggcctggcc	10680
ctgcccggg ggcgggaagg ctggcggttg gagtgtccc gcacgggcag cctggacggg	10740
ttgaccacga ccgagtcgcc ggcccgggcg gtgccgtcg gccggggcga ggtacgggtc	10800
gcggtcggg ccgcggggct gaacttcgc gatgtgtga tcgcgtggg cgtggtgcc	10860
ggcgggacg cgtgggcag tgagggggcg gggatcgtcc tcgaggtcgg ggcggaggtc	10920

cgcgatctca cgccccggga ccgggtggtg ggtatcttcc ccgaggcggtt cgccccggtg	10980
gccgtggccg agcgggcgac cttggcgcgg atccccgacg gctggtcgtt cgcccaggcc	11040
gcgtcggtcc ccatcggtt cgccaccgcg taccacggcc tggtcgatct ggccgcctg	11100
cgccggggg aatcggtgct gatccatgcc gcggccggcg gggtgggcat ggccgccgtg	11160
caactggcgc gccatctggg ggccgaggtg tacgccacag ccggccccgg caagtggcac	11220
atcctgcgct cccaaggcat cgacgacgac catctggcgt cgtcgcgac gctggagttc	11280
gagcagcgct tcgccgac ccacggcggg cggggcatcg atgtcgtcct ggactgtctg	11340
gccccatgagt tcgtcgacgc ctcgtgcgc ctggtggcgc gtgacggcgg ccggttcctg	11400
gagatgggca agagcgacat ccgtgaccgc cggcaggtgg cgctggacca tccgggcgtg	11460
ctctaccggg cggtcgacct gttggaggcc ggcccgagc gggtcgggca gatcctgcgc	11520
accgtactgg acctgttcga gcgcggtgtc ctggcgcacc tgccgacgac ctgctgggac	11580
atccggcagg cggagcaggc cttccgcat ctgcagcagg gccgccacat cggaaagaac	11640
gtgctcaccg tcccgccgg ctggaacgcc gagggcaccg tactgatcac cggcggtaacg	11700
ggcacccctg gtgccgcct cgctcgccat ctggcgggta ccgggcgcgc ccgccatctg	11760
ctgctggtcg gccgacgcgg ccccgacgcc ccgggcgcgg aggagctgcg agaggagctg	11820
accgagctgg gcgcgcgggt caccatcgcc gcatgcgac tcggcgaccg ggccggcggtc	11880
gcccggctcc tggggcgcat cccggccgag cggccgtga ccgccgtcat ccacgcggcg	11940
ggtgtcgtcg acgatgccac cctcgggtcc ctaccccc gccacctgga cgccgcctg	12000
gccgccaagg ccgacgcgc ctggcatctg cacacctca cccgccacgc cgactggcc	12060
gcgttcgtcc tcttctctc ggtcgccggt ctgctcggt cgcccgggca gggcaactac	12120
gccgcggcca acgcttctt ggacgcgtc gccaccacc ggcgctgctc tggccttcg	12180
gcggtgtcgc tggcgtggg gctgtgggag cagaccagcg gcatgaccg agacctggac	12240
caggccgacc gcgcccggct ggcccggctc ggcatcagcc cgctcacgac cgggcaggcg	12300
ctcgaacttt tcgacaccgc cctcggccac caccgccccg tgcctgtccc cgccgcctc	12360
gacgtgcccc accgcaccc cggtcgtcg accgtgccgc cctgtaccg gggcctggtc	12420
ggatccagga cccggcggac accccccgcg tccgcccca ccgggcggtt cccctgcat	12480
accgcctcg acggtcacgc cccggccgag cagcacgaga tgctgctc gctggtcgc	12540
tcgcacgcc ctctcgtgt gggccgcgac gatccggaca cggccatcc cggcgcgac	12600
ttccgcggtc tgggcttcga ctccctgacc gcggtcgagc tccgcaatcg gctcaacgcc	12660
gccaccggcc tccggtctc caccacctc gtcttcgacc accccacgcc cgacgaactc	12720
gcccgtcacg tccgggagca ggtgtgggc gacggcgaag cggcgcggtt ggccccggtg	12780

ctggccgagc tcgacaggct ggaggccgcg ctgtcccggg tgaacgggga cgatgcgctc 12840

cgggcgaggg tgacggcccg gctgcaggcc ctctctctga agtggaaaca gtccgatggt 12900

ccggcgacgg gcgcagacgg tgcgggcagg ctggcgctcg ccacggccgc cgagggtctg 12960

gatttcatca ggaacgacct cggcctctcc tga 12993

<210> 70

<211> 13032

<212> DNA

<213> Artificial Sequence

<220><223> milA1 gene of Streptomyces milbemycinicus

<400> 70

ttgccaaag ccagaaaca gttcgagtg gccggtcatc cgtggatcct ctccgggcac 60

accggaaccg cgctcggggc ccaggcacgc cggctccacg accatgtcgc cgaccacct 120

cggtccgtc cggaagacat cggccacacg ctggcgagca gcggcccggc gtcacccat 180

cgcgccggcg tgatcgccgc ggaccgggaa ggacatctcc gggggctcga cgcggtggcc 240

cggggtgagg acacccccg tgtctgacgg ggcacggcgg ccgcgggcgg cgacggggtc 300

gcgttctgt tccccggcca gggcaccacg tggccccgta tggccgccga tctgtgacg 360

gtctccccg cttcagccg ggcggtcgac gcctgcgcgg aggcgttcga accgtatgtc 420

tcctgtgtac cggaggccgt gctgcggggc gctccgggcg cgcccccct ggaggggacc 480

gatgtggtgc agccgacgt gttcgccgtc atggtggggc tggccgagct gtggcggact 540

cttgggggtga gcccgacgtc gatcgtgggc cactgcatcg gggagatcgc ggcagcccat 600

ctctcgccgc cctgtcgtc gtccgacgcg gcgcgcgtgg tgatcgagag cagccgggcc 660

caggcgacgc tctccgggtc ggggtcgtc atcgcggtcg cgcggtccga ggcgcagctg 720

cttccgttgc tgcggcggtg gccgggcagg ctgacgatcg ccgcggtcaa cggcccgatg 780

gccacggtcg tctccggcga tcggccggcc gccgacgagc tgttggcgga gttcgcccgt 840

gccggtgtcc gggcccgcga ggtggcgatc gacatccccg cgactcgcg gttcatggcc 900

cccctcaggg acggtctgtc cgactcgtc tcatcggtca ccgcgggtgc gtcgcggtg 960

ccgttccact cctcggtcat cggggggccg ctggagaccc aagggtcga cgcggttac 1020

tggtagcgga acctcgccga cacggtccgc ttcgaaagcg tcgtcacggg gctgctcgcg 1080

cagggcacac gctgcttctg ggagctgagc ccgcacccga tgctgacct gtgtgtgcag 1140

gccacgcgcg aggaggtggt cggcggtgag cgcgtctgta tctgcccac gctgcatcgc 1200

gggcaggccg ccgtcgagtc cgttcgacc acgtggccg agctgtacgt acggggcgca 1260

ctggatgacc atcgggcggc gttctcggtg ccgggcggcc gcctgacac cctgcctctc	1320
gagccgcccc cggacacgtc cgtagagctc gccgacgccc cggacccggc ggaggcctgc	1380
cgccccccct tggtaggagc gcttccccgg ctctccaccg cggagcggaa gcggcggctg	1440
cgcgagctgg tgggcgtcga ggcggccaag gtctctgagg acgtcgccgg gcgggacgcg	1500
ccgggccacg gcatcgcgga gcaggagcac ttctgtcactt cgggcttcga ctccgcggcc	1560
gcggtcgcgc tgcgcaaccg cctgaacgac gccaccggtt tgcgtctgcc cttcaccttg	1620
gccttcgacc atccgacacc cggcgccgtc gccgaccatc tgcactcccg gctcttcgat	1680
caccaggggc gcgggcagcc gggcgccgac ggccggcccc accccgcggc gcgggcccgt	1740
ccggccaggg ccgacgacga gccgatcgcc gtcactggca tggcggggcg cttccccggg	1800
ggcggccgta ccccgaggga gctgtgggaa ctggtcgccg aaggcaccga cgcctctcgc	1860
cccttcccgg agggcggggg ctgggatacg ctgcggctct acgatccgga ccccgcccgg	1920
cccggcacgt actaccagcg cgaagcggga ttctctcacg acgccgacaa gttcgacgcc	1980
gagttcttcg gcatcgcgcc acgcgaggcc accgcaatgg atccccagca gcggctgctc	2040
ctggagacct cctgggaggg gctcgaacgg gcgcggatcg acccgaccgc gctgcggggc	2100
agccgcaccg ggggtgttcgt cggcgtggcc ccgctggact acagccccg aatgcaccag	2160
gcgtcgccgg agctggaggg ccatctgctg accggcaaca tcggcgccgc ggcctcgggg	2220
cggatctcct acgtactcgg gcttgagggg cccgcggtgt ccgtggacac ggctgtctcg	2280
tcgtccctgg tcgcccgcga tctggcggcc caggcgctgc gggccgggga gtgctcgctg	2340
gcccctggcg gcggggcgac ggtctctctg acccccggca tgttcacga gttctcgcg	2400
cagcgcggtc tggctccgga cggccgctgc aaggcgtacg cggccgcgcg ggacggcacc	2460
ggctggtccg aggggtgtgg catgtgtctc gtcgagcggc tgtccgacgc gcgacggctc	2520
ggacaccagg tgcctgcggt ggtacggggc tccgccgtca accaggacgg ggcgagcaac	2580
ggcttcacgg cgcccagcgg tccatcacag caacaggtca tccgggcggc cctggccaat	2640
gccggggtgt cggtccgga ggtcgacgcg gtggaggggc acggcaccgg caccgggttg	2700
ggcgatccga tcgaggcgca ggcgtgtctg gccgcctacg ggcaggggcg gcgggccgac	2760
cggccgctgt ggctggggtc gatcaagtgc aacatcggac acaccagtg gcccgcggtt	2820
gtcatcgggg tcatcaaaat ggtgtctcgc ctccagcacg gtgtgtctgc gcgcacgtg	2880
cacgtggaca agccgtcgga ttactggac tggtcggccg gggccgtacg gctgttgacg	2940
gagccggtgc cctggccgga gcggggccac ccgcgccggg cgggggtgtc gtccttcggg	3000

gtgagcggca ccaacgcccc tgtcatcttc gagcaggcaa cgccatcgtc cacggtggct	3060
cccggggggc ataccgccga ggccgggcct ccctgccgt gggtgggtctc ggcgaagacg	3120
ccccaggcac tgcgcgacca ggcccgccgc ctgcacgaac acctcaccgc ccagccacag	3180
ctccaaccgg ccgacgtcgg ccacaccttc gccaccggcc gcgccacctt cgaccaccgg	3240
gccgtctca tgggtccga ccgcgaacaa ctctccacg gcctggacgc gctcgccacc	3300
ggccggcccc acccagcggc ccaccagacg tcggaccgtc ccgccaccgc cgacggccgt	3360
atcgtcttcg tcttccccgg acaaggcggc caatgggcgg gcatgggcct acggctgctg	3420
aacgcctcac ccgtcttcac cgagcggatg gccgcctgcg aacaggccct ctccccctac	3480
gtcgactggt cactcacgga catctccac cgccggcccg acgacgccgt atggcaacgc	3540
gccgacatcg tccagccgc cctgttctcg atcatggtgt ccctggccgc gctctggcgc	3600
tcttgcgga tcgaaccgga cgccgtctc ggccactccc aaggcgagat cgccgcggcc	3660
cacgtctgcg gcgccctgac gctccacgac gcggccaagg tcatcgccct gcgcagccag	3720
gccctccaag ccgtacgcgg cgccgggggc atggcctccg taccctgcc cgccgaccag	3780
gtcacgagg atctgcgac cactggccc gaccggctat gggtggccgc caccaactcc	3840
cccacggcaa ccgtcatctc gggaacacac gacgcgcttg acgaagcgct cgaccactac	3900
cacgcccacg acgtacgggc caagcgcatc ccggtcgact acgcctccca ctgccccat	3960
atcgacgcgg tggccgagcg actgcccgc ctgctgggcg gcatcgctcc gcgcgccgc	4020
gacatccct tctactccac ggttgacggc cgatgggcgg agccgaccga gctcgacgcc	4080
gactactggt accgcaacct ccgcagcccg gtacggttcg cccacgccgt ccacgccctc	4140
accgagaccg accaccgcac ctttctgaa gtcagccac accccacgt cccccgcc	4200
atcacggcca ccaccgaac caccgaccgc accaccaccg tcatgcctc gctccaccgc	4260
gaccacgacg acaccacca catctcacc aacctgccc aggccacat ccacggccac	4320
accatcgact ggcgacacca ctaccagact ctgcgcccc cccacccca tatcgacctc	4380
ccacctaacc cttccaaca ccaccactac tggctccag actccaccga ggacaaggcg	4440
gtgggtacgg acctcgccgc ggcccgttc tgggaggcgg tccacggcga ggacaccaac	4500
gccgtcgccg cgctctcga cgtcgagccg ggcacctcac tggacgcgt gctgccggcc	4560
ctgtccgctt ggcacggtcg gcgtcgcgac caggccatca ccgacacctg gtgttacgc	4620
gacatctgga agccggccga cctcaccgcc gcgcgcccc ggccgtccgg ccgatggctt	4680
gtcgcgatct ccgcaggcg ggccgatcac ctccacgtca gtgccgtcct ggacgtctg	4740

gaacgccagg gtctgcccac cgccaccctc gtcctcgacg acaccacac cgaactcccc	4800
ctgtctggagc ggcatctcgc acaggcgatc gcgagcgatg ggccggccat cggcggcgtg	4860
ctctcgctgc tcgccctcga cgaggggcca catccgcgcc acccgaggt gcccgctggc	4920
accgccctca cctcagcct gatccaggcg ctcatcgac gcgaggacat ggcgccccgg	4980
ctgtggctgg ccaccacga ggccgctgcc acctcgtccg cggatacgt cgatacccc	5040
ctccaggcga tggctctggg gctgggacgc accgccgac tcgaacacc cgatctgtgg	5100
ggcgccctca tcgacctcc ggacactctc accgaacggg tcctccacgg cctcgtcacg	5160
gcgctgacca cctgtcacga cgaggacgaa ctcgcgctgc gcgccaccgg cccacgcacc	5220
cggcgccctga tccggacgcc gtccaccgcc gcagcggagg acacccgcc gtggacgccc	5280
cgtggcaccg tcctcatcac cggcggcacc ggggccctgg gctcccgt cgccaccgc	5340
atcgccgaac gccacccga ctgccactg ctgctggtga gccggcgagg gcccaggcc	5400
cccggcgcca ccgcgtccg cgaccagctc atcgaactcg gcgccacgt gaccctcgcc	5460
gcatgtgaca ccgccaccc cggcgcgtc gcggatctcc tcgccgatgt cccctcgac	5520
cgccccctca ccgcgtcgt ccacaccgcg ggcgctctgg acgacagcac cctcgccgta	5580
cagaccccg accacctgc cgccgttctg gggcccaagt cccatgccgc acaccatctg	5640
cacgccctcg ccagcacca cccctcgac gcgttcgtcc tcttctctc cgtcgcggcg	5700
cccttcggcg ccgcgggcca ggccaactac gcggccgcca acgctacct cgacgcctc	5760
gcccagcacc gccgggcca ggggctggcc gccacctca tcgctgggg caactgggac	5820
ggcgacggcg tcgcgagcac acagtccgc cagacgtacc tgcgcaacc cggtttcct	5880
ccgatgccg cacacctggc gctggccgcc ctggagcgag ccatcgtctc gcccacgcc	5940
cagctcgtcg tcgccagct cgactggaag aagctcaagc cggcgccgca caccgcgac	6000
atccccggaa gccgcgccc gggcccgcc gccaccgacg gcgcggacag gacggccgac	6060
gccaccgca gcctccgtac ccgcctcgc ggtcagagc cggccgaac gcaccagacg	6120
ctctcgacc tcacagctc tcatacagc gccgtctctg ggcacgccac gcccagacg	6180
atccccacgg accgggcctt ccgcgacctg ggtttcacct cgtgacggc catcgagctc	6240
cgcaaccgcc tcgcggcgcc caccgggctc gcctgccga ccaccgtcg cttcgaccgt	6300
ccgacgccg acaagctcg ggccgacctg ctggcgcggt gcgcgccgac agggccggac	6360
ggcatcgggg tgacgccga cgcgacggc acgagtggca gttcgcccg tcggcgcat	6420
ggcgcgccgg accccgccga gcccacgcc atcgtcggt tggcctgcc ctacccggc	6480
gggatcggt ccccgagga cctgtgggag ttcatcacg cacaccgga cggcgtcgga	6540
gacttcccga ccgaccggg ctgggacctg gcgaggctct tcgacccga tccggaccgg	6600

ccgggcacct cgtacagccg acagggcgcc ttcttcgcg acgcgggcga cttcgacccg	6660
gagttcttcg ggatcagccc acgggaggcg acggcgacgg acccccagca gcgactgtc	6720
ctggaggcgt cctgggaagc cctcgaacga gccgggatca acccccacga tctccacggc	6780
agtccgacgg gcgtcttcac cggcagcaac gcgcaggact tcagcgcgcg gctgcggcag	6840
acgccgtcgg agctggcgga gctgtgcgag ggctatgcgc tgactggcag caacaacagc	6900
gtcgctcgg ggcgctctc gtacgcgtc ggcttgaag gcccgcggt cagcatcgac	6960
accgcctgct cgtcctcgt cgtggcgctc catctggcct gccagtgcgt ccgggcccgc	7020
gaatgtctgc tggccctggc gggcgcgctc acggtcatga tgacccgtt caacttcgtg	7080
gagttctccc ggcagcgggg cctggcgcg gacggccggt gcaaggcgtt ctccgccaca	7140
gccgatggca ccggctgggg cgaggcggtg ggcatggtgg tggaggagcg gctgtcggac	7200
gcgcggcgca acggccatcg tgtgtggcc ctggtccgcg gcagcgccgt caaccaggac	7260
ggtgccagca atgggctgac tggccgaac ggccctcgc agcagcgggt catccgcgc	7320
gccctggccg ccgccgggt cgcccgga gaagtggacg cggtcgagc gcacggcacg	7380
gggacgacgc tcggcatcc gatcgaggcc caggccctgc tcgccaccta cgggcagggg	7440
cggccggcgg accgggcgt gtggctcgt acggtcaagt ccaacatcgg acacgccag	7500
tcggcccgcg gtatcgccgg ggtcatcaag atggtgtcgg ccctgcggca cgggatgctg	7560
ccgctacgc tgcatgtgc cgagccgtcg ccgcatgtgg actggtcggc ggggtcggta	7620
cggctgctga ccgaggacca gccgtggcg gacaccgggc gccccggcg ggcgggggtg	7680
tcgtccttcg gcgtgagcgg caccaatgcc catgtgatcc tggagcagc ggagccgggg	7740
ccggaccggg caccgacggc ctccgcgcc tccttgccc cctggccct ctccgccagg	7800
tcggcgagg ccttcgggg ccaggccgt aggttgctgg cgtacgtggc cgagcaccg	7860
gatgtcgacc ccgccagct ggggtactcc ctgcgcgcg gacgggccgt gttcgagcac	7920
cgggcctgc tcctcgccac cggccacgac gacttcggc gcgccctgga cgccctggcg	7980
tcgggcgcgc ccgacggcg ggtcgtccag ggcgcgcg tggggcgga gggcaaggtc	8040
gtctttgtgt gctcggggca gggcacccag cgccccgga tgggccgcg gctctaccg	8100
tcgtccacgg cgttcgccg ggcgtggag gaggtgtgc cccatctgga cccgtatctg	8160
gaacaccctc tgatggaggt gatgttcgcc gatgagaaga gcgatacgtc ggcgtgctg	8220
catctgaccg cctacgccc accggccctc ttgcctcc agaccgcgt gcaccgcatg	8280
gtaccgagg agttcgggt cacccccac tatctggcc gccactcct gggcgagctg	8340
accgcgccc atctggcgg catctcagc ctgccgacg ccgcggcgt ggtggcgcc	8400
cgccccgcg ccatcgggga ccttcagcg accggagcca tggtcgccgt cgaggccacc	8460

gaggcggagc tgcggccccc gctcggcagc ttggcggacc gggtcggcat cgccgcccgc	8520
aacgcccccg cgtccctggc catcaccggc gaccacgacg ccgtgcacca gatcgccgac	8580
gacttccgcg ggcagggcag gaaggtcact cccctccagg tcagcggcgc ctccactct	8640
ccccatatgg agccccctgt cgacgagatc gggcgccaccg ccgaaaccct cacctaccac	8700
cggcccccaca ctccccctgt caccgcgtcg gccgacggcg gcgacgacac gaccgagccg	8760
cgggcccagc acgaccgggg caccggccgcg ttctggcctc tccaggcccg gcgcaccgtc	8820
cactacgcgc gggccgtgga gcggctgcgc gcccgccggcg tcaccacgtt cctggaactc	8880
ggccccgact ccacctcac taccctcgtc caccacaatc tcgccgcgca cgatcccgtg	8940
gccgtctccc tgctccatcc ggagcgggtgc gagacgcaca gcgtcctcgg cgcactcgcc	9000
gcggtccacg cccacagccg ccccgctgac tggacacgcc actacaccgc acggccgcgg	9060
ccgacgccac accagatcga cgtgcccacc tatgccttcc ggacccggcg ctactggctg	9120
ccgccccggg cggcgggtcgg cgatgtgacg gcccgccggcg tcgacgcggc ggagcaccgc	9180
ctgatcggcg ccgccgtgtg gctcgcggag ggcgacggct gtctgtgac cggcaggatc	9240
tcgcccgcta cgacccgtg gctggccgac catgtcatcg ccggcactgt gtccttccg	9300
ggcaccgcgt tcgtggagct ggcgctgcgg gccggggcgt acgtgggctg cgaccgtgtg	9360
gaggagctga cctgcacgc gccactcccg ctccccccg acggtgaggt ggtgctgcag	9420
gtggcgggtg gggccgccga cgagtcggc cgcctgagc tgagcatcca cgccggccg	9480
gcggacgacg gtacatggac acggcacgcc atcggcacgc tggcatcggc ccgcggcgtc	9540
ggcctcgacg atggcacggg gcacaatggc cagccccgg cgggcgacga gccgttcggg	9600
tcgtgggcca cggcctggcc gccgcccgt gccgagccct tggacgtcac cggggtctac	9660
gaccggtttg ccgacgccga gttcacgtac ggggagcagc tccaggggct ggtcgcggct	9720
tggcggcacg gcgacgagac gctggcggag gtccgcctcc ccgaccagcc ggccggtgac	9780
gccctccgct tcgggtgca ccccgcgctg ctcgacgcgg cactgcagac catgtggctc	9840
gtggagcccc acggcacacg gccgagcggg gccctgggcg gccccgatcg gggcctgccg	9900
ttcgctggc aggggggtctc gctgcgtacg gcgggcccgt cggccctgcg ggtacggctg	9960
cgacggccgg cggcgacac cgtggccgtc gccgtggccg acgcggccgg ccggccggtc	10020
gcgtcgggtg agtcgtgac gctgcggccg gtccgcggg gcgccttgcg cggcaccgag	10080
acggcgggtg gcacctcgtt gtacggcctg gactggacgg atgtgccgt gccgacgccg	10140
cagacggccc tgccccggtg tgcgtgatc ggagcggaca cgctcgacct ggtccccgcg	10200

ctcgaggccg cggcgcccga ccgcatcacc gacggcgtgg agcgctacgc cgacctggag	10260
gagctggtgc gctccgtggc ggccggcgcc cccgccccgg acctcgtcat cgccggctgc	10320
cacgcagccc ctgaagccga cggcgcgagc gaacagccac agcccagagc ggtgcgcaca	10380
aggacgggtc aggtgctgga gctgcttcag cggtagctcg gcgcggacgg gctcgccgac	10440
gcacacctgg tgcgtgtcac ctacggcgcg gtcgccaccc ggccggcgga gccggtgcgg	10500
gacctggcgg ggccggcggt ctggggcttg gtgcgctccg gccagtcgga gcatccggag	10560
tgcttcaccg tggtagacat ggacggcgcc caggagtccc gcgcggcgct gctcggcgcg	10620
ctcggcctcg gcgagccga actggcgggt gcggcgccgg gggcgctggc gccgcgcctg	10680
gtgcgcccgg gtgacgccga cgacgacagc ggccctggccc tgccgcaggc gccggaaggc	10740
tggcggttgg agtgtcccg cacgggcagc ctggacgggt tgaccacgac cgagtccccg	10800
gcccgggcgg tgccgctcgg cccggcgag gtacgggtcg cggtagggc gcggggctg	10860
aacttcgcg atgtgctgat cgcgctgggc gtggtgccc ggccggacggc gctgggcagt	10920
gagggggcgg ggatcgtcct cgaggtcggg gcggaggtcc gcgatctcg gcccggggac	10980
cgggtggtgg gtatcttccc cgaggcgttc ggcccggtgg ccgtggccga gcgggcgacc	11040
ctggcgcggg tccccagcg ctggtcgttc gccagggccg cgtcgggtccc catcgtgttc	11100
gccaccgct accacggcct ggtcgatctg gcgcgcctgc ggccggggga atcgggtctg	11160
atccatgccg cggccggcgg ggtgggcatg gccgccgtgc aactggcgcg ccatttgggg	11220
gccgaggtgt acgccacggc cggccccggc aagtggcaca tctcgttc ccaaggcatc	11280
gacgacgacc atctggcctc gtcgcgcacg ctggagtctg agcagcgctt cgcccgacc	11340
cgccggggc gggggatcga tgtcgtcctg gactgtctgg cccatgagtt cgtcgacgcc	11400
tcgtcgccc tggtagcgcg tgacggcggc cggttcctgg agatgggcaa gagcgacatc	11460
cgtgaccgc ggaggtggc gctggacat cgggcgtgc tctaccgggc gttcgacctg	11520
ctggaggccg ggccggagcg ggtcgggcag atcctgcga ccgtactgga cctgttcgag	11580
cgcggtgtcc tggcgacct gccgacgacc tgctgggaca tccggcagc ggagcacgcc	11640
ttccgccatc tgcagcaggc ccgtcacatc ggaagaacg tgcacaccg cccggccggc	11700
tggaacgcc agggcacctg actgatcacc ggccgtatgg gcacctggg gccgcctc	11760
gcccgtcatc tggcgggtac cgggcgcgcc cgcatctgc tgcaggcgg ccgacggc	11820
cccagcggc cgggcgcga ggagctgca gaggagtga ccgagctgg gcgcgggtc	11880
accatgccg catcgatct cggcgaccgg gcggcggtcg cccggctcct gggggcgatc	11940
ccggccgagc ggccgtgac cgctgtcatc cagcgggcgg gtgtcgtcga cgatgccacc	12000
ctcgggtccc tcacccccg ccacctggac gccgcctgg ccgccaaggc cgacgccgc	12060

tggcatctgc acaccctcac ccgccacgcc gacgtggccg cgttcgtcct ctctcctcg 12120
gtcgcgggtc tgctcggctc gcccgggcag ggcaactacg ccgcggccaa cgccttcttg 12180
gacgcgtcgc cccaccaccg gcgcggctct ggccttcagg cgggtgctgc ggcgtggggg 12240
ctgtgggagc agaccagcgg catgaccggg cacctggacc aggccgaccg cggccggctg 12300

gcccggctcg gcatcagccc gctcagacc gggcaggcgc tcggcctttt cgacgccgcc 12360
ctcggccacc accgcccgtg gctcgtcccc gccgcctcg acgtgcccga tccgcacccc 12420
ggctcgtcga ccgtgccgcc cctgtaccgg ggcttggtcg gatccaggac ccggcggaca 12480
ccccccggc ccgccccac cgggcggctt cccctgcata ccgcctcgg cggtcacgcc 12540
ccggccgagc agcacagat gctgctctcg ctggtccgt cccacgccgc ctcgtgctg 12600
ggccgcgacg atccggacac ggtccatccc ggcgcgact tccgcggcct gggttcgac 12660
tcctgaccg cggtcgagct ccgcaaccgg ctcaacgcc ccaccggcct cgggtctcc 12720

accacctcg tcttcgacca cccacgccc gacgaactcg cccgtcacgt ccgggagcag 12780
gtgctgggcg acggcgaagc ggcggggtg gccccggtgc tggccgagct cgacaggctg 12840
gaagcggcgc tgtccgggt ggacggggac gatgcgtcc gggcgagggt gacggcccg 12900
ttgcaggccc ttctcctgaa gtggaacgag tccgatggtc cggcgacggg cggtgacggt 12960
gcgggcaggc tggcgtccgc cacggccgc gaggtgctgg atttcacag gaacgacctc 13020
ggcctctcct ga 13032

<210> 71

<211> 16599

<212> DNA

<213> Artificial Sequence

<220><223> aveA3 gene of Streptomyces avermitilis MA-4680

<400> 71

atggacacgt ccagcgaaaa gtcgtcgcac gcgcttaggg cgtctctgaa ggcgaaccag 60
accctgcggg cacgtaatga gcaactggca gccgccatgg aggcgtccag cgagccgatt 120
gcgattgtgg ggatggcgtg tcgttttccg ggtggggtgt gttcgccgga ggagtgtgg 180
gagctggttg cgtcgggtgg ggatgcgatt ggtgaatttc cggccggtcg ggggtgggat 240
ctggaggggt tgtttgattc ggacctgac cggtcgggga cgtcgtacgc gcggtatggc 300
gggtttttgt atgaggcggg ggagttcgat gcggacttct tcgggatcag tccgcgtgag 360

gcgttggcga tggatccgca gcagcggttg ttgctggaga cgtcgtggga ggcgttcgag 420
cgggcgggta tcgatccgct gtcgatgcgt ggctcccgt cgggtgtctt cgccggggtg 480

atgtaccacg actacggatc ccgcctgggt accatccccg agggattcga gggctacatc	540
ggcaacggta gggcgggcgc cgtcgcgtcg ggccgcgtcg cctacacgct cgggtctcgag	600
ggccctgccg tctcgggtga cacggcatgt tcgtcgtcgt tgggtggcgt gcatctggcg	660
tgccagtgcg tgcggtcggg tgagtgacg ctgcgcgtgg ccggcgggtg gacgggatg	720
tcgacccgc accctctcgt cgagttctca cgccagcgcg gactgtcggg ggacggccgc	780
tgcaagtcc tgcgggtgg agccgacggc accggcatgg gcgagggcgt cgggatgctg	840
ttggtggagc ggttgtcgga tgcggtgcgg ctggggcatc ggggtgctggc ggtgctgcgc	900
ggcagtgcgg tcaatcagga cgggtcgtcg aatgggtga ccggcggcga tgggtccggt	960
caggagcggg tgatccggca ggcgttggcg aacgcggggt tgtccgtggc ggatgtggat	1020
gtggtggagg ggcattggac gggcacgacg ctgggtgatc cgatcgaggc gcaggcgttg	1080
ctgccacgt acgggcagcg ggccggtaac aggcgcgtgt ggctgggatc ggtgaagtcg	1140
aacatcgcc atgcgcaggc tgcccggggt gtgggtgggg tcatcaagat ggtgatggcg	1200
ttgcgggagg ggggtgtgcc gcggacgttg catgtggatg agccgtcgcc gcaggaggac	1260
tgggtccggg gggcgggtcg gctgctgacg gaggcgtgc cgtggccggg ggacgcggca	1320
gggcgggttc ggcggggcgg agtgtcgtcg ttccgggtca gtggcacgaa tgcgcatgtg	1380
atgttgagg aggcgccggc ggcggggggc tgtgttgcg ggggtggggt gttggagggt	1440
gtccgggtc ttgccatttc ggtggctgag tcggtggcgg ctccagtggc tgtgtctgcg	1500
ccggtggctg agtcggtgcc ggtgccgtg ccggtgccgg ttctgtgcc ggtgtcggct	1560
aggtctgagg ctgggttgcg ggccgaggcg gaggcgttc gtcagtacgt ggcagtcgg	1620
ccggacgttt cgcttgccga tgtgggtgcg ggtctggcct gtgggcgggc tgtgctggag	1680
catcgtgcgg tcgtcttgcc gcgcgaccgt gaggagctgg tgcaagggtt gggggcgtcg	1740
gcggcgggtg agccggatcg gcgggtgacc acgggtcatg ccgccgggtg tgaccggggc	1800
ggtgtcgtct tcgtgtttcc cggacagggt gggcagtggt ccgggatggg tgtgcgtctg	1860
ctgcctcct ctccggtgtt gcgccggcgg atgcaggcgt gcgaggagc tctggcgccg	1920
tgggtggact ggtctgtgtt ggacatcctg cgccgggacg cgggggatgc ggtgtgggag	1980
cgggccgatg tggccagcc tgtgtgttc agcgtcatgg tgtctttggc tgctctgtgg	2040
cgttcctacg gtatcgaacc cgacgcgtc ctgggccatt cccagggcga gatcgcgcc	2100
gcgcatgtgt gtggggcgt gagcctgaag gacgcggcga agactgttgc gctgcgcagc	2160
cgggcgtcgg ccgtgtgcg gggccggggc ggcatggcct cagtcccgt gcctgcccag	2220
gaggtggagc agctcattgg tgagcgggtg gcggggcggg tgtgggtggc ggcggtaac	2280
ggccccgct ccaccgccgt ctgggggat gccgagcggt tggacgaggt gctggcgta	2340

tgtagccgga ccggggtgcg ggcccggcgg atcccggctg actatgcctc gactgcccc	2400
catgtgcagc ccctgcggga ggagtgtctg gagctgtctgg gggacatcag cccgcagccg	2460
tccggcgtgc cgttcttctc cacggtggag ggcacctggc tggacaccac aaccctggac	2520
gccgcctact ggtaccgcaa cctgcaccag cctgtccgtt tcagcgatgc cgtccaggcc	2580
ctggcggatg acggacaccg cgtcttctgc gaagtacgcc cccacccac cctcgtcccc	2640
gccatcgaag acaccaccga agacaccgcc gaagacgtca ccgcgatcgg cagcctccgc	2700
cgcggcgaca acgacaccg ccgcttctc accgccctcg cccacacca caccaccggc	2760
atcggcacac ccaccacctg gcaccaccac tacaccacc accacacca cccccacaac	2820
caccacctcg acctccccac ttatcccttc caacgccagc actactggct cgacgtccc	2880
acgggagcag gtgacgtcgc cgctgtctggc ttggagccgg ccgaacaccc tctgctcgcg	2940
gcaacagtcc aactgcaga cacggacggc tgcctactga cgggtgcct gtccttgcg	3000
tcgcatccgt ggctgggcga ttacagggtg gggggtgcgg tccgtctgc ggggtcggcg	3060
ttcgtggagc tggcggcca ggttggcgaa cgcgtgggct gcaccgaat cgagcaactc	3120
actgtcatg cgcgcgtggt ggttctctgt ggtgggggtg tgagtgtgca ggttggggtt	3180
gcggctcggc atggggaggc gcggcgtttg gtgagtgtgt atgcgcgggg tgggagtgt	3240
tgtggtgggg gtggtgcgtc ggggtggggtg tggacgtgtc atgcctcggg ggtgctggtt	3300
gaggctgctg ctggtgtggg tgtggtgtgt gatggtctgg cgggggtgtg gccgccgcgg	3360
ggtgcggtgg cgggtggaat cgatggtgtc cgtgaccgtt tggctggggc tggttgtgtt	3420
ttggggccgg tgttttcggg gctgcgtgcg gtgtggcgtg atggggggga ttgtctggt	3480
gagggtgtgtc tgccggagga ggcgtggggt gatgcgctg gttttgggt gcatccggcg	3540
ttgttgatg gtgtgttcca gccgttgtc gtgttcttc cgggtgggac ggggtttggg	3600
gagggggcgg ggttcgggga ggggtttcgg gtgccgctg tgtggggtgg tgtgtcgctt	3660
caccgggcgg gtgtgaccgg tgtgcgggtg cgtgtgtggg ctgtaggcg gggcggcggg	3720
cgtgaggcgg tgtcgtcgt ggtcggggat gaggcgggtg tgccggtggc gtcggtcgat	3780
cgtcttgagt tgcggcctgt ggatatgggt cagttgcgtg ctgtctcggt ttcggcgggg	3840
cggcggggtt cgtgtatgc ggtgcagtgg gctgaggtgg gtcctgtgcc ggtgtgtggg	3900
caggcgtggg cgtggcacga ggacgtgggt gagagcgtg gtgggcctgt gccgggggtg	3960
gtggtgttgc ggtgcccgga tgccggtgcc ggtggcggcg gtggcgggtg tgtgggtgag	4020
gttgttggtg ggggtgtggg tgtggtgcag ggggtggctgg ggctggagcg gtttgcgggt	4080

tcgcggctgg tggtagtgac ccgggggtgcg gtggtaggccc gccagaaga cggcccgggtg	4140
gatgtggtgg gtgcggcggt gtggggggctg gtgcggctcg cgcaggctga gcatccggac	4200
cggtttgtcc tctcgcacct cgacaccgac accgacaccg gcaccgacct cgacaccggt	4260
gctggtgctg gtgctggtgc tggtaggggc gtggatggtg ggcatgtggc ggccggtggtg	4320
gcgtgtggtg agccgcagtt ggccgtgctg gtgagcggg tgcaggccgc acgcctgacg	4380
cgacttgagt cgtccgttga tgtacctgct cagcgggtccg gtgatgttgc tggtagggag	4440
gtgttgccgt ggtgtcggg tgggtcgggt ttggtgacgg gtgggacggg tgtgctgggt	4500
gcggcgggtg cgcggcatct ggctggtgtg tgtgggtgct gggatctgct gttggtgagc	4560
cggcgtggtc cggatgctcc gggtagcggg ggtttgcggg cggagctggc cgcgttgggg	4620
gcggaggtgc ggattgttgc gtgtgatgtg ggggagcggc gggaggtggt ccggctgctg	4680
gagggtgttc ctccgggtg tccgtgacg ggtgtcgtgc atcgccgtgg tgtgctggac	4740
gatgcgacga tcgcctctct cagcccgag cggctgggca cgggtttcgc ggccaaggtg	4800
gatgccgtc ttttctgga tgagctgacg cggggtatgg agctgtcggc gttcgtgctg	4860
ttctctcgg ccgcggggat cctggggtcg gccgggcagg gcaactacgc cgcggccaat	4920
gccgtcttgg acgcgtggc gtaccggcgg cgggcggcgg gtctgccggg ggtgtcgtg	4980
gcgtgggggc tgtgggaaga ggccagcggg atgaccgggc acctggccgg caccgaccac	5040
cggcgcata tccgttccgg tctgcatcc atgtcgacc cggacgcact ggctctcttc	5100
gatgcggccc tggtctgga ccggccggtc ctgctgccg ccgacctgcg tcccggccc	5160
ccccctccgc cctgtctga ggacctctg ccgcaccac gccgcgcac caccgcacc	5220
accactaccg gtggtgcgga caacggcgcc cagctgcatg cccggctggc cggccagaca	5280
cacgaacaac agcacaccac cctctctgcc ctggtcgtc cccacatgc caccgtctc	5340
ggccacacca ccccgacac catcccccc gaccgcgct tccgcacct cggcttcgac	5400
tcctcaccg ccgtcgaact acgcaaccgg ctctcccga ccaccggact ccgcctcccc	5460
accacctctg cttcgacca ccccaacccc accacctca cccaccacct ccacacaaa	5520
cttctgggtc cggacagcac tgcctccatc ccagctcccc gtgctgcggc tgtgcctgca	5580
gaccaggacg agcccgctgc gatcattggc atggcgtgcc gctatcccgg aggcgtcacc	5640
tcagccgagg agctgtggga actgtctgca tcggggaggg acacggctcg cgagtttccg	5700
acggaccgtg ggtgggacct ggaagcactg ttcatccgg aaccgggtcg gccgggcacc	5760
tcgtacacc gctgtgggag ttctctctac gacgcggggg agttcgacgc cggcttcttc	5820

gggatcagtc cgctgaggc actggcgaig gacccgcagc agcgattgct gctggaggcc	5880
tcatgggagg ccatggagca ggcaggtatt gaccctacga ccgtacgcgg gagccagaca	5940
ggcgtgttcg cgggcctcat tccgcaggcc tatggacca ggctgcacga aaacgccga	6000
gccgacaccg agggctatgt cctgaccggc acatccggga gtgtggcctc cggctgtatc	6060
tcgtacacgt ttggttttga gggctctgcg gtgtcgggtg acacggcttg ttcctcgtcg	6120
ttggtggctt tacatctggc ctgtcaggcg ttgcgtcgg gtgagtgtc gatggcgctt	6180
gccgggggtg tgacggtgat gtcgtctccg ggtgccttcg tggagttttc gcggcagcgg	6240
ggctctggccg cggacgggca ttgcaaggcg ttctcggcgg cggcggacgg gaccggctgg	6300
ggtgagggtg tgggatgct gctggtggag cggctctccg acgcccgtcg caacggtcac	6360
cgtgtcctgg ccgtggtgcg tggcagtgcg gtcaaccagg acggtgcgag caacgggctg	6420
accgcgcccc acgggcctc ccagcagcgt gtcatccgcc aggccctcgc caacgccggc	6480
ttgtcggcgg gtgatgtcga tgcggtggag gccacaggca ccggcaccac ttgggagac	6540
ccgatcgagg ccaggccct ccttcgcacc tacgggcagg accgtgccgg cgagggggcg	6600
ctgtggctgg gtcggtcaa gtccaatgtc ggtcacacac aggtgccgc gggcgtcgcc	6660
ggggtgatca agatggtgat ggcgtgcgg aatggtctgc tgccgcggac gttgcatgtg	6720
gatgagccgt cgcgcgatgt ggactggtcc gcgggtgcgg tgcagctgct gacggagacg	6780
gtccctggc cggcgggga gggcgggcta cggcggcag gagtgtcatc attcggcgtc	6840
agcggacca acgcccagt catctcgaa gaagcaccg ccacaacat ccgctcagac	6900
acaccgccg acgacgttc ggggggacca ccgccggcg aggatgccgg tagtgccgag	6960
gaggctgctg ccggcagtc aggggtgtgg ccgtggctgg tgtcgccaa gtcgcagccg	7020
gccctgcgcg ccaggccca ggcctgcac gccacctca ccgaccacc cggcctcgac	7080
ctcgcgacg tcggatacac ctcgcccac gccgcgccc tgttcgacca ccgcgccacc	7140
ctcatcgccg ccgaccgca cacttcttg caagcactcc aggcactcgc cgcaggcgaa	7200
ccccacccg ccgtcatcca cagcagcgc ccaggcggga ccgggaccgg ggaggccga	7260
ggaaagaccg cattcatctg ctccggacag ggcaccaac gccccggcat ggcccacggc	7320
ctctaccaca cccacccgt ctccgccgc gactcaacg acatctgcac ccacctgac	7380
ccccacctg accacccct cctccctc ctacccagg acccaacac ccaggacacc	7440
accacctcg aagaagcggc cgactgtc cagcagacc cgtacgcca gccgcctc	7500
ttcgcttcc aggtcgccct ccaccgctc ctaccgacg gctaccacat cccccccac	7560
tactacgcg gacactcct cggcgaaac accgccccc acctcgccgg catctcacc	7620
ctaccgacg ccaccacct catcaccaa cgcgccacc tcatgcaaac catgcccc	7680

ggcaccatga ccacctcca caccaccccc caccacatca cccaccacat caccgcccac	7740
gaaaacgacc tcgcatcg cgccatcaac acccccacct ccctcgtcat cagcggcacc	7800
ccccacaccg tccaacacat caccacctc tgccaacaac aaggcatcaa aacaaaaacc	7860
ctccccacca accacgcctt ccactcccc cacaccaacc ccatcctcaa ccaactccac	7920
cagcacaccc aaacctcac ctaccacca ccccacaccc ccctcatcac cgccaacacc	7980
ccaccgacc aactcctcac ccccactac tggaccaaac aagcccgaa caccgtcgac	8040
atagccacca ccaccaaac cctccacca caggcgctca ccactacat cgaactcgga	8100
cccgacaaca cctcaccac cctcaccac cacaacctc ccaacacccc caccaccacc	8160
ctcacctca cccacccca ccaccaccc caaacccacc tcctcacaa cctcgccaaa	8220
accaccacca cctggcacc ccaccactac acccaccacc acaaccaacc ccacaccac	8280
accacctcg acctccccc ctacccttc caacaccacc actactggt cgaagcaca	8340
cagcccggtg ccggcaacgt gtcagcagc ggactcgacc ccaccgaaca cccctactc	8400
ggcgccacat tggaactggc cgaaggggac ggctgcctac tgacggggcg cctctcgtt	8460
cgcacgcac cctggctcg cgccatcg gtaggcggtg tcgtgctgt gccgggtacg	8520
gccttcgagg aactggccct tcatgccgga gaaagtgtg gttgcgacca cgtggacgag	8580
ctgacgtcc acacaccgtt ggtcattct gaggtcggag acgtgacct tcaggttgc	8640
attgcggcgc cggacgagtc gggtcgccgc atgatgacca tccactcac cggtagggc	8700
ggcagtggtg gagccgatgc gtcggccagt gcgtggacgc gtcatgccgc gggtgtgctg	8760
agccctgcca aggacgatga cactgcctcg tacgagctgc ttgcgggacc ctggcctccc	8820
gttgagcta cgctgtcga cctgaacacg gcttacgac aaatggccga cgccggcttt	8880
gcttatggcc tggcattcca agggttgcgc gcggcctggc gctacggcga cgacatcctc	8940
gtcgaggcac gtcttccga agaagtgtcg ggagacgagg cggcgtacgg tctgcacccg	9000
gccctgctcg acgtgcctt tcagggcacc ggctgtctt ctgtggcggg tccggggacg	9060
cccgctgtgc ccatgtgtg gaacggtctg cggttccgta cgcatggtgc agtctccgtg	9120
cgcgctgcc tgtcgacgt tggagcgaca ggggcggccg tgtgcgtgc catcaccgac	9180
gacaccgggg tgccggtggc gtcggtcgat cgtcttgagt tgcggcctgt ggatatgggt	9240
cagttcgtg ctgtctcggt ttcggcgggg cgccggggtt cgctgtatgc ggtgcagtgg	9300
gctgagggtg gtctgtgcc ggtgtgtggg caggcgtggg cgtggcacga ggacgtgggt	9360
gagagcgggt gtgggctgt gccgggggtg gtggtgttc ggtgcccgga tgccggtgcc	9420
gatggcggcg gtggcgggtg tgtgggtgag gttgttgtg ggggtgtggg tgtggtgcag	9480
gggtggctgg ggctggagcg gtttcgggt tcgcggctgg tgggtgtgac cgggggtgcg	9540

gtggtggccg gcccgaggga cggcccggtg gatgtggtgg gtgcggcggt gtgggggctg	9600
gtgcggtcgg cgcaggctga gcatccggac cgttttgtcc tcctcgacct ggacaccgac	9660
ctcgacacgg gcgctgacgc cgatcccggc aacgaggccg gtatggggtc tggctctggat	9720
ggtgggcgtg tggctgcggt ggtggcgtgt ggtgagccgc agttggcggt gcgtggtgag	9780
cgggtgctgg ccgcacgcct gacacgactt gagtcgccgg ttgatgtatc gggtcgggag	9840
gtgttgccgt ggttgtcggg tgggtcgggt ttggtgacgg gtgggacggg tgtgctgggt	9900
gcggcggtgg cgcggcatct ggctgggtgtg tgtggggtgc gggatctgtt gttggtgagc	9960
cggcgtggtc cggatgctcc ggggtcgggag ggtttgcggg cggagctggc cgcgttgggg	10020
gcggagggtgc ggattgttgc gtgtgatgtg ggggagcggc gggagggtgt cggctgctg	10080
gagggtgttc ctgccgggtg tccgctgacg ggtgtcgtgc atgcggctgg tgtgctggac	10140
gatgcgacga tcgctctct cagccccgag cggctgggca cggtgttcgc ggccaagggtg	10200
gatgccgttc ttttgcctga tgagctgacg cggggatagg agctgtcggc gttcgtgctg	10260
ttctcctcgg ccgcggggat cctggggctg gccgggcagg gcaactacgc cgcggccaat	10320
gccgctctgg acgcgtggc gtaccggcgg cgggcggcgg gtctgccggg ggtgtcgctg	10380
gcgtgggggc tgtgggaaga ggccagcggg atgaccgggc acctggccgg caccgaccac	10440
cggcgcatca tccgttccgg tctgcatccc atgtcgacce cggacgcaact ggctctcttc	10500
gatgcggccc tggctctgga ccggccggtc ctgctgcccg ccgacctgcg tcccgcgccg	10560
cccctgccgc cctgtctgca ggacctctg cccgccaccc gccgccgcac caccgcacc	10620
accactaccg gtggtgcgga caacggcgcc cagctgcatg cccggctggc cggccagaca	10680
cacgaacaac agcacaccac cctctctgcc ctggtccgct cccacatcgc caccgtctc	10740
ggccacaacg cgcggagat gatccccgtt gactcgcgct tccgcgacct aggtctcgac	10800
tccttgacag cggatggaact ccgtaaccgc ctgggtgagg caacgggact gcgactgccg	10860
accagtctgg tcttcgacca gccgaatgca gcgacctgg cgcgtcacct acgtcgtgag	10920
ctgatgggag acgacgcgga aggcgagacg ccatcgagg tcgcacttca tcaggttgcc	10980
gcggtgagc cgattgcgat tgtggggtg gcgtgtcgtt ttccgggtgg ggtgtgttcg	11040
ccggaggagt tgtgggagct ggttgcgtcg ggtggggtg cgattggtga atttcggcc	11100
ggtcgggggt gggatctgga ggggttgttt gattcggaac ctgaccggtc ggggacgtcg	11160
tacgcgcggt atggcgggtt tttgtatgag gcgggggagt tcgatgcgga cttcttcggg	11220
atcagtcgcg gtgaggcgtt ggcgatggat ccgcagcagc ggttgttgct ggagacgtcg	11280

tgggaggcgt tcgagcgggc gggatatcat ccgctgtcga tgcgtggctc ccgtacgggt	11340
gtcttcgccg gggtagatga ccacgactac gccgcgcgtc tccacatgt ccccgagggt	11400
ttcgaaggcc tcatcgccaa cggcagcgca ggcagcgtcg cgaccggccg ggtggcctac	11460
agctttggcc ttgagggtcc ggccgtgacc gtcgatacgg cgtgttcgtc gtcgttgggtg	11520
gcgttgcatc gggcggcgca ggcgttgctg gcgggtgagt gttcgatggc gcttgccggg	11580
ggtgtgacgg tgatgtcgtc tccgggtacg tttgtggagt tctcacgtca gcggggtctg	11640
gccgcggacg ggcggtgcaa ggcctattcg gcggctgctg acggtaccgg ctgggcccag	11700
ggtgtgggga tgctgctggt ggagcggtc tccgacgcc gtcgcaacgg tcaccgtgtc	11760
ctggccgtgg tgcgtggcag tgcggtcaac caggacggtg cgagcaacgg tctgaccg	11820
cccaacgggc cctcccagca gcgtgtcatc cgtcaggccc tggccaatgc gggactgacc	11880
ccggccgatg tcgacgcagt ggaggggccac ggcaccggga ccactctggg ggaccgatc	11940
gaggcccagg cactcctggc cgcctacgga caacaccgcc cccaccaccg ccccttgtgg	12000
ctgggatccc tcaaatcaa catcgggcac gcacaggccg ccgcgggctg gggcggagtc	12060
atcaagatgg tgatggccct gcgcaacggg ctgctgccac agaccctcca cgtggacgag	12120
ccccccccc aggtcgactg gtccacaggc gcagtacaac tcctgacaca accggtgccc	12180
tggcccgccg acccgccggg ccggccacgc cagccggcg tgtcatcatt cggcgtcagc	12240
ggcaccaacg ccatgtgat tttggaggag ggccttcggg cggcgggagg tgctgccgtg	12300
ggtgggggtg cgggtgggtg tccgaatca gcccttcggg tggctgagtc tgagccgggtg	12360
ccggtgccgg tgccggtgtc ggcgaggtct gaggccgggt tgcgggcgca ggcacaggcg	12420
ttgcgccagt acgtggcagc ccgccggac atgtcacctg ccgacatcgg tgcgggtctg	12480
gcccgcggcc gggccgtact ggaacaccgc gccgtcatcc tggccgcgga ccgcgaggaa	12540
ctggcgcagg cactgacagc cctggcagcc gggaacccc accccacat caccacaggc	12600
cacaccgggg gcagtgaccg cggcggcgctc gtcttcgtct tcccggaca gggcgccag	12660
tgggcgggga tgggcctgac cctgtcacc tctcaccg tgttcgccga acacatcgac	12720
gcatgcgaga aagccctcac cccctgggtg ccttggtccc tgaccgacat cctgcaccgc	12780
gaccccgacg accccgatg gcaacaagc gacgtggtcc agcccgctgt cttcagcatc	12840
atggtctccc tcgccccct gtggcgctcc tacggcatcg aaccgacgc ggtcctcggc	12900
cactcccagg gagaaatgc cgcccccac atctcgggc cactcagcct gaaagacgcc	12960
gccaaaaccg ttgactgcg cagccaggca ctggccgccg tacgaggccg gggcgccatg	13020
gtctcactgc cctgcgccg ccaggacgtg cagcagctca tttccgaac gtgggaaggg	13080
cagttgtggg tggcagccct caacggcccc cactccacca ccgtctccgg cgacaccacc	13140

gcagtagaag aactcctcac ccactgtgcc gacaccggcc tacgggccaa acgcatcccc	13200
gtcgactacg cctccactg cccccactg caaccctcc acgacgaact cctgcacctg	13260
ctgggagaca tcacccccca gccgtccacc atgccgttct tctccaccgt cgtagggcac	13320
ctggtctggt acaccacaac cctggacgcc gcctactggt accgcaacct ccaccagccc	13380
gtccgcttca gccacgcat ccagaccctg accgacgacg gacaccgccc cttcacgaa	13440
atcagtcccc accccacct cgtecccgcc atcgaagaca ccaccgaaaa caccaccgaa	13500
aacatcaccg cgaccggcag cctccgccgc ggcgacaacg acaccaccg ctccctcacc	13560
gccctcgccc acaccacac caccggcatt cggacacca ccacctggca ccaccactac	13620
acccaaacc accccaccc ccacaaccac cactcgacc tgcccaccta ccccttcaa	13680
caccagcact actggctcca accaccacc acgacaaccg acctcaccac caccggcctc	13740
accccaacc accacccct cctcaccgca aactcacc tcgccaacaa caacacaaa	13800
ctactcaccg gccgcctctc cctacgcacc caccctggc tcaccgacca caccgtcgtc	13860
ggtaccactc ttgtgccagg aaccgccctc ctggaactcg ccctccaagc aaccacgacc	13920
gaccacctcg aagaactcg cctccacag cctctcgtca tccccgtga ggggtgccgtc	13980
gacgttcagg tgcacatcaa tcaccggac gacaccgaca ctcgttcact gacgatctac	14040
tcggaagcg agaagcccc cgcagcggct cctggcgtc atcacccac gccggttctg	14100
ggaaccaaga cctcgcgat tgagacaggc cgtagccacg atgatctgtc gatgtggccg	14160
ccagcgggcg cagttcgctg tgctgatgag gaattggcag ccttgtatgg cgactacgag	14220
gcaaattggt ttgtctatgg ccccgattc cgggggctga ctgtgcctg gcgtctggga	14280
gacgaggtgt ttgccgaggt tcgccttcca gaacaggtgc acggcgaggc atccgcgtac	14340
aacctgcacc cggcactgct ggatgctgcc ttgcacgag cggcctttgc gccgtcgggc	14400
agttcgccgc agggatccgt accgttctcc ttaccgggtg tgacgctgca ccccaccaat	14460
gcgtcgtcgt tgcgcgtgcg actctgcgg gccgatccga acagcgcca cggcgagtt	14520
tccgtgctgg tcacggatga caccgtacg cccgtggcgt ccgtcgaggc gttggcggtg	14580
cggcgttgg cggcgagca attgcgagct gccgagcgg ccgtacagcg cgctgagctc	14640
ttcgacatga agtgggttga ggtgccctca gatgtactgg tgtcgggcgg ggcacggtg	14700
gtggtgctgg atggtgccga cgacctggt ggtctggcgg ctgaggagga tgggtgtccg	14760
ggggtggtgg tgttgcggtg cccgatgcc ggtgccgatg gcggcggtgg tggcggtggt	14820
gtgggtgagg ttgttgggtg ggtgttgggt gtggtgcagg ggtggctggg gctggagcgg	14880
tttgcgggtt cgcggctggt ggtggtgacc cggggtgcgg tgggtggcgg cccggaggac	14940
ggcccggtgg atggcccggt ggatgtggtg ggtgcggcgg tgtgggggct ggtgcggtcg	15000

gcgagggctg agcatccgga ccggtttgtc ctccctcgacc tggacaccga cctcgacagc 15060

ggcgctgacc gcgatgccgg caacgaggcc ggtatggggc ctggctctga tggtagggcgt 15120

gtggctgcgg tggtagggctg tggtagccg cagttggcgg tgcgtggga gcgggtgctg 15180

gccgcacgcc tgacacgact tgagtcgccg gttgatgtat cgggtcgga ggtgttgccg 15240

tggttgtcgg gtgggtcggg gttggtgacg ggtgggacgg gtgtgctggg tgcggcggtg 15300

gcgcggcatc tggctgggtg gtgtgggggtg cgggatctgt tgttggtag ccggcgtggt 15360

ccggatgctc cgggtcgga gggtttgcgg gcggagctgg ccgcgttggg ggcggaggtg 15420

cggattgttg cgtgtgatgt gggggagcgg cgggaggtgg tccggctgct ggagggtgtt 15480

ctgcccgggt gtccgctgac ggggtgtcgt catgcggctg gtgtgctgga cgatgcgacg 15540

atgcctctc tcacgccga gcggctgggc acggtgttcg cggccaaggt ggatgccgt 15600

cttttctgg atgagctgac gcgggggatg gagctgtcgg cgttcgtgct gttctcctcg 15660

gccgcgggga tcttggggtc ggccgggcag ggcaactacg ccgcggccaa tgccgctctg 15720

gacgcgtcgg cgtaccggcg gcgggcggcg ggtctgccgg ggggtgtcgt ggcgtggggg 15780

ctgtgggaag agccacggg gatgaccggg catctggcgg gcaccgacca ccggcgcac 15840

atccgttcg gtctgcatcc catgtcgacc ccggacgcac tggccctctt cgatgcggcc 15900

ctggctctgg accggccgggt cctgtgtccc gccgacctgc gtcccgcgcc gccctgccg 15960

cccctgtgc aggacctct ccccgccacc cgcgcgccga ccaccgcac caccactacc 16020

ggtggtgcgg acaacggcgc ccagctgcac ggccggctgg ccggccagac acacgaacaa 16080

cagcacacca cctcctcgc cctggtccgc tcccacatcg ccaccgtctt gggccacacc 16140

accccgaca ccatccccc cgaccgcgcg ttccgcgacc tcggcttcga ctccctcacc 16200

gccgtgaac tacgaaccg gctctccac accaccggac tccgcctccc caccacctc 16260

gccttcgacc accccaacc caccacctc accaccacc tccacacaca actcgtcagc 16320

aagggactca ccgccgggc cgagccggac gccgaacga cccccggg gctgccctcg 16380

ctgtctcgg agctcgagcg gctggaggcg gtagtgtct cctccaccac atcctccgt 16440

gccccgtgg acgacggcgc gcgcacgcgg ctggcctccc gactgcattc cctcgcccag 16500

aagtgaacg gcgacgacac ccccccgac ctgcagaga catcgacga ggagatgtt 16560

gtctcatcg acagggaagt cggattcgaa tctcaatga 16599

<210> 72

<211> 5532

<212> PRT

<213> Artificial Sequence

<220><223> type I polyketide synthase AVES 3 (BAA84478.1)

<400> 72

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

1 5 10 15

Lys Ala Asn Gln Thr Leu Arg Ala Arg Asn Glu Gln Leu Ala Ala Ala

20 25 30

Met Glu Ala Ser Ser Glu Pro Ile Ala Ile Val Gly Met Ala Cys Arg

35 40 45

Phe Pro Gly Gly Val Cys Ser Pro Glu Glu Leu Trp Glu Leu Val Ala

50 55 60

Ser Gly Gly Asp Ala Ile Gly Glu Phe Pro Ala Gly Arg Gly Trp Asp

65 70 75 80

Leu Glu Gly Leu Phe Asp Ser Asp Pro Asp Arg Ser Gly Thr Ser Tyr

85 90 95

Ala Arg Tyr Gly Gly Phe Leu Tyr Glu Ala Gly Glu Phe Asp Ala Asp

100 105 110

Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln

115 120 125

Arg Leu Leu Leu Glu Thr Ser Trp Glu Ala Phe Glu Arg Ala Gly Ile

130 135 140

Asp Pro Leu Ser Met Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Val

145 150 155 160

Met Tyr His Asp Tyr Gly Ser Arg Leu Gly Thr Ile Pro Glu Gly Phe

165 170 175

Glu Gly Tyr Ile Gly Asn Gly Ser Gly Gly Ala Val Ala Ser Gly Arg

180 185 190

Val Ala Tyr Thr Leu Gly Leu Glu Gly Pro Ala Val Ser Val Asp Thr

195 200 205

Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Cys Gln Ser Leu

210 215 220

Arg Ser Gly Glu Cys Thr Leu Ala Leu Ala Gly Gly Val Thr Val Met

225 230 235 240
 Ser Thr Pro His Leu Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ser
 245 250 255
 Val Asp Gly Arg Cys Lys Ser Phe Ala Gly Gly Ala Asp Gly Thr Gly
 260 265 270
 Met Gly Glu Gly Val Gly Met Leu Leu Val Glu Arg Leu Ser Asp Ala
 275 280 285

 Val Arg Leu Gly His Arg Val Leu Ala Val Leu Arg Gly Ser Ala Val
 290 295 300
 Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ala
 305 310 315 320
 Gln Glu Arg Val Ile Arg Gln Ala Leu Ala Asn Ala Gly Leu Ser Val
 325 330 335
 Ala Asp Val Asp Val Val Glu Gly His Gly Thr Gly Thr Thr Leu Gly
 340 345 350
 Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Arg Ala

 355 360 365
 Gly Asn Arg Pro Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His
 370 375 380
 Ala Gln Ala Ala Ala Gly Val Gly Gly Val Ile Lys Met Val Met Ala
 385 390 395 400
 Leu Arg Glu Gly Val Leu Pro Arg Thr Leu His Val Asp Glu Pro Ser
 405 410 415
 Pro Gln Val Asp Trp Ser Ala Gly Ala Val Arg Leu Leu Thr Glu Ala
 420 425 430

 Val Pro Trp Pro Gly Asp Ala Ala Gly Arg Leu Arg Arg Ala Gly Val
 435 440 445
 Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Val Ile Leu Glu Glu
 450 455 460
 Ala Pro Ala Ala Gly Gly Cys Val Ala Gly Gly Gly Val Leu Glu Gly
 465 470 475 480
 Ala Pro Gly Leu Ala Ile Ser Val Ala Glu Ser Val Ala Ala Pro Val

485 490 495
Ala Val Ser Ala Pro Val Ala Glu Ser Val Pro Val Pro Val Pro Val

500 505 510
Pro Val Pro Val Pro Val Ser Ala Arg Ser Glu Ala Gly Leu Arg Ala

515 520 525
Gln Ala Glu Ala Leu Arg Gln Tyr Val Ala Val Arg Pro Asp Val Ser

530 535 540
Leu Ala Asp Val Gly Ala Gly Leu Ala Cys Gly Arg Ala Val Leu Glu

545 550 555 560
His Arg Ala Val Val Leu Ala Ala Asp Arg Glu Glu Leu Val Gln Gly

565 570 575
Leu Gly Ala Leu Ala Ala Gly Glu Pro Asp Arg Arg Val Thr Thr Gly

580 585 590
His Ala Pro Gly Gly Asp Arg Gly Gly Val Val Phe Val Phe Pro Gly

595 600 605
Gln Gly Gly Gln Trp Ala Gly Met Gly Val Arg Leu Leu Ala Ser Ser

610 615 620
Pro Val Phe Ala Arg Arg Met Gln Ala Cys Glu Glu Ala Leu Ala Pro

625 630 635 640
Trp Val Asp Trp Ser Val Val Asp Ile Leu Arg Arg Asp Ala Gly Asp

645 650 655
Ala Val Trp Glu Arg Ala Asp Val Val Gln Pro Val Leu Phe Ser Val

660 665 670
Met Val Ser Leu Ala Ala Leu Trp Arg Ser Tyr Gly Ile Glu Pro Asp

675 680 685
Ala Val Leu Gly His Ser Gln Gly Glu Ile Ala Ala Ala His Val Cys

690 695 700
Gly Ala Leu Ser Leu Lys Asp Ala Ala Lys Thr Val Ala Leu Arg Ser

705 710 715 720
Arg Ala Leu Ala Ala Val Arg Gly Arg Gly Gly Met Ala Ser Val Pro

725 730 735
Arg Ala Leu Ala Ala Val Arg Gly Arg Gly Gly Met Ala Ser Val Pro

Leu Pro Ala Gln Glu Val Glu Gln Leu Ile Gly Glu Arg Trp Ala Gly
 740 745 750
 Arg Leu Trp Val Ala Ala Val Asn Gly Pro Arg Ser Thr Ala Val Ser
 755 760 765
 Gly Asp Ala Glu Ala Val Asp Glu Val Leu Ala Tyr Cys Ala Gly Thr
 770 775 780
 Gly Val Arg Ala Arg Arg Ile Pro Val Asp Tyr Ala Ser His Cys Pro

 785 790 795 800
 His Val Gln Pro Leu Arg Glu Glu Leu Leu Glu Leu Leu Gly Asp Ile
 805 810 815
 Ser Pro Gln Pro Ser Gly Val Pro Phe Phe Ser Thr Val Glu Gly Thr
 820 825 830
 Trp Leu Asp Thr Thr Thr Leu Asp Ala Ala Tyr Trp Tyr Arg Asn Leu
 835 840 845
 His Gln Pro Val Arg Phe Ser Asp Ala Val Gln Ala Leu Ala Asp Asp
 850 855 860

 Gly His Arg Val Phe Val Glu Val Ser Pro His Pro Thr Leu Val Pro
 865 870 875 880
 Ala Ile Glu Asp Thr Thr Glu Asp Thr Ala Glu Asp Val Thr Ala Ile
 885 890 895
 Gly Ser Leu Arg Arg Gly Asp Asn Asp Thr Arg Arg Phe Leu Thr Ala
 900 905 910
 Leu Ala His Thr His Thr Thr Gly Ile Gly Thr Pro Thr Thr Trp His
 915 920 925
 His His Tyr Thr His His His Thr His Pro His Asn His His Leu Asp

 930 935 940
 Leu Pro Thr Tyr Pro Phe Gln Arg Gln His Tyr Trp Leu Asp Ala Pro
 945 950 955 960
 Thr Gly Ala Gly Asp Val Ala Ala Ala Gly Leu Glu Pro Ala Glu His
 965 970 975
 Pro Leu Leu Ala Ala Thr Val Gln Leu Ala Asp Thr Asp Gly Cys Leu
 980 985 990

Leu Thr Gly Arg Leu Ser Leu Arg Ser His Pro Trp Leu Gly Asp Tyr

995 1000 1005

Glu Val Gly Gly Ala Val Leu Leu Ser Gly Ser Ala Phe Val Glu Leu

1010 1015 1020

Ala Val Gln Val Gly Glu Arg Val Gly Cys Thr Arg Ile Glu Gln Leu

1025 1030 1035 1040

Thr Val His Ala Pro Leu Val Val Pro Val Gly Gly Gly Val Ser Val

1045 1050 1055

Gln Val Gly Val Ala Ala Ala Asp Gly Glu Gly Arg Arg Leu Val Ser

1060 1065 1070

Val Tyr Ala Arg Gly Gly Ser Ala Cys Gly Gly Gly Gly Ala Ser Gly

1075 1080 1085

Gly Val Trp Thr Cys His Ala Ser Gly Val Leu Val Glu Ala Ala Ala

1090 1095 1100

Gly Gly Gly Val Val Val Asp Gly Leu Ala Gly Val Trp Pro Pro Arg

1105 1110 1115 1120

Gly Ala Val Ala Val Asp Val Asp Gly Val Arg Asp Arg Leu Ala Gly

1125 1130 1135

Ala Gly Cys Val Leu Gly Pro Val Phe Ser Gly Leu Arg Ala Val Trp

1140 1145 1150

Arg Asp Gly Gly Asp Leu Leu Ala Glu Val Cys Leu Pro Glu Glu Ala

1155 1160 1165

Trp Gly Asp Ala Ala Gly Phe Gly Leu His Pro Ala Leu Leu Asp Gly

1170 1175 1180

Val Val Gln Pro Leu Ser Val Leu Leu Pro Gly Gly Thr Gly Phe Gly

1185 1190 1195 1200

Glu Gly Ala Gly Phe Gly Glu Gly Val Arg Val Pro Ala Val Trp Gly

1205 1210 1215

Gly Val Ser Leu His Arg Ala Gly Val Thr Gly Val Arg Val Arg Val

1220 1225 1230

Trp Ala Val Gly Arg Gly Gly Gly Arg Glu Ala Val Ser Val Val Val

1235 1240 1245
 Gly Asp Glu Ala Gly Val Pro Val Ala Ser Val Asp Arg Leu Glu Leu
 1250 1255 1260
 Arg Pro Val Asp Met Gly Gln Leu Arg Ala Val Ser Val Ser Ala Gly
 1265 1270 1275 1280
 Arg Arg Gly Ser Leu Tyr Ala Val Gln Trp Ala Glu Val Gly Pro Val
 1285 1290 1295

 Pro Val Cys Gly Gln Ala Trp Ala Trp His Glu Asp Val Gly Glu Ser
 1300 1305 1310
 Gly Gly Gly Pro Val Pro Gly Val Val Val Leu Arg Cys Pro Asp Ala
 1315 1320 1325
 Gly Ala Gly Gly Gly Gly Gly Gly Gly Val Gly Glu Val Val Gly Gly
 1330 1335 1340
 Val Leu Gly Val Val Gln Gly Trp Leu Gly Leu Glu Arg Phe Ala Gly
 1345 1350 1355 1360
 Ser Arg Leu Val Val Val Thr Arg Gly Ala Val Val Ala Gly Gln Glu

 1365 1370 1375
 Asp Gly Pro Val Asp Val Val Gly Ala Ala Val Trp Gly Leu Val Arg
 1380 1385 1390
 Ser Ala Gln Ala Glu His Pro Asp Arg Phe Val Leu Leu Asp Leu Asp
 1395 1400 1405
 Thr Asp Thr Asp Thr Gly Thr Asp Leu Asp Thr Gly Ala Gly Ala Gly
 1410 1415 1420
 Ala Gly Ala Gly Trp Gly Val Asp Gly Gly His Val Ala Ala Val Val
 1425 1430 1435 1440

 Ala Cys Gly Glu Pro Gln Leu Ala Val Arg Gly Glu Arg Val Leu Ala
 1445 1450 1455
 Ala Arg Leu Thr Arg Leu Glu Ser Ser Val Asp Val Pro Ala Gln Arg
 1460 1465 1470
 Ser Gly Asp Val Ala Gly Arg Glu Val Leu Pro Trp Leu Ser Gly Gly
 1475 1480 1485
 Ser Val Leu Val Thr Gly Gly Thr Gly Val Leu Gly Ala Ala Val Ala

1490 1495 1500
 Arg His Leu Ala Gly Val Cys Gly Val Arg Asp Leu Leu Leu Val Ser

 1505 1510 1515 1520
 Arg Arg Gly Pro Asp Ala Pro Gly Ala Glu Gly Leu Arg Ala Glu Leu

 1525 1530 1535
 Ala Ala Leu Gly Ala Glu Val Arg Ile Val Ala Cys Asp Val Gly Glu

 1540 1545 1550
 Arg Arg Glu Val Val Arg Leu Leu Glu Gly Val Pro Ala Gly Cys Pro

 1555 1560 1565
 Leu Thr Gly Val Val His Ala Ala Gly Val Leu Asp Asp Ala Thr Ile

 1570 1575 1580

 Ala Ser Leu Thr Pro Glu Arg Leu Gly Thr Val Phe Ala Ala Lys Val
 1585 1590 1595 1600
 Asp Ala Ala Leu Leu Leu Asp Glu Leu Thr Arg Gly Met Glu Leu Ser

 1605 1610 1615
 Ala Phe Val Leu Phe Ser Ser Ala Ala Gly Ile Leu Gly Ser Ala Gly

 1620 1625 1630
 Gln Gly Asn Tyr Ala Ala Ala Asn Ala Ala Leu Asp Ala Leu Ala Tyr

 1635 1640 1645
 Arg Arg Arg Ala Ala Gly Leu Pro Gly Val Ser Leu Ala Trp Gly Leu

 1650 1655 1660
 Trp Glu Glu Ala Ser Gly Met Thr Gly His Leu Ala Gly Thr Asp His
 1665 1670 1675 1680
 Arg Arg Ile Ile Arg Ser Gly Leu His Pro Met Ser Thr Pro Asp Ala

 1685 1690 1695
 Leu Ala Leu Phe Asp Ala Ala Leu Ala Leu Asp Arg Pro Val Leu Leu

 1700 1705 1710
 Pro Ala Asp Leu Arg Pro Ala Pro Pro Leu Pro Pro Leu Leu Gln Asp

 1715 1720 1725

 Leu Leu Pro Ala Thr Arg Arg Arg Thr Thr Arg Thr Thr Thr Thr Gly

 1730 1735 1740

Gly Ala Asp Asn Gly Ala Gln Leu His Ala Arg Leu Ala Gly Gln Thr
 1745 1750 1755 1760
 His Glu Gln Gln His Thr Thr Leu Leu Ala Leu Val Arg Ser His Ile
 1765 1770 1775
 Ala Thr Val Leu Gly His Thr Thr Pro Asp Thr Ile Pro Pro Asp Arg
 1780 1785 1790
 Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Thr Ala Val Glu Leu Arg
 1795 1800 1805
 Asn Arg Leu Ser Arg Thr Thr Gly Leu Arg Leu Pro Thr Thr Leu Ala
 1810 1815 1820
 Phe Asp His Pro Asn Pro Thr Thr Leu Thr His His Leu His Thr Gln
 1825 1830 1835 1840
 Leu Leu Gly Ser Asp Ser Thr Ala Ser Ile Pro Ala Pro Arg Ala Ala
 1845 1850 1855
 Ala Val Pro Ala Asp Gln Asp Glu Pro Val Ala Ile Ile Gly Met Ala
 1860 1865 1870

 Cys Arg Tyr Pro Gly Gly Val Thr Ser Ala Glu Glu Leu Trp Glu Leu
 1875 1880 1885
 Leu Ala Ser Gly Arg Asp Thr Val Gly Glu Phe Pro Thr Asp Arg Gly
 1890 1895 1900
 Trp Asp Leu Glu Ala Leu Phe Asp Pro Glu Pro Gly Arg Pro Gly Thr
 1905 1910 1915 1920
 Ser Tyr Thr Arg Cys Gly Ser Phe Leu Tyr Asp Ala Gly Glu Phe Asp
 1925 1930 1935
 Ala Gly Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro
 1940 1945 1950
 Gln Gln Arg Leu Leu Leu Glu Ala Ser Trp Glu Ala Met Glu Gln Ala
 1955 1960 1965
 Gly Ile Asp Pro Thr Thr Val Arg Gly Ser Gln Thr Gly Val Phe Ala
 1970 1975 1980
 Gly Leu Ile Pro Gln Ala Tyr Gly Pro Arg Leu His Glu Asn Ala Ala
 1985 1990 1995 2000

Ala Asp Thr Glu Gly Tyr Val Leu Thr Gly Thr Ser Gly Ser Val Ala
2005 2010 2015

Ser Gly Arg Ile Ser Tyr Thr Phe Gly Phe Glu Gly Pro Ala Val Ser
2020 2025 2030

Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Cys
2035 2040 2045

Gln Ala Leu Arg Ala Gly Glu Cys Ser Met Ala Leu Ala Gly Gly Val
2050 2055 2060

Thr Val Met Ser Ser Pro Gly Ala Phe Val Glu Phe Ser Arg Gln Arg
2065 2070 2075 2080

Gly Leu Ala Ala Asp Gly His Cys Lys Ala Phe Ser Ala Ala Ala Asp
2085 2090 2095

Gly Thr Gly Trp Gly Glu Gly Val Gly Met Leu Leu Val Glu Arg Leu
2100 2105 2110

Ser Asp Ala Arg Arg Asn Gly His Arg Val Leu Ala Val Val Arg Gly
2115 2120 2125

Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn
2130 2135 2140

Gly Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Asn Ala Gly
2145 2150 2155 2160

Leu Ser Ala Gly Asp Val Asp Ala Val Glu Ala His Gly Thr Gly Thr
2165 2170 2175

Thr Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly
2180 2185 2190

Gln Asp Arg Ala Gly Glu Gly Pro Leu Trp Leu Gly Ser Val Lys Ser
2195 2200 2205

Asn Val Gly His Thr Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys
2210 2215 2220

Met Val Met Ala Leu Arg Asn Gly Leu Leu Pro Arg Thr Leu His Val
2225 2230 2235 2240

Asp Glu Pro Ser Pro His Val Asp Trp Ser Ala Gly Ala Val Gln Leu

2245 2250 2255
 Leu Thr Glu Thr Val Pro Trp Pro Gly Gly Glu Gly Arg Leu Arg Arg
 2260 2265 2270
 Ala Gly Val Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Val Ile
 2275 2280 2285
 Leu Glu Glu Ala Pro Ala His Asn Ile Pro Ser Asp Thr Pro Ala Asp
 2290 2295 2300

 Asp Val Pro Gly Gly Pro Pro Ala Gly Glu Asp Ala Gly Ser Gly Glu
 2305 2310 2315 2320
 Glu Ala Ala Ala Gly Ser Pro Gly Val Trp Pro Trp Leu Val Ser Ala
 2325 2330 2335
 Lys Ser Gln Pro Ala Leu Arg Ala Gln Ala Gln Ala Leu His Ala His
 2340 2345 2350
 Leu Thr Asp His Pro Gly Leu Asp Leu Ala Asp Val Gly Tyr Thr Leu
 2355 2360 2365
 Ala His Ala Arg Ala Val Phe Asp His Arg Ala Thr Leu Ile Ala Ala

 2370 2375 2380
 Asp Arg Asp Thr Phe Leu Gln Ala Leu Gln Ala Leu Ala Ala Gly Glu
 2385 2390 2395 2400
 Pro His Pro Ala Val Ile His Ser Ser Ala Pro Gly Gly Thr Gly Thr
 2405 2410 2415
 Gly Glu Ala Ala Gly Lys Thr Ala Phe Ile Cys Ser Gly Gln Gly Thr
 2420 2425 2430
 Gln Arg Pro Gly Met Ala His Gly Leu Tyr His Thr His Pro Val Phe
 2435 2440 2445

 Ala Ala Ala Leu Asn Asp Ile Cys Thr His Leu Asp Pro His Leu Asp
 2450 2455 2460
 His Pro Leu Leu Pro Leu Leu Thr Gln Asp Pro Asn Thr Gln Asp Thr
 2465 2470 2475 2480
 Thr Thr Leu Glu Glu Ala Ala Ala Leu Leu Gln Gln Thr Pro Tyr Ala
 2485 2490 2495
 Gln Pro Ala Leu Phe Ala Phe Gln Val Ala Leu His Arg Leu Leu Thr

2500 2505 2510
Asp Gly Tyr His Ile Thr Pro His Tyr Tyr Ala Gly His Ser Leu Gly

2515 2520 2525
Glu Ile Thr Ala Ala His Leu Ala Gly Ile Leu Thr Leu Thr Asp Ala

2530 2535 2540
Thr Thr Leu Ile Thr Gln Arg Ala Thr Leu Met Gln Thr Met Pro Pro

2545 2550 2555 2560
Gly Thr Met Thr Thr Leu His Thr Thr Pro His His Ile Thr His His

2565 2570 2575
Ile Thr Ala His Glu Asn Asp Leu Ala Ile Ala Ala Ile Asn Thr Pro

2580 2585 2590

Thr Ser Leu Val Ile Ser Gly Thr Pro His Thr Val Gln His Ile Thr

2595 2600 2605

Thr Leu Cys Gln Gln Gln Gly Ile Lys Thr Lys Thr Leu Pro Thr Asn

2610 2615 2620

His Ala Phe His Ser Pro His Thr Asn Pro Ile Leu Asn Gln Leu His

2625 2630 2635 2640

Gln His Thr Gln Thr Leu Thr Tyr His Pro Pro His Thr Pro Leu Ile

2645 2650 2655

Thr Ala Asn Thr Pro Pro Asp Gln Leu Leu Thr Pro His Tyr Trp Thr

2660 2665 2670

Gln Gln Ala Arg Asn Thr Val Asp Ile Ala Thr Thr Thr Gln Thr Leu

2675 2680 2685

His Gln His Gly Val Thr Thr Tyr Ile Glu Leu Gly Pro Asp Asn Thr

2690 2695 2700

Leu Thr Thr Leu Thr His His Asn Leu Pro Asn Thr Pro Thr Thr Thr

2705 2710 2715 2720

Leu Thr Leu Thr His Pro His His His Pro Gln Thr His Leu Leu Thr

2725 2730 2735

Asn Leu Ala Lys Thr Thr Thr Thr Trp His Pro His His Tyr Thr His

2740 2745 2750

His His Asn Gln Pro His Thr His Thr His Leu Asp Leu Pro Thr Tyr
2755 2760 2765

Pro Phe Gln His His His Tyr Trp Leu Glu Ser Thr Gln Pro Gly Ala
2770 2775 2780

Gly Asn Val Ser Ala Ala Gly Leu Asp Pro Thr Glu His Pro Leu Leu
2785 2790 2795 2800

Gly Ala Thr Leu Glu Leu Ala Glu Gly Asp Gly Cys Leu Leu Thr Gly
2805 2810 2815

Arg Leu Ser Leu Arg Thr His Pro Trp Leu Ala Gly His Ala Val Gly
2820 2825 2830

Gly Val Val Leu Leu Pro Gly Thr Ala Phe Ala Glu Leu Ala Leu His
2835 2840 2845

Ala Gly Glu Ser Val Gly Cys Asp His Val Asp Glu Leu Thr Leu His
2850 2855 2860

Thr Pro Leu Val Ile Pro Glu Val Gly Asp Val Thr Leu Gln Val Ala
2865 2870 2875 2880

Ile Ala Ala Pro Asp Glu Ser Gly Arg Arg Met Met Thr Ile His Ser
2885 2890 2895

Arg Gly Glu Gly Gly Ser Gly Gly Ala Asp Ala Ser Ala Ser Ala Trp
2900 2905 2910

Thr Arg His Ala Ala Gly Val Leu Ser Pro Ala Lys Asp Asp Asp Thr
2915 2920 2925

Ala Ser Tyr Glu Leu Leu Ala Gly Pro Trp Pro Pro Val Gly Ala Thr
2930 2935 2940

Pro Val Asp Leu Asn Thr Ala Tyr Asp Gln Met Ala Asp Ala Gly Phe
2945 2950 2955 2960

Ala Tyr Gly Leu Ala Phe Gln Gly Leu Arg Ala Ala Trp Arg Tyr Gly
2965 2970 2975

Asp Asp Ile Leu Val Glu Ala Arg Leu Pro Glu Glu Val Ser Gly Asp
2980 2985 2990

Ala Ala Ala Tyr Gly Leu His Pro Ala Leu Leu Asp Ala Ala Leu Gln
2995 3000 3005

Gly Thr Gly Leu Leu Ser Val Ala Gly Pro Gly Thr Pro Val Val Pro
3010 3015 3020

His Val Trp Asn Gly Leu Arg Phe Arg Thr His Gly Ala Val Ser Val
3025 3030 3035 3040

Arg Ala Cys Leu Ser Thr Leu Gly Ala Thr Gly Ala Ala Val Cys Val
3045 3050 3055

Arg Ile Thr Asp Asp Thr Gly Val Pro Val Ala Ser Val Asp Arg Leu
3060 3065 3070

Glu Leu Arg Pro Val Asp Met Gly Gln Leu Arg Ala Val Ser Val Ser
3075 3080 3085

Ala Gly Arg Arg Gly Ser Leu Tyr Ala Val Gln Trp Ala Glu Val Gly
3090 3095 3100

Pro Val Pro Val Cys Gly Gln Ala Trp Ala Trp His Glu Asp Val Gly
3105 3110 3115 3120

Glu Ser Gly Gly Gly Pro Val Pro Gly Val Val Val Leu Arg Cys Pro
3125 3130 3135

Asp Ala Gly Ala Asp Gly Gly Gly Gly Gly Val Gly Glu Val Val
3140 3145 3150

Gly Gly Val Leu Gly Val Val Gln Gly Trp Leu Gly Leu Glu Arg Phe
3155 3160 3165

Ala Gly Ser Arg Leu Val Val Val Thr Arg Gly Ala Val Val Ala Gly
3170 3175 3180

Pro Glu Asp Gly Pro Val Asp Val Val Gly Ala Ala Val Trp Gly Leu
3185 3190 3195 3200

Val Arg Ser Ala Gln Ala Glu His Pro Asp Arg Phe Val Leu Leu Asp
3205 3210 3215

Leu Asp Thr Asp Leu Asp Ser Gly Ala Asp Ala Asp Ala Gly Asn Glu
3220 3225 3230

Ala Gly Met Gly Ser Gly Leu Asp Gly Gly Arg Val Ala Ala Val Val
3235 3240 3245

Ala Cys Gly Glu Pro Gln Leu Ala Val Arg Gly Glu Arg Val Leu Ala

3250 3255 3260
 Ala Arg Leu Thr Arg Leu Glu Ser Pro Val Asp Val Ser Gly Arg Glu
 3265 3270 3275 3280
 Val Leu Pro Trp Leu Ser Gly Gly Ser Val Leu Val Thr Gly Gly Thr
 3285 3290 3295
 Gly Val Leu Gly Ala Ala Val Ala Arg His Leu Ala Gly Val Cys Gly
 3300 3305 3310

 Val Arg Asp Leu Leu Leu Val Ser Arg Arg Gly Pro Asp Ala Pro Gly
 3315 3320 3325
 Ala Glu Gly Leu Arg Ala Glu Leu Ala Ala Leu Gly Ala Glu Val Arg
 3330 3335 3340
 Ile Val Ala Cys Asp Val Gly Glu Arg Arg Glu Val Val Arg Leu Leu
 3345 3350 3355 3360
 Glu Gly Val Pro Ala Gly Cys Pro Leu Thr Gly Val Val His Ala Ala
 3365 3370 3375
 Gly Val Leu Asp Asp Ala Thr Ile Ala Ser Leu Thr Pro Glu Arg Leu

 3380 3385 3390
 Gly Thr Val Phe Ala Ala Lys Val Asp Ala Ala Leu Leu Leu Asp Glu
 3395 3400 3405
 Leu Thr Arg Gly Met Glu Leu Ser Ala Phe Val Leu Phe Ser Ser Ala
 3410 3415 3420
 Ala Gly Ile Leu Gly Ser Ala Gly Gln Gly Asn Tyr Ala Ala Ala Asn
 3425 3430 3435 3440
 Ala Ala Leu Asp Ala Leu Ala Tyr Arg Arg Arg Ala Ala Gly Leu Pro
 3445 3450 3455

 Gly Val Ser Leu Ala Trp Gly Leu Trp Glu Glu Ala Ser Gly Met Thr
 3460 3465 3470
 Gly His Leu Ala Gly Thr Asp His Arg Arg Ile Ile Arg Ser Gly Leu
 3475 3480 3485
 His Pro Met Ser Thr Pro Asp Ala Leu Ala Leu Phe Asp Ala Ala Leu
 3490 3495 3500
 Ala Leu Asp Arg Pro Val Leu Leu Pro Ala Asp Leu Arg Pro Ala Pro

3505 3510 3515 3520
 Pro Leu Pro Pro Leu Leu Gln Asp Leu Leu Pro Ala Thr Arg Arg Arg

 3525 3530 3535
 Thr Thr Arg Thr Thr Thr Thr Gly Gly Ala Asp Asn Gly Ala Gln Leu
 3540 3545 3550
 His Ala Arg Leu Ala Gly Gln Thr His Glu Gln Gln His Thr Thr Leu
 3555 3560 3565
 Leu Ala Leu Val Arg Ser His Ile Ala Thr Val Leu Gly His Asn Ala
 3570 3575 3580
 Pro Glu Met Ile Pro Val Asp Ser Ala Phe Arg Asp Leu Gly Phe Asp
 3585 3590 3595 3600

 Ser Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Gly Glu Ala Thr Gly
 3605 3610 3615
 Leu Arg Leu Pro Thr Ser Leu Val Phe Asp Gln Pro Asn Ala Ala Thr
 3620 3625 3630
 Leu Ala Arg His Leu Arg Arg Glu Leu Met Gly Asp Asp Ala Glu Gly
 3635 3640 3645
 Glu Thr Pro Ser Gln Val Ala Leu His Gln Val Ala Ala Asp Glu Pro
 3650 3655 3660
 Ile Ala Ile Val Gly Met Ala Cys Arg Phe Pro Gly Gly Val Cys Ser

 3665 3670 3675 3680
 Pro Glu Glu Leu Trp Glu Leu Val Ala Ser Gly Gly Asp Ala Ile Gly
 3685 3690 3695
 Glu Phe Pro Ala Gly Arg Gly Trp Asp Leu Glu Gly Leu Phe Asp Ser
 3700 3705 3710
 Asp Pro Asp Arg Ser Gly Thr Ser Tyr Ala Arg Tyr Gly Gly Phe Leu
 3715 3720 3725
 Tyr Glu Ala Gly Glu Phe Asp Ala Asp Phe Phe Gly Ile Ser Pro Arg
 3730 3735 3740

 Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Thr Ser
 3745 3750 3755 3760

Trp Glu Ala Phe Glu Arg Ala Gly Ile Asp Pro Leu Ser Met Arg Gly
 3765 3770 3775
 Ser Arg Thr Gly Val Phe Ala Gly Val Met Tyr His Asp Tyr Ala Ala
 3780 3785 3790
 Arg Leu His His Val Pro Glu Gly Phe Glu Gly Leu Ile Ala Asn Gly
 3795 3800 3805
 Ser Ala Gly Ser Val Ala Thr Gly Arg Val Ala Tyr Ser Phe Gly Leu
 3810 3815 3820
 Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val
 3825 3830 3835 3840
 Ala Leu His Trp Ala Ala Gln Ala Leu Arg Ala Gly Glu Cys Ser Met
 3845 3850 3855
 Ala Leu Ala Gly Gly Val Thr Val Met Ser Ser Pro Gly Thr Phe Val
 3860 3865 3870
 Glu Phe Ser Arg Gln Arg Gly Leu Ala Ala Asp Gly Arg Cys Lys Ala
 3875 3880 3885
 Tyr Ser Ala Ala Ala Asp Gly Thr Gly Trp Ala Glu Gly Val Gly Met
 3890 3895 3900
 Leu Leu Val Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His Arg Val
 3905 3910 3915 3920
 Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn
 3925 3930 3935
 Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Gln
 3940 3945 3950
 Ala Leu Ala Asn Ala Gly Leu Thr Pro Ala Asp Val Asp Ala Val Glu
 3955 3960 3965
 Gly His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu Ala Gln Ala
 3970 3975 3980
 Leu Leu Ala Ala Tyr Gly Gln His Arg Pro His His Arg Pro Leu Trp
 3985 3990 3995 4000
 Leu Gly Ser Leu Lys Ser Asn Ile Gly His Ala Gln Ala Ala Ala Gly
 4005 4010 4015

Val Gly Gly Val Ile Lys Met Val Met Ala Leu Arg Asn Gly Leu Leu
4020 4025 4030

Pro Gln Thr Leu His Val Asp Glu Pro Thr Pro Gln Val Asp Trp Ser
4035 4040 4045

Thr Gly Ala Val Gln Leu Leu Thr Gln Pro Val Pro Trp Pro Ala Asp
4050 4055 4060

Pro Ala Gly Arg Pro Arg His Ala Gly Val Ser Ser Phe Gly Val Ser
4065 4070 4075 4080

Gly Thr Asn Ala His Val Ile Leu Glu Glu Ala Pro Ala Ala Ala Gly
4085 4090 4095

Gly Ala Ala Gly Gly Gly Val Ser Val Gly Ala Pro Asn Pro Ala Leu
4100 4105 4110

Pro Val Ala Glu Ser Glu Pro Val Pro Val Pro Val Pro Val Ser Ala
4115 4120 4125

Arg Ser Glu Ala Gly Leu Arg Ala Gln Ala Gln Ala Leu Arg Gln Tyr
4130 4135 4140

Val Ala Ala Arg Pro Asp Met Ser Pro Ala Asp Ile Gly Ala Gly Leu
4145 4150 4155 4160

Ala Arg Gly Arg Ala Val Leu Glu His Arg Ala Val Ile Leu Ala Ala
4165 4170 4175

Asp Arg Glu Glu Leu Ala Gln Ala Leu Thr Ala Leu Ala Ala Gly Glu
4180 4185 4190

Pro His Pro His Ile Thr Thr Gly His Thr Arg Gly Ser Asp Arg Gly
4195 4200 4205

Gly Val Val Phe Val Phe Pro Gly Gln Gly Gly Gln Trp Ala Gly Met
4210 4215 4220

Gly Leu Thr Leu Leu Thr Ser Ser Pro Val Phe Ala Glu His Ile Asp
4225 4230 4235 4240

Ala Cys Glu Lys Ala Leu Thr Pro Trp Val Pro Trp Ser Leu Thr Asp
4245 4250 4255

Ile Leu His Arg Asp Pro Asp Asp Pro Ala Trp Gln Gln Ala Asp Val

4260 4265 4270
 Val Gln Pro Val Leu Phe Ser Ile Met Val Ser Leu Ala Ala Leu Trp
 4275 4280 4285
 Arg Ser Tyr Gly Ile Glu Pro Asp Ala Val Leu Gly His Ser Gln Gly
 4290 4295 4300
 Glu Ile Ala Ala Ala His Ile Cys Gly Ala Leu Ser Leu Lys Asp Ala
 4305 4310 4315 4320

 Ala Lys Thr Val Ala Leu Arg Ser Gln Ala Leu Ala Ala Val Arg Gly
 4325 4330 4335
 Arg Gly Ala Met Val Ser Leu Pro Leu Pro Ala Gln Asp Val Gln Gln
 4340 4345 4350
 Leu Ile Ser Glu Arg Trp Glu Gly Gln Leu Trp Val Ala Ala Leu Asn
 4355 4360 4365
 Gly Pro His Ser Thr Thr Val Ser Gly Asp Thr Thr Ala Val Glu Glu
 4370 4375 4380
 Leu Leu Thr His Cys Ala Asp Thr Gly Leu Arg Ala Lys Arg Ile Pro

 4385 4390 4395 4400
 Val Asp Tyr Ala Ser His Cys Pro His Val Gln Pro Leu His Asp Glu
 4405 4410 4415
 Leu Leu His Leu Leu Gly Asp Ile Thr Pro Gln Pro Ser Thr Met Pro
 4420 4425 4430
 Phe Phe Ser Thr Val Val Gly His Leu Val Trp Tyr Thr Thr Thr Leu
 4435 4440 4445
 Asp Ala Ala Tyr Trp Tyr Arg Asn Leu His Gln Pro Val Arg Phe Ser
 4450 4455 4460

 His Ala Ile Gln Thr Leu Thr Asp Asp Gly His Arg Pro Phe Ile Glu
 4465 4470 4475 4480
 Ile Ser Pro His Pro Thr Leu Val Pro Ala Ile Glu Asp Thr Thr Glu
 4485 4490 4495
 Asn Thr Thr Glu Asn Ile Thr Ala Thr Gly Ser Leu Arg Arg Gly Asp
 4500 4505 4510
 Asn Asp Thr His Arg Phe Leu Thr Ala Leu Ala His Thr His Thr Thr

4515 4520 4525
Gly Ile Arg Thr Pro Thr Thr Trp His His His Tyr Thr Gln Thr His

4530 4535 4540
Pro His Pro His Asn His His Leu Asp Leu Pro Thr Tyr Pro Phe Gln
4545 4550 4555 4560

His Gln His Tyr Trp Leu Gln Pro Pro Thr Thr Thr Thr Asp Leu Thr
4565 4570 4575

Thr Thr Gly Leu Thr Pro Thr His His Pro Leu Leu Thr Ala Thr Leu
4580 4585 4590

Thr Leu Ala Asn Asn Asn Thr Gln Leu Leu Thr Gly Arg Leu Ser Leu
4595 4600 4605

Arg Thr His Pro Trp Leu Thr Asp His Thr Val Val Gly Thr Thr Leu
4610 4615 4620

Val Pro Gly Thr Ala Leu Leu Glu Leu Ala Leu Gln Ala Thr Thr Thr
4625 4630 4635 4640

Asp His Leu Glu Glu Leu Ala Leu His Thr Pro Leu Val Ile Pro Arg
4645 4650 4655

Glu Gly Ala Val Asp Val Gln Val His Ile Asn Pro Pro Asp Asp Thr
4660 4665 4670

Asp Thr Arg Ser Leu Thr Ile Tyr Ser Arg Ser Glu Asn Ala Pro Ala

4675 4680 4685
Ala Ala Pro Trp Arg His His Ala Thr Ala Val Leu Gly Thr Lys Thr

4690 4695 4700
Ser Arg Ile Glu Thr Gly Arg Ser His Asp Asp Leu Ser Met Trp Pro

4705 4710 4715 4720
Pro Ala Gly Ala Val Arg Cys Ala Asp Glu Glu Leu Ala Ala Leu Tyr

4725 4730 4735
Gly Asp Tyr Glu Ala Asn Gly Phe Val Tyr Gly Pro Ala Phe Arg Gly

4740 4745 4750
Leu Thr Ala Ala Trp Arg Leu Gly Asp Glu Val Phe Ala Glu Val Arg

4755 4760 4765

Leu Pro Glu Gln Val His Gly Glu Ala Ser Ala Tyr Asn Leu His Pro
 4770 4775 4780
 Ala Leu Leu Asp Ala Ala Leu His Ala Ala Ala Phe Ala Pro Ser Gly
 4785 4790 4795 4800
 Ser Leu Pro Gln Gly Ser Val Pro Phe Ser Phe Thr Gly Val Thr Leu
 4805 4810 4815
 His Ala Ala Asn Ala Ser Ser Leu Arg Val Arg Leu Ser Pro Ala Asp
 4820 4825 4830
 Pro Asn Ser Gly His Ala Ala Val Ser Val Leu Val Thr Asp Asp Thr
 4835 4840 4845
 Gly Thr Pro Val Ala Ser Val Glu Ala Leu Ala Val Arg Pro Leu Ala
 4850 4855 4860
 Ala Asp Glu Leu Arg Ala Ala Glu Arg Ala Val Gln Arg Ala Glu Leu
 4865 4870 4875 4880
 Phe Asp Met Lys Trp Val Glu Val Pro Ser Asp Val Leu Val Ser Gly
 4885 4890 4895
 Gly Ala Ser Val Val Val Leu Asp Gly Ala Asp Asp Leu Val Gly Leu
 4900 4905 4910
 Ala Ala Glu Glu Asp Gly Val Pro Gly Val Val Val Leu Arg Cys Pro
 4915 4920 4925
 Asp Ala Gly Ala Asp Gly Gly Gly Gly Gly Gly Val Gly Glu Val
 4930 4935 4940
 Val Gly Gly Val Leu Gly Val Val Gln Gly Trp Leu Gly Leu Glu Arg
 4945 4950 4955 4960
 Phe Ala Gly Ser Arg Leu Val Val Val Thr Arg Gly Ala Val Val Ala
 4965 4970 4975
 Gly Pro Glu Asp Gly Pro Val Asp Gly Pro Val Asp Val Val Gly Ala
 4980 4985 4990
 Ala Val Trp Gly Leu Val Arg Ser Ala Gln Ala Glu His Pro Asp Arg
 4995 5000 5005
 Phe Val Leu Leu Asp Leu Asp Thr Asp Leu Asp Ser Gly Ala Asp Arg
 5010 5015 5020

Asp Ala Gly Asn Glu Ala Gly Met Gly Ser Gly Leu Asp Gly Gly Arg
 5025 5030 5035 5040

Val Ala Ala Val Val Ala Cys Gly Glu Pro Gln Leu Ala Val Arg Gly
 5045 5050 5055

Glu Arg Val Leu Ala Ala Arg Leu Thr Arg Leu Glu Ser Pro Val Asp
 5060 5065 5070

Val Ser Gly Arg Glu Val Leu Pro Trp Leu Ser Gly Gly Ser Val Leu
 5075 5080 5085

Val Thr Gly Gly Thr Gly Val Leu Gly Ala Ala Val Ala Arg His Leu
 5090 5095 5100

Ala Gly Val Cys Gly Val Arg Asp Leu Leu Leu Val Ser Arg Arg Gly

5105 5110 5115 5120

Pro Asp Ala Pro Gly Ala Glu Gly Leu Arg Ala Glu Leu Ala Ala Leu
 5125 5130 5135

Gly Ala Glu Val Arg Ile Val Ala Cys Asp Val Gly Glu Arg Arg Glu
 5140 5145 5150

Val Val Arg Leu Leu Glu Gly Val Pro Ala Gly Cys Pro Leu Thr Gly
 5155 5160 5165

Val Val His Ala Ala Gly Val Leu Asp Asp Ala Thr Ile Ala Ser Leu
 5170 5175 5180

Thr Pro Glu Arg Leu Gly Thr Val Phe Ala Ala Lys Val Asp Ala Ala
 5185 5190 5195 5200

Leu Leu Leu Asp Glu Leu Thr Arg Gly Met Glu Leu Ser Ala Phe Val
 5205 5210 5215

Leu Phe Ser Ser Ala Ala Gly Ile Leu Gly Ser Ala Gly Gln Gly Asn
 5220 5225 5230

Tyr Ala Ala Ala Asn Ala Ala Leu Asp Ala Leu Ala Tyr Arg Arg Arg
 5235 5240 5245

Ala Ala Gly Leu Pro Gly Val Ser Leu Ala Trp Gly Leu Trp Glu Glu

5250 5255 5260

Ala Ser Gly Met Thr Gly His Leu Ala Gly Thr Asp His Arg Arg Ile

5265 5270 5275 5280
 Ile Arg Ser Gly Leu His Pro Met Ser Thr Pro Asp Ala Leu Ala Leu
 5285 5290 5295
 Phe Asp Ala Ala Leu Ala Leu Asp Arg Pro Val Leu Leu Pro Ala Asp
 5300 5305 5310
 Leu Arg Pro Ala Pro Pro Leu Pro Pro Leu Leu Gln Asp Leu Leu Pro
 5315 5320 5325

 Ala Thr Arg Arg Arg Thr Thr Arg Thr Thr Thr Thr Gly Gly Ala Asp
 5330 5335 5340
 Asn Gly Ala Gln Leu His Gly Arg Leu Ala Gly Gln Thr His Glu Gln
 5345 5350 5355 5360
 Gln His Thr Thr Leu Leu Ala Leu Val Arg Ser His Ile Ala Thr Val
 5365 5370 5375
 Leu Gly His Thr Thr Pro Asp Thr Ile Pro Pro Asp Arg Ala Phe Arg
 5380 5385 5390
 Asp Leu Gly Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Arg Leu

 5395 5400 5405
 Ser His Thr Thr Gly Leu Arg Leu Pro Thr Thr Leu Ala Phe Asp His
 5410 5415 5420
 Pro Asn Pro Thr Thr Leu Thr His His Leu His Thr Gln Leu Val Ser
 5425 5430 5435 5440
 Lys Gly Leu Thr Ala Ala Ala Glu Pro Asp Ala Ala Thr Thr Pro Pro
 5445 5450 5455
 Gly Leu Pro Ser Leu Leu Ser Glu Leu Glu Arg Leu Glu Ala Val Val
 5460 5465 5470

 Leu Ser Ser Thr Thr Ser Ser Ala Ala Pro Leu Asp Asp Gly Ala Arg
 5475 5480 5485
 Thr Arg Leu Ala Ser Arg Leu His Ser Leu Ala Gln Lys Leu Asn Gly
 5490 5495 5500
 Asp Asp Thr Ala Pro Asp Leu Ala Glu Thr Ser Asp Glu Glu Met Phe
 5505 5510 5515 5520
 Ala Leu Ile Asp Arg Glu Val Gly Phe Glu Ser Gln

5525	5530	
<210>	73	
<211>	17460	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	milA3 gene of Streptomyces milbemycinicus	
<400>	73	
atggccgctg gccacgacaa ggtgatcgag gcgctgcggg cgtccctcaa gaccaacgag		60
cggcagaggg aacagatcca ccggtcact acggcggcgc gggaacccat cgccatcatc		120
ggcatggcct gccgctatcc gggcggagtg ggatcgccgg aggacctgtg ggagctggtg		180
gccgccggtc gtgacccat ccggcaccttc ccgaggacc ggggctggga cgtggagcgg		240
ctgtacgacc ccgatccga gcgggccggc acctcgtga cccagcatgg cggattcctg		300
taccaggcag gggagttcga ccccggtttc ttcgggatca gcccgcgca ggcgctggcg		360
atggaccgc agcagcggct gctgtggag atctcctggg aggtgttcga gcgggccggg		420
atcgaccgg cctcggtagc cggcagccgc accggggtgt tcgcgggcgt catgtaccac		480
gactacggct cccggctgca caccgtccc gaaggcttcg agggctatgt cggcaacggc		540
agcggcggcg gcgtggcgtc cggccgggtc gcctacacc tcggcctcga aggcccggcc		600
gtgaccgtgg acaccgctg ctctcctcgt ttggtcgccc tgcacctggc ctgccaggcg		660
ctgcgggcgg gcgagtgtc actcgccctg gcgggcgggg tgacggtgat gtccaccccc		720
agcctgttcg tcgagtactc ccggcagcgc gcgctcgcg cagacggccg gtgcaaggcg		780
tacggggcgg gggcggacgg caccggctgg gcagaaggcg ccgggatgct gctggtgaa		840
cggctcacgg acgcacagcg cctcggccac cgggtgtctg cggtagtcg gggcagcgcg		900
gtcaaccagg acggcgcgag caacggcctc accgccccca acggccccgc gcaacaacgg		960
ggcatccggc aggcactggc gagcgccggg gtgtcgcggt ccgaggtcga cgccgtggag		1020
gggcatggga cggggacgcg gctgggcgat ccgatcagg cgacggcgtt gctggcgacc		1080
tacggtcagc agcggccccg ggaccggccg ctgtggctcg ggtcgatgaa gtccaacgtc		1140
ggccatgcgc aggcggccgc cggcgtgggc gggatcatca agatggtgat ggccatgcgg		1200
agcgggacgc tgccgcgcac cctgcacgcg gacgagccgt cgccacacat cgactgggac		1260
tcgggcgcgg tacggctgct gaccgagccg gtgcctggc cggagcgcga ccggccccgc		1320
cgcgcgcggg tgtcctctt cggggtcagc ggcaccaacg cccatgtgat cctcgaggec		1380
gcatcgcaga cggcgccgca gacggattcc gcgtcgagg cggaaccga cgacgtccc		1440

gcaccgcacg gcgcgccggg ccatgccgtg gcggggccgc tgctctggcc ctgtcgggc	1500
gcgacggccg aggcgtcg gcgccaggcc ggggagctgc gtcgttctgt ggcggccgat	1560
gagctgctgc gccccccga cgtcgggcac accctggtct tcggccgctc ggacctcga	1620
caccgcgacg tcgtcctcgg ctccgaccgg gaaaccctgc tgcgcgtctt ggacactctg	1680
gcaggggagg ggccggacga cggctcggtc gtacggggca tggcggccgc cggggccggt	1740
gcgggcgtgg tggtcgtctt cccgggacag ggcggccagt gggccggcat ggggctcgcg	1800
ctgctggaga cctcgtcgtt cttcgccgag cggatggcgg agtgcgaggc ggcgttggca	1860
ccgtatgccg actggctcgt gctcgacgtt ctgcgccggg accccgggga cccggtctgg	1920
gagcgggccc atgtcgtcca gccgatgctg ttctcggtga tgggtgcgt ggcgcagctg	1980
tggcgtcgt acggcgtcga accggacgcc gtactcggcc actcccaggg cgagatcgcc	2040
gccgcccaca tctcgggcgc gctgacctg gacgacgcc cgaagttgt cgcgtcgc	2100
agccggggcc tgcagacct gcgcggttcg ggcggcatgg cctccgtacc actgccggcg	2160
gacgaggtea cgggctgct gcggaccga tggccggacc ggctgtgggt ggccgccgtc	2220
aacgccccca cggccacggt gatctccgc gacgcggact ctctggcgga ggcgtggaa	2280
cactaccggg accagggcgt cgaagcgaag cgggtcccgg tcgactacgc ctccactgc	2340
ccgcatacgc aagccgtgga gcaggagctg ctgggcctgt tgcgggggat cgtccaagg	2400
gccgccgaca tccccttcta ctccaccgtg gacaaccagt gggccgacac catgggactc	2460
gacgcccgt actggtaccg caatctgcgc cgcccgtac gttcgcga agcgtccgc	2520
gccctcggcg ccgccagta ccggacgtat gtcgaggtcg gccgcaccc caccctcacc	2580
cccgccatcg aggacaccac tgaggccgc ggcgtcgcgg ccacggttgt cggatccctg	2640
cgcgcggcg aggacgacgc ccaccgcac ctgacctgc tggcccgggc tcataatcat	2700
ggcctgcccc tggcgtggga ccgccactac cgggcgtcgc ccccagagc gaacctgtc	2760
gacctgccc cctacgcctt ccagcgcgc cgctactggc tggacgccc ggcgaccacc	2820
ggggacgtga cggccgcggg gctggccccg gtcggacacc cactgctcg cgcggcggtc	2880
ggactcgccg agggcgacgg atatctgtc accggccggc tcgccccga cccccccc	2940
tggctcaccg accacgcgt cgccggcacc gtctgtctgc cgggcaccgc atactggaa	3000
ctggccgtgc acgtcggcg acacctcggc tcccccggc tggaggagct caccctgcac	3060
gccccgctcg tctccccga caccggcggc gtggcgctcc aggtggccgt cggggcaccg	3120
gacgagaccg gccgccgcgc actgagcgtc tacgcacagc gcgacgacga ccccgctgg	3180
gagggggcgg cccggggcgc gtggacacgg catgcgacc gcacactggc ggccgagcc	3240

ccgactgatg gcatcagcgg tgccgacggt gccgggaccc tggcgggggc gtggcctccg	3300
ccgggcgcgg agccccctgga catcagcggc ctctacgaca cgctggccgc cgcagacttc	3360
ggctacggcc cggccttcca ggggctgcgc gccgtctggc ggcaaggcga ggagacctac	3420
gccgaggtgc ggctccccga ccaggtggcc gccgacgcc cacgttctg cctccacccc	3480
gcgctgctcg acgccgcgt ccacccgctg gcactcgaca gcggccgaag cgaggagaat	3540
ccagcgggac atggcctgct gccgttcgcc tggcgcggcg tcagcctgcg ctccccgggc	3600
acaccgacgc tgcgcgtacg gctgcggccg cagggcccgg actcgattgc cgtcgacgtg	3660
gccgacgaga cgggcgcgcc ggtggcctcg gccgaatcg tcacgtgcg gccggtggcc	3720
ctggaggacc tgcgggccct cggcggccag gcgggcgaca ccctctacgc cctggagtgg	3780
accgccgcgc ccgagcccc ggcgacggcc ctcgggcggg gcgctgtgat tggccaagcc	3840
attcttggat gggtgcgcg gctggagacg gcggcagcgg ggcccgtacg gcggtaccgc	3900
gaccttgccg gactggtgac ggccctggac gcgggcgac cgcctccgga cctggtgttc	3960
gtgggctgcc ctccggctgc cgccgggccc gacgacacga cggtcgccga cgtccacacc	4020
gcccgtaccc gtgtccgtac ccgacaagcg ctggacctgc ttcagggtcg gtcggcgaa	4080
gcgcggctgg ccggcgcgag gctggtgctg gtcacctgcg gcgcggtggc caccgggccc	4140
gcggagggag tgatggacct ggccgggcgcg gcgactcgcg gactggtgcg atccgcgcag	4200
gccgaggagc ccgaccgtat cctcctggtg gacctggacg cggccgagga gtcgtgggcg	4260
gcgctaccac gggcggtcgc gctgggcgaa ccgcagatgg ccatccgggc cggccagccg	4320
cacatggccc ggttggttcg agccgacacc gagggggggc ccctgctcac gccgccacag	4380
gggagcggcg gctggcggtc cgactgcgc gacgcgggca cggccaggg gctggcgcct	4440
gtggcgtcct cgcccgaccg cgacccgctg ggccgcacc aggtacggat cgaggtgcgt	4500
gcggccgggc tgaacttccg cgatgtcctg gtggccctgg ggatggtccc tgggcagcgg	4560
gggctgggca gcgagggcgc cggggtggtg ctgaagccg ggccctgaagt ggccgacctg	4620
gcgccggggg accgggtgat gggcgtgttc gcggatgcgt tcggcccgtt cgcgatgcc	4680
gaccgggcca ccgtgatccg cgtccccgac cactggacct tcggccaggc cgccgcgctc	4740
cccgtcgtgt tcgccaccgc ctattacggg ctggtggacc tggcaggact gcgccgggt	4800
gagtcggtgc tggcgcacgc tgcggccggc ggagtgggac tggccgctgt ccaactggcc	4860
cgccacctgg gcgtgaggt ctacgccacg gcgagccccg gcaaatggga caccctacgc	4920
gcccacggca tcccccgga gcgcacgcc tcgtcccgca ccctcgactt cgagagccgg	4980

ttcaccggcc ggaacatcga cgtcgtcctc aactccctgg cccatgagta cgtcgacgcc	5040
tcgctgcgcc tgggtgccgc cgacagcggc cggttcctgg agatgggcaa gaccgacctc	5100
cgcgacccgg aggaggtggc gcaggcgtac cccggtgtcg cctaccgggc gtacgacctg	5160
atggaggccg gacccgagcg catcggggag atcctgcgca ccgtgttcgc gctgttcgac	5220
gagggcgtgc tcaccccgct gccgctcacc tgctgggaca tccggcaggc cagggatgcc	5280
ttccgccaac tcacgagggc ccgcaccgtc ggaaagaatg tgctcacgct ggaccgcacc	5340
cccgaccccg acggcacctg cctcatcacc ggtggcaccc gtaccctcgg cgccgcgctc	5400
gccccccatc tcgccccac cggccgagca cggcatctgc tactgatcag ccgccgtggc	5460
ctcgatgcgc caggcgctcc cgaactcatc gctgagattg acgagttggg cgccacggcg	5520
accgtcgcca cctgcgacgt cggcgaccgt gccgcgtcg ccgaactgct cgggcggatc	5580
cccgccgagc acccgctgac gccgctcgtc cacgccggg gcaccctcga cgacgccacg	5640
ctcggtcccc tcaccgcgcg ccacctgac accgttctgc ccgcgaaggc cgatgccgcc	5700
tggcatctgc acgacctgac ctgccggtg gatctggccg cgttcgtgct gttctcgtcc	5760
gccgcgggtg tcttgggctc gccggggcag ggcaactacg ccgccgcaa cgcctttctc	5820
gacgcgctcg ccttcacgag cggggcgatg ggactccccg ccgtgtccct ggcatgggga	5880
ctgtgggagg aggccagcgg aatgaccggc cacctcgacc agaccgaccg caccgcgatg	5940
gccccgctcg gcttcggcc actggccacg gacgaggccc tggcgctgtt cgacaacgt	6000
ctcgtcgacg gcccaccgt gctgtctccg gcccgatcg acaccaaggc gctacggggc	6060
accaccgcac cgcctctgtt ccagagctc gtacgcccc aaccggcca ccggccacgc	6120
cccgcgacac ccgacggccg ctctctctc cgagcccggc tcgcccgggt cgaccccgcc	6180
gcacagcacg aggtcctgct caccctcgtc cgcggccacg ccgccacggt cctcgccac	6240
ccgagccccc acgcatcgc ccgcgaggcg gccttcctgt acctcggtt cgactccctc	6300
accgccgtgg agtccgcaa ccgcctcaag gaggcaaccg gcctgcggct ccccgccacc	6360
atcgtcttcg accatccac tctgcccgt ctcgcccagc acctgcggga cggcctcatc	6420
ggcggcgccg atacggtcac cctggctgcg gctcctgctc cgagcaaggt ggcgatggtg	6480
gcggatgagg ccatcgcat catcgccatg gcctgccggt atccgggggg cgtgcggtcg	6540
gccgaggggc tgtgggatct ggtcgctcc ggacccagc ccatgagcgg attccccagc	6600
gaccgcgggt gggacctcga ccgcctctac gccccccagg accaggacgt gccgggcacc	6660
acatacacc gccacggggg ctctctccac gacgcgggca agttcgacgc gggattcttc	6720
ggcatcgcc cactgaggc gctggcgatg gatccgcagc agcggctgct gctggagacc	6780
tcttgggagg ttttcgaaca cgcgggaatc gaccctcgt cgggtacggc gagccggacc	6840

ggagtcttcg ccggtgtgat gccgacggac tacggccccc ggctgcaaga caccgtggcc	6900
gaggtcgagg gctatgtcct caccggaaac tccggcagcg tcgcctcggg ccgtatcgcc	6960
tacaccttcg gcctggaagg ccccgcggtg tcggtggaca cggcgtgttc gtcgtctctg	7020
gtggcgttgc atctggcgtg tcaggcgctg cgtgcggggg agtgctccat ggcgctggcc	7080
ggcgggggtga cggatgatgc gacgcctggt gccttcgtgg agtttgcgcg gcagcggggg	7140
ttgtcggtag atgggcggtg caaggcggtt ggggtgggtg cggatggtac ggggtggcg	7200
gagggggtgg ggatgctgtt ggtggagcgg ttgtctgatg cgcggcgggt ggggcatcgg	7260
gtgttgccgg tggtagcggg ttctcggtg aatcaggatg gtgcgtcgaa tggtttgacg	7320
gcgccaatg gtccgtcgca gcagcgggtg atccggcagg cgttgccag tgcgcgggtt	7380
ggcggggcgg atgtggatgt ggtggagggg cacggtacgg ggacgcggct gggtagatccg	7440
atcaggcgcg aggcgttgct ggcgacctac ggtcaggagc ggggtggggg cggtcgttg	7500
tggttggggt cggatgaagtc gaatatcggg catgcgcagg ccgcggcggg ggttgcggt	7560
gtcatcaaga tggatgatgc gatgcggtat ggggtgttgc cgcggacgtt gcatgtgcag	7620
gagccgtcgc cgcatgtgga ctggtcctcg ggccgggtgc ggctgctgac ggaggcgggtg	7680
ccgtggccgg agacggggcg tgcgcggcgt gccgggggtgt cgtcgttcgg ggtcagtggc	7740
accaacgcgc acatcatcct cgaacaggcg ccgcctgagg agcacgacga tccggcggac	7800
gtctcgtccg ggtcgtttcc gtggatggtg tcggccaagt ccgaacaggc actacaggcg	7860
caggcagcac agttgcgcgc gtatctggcg gcacatcctg agctggggct ggctgatgtc	7920
gggtatgcgc tggcctccgg ccgcacggcc ttccggccacc gtgccgtgct cctgggccccg	7980
gaccgcgaag ccttcgtcga agagctggga gctctggagg ccggtgagga acacgccggg	8040
ctggtacggg gcgtggcgac gggtagcggg aagctggcgt ttgtgtgttc cgggcaggga	8100
acgcaacgtc cccgtatggg acacgggctg tactacgcct tcccgctgtt cggcgagcc	8160
atggacgaag cctgcgcaca cctggacca cacctcgacc atcccctgcg ggatgtcatg	8220
ttcgccgagc cgggcaccga caccgcccag ctgtccacc agaccgcta cggccagccc	8280
gccctgttcg cctccagat cgccctgcac gccttggtca ccgaacacca cggccttacc	8340
ccccactact acgccggcca ttccctcgga gagatcacg cggcccacct cgccgggatc	8400
ctacccctcc ccgacgcggc ccgcctggtc accaccgcg ccgcctcat gcaatctctc	8460
cccgccaccg gcgcatgac caccctcaa gcagaccccg acgaactcca cgaacacctc	8520
acacgatgcg aaggacgggt ctactcgcg gccgtgaacg cggccgggtc cgtggtcatc	8580
agcggtagtc gccacgacgt agacgtacg gccgaaaacc tccgcgcat gggacgaag	8640
accactgcgc tgaaggtag cggcgcttcc cactcacacc acatcgacc actcctcaac	8700

gaactccgca acacggcaga aacctcacc taccacccac cccacacccc cctcatcacc	8760
accaacccca ccgaccacga ccccaccaca cccactact ggggccggca agcgcgcgag	8820
acgggccact acgcccacac caccacaaca ctcacacccc acggcggtcac cgcctacctc	8880
gaactcggcc ccgaccacac cctcaccgcc ctcaccaccc acaacctccc cgaccacacc	8940
ccgctagccg tcccgtttct ccaccccgac caatccgaga cccacaccac ccacaccgcc	9000
ctcgcccacc tcacaccca cggecacccc accacctggc accaccatca caccceccacc	9060
cactaccacc caaacctccc cactacccc ttccaacacc accactactg gctcaacacc	9120
accactgcc a cgggtgatat gtccggtgca ggccttgagc cggcgccggca tcccctgttg	9180
ggcgccggcg tcgggttggc cgatggtgag gggttgctgt tccactggcg gattttctctc	9240
cgtacgcatc cctggctggc cgaccacgcc gtccggcgcg ccgtgttgct ccccgttacg	9300
gcctttctcg aactcgccct ccaagccgcc gccatgccg actgccgtcg ggtcgaggag	9360
cttacgtcc acaccccgt cgtcgtaccg gatagcccg gcgtagtgt gcaggctact	9420
gtggccgcgc cgaacgaagc aggaaccgg gcggtggata tctactcgcg aatcgatgtc	9480
ggcgccctca ccgccattc ggctggcgag ccgtggacgc gccatgccgc cgggtacctt	9540
gccgacaagc ctgaccaga ctgcggtgac tcggcggtatg gtgtcatgcc cgcggcgca	9600
tggccgccgc cgggtgcggt cgccgtggat ctggaggac tgtacgagca actggccgag	9660
gggggtttcc actacgtgc ggccttccgt tgctggacg ccgcctggca acgcggggac	9720
gaggtcttcg cgaccgcgt tatgtcagag gatcagctgg gcgacacggc tgcggtcgg	9780
ttcgcgctgc accccgcgt gctggattcc gcactgcaca ccattccact ttgcccctcc	9840
ctacggggac aacaggacag cgggctgccg ttacgtgga caggagtcac cctgcgtgca	9900
tccggggcga cgctctgcg cgtccggctg aggccggacg gccatggccc gggggcggtg	9960
tcggtcgacg tgtccgaga ggccgggtgag ccgtagcat cggtcggtc gttggccctg	10020
cggccggtga ccaggccga gttgcatacg gccgagttgc gcacagccgc cccggttgcc	10080
ccccatggt cgtctctga ggtgcgatgg gaaccgtcc ccagccttc agcggccgaa	10140
gaagccgccc catgggtgat gatcgggacc gggccgacgc tgcgccggt cgaggacttc	10200
gtcactccgc cggagcggac gtacccgac ctggccgcgc tgtgcgtggc aatcgccgat	10260
gacgcgccgc ttccccggac ggtcgtggcc tggteccag ccgggagcga agacgagtcg	10320
agtgaggcgc tgcgccaggc cacacaccac atgctgggcc tactgcagca gtggttggcg	10380
gacagccggt tcgccgacag tcgcctggtg atcctcacc gagccgcggt ggccactgcg	10440

ccggacgagg aggtagaaga cctggcggga gcggcggcgc ggggtctgat ccgctccgcc	10500
cagtcggagc accctgaccg attcgtcctg ctcgacctgg acgacctcc cgtgacgcg	10560
aaagaccacg accgaatgct gtcgatggcc ctggcctgcg gggaaccgga agtggccgta	10620
cgcgatggag ccctgcgcac accccggctg agcccgtgg ccggcaccgc caccgaggcc	10680
atggacgagc atccctggga tcaggacggc accgtactca tcaccggcgg caccggcagc	10740
ctcggcgcca tgcctgcccc ccacttggtg gccacccatg gcgtacggca tctgatgctg	10800
atcagccgac gtggcctcga cgccccgggg gccaggcgac tgggggtcga acttgcgag	10860
ctcggggcgc aggtgacgat caccgcgtgc gatgccgag accaaaggca acttgcaac	10920
gtattgtcgg agatctccgt cgaccatccg ctgaccgtg tggatcatgc ggcaggcgta	10980
ctggacgacg gggatgatcac atccctcaca ccggagggcc tgacctatgt cctgcgggcc	11040
aaggtcgatt cggcgctcaa tctccaccag ctcacacgcg acctgccgt gtccgcgttt	11100
gtgctcttct cctcgctggc cggggtgatg ggttcggcag ggcagggcaa ctacccgcc	11160
gccaacgcag ccctggacgc gctggcgagt caccggaggg ccgctcggct gccggcggtg	11220
tccttggcct ggggagtttg ggagcagacc gagggcata cgggcagtt ggaggccacg	11280
gaccacgcgc ggctccgccg ctcgggcctg aggccgtgg ccatcagcga gggcctggag	11340
ctcttcgaca aggccctgag ctgtggacac gccctggtgg tgcccgcgc actcagcacg	11400
agggagcttc agacatccgg atccgtccc ccatctctgc gccacctgac gggatgtcgt	11460
ccggccccgc cgtcccgac ccgcgacgc tcggccggtg agccgacct cctgcggcgg	11520
cggttgaccg gcctcgggcc ggaagaacgg ctacgcgagg tgctcggct ggtgcgtcc	11580
cgggcggctg cggatgtgg gcacggcacg gccgaatcgg tcccggcgga ctcggcgttc	11640
cgcgacctgg ggttcgactc cctcgccgcg gtggacctgc ggaaccggt gcagcaggcc	11700
accgggctgc gcctgccgc cggcttgatc ttcgaccggc cgcgtccgga cgtgctgcc	11760
cgtttctgt gtgacgagtt ggccggcgcc ggcggtacgt cggcggccac ggccgcccc	11820
cccgttcgg ccgtcggcg ggcagccggc gagccgtgg ccatcgctgg catggcatgc	11880
cggtttcgg gaggtgtgcg gtcggccgag ggcctgtggg atctggtcgc ctccgtatg	11940
gacgcggtgg gtgacttccc cgcagaccga ggctgggagg tggaacggct ctacacccc	12000
gaccggacc gaaccggcac ctctacacc cggcaaggcg ggttccttta cgacgcgggt	12060
gagttcgacg cggcattctt cgggatcggc ccgctgagg cggtagccat ggatccacag	12120
cagcggtgc tgcctggagat ctctgggag gcgttggaac gtgcggggat cgaccggcg	12180
tcgtcgggg ggagttcgac cggggtgttc gctggggtga tgtaccacga ctacgcacc	12240
cgcctgcgcg agatcccaga gggctacgag ggctatatcg gcaatggaaa cgcgggcagc	12300

gtcgcgtcgg gacgtgtcgc ctacaccttc ggccctggagg ggccggcggt caccgtggac	12360
acggcgtgtt cgtcgtccct ggtcgccctg catctggcct gccaggcgct gcggtcaggg	12420
gagtgtccca tggcgtggc cggcggggtc accgtcatgt ccacccccac cacttttgtc	12480
gagttctcgc gccagcgggg actggccccg gacgggcggt gcaagtcctt cggggccggc	12540
gcggacggaa caggctgggc ggagggggcg gggatgtctc tgggtgaacg gctttcggac	12600
gcccggcgca acggccaccg ggtcctggcg gtggtacggg ggagtgcggt caaccaggac	12660
ggggcgagca atgggctgac ggccgccaac ggcccgtcgc aagagcgggt gatccgccag	12720
gcgtgggcaa acgcgggtgt ggcccgcatg gacatcgacg cgggtggagg acacggcacg	12780
gggacgacgc tcggtgacct catcgaggcc caggcgctgc tggggacgta cggacaggga	12840
cggtcggccg atcgccgtt gtggttggga tcgatcaagt ccaacgtcgg acacaccag	12900
gccgcccgcg ggttggcgcg cgtcatcaag atggtgatgg ccatgcgcca cgggctgctc	12960
ccgcagacct tgcacgccga ggagccctca cctcatgtgg actggtcggg cgggacggtg	13020
cggttgtcta ccgagtcggt ggccctggccc gagcaggggc ggatgcgccg tgcgggcgtc	13080
tcctctttcg gtgtcagcgg taccaacgcc cactcatcc tggaaacaagc accgcctgcc	13140
gcggagacct acgaaccggc agagcccaac accgcgcag gccactgcc ctgggcgac	13200
tccgcgaaga gcccgaagc gctacgtgcc caggcgcgcc aactgcacac gtacctgacc	13260
aacgcccccg aggcgaacct cgccgacgtc ggccacacce tcgcgacggg ccgcgcctct	13320
ttcgagcatc gtgtgtggt catcggtcc gaccgagcgg agttcctggg tggcctggat	13380
gtcttggcgg ccgacgagc ccacaccgcc gtctgcacgg ggatcgcgag gaaggccggt	13440
gaccagggca aggtggtgtt cgtgttcccc gggcagggcg gtcagtgggc cgggatggga	13500
ctgcgcgtgc ttaagacct acccgtcttc gcccaatcga tccaggcctg cgaacaagcc	13560
ctgccccc acaccactg gacctgacc gacatctgc accggcccca caccgaccc	13620
ctgtggcagc gcgccgacgt catccagccc gtctcttctg cctcatgac ctccctgcc	13680
gccctctggc aatcccacgg ccttaacccc gacgccgtca tcggccactc ccaaggcgaa	13740
atcacgcgg ccacatcag cggagcgctg agcctggagg acgcccgaa aaccgtcgcg	13800
ctgcgcagcc gggccctgca gacctgcgc ggttcggcg gcatggcctc cgtaccactg	13860
ccggcgacc aggtcaccgg gctgtgcag accatgtggc cggaccggt gtgggtggcc	13920
gccgtcaacg ccctaccgc cacggtgatc tccggcaacg cggaagctct cacacaggcg	13980
ctggaacct accgggacca aggcgtcgac gcgaaacgga tccggctga ctacgcctc	14040
cactgcccc acatccagc cgtggaacag gaactgtcac ggctgttgcg gggcatcacc	14100
ccacggcgcc ccaccacccc cttctactcc accaccgaca accaatggac cgacaccacc	14160

accctcaacg cccactactg gtaccgaaac ctccgccaac ccgtccacct cgccgacgcc	14220
atcaccaacc tcaccacca aggccaccac accttcacg aaatcagccc ccaccccacc	14280
ctcaccctcg ccatccaaga aaccaccgac accaccacaca ccccaccac cgtcatcagc	14340
acactccgcc gcaaccacaa cgacaccac caaatcctcc acgcctcgc ccacgccac	14400
accaccggcc accccatcaa ctggcacacc accaccaac accacacccc aaccccccaa	14460
cacatcgacc taccaccta ccccttccaa caccaccact actgggtcaa cacccccacc	14520
cagacagggg atgcggcggc cgtcggcctg gaccggcac atcacccgtt gctgggcgcg	14580
gcggtcgcgg tggccgaggg ggagggtat ctgctaccg gtcggctcgc cctgtccacc	14640
caccctggc tcgccatca caccatgcc ggcgcggttg tcctccctgg aactgccctt	14700
ctcgagatcg ccttcaggc gggccatcgt gtggactgct ggcgcatcga agaactcacc	14760
ctccaatcac cgtgttcat cccggaagag ggagcagtac aggtgcaggc atgggtggcg	14820
gcaccggatg agaacgggtg ccgaagcctg acggtgtcct cccgacgcga gggtacgtac	14880
gaggacgcca cgtgggtgcg ccatgccacg ggccgggtcg gcccgcacc ggccgaccag	14940
gatgaagcca tcgcaggct caccgacca caaggcgacg gagcggcggc ggcggtctgg	15000
ccaccgcagg gcgtgtcgc gttcaccgca gacgatctgg agggcctgta cgacgggtac	15060
gcggcgcggg gattcgagta cggcccgtg ttccgaggcc tgcggcggcg ctggcgacgt	15120
ggcgaggaca tcttcgccga ggtgcgcctt cccgacacgg cggacggcga cgcctcccag	15180
ttctccgtac accccgcctt gctggacgcc gactgcacg ccgcggcctt ccgccggcc	15240
gacaaactcc cgcacggcgc cctgccgttc tccttcagcg gggtaggct gcacgggccc	15300
ggagcgtcga ccttcgggt gcgcctcacc ccggacggcc aggcgcggga cacgcacgca	15360
tggtcggtcg cggtggtcga cggcgagggg cggccgttg cctcgatcgc atcgctcgcg	15420
gtccgcccg tgcgacgca ggagtgtctg gcggcctccg gtacggcgcg gcgggactcg	15480
ctcttcgcgg tcgagtgggt gaccgccctg gcgccacct cgtcgtccgt tccgcaacgc	15540
ctggccacgg tggggcccag gcaccgcctc ccctcggcag acgcgtacgc gaacctgcc	15600
gacctggccg ccgcagtgct ggaggcgggg gccccggcg ccgatgcggt cgtggtcgac	15660
tgcggccgcc gcgatgcgcg gccaccgcc gtgccggagg acgtaaggac cctcaccgg	15720
cgcactctgg gtctgtcga ggagtggctg gcggacgaga ggccggcctc gagccggatg	15780
gtcgtactga cccgtgggtg ggtggccacc actccggggg aggacgtggc ggacctggcg	15840
ggcgcgcgcg tgtcggcat ggtgcgctcc gcgcagtcgg aacatcccgg ccggttcgtc	15900

ctgctggacc tcgaccccg cccggacctc gacggcgggg aagtgccacc gaccgtcgtt 15960
ccggcggctc tcgcctgtgg tgagccgcag atcgcggtgc gtgcgaaccg gcacctggtg 16020
ccccggctga cccgcgttcc ggcgctccgtc cccgtccccg ggcggtgttcc cgttccccgc 16080
gccgaggcag ccgacccgga caccacgccc acggcggttcg accccgacgg caccgtagtg 16140
atcacccggc gcaccggcac ccttggcgcg atgctcgcgc gccatctggt cagccgtcac 16200
ggtgtacgac acctcctgct ggcatcgcg cgcggaccgc acgcacccgg cgccaccgag 16260
ctgcgggcgg aactggccga gctcggcgcc gaggtgacgg tgcgcgcttg tgacaccggt 16320

gaccgaggcg cgctggcgga tctcatcgcg gggattccca ccggccaccc tttagaccgt 16380
gtggtccacg ctgcgggctt cctggacgac gccaccgtcg cctcgctcac cccccacac 16440
ctggacaccg cgctgacacc caaggccgac gccgccttcc atctgcacga gtcacccgc 16500
cacgcccggc cgcgcgcctt cgtctgttcc tcctcgcccg ccggtgtcct cggcgcagcc 16560
gggcagggca actatgcggc cgcaacgct ttctcgcag ccctcgccga acaccgagg 16620
gcgcagggcc tgccggcctt gtcgctcgcg tggggcctgt gggagcaggg cagcgcatg 16680
accgggcatc tcgaccgac cgaccgggccc cgcataaacc gtcgggact cgccccctc 16740

gccacggagg acgtctcgc gctcttcgac gccgcctcg ccggcgatcg gccgttctg 16800
gtgcccgcc ggcctggacct gcggggttca agcgcgcgcg agaccccgcc gccgtgttc 16860
tccaggatcg ccccggtcgt tacgacccgg ggccgggtccc ccggcgccga gggcgccgt 16920
gaccttcgta cccgtctcgc ggcccaggac gccgcgcgagc agcgcgacac gcttctcacg 16980
atcgtccga cccacaccgc cgccgtctg gggcatgaca ccgctgccgc cgtgcggccg 17040
gacggggcct tccgtgaact gggtttcgac tcctcgcgc ccgtggaact ccgtaaccgc 17100
cttcaaagca ccaccgcct caccctgcc gcgaccaccg tcttcgacca cccaccccc 17160

gctgcctcg ccgateatct gcgtactcag ctctgccagg acgtcagtc ctcggcgcg 17220
gccacggcca tggcggcgat ggcgagctg gccaggctgg agtcgcgcgt ctccgattcg 17280
gtggcgctcg acgacgac gcgcagcggc ctgcgcgagc gcctgcggtc cctcgcccgc 17340
aagatgagca gtggcgtgt cgtcgaccac gacggcggcg gcgtgcgga cctggatctt 17400
cagtcggtca cggacgatga gatgttcgag ctgatcgaca aggaggtcag ccgagactga 17460
17460

<210> 74
<211> 5819
<212> PRT

<213> Artificial Sequence

<220><223> milA3 protein of Streptomyces milbemycinicus

<400> 74

Met Ala Ala Gly His Asp Lys Val Ile Glu Ala Leu Arg Ala Ser Leu

1 5 10 15

Lys Thr Asn Glu Arg Gln Arg Glu Gln Ile His Arg Leu Thr Thr Ala

20 25 30

Ala Arg Glu Pro Ile Ala Ile Ile Gly Met Ala Cys Arg Tyr Pro Gly

35 40 45

Gly Val Gly Ser Pro Glu Asp Leu Trp Glu Leu Val Ala Ala Gly Arg

50 55 60

Asp Ala Ile Gly Thr Phe Pro Glu Asp Arg Gly Trp Asp Val Glu Arg

65 70 75 80

Leu Tyr Asp Pro Asp Pro Glu Arg Ala Gly Thr Ser Cys Thr Gln His

85 90 95

Gly Gly Phe Leu Tyr Gln Ala Gly Glu Phe Asp Pro Gly Phe Phe Gly

100 105 110

Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu

115 120 125

Leu Glu Ile Ser Trp Glu Val Phe Glu Arg Ala Gly Ile Asp Pro Ala

130 135 140

Ser Val Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Val Met Tyr His

145 150 155 160

Asp Tyr Gly Ser Arg Leu His Thr Val Pro Glu Gly Phe Glu Gly Tyr

165 170 175

Val Gly Asn Gly Ser Gly Gly Gly Val Ala Ser Gly Arg Val Ala Tyr

180 185 190

Thr Leu Gly Leu Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser

195 200 205

Ser Ser Leu Val Ala Leu His Leu Ala Cys Gln Ala Leu Arg Ala Gly

210 215 220

Glu Cys Ser Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser Thr Pro

225 230 235 240

Ser Leu Phe Val Glu Tyr Ser Arg Gln Arg Ala Leu Ala Ala Asp Gly
245 250 255

Arg Cys Lys Ala Tyr Gly Ala Gly Ala Asp Gly Thr Gly Trp Ala Glu
260 265 270

Gly Ala Gly Met Leu Leu Val Glu Arg Leu Thr Asp Ala Gln Arg Leu
275 280 285

Gly His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp
290 295 300

Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ala Gln Gln Arg
305 310 315 320

Ala Ile Arg Gln Ala Leu Ala Ser Ala Gly Val Ser Ala Ser Glu Val
325 330 335

Asp Ala Val Glu Gly His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile
340 345 350

Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Gln Arg Pro Ala Asp
355 360 365

Arg Pro Leu Trp Leu Gly Ser Met Lys Ser Asn Val Gly His Ala Gln
370 375 380

Ala Ala Ala Gly Val Gly Gly Ile Ile Lys Met Val Met Ala Met Arg
385 390 395 400

Ser Gly Thr Leu Pro Arg Thr Leu His Ala Asp Glu Pro Ser Pro His
405 410 415

Ile Asp Trp Asp Ser Gly Ala Val Arg Leu Leu Thr Glu Pro Val Ala
420 425 430

Trp Pro Glu Arg Asp Arg Pro Arg Arg Ala Ala Val Ser Ser Phe Gly
435 440 445

Val Ser Gly Thr Asn Ala His Val Ile Leu Glu Ala Ala Ser Gln Thr
450 455 460

Ala Pro Gln Thr Asp Ser Ala Ser Gln Ala Glu Thr Asp Asp Ala Pro
465 470 475 480

Ala Pro His Gly Ala Pro Gly His Ala Val Ala Gly Pro Leu Leu Trp

485	490	495
Pro Leu Ser Gly Ala Thr	Ala Glu Ala Leu Arg Ala Gln Ala Gly Glu	
500	505	510
Leu Arg Arg Phe Val	Ala Ala Asp Glu Leu Leu Arg Pro Ala Asp Val	
515	520	525
Gly His Thr Leu Val Phe	Gly Arg Ser Asp Leu Ala His Arg Ala Val	
530	535	540
Val Leu Gly Ser Asp	Arg Glu Thr Leu Leu Arg Ala Leu Asp Thr Leu	
545	550	555
		560
Ala Gly Glu Gly Pro Asp	Asp Gly Ser Val Val Arg Gly Met Ala Ala	
565	570	575
Ala Gly Ala Gly Ala Gly	Val Val Phe Val Phe Pro Gly Gln Gly Gly	
580	585	590
Gln Trp Ala Gly Met Gly	Leu Arg Leu Leu Glu Thr Ser Ser Phe Phe	
595	600	605
Ala Glu Arg Met Ala Glu	Cys Glu Ala Ala Leu Ala Pro Tyr Ala Asp	
610	615	620
Trp Ser Leu Leu Asp Val	Leu Arg Arg Asp Pro Gly Asp Pro Val Trp	
625	630	635
Glu Arg Ala Asp Val Val	Gln Pro Met Leu Phe Ser Val Met Val Ser	
645	650	655
Leu Ala Gln Leu Trp Arg	Ser Tyr Gly Val Glu Pro Asp Ala Val Leu	
660	665	670
Gly His Ser Gln Gly Glu	Ile Ala Ala Ala His Ile Cys Gly Ala Leu	
675	680	685
Thr Leu Asp Asp Ala Ala	Lys Val Val Ala Leu Arg Ser Arg Ala Leu	
690	695	700
Gln Thr Leu Arg Gly Ser	Gly Gly Met Ala Ser Val Pro Leu Pro Ala	
705	710	715
Asp Glu Val Thr Gly Leu	Leu Arg Thr Ala Trp Pro Asp Arg Leu Trp	
725	730	735
Val Ala Ala Val Asn Ala	Pro Thr Ala Thr Val Ile Ser Gly Asp Ala	

740 745 750
 Asp Ser Leu Ala Glu Ala Leu Glu His Tyr Arg Asp Gln Gly Val Glu
 755 760 765
 Ala Lys Arg Val Pro Val Asp Tyr Ala Ser His Cys Pro His Ile Glu

 770 775 780
 Ala Val Glu Gln Glu Leu Leu Gly Leu Leu Arg Gly Ile Ala Pro Arg
 785 790 795 800
 Ala Ala Asp Ile Pro Phe Tyr Ser Thr Val Asp Asn Gln Trp Ala Asp
 805 810 815
 Thr Met Gly Leu Asp Ala Arg Tyr Trp Tyr Arg Asn Leu Arg Arg Pro
 820 825 830
 Val Arg Phe Ala Glu Ala Leu Arg Ala Leu Gly Ala Ala Glu Tyr Arg
 835 840 845

 Thr Tyr Val Glu Val Gly Pro His Pro Thr Leu Thr Pro Ala Ile Glu
 850 855 860
 Asp Thr Thr Glu Ala Ala Gly Val Ala Ala Thr Val Val Gly Ser Leu
 865 870 875 880
 Arg Arg Gly Glu Asp Asp Ala His Arg Ile Leu Thr Ser Leu Ala Arg
 885 890 895
 Ala His Ile His Gly Leu Pro Val Ala Trp Asp Arg His Tyr Arg Ala
 900 905 910
 Leu Ala Pro Glu Ala Asn His Val Asp Leu Pro Thr Tyr Ala Phe Gln

 915 920 925
 Arg Arg Arg Tyr Trp Leu Asp Ala Pro Ala Thr Thr Gly Asp Val Thr
 930 935 940
 Ala Ala Gly Leu Ala Pro Val Gly His Pro Leu Leu Gly Ala Ala Val
 945 950 955 960
 Gly Leu Ala Glu Gly Asp Gly Tyr Leu Leu Thr Gly Arg Leu Ala Pro
 965 970 975
 His Thr His Pro Trp Leu Thr Asp His Ala Val Ala Gly Thr Val Leu
 980 985 990

Leu Pro Gly Thr Ala Tyr Val Glu Leu Ala Val His Val Gly Gly His
 995 1000 1005
 Leu Gly Cys Pro Arg Leu Glu Glu Leu Thr Leu His Ala Pro Leu Val
 1010 1015 1020
 Leu Pro Asp Thr Gly Gly Val Ala Leu Gln Val Ala Val Gly Ala Pro
 1025 1030 1035 1040
 Asp Glu Thr Gly Arg Arg Ala Leu Ser Val Tyr Ala Gln Arg Asp Asp
 1045 1050 1055
 Asp Pro Ala Trp Glu Gly Ala Ala Arg Gly Ala Trp Thr Arg His Ala
 1060 1065 1070
 Thr Gly Thr Leu Ala Ala Glu Ala Pro Thr Asp Gly Ile Ser Gly Ala
 1075 1080 1085
 Asp Gly Ala Gly Thr Leu Ala Gly Ala Trp Pro Pro Gly Ala Glu
 1090 1095 1100
 Pro Leu Asp Ile Ser Gly Leu Tyr Asp Thr Leu Ala Ala Ala Asp Phe
 1105 1110 1115 1120
 Gly Tyr Gly Pro Ala Phe Gln Gly Leu Arg Ala Val Trp Arg Gln Gly
 1125 1130 1135
 Glu Glu Thr Tyr Ala Glu Val Arg Leu Pro Asp Gln Val Ala Ala Asp
 1140 1145 1150
 Ala Pro Arg Phe Cys Leu His Pro Ala Leu Leu Asp Ala Ala Leu His
 1155 1160 1165
 Pro Leu Ala Leu Asp Ser Gly Arg Ser Glu Glu Asn Pro Ala Gly His
 1170 1175 1180
 Gly Leu Leu Pro Phe Ala Trp Arg Gly Val Ser Leu Arg Ser Pro Gly
 1185 1190 1195 1200
 Thr Pro Thr Leu Arg Val Arg Leu Arg Pro Gln Gly Pro Asp Ser Ile
 1205 1210 1215
 Ala Val Asp Val Ala Asp Glu Thr Gly Ala Pro Val Ala Ser Ala Glu
 1220 1225 1230
 Ser Leu Thr Leu Arg Pro Val Ala Leu Glu Asp Leu Arg Ala Leu Gly
 1235 1240 1245

Gly Gln Ala Gly Asp Thr Leu Tyr Ala Leu Glu Trp Thr Ala Ala Pro
1250 1255 1260

Glu Pro Pro Ala Thr Ala Leu Gly Arg Cys Ala Val Ile Gly Gln Ala
1265 1270 1275 1280

Ile Pro Gly Trp Ala Ala Ala Leu Glu Thr Ala Ala Ala Gly Pro Val
1285 1290 1295

Arg Arg Tyr Pro Asp Leu Ala Gly Leu Val Thr Ala Leu Asp Ala Gly
1300 1305 1310

Asp Pro Pro Pro Asp Leu Val Phe Val Gly Cys Pro Pro Ala Ala Ala
1315 1320 1325

Gly Pro Asp Asp Thr Thr Val Ala Asp Val His Thr Ala Arg Thr Arg
1330 1335 1340

Val Arg Thr Arg Gln Ala Leu Asp Leu Leu Gln Gly Trp Leu Gly Glu
1345 1350 1355 1360

Ala Arg Leu Ala Gly Ala Arg Leu Val Leu Val Thr Cys Gly Ala Val
1365 1370 1375

Ala Thr Gly Pro Ala Glu Gly Val Met Asp Leu Ala Gly Ala Ala Ile
1380 1385 1390

Cys Gly Leu Val Arg Ser Ala Gln Ala Glu Glu Pro Asp Arg Ile Leu
1395 1400 1405

Leu Val Asp Leu Asp Ala Ala Glu Glu Ser Trp Ala Ala Leu Pro Arg
1410 1415 1420

Ala Val Ala Leu Gly Glu Pro Gln Met Ala Ile Arg Ala Gly Gln Pro
1425 1430 1435 1440

His Met Ala Arg Leu Val Arg Ala Asp Thr Glu Gly Gly Ala Leu Leu
1445 1450 1455

Thr Pro Pro Gln Gly Ser Gly Gly Trp Arg Leu Asp Cys Ala Asp Ala
1460 1465 1470

Gly Thr Val Gln Gly Leu Ala Pro Val Ala Ser Ser Ala Asp Arg Asp
1475 1480 1485

Pro Leu Gly Pro His Gln Val Arg Ile Glu Val Arg Ala Ala Gly Leu

1490 1495 1500
 Asn Phe Arg Asp Val Leu Val Ala Leu Gly Met Val Pro Gly Gln Arg
 1505 1510 1515 1520
 Gly Leu Gly Ser Glu Gly Ala Gly Val Val Leu Glu Ala Gly Pro Glu
 1525 1530 1535
 Val Ala Asp Leu Ala Pro Gly Asp Arg Val Met Gly Val Phe Ala Asp
 1540 1545 1550
 Ala Phe Gly Pro Phe Ala Ile Ala Asp Arg Ala Thr Val Ile Arg Val
 1555 1560 1565

 Pro Asp His Trp Thr Phe Gly Gln Ala Ala Ala Val Pro Val Val Phe
 1570 1575 1580
 Ala Thr Ala Tyr Tyr Gly Leu Val Asp Leu Ala Gly Leu Arg Pro Gly
 1585 1590 1595 1600
 Glu Ser Val Leu Val His Ala Ala Ala Gly Gly Val Gly Leu Ala Ala
 1605 1610 1615
 Val Gln Leu Ala Arg His Leu Gly Ala Glu Val Tyr Ala Thr Ala Ser
 1620 1625 1630
 Pro Gly Lys Trp Asp Thr Leu Arg Ala His Gly Ile Pro Pro Glu Arg

 1635 1640 1645
 Ile Ala Ser Ser Arg Thr Leu Asp Phe Glu Ser Arg Phe Thr Gly Arg
 1650 1655 1660
 Asn Ile Asp Val Val Leu Asn Ser Leu Ala His Glu Tyr Val Asp Ala
 1665 1670 1675 1680
 Ser Leu Arg Leu Val Ser Gly Asp Ser Gly Arg Phe Leu Glu Met Gly
 1685 1690 1695
 Lys Thr Asp Leu Arg Asp Pro Glu Glu Val Ala Gln Ala Tyr Pro Gly
 1700 1705 1710

 Val Ala Tyr Arg Ala Tyr Asp Leu Met Glu Ala Gly Pro Glu Arg Ile
 1715 1720 1725
 Gly Glu Ile Leu Arg Thr Val Leu Arg Leu Phe Asp Glu Gly Val Leu
 1730 1735 1740
 Thr Pro Leu Pro Leu Thr Cys Trp Asp Ile Arg Gln Ala Arg Asp Ala

1745 1750 1755 1760
 Phe Arg Gln Leu Gln Gln Gly Arg Thr Val Gly Lys Asn Val Leu Thr
 1765 1770 1775
 Leu Asp Arg Thr Pro Asp Pro Asp Gly Thr Val Leu Ile Thr Gly Gly

 1780 1785 1790
 Thr Gly Thr Leu Gly Ala Ala Leu Ala Arg His Leu Ala Ala Thr Gly
 1795 1800 1805
 Arg Ala Arg His Leu Leu Leu Ile Ser Arg Arg Gly Leu Asp Ala Pro
 1810 1815 1820
 Gly Ala Pro Glu Leu Ile Ala Glu Ile Asp Glu Leu Gly Ala Thr Ala
 1825 1830 1835 1840
 Thr Val Ala Thr Cys Asp Val Gly Asp Arg Ala Ala Leu Ala Glu Leu
 1845 1850 1855

 Leu Gly Arg Ile Pro Ala Glu His Pro Leu Thr Ala Val Val His Ala
 1860 1865 1870
 Ala Gly Thr Leu Asp Asp Ala Thr Leu Gly Ser Leu Thr Ala Arg His
 1875 1880 1885
 Leu Asp Thr Val Leu Pro Ala Lys Ala Asp Ala Ala Trp His Leu His
 1890 1895 1900
 Asp Leu Thr Cys Arg Leu Asp Leu Ala Ala Phe Val Leu Phe Ser Ser
 1905 1910 1915 1920
 Ala Ala Gly Val Leu Gly Ser Pro Gly Gln Gly Asn Tyr Ala Ala Ala

 1925 1930 1935
 Asn Ala Phe Leu Asp Ala Leu Ala Phe Gln Arg Arg Ala Met Gly Leu
 1940 1945 1950
 Pro Ala Val Ser Leu Ala Trp Gly Leu Trp Glu Glu Ala Ser Gly Met
 1955 1960 1965
 Thr Gly His Leu Asp Gln Thr Asp Arg Thr Arg Met Ala Arg Val Gly
 1970 1975 1980
 Leu Arg Pro Leu Ala Thr Asp Glu Ala Leu Ala Leu Phe Asp Asn Ala
 1985 1990 1995 2000

Leu Val Asp Gly Pro Pro Leu Leu Leu Pro Ala Arg Ile Asp Thr Lys
 2005 2010 2015
 Ala Leu Arg Gly Thr Thr Ala Pro Pro Leu Phe Gln Ser Leu Val Arg
 2020 2025 2030
 Pro Thr Thr Gly His Arg Pro Arg Pro Ala Thr Pro Asp Gly Arg Ser
 2035 2040 2045
 Ser Leu Arg Ala Arg Leu Ala Gly Leu Asp Pro Ala Ala Gln His Glu
 2050 2055 2060
 Val Leu Leu Thr Leu Val Arg Gly His Ala Ala Thr Val Leu Gly His
 2065 2070 2075 2080
 Pro Ser Pro Asp Ala Ile Ala Arg Glu Ala Ala Phe Arg Asp Leu Gly
 2085 2090 2095
 Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Lys Glu Ala
 2100 2105 2110
 Thr Gly Leu Arg Leu Pro Ala Thr Ile Val Phe Asp His Pro Thr Pro
 2115 2120 2125
 Ala Ala Leu Ala Gln His Leu Arg Asp Gly Leu Ile Gly Gly Ala Asp
 2130 2135 2140
 Thr Val Thr Leu Ala Ala Ala Pro Ala Pro Ser Lys Val Ala Met Val
 2145 2150 2155 2160
 Ala Asp Glu Ala Ile Ala Ile Ile Gly Met Ala Cys Arg Tyr Pro Gly
 2165 2170 2175
 Gly Val Arg Ser Ala Glu Gly Leu Trp Asp Leu Val Ala Ser Gly Thr
 2180 2185 2190
 Asp Ala Met Ser Gly Phe Pro Ser Asp Arg Gly Trp Asp Leu Asp Arg
 2195 2200 2205
 Leu Tyr Ala Pro Gln Asp Gln Asp Val Pro Gly Thr Thr Tyr Thr Arg
 2210 2215 2220
 His Gly Gly Phe Leu His Asp Ala Gly Lys Phe Asp Ala Gly Phe Phe
 2225 2230 2235 2240
 Gly Ile Gly Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu
 2245 2250 2255

Leu Leu Glu Thr Ser Trp Glu Val Phe Glu His Ala Gly Ile Asp Pro
 2260 2265 2270
 Ser Ser Val Arg Arg Ser Arg Thr Gly Val Phe Ala Gly Val Met Pro
 2275 2280 2285

 Thr Asp Tyr Gly Pro Arg Leu Gln Asp Thr Val Ala Glu Val Glu Gly
 2290 2295 2300
 Tyr Val Leu Thr Gly Asn Ser Gly Ser Val Ala Ser Gly Arg Ile Ala
 2305 2310 2315 2320
 Tyr Thr Phe Gly Leu Glu Gly Pro Ala Val Ser Val Asp Thr Ala Cys
 2325 2330 2335
 Ser Ser Ser Leu Val Ala Leu His Leu Ala Cys Gln Ala Leu Arg Ala
 2340 2345 2350
 Gly Glu Cys Ser Met Ala Leu Ala Gly Gly Val Thr Val Met Ala Thr

 2355 2360 2365
 Pro Gly Ala Phe Val Glu Phe Ala Arg Gln Arg Gly Leu Ser Val Asp
 2370 2375 2380
 Gly Arg Cys Lys Ala Phe Gly Val Gly Ala Asp Gly Thr Gly Trp Ala
 2385 2390 2395 2400
 Glu Gly Val Gly Met Leu Leu Val Glu Arg Leu Ser Asp Ala Arg Arg
 2405 2410 2415
 Leu Gly His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln
 2420 2425 2430

 Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln
 2435 2440 2445
 Arg Val Ile Arg Gln Ala Leu Ala Ser Ala Arg Val Gly Gly Ala Asp
 2450 2455 2460
 Val Asp Val Val Glu Gly His Gly Thr Gly Thr Arg Leu Gly Asp Pro
 2465 2470 2475 2480
 Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Glu Arg Val Gly
 2485 2490 2495
 Asp Gly Ser Leu Trp Leu Gly Ser Val Lys Ser Asn Ile Gly His Ala

2500	2505	2510	
Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met Val Met Ala Met			
2515	2520	2525	
Arg Tyr Gly Val Leu Pro Arg Thr Leu His Val Gln Glu Pro Ser Pro			
2530	2535	2540	
His Val Asp Trp Ser Ser Gly Gly Val Arg Leu Leu Thr Glu Ala Val			
2545	2550	2555	2560
Pro Trp Pro Glu Thr Gly Arg Ala Arg Arg Ala Gly Val Ser Ser Phe			
2565	2570	2575	
Gly Val Ser Gly Thr Asn Ala His Ile Ile Leu Glu Gln Ala Pro Pro			
2580	2585	2590	
Glu Glu His Asp Asp Pro Ala Asp Val Ser Ser Gly Ser Phe Pro Trp			
2595	2600	2605	
Met Val Ser Ala Lys Ser Glu Gln Ala Leu Gln Ala Gln Ala Ala Gln			
2610	2615	2620	
Leu Arg Ala Tyr Leu Ala Ala His Pro Glu Leu Gly Leu Ala Asp Val			
2625	2630	2635	2640
Gly Tyr Ala Leu Ala Ser Gly Arg Thr Ala Phe Gly His Arg Ala Val			
2645	2650	2655	
Leu Leu Gly Pro Asp Arg Glu Ala Phe Val Glu Glu Leu Gly Ala Leu			
2660	2665	2670	
Glu Ala Gly Glu Glu His Ala Gly Leu Val Arg Gly Val Ala Thr Gly			
2675	2680	2685	
Ala Gly Lys Leu Ala Phe Val Cys Ser Gly Gln Gly Thr Gln Arg Pro			
2690	2695	2700	
Arg Met Gly His Gly Leu Tyr Tyr Ala Phe Pro Leu Phe Ala Ala Ala			
2705	2710	2715	2720
Met Asp Glu Ala Cys Ala His Leu Asp Pro His Leu Asp His Pro Leu			
2725	2730	2735	
Arg Asp Val Met Phe Ala Glu Pro Gly Thr Asp Thr Ala Gln Leu Leu			
2740	2745	2750	
His Gln Thr Arg Tyr Ala Gln Pro Ala Leu Phe Ala Leu Gln Ile Ala			

2755 2760 2765
 Leu His Arg Leu Val Thr Glu His His Gly Leu Thr Pro His Tyr Tyr
 2770 2775 2780
 Ala Gly His Ser Leu Gly Glu Ile Thr Ala Ala His Leu Ala Gly Ile

 2785 2790 2795 2800
 Leu Thr Leu Pro Asp Ala Ala Arg Leu Val Thr Thr Arg Ala Arg Leu
 2805 2810 2815
 Met Gln Ser Leu Pro Ala Thr Gly Ala Met Thr Thr Leu Gln Ala Asp
 2820 2825 2830
 Pro Asp Glu Leu His Glu His Leu Thr Arg Cys Glu Gly Arg Val Ser
 2835 2840 2845
 Leu Ala Ala Val Asn Ala Pro Gly Ser Val Val Ile Ser Gly Asp Arg
 2850 2855 2860

 His Asp Val Asp Ala Thr Ala Glu Asn Leu Arg Ala Met Gly Arg Lys
 2865 2870 2875 2880
 Thr Thr Ala Leu Lys Val Ser Gly Ala Phe His Ser His His Ile Asp
 2885 2890 2895
 Pro Leu Leu Asn Glu Leu Arg Asn Thr Ala Glu Thr Leu Thr Tyr His
 2900 2905 2910
 Pro Pro His Thr Pro Leu Ile Thr Thr Asn Pro Thr Asp His Asp Pro
 2915 2920 2925
 Thr Thr Pro His Tyr Trp Val Arg Gln Ala Arg Glu Thr Val His Tyr

 2930 2935 2940
 Ala His Thr Thr Gln Gln Leu His Thr His Gly Val Thr Ala Tyr Leu
 2945 2950 2955 2960
 Glu Leu Gly Pro Asp His Thr Leu Thr Ala Leu Thr His His Asn Leu
 2965 2970 2975
 Pro Asp His Thr Pro Leu Ala Val Pro Leu Leu His Pro Asp Gln Ser
 2980 2985 2990
 Glu Thr His Thr Thr His Thr Ala Leu Ala His Leu His Thr His Gly
 2995 3000 3005

His Pro Thr Thr Trp His His His His Thr Pro Thr His Tyr His Pro
3010 3015 3020

Asn Leu Pro Thr Tyr Pro Phe Gln His His His Tyr Trp Leu Asn Thr
3025 3030 3035 3040

Thr Thr Ala Thr Gly Asp Met Ser Ala Ala Gly Leu Glu Pro Ala Arg
3045 3050 3055

His Pro Leu Leu Gly Ala Ala Val Gly Leu Ala Asp Gly Glu Gly Leu
3060 3065 3070

Leu Phe Thr Gly Arg Ile Ser Leu Arg Thr His Pro Trp Leu Ala Asp
3075 3080 3085

His Ala Val Gly Gly Ala Val Leu Leu Pro Gly Thr Ala Phe Leu Glu
3090 3095 3100

Leu Ala Leu Gln Ala Ala Ala His Ala Asp Cys Arg Arg Val Glu Glu
3105 3110 3115 3120

Leu Thr Leu His Thr Pro Leu Val Val Pro Asp Ser Ala Gly Val Val
3125 3130 3135

Leu Gln Val Thr Val Ala Ala Pro Asn Glu Ala Gly Asn Arg Ala Val
3140 3145 3150

Asp Ile Tyr Ser Arg Ile Asp Val Gly Gly Leu Thr Ala Asp Ser Ala
3155 3160 3165

Gly Glu Pro Trp Thr Arg His Ala Ala Gly Tyr Leu Ala Asp Lys Pro
3170 3175 3180

Asp Pro Asp Cys Gly Asp Ser Ala Asp Gly Val Met Pro Ala Gly Ala
3185 3190 3195 3200

Trp Pro Pro Pro Gly Ala Val Ala Val Asp Leu Glu Gly Leu Tyr Glu
3205 3210 3215

Gln Leu Ala Glu Gly Gly Phe His Tyr Gly Ala Ala Phe Arg Cys Leu
3220 3225 3230

Asp Ala Ala Trp Gln Arg Gly Asp Glu Val Phe Ala Thr Ala Tyr Met
3235 3240 3245

Ser Glu Asp Gln Leu Gly Asp Thr Ala Ala Ala Arg Phe Ala Leu His
3250 3255 3260

Pro Ala Leu Leu Asp Ser Ala Leu His Thr Ile Pro Leu Leu Pro Ser
 3265 3270 3275 3280
 Leu Arg Gly Gln Gln Asp Ser Gly Leu Pro Phe Thr Trp Thr Gly Val
 3285 3290 3295
 Thr Leu Arg Ala Ser Gly Ala Thr Ala Leu Arg Val Arg Leu Arg Pro
 3300 3305 3310
 Asp Gly His Gly Pro Gly Ala Val Ser Val Asp Val Ser Asp Glu Ala
 3315 3320 3325
 Gly Glu Pro Val Ala Ser Val Arg Ser Leu Ala Leu Arg Pro Val Thr
 3330 3335 3340
 Arg Ala Glu Leu His Thr Ala Glu Leu Arg Thr Ala Ala Pro Val Ala
 3345 3350 3355 3360
 Pro His Gly Ser Leu Phe Glu Val Arg Trp Glu Pro Val Pro Gln Pro
 3365 3370 3375
 Ser Ala Ala Glu Glu Ala Ala Pro Trp Val Met Ile Gly Thr Gly Pro
 3380 3385 3390
 Thr Leu Arg Pro Val Glu Asp Phe Val Thr Pro Pro Glu Arg Thr Tyr
 3395 3400 3405
 Ala Asp Leu Ala Ala Leu Cys Val Ala Ile Ala Asp Asp Ala Pro Val
 3410 3415 3420
 Pro Arg Thr Val Val Ala Trp Ser Pro Ala Gly Ser Glu Asp Glu Ser
 3425 3430 3435 3440
 Ser Glu Ala Leu Arg Gln Ala Thr His His Met Leu Gly Leu Leu Gln
 3445 3450 3455
 Gln Trp Leu Ala Asp Ser Arg Phe Ala Asp Ser Arg Leu Val Ile Leu
 3460 3465 3470
 Thr Arg Ala Ala Val Ala Thr Ala Pro Asp Glu Glu Val Glu Asp Leu
 3475 3480 3485
 Ala Gly Ala Ala Ala Arg Gly Leu Ile Arg Ser Ala Gln Ser Glu His
 3490 3495 3500
 Pro Asp Arg Phe Val Leu Leu Asp Leu Asp Asp Arg Pro Ala Asp Ala

3505 3510 3515 3520
 Lys Asp His Asp Arg Met Leu Ser Met Ala Leu Ala Cys Gly Glu Pro
 3525 3530 3535
 Glu Val Ala Val Arg Asp Gly Ala Leu Arg Thr Pro Arg Leu Ser Pro
 3540 3545 3550
 Leu Ala Gly Thr Ala Thr Glu Ala Met Asp Glu His Pro Trp Asp Gln
 3555 3560 3565
 Asp Gly Thr Val Leu Ile Thr Gly Gly Thr Gly Ser Leu Gly Ala Met
 3570 3575 3580

 Leu Ala Arg His Leu Val Ala Thr His Gly Val Arg His Leu Met Leu
 3585 3590 3595 3600
 Ile Ser Arg Arg Gly Leu Asp Ala Pro Gly Ala Arg Arg Leu Gly Val
 3605 3610 3615
 Glu Leu Ala Glu Leu Gly Ala Gln Val Thr Ile Thr Ala Cys Asp Ala
 3620 3625 3630
 Ala Asp Gln Arg Gln Leu Ala Asn Val Leu Ser Glu Ile Ser Val Asp
 3635 3640 3645
 His Pro Leu Thr Ala Val Val His Ala Ala Gly Val Leu Asp Asp Gly

 3650 3655 3660
 Val Ile Thr Ser Leu Thr Pro Glu Gly Leu Thr His Val Leu Arg Ala
 3665 3670 3675 3680
 Lys Val Asp Ser Ala Leu Asn Leu His Gln Leu Thr Arg Asp Leu Pro
 3685 3690 3695
 Leu Ser Ala Phe Val Leu Phe Ser Ser Leu Ala Gly Val Met Gly Ser
 3700 3705 3710
 Ala Gly Gln Gly Asn Tyr Ala Ala Ala Asn Ala Ala Leu Asp Ala Leu
 3715 3720 3725

 Ala Ser His Arg Arg Ala Ala Arg Leu Pro Ala Val Ser Leu Ala Trp
 3730 3735 3740
 Gly Val Trp Glu Gln Thr Glu Gly Met Thr Gly Gln Leu Glu Ala Thr
 3745 3750 3755 3760
 Asp His Ala Arg Leu Arg Arg Ser Gly Leu Arg Pro Leu Ala Ile Ser

3765	3770	3775	
Glu Gly Leu Glu Leu Phe Asp Lys Ala Leu Ser Cys Gly His Ala Leu			
3780	3785	3790	
Val Val Pro Ala Ala Leu Ser Thr Arg Glu Leu Gln Thr Ser Gly Ser			
3795	3800	3805	
Val Pro Pro Phe Leu Arg His Leu Thr Gly Val Ala Pro Ala Arg Pro			
3810	3815	3820	
Ser Arg Thr Arg Asp Ala Ser Ala Gly Glu Pro Thr Ser Leu Arg Arg			
3825	3830	3835	3840
Arg Leu Thr Gly Leu Gly Pro Glu Glu Arg Leu Arg Glu Val Leu Arg			
3845	3850	3855	
Leu Val Arg Ser Arg Ala Ala Ala Val Leu Gly His Gly Thr Ala Glu			
3860	3865	3870	
Ser Val Pro Ala Asp Ser Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu			
3875	3880	3885	
Ala Ala Val Asp Leu Arg Asn Arg Leu Gln Gln Ala Thr Gly Leu Arg			
3890	3895	3900	
Leu Pro Ala Gly Leu Ile Phe Asp Arg Pro Arg Pro Asp Val Leu Ala			
3905	3910	3915	3920
Arg Phe Leu Cys Asp Glu Leu Ala Gly Ala Gly Gly Thr Ser Ala Ala			
3925	3930	3935	
Thr Ala Ala Pro Pro Val Ala Ala Val Gly Gly Ala Ala Gly Glu Pro			
3940	3945	3950	
Val Ala Ile Val Gly Met Ala Cys Arg Phe Pro Gly Gly Val Arg Ser			
3955	3960	3965	
Ala Glu Gly Leu Trp Asp Leu Val Ala Ser Gly Met Asp Ala Val Gly			
3970	3975	3980	
Asp Phe Pro Ala Asp Arg Gly Trp Glu Val Glu Arg Leu Tyr Asp Pro			
3985	3990	3995	4000
Asp Pro Asp Arg Thr Gly Thr Ser Tyr Thr Arg Gln Gly Gly Phe Leu			
4005	4010	4015	

Tyr Asp Ala Gly Glu Phe Asp Ala Ala Phe Phe Gly Ile Gly Pro Arg
 4020 4025 4030
 Glu Ala Val Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Ile Ser
 4035 4040 4045
 Trp Glu Ala Leu Glu Arg Ala Gly Ile Asp Pro Ala Ser Leu Arg Gly
 4050 4055 4060
 Ser Ser Thr Gly Val Phe Ala Gly Val Met Tyr His Asp Tyr Gly Thr
 4065 4070 4075 4080
 Arg Leu Arg Glu Ile Pro Glu Gly Tyr Glu Gly Tyr Ile Gly Asn Gly
 4085 4090 4095
 Asn Ala Gly Ser Val Ala Ser Gly Arg Val Ala Tyr Thr Phe Gly Leu
 4100 4105 4110
 Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val
 4115 4120 4125
 Ala Leu His Leu Ala Cys Gln Ala Leu Arg Ser Gly Glu Cys Ser Met
 4130 4135 4140
 Ala Leu Ala Gly Gly Val Thr Val Met Ser Thr Pro Thr Thr Phe Val
 4145 4150 4155 4160
 Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser
 4165 4170 4175
 Phe Gly Ala Gly Ala Asp Gly Thr Gly Trp Ala Glu Gly Ala Gly Met
 4180 4185 4190
 Leu Leu Val Glu Arg Leu Ser Asp Ala Arg Arg Asn Gly His Arg Val
 4195 4200 4205
 Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn
 4210 4215 4220
 Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Glu Arg Val Ile Arg Gln
 4225 4230 4235 4240
 Ala Trp Ala Asn Ala Gly Val Ala Ala Met Asp Ile Asp Ala Val Glu
 4245 4250 4255
 Gly His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile Glu Ala Gln Ala
 4260 4265 4270

Leu Leu Gly Thr Tyr Gly Gln Gly Arg Ser Ala Asp Arg Pro Leu Trp
 4275 4280 4285
 Leu Gly Ser Ile Lys Ser Asn Val Gly His Thr Gln Ala Ala Ala Gly
 4290 4295 4300

 Val Gly Gly Val Ile Lys Met Val Met Ala Met Arg His Gly Leu Leu
 4305 4310 4315 4320
 Pro Gln Thr Leu His Ala Glu Glu Pro Ser Pro His Val Asp Trp Ser
 4325 4330 4335
 Gly Gly Thr Val Arg Leu Leu Thr Glu Ser Val Ala Trp Pro Glu Gln
 4340 4345 4350
 Gly Arg Met Arg Arg Ala Gly Val Ser Ser Phe Gly Val Ser Gly Thr
 4355 4360 4365
 Asn Ala His Val Ile Leu Glu Gln Ala Pro Pro Ala Ala Glu Thr His

 4370 4375 4380
 Glu Pro Ala Glu Pro Asn Thr Ala Pro Gly Pro Leu Pro Trp Ala Ile
 4385 4390 4395 4400
 Ser Ala Lys Ser Pro Gln Ala Leu Arg Ala Gln Ala Arg Gln Leu His
 4405 4410 4415
 Thr Tyr Leu Thr Asn Ala Pro Glu Ala Asn Pro Ala Asp Val Gly His
 4420 4425 4430
 Thr Leu Ala Thr Gly Arg Ala Ser Phe Glu His Arg Ala Val Val Ile
 4435 4440 4445

 Gly Ser Asp Arg Ala Glu Phe Leu Gly Gly Leu Asp Ala Leu Ala Ala
 4450 4455 4460
 Asp Glu Ala His Thr Ala Val Val Thr Gly Ile Ala Arg Lys Ala Gly
 4465 4470 4475 4480
 Asp Gln Gly Lys Val Val Phe Val Phe Pro Gly Gln Gly Gly Gln Trp
 4485 4490 4495
 Ala Gly Met Gly Leu Arg Leu Leu Lys Thr Ser Pro Val Phe Ala Gln
 4500 4505 4510
 Ser Ile Gln Ala Cys Glu Gln Ala Leu Ala Pro His Thr Asp Trp Thr

4515 4520 4525
 Leu Thr Asp Ile Leu His Arg Pro His Thr Asp Pro Leu Trp Gln Arg
 4530 4535 4540
 Ala Asp Val Ile Gln Pro Val Leu Phe Ala Leu Met Thr Ser Leu Ala
 4545 4550 4555 4560
 Ala Leu Trp Gln Ser His Gly Leu Asn Pro Asp Ala Val Ile Gly His
 4565 4570 4575
 Ser Gln Gly Glu Ile Thr Ala Ala His Ile Ser Gly Ala Leu Ser Leu
 4580 4585 4590

 Glu Asp Ala Ala Lys Thr Val Ala Leu Arg Ser Arg Ala Leu Gln Thr
 4595 4600 4605
 Leu Arg Gly Ser Gly Gly Met Ala Ser Val Pro Leu Pro Ala Asp Gln
 4610 4615 4620
 Val Thr Gly Leu Leu Gln Thr Met Trp Pro Asp Arg Leu Trp Val Ala
 4625 4630 4635 4640
 Ala Val Asn Ala Pro Thr Ala Thr Val Ile Ser Gly Asn Ala Glu Ala
 4645 4650 4655
 Leu Thr Gln Ala Leu Glu His Tyr Arg Asp Gln Gly Val Asp Ala Lys

 4660 4665 4670
 Arg Ile Pro Val Asp Tyr Ala Ser His Cys Pro His Ile Gln Ala Val
 4675 4680 4685
 Glu Gln Glu Leu Ser Arg Leu Leu Arg Gly Ile Thr Pro Arg Ala Ala
 4690 4695 4700
 Thr Thr Pro Phe Tyr Ser Thr Thr Asp Asn Gln Trp Thr Asp Thr Thr
 4705 4710 4715 4720
 Thr Leu Asn Ala His Tyr Trp Tyr Arg Asn Leu Arg Gln Pro Val His
 4725 4730 4735

 Leu Ala Asp Ala Ile Thr Asn Leu Thr His Gln Gly His His Thr Phe
 4740 4745 4750
 Ile Glu Ile Ser Pro His Pro Thr Leu Thr Pro Ala Ile Gln Glu Thr
 4755 4760 4765
 Thr Asp Thr Thr His Thr Pro Thr Thr Val Ile Ser Thr Leu Arg Arg

4770 4775 4780
 Asn His Asn Asp Thr His Gln Ile Leu His Ala Leu Ala His Ala His
 4785 4790 4795 4800
 Thr Thr Gly His Pro Ile Asn Trp His Thr Thr His Gln His His Thr

 4805 4810 4815
 Pro Thr Pro Gln His Ile Asp Leu Pro Thr Tyr Pro Phe Gln His His
 4820 4825 4830
 His Tyr Trp Leu Asn Thr Pro Thr Gln Thr Gly Asp Ala Ala Ala Val
 4835 4840 4845
 Gly Leu Asp Pro Ala His His Pro Leu Leu Gly Ala Ala Val Ala Val
 4850 4855 4860
 Ala Glu Gly Glu Gly Tyr Leu Leu Thr Gly Arg Leu Ala Leu Ser Thr
 4865 4870 4875 4880

 His Pro Trp Leu Ala Asp His Thr Ile Ala Gly Ala Val Val Leu Pro
 4885 4890 4895
 Gly Thr Ala Leu Leu Glu Ile Ala Leu Gln Ala Gly His Arg Val Asp
 4900 4905 4910
 Cys Trp Arg Ile Glu Glu Leu Thr Leu Gln Ser Pro Leu Phe Ile Pro
 4915 4920 4925
 Glu Glu Gly Ala Val Gln Val Gln Ala Trp Val Ala Ala Pro Asp Glu
 4930 4935 4940
 Asn Gly Cys Arg Ser Leu Thr Val Ser Ser Arg Arg Glu Gly Thr Tyr

 4945 4950 4955 4960
 Glu Asp Ala Thr Trp Val Arg His Ala Thr Gly Arg Val Gly Pro Ala
 4965 4970 4975
 Pro Ala Asp Gln Asp Glu Ala Ile Ala Arg Leu Thr Asp Pro Gln Gly
 4980 4985 4990
 Asp Gly Ala Ala Ala Ala Val Trp Pro Pro Gln Gly Ala Val Ala Phe
 4995 5000 5005
 Thr Ala Asp Asp Leu Glu Gly Leu Tyr Asp Gly Tyr Ala Ala Arg Gly
 5010 5015 5020

Phe Glu Tyr Gly Pro Val Phe Arg Gly Leu Arg Ala Ala Trp Arg Arg
 5025 5030 5035 5040
 Gly Glu Asp Ile Phe Ala Glu Val Arg Leu Pro Asp Thr Ala Asp Gly
 5045 5050 5055
 Asp Ala Ser Gln Phe Ser Val His Pro Ala Leu Leu Asp Ala Ala Leu
 5060 5065 5070
 His Ala Ala Ala Phe Arg Pro Ala Asp Lys Leu Pro His Gly Ala Leu
 5075 5080 5085
 Pro Phe Ser Phe Ser Gly Val Arg Leu His Gly Pro Gly Ala Ser Thr

 5090 5095 5100
 Leu Arg Val Arg Leu Thr Pro Asp Gly Gln Ala Arg Asp Thr His Ala
 5105 5110 5115 5120
 Trp Ser Val Ala Val Val Asp Gly Glu Gly Arg Pro Val Ala Ser Ile
 5125 5130 5135
 Ala Ser Leu Ala Val Arg Pro Val Ser Thr Gln Glu Leu Leu Ala Ala
 5140 5145 5150
 Ser Gly Thr Ala Arg Arg Asp Ser Leu Phe Ala Val Glu Trp Val Thr
 5155 5160 5165

 Ala Leu Ala Pro Thr Ser Ser Ser Val Pro Gln Arg Leu Ala Thr Val
 5170 5175 5180
 Gly Pro Ser Asp Arg Leu Pro Ser Ala Asp Ala Tyr Ala Asn Leu Ala
 5185 5190 5195 5200
 Asp Leu Ala Ala Ala Val Leu Glu Ala Gly Ala Pro Ala Pro Asp Ala
 5205 5210 5215
 Val Val Val Asp Cys Gly Arg Arg Asp Ala Arg Ala Thr Ala Val Pro
 5220 5225 5230
 Glu Asp Val Arg Thr Leu Thr Arg Arg Ile Leu Gly Leu Leu Gln Glu

 5235 5240 5245
 Trp Leu Ala Asp Glu Arg Pro Ala Ser Ser Arg Met Val Val Leu Thr
 5250 5255 5260
 Arg Gly Ala Val Ala Thr Thr Pro Gly Glu Asp Val Ala Asp Leu Ala
 5265 5270 5275 5280

Gly Ala Ala Val Cys Gly Met Val Arg Ser Ala Gln Ser Glu His Pro
5285 5290 5295

Gly Arg Phe Val Leu Leu Asp Leu Asp Pro Asp Pro Asp Leu Asp Gly
5300 5305 5310

Gly Glu Val Pro Pro Thr Val Val Pro Ala Ala Leu Ala Cys Gly Glu
5315 5320 5325

Pro Gln Ile Ala Val Arg Ala Asn Arg His Leu Val Pro Arg Leu Thr
5330 5335 5340

Arg Val Pro Ala Ser Val Pro Val Pro Gly Arg Val Pro Val Pro Ala
5345 5350 5355 5360

Ala Glu Ala Ala Asp Pro Asp Thr Thr Pro Thr Ala Phe Asp Pro Asp
5365 5370 5375

Gly Thr Val Val Ile Thr Gly Gly Thr Gly Thr Leu Gly Ala Met Leu
5380 5385 5390

Ala Arg His Leu Val Ser Arg His Gly Val Arg His Leu Leu Leu Ala
5395 5400 5405

Ser Arg Arg Gly Pro Asp Ala Pro Gly Ala Thr Glu Leu Arg Ala Glu
5410 5415 5420

Leu Ala Glu Leu Gly Ala Glu Val Thr Val Arg Ala Cys Asp Thr Gly
5425 5430 5435 5440

Asp Arg Gly Ala Leu Ala Asp Leu Ile Ala Gly Ile Pro Thr Gly His
5445 5450 5455

Pro Leu Thr Gly Val Val His Ala Ala Gly Val Leu Asp Asp Ala Thr
5460 5465 5470

Val Ala Ser Leu Thr Pro Arg His Leu Asp Thr Ala Leu Thr Pro Lys
5475 5480 5485

Ala Asp Ala Ala Phe His Leu His Glu Leu Thr Arg His Ala Arg Pro
5490 5495 5500

Arg Ala Phe Val Leu Phe Ser Ser Ala Ala Gly Val Leu Gly Ala Ala
5505 5510 5515 5520

Gly Gln Gly Asn Tyr Ala Ala Ala Asn Ala Phe Leu Asp Ala Leu Ala

5525 5530 5535
 Glu His Arg Arg Ala Gln Gly Leu Pro Ala Leu Ser Leu Ala Trp Gly
 5540 5545 5550
 Leu Trp Glu Gln Gly Ser Gly Met Thr Gly His Leu Asp Arg Thr Asp
 5555 5560 5565
 Arg Ala Arg Ile Asn Arg Ser Gly Leu Ala Pro Leu Ala Thr Glu Asp
 5570 5575 5580
 Ala Leu Ala Leu Phe Asp Ala Ala Leu Ala Gly Asp Arg Pro Phe Leu
 5585 5590 5595 5600

 Val Pro Ala Arg Leu Asp Leu Arg Gly Ser Ser Ala Ala Glu Thr Pro
 5605 5610 5615
 Ala Pro Leu Phe Ser Arg Ile Ala Pro Ala Arg Thr Thr Arg Gly Arg
 5620 5625 5630
 Ser Pro Gly Ala Glu Gly Ala Ala Asp Leu Arg Thr Arg Leu Ala Ala
 5635 5640 5645
 Gln Asp Ala Ala Glu Gln Arg Asp Thr Leu Leu Thr Ile Val Arg Thr
 5650 5655 5660
 His Thr Ala Ala Val Leu Gly His Asp Thr Ala Ala Ala Val Arg Pro

 5665 5670 5675 5680
 Asp Gly Ala Phe Arg Glu Leu Gly Phe Asp Ser Leu Ala Ala Val Glu
 5685 5690 5695
 Leu Arg Asn Arg Leu Gln Thr Thr Thr Ala Leu Thr Leu Pro Ala Thr
 5700 5705 5710
 Thr Val Phe Asp His Pro Thr Pro Ala Ala Leu Ala Asp His Leu Arg
 5715 5720 5725
 Thr Gln Leu Cys Gln Asp Ala Gln Ser Ser Ala Ala Ala Thr Ala Met
 5730 5735 5740

 Ala Ala Met Ala Glu Leu Ala Arg Leu Glu Ser Ala Val Ser Asp Ser
 5745 5750 5755 5760
 Val Ala Leu Asp Asp Asp Thr Arg Ser Gly Leu Ala Glu Arg Leu Arg
 5765 5770 5775
 Ser Leu Ala Arg Lys Met Ser Ser Gly Arg Val Val Asp His Asp Gly

5780	5785	5790
Gly Gly Ala Ala Asp Leu Asp Leu Gln Ser Val Thr Asp Asp Glu Met		
5795	5800	5805
Phe Glu Leu Ile Asp Lys Glu Val Ser Arg Asp		
5810	5815	
<210>	75	
<211>	17469	
<212>	DNA	
<213>	Artificial Sequence	
<220><223>	meiA3 gene of Streptomyces nanchangensis	
<400>	75	
ttggagatac cgatggccgc tggccacgac aaggtgatcg aggcgctgcg ggcgtccctc		60
aagaccaacg agcggcagag ggaacagatc caccggctca ctacggcggc gcgggaaccc		120
atgccatca tcggcatggc ctgccgctat cccggcggag tgggatcgcc ggaggacctg		180
tgggagctgg tggccgccgg ccgtgacgcc atcggcacct tccccagga cggggctgg		240
gacgcggcgc ggcgtgacga ccccgatccg gagcgggccg gcacctcgta caccagcat		300
ggcggattcc tttaccagc aggggagtgc gacccgggtt tcttcgggat cagcccgcgc		360
gaggcgtg cgatggaccc gcagcagcgg ctgctgtgag agatctcctg ggaggcggtc		420
gagcgggccg ggatcgaccc ggcctcggtg cgcggcagcc gcaccggggt cttcgcgggc		480
gtcatgtacc acgactacgg ctcccggctg cacaccgtcc ccgaaggctt cgagggtctac		540
gtcggcaacg gcagcggcgg cggcgtggcg tccggccggg tcgcctacac cctcggcctc		600
gaaggcccgg ccgtgaccgt ggacaccgcc tgcctcctct cactggtcgc cctgcacctg		660
gcctgccagg cgctgcgggc cggcgagtgc tcactcgccc tcgcgggcgg ggtgacggtg		720
atgtccaccc ccagcctgtt cgctcagtag tcccggcagc gcgcgctcgc ggcggatggc		780
cggtgcaagg cgtacggggc gggggcggac ggcaccggct gggcagaagg cggcgggatg		840
ttgctggtgg aacggctcac ggacgcacag cggctcggcc accgggtgct ggcagtggtc		900
cggggcagcg cggtaacca ggacggcgcg agcaacggcc tcaccgcccc caacggcccc		960
gcgcagcaac gggatcatccg gcaggcactg gcgagcggcg ggggtgcggc gtccgaggtc		1020
gacgccgtgg aggggcatgg gacggggacg cggctgggcg atccgatcga ggcgcaggcg		1080
ttgctggcga cctacggta gcagcggccc gcggaccggc cgctgtggct cgggtcgatg		1140
aagtccaacg tcggccatgc gcaggcggcc gccggcgtgg gcgggatcat caagatggtg		1200

atggccatgc ggagcgggac gctgccgcgc accctgcacg cggacgagcc gtcgccgcac	1260
atcgactggg actcgggcgc ggtgcggctg ctgaccgagc cggtcgcctg gccggagcgc	1320
gaccggccgc gccgcgccgc ggtgtcctcc ttcggggtca gcggcaccaa cgcccatgtg	1380
atcctcgagg ccgcgtcgca gacggcgccg cacacggaat ccgcgtcgca gacggaaacc	1440
gacgacgtc ccgcgccga cggcgccgcg ggccatgccg tggcggggcc gctgccctgg	1500
ccccgtcgg gcgcgacgc cgaggcgctg cgggccagg ccaggagct gcgtcgttc	1560
gtggcgccg atgagctgt gcgccccgc gacgtcggc acacctggt ctgggccgc	1620
tcggacctc cacaccgc agtcgtctc ggctccgacc gggaaacct gctgcgcgt	1680
ctggacactc tgacaggga gggccggac ggcggtcgg tcgtacggg cgtggcgcg	1740
gcaggggccg gtgcggcgt ggtgttcgtc tccccggac agggcgcca gtggccggc	1800
atgggctgc ggctgtgga gacctgtc ttttcgcg agcgatggc ggagtgcgag	1860
gcggccttg caccgtatgt cgactgtc ctgctcagc tgctgcgcg ggacccggg	1920
gacccggtgt gggagcggc cgatgtcgc cagccgatgc tgttctcgt gatggtgtc	1980
ctggcgagc tgtggcgtc gtatggcgtc gaaccggac ccgtactcg cactcccag	2040
ggcgagatc ccgcccca catctgcgc gcctgacc tggacgagc cgcgaaggt	2100
gtgccctgc gcagccggc cctgcagacc ctgcgcggt cggcgccat ggcctccgt	2160
ccactgacg cggacgaggt ccgccggctg ctgcggacc catggccga ccggtgtg	2220
gtggccgcc tcaaccccc caggccacg gtgatctcc gcgacgcga ctctctggc	2280
gaggcgttg aactactac ggaccaggc gtgcagcga agcgggtccc ggtcgactac	2340
gcctccact gcccgatat cgaggccgt gagcaggagc tgctgagct gttcggggg	2400
atcgctcaa gggccgccga cattccctc tactccactg tggacaacca gtggccgac	2460
accatgggac tcgacggcg gtactggtac cgcaatctg gccggccgt acgcttcgc	2520
gaagcgctc gcgtctcgg tgccgccgag taccggtcgt atgtcgaggt cggccgcac	2580
cccacctca ccccgccat cgaggacacc actgaggcc cggcgccgc ggccacggt	2640
gtcggtccc tgcgcgcgg cgaggacgac gccaccgca tctcacctc gctggccgg	2700
gtcatattc atggcctgcc cgtggcgtg gaccgccact accggcgct cggcccgag	2760
gcgaaccatg tcgacctgc cactacgcc ttccagcgc gccgtactg gctggacgc	2820
ccggcgacca ccggggacgt gacggccgc gggtggccc cggtcggaca cccactgctc	2880
ggcgcgccg tcggactgc cgaggcgac ggatatctg tcaccggcc gctcgcccc	2940
cacaccacc cctggctcac cgaccacgc gtgcgggca ccgtcctgt gccggcacc	3000
gcatactgg aactggcgt gcacgtcgc gaacacctg gctgccccg gctggaggag	3060

ctcacctgc acgccccgct cgtctctccc gacacgggcg gtgtggcgct ccaggtggcc	3120
gtcggcgcac cggacgagac cggccggcgc gactgagcg tctacgaca gcgcgacgac	3180
gacccacgt gggaaggggc ggcccggggc gctgggacac ggcatgacgac cggcacactg	3240
gcggccgagg ccgcgaccga tggcatcaac ggtgccgacg gtgccgggcc cctggcgggg	3300
gcgtggcctc cgccggggcg ggagcccctg gacatcagcg gcctctacga cacgtggcc	3360
gccgcggact tcggctacgg cccggccttc caggggctgc gcgcgctctg gcggcacggc	3420
gaggagacct acgccaggt gcggtctccc gaccaggtgg ccgccgacgc cccacgttc	3480
tgctccacc ccgcttgct cgacgccgcg ctccaccgc tggcactcga cagcgccga	3540
agcaggaga atccagcggg acatggcctg ctccattcg cctggcgagg cgtcagcctg	3600
cgctccccgg gcacaccgac gctgcgcgta cggctgcggc cgcagggccc ggactcgatt	3660
gccgtcgacg tggccgacga gacgggcgcg tcggtggtct cggccgaatc gctcacgctg	3720
cgaccgttg ccttgagga cctgcgggtc ctggcgggc aggcgaacga cccctctat	3780
gccctggagt ggaccgccgc gcccagccc ctgacaacag ccctcgggcg gtgcgccgtg	3840
cttggccacg ccacccccg atgggccgccc gcgttgagga cggcggcagc ggagcccga	3900
cggcggtacc cggacctgc cggactggtg gcggccctgg acgccggcga tccgcctccg	3960
gacctggtgt tcgtgggctg cctccggct gccgccggc ccgacgacac gacggtgcc	4020
gacgttcaca ccaccgtac ccgtgtcgt acccgacaag cgtggagct gcttaaggc	4080
tggtcggcg aagcgcggct ggccggcgcg cggctggtgc tggtcacccg cggcgcggtg	4140
gccaccgggc cggcggggg agggatggac ctggcgggcg cggcgatctg cggactggtg	4200
cgatcgcac aggccgagga gcccgatcg atcctcttg tggacttga cacggccgag	4260
gagtcgtgg cggcgctgcc acggcggtc gcgtgggcg aaccgagat ggccatccg	4320
gccggccagc cgcacatggc ccggtggtg cgagccgaca ccgagaggga cggcctgctc	4380
acgccccac gggggagcgg cggctggcgg ctgactgcg ccgatgcggg cacgtccag	4440
gggttggcg cggtggcgtc ctggccgac cagacccgc tgggcccga gcaggtacgg	4500
atcaggtgc gtgcggccg gctgaacttc cgcgatgtcc tggtgccct ggggatggtc	4560
cctgggcagc aggggtggg cagcgaggc gccgggtgg tgctcgaagc cgggcctgaa	4620
gtggccgacc tggcggggc agaccgggtg atggcggtgt tcgcggacgc gttcgcccc	4680
ttcgcatcg ccgaccggc cacagtgat cgcgtcccc agcactggac cttcgcccag	4740
gccgcccg tccccgtgt gttcgccacc gcctactac ggctggtgga cctggcagga	4800
ctgcgccgg gcgagtcgt gctggtgcac gcccgggccg gcggagtggg actggccgc	4860

gtccaactgg cccgccacct gggcgctgag gtctacgcca cggcgagccc cggcaaatgg	4920
gacaccctac gcgcccacgg catccccccg gagcgcatcg cctcgtcccg caccctcgac	4980
ttcgagagcc ggttcaccgg ccggaacatc gacgtcgtcc tcaactccct ggcccatgag	5040
tacgtcgacg cctcgtcgcg cctgggtgtcc ggcgacagcg gccggttcct cgagatgggc	5100
aagaccgacc tccgtgacct ggaggagggtg gcggaggcgt accccggtgt cgcctaccgg	5160
gcgtacgacc tgatggaggc cggacccgag cgcctcgggg agatcctgcg caccgtgctg	5220
cggctgttcg acgaggcggt gctcaccgcc ctgccgtca cctgctggga catccggcag	5280
gccagggatg ctttcgcca actccagcag ggccgcaccg tcggaaagaa tgtgtcacg	5340
ctggaccgca cccccagccc cgacggcacc gtctcatca ccggtggcac cggcacccctc	5400
ggcgccgcgc tcgcccga tctcgccgc accggccgag cacggcatct gctgtgatc	5460
agccgccgtg gcctcgatgc gccaggcgt cccgaactca tcgtgagat cgacgagttg	5520
ggcgccgga cgaccgtcg caccgtcgac gtccggcacc gtgccgcgt cccgaactg	5580
ctcgccgga tccccgcca gacccgctg accgccgtc tccacgccgc ggccacactc	5640
gacgagcca cgtcggctc cctcaccgcg cgcacctcg acaccgttct gcccggaag	5700
gccgatgccg cctggcatct gcacgagctg acctgccggc tggatctggc cgcgttcgtg	5760
ctgttctgt cgcgccggg cgtcttgggc tcgccgggc agggcaacta cgccgccgc	5820
aatgcctttc tcgacgcgt cgccttccag cgacgggga tgggactccc cgccgtgtcc	5880
ctggcatggg gactgtggga ggagccagc gggatgaccg gccacctga ccagaccgac	5940
cgcaccgca tggcccggt cggcctccgg cactggcca cgaacgagc cctggcgctg	6000
ttcgacaacg ctctgtga tggccaccg ctgtgtcc cgcccgtat cgacaccaag	6060
gcgtacggg gcaccaccgc accgcccctg ttccagagcc tcgtactcc caccaccgc	6120
caccggccac gcccgcgac accgacggc cgtctctccc tccgagccg gctcgcggg	6180
ctcgaccctg ccgcacagca cgaggctctg ctaccctcg tcccggccca cgccccacg	6240
gtctcgcc acccgagccc cgacgccatc gccccgagg cggccttcg tgacctggc	6300
ttcgactccc tcaccgcgt agagctccgc aaccgcctca aggaggcaac cggctcgcg	6360
ctccccgca cctcgtctt cgaccacccc actctgccg ctctcgcca gcacctgcg	6420
gacggcctca tcggcggcgc cgatcgggc accttggtt cggctctgc tccgagcgag	6480
gtggcgacgg tggcgatga ggccatcgcg atcatcgga tggcctgcc gtatccggg	6540
ggcgtgcgt cggccgaagg gctgtggat ctggtgcct ccggcaccga cgccatgagc	6600

ggattcccca ccgaccgcgg ctgggacctc gaccgcctct acgccccca ggaccaggac	6660
cggccgggca ccacatacac ccgccacggg ggcttcctcc acgacgcggg caagttcgac	6720
gcgggattct tcggcatcgg ccacacgtgag gcgctggcga tggatccaca gcagcggctg	6780
ctgctggaga cctcctggga ggttttcgaa cacgcgggaa tcgaccctc gtcggtacgg	6840
cggagccgga ccggagtctt cgccggtgtg atgccgacgg actacggccc ccggctgcaa	6900
gacaccgtgg ccgaggtcga gggctatgtc ctaccggaa actccggcag cgtcgctcg	6960
ggccgtatcg cctacacctt cggctcggaa gggccgcgg tgcgggtgga cacggcgtgt	7020
tcgtcgtctc tggtagcggt gcactcggcg tgcaggcgc tgcgtgcggg ggagtctcc	7080
atggcgtcgg ccggcggggt gacggtgatg gcgacgcctg gtgccttcgt ggagtttgcg	7140
cggcagcggg ggttgtcggg ggatgggcgg tgcaaggcgt ttggggtggg tgcggatggt	7200
acggggtggg ccgaggggggt ggggatgctg ttggtggagc ggttgtctga tgcgcggcgg	7260
ttggggcatc ggggtgttggc ggtggtgcgg ggttctcgg tgaatcagga cggggcgagc	7320
aatggtttga cggcgccgaa tggccgtcg cagcagcggg tgatccggca ggcgttggcc	7380
agtgcgcggg ttggtggggc ggatgtggat gtggtggagg ggcacggtac ggggacgcgg	7440
ctgggtgatc cgatcgaggc gcaggcgttg ctggcgacct acggtcagga gcggtcgggg	7500
gatgaaccgt tgtggttggg gtcggtgaag tcgaatatcg ggcatgcga ggctgcggcg	7560
ggtgttgcgg gtgtcatcaa gatggtgatg gcgatgcggg gtggggtgtt gcccgggacg	7620
ttgatgtgc aggagccgtc gccgatgtg gactggtcct cgggtggggg gcggctgctg	7680
acggaggcgg tccgtggcc ggagacgggt cgtgcgcggc gtgcgggggt gtcgtcgttc	7740
ggggtcagcg gcaccaacgc gcacatcatc ctgaacagg caccgccgga ggagcacgac	7800
gatccggcgg acgtttcgtc cgggtcgttt ccgtggatgg tgcggccaa gtccgaacag	7860
gcactacagg cacaggcagc gcagctgcgc gcgtatctgg cggcacgtcc cggggtgggg	7920
ctggctgatg tcgggtatgc gctggccgcc ggccgtaccg ctttcgacca ccgtgccgtg	7980
ctcctgggcc cggaccgcga agccttcctc gaagggtcgg gggctctggg gcccggtgag	8040
gaacacgccg ggctcgtacg gggcgtggcg acgggtgcgg ggaagctggc gttcgtgtgt	8100
tccgggcagg gcacgcagcg ccctcgtatg gggcacgagc tgtaccgcgc cttcccgtg	8160
ttcgcccgag ccatggacga agcctgcgca tacctggacc cgcactcga ccggcctctg	8220
cgggatgtcg tgttcgcga gccggactcc ggtacggccc ggctgctgca gcagacgcgc	8280
tatgcccage ccgcgtgtt cgccctccag gtcgcctgc atgcctggt caccgaacac	8340
tacggcctca cggccacta ctacggggc cattccctgg gggagatcac cgcggccac	8400
ctcgccggga tctgacct ctgcgacgcg gcgcgtctgg tcaccaccg cggccgctg	8460

atgcagtctc tccccgccac cggcgcgatg accaccctcc aagcagaccc cgacgaactc	8520
cacgaacacc tcgcacgatg cgaggacgg gtgtcgctcg cggccgtgaa cgcgcctggg	8580
tccgtggtea tcagcgtga cggccacgac gtagacgcca cggccgaaaa ctccgcgcc	8640
atggggcgca agaccacccc gttgaaggta agcggcgctt tccactcaca ccacatcgac	8700
ccactcctcg acgaactccg cgccaccgcc gaaacctca cctaccaccc accccacacc	8760
cccccatca cgaccgacct gaccgaccag gacccacca cactggcta ttgggtccgg	8820
caaacacgag agaccgtcca ctacggccac accaccaac aactccacac ccacggcgtc	8880
accgcctacc tcgaactcgg ccccgacacc acactacca cctcaccca ccacaacctc	8940
ccccaccaca cccccctagc catccccctc ctccaccccg accaacccga aaccacacc	9000
accacaccg ccttcgcca cctccacacc cagggccacc ccaccacctg gcaccaccac	9060
cacaccccca cccaccacca cccaaacctc cccacctacc cttccaaca ccaccactac	9120
tggctcaaca ccaccactgc caccggtgat atgtcggcgg caggccttga gccggcgagg	9180
catccctgt tgggcggggc ggtcgagttg gccgatggtg aggggttgct gttactggg	9240
cggatttca tccgtacgca tccctggttg gccgaccacg ccgtcggcgg cggcgtgttg	9300
ctccccggtc cggcctttct cgaactcgcc ctggaagccg ccgcccattg cgactgcat	9360
cggatcgagg agcttacgt ccacaccccg ctgctgtac cggagagcgg cggcgtagt	9420
ctgcaggta ccgtggcgg gccgaacgaa gcaggaaacc gggcgggtga tatctactg	9480
cgaatcgat tcggcgccct caccgccgat tcggtgggag agcgtggac gcgcatgcc	9540
gccgggtacc ttgccgaca gcctggccca gactgcggtg actcggcgga tgggtgcatg	9600
cctgcgggag catggccgcc gccgggtgag gtgcgctgg atctggagga actgtacgag	9660
cagctggcgg aggggggttt ccactacggt gggccttcc gttgcctgga cggcgctgg	9720
caacgcggcg acgaggtctt cgcgactgtg catatgtcag agaatcagct gggcgacacg	9780
gccgcggctc ggttcgcgt gcaccccgcg ctgctggatt ccgactgca caccattcca	9840
ctctccct ccttcaggga acaacaggac agcgggctgc cgttcacgt ggcaggagtc	9900
acctgcgag catccggggc caggccctg cgcgtccggc tgaggccgga tggccatggc	9960
ccggggggcg tgtccgtcga cgtgtccgac gaggcgggtg agcccgtagc atcagttcgg	10020
tcgttggccc tcggcccggt gaccaggttc gatttgata cggccgagtt gcgcacagcc	10080
gccccagttg cccccatag ctgctcttc gaggtgcgat gggaacccgt cccccagccc	10140
tcagcggcgg aagaagccga tccatgggtg atgatcggga ccggaccgac gctgcgccc	10200
gacgaggact tcgcaactcc gccggagcgg acgtacgccc acctggccgc gctgtcgcg	10260
gcagtcgccc atggcgccc cgttccccgg acggtcgtgg cctggtcca ggccgggagc	10320

gaagacgagt cgagtgaggc gctgcgccac gccacacacc acatgctggg cctactgcag	10380
cagtggtttg cgacagccg gtctctcgac agtcgcctgg tgatcctcac ccgagccgcg	10440
gtggccactg cgccggagga ggaggtaaaa gacctggcgg gagcggcgac gcggggtctg	10500
atccgctccg ccagtcgga gacccccgac cgattcgtcc tgctcgacct ggacgacct	10560
cccgctgacg cgaagacca cgaccgaatg ctgtcggtag ccctggcctg cggggaaccg	10620
gaagtggcgg tacgcatgg agccctgcgc acacccggc tgagcccgt tgcggcacc	10680
gccaccgagg ccatggacga gcatccctgg gatccggacg gcaccgtact catcacggc	10740
ggcaccggca gcctcggcgc catgtctgcc cgccacttgg tggccacca tggcgtacgg	10800
catctgctgc tgatcagccg acgtggcctc gacgccccgg gggccaggcg acaggggaac	10860
gaactcgtcg agctcggagc gcagttgacc atcgccgcgt gcgatgccgc agaccaaagg	10920
caacttgcaa acgattgtc ggagatctcc gtgcaccatc cgctgaccgc tgtggtgcat	10980
gcggcaggcg tactggacga cggggtgatc acatccctca caccggagga cctgacccat	11040
gtcctgcggg ccaaggtcga ttcggcgtc aatctccacc agctcacacg cgacctgccg	11100
ctgtccgcgt ttgtgtctt ctctcgtcg gccggggtga tgggttcggc agggcagggc	11160
aactacgccg ccgccaacgc cgccctggac gcgtggcga gtaccgcag ggccactcgg	11220
ctgccggcgg tgtccctggc ctggggagt tgggagcaga ccgaggcat gaccgggcag	11280
ttggaggcca cgggccacgc gaggtccgc cgctcgggcc tgaggccgt gccaccagc	11340
gagggccttg agctcttca caaggccttg agctgtggac acgccctgt ggtgcccgc	11400
gcactcagca cgaaggagct tcagacatcc ggatccgtcc caccattct gcgccacgtg	11460
acgggcgtcg ctccggcccc gccgtcccgg acccgcgacg cctcggccgg tgagccgacc	11520
cccctgcggc ggcggttag cggcctcggg ccggaagagc ggctacgca ggtgctcgg	11580
ctggtgcgt cccgggcggc tgcgtgctg gggcacggca cggccgaagc ggtcccggcg	11640
gactcggcgt tccgcacct ggggttcgac tcctcgcg cggtggacct gcggaaccgg	11700
ttgcagcagg ccaccgggt gcgctgccg gccggttga tcttcgaccg gccgcgtccg	11760
gacgtactcg cccgtttct gtgtgacgag ttggccggtg tcggcggtag gtcggcgcc	11820
acggccccc caccgttgc ggccgtcggc ggggcagccg gcgagccgt gccatcgtc	11880
ggcatggcat gccggttcc gggaggtgtg cggtcggccg agggcctgt ggtatgtgtc	11940
gcctccgta tggacgcgt gggtgacttc cccacagacc gaggttgga ggtggaacgg	12000
ctctacgacc ccgaccgga ccgaaccggc acctctata cccggcaagg cgggttcctc	12060

tacgacgcgg gtgagttcga cgcggcggttc ttcgggattg gcccgcgtga ggccgtggcg	12120
atggatccac agcagcggtt gctgctggag atttctggg aggcgtgga acggcgggga	12180
atcgaccggc cgtcgtcgcg ggggagttcg actggagtgt tcgctggggt gatgtaccac	12240
gactacggca cccgcttgcg cgagatccca gagggctacg agggctatat cggcaatgga	12300
aacgcgggca gcgtcgcttc gggacgtgtc tcctacactt tcggcctgga ggggcccggc	12360
gtcacgtgg acacggcgtg ttctgctgcc ctggtcgccc tgcactggc ctgccaggcg	12420
ctgcggtcag gggagtgtc catggcgctg gcgggcccggg tcacgtcat gtccaccccc	12480
accacttttg tcgagttctc gcgccaacgg ggactggccc cggacgggcg gtgcaagtc	12540
ttcggggcgg gcgcggacgg aacgggctgg gcggaggcg cgggatgct cctggtggag	12600
cggctttcgg acgcccggcg caacggccac cgggtcctgg cggtggtacg ggggagcgcg	12660
gtcaaccagg acggggcgag caatgggctg acggcgccga acggccgtc gcaagagcgg	12720
gtgatccgcc aggcgtgggc aaatgcgggt gtggccgca tggacatcga cgcgtggag	12780
ggacacggca cggggacgac gctcggtagc ccattgagg ccaggcgct gctggggacg	12840
tatggacagg gacggtcggc cgatcgccg ttgtggttgg gatcgatcaa gtccaacgtc	12900
ggacacaccc aggcccgccg gggggtgggc ggcgtcatca agatggtgat ggccatgcgc	12960
cacgggctgc tcccgcagac cctgcacgc gaggagccct cacctcatgt ggactggtcg	13020
ggcgggacgg tgcggttgc gaccgagccg gtggcctggc ctgagcgggg gcggatgcgc	13080
cgcgcaggcg tctcctcttt cgggtgtcagc ggtaccaacg cccacgtcat cttggaacaa	13140
gcaccaccta acgcgagac ccacgaaccg gcagagcccc acaccgcgc agggccactg	13200
ccctggacga tctccgcgaa gagcccga ggcgtacgtg cccaggcgcg tcagttgcac	13260
acgtacctga ccaacacccc cgaggcgaac cccgccgacg tcggccacac cctcgcgatg	13320
ggccgcgcct ctttcgagca tcgtcgggtg gttatcggct ccgatcagg ggagtttctg	13380
ggtggtctgg atgctgtggc ggcagatgag gccactctg ctgtggtcac gggtatcgcg	13440
aggaaggccg gtgacctggg gaaggtggtg ttctcttcc cgggcaggg tggtcagtgg	13500
gccgggatgg gactgcggct gctcaagacc tcgcccgtct tcgcaatc catccagcc	13560
tgcaacaag ccttcgcccc ccacaccgac tggacctga ccgacatcct gcaccgcccc	13620
cacaccgacc cctgtggca gcgcgccgac gtcattcagc ccgcccctct cgcctcatg	13680
acctccctca ccacctctg gcaatccac ggcctcaacc ccgacgccgt catcgccac	13740
tcccaaggcg aaatcaccgc cggccacgcc tgcggagcac tgagcctgga agacgccgcg	13800
aaaatcgtcg ccttcgcag ccagacctg caaacctcc aaggtcagg cggcatggcc	13860
tccgtaccac tgcccgcaga ccaggtcacc gcaactgtgc acaccatgtg gcccgaccg	13920

ctatgggtcg ccgcatcaa cgccccacc accacagtca tctccggcga cacacaagcc	13980
ctcacacaag cgctgaacca ctaccgggac caagacatcg acgcgaaacg catcccggtc	14040
gactacgcct cccactgccc ccacatccag gccgtccaac acgaactctc agacctgttg	14100
caggacatca cccacagggc cgcgaccacc cctttctact ccaccaccga caaccaatgg	14160
accgacacca ccacctcaa cgcccactac tggtagcgaa acctccgcca acccgteccac	14220
ctcaccaacg ccatacacia cctcacccac caaggccacc acacctacat cgaaatcagc	14280
ccccaccca cctcacccc cgccatccag gaaaccaccc acaccacca cccccacc	14340
accgtcatca gcacactcg cgcaaccac aacgacaccc accaactcct ccacgcctc	14400
gcccacgccc acaccacgg ccaccccatc aactggcacc ccaccacca acaccacacc	14460
ccaaccccc aacacaccga cctccccacc tacccttcc aacaccaacg ctactggctc	14520
aacaccccca cccaacagg agacgcagca gccatcggcc tggaccgggc acatcacccg	14580
ctgctcggcg cggcggtcgc agtggccgag ggggagggt atctgctcac cggtcggctc	14640
gccctgtcca cccaccctg gcttgccgat cacaccatcg cgggcgcggt cgtccttccc	14700
ggaactgccc ttcttgagat cgcccttcag gcgggccatc gtgtggactg ccatgcac	14760
gaagaactca cctccaatc gccgtgttc atcccgaag agggagcagt acaggtgcag	14820
gcatgggtgg cggcgccgga tgagaacggg taccgaagcc tgacggtgtc ctcccgacgt	14880
gagggtacgt acgaggacgc cagtggtgtg cgccatgcca cgggcgggt cggteccga	14940
ccggccgacc aggatgatgc catcgcgcg ctcaccgacc cacaaggcga cggagcggcg	15000
gcggtctggc caccgcagg cgctgtcgcg ttcacagcag acgatctgga gggcctgtac	15060
gacgggtacg cggcgcgggg attcgagtac ggcccgtgt tccgaggact gcggcgggcc	15120
tggcgacgtg gcgaggacat cttcgccgag gtgcgcctc cgcacacgc ggacggcgac	15180
gcctcccagt tctccgtaca ccccgccctg ctggacgccg cctgcacgc cggcgcttc	15240
cggcgggcgc acgaactccc gcacggggct ctgcccttct ccttcagcgg ggtgaggctg	15300
cacgggcccc gagcgtcgac cctgcgggtg cgcctaccc cgcatggcca ggcgcgggac	15360
acgcacgcat ggtcggtcgc ggtggtcgac ggcgaggggc ggccggtggc ctcgatcgcg	15420
tcgctcgcg tccgcccgt gtgcagcag gattgtctgg cggcctccg tacggcgcg	15480
cgggactcgc tcttcggtg cgagtgggtg accgccccg cgccgacctc gtcgtccgt	15540
ccgcgacgcc tggccacgtt ggggcccagc gaccgcctcc cctcggcaga cgcgtacgcg	15600
aacctcgccg acctggccgc cgcagtgtg gaggcggagg ccccgcgcc cgatgcggtc	15660
gtggtcgact gcggcgccg cgacgcgcgc gccacggccg tggcgaggga cgtacggacc	15720
ctcacccgac gcatcctggg tctgtgcag gattggctgg cggacgagag gccggcctcg	15780

agccggatgg tcgtactgac ccgtgggtgcg gtggccacca caccggggga ggacgtggcg	15840
gacctggcgg gcgcggcggt gtgcggcatg gtgcgggtccg cgcagtcgga acatcccggc	15900
cggttcgtcc tgctggacct cgaccccgac ccggacctcg acggcgggga agtgccaccg	15960
accgtcgtac cggcggctct cgcctgtggt gagccgcaga tcgcgggtgcg tgcgaaccgg	16020
cacctgggtgc cccggctgac ccgcgttcg gtgtccgtcc ccgtccccgg gcctgttccc	16080
gttcccgccg ccgaggcagc cgaccaggac accacgcca cggcgttcga ccccgacggc	16140
accgtactga tcaccggcgg caccggcacc ctccggcggg tgctcgcgcg ccatctggtc	16200
agccgtcacg gcgtacggca cctgtctgtg gcatcgcgac gcggggcccga cgcaccggc	16260
gccaccgagc tgcgggcgga actggccgag ctccggggccg aggcgacggt gcgcgttgt	16320
gacaccggtg accgaggcgc gctggcggat ctcatcgcgg ggattccac cggccaccct	16380
ttgaccggtg tgggtccacg cgcgggcgtc ctggatgacg ccaccgtcgc ctccctcacc	16440
ccccgacacc tggacaccgc gctgacacc aaggccgacg ccgccttcca tctgcacgag	16500
ctcaccgcc acgcccggcc gcgcgccttc gtctgttct cctcggccgc cgggtgtctc	16560
ggcgcagccg ggcagggcaa ctacgtgcc gccaacgcct tcctcgacgc cctcgccgaa	16620
caccgcaggg cgcagggcct gccgccttg tcgtcgcgt ggggcttgt ggagcagggc	16680
agcggcatga ccgggcatct cgaccgcacc gaccgggccc gcatcaaccg ctccggactc	16740
gccccctcg ccaccgagga cgctctcgcg ctcttcgacg ccgcctcgc cggcgatcgg	16800
ccgttcttg tgcccgccc gctggacct cggggttcaa gcgccgccga gacccggcg	16860
ccgtgttct ccaggatcg cccggctcgt acgaccggg gccgtaccc cggcgtgag	16920
ggcgcgcgtg accttcgtac ccgtctcgcg gccaggatg ccaccgagca gcgcgacacg	16980
cttctcaga tcgtccgcac ccacaccgcc gccgtcctgg ggcatgacac ggctgccgcc	17040
gtgcggccgg acgcgccctt ccgtgagctg ggtttcgact ccctcgccgc cgtggaactc	17100
cgtaaccgcc ttcaaagac caccgccctc accctgccc cgaccaccgt ttctgaccac	17160
cccacgccc ctgcctcgc gatcatctg cgtactcagc tctgccagga cgctccgtcc	17220
ccggcggcgg ccacggccat ggccggcatg gcggagctgg ccaggctgga gtccgccgtc	17280
tccgattcgg cggcgtcga cgacgacac gcgcagggcc tcgcggagcg cctgcggtcc	17340
ctgcgccga agatgagcag tggccgtgtc gtgcaccaca acggcggcgg cgtgcgggc	17400
ctggatctcc agtcggccac ggacgatgag atgttcgagc tgatcgaaa ggaggtcagc	17460
cgagactga	17469

<210> 76
 <211> 5822
 <212> PRT
 <213> Artificial Sequence
 <220><223> meiA3 protein of Streptomyces nanchangensis
 <400> 76
 Met Glu Ile Pro Met Ala Ala Gly His Asp Lys Val Ile Glu Ala Leu
 1 5 10 15
 Arg Ala Ser Leu Lys Thr Asn Glu Arg Gln Arg Glu Gln Ile His Arg
 20 25 30
 Leu Thr Thr Ala Ala Arg Glu Pro Ile Ala Ile Ile Gly Met Ala Cys
 35 40 45

 Arg Tyr Pro Gly Gly Val Gly Ser Pro Glu Asp Leu Trp Glu Leu Val
 50 55 60
 Ala Ala Gly Arg Asp Ala Ile Gly Thr Phe Pro Glu Asp Arg Gly Trp
 65 70 75 80
 Asp Ala Ala Arg Leu Tyr Asp Pro Asp Pro Glu Arg Ala Gly Thr Ser
 85 90 95
 Tyr Thr Gln His Gly Gly Phe Leu Tyr Gln Ala Gly Glu Phe Asp Pro
 100 105 110
 Gly Phe Phe Gly Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln

 115 120 125
 Gln Arg Leu Leu Leu Glu Ile Ser Trp Glu Ala Phe Glu Arg Ala Gly
 130 135 140
 Ile Asp Pro Ala Ser Val Arg Gly Ser Arg Thr Gly Val Phe Ala Gly
 145 150 155 160
 Val Met Tyr His Asp Tyr Gly Ser Arg Leu His Thr Val Pro Glu Gly
 165 170 175
 Phe Glu Gly Tyr Val Gly Asn Gly Ser Gly Gly Gly Val Ala Ser Gly
 180 185 190

 Arg Val Ala Tyr Thr Leu Gly Leu Glu Gly Pro Ala Val Thr Val Asp
 195 200 205

Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Cys Gln Ala
 210 215 220
 Leu Arg Ala Gly Glu Cys Ser Leu Ala Leu Ala Gly Gly Val Thr Val
 225 230 235 240
 Met Ser Thr Pro Ser Leu Phe Val Glu Tyr Ser Arg Gln Arg Ala Leu
 245 250 255
 Ala Ala Asp Gly Arg Cys Lys Ala Tyr Gly Ala Gly Ala Asp Gly Thr
 260 265 270
 Gly Trp Ala Glu Gly Ala Gly Met Leu Leu Val Glu Arg Leu Thr Asp
 275 280 285
 Ala Gln Arg Leu Gly His Arg Val Leu Ala Val Val Arg Gly Ser Ala
 290 295 300
 Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro
 305 310 315 320
 Ala Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Ser Ala Gly Val Ser
 325 330 335
 Ala Ser Glu Val Asp Ala Val Glu Gly His Gly Thr Gly Thr Arg Leu
 340 345 350
 Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Gln
 355 360 365
 Arg Pro Ala Asp Arg Pro Leu Trp Leu Gly Ser Met Lys Ser Asn Val
 370 375 380
 Gly His Ala Gln Ala Ala Ala Gly Val Gly Gly Ile Ile Lys Met Val
 385 390 395 400
 Met Ala Met Arg Ser Gly Thr Leu Pro Arg Thr Leu His Ala Asp Glu
 405 410 415
 Pro Ser Pro His Ile Asp Trp Asp Ser Gly Ala Val Arg Leu Leu Thr
 420 425 430
 Glu Pro Val Ala Trp Pro Glu Arg Asp Arg Pro Arg Arg Ala Ala Val
 435 440 445
 Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Val Ile Leu Glu Ala
 450 455 460

Ala Ser Gln Thr Ala Pro His Thr Glu Ser Ala Ser Gln Thr Glu Thr
465 470 475 480

Asp Asp Ala Pro Ala Pro His Gly Ala Pro Gly His Ala Val Ala Gly
485 490 495

Pro Leu Pro Trp Pro Leu Ser Gly Ala Thr Ala Glu Ala Leu Arg Ala
500 505 510

Gln Ala Arg Glu Leu Arg Arg Phe Val Ala Ala Asp Glu Leu Leu Arg
515 520 525

Pro Ala Asp Val Gly His Thr Leu Val Leu Gly Arg Ser Asp Leu Ala
530 535 540

His Arg Ala Val Val Leu Gly Ser Asp Arg Glu Thr Leu Leu Arg Gly

545 550 555 560
Leu Asp Thr Leu Thr Gly Glu Gly Pro Asp Gly Gly Ser Val Val Arg
565 570 575

Gly Val Ala Ala Ala Gly Ala Gly Ala Gly Val Val Phe Val Phe Pro
580 585 590

Gly Gln Gly Gly Gln Trp Ala Gly Met Gly Leu Arg Leu Leu Glu Thr
595 600 605

Ser Ser Phe Phe Ala Glu Arg Met Ala Glu Cys Glu Ala Ala Leu Ala
610 615 620

Pro Tyr Val Asp Trp Ser Leu Leu Asp Val Leu Arg Arg Asp Pro Gly
625 630 635 640

Asp Pro Val Trp Glu Arg Ala Asp Val Val Gln Pro Met Leu Phe Ser
645 650 655

Val Met Val Ser Leu Ala Gln Leu Trp Arg Ser Tyr Gly Val Glu Pro
660 665 670

Asp Ala Val Leu Gly His Ser Gln Gly Glu Ile Ala Ala Ala His Ile
675 680 685

Cys Gly Ala Leu Thr Leu Asp Asp Ala Ala Lys Val Val Ala Leu Arg

690 695 700

Ser Arg Ala Leu Gln Thr Leu Arg Gly Ser Gly Gly Met Ala Ser Val

705 710 715 720
 Pro Leu Thr Ala Asp Glu Val Ala Gly Leu Leu Arg Thr Ala Trp Pro
 725 730 735
 Asp Arg Leu Trp Val Ala Ala Val Asn Ala Pro Thr Ala Thr Val Ile
 740 745 750
 Ser Gly Asp Ala Asp Ser Leu Ala Glu Ala Leu Glu His Tyr Arg Asp
 755 760 765

 Gln Gly Val Asp Ala Lys Arg Val Pro Val Asp Tyr Ala Ser His Cys
 770 775 780
 Pro His Ile Glu Ala Val Glu Gln Glu Leu Leu Ser Leu Leu Arg Gly
 785 790 795 800
 Ile Ala Pro Arg Ala Ala Asp Ile Pro Phe Tyr Ser Thr Val Asp Asn
 805 810 815
 Gln Trp Ala Asp Thr Met Gly Leu Asp Ala Arg Tyr Trp Tyr Arg Asn
 820 825 830
 Leu Arg Arg Pro Val Arg Phe Ala Glu Ala Leu Arg Ala Leu Gly Ala

 835 840 845
 Ala Glu Tyr Arg Ser Tyr Val Glu Val Gly Pro His Pro Thr Leu Thr
 850 855 860
 Pro Ala Ile Glu Asp Thr Thr Glu Ala Ala Gly Ala Ala Ala Thr Val
 865 870 875 880
 Val Gly Ser Leu Arg Arg Gly Glu Asp Asp Ala His Arg Ile Leu Thr
 885 890 895
 Ser Leu Ala Arg Ala His Ile His Gly Leu Pro Val Ala Trp Asp Arg
 900 905 910

 His Tyr Arg Ala Leu Ala Pro Glu Ala Asn His Val Asp Leu Pro Thr
 915 920 925
 Tyr Ala Phe Gln Arg Arg Arg Tyr Trp Leu Asp Ala Pro Ala Thr Thr
 930 935 940
 Gly Asp Val Thr Ala Ala Gly Leu Ala Pro Val Gly His Pro Leu Leu
 945 950 955 960
 Gly Ala Ala Val Gly Leu Ala Glu Gly Asp Gly Tyr Leu Leu Thr Gly

965 970 975
Arg Leu Ala Pro His Thr His Pro Trp Leu Thr Asp His Ala Val Ala

980 985 990
Gly Thr Val Leu Leu Pro Gly Thr Ala Tyr Val Glu Leu Ala Val His

995 1000 1005
Val Gly Glu His Leu Gly Cys Pro Arg Leu Glu Glu Leu Thr Leu His

1010 1015 1020
Ala Pro Leu Val Leu Pro Asp Thr Gly Gly Val Ala Leu Gln Val Ala

1025 1030 1035 1040
Val Gly Ala Pro Asp Glu Thr Gly Arg Arg Ala Leu Ser Val Tyr Ala

1045 1050 1055

Gln Arg Asp Asp Asp Pro Thr Trp Glu Gly Ala Ala Arg Gly Ala Trp
1060 1065 1070

Thr Arg His Ala Thr Gly Thr Leu Ala Ala Glu Ala Ala Thr Asp Gly
1075 1080 1085

Ile Asn Gly Ala Asp Gly Ala Gly Pro Leu Ala Gly Ala Trp Pro Pro
1090 1095 1100

Pro Gly Ala Glu Pro Leu Asp Ile Ser Gly Leu Tyr Asp Thr Leu Ala
1105 1110 1115 1120

Ala Ala Asp Phe Gly Tyr Gly Pro Ala Phe Gln Gly Leu Arg Ala Val

1125 1130 1135
Trp Arg His Gly Glu Glu Thr Tyr Ala Glu Val Arg Leu Pro Asp Gln

1140 1145 1150
Val Ala Ala Asp Ala Pro Arg Phe Cys Leu His Pro Ala Leu Leu Asp

1155 1160 1165
Ala Ala Leu His Pro Leu Ala Leu Asp Ser Gly Arg Ser Glu Glu Asn

1170 1175 1180
Pro Ala Gly His Gly Leu Leu Pro Phe Ala Trp Arg Gly Val Ser Leu

1185 1190 1195 1200

Arg Ser Pro Gly Thr Pro Thr Leu Arg Val Arg Leu Arg Pro Gln Gly
1205 1210 1215

Pro Asp Ser Ile Ala Val Asp Val Ala Asp Glu Thr Gly Ala Ser Val
1220 1225 1230
Val Ser Ala Glu Ser Leu Thr Leu Arg Pro Val Ala Leu Glu Asp Leu
1235 1240 1245
Arg Val Leu Gly Gly Gln Ala Asn Asp Pro Leu Tyr Ala Leu Glu Trp
1250 1255 1260
Thr Ala Ala Pro Glu Pro Leu Thr Thr Ala Leu Gly Arg Cys Ala Val

1265 1270 1275 1280
Leu Gly His Ala Thr Pro Gly Trp Ala Ala Ala Leu Glu Thr Ala Ala
1285 1290 1295
Ala Glu Pro Val Arg Arg Tyr Pro Asp Leu Ala Gly Leu Val Ala Ala
1300 1305 1310
Leu Asp Ala Gly Asp Pro Pro Pro Asp Leu Val Phe Val Gly Cys Pro
1315 1320 1325
Pro Ala Ala Ala Gly Pro Asp Asp Thr Thr Val Ala Asp Val His Thr
1330 1335 1340

Thr Arg Thr Arg Val Arg Thr Arg Gln Ala Leu Glu Leu Leu Gln Gly
1345 1350 1355 1360
Trp Leu Gly Glu Ala Arg Leu Ala Gly Ala Arg Leu Val Leu Val Thr
1365 1370 1375
Arg Gly Ala Val Ala Thr Gly Pro Ala Gly Gly Gly Met Asp Leu Ala
1380 1385 1390
Gly Ala Ala Ile Cys Gly Leu Val Arg Ser Ala Gln Ala Glu Glu Pro
1395 1400 1405
Asp Arg Ile Leu Leu Val Asp Leu Asp Thr Ala Glu Glu Ser Trp Ala

1410 1415 1420
Ala Leu Pro Arg Ala Val Ala Leu Gly Glu Pro Gln Met Ala Ile Arg
1425 1430 1435 1440
Ala Gly Gln Pro His Met Ala Arg Leu Val Arg Ala Asp Thr Glu Arg
1445 1450 1455
Asp Ala Leu Leu Thr Pro Pro Arg Gly Ser Gly Gly Trp Arg Leu Asp
1460 1465 1470

Cys Ala Asp Ala Gly Thr Leu Gln Gly Leu Ala Pro Val Ala Ser Ser

1475 1480 1485

Ala Asp His Asp Pro Leu Gly Pro Gln Gln Val Arg Ile Glu Val Arg

1490 1495 1500

Ala Ala Gly Leu Asn Phe Arg Asp Val Leu Val Ala Leu Gly Met Val

1505 1510 1515 1520

Pro Gly Gln Gln Gly Leu Gly Ser Glu Gly Ala Gly Val Val Leu Glu

1525 1530 1535

Ala Gly Pro Glu Val Ala Asp Leu Ala Pro Gly Asp Arg Val Met Gly

1540 1545 1550

Val Phe Ala Asp Ala Phe Gly Pro Phe Ala Ile Ala Asp Arg Ala Thr

1555 1560 1565

Val Ile Arg Val Pro Glu His Trp Thr Phe Ala Gln Ala Ala Val

1570 1575 1580

Pro Val Val Phe Ala Thr Ala Tyr Tyr Gly Leu Val Asp Leu Ala Gly

1585 1590 1595 1600

Leu Arg Pro Gly Glu Ser Val Leu Val His Ala Ala Ala Gly Gly Val

1605 1610 1615

Gly Leu Ala Ala Val Gln Leu Ala Arg His Leu Gly Ala Glu Val Tyr

1620 1625 1630

Ala Thr Ala Ser Pro Gly Lys Trp Asp Thr Leu Arg Ala His Gly Ile

1635 1640 1645

Pro Pro Glu Arg Ile Ala Ser Ser Arg Thr Leu Asp Phe Glu Ser Arg

1650 1655 1660

Phe Thr Gly Arg Asn Ile Asp Val Val Leu Asn Ser Leu Ala His Glu

1665 1670 1675 1680

Tyr Val Asp Ala Ser Leu Arg Leu Val Ser Gly Asp Ser Gly Arg Phe

1685 1690 1695

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

1700 1705 1710

Ala Tyr Pro Gly Val Ala Tyr Arg Ala Tyr Asp Leu Met Glu Ala Gly

1715 1720 1725
 Pro Glu Arg Ile Gly Glu Ile Leu Arg Thr Val Leu Arg Leu Phe Asp
 1730 1735 1740
 Glu Gly Val Leu Thr Pro Leu Pro Leu Thr Cys Trp Asp Ile Arg Gln
 1745 1750 1755 1760
 Ala Arg Asp Ala Phe Arg Gln Leu Gln Gln Gly Arg Thr Val Gly Lys
 1765 1770 1775

 Asn Val Leu Thr Leu Asp Arg Thr Pro Asp Pro Asp Gly Thr Val Leu
 1780 1785 1790
 Ile Thr Gly Gly Thr Gly Thr Leu Gly Ala Ala Leu Ala Arg His Leu
 1795 1800 1805
 Ala Ala Thr Gly Arg Ala Arg His Leu Leu Leu Ile Ser Arg Arg Gly
 1810 1815 1820
 Leu Asp Ala Pro Gly Ala Pro Glu Leu Ile Ala Glu Ile Asp Glu Leu
 1825 1830 1835 1840
 Gly Ala Ala Thr Thr Val Ala Thr Cys Asp Val Gly Asp Arg Ala Ala

 1845 1850 1855
 Leu Ala Glu Leu Leu Gly Arg Ile Pro Ala Glu His Pro Leu Thr Ala
 1860 1865 1870
 Val Val His Ala Ala Gly Thr Leu Asp Asp Ala Thr Leu Gly Ser Leu
 1875 1880 1885
 Thr Ala Arg His Leu Asp Thr Val Leu Pro Ala Lys Ala Asp Ala Ala
 1890 1895 1900
 Trp His Leu His Glu Leu Thr Cys Arg Leu Asp Leu Ala Ala Phe Val
 1905 1910 1915 1920

 Leu Phe Ser Ser Ala Ala Gly Val Leu Gly Ser Pro Gly Gln Gly Asn
 1925 1930 1935
 Tyr Ala Ala Ala Asn Ala Phe Leu Asp Ala Leu Ala Phe Gln Arg Arg
 1940 1945 1950
 Ala Met Gly Leu Pro Ala Val Ser Leu Ala Trp Gly Leu Trp Glu Glu
 1955 1960 1965
 Ala Ser Gly Met Thr Gly His Leu Asp Gln Thr Asp Arg Thr Arg Met

1970 1975 1980
 Ala Arg Val Gly Leu Arg Pro Leu Ala Thr Asn Glu Ala Leu Ala Leu

 1985 1990 1995 2000
 Phe Asp Asn Ala Leu Val Asp Gly Pro Pro Leu Leu Leu Pro Ala Arg

 2005 2010 2015
 Ile Asp Thr Lys Ala Leu Arg Gly Thr Thr Ala Pro Pro Leu Phe Gln

 2020 2025 2030
 Ser Leu Val Arg Pro Thr Thr Gly His Arg Pro Arg Pro Ala Thr Pro

 2035 2040 2045
 Asp Gly Arg Ser Ser Leu Arg Ala Arg Leu Ala Gly Leu Asp Pro Ala

 2050 2055 2060

 Ala Gln His Glu Val Leu Leu Thr Leu Val Arg Gly His Ala Ala Thr
 2065 2070 2075 2080
 Val Leu Gly His Pro Ser Pro Asp Ala Ile Ala Pro Glu Ala Ala Phe

 2085 2090 2095
 Arg Asp Leu Gly Phe Asp Ser Leu Thr Ala Val Glu Leu Arg Asn Arg

 2100 2105 2110
 Leu Lys Glu Ala Thr Gly Leu Arg Leu Pro Ala Thr Leu Val Phe Asp

 2115 2120 2125
 His Pro Thr Pro Ala Ala Leu Ala Gln His Leu Arg Asp Gly Leu Ile

 2130 2135 2140
 Gly Gly Ala Asp Ala Ala Thr Leu Ala Ser Ala Pro Ala Pro Ser Glu
 2145 2150 2155 2160
 Val Ala Thr Val Ala Asp Glu Ala Ile Ala Ile Ile Gly Met Ala Cys

 2165 2170 2175
 Arg Tyr Pro Gly Gly Val Arg Ser Ala Glu Gly Leu Trp Asp Leu Val

 2180 2185 2190
 Ala Ser Gly Thr Asp Ala Met Ser Gly Phe Pro Thr Asp Arg Gly Trp

 2195 2200 2205

 Asp Leu Asp Arg Leu Tyr Ala Pro Gln Asp Gln Asp Arg Pro Gly Thr

 2210 2215 2220

Thr Tyr Thr Arg His Gly Gly Phe Leu His Asp Ala Gly Lys Phe Asp
 2225 2230 2235 2240
 Ala Gly Phe Phe Gly Ile Gly Pro Arg Glu Ala Leu Ala Met Asp Pro
 2245 2250 2255
 Gln Gln Arg Leu Leu Leu Glu Thr Ser Trp Glu Val Phe Glu His Ala
 2260 2265 2270
 Gly Ile Asp Pro Ser Ser Val Arg Arg Ser Arg Thr Gly Val Phe Ala
 2275 2280 2285
 Gly Val Met Pro Thr Asp Tyr Gly Pro Arg Leu Gln Asp Thr Val Ala
 2290 2295 2300
 Glu Val Glu Gly Tyr Val Leu Thr Gly Asn Ser Gly Ser Val Ala Ser
 2305 2310 2315 2320
 Gly Arg Ile Ala Tyr Thr Phe Gly Leu Glu Gly Pro Ala Val Ser Val
 2325 2330 2335
 Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu Ala Cys Gln
 2340 2345 2350
 Ala Leu Arg Ala Gly Glu Cys Ser Met Ala Leu Ala Gly Gly Val Thr
 2355 2360 2365
 Val Met Ala Thr Pro Gly Ala Phe Val Glu Phe Ala Arg Gln Arg Gly
 2370 2375 2380
 Leu Ser Val Asp Gly Arg Cys Lys Ala Phe Gly Val Gly Ala Asp Gly
 2385 2390 2395 2400
 Thr Gly Trp Ala Glu Gly Val Gly Met Leu Leu Val Glu Arg Leu Ser
 2405 2410 2415
 Asp Ala Arg Arg Leu Gly His Arg Val Leu Ala Val Val Arg Gly Ser
 2420 2425 2430
 Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly
 2435 2440 2445
 Pro Ser Gln Gln Arg Val Ile Arg Gln Ala Leu Ala Ser Ala Arg Val
 2450 2455 2460
 Gly Gly Ala Asp Val Asp Val Val Glu Gly His Gly Thr Gly Thr Arg
 2465 2470 2475 2480

Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln
 2485 2490 2495

Glu Arg Ser Gly Asp Glu Pro Leu Trp Leu Gly Ser Val Lys Ser Asn
 2500 2505 2510

Ile Gly His Ala Gln Ala Ala Ala Gly Val Ala Gly Val Ile Lys Met
 2515 2520 2525

Val Met Ala Met Arg Cys Gly Val Leu Pro Arg Thr Leu His Val Gln
 2530 2535 2540

Glu Pro Ser Pro His Val Asp Trp Ser Ser Gly Gly Val Arg Leu Leu
 2545 2550 2555 2560

Thr Glu Ala Val Pro Trp Pro Glu Thr Gly Arg Ala Arg Arg Ala Gly
 2565 2570 2575

Val Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Ile Ile Leu Glu
 2580 2585 2590

Gln Ala Pro Pro Glu Glu His Asp Asp Pro Ala Asp Val Ser Ser Gly
 2595 2600 2605

Ser Phe Pro Trp Met Val Ser Ala Lys Ser Glu Gln Ala Leu Gln Ala
 2610 2615 2620

Gln Ala Ala Gln Leu Arg Ala Tyr Leu Ala Ala Arg Pro Gly Val Gly
 2625 2630 2635 2640

Leu Ala Asp Val Gly Tyr Ala Leu Ala Ala Gly Arg Thr Ala Phe Asp
 2645 2650 2655

His Arg Ala Val Leu Leu Gly Pro Asp Arg Glu Ala Phe Leu Glu Gly
 2660 2665 2670

Leu Gly Ala Leu Gly Ala Gly Glu Glu His Ala Gly Leu Val Arg Gly
 2675 2680 2685

Val Ala Thr Gly Ala Gly Lys Leu Ala Phe Val Cys Ser Gly Gln Gly
 2690 2695 2700

Thr Gln Arg Pro Arg Met Gly His Glu Leu Tyr Arg Ala Phe Pro Leu
 2705 2710 2715 2720

Phe Ala Ala Ala Met Asp Glu Ala Cys Ala Tyr Leu Asp Pro His Leu

2725 2730 2735
 Asp Arg Pro Leu Arg Asp Val Val Phe Ala Glu Pro Asp Ser Gly Thr
 2740 2745 2750
 Ala Arg Leu Leu Gln Gln Thr Arg Tyr Ala Gln Pro Ala Leu Phe Ala
 2755 2760 2765
 Leu Gln Val Ala Leu His Arg Leu Val Thr Glu His Tyr Gly Leu Thr
 2770 2775 2780

 Pro His Tyr Tyr Ala Gly His Ser Leu Gly Glu Ile Thr Ala Ala His
 2785 2790 2795 2800
 Leu Ala Gly Ile Leu Thr Leu Cys Asp Ala Ala Arg Leu Val Thr Thr
 2805 2810 2815
 Arg Ala Arg Leu Met Gln Ser Leu Pro Ala Thr Gly Ala Met Thr Thr
 2820 2825 2830
 Leu Gln Ala Asp Pro Asp Glu Leu His Glu His Leu Ala Arg Cys Glu
 2835 2840 2845
 Gly Arg Val Ser Leu Ala Ala Val Asn Ala Pro Gly Ser Val Val Ile

 2850 2855 2860
 Ser Gly Asp Arg His Asp Val Asp Ala Thr Ala Glu Asn Phe Arg Ala
 2865 2870 2875 2880
 Met Gly Arg Lys Thr Thr Pro Leu Lys Val Ser Gly Ala Phe His Ser
 2885 2890 2895
 His His Ile Asp Pro Leu Leu Asp Glu Leu Arg Ala Thr Ala Glu Thr
 2900 2905 2910
 Leu Thr Tyr His Pro Pro His Thr Pro Leu Ile Thr Thr Asp Leu Thr
 2915 2920 2925

 Asp Gln Asp Pro Thr Thr Pro Gly Tyr Trp Val Arg Gln Thr Arg Glu
 2930 2935 2940
 Thr Val His Tyr Ala His Thr Thr Gln Gln Leu His Thr His Gly Val
 2945 2950 2955 2960
 Thr Ala Tyr Leu Glu Leu Gly Pro Asp Thr Thr Leu Thr Thr Leu Thr
 2965 2970 2975
 His His Asn Leu Pro His His Thr Pro Leu Ala Ile Pro Leu Leu His

2980 2985 2990
Pro Asp Gln Pro Glu Thr His Thr Thr His Thr Ala Leu Ala His Leu

2995 3000 3005
His Thr His Gly His Pro Thr Thr Trp His His His His Thr Pro Thr

3010 3015 3020
His His His Pro Asn Leu Pro Thr Tyr Pro Phe Gln His His His Tyr

3025 3030 3035 3040
Trp Leu Asn Thr Thr Thr Ala Thr Gly Asp Met Ser Ala Ala Gly Leu

3045 3050 3055
Glu Pro Ala Arg His Pro Leu Leu Gly Ala Ala Val Glu Leu Ala Asp

3060 3065 3070
Gly Glu Gly Leu Leu Phe Thr Gly Arg Ile Ser Leu Arg Thr His Pro

3075 3080 3085
Trp Leu Ala Asp His Ala Val Gly Gly Ala Val Leu Leu Pro Gly Thr

3090 3095 3100
Ala Phe Leu Glu Leu Ala Leu Glu Ala Ala Ala His Val Asp Cys His

3105 3110 3115 3120
Arg Ile Glu Glu Leu Thr Leu His Thr Pro Leu Val Val Pro Glu Ser

3125 3130 3135
Gly Gly Val Val Leu Gln Val Thr Val Ala Gly Pro Asn Glu Ala Gly

3140 3145 3150
Asn Arg Ala Val Asp Ile Tyr Ser Arg Ile Asp Val Gly Gly Leu Thr

3155 3160 3165
Ala Asp Ser Val Gly Glu Pro Trp Thr Arg His Ala Ala Gly Tyr Leu

3170 3175 3180
Ala Asp Lys Pro Gly Pro Asp Cys Gly Asp Ser Ala Asp Gly Val Met

3185 3190 3195 3200
Pro Ala Gly Ala Trp Pro Pro Pro Gly Ala Val Ala Val Asp Leu Glu

3205 3210 3215
Glu Leu Tyr Glu Gln Leu Ala Glu Gly Gly Phe His Tyr Gly Ala Ala

3220 3225 3230
Glu Leu Tyr Glu Gln Leu Ala Glu Gly Gly Phe His Tyr Gly Ala Ala

Phe Arg Cys Leu Asp Ala Ala Trp Gln Arg Gly Asp Glu Val Phe Ala
 3235 3240 3245
 Thr Val His Met Ser Glu Asn Gln Leu Gly Asp Thr Ala Ala Ala Arg
 3250 3255 3260
 Phe Ala Leu His Pro Ala Leu Leu Asp Ser Ala Leu His Thr Ile Pro
 3265 3270 3275 3280
 Leu Leu Pro Ser Leu Gln Gly Gln Gln Asp Ser Gly Leu Pro Phe Thr
 3285 3290 3295
 Trp Ala Gly Val Thr Leu Arg Ala Ser Gly Ala Thr Ala Leu Arg Val
 3300 3305 3310
 Arg Leu Arg Pro Asp Gly His Gly Pro Gly Ala Val Ser Val Asp Val
 3315 3320 3325
 Ser Asp Glu Ala Gly Glu Pro Val Ala Ser Val Arg Ser Leu Ala Leu
 3330 3335 3340
 Arg Pro Val Thr Arg Val Glu Leu His Thr Ala Glu Leu Arg Thr Ala
 3345 3350 3355 3360
 Ala Pro Val Ala Pro His Ser Ser Leu Phe Glu Val Arg Trp Glu Pro
 3365 3370 3375
 Val Pro Gln Pro Ser Ala Ala Glu Glu Ala Asp Pro Trp Val Met Ile
 3380 3385 3390
 Gly Thr Gly Pro Thr Leu Arg Pro Asp Glu Asp Phe Ala Thr Pro Pro
 3395 3400 3405
 Glu Arg Thr Tyr Ala Asp Leu Ala Ala Leu Cys Ala Ala Val Ala Asp
 3410 3415 3420
 Gly Ala Pro Val Pro Arg Thr Val Val Ala Trp Ser Gln Ala Gly Ser
 3425 3430 3435 3440
 Glu Asp Glu Ser Ser Glu Ala Leu Arg His Ala Thr His His Met Leu
 3445 3450 3455
 Gly Leu Leu Gln Gln Trp Leu Ala Asp Ser Arg Phe Val Asp Ser Arg
 3460 3465 3470
 Leu Val Ile Leu Thr Arg Ala Ala Val Ala Thr Ala Pro Glu Glu Glu
 3475 3480 3485

Val Lys Asp Leu Ala Gly Ala Ala Thr Arg Gly Leu Ile Arg Ser Ala
3490 3495 3500

Gln Ser Glu His Pro Asp Arg Phe Val Leu Leu Asp Leu Asp Asp Arg
3505 3510 3515 3520

Pro Ala Asp Ala Lys Asp His Asp Arg Met Leu Ser Val Ala Leu Ala
3525 3530 3535

Cys Gly Glu Pro Glu Val Ala Val Arg Asp Gly Ala Leu Arg Thr Pro
3540 3545 3550

Arg Leu Ser Pro Leu Ala Gly Thr Ala Thr Glu Ala Met Asp Glu His
3555 3560 3565

Pro Trp Asp Pro Asp Gly Thr Val Leu Ile Thr Gly Gly Thr Gly Ser
3570 3575 3580

Leu Gly Ala Met Leu Ala Arg His Leu Val Ala Thr His Gly Val Arg
3585 3590 3595 3600

His Leu Leu Leu Ile Ser Arg Arg Gly Leu Asp Ala Pro Gly Ala Arg
3605 3610 3615

Arg Gln Gly Asn Glu Leu Val Glu Leu Gly Ala Gln Leu Thr Ile Ala
3620 3625 3630

Ala Cys Asp Ala Ala Asp Gln Arg Gln Leu Ala Asn Ala Leu Ser Glu
3635 3640 3645

Ile Ser Val Asp His Pro Leu Thr Ala Val Val His Ala Ala Gly Val
3650 3655 3660

Leu Asp Asp Gly Val Ile Thr Ser Leu Thr Pro Glu Asp Leu Thr His
3665 3670 3675 3680

Val Leu Arg Ala Lys Val Asp Ser Ala Leu Asn Leu His Gln Leu Thr
3685 3690 3695

Arg Asp Leu Pro Leu Ser Ala Phe Val Leu Phe Ser Ser Leu Ala Gly
3700 3705 3710

Val Met Gly Ser Ala Gly Gln Gly Asn Tyr Ala Ala Ala Asn Ala Ala
3715 3720 3725

Leu Asp Ala Leu Ala Ser His Arg Arg Ala Thr Arg Leu Pro Ala Val

3730 3735 3740
 Ser Leu Ala Trp Gly Val Trp Glu Gln Thr Glu Gly Met Thr Gly Gln
 3745 3750 3755 3760
 Leu Glu Ala Thr Gly His Ala Arg Leu Arg Arg Ser Gly Leu Arg Pro
 3765 3770 3775
 Leu Ala Thr Ser Glu Gly Leu Glu Leu Phe Asp Lys Ala Leu Ser Cys
 3780 3785 3790

 Gly His Ala Leu Val Val Pro Ala Ala Leu Ser Thr Lys Glu Leu Gln
 3795 3800 3805
 Thr Ser Gly Ser Val Pro Pro Phe Leu Arg His Val Thr Gly Val Ala
 3810 3815 3820
 Pro Ala Arg Pro Ser Arg Thr Arg Asp Ala Ser Ala Gly Glu Pro Thr
 3825 3830 3835 3840
 Pro Leu Arg Arg Arg Leu Thr Gly Leu Gly Pro Glu Glu Arg Leu Arg
 3845 3850 3855
 Glu Val Leu Arg Leu Val Arg Ser Arg Ala Ala Ala Val Leu Gly His

 3860 3865 3870
 Gly Thr Ala Glu Ala Val Pro Ala Asp Ser Ala Phe Arg Asp Leu Gly
 3875 3880 3885
 Phe Asp Ser Leu Ala Ala Val Asp Leu Arg Asn Arg Leu Gln Gln Ala
 3890 3895 3900
 Thr Gly Leu Arg Leu Pro Ala Gly Leu Ile Phe Asp Arg Pro Arg Pro
 3905 3910 3915 3920
 Asp Val Leu Ala Arg Phe Leu Cys Asp Glu Leu Ala Gly Val Gly Gly
 3925 3930 3935

 Thr Ser Ala Ala Thr Ala Ala Pro Pro Val Ala Ala Val Gly Gly Ala
 3940 3945 3950
 Ala Gly Glu Pro Val Ala Ile Val Gly Met Ala Cys Arg Phe Pro Gly
 3955 3960 3965
 Gly Val Arg Ser Ala Glu Gly Leu Trp Asp Leu Val Ala Ser Gly Met
 3970 3975 3980
 Asp Ala Val Gly Asp Phe Pro Thr Asp Arg Gly Trp Glu Val Glu Arg

3985 3990 3995 4000
 Leu Tyr Asp Pro Asp Pro Asp Arg Thr Gly Thr Ser Tyr Thr Arg Gln

 4005 4010 4015
 Gly Gly Phe Leu Tyr Asp Ala Gly Glu Phe Asp Ala Ala Phe Phe Gly
 4020 4025 4030
 Ile Gly Pro Arg Glu Ala Val Ala Met Asp Pro Gln Gln Arg Leu Leu
 4035 4040 4045
 Leu Glu Ile Ser Trp Glu Ala Leu Glu Arg Ala Gly Ile Asp Pro Ala
 4050 4055 4060
 Ser Leu Arg Gly Ser Ser Thr Gly Val Phe Ala Gly Val Met Tyr His
 4065 4070 4075 4080

 Asp Tyr Gly Thr Arg Leu Arg Glu Ile Pro Glu Gly Tyr Glu Gly Tyr
 4085 4090 4095
 Ile Gly Asn Gly Asn Ala Gly Ser Val Ala Ser Gly Arg Val Ser Tyr
 4100 4105 4110
 Thr Phe Gly Leu Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser
 4115 4120 4125
 Ser Ser Leu Val Ala Leu His Leu Ala Cys Gln Ala Leu Arg Ser Gly
 4130 4135 4140
 Glu Cys Ser Met Ala Leu Ala Gly Gly Val Thr Val Met Ser Thr Pro

 4145 4150 4155 4160
 Thr Thr Phe Val Glu Phe Ser Arg Gln Arg Gly Leu Ala Pro Asp Gly
 4165 4170 4175
 Arg Cys Lys Ser Phe Gly Ala Gly Ala Asp Gly Thr Gly Trp Ala Glu
 4180 4185 4190
 Gly Ala Gly Met Leu Leu Val Glu Arg Leu Ser Asp Ala Arg Arg Asn
 4195 4200 4205
 Gly His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp
 4210 4215 4220

 Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Glu Arg
 4225 4230 4235 4240

Val Ile Arg Gln Ala Trp Ala Asn Ala Gly Val Ala Ala Met Asp Ile
4245 4250 4255

Asp Ala Val Glu Gly His Gly Thr Gly Thr Thr Leu Gly Asp Pro Ile
4260 4265 4270

Glu Ala Gln Ala Leu Leu Gly Thr Tyr Gly Gln Gly Arg Ser Ala Asp
4275 4280 4285

Arg Pro Leu Trp Leu Gly Ser Ile Lys Ser Asn Val Gly His Thr Gln
4290 4295 4300

Ala Ala Ala Gly Val Gly Gly Val Ile Lys Met Val Met Ala Met Arg
4305 4310 4315 4320

His Gly Leu Leu Pro Gln Thr Leu His Ala Glu Glu Pro Ser Pro His
4325 4330 4335

Val Asp Trp Ser Gly Gly Thr Val Arg Leu Leu Thr Glu Pro Val Ala
4340 4345 4350

Trp Pro Glu Arg Gly Arg Met Arg Arg Ala Gly Val Ser Ser Phe Gly
4355 4360 4365

Val Ser Gly Thr Asn Ala His Val Ile Leu Glu Gln Ala Pro Pro Asn
4370 4375 4380

Ala Glu Thr His Glu Pro Ala Glu Pro His Thr Ala Pro Gly Pro Leu
4385 4390 4395 4400

Pro Trp Thr Ile Ser Ala Lys Ser Pro Gln Ala Leu Arg Ala Gln Ala
4405 4410 4415

Arg Gln Leu His Thr Tyr Leu Thr Asn Thr Pro Glu Ala Asn Pro Ala
4420 4425 4430

Asp Val Gly His Thr Leu Ala Met Gly Arg Ala Ser Phe Glu His Arg
4435 4440 4445

Ala Val Val Ile Gly Ser Asp Arg Gly Glu Phe Leu Gly Gly Leu Asp
4450 4455 4460

Ala Val Ala Ala Asp Glu Ala His Ser Ala Val Val Thr Gly Ile Ala
4465 4470 4475 4480

Arg Lys Ala Gly Asp Leu Gly Lys Val Val Phe Val Phe Pro Gly Gln
4485 4490 4495

Gly Gly Gln Trp Ala Gly Met Gly Leu Arg Leu Leu Lys Thr Ser Pro
4500 4505 4510

Val Phe Ala Gln Ser Ile Gln Ala Cys Glu Gln Ala Leu Ala Pro His
4515 4520 4525

Thr Asp Trp Thr Leu Thr Asp Ile Leu His Arg Pro His Thr Asp Pro
4530 4535 4540

Leu Trp Gln Arg Ala Asp Val Ile Gln Pro Ala Leu Phe Ala Leu Met
4545 4550 4555 4560

Thr Ser Leu Thr Thr Leu Trp Gln Ser His Gly Leu Asn Pro Asp Ala
4565 4570 4575

Val Ile Gly His Ser Gln Gly Glu Ile Thr Ala Ala His Ala Cys Gly
4580 4585 4590

Ala Leu Ser Leu Glu Asp Ala Ala Lys Ile Val Ala Leu Arg Ser Gln
4595 4600 4605

Thr Leu Gln Thr Leu Gln Gly Ser Gly Gly Met Ala Ser Val Pro Leu
4610 4615 4620

Pro Ala Asp Gln Val Thr Ala Leu Leu His Thr Met Trp Pro Asp Gln
4625 4630 4635 4640

Leu Trp Val Ala Ala Ile Asn Ala Pro Thr Thr Thr Val Ile Ser Gly
4645 4650 4655

Asp Thr Gln Ala Leu Thr Gln Ala Leu Asn His Tyr Arg Asp Gln Asp
4660 4665 4670

Ile Asp Ala Lys Arg Ile Pro Val Asp Tyr Ala Ser His Cys Pro His
4675 4680 4685

Ile Gln Ala Val Gln His Glu Leu Ser Asp Leu Leu Gln Asp Ile Thr
4690 4695 4700

Pro Arg Ala Ala Thr Thr Pro Phe Tyr Ser Thr Thr Asp Asn Gln Trp
4705 4710 4715 4720

Thr Asp Thr Thr Thr Leu Asn Ala His Tyr Trp Tyr Arg Asn Leu Arg
4725 4730 4735

Gln Pro Val His Leu Thr Asn Ala Ile Thr Asn Leu Thr His Gln Gly

4740 4745 4750
 His His Thr Tyr Ile Glu Ile Ser Pro His Pro Thr Leu Thr Pro Ala
 4755 4760 4765
 Ile Gln Glu Thr Thr His Thr Thr His Thr Pro Thr Thr Val Ile Ser
 4770 4775 4780
 Thr Leu Arg Arg Asn His Asn Asp Thr His Gln Leu Leu His Ala Leu
 4785 4790 4795 4800

 Ala His Ala His Thr Thr Gly His Pro Ile Asn Trp His Pro Thr His
 4805 4810 4815
 Gln His His Thr Pro Thr Pro Gln His Thr Asp Leu Pro Thr Tyr Pro
 4820 4825 4830
 Phe Gln His Gln Arg Tyr Trp Leu Asn Thr Pro Thr Gln Thr Gly Asp
 4835 4840 4845
 Ala Ala Ala Ile Gly Leu Asp Pro Ala His His Pro Leu Leu Gly Ala
 4850 4855 4860
 Ala Val Ala Val Ala Glu Gly Glu Gly Tyr Leu Leu Thr Gly Arg Leu

 4865 4870 4875 4880
 Ala Leu Ser Thr His Pro Trp Leu Ala Asp His Thr Ile Ala Gly Ala
 4885 4890 4895
 Val Val Leu Pro Gly Thr Ala Leu Leu Glu Ile Ala Leu Gln Ala Gly
 4900 4905 4910
 His Arg Val Asp Cys His Arg Ile Glu Glu Leu Thr Leu Gln Ser Pro
 4915 4920 4925
 Leu Phe Ile Pro Glu Glu Gly Ala Val Gln Val Gln Ala Trp Val Ala
 4930 4935 4940

 Ala Pro Asp Glu Asn Gly Tyr Arg Ser Leu Thr Val Ser Ser Arg Arg
 4945 4950 4955 4960
 Glu Gly Thr Tyr Glu Asp Ala Thr Trp Val Arg His Ala Thr Gly Arg
 4965 4970 4975
 Val Gly Pro Ala Pro Ala Asp Gln Asp Asp Ala Ile Ala Arg Leu Thr
 4980 4985 4990
 Asp Pro Gln Gly Asp Gly Ala Ala Ala Val Trp Pro Pro Gln Gly Ala

4995 5000 5005
Val Ala Phe Thr Ala Asp Asp Leu Glu Gly Leu Tyr Asp Gly Tyr Ala

5010 5015 5020
Ala Arg Gly Phe Glu Tyr Gly Pro Val Phe Arg Gly Leu Arg Ala Ala

5025 5030 5035 5040
Trp Arg Arg Gly Glu Asp Ile Phe Ala Glu Val Arg Leu Pro Asp Thr

5045 5050 5055
Ala Asp Gly Asp Ala Ser Gln Phe Ser Val His Pro Ala Leu Leu Asp

5060 5065 5070
Ala Ala Leu His Ala Ala Ala Phe Arg Pro Ala Asp Glu Leu Pro His

5075 5080 5085

Gly Ala Leu Pro Phe Ser Phe Ser Gly Val Arg Leu His Gly Pro Gly

5090 5095 5100

Ala Ser Thr Leu Arg Val Arg Leu Thr Pro Asp Gly Gln Ala Arg Asp

5105 5110 5115 5120

Thr His Ala Trp Ser Val Ala Val Val Asp Gly Glu Gly Arg Pro Val

5125 5130 5135

Ala Ser Ile Ala Ser Leu Ala Val Arg Pro Val Ser Thr Gln Glu Leu

5140 5145 5150

Leu Ala Ala Ser Gly Thr Ala Arg Arg Asp Ser Leu Phe Ala Val Glu

5155 5160 5165

Trp Val Thr Ala Pro Ala Pro Thr Ser Ser Ser Ala Pro Arg Arg Leu

5170 5175 5180

Ala Thr Val Gly Pro Ser Asp Arg Leu Pro Ser Ala Asp Ala Tyr Ala

5185 5190 5195 5200

Asn Leu Ala Asp Leu Ala Ala Ala Val Leu Glu Ala Glu Ala Pro Ala

5205 5210 5215

Pro Asp Ala Val Val Val Asp Cys Gly Arg Arg Asp Ala Arg Ala Thr

5220 5225 5230

Ala Val Ala Glu Asp Val Arg Thr Leu Thr Arg Arg Ile Leu Gly Leu

5235 5240 5245

Leu Gln Glu Trp Leu Ala Asp Glu Arg Pro Ala Ser Ser Arg Met Val
 5250 5255 5260
 Val Leu Thr Arg Gly Ala Val Ala Thr Thr Pro Gly Glu Asp Val Ala
 5265 5270 5275 5280
 Asp Leu Ala Gly Ala Ala Val Cys Gly Met Val Arg Ser Ala Gln Ser
 5285 5290 5295
 Glu His Pro Gly Arg Phe Val Leu Leu Asp Leu Asp Pro Asp Pro Asp
 5300 5305 5310
 Leu Asp Gly Gly Glu Val Pro Pro Thr Val Val Pro Ala Ala Leu Ala
 5315 5320 5325
 Cys Gly Glu Pro Gln Ile Ala Val Arg Ala Asn Arg His Leu Val Pro
 5330 5335 5340
 Arg Leu Thr Arg Val Pro Val Ser Val Pro Val Pro Gly Pro Val Pro
 5345 5350 5355 5360
 Val Pro Ala Ala Glu Ala Ala Asp Gln Asp Thr Thr Pro Thr Ala Phe
 5365 5370 5375
 Asp Pro Asp Gly Thr Val Leu Ile Thr Gly Gly Thr Gly Thr Leu Gly
 5380 5385 5390
 Ala Val Leu Ala Arg His Leu Val Ser Arg His Gly Val Arg His Leu
 5395 5400 5405
 Leu Leu Ala Ser Arg Arg Gly Pro Asp Ala Pro Gly Ala Thr Glu Leu
 5410 5415 5420
 Arg Ala Glu Leu Ala Glu Leu Gly Ala Glu Ala Thr Val Arg Ala Cys
 5425 5430 5435 5440
 Asp Thr Gly Asp Arg Gly Ala Leu Ala Asp Leu Ile Ala Gly Ile Pro
 5445 5450 5455
 Thr Gly His Pro Leu Thr Gly Val Val His Ala Ala Gly Val Leu Asp
 5460 5465 5470
 Asp Ala Thr Val Ala Ser Leu Thr Pro Arg His Leu Asp Thr Ala Leu
 5475 5480 5485
 Thr Pro Lys Ala Asp Ala Ala Phe His Leu His Glu Leu Thr Arg His
 5490 5495 5500

Ala Arg Pro Arg Ala Phe Val Leu Phe Ser Ser Ala Ala Gly Val Leu
5505 5510 5515 5520

Gly Ala Ala Gly Gln Gly Asn Tyr Ala Ala Ala Asn Ala Phe Leu Asp
5525 5530 5535

Ala Leu Ala Glu His Arg Arg Ala Gln Gly Leu Pro Ala Leu Ser Leu
5540 5545 5550

Ala Trp Gly Leu Trp Glu Gln Gly Ser Gly Met Thr Gly His Leu Asp
5555 5560 5565

Arg Thr Asp Arg Ala Arg Ile Asn Arg Ser Gly Leu Ala Pro Leu Ala
5570 5575 5580

Thr Glu Asp Ala Leu Ala Leu Phe Asp Ala Ala Leu Ala Gly Asp Arg

5585 5590 5595 5600
Pro Phe Leu Val Pro Ala Arg Leu Asp Leu Arg Gly Ser Ser Ala Ala
5605 5610 5615

Glu Thr Pro Ala Pro Leu Phe Ser Arg Ile Ala Pro Ala Arg Thr Thr
5620 5625 5630

Arg Gly Arg Thr Pro Gly Ala Glu Gly Ala Ala Asp Leu Arg Thr Arg
5635 5640 5645

Leu Ala Ala Gln Asp Ala Thr Glu Gln Arg Asp Thr Leu Leu Thr Ile
5650 5655 5660

Val Arg Thr His Thr Ala Ala Val Leu Gly His Asp Thr Ala Ala Ala
5665 5670 5675 5680

Val Arg Pro Asp Ala Ala Phe Arg Glu Leu Gly Phe Asp Ser Leu Ala
5685 5690 5695

Ala Val Glu Leu Arg Asn Arg Leu Gln Thr Thr Thr Ala Leu Thr Leu
5700 5705 5710

Pro Ala Thr Thr Val Phe Asp His Pro Thr Pro Ala Ala Leu Ala Asp
5715 5720 5725

His Leu Arg Thr Gln Leu Cys Gln Asp Ala Pro Ser Pro Ala Ala Ala

5730 5735 5740

Thr Ala Met Ala Ala Met Ala Glu Leu Ala Arg Leu Glu Ser Ala Val

5745 5750 5755 5760
 Ser Asp Ser Ala Ala Leu Asp Asp Asp Thr Arg Ser Gly Leu Ala Glu
 5765 5770 5775
 Arg Leu Arg Ser Leu Ala Arg Lys Met Ser Ser Gly Arg Val Val Asp
 5780 5785 5790
 His Asn Gly Gly Gly Ala Ala Gly Leu Asp Leu Gln Ser Ala Thr Asp
 5795 5800 5805

Asp Glu Met Phe Glu Leu Ile Asp Lys Glu Val Ser Arg Asp
 5810 5815 5820

<210> 77
 <211> 17481
 <212> DNA
 <213> Artificial Sequence
 <220><223> mila3 gene of Streptomyces bingchenggensis
 <400> 77

atggccgctg gccacgacaa ggtgatcgag gcgctgcggg cgtccctcaa gaccaacgag	60
cggcagaggg aacagatcca ccggctcact acggcggcgc gggaacccat cgccatcatc	120
ggcatggcct gccgctatcc gggcggagtg ggatcgccgg aggacctgtg ggagctggtg	180
gccgccggtc gtgacgcat ccggcaccttc cccgaggacc ggggctggga cgtggagcgg	240
ctgtacgacc ccgatccgga gcgggccggc acctcgtgta cccagcatgg cggattcctg	300
taccaggcag gggagttcga ccccggtttc ttcgggatca gcccgcgcga ggcgctggcg	360
atggaccgcg agcagcggtt gctgctggag atctcctggg aggtgttcga gcgggccggg	420
atcgaccggg cctcgggtcg cggcagccgc accggggtgt tcgcgggcgt catgtaccac	480
gactacggct cccgggtgca caccgtcccc gaaggtctcg agggctatgt cggcaacggc	540
agcggcggcg gcgtggcgtc cggccgggtc gcctacaccc tcggcctcga aggcccggcc	600
gtgacctggg acaccgcctg ctctcctcgt ttggtcgccc tgcacctggc ctgccaggcg	660
ctgcgggccg gcgagtgtc actcgccctg gcgggcgggg tgacggtgat gtccaccccc	720
agcctgttcg tcgagtactc ccggcagcgc gcgctcgcg cagacggcgc gtgcaaggcg	780
tacggggcgg gggcggacgg caccggctgg gcagaaggcg ccgggatgct gctggtggaa	840
cggctcagcg acgcacagcg cctcggccac cgggtgctgg cggtggtccg gggcagcgcg	900
gtcaaccagg acggcgcgag caacggcctc accgccccca acggccccgc gcaacaacgg	960
gccatccggc aggcactggc gagcgccggg gtgtcggcgt ccgaggtcga cgccgtggag	1020

gggcatggga cggggacgcg gctgggcat ccgatcgagg cgcaggcgtt gctggcgacc 1080

tacggtcagc agcggcccg ggaccggcg ctgtggctcg ggtcgatgaa gtccaacgtc 1140

ggccatgcgc aggcggccgc cggcgtgggc gggatcatca agatggtgat ggccatgcgg 1200

agcgggacgc tgcgcgcac cctgcacgcg gacgagccgt cgccacacat cgactgggac 1260

tcgggcgccg tacggctgct gaccgagccg gtgcctggc cggagcgca cggccccgc 1320

cgcgcgcgg tgctctctt cggggtcagc ggcaccaacg cccatgtgat cctcgaggcc 1380

gcatcgaga cggcgccga gacggattcc gcgtcgagg cggaaaccga cgacgtccc 1440

gcaccgcac gcgcgcggg ccatgccgtg gcggggccgc tgcctggcc ctgtcgggc 1500

gcgacggccg aggcgtcgc gggcaggcc ggggagctgc gtcgttcgt ggccggccgat 1560

gagctgctgc gccccccga cgtcgggcac accctggtct tcggccgctc ggacctcga 1620

caccgcgcag tgcctcgcg ctccgaccgg gaaacctgc tgcgcgtct ggacactctg 1680

gcaggggagg ggccggaca cggctcgtc gtacgggca tggcggccgc cggggccggt 1740

gcgggctgg tgctcgtct cccgggacag ggccggcagt gggccggcat ggggctgcgg 1800

ctgctggaga cctcgtcgt ctccgcccag cggatggcgg agtcgaggc ggcgttgca 1860

ccgtatgccg actggtcgt gctcgacgt ctgcgcggg accccgggga cccggtctgg 1920

gagcgggccc atgtcgtcca gccgatgctg ttctcggtga tgggtgctg ggccgagctg 1980

tggcgctcgt acggcgtca accggacgc gtactcggcc actcccagg cgagatcgcc 2040

gccgcccaca tctcggcgc gctgacctg gacgacccg cgaaggttgt cgcgtcgc 2100

agccggggcc tgcagacct gcgcggttcg ggccgcatgg cctccgtacc actgccggcg 2160

gacgaggta cgggctgct gcggaccgac tctctgtggg tggccgccgt caacgcccc 2220

acggccacgg tgatctccg gcacgcggac tctctggcgg aggcgtgga aactaccgg 2280

gaccaggcgc tcgaagcga gcgggtccc gtcgactac cctccactg cccgcatatc 2340

gaagccgtgg agcaggagct gctgggcctg ttgcgggga tcgtccaag ggccgccgac 2400

atcccttct actccacct ggacaaccag tggccgaca ccatgggact cgacgcccg 2460

tactggtacc gcaatctgc cggcccgta cgttcgccg aagcgtccg cgccctcgcc 2520

gccgccgagt accggacgta tgcgaggtc ggccgcacc ccacctcac cccgccatc 2580

gaggacacca ctgaggccg cggcgtcgc gccacgttg tcggatccct gcgcgcggc 2640

gaggacgacg cccaccgat cctgacctg ctggccggg ctcataattca tggcctgccc 2700

gtggcgtgg accgccata cgggcgctc gccccgagg cgaacctgt cgacctgcc 2760

acctacgcct tccagcgccg ccgctactgg ctggacgccc cggcgaccac cggggacgtg	2820
acggccgcgg ggctggcccc ggtcggacac ccactgctcg gcgcggcggt cggactcgcc	2880
gaggcgacg gatatctgct caccggccgg ctgccccgc acaccaccc ctggctcacc	2940
gaccacgcgg tcgccggcac cgtcctgctg ccgggcaccg catacgtgga actggccgtg	3000
cacgtcggcg gacacctcgg ctgccccgg ctggaggagc tcacctgca cgcgccgctc	3060
gtctccccg acaccggcgg cgtggcgctc caggtggccg tcggggcacc ggacgagacc	3120
ggccgccgcg cactgagcgt ctacgcacag cgcgacgacg accccgcgtg ggagggggcg	3180
gccccggggc cgtggacacg gcatgcgacc ggcacactgg cggccgaggc cccgactgat	3240
ggcatcagcg gtgccgacgg tgccgggacc ctggcggggg cgtggcctcc gccgggcgcg	3300
gagcccttgg acatcagcgg cctctacgac acgttgccg ccgcagactt cggctacggc	3360
ccggccttcc aggggctcg cgcgtctgg cggcaaggcg aggagacctt cggcgaggtg	3420
cggctccccg accagggtgg cgcgcagcc ccacgttct gcctccacc cgcgtgctc	3480
gacgccgcg tccaccgct ggcactcgac agcggccgaa gcgaggagaa tccagcggga	3540
catggcctgc tgccgttcgc ctggcgggc gtcagcctgc gctccccggg cacaccgacg	3600
ctgcgcgtac ggctgcggcc gcaggggccc gactcgattg ccgtcgacgt ggccgacgag	3660
acgggcgcgc cggctggcctc ggccgaatcg ctacgctgc ggccggtggc cctggaggac	3720
ctgcggggcc tcggcgccca ggccggcgac accctctacg ccctggagtg gaccgccgcg	3780
cccgagcccc cggcgacggc cctcggggcg tgcgtgtga ttggccaagc cattcctgga	3840
tgggctgccg cgttgagac ggccgcagcg gggcccgtac ggcggtaccc ggaccttgc	3900
ggactggtga cggccctgga cgcggcgat ccgcctccgg acctggtgtt cgtgggctgc	3960
cctccggctg ccgccgggcc cgacgacacg acggtcgccg acgtccacac cgccttacc	4020
cgtgtccgta cccgacaagc gctggacctg cttcagggt ggctcggcga agcgcggctg	4080
gccggcgca ggctggtgct ggtcacctgc ggcgcggtgg ccaccgggcc ggcggaggga	4140
gtgatggacc tgccgggcgc ggcgatctgc ggactggtgc gatccgcga gcccgaggag	4200
cccgaccgta tctcctggt ggacctggac gcggccgagg agtcgtgggc ggcgctacca	4260
cgggcggtcg cgttggcga accgcagatg gccatccggg ccggccagcc gcacatggcc	4320
cggctggttc gagccgacac cgaggggggc gccctgtca cgcgccaca ggggagcggc	4380
ggctggcggc tcgactgcgc cgacggggc acggtccagg ggctggcc tgtggcgtcc	4440
tcggccgacc gcgaccgct gggccgcac caggtacgga tcgaggtgcg tcggccggg	4500
ctgaacttcc gcgatgtct ggtggccctg gggatggtcc ctgggcagcg ggggctgggc	4560
agcgaggcg ccggggtggt gctcgaagc gggcctgaag tggccgacct ggcgccggg	4620

gaccgggtga tggcggtgtt cgcggatgcg ttcggcccgt tcgcgatcgc cgaccgggcc	4680
accgtgatcc ggcgtcccca ccactggacc ttcggccagg ccgccgccgt ccccgctctg	4740
ttcgccaccg cctattacgg gctggtggac ctggcaggac tgcgcccggg tgagtcggtg	4800
ctggtgcacg ctgcggccgg cggagtggga ctggccgctg tccaactggc ccgccacctg	4860
ggcgctgagg tctacgccac ggcgagcccc ggcaaatggg acaccctacg cccccacggc	4920
atccccccgg agcgcatcgc ctcgccccgc accctcgact tcgagagccg gttcacccgc	4980
cggaacatcg acgtcgtcct caactccctg gcccatgagt acgtcgacgc ctcgctgcgc	5040
ctggtgtccg gcgacagcgg ccggttctcg gagatgggca agaccgacct ccgcgacccg	5100
gaggaggtgg cgcaggcgta ccccggtgtc gcctaccggg cgtacgacct gatggaggcc	5160
ggacccgagc gcatcgggga gatcctgcgc accgtgttgc ggctgttcga cgagggcgtg	5220
ctcacccgcg tgccgctcac ctgctgggac atccggcagg ccagggatgc ctccgccaa	5280
ctccagcagg gccgcacctg cggaaagaat gtgctcacgc tggaccgcac ccccgacccc	5340
gacggcaccg tctcatcac cggtaggcacc ggtaccctcg gcgccgcgt cgcgcccat	5400
ctgcgcgcca ccggcggagc acggcatctg ctactgatca gccgccgtgg ctcgatgcg	5460
ccaggcgctc ccgaactcat cgctgagatt gacgagttgg gcgccacggc gaccgtcgcc	5520
acctgcgacg tcggcgaccg tgccgcgtc gccgaactgc tcgggcggat ccccgccgag	5580
caccgcgtga ccgccgtcgt ccacgccgcg ggacacctcg acgacgccac gctcggtcc	5640
ctcacccgcg gccacctga caccgttctg ccgcgaagg ccgatgccgc ctggcatctg	5700
cacgacctga cctgccggct ggatctggcc gcgttcgtgc tgttctcgtc cgcgcgggt	5760
gtcttgggt cgcgggggca gggcaactac gccgccgcca acgcctttct cgacgcgtc	5820
gccttcacg gacggcgat gggactcccc gccgtgtccc tggcatgggg actgtgggag	5880
gaggccagcg gaatgaccgg ccacctgac cagaccgacc gcacccgcat ggccccgctc	5940
ggcctccggc cactggccac ggacgaggcc ctggcgctgt tcgacaacgc tctcgtcgc	6000
ggccccaccg tctcgtccc ggccccgtatc gacaccaagg cgtacgggg caccaccga	6060
ccgcccctgt tcagagcct cgtacgcccc accaccggcc accggccacg ccccgcgaca	6120
cccgacggcc gtcctccct ccgagcccgg ctgcggggc tcgacccgc cgcacagcac	6180
gaggctctgc tcacctcgt ccgcggccac gccgccacgg tctcggcca cccgagcccc	6240
gacgccatcg ccccgaggc ggccttccgt gacctcggct tcgactccct caccgcgtg	6300
gagctccga accgcctcaa ggaggcaacc ggctgcggc tcccccccc cccccgctc	6360
aaggaggcaa ccggcctcgc gctccccgcc accatcgtct tcgacctcc cactcctgcc	6420
gctctcgccc agcacctcgc ggacggcctc atcggcggcg ccgatacggt caccctggct	6480

gcggctcctg ctccgagcaa ggtggcgaatg gtggcggatg aggccatcgc gatcatcggc	6540
atggcctgcc ggtatccggg gggcgtgcgg tcggccgagg ggctgtggga tctggtcgcc	6600
tccggcaccg acgcatgag cggattcccc agcgaccgcg gctgggacct cgaccgcctc	6660
tacgcccccc aggaccagga cgtgccgggc accacataca cccgccacgg gggttctctc	6720
cacgacgcgg gcaagttcga cgcgggattc ttccgcatcg gccacgtga ggcgctggcg	6780
atggatccgc agcagcggtc gctgctggag acctcctggg aggttttcga acacgcggga	6840
atcgaccctt cgtcgttacg gcggagccgg accggagtct tcgccggtgt gatgccgacg	6900
gactacggcc cccggctgca agacaccgtg gccgaggtcg agggctatgt cctcaccgga	6960
aactccggca gcgtcgctc gggccgtatc gcctacacct tcggcctgga agggcccgcg	7020
gtgtcgttgg acacggcgtg ttctgtctct ctggtggcgt tgcattcgcc gtgtcaggcg	7080
ctgcgtgcgg gggagtgtc catggcgtg gccggcgggg tgacggtgat ggcgacgcct	7140
ggtgccttcg tggagtttgc gcggcagcgg gggttgtcgg tggatggcg gtgcaaggcg	7200
tttggggtgg gtgcggatgg tacggggtgg gcggaggggg tgggatgct gtgtgtggag	7260
cggttgtctg atgcgcggcg gttggggcat cgggtgttgg cggtggtgcg gggttctgcg	7320
gtgaatcagg atggtgcgtc gaatggtttg acggcgccga atggtccgtc gcagcagcgg	7380
gtgatccggc aggcgttggc cagtgcgcgg gtggcgggg cggatgtgga tgtgtggag	7440
gggcacggta cggggacgcg gctgggtgat ccgatcagg cgaggcggtt gctggcgacc	7500
tacggtcagg aacccccga tgatcgacct gtctggttgg ggtcggtgaa gtgcaatc	7560
gggcattgcg aggccgcgc gggggttgcg ggtgtcatca agatggtgat ggcgatgcgg	7620
tatggggtgt tgccgcggac gttgcatgtg caggagccgt cgcccatgt ggactggtcc	7680
tcgggcgggg tcggctgct gacggaggcg gtccgtggc cggagacggg gcgtgcgcgg	7740
cgtgcggggg tgtcgtcgtt cggggtcagt ggcaccaacg cgcacatcat cctcgaacag	7800
gcgccgcctg aggagcacga cgatccggcg gacgtctcgt ccgggtcgtt tccgtggatg	7860
gtgtcggcca agtccgaaca ggcactacag gcgcaggcag cacagttagc cgcgtatctg	7920
gcggcacatc ctgagctggg gctggctgat gtcgggtatg cgctggcctc cggccgcacg	7980
gccttcggcc accgtgccgt gtccttgggc ccggaccgcg aagccttcgt cgaagagctg	8040
ggagctctgg aggccgtga ggaacacgcc gggctgttac ggggcgtggc gacgggtgcg	8100
gggaagctgg cgtttgtgtg ttccgggcag ggaacgaac gtccccgtat gggacacggg	8160
ctgtactcgc cttcccgtg ttccgcgcag ccatggacga agcctgcgca cacctggacc	8220

cacacctga ccatccctg cgggatgtca tgttcgccga gccgggcacc gacaccgccc	8280
agctgtctca ccagaccgc tacgcccagc ccgcgtgtt cgccctccag gtcgacctgc	8340
accgctggt caccgaacac caggccctta cccccacta ctacgccgc cattccctcg	8400
gagagatcac cgcggccac ctgcgggga tctcacctt ccccgacgc gcccgctgg	8460
tcaccaccg cgcgcctc atgcaatct tccccccac cggcgccaat gaccaccctc	8520
caagcagacc ccgacgaact ccacgaacac ctcacacgat gcgaaggac ggtctcactc	8580
gcggcgtga acgcgcccg gtccgtgttc atcagcgtg atcgccacga cgtagacgt	8640
acggccgaaa acctccgcg catgggacgc aagaccactg cgctgaaggt cagcggcgt	8700
ttccactcac accacatga cccactctc aacgaactcc gcaacacggc agaaacctc	8760
acctaccacc caccacac cccctcatc accaccaacc ccaccgacca cgacccacc	8820
acacccact actgggtcg gcaagcgcg gagacgttc actacgcca caccaccaa	8880
caactccaca cccagggct caccgcctac ctggaactcg gcccgacca caccctacc	8940
gccctaccc accacaact ccccgaccac acccgctag ccgtccgct tctccacc	9000
gaccaatcg agaccacac caccacacc gccctcgcc acctccac ccacggccac	9060
cccaccact ggcaccacca tcacacccc acccactacc acccaaact cccacctac	9120
cccttcaac accaccacta ctggctcaac accaccactg ccaccggtga tatgtcggt	9180
gcaggccttg agcggcgcg gcatccctg ttggcgcggg cggtcgggtt ggccgatggt	9240
gaggggttgc tgttcaactg gcggatttct ctccgtacgc atccctggct ggccgaccac	9300
gccgtcggcg gcgcgtgtt gctccccgt acggccttct tcgaactcgc cctccaagcc	9360
gccgcccatt ccgactgccg tcgggtcgag gagcttacgc tccacacccc gtcgtcgta	9420
ccgatagcg cggcgtagt gctgcaggtc actgtggccg cgccgaacga agcaggaaac	9480
cggcggtgg atatctact gcgaatgat gtgcgggcc tcaccgccga ttcggtggc	9540
gagcgtgga cgcgccatgc cgcgggtac ctgcccaca agcctgacc agactgcgt	9600
gactcggcg atggtgtcat gcccggggc gcatggccgc cgcgggtgc ggtcgccgtg	9660
gatctggagg gactgtacga gcaactggc gaggggggtt tccactacgg tgcggccttc	9720
cgttgcttg acgccgctg gcaacgcggg gacgaggtct tcgacccgc gtatatgtca	9780
gaggatcag tggcgacac ggctcggct cggttcgcg tgcacccgc gctgctggat	9840
tccgactgc acaccattcc acttttgc tccctacggg gacaacagga cagcgggctg	9900
ccgttcacgt ggacaggagt caccctgcgt gcatccgggg cgacggctct gcgcgtccg	9960
ctgagccgg acggccatgg cccggggcg gtgtcggtcg acgtgtccga cgaggcgggt	10020
gagccgtag catcggtccg gtcgttgcc ctgcggccgg tgaccaggc cgagtgcac	10080

acggccgagt tgcgcacagc cgccccggtt gcccccatg gctcgctctt cgaggtagcga	10140
tgggaacccg tccccagcc ttcagcgcc gaagaagccg ccccatgggt gatgatcggg	10200
accgggccga cgctgcgccc ggtcgaggac ttcgtcactc cgccggagcg gacgtacgcc	10260
gacctggccg cgctgtgcgt ggcaatcgcc gatgacgcgc ccgttccccg gacggtagtg	10320
gcctggtccc cagccgggag cgaagacgag tcgagtgagg cgctgcgcca ggccacacac	10380
cacatgctgg gctactgca gcagtgggtt gcggacagcc ggttcgccga cagtcgcctg	10440
gtgatcctca cccgagccgc ggtggccact gcgccggacg aggaggtaga agacctggcg	10500
ggagcggcgg cgcggggtct gatccgctcc gccagtcgg agcacctga ccgattcgtc	10560
ctgctcgacc tggacgaccg tcccgtgac gcgaaagacc acgaccgaat gctgtcgatg	10620
gccctggcct gcggggaacc ggaagtggcc gtacgcgatg gagccctgcg cacaccccg	10680
ctgagcccg cgccggcac gccaccgag gccatggacg agcatccctg ggatcaggac	10740
ggcacctac tcatcacgg cggcaccggc agcctcggcg ccatgcttgc ccgccacttg	10800
gtggccacc atggcgtacg gcactgatg ctgatcagcc gacgtggcct cgacgccccg	10860
ggggccaggc gactgggggt cgaacttgcg gagctcgggg cgaggtgac gatcacgcg	10920
tgcgatccg cagaccaag gcaacttgcg aacgtattgt cggagatctc cgtcgacat	10980
ccgtgaccg ctgtggtgca tgcggcaggc gtactggacg acggggtgat cacatccctc	11040
acaccggagg gctgaccca tgtctgcgg gccaaagtgc attcggcgct caatctccac	11100
cagctcacac gcgacctgcc gctgtccgcg tttgtgctct tctcctcgct ggccggggtg	11160
atgggttcgg cagggcagg caactacgc gccgccaacg cagccctgga cgcgtggcg	11220
agtcaccgga gggccgctcg gctgccggcg gtgtccttgg cctggggagt ttgggagcag	11280
accgagggca tgaccggca gttggaggcc acggaccag cgcggtccg ccgctcgggc	11340
ctgaggccgc tggccatcag cgaggccctg gagctcttcg acaaggccct gagctgtgga	11400
cacgccctgg tggtagccgc cgcactcagc acgagggagc ttcagacatc cggatccgtc	11460
ccgccattcc tgcgccacct gacgggtgtc gctccggccc ggccgtccc gaccccgac	11520
gcctcggccg gtgagccgac ctccctgcgg cggcggttga ccggcctcg gccggaagaa	11580
cggtacgcg aggtgctgcg gctggtgcgc tcccggcgg ctgcggtgct ggggcacggc	11640
acggccgaat cggtagccgc ggaactcggc ttcgcgacc tggggttcga ctccctgcc	11700
gcggtggacc tgcggaaccg gttgcagcag gccaccgggc tgcgcctgcc ggccggcttg	11760
atcttcgacc ggccgcgtcc ggacgtgtc gcccgtttc tgtgtgacga gttggccggc	11820
gccggcggtg cgtcggcgcc caggccgcc ccaccgttg cggccggcgg gggggggggc	11880
cgcgggggag ccggtggcca tcgtcggcat ggcatgccgg tttccgggag gtgtgcggtc	11940

ggccgagggc ctgtgggac tggtcgctc cggtatggac gcgtgggtga cttccccgca	12000
gaccgaggct gggagggtgga acggctctac gaccccgacc cggaccgaac cggcacctcc	12060
tacacccggc aaggcgggtt cctctacgac gcgggtgagt tcgacgcggc attcttcggg	12120
atcgcccgcc gtgaggcggc agccatggat ccacagcagc ggctgctgct ggagatctcc	12180
tgggaggcgc tggaaactgc ggggatcgac ccggcgtcgc tgcgggggag ttcgaccggg	12240
gtgttcgctg gggatgatga ccacgactac ggcacccgcc tgcgcgagat cccagagggc	12300
tacgagggct atatcgcaa tggaaacgcg ggcagcgtcg cgtcgggacg tgcgcctac	12360
accttcggcc tggagggggc ggcggtcacc gtggacacgg cgtgttcgct gtccttggtc	12420
gccctgcatc tggcctgcca ggcgtgctgg tcaggggagt gtcctatggc gctggccggc	12480
ggggtcaccg tcatgtccac cccaccact ttgtcgagt tctcgcgcca gcggggactg	12540
gccccggacg ggcggtgcaa gtccttcggg gccggcgcgg acggaacagg ctgggcggag	12600
ggggcgggga tgctcctggt ggaacggctt tcggacgccc ggcgcaacgg ccaccgggtc	12660
ctggcgggtg tacgggggag tgcggtcaac caggacgggg cgagcaatgg gctgacggcg	12720
ccgaacggcc cgtcgcaaga gcgggtgac cgccaggcgt gggcaaacgc ggggtgtggcc	12780
gcgatggaca tcgacgcggt ggaggggacac ggcacgggga cgacgctcgg tgacccatc	12840
gagggccagg cgtgctggg gacgtacgga caggacggc cggccgatcg gccgttgtgg	12900
ttgggatcga tcaagtcaa cgtcggacac acccaggccg ccgcgggggt gggcggcgtc	12960
atcaagatgg tgatggccat gcgccacggg ctgctcccg agaccctgca cgcgaggag	13020
ccctcacctc atgtggactg gtcgggcggg acggtgcggt tgctgaccga gtcggtggcc	13080
tggcccgagc aggggcggat gcgccgtgcg ggcgtctcct ctttcggtgt cagcggtacc	13140
aacgcccacg tcatcctgga acaagcaccg cctgcccgcg agaccacga accggcagag	13200
cccaacaccg cgccaggccc actgccctgg gcgatctccg cgaagagccc gcaagcgcta	13260
cgtgcccagg cgcgccaact gcacacgtac ctgaccaacg cccccgagc gaaccccgcc	13320
gacgtcgcc acaccctcgc gacgggccgc gcctctttcg agcatcgtgc tgtggtcatc	13380
ggctccgacc gagcggagt cctgggtggc ctggatgctc tggcggccga cgaggccac	13440
accgccgtcg tcacggggat cgcgaggaag gccggtgacc agggcaaggt ggtgttcgtg	13500
ttccccgggc agggcggtca gtgggcccgg atgggactgc ggctgcttaa gacctaccc	13560
gtcttcgccc aatcgatcca ggcctgcgaa caagccctcg cccccacac cgactggacc	13620
ctgaccgaca tctgcaccg gccccacac gacccctgt ggcagcgcgc cgacgtcatc	13680

cagcccgctcc tcttcgcctt catgacctcc ctgcgcgccc tctggcaatc ccacggcctt	13740
aaccccgacg ccgtcatcgg ccactcccaa ggcgaaatca ccgccgccca catcagcgga	13800
gcgctgagcc tggaggacgc cgcgaaaacc gtgcgcgtgc gcagccgggc cctgcagacc	13860
ctgcgcggtt cgggcggcat ggcctccgta ccactgccgg cggacgaggt caccgggctg	13920
ctgcggaccg gactctctgg cggaggcgcc ccccccccg ccacgggtgat ctccggcaac	13980
gcggaagctc tcacacaggc gctggaacac taccgggacc aaggcgctga cgcgaaacgg	14040
atcccggtcg actacgcctc ccactgcccc cacatccagg ccgtggaaca ggaactgtca	14100
cggtgtgtgc ggggcatcac cccacgggcc gccaccaccc ctttctactc caccaccgac	14160
aaccaatgga ccgacaccac caccctcaac gccactact ggtaccgaaa cctccgcaa	14220
cccgtccacc tcgccgacgc catcaccaac ctacccacc aaggccacca caccttcac	14280
gaaatcagcc cccacccac cctcaccccc gccatccaag aaaccaccga caccacccac	14340
acccccacca ccgtcatcag cacactccgc cgcaaccaca acgacacca ccaaactctc	14400
cacgcccctg cccacgcca caccaccggc caccatca actggcacac caccaccaa	14460
caccacacc caaccccca acacatcgac ctaccacct acccttcca acaccaccac	14520
tactggctca acacccccc ccagacaggg gatgcggcgg ccgtcggcct ggacccggca	14580
catcacccgt tgcctggcgc ggcggctcgc gtggccgagg gggagggcta tctgtcacc	14640
ggtcggctcg cctgtccac ccacccctgg ctgcgcgac acaccatcg cggcgcggtt	14700
gtcctccctg gaactgcct tctcgagatc gcccttcagg cgggccatcg tgtggactgc	14760
tggcgcatcg aagaactcac cctccaatca ccgtgttca tcccgaaga gggagcagta	14820
caggtgcagg catgggtggc ggcaccggat gagaacgggt gccgaagcct gacgggtgcc	14880
tcccgacgcg aggttacgta cgaggacgcc acgtgggtgc gccatgccac gggccgggtc	14940
ggccccgcac cggccgacca ggatgaagcc atgcacggc tcaccgacc acaaggcgac	15000
ggagcggcgg cggcggcttg gccaccgcag ggcgtgtcg cgttcaccgc agacgatctg	15060
gagggcctgt acgacgggta cgcggcgcgg ggattcgagt acggcccgt gttccgagc	15120
ctgcggcggg cctggcgacg tggcgaggac atcttcgcc aggtgcgcct tcccgcacg	15180
gcggacggcg acgcctccca gttctccgta cccccgcc tgctggacgc cgcactgcac	15240
gccgcggcct tccggcggc cgacaaactc ccgcacggcg cctgcggtt ctcttcagc	15300
ggggtgaggc tgcacgggcc cggagcgctg accctgcggg tgcgcctcac cccggacggc	15360
caggcgcggg acacgcacgc atggtcggtc gcggtggtcg acggcgagg ggcggcggtg	15420
gcctcgatcg catcgctcgc ggtccgccc gtgtcgacgc aggagtgtg ggcggcctcc	15480
ggtacggcgc ggcgggactc gctcttcgc gtcgagtggg tgaccgcct ggcgcggacc	15540

tcgtcgtccg ttccgcaacg cctggccacg gtggggccca gcgaccgcct cccctcggca	15600
gacgcgtacg cgaacctcgc cgacctggcc gccgcagtgc tggaggcggg gggcccgcg	15660
cccgatgcgg tcgtggtcga ctgcggccgc cgcgatgcgc ggcaccgc cgtgccggag	15720
gacgtaagga cctcacccg gcgcatactg ggtctgctgc aggagtggct ggcggacgag	15780
aggccggcct cgagccggat ggtcgtactg acccgtggtg cggtagccac cactccgggg	15840
gaggacgtgg cggacctggc gggcgcggcg gtgtgcggca tggtagctc gcgcagtcg	15900
gaacatcccg gccggttcgt cctgctggac ctgcacccg acccggacct cgacggcggg	15960
gaagtccac cgaccgtcgt tccggcggct ctgcctgtg gtgagccga gatcgcggtg	16020
cgtgcgaacc ggcacctggt gcccggctg acccgcgttc cggcgtccgt ccccgcccc	16080
gggcgtgttc ccgttccgc cgccgaggca gccgaccgg acaccacgc cacggcgttc	16140
gaccccgacg gcaccgtagt gatcacggc ggcaccggca cccttggcgc gatgctcgcg	16200
cgccatctgg tcagccgtca cgggttacga cacctctgc tggcatcgc acgcggaccc	16260
gacgcacccg gcgccaccga gctgcggcg gaactggccg agctcggcgc cgaggtgacg	16320
gtgcgcgtt gtgacacgg tgaccgaggc gcgctggcgg atctcatcgc ggggattccc	16380
accggccacc ctttgaccg tgtgttccac gctgcggcg tcctggacga cgccaccgtc	16440
gcctcgtca cccccgaca cctggacacc gcgctgacac ccaaggccga cgccgccttc	16500
catctgcacg agtcacccg ccacgcccgg ccgcgcct tcgtcctgtt ctctcggcc	16560
gccggtgtcc tcggcgacg cgggcagggc aactatgcgg ccgccaacgc tttcctcgac	16620
gccctcgccg aacaccgcag ggcgcagggc ctgccgcct tgtcgtcgc gtggggcctg	16680
tgggagcagg gcagcggcat gaccgggcat ctgcaccga ccgaccgggc cgcataaac	16740
cgctccggac tcgccccct gccacggag gacgtctcg cgctcttga cgccgcctc	16800
gccggcgatc ggccgttct ggtgcccgc cggctggacc tgcggggttc aagcgccgc	16860
gagaccccg cgccgtgtt ctccaggatc gcccggctc gtacgaccg gggccggtcc	16920
cccggcgccg agggcgccgc tgaccttctg acccgtctcg cggcccagga cgccgccgag	16980
cagcgcgaca cgtttctac gatcgtcgc acccacacc cgcgcgtct ggggcatgac	17040
acggctgccg ccgtgcggcc ggacggggcc ttccgtgaac tgggtttcga ctccctgcc	17100
gccgtggaac tccgtaaccg ctttcaaag accaccgcc tcacctgcc cgcgaccacc	17160
gtcttcgacc accccaccc cgctgccctc gccgatcatc tgcgtactca gctctgccag	17220
gacgtcagt cctcggcggc ggccacggcc atggcgcgga tggcggagct ggccaggctg	17280
gagtcgcgg tctccgattc ggtggcgctc gacgacgaca cgcgacggc cctcgcgag	17340
cgctgcgggt ccctcgcccc caagatgagc agtggccgtg tcgtcgacca cgacggcggc	17400

ggcgctgcgg acctggatct tcagtcggtc acggacgatg agatgttcga gctgatcgac 17460

aaggaggtca gccgagactg a 17481

<210> 78

<211> 5826

<212> PRT

<213> Artificial Sequence

<220><223> milA3 protein of Streptomyces bingchenggensis

<400> 78

Met Ala Ala Gly His Asp Lys Val Ile Glu Ala Leu Arg Ala Ser Leu

1 5 10 15

Lys Thr Asn Glu Arg Gln Arg Glu Gln Ile His Arg Leu Thr Thr Ala

20 25 30

Ala Arg Glu Pro Ile Ala Ile Ile Gly Met Ala Cys Arg Tyr Pro Gly

35 40 45

Gly Val Gly Ser Pro Glu Asp Leu Trp Glu Leu Val Ala Ala Gly Arg

50 55 60

Asp Ala Ile Gly Thr Phe Pro Glu Asp Arg Gly Trp Asp Val Glu Arg

65 70 75 80

Leu Tyr Asp Pro Asp Pro Glu Arg Ala Gly Thr Ser Cys Thr Gln His

85 90 95

Gly Gly Phe Leu Tyr Gln Ala Gly Glu Phe Asp Pro Gly Phe Phe Gly

100 105 110

Ile Ser Pro Arg Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu

115 120 125

Leu Glu Ile Ser Trp Glu Val Phe Glu Arg Ala Gly Ile Asp Pro Ala

130 135 140

Ser Val Arg Gly Ser Arg Thr Gly Val Phe Ala Gly Val Met Tyr His

145 150 155 160

Asp Tyr Gly Ser Arg Leu His Thr Val Pro Glu Gly Phe Glu Gly Tyr

165 170 175

Val Gly Asn Gly Ser Gly Gly Gly Val Ala Ser Gly Arg Val Ala Tyr

180	185	190	
Thr Leu Gly Leu Glu Gly Pro Ala Val Thr Val Asp Thr Ala Cys Ser			
195	200	205	
Ser Ser Leu Val Ala Leu His Leu Ala Cys Gln Ala Leu Arg Ala Gly			
210	215	220	
Glu Cys Ser Leu Ala Leu Ala Gly Gly Val Thr Val Met Ser Thr Pro			
225	230	235	240
Ser Leu Phe Val Glu Tyr Ser Arg Gln Arg Ala Leu Ala Ala Asp Gly			
245	250	255	
Arg Cys Lys Ala Tyr Gly Ala Gly Ala Asp Gly Thr Gly Trp Ala Glu			
260	265	270	
Gly Ala Gly Met Leu Leu Val Glu Arg Leu Thr Asp Ala Gln Arg Leu			
275	280	285	
Gly His Arg Val Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp			
290	295	300	
Gly Ala Ser Asn Gly Leu Thr Ala Pro Asn Gly Pro Ala Gln Gln Arg			
305	310	315	320
Ala Ile Arg Gln Ala Leu Ala Ser Ala Gly Val Ser Ala Ser Glu Val			
325	330	335	
Asp Ala Val Glu Gly His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile			
340	345	350	
Glu Ala Gln Ala Leu Leu Ala Thr Tyr Gly Gln Gln Arg Pro Ala Asp			
355	360	365	
Arg Pro Leu Trp Leu Gly Ser Met Lys Ser Asn Val Gly His Ala Gln			
370	375	380	
Ala Ala Ala Gly Val Gly Gly Ile Ile Lys Met Val Met Ala Met Arg			
385	390	395	400
Ser Gly Thr Leu Pro Arg Thr Leu His Ala Asp Glu Pro Ser Pro His			
405	410	415	
Ile Asp Trp Asp Ser Gly Ala Val Arg Leu Leu Thr Glu Pro Val Ala			
420	425	430	
Trp Pro Glu Arg Asp Arg Pro Arg Arg Ala Ala Val Ser Ser Phe Gly			

435 440 445
 Val Ser Gly Thr Asn Ala His Val Ile Leu Glu Ala Ala Ser Gln Thr
 450 455 460
 Ala Pro Gln Thr Asp Ser Ala Ser Gln Ala Glu Thr Asp Asp Ala Pro

 465 470 475 480
 Ala Pro His Gly Ala Pro Gly His Ala Val Ala Gly Pro Leu Leu Trp
 485 490 495
 Pro Leu Ser Gly Ala Thr Ala Glu Ala Leu Arg Ala Gln Ala Gly Glu
 500 505 510
 Leu Arg Arg Phe Val Ala Ala Asp Glu Leu Leu Arg Pro Ala Asp Val
 515 520 525
 Gly His Thr Leu Val Phe Gly Arg Ser Asp Leu Ala His Arg Ala Val
 530 535 540

 Val Leu Gly Ser Asp Arg Glu Thr Leu Leu Arg Ala Leu Asp Thr Leu
 545 550 555 560
 Ala Gly Glu Gly Pro Asp Asp Gly Ser Val Val Arg Gly Met Ala Ala
 565 570 575
 Ala Gly Ala Gly Ala Gly Val Val Phe Val Phe Pro Gly Gln Gly Gly
 580 585 590
 Gln Trp Ala Gly Met Gly Leu Arg Leu Leu Glu Thr Ser Ser Phe Phe
 595 600 605
 Ala Glu Arg Met Ala Glu Cys Glu Ala Ala Leu Ala Pro Tyr Ala Asp

 610 615 620
 Trp Ser Leu Leu Asp Val Leu Arg Arg Asp Pro Gly Asp Pro Val Trp
 625 630 635 640
 Glu Arg Ala Asp Val Val Gln Pro Met Leu Phe Ser Val Met Val Ser
 645 650 655
 Leu Ala Gln Leu Trp Arg Ser Tyr Gly Val Glu Pro Asp Ala Val Leu
 660 665 670
 Gly His Ser Gln Gly Glu Ile Ala Ala Ala His Ile Cys Gly Ala Leu
 675 680 685

Thr Leu Asp Asp Ala Ala Lys Val Val Ala Leu Arg Ser Arg Ala Leu
 690 695 700
 Gln Thr Leu Arg Gly Ser Gly Gly Met Ala Ser Val Pro Leu Pro Ala
 705 710 715 720
 Asp Glu Val Thr Gly Leu Leu Arg Thr Asp Ser Leu Trp Val Ala Ala
 725 730 735
 Val Asn Ala Pro Thr Ala Thr Val Ile Ser Gly Asp Ala Asp Ser Leu
 740 745 750
 Ala Glu Ala Leu Glu His Tyr Arg Asp Gln Gly Val Glu Ala Lys Arg
 755 760 765
 Val Pro Val Asp Tyr Ala Ser His Cys Pro His Ile Glu Ala Val Glu
 770 775 780
 Gln Glu Leu Leu Gly Leu Leu Arg Gly Ile Ala Pro Arg Ala Ala Asp
 785 790 795 800
 Ile Pro Phe Tyr Ser Thr Val Asp Asn Gln Trp Ala Asp Thr Met Gly
 805 810 815
 Leu Asp Ala Arg Tyr Trp Tyr Arg Asn Leu Arg Arg Pro Val Arg Phe
 820 825 830
 Ala Glu Ala Leu Arg Ala Leu Gly Ala Ala Glu Tyr Arg Thr Tyr Val
 835 840 845
 Glu Val Gly Pro His Pro Thr Leu Thr Pro Ala Ile Glu Asp Thr Thr
 850 855 860
 Glu Ala Ala Gly Val Ala Ala Thr Val Val Gly Ser Leu Arg Arg Gly
 865 870 875 880
 Glu Asp Asp Ala His Arg Ile Leu Thr Ser Leu Ala Arg Ala His Ile
 885 890 895
 His Gly Leu Pro Val Ala Trp Asp Arg His Tyr Arg Ala Leu Ala Pro
 900 905 910
 Glu Ala Asn His Val Asp Leu Pro Thr Tyr Ala Phe Gln Arg Arg Arg
 915 920 925
 Tyr Trp Leu Asp Ala Pro Ala Thr Thr Gly Asp Val Thr Ala Ala Gly
 930 935 940

Leu Ala Pro Val Gly His Pro Leu Leu Gly Ala Ala Val Gly Leu Ala
 945 950 955 960
 Glu Gly Asp Gly Tyr Leu Leu Thr Gly Arg Leu Ala Pro His Thr His
 965 970 975

 Pro Trp Leu Thr Asp His Ala Val Ala Gly Thr Val Leu Leu Pro Gly
 980 985 990
 Thr Ala Tyr Val Glu Leu Ala Val His Val Gly Gly His Leu Gly Cys
 995 1000 1005
 Pro Arg Leu Glu Glu Leu Thr Leu His Ala Pro Leu Val Leu Pro Asp
 1010 1015 1020
 Thr Gly Gly Val Ala Leu Gln Val Ala Val Gly Ala Pro Asp Glu Thr
 1025 1030 1035 1040
 Gly Arg Arg Ala Leu Ser Val Tyr Ala Gln Arg Asp Asp Asp Pro Ala

 1045 1050 1055
 Trp Glu Gly Ala Ala Arg Gly Ala Trp Thr Arg His Ala Thr Gly Thr
 1060 1065 1070
 Leu Ala Ala Glu Ala Pro Thr Asp Gly Ile Ser Gly Ala Asp Gly Ala
 1075 1080 1085
 Gly Thr Leu Ala Gly Ala Trp Pro Pro Pro Gly Ala Glu Pro Leu Asp
 1090 1095 1100
 Ile Ser Gly Leu Tyr Asp Thr Leu Ala Ala Ala Asp Phe Gly Tyr Gly
 1105 1110 1115 1120

 Pro Ala Phe Gln Gly Leu Arg Ala Val Trp Arg Gln Gly Glu Glu Thr
 1125 1130 1135
 Tyr Ala Glu Val Arg Leu Pro Asp Gln Val Ala Ala Asp Ala Pro Arg
 1140 1145 1150
 Phe Cys Leu His Pro Ala Leu Leu Asp Ala Ala Leu His Pro Leu Ala
 1155 1160 1165
 Leu Asp Ser Gly Arg Ser Glu Glu Asn Pro Ala Gly His Gly Leu Leu
 1170 1175 1180
 Pro Phe Ala Trp Arg Gly Val Ser Leu Arg Ser Pro Gly Thr Pro Thr

1185 1190 1195 1200
 Leu Arg Val Arg Leu Arg Pro Gln Gly Pro Asp Ser Ile Ala Val Asp
 1205 1210 1215
 Val Ala Asp Glu Thr Gly Ala Pro Val Ala Ser Ala Glu Ser Leu Thr
 1220 1225 1230
 Leu Arg Pro Val Ala Leu Glu Asp Leu Arg Ala Leu Gly Gly Gln Ala
 1235 1240 1245
 Gly Asp Thr Leu Tyr Ala Leu Glu Trp Thr Ala Ala Pro Glu Pro Pro
 1250 1255 1260

 Ala Thr Ala Leu Gly Arg Cys Ala Val Ile Gly Gln Ala Ile Pro Gly
 1265 1270 1275 1280
 Trp Ala Ala Ala Leu Glu Thr Ala Ala Ala Gly Pro Val Arg Arg Tyr
 1285 1290 1295
 Pro Asp Leu Ala Gly Leu Val Thr Ala Leu Asp Ala Gly Asp Pro Pro
 1300 1305 1310
 Pro Asp Leu Val Phe Val Gly Cys Pro Pro Ala Ala Ala Gly Pro Asp
 1315 1320 1325
 Asp Thr Thr Val Ala Asp Val His Thr Ala Arg Thr Arg Val Arg Thr

 1330 1335 1340
 Arg Gln Ala Leu Asp Leu Leu Gln Gly Trp Leu Gly Glu Ala Arg Leu
 1345 1350 1355 1360
 Ala Gly Ala Arg Leu Val Leu Val Thr Cys Gly Ala Val Ala Thr Gly
 1365 1370 1375
 Pro Ala Glu Gly Val Met Asp Leu Ala Gly Ala Ala Ile Cys Gly Leu
 1380 1385 1390
 Val Arg Ser Ala Gln Ala Glu Glu Pro Asp Arg Ile Leu Leu Val Asp
 1395 1400 1405

 Leu Asp Ala Ala Glu Glu Ser Trp Ala Ala Leu Pro Arg Ala Val Ala
 1410 1415 1420
 Leu Gly Glu Pro Gln Met Ala Ile Arg Ala Gly Gln Pro His Met Ala
 1425 1430 1435 1440
 Arg Leu Val Arg Ala Asp Thr Glu Gly Gly Ala Leu Leu Thr Pro Pro

1445 1450 1455
 Gln Gly Ser Gly Gly Trp Arg Leu Asp Cys Ala Asp Ala Gly Thr Val
 1460 1465 1470
 Gln Gly Leu Ala Pro Val Ala Ser Ser Ala Asp Arg Asp Pro Leu Gly

 1475 1480 1485
 Pro His Gln Val Arg Ile Glu Val Arg Ala Ala Gly Leu Asn Phe Arg
 1490 1495 1500
 Asp Val Leu Val Ala Leu Gly Met Val Pro Gly Gln Arg Gly Leu Gly
 1505 1510 1515 1520
 Ser Glu Gly Ala Gly Val Val Leu Glu Ala Gly Pro Glu Val Ala Asp
 1525 1530 1535
 Leu Ala Pro Gly Asp Arg Val Met Gly Val Phe Ala Asp Ala Phe Gly
 1540 1545 1550

 Pro Phe Ala Ile Ala Asp Arg Ala Thr Val Ile Arg Val Pro Asp His
 1555 1560 1565
 Trp Thr Phe Gly Gln Ala Ala Ala Val Pro Val Val Phe Ala Thr Ala
 1570 1575 1580
 Tyr Tyr Gly Leu Val Asp Leu Ala Gly Leu Arg Pro Gly Glu Ser Val
 1585 1590 1595 1600
 Leu Val His Ala Ala Ala Gly Gly Val Gly Leu Ala Ala Val Gln Leu
 1605 1610 1615
 Ala Arg His Leu Gly Ala Glu Val Tyr Ala Thr Ala Ser Pro Gly Lys

 1620 1625 1630
 Trp Asp Thr Leu Arg Ala His Gly Ile Pro Pro Glu Arg Ile Ala Ser
 1635 1640 1645
 Ser Arg Thr Leu Asp Phe Glu Ser Arg Phe Thr Gly Arg Asn Ile Asp
 1650 1655 1660
 Val Val Leu Asn Ser Leu Ala His Glu Tyr Val Asp Ala Ser Leu Arg
 1665 1670 1675 1680
 Leu Val Ser Gly Asp Ser Gly Arg Phe Leu Glu Met Gly Lys Thr Asp
 1685 1690 1695

Leu Arg Asp Pro Glu Glu Val Ala Gln Ala Tyr Pro Gly Val Ala Tyr
 1700 1705 1710
 Arg Ala Tyr Asp Leu Met Glu Ala Gly Pro Glu Arg Ile Gly Glu Ile
 1715 1720 1725
 Leu Arg Thr Val Leu Arg Leu Phe Asp Glu Gly Val Leu Thr Pro Leu
 1730 1735 1740
 Pro Leu Thr Cys Trp Asp Ile Arg Gln Ala Arg Asp Ala Phe Arg Gln
 1745 1750 1755 1760
 Leu Gln Gln Gly Arg Thr Val Gly Lys Asn Val Leu Thr Leu Asp Arg
 1765 1770 1775
 Thr Pro Asp Pro Asp Gly Thr Val Leu Ile Thr Gly Gly Thr Gly Thr
 1780 1785 1790
 Leu Gly Ala Ala Leu Ala Arg His Leu Ala Ala Thr Gly Arg Ala Arg
 1795 1800 1805
 His Leu Leu Leu Ile Ser Arg Arg Gly Leu Asp Ala Pro Gly Ala Pro
 1810 1815 1820
 Glu Leu Ile Ala Glu Ile Asp Glu Leu Gly Ala Thr Ala Thr Val Ala
 1825 1830 1835 1840
 Thr Cys Asp Val Gly Asp Arg Ala Ala Leu Ala Glu Leu Leu Gly Arg
 1845 1850 1855
 Ile Pro Ala Glu His Pro Leu Thr Ala Val Val His Ala Ala Gly Thr
 1860 1865 1870
 Leu Asp Asp Ala Thr Leu Gly Ser Leu Thr Ala Arg His Leu Asp Thr
 1875 1880 1885
 Val Leu Pro Ala Lys Ala Asp Ala Ala Trp His Leu His Asp Leu Thr
 1890 1895 1900
 Cys Arg Leu Asp Leu Ala Ala Phe Val Leu Phe Ser Ser Ala Ala Gly
 1905 1910 1915 1920
 Val Leu Gly Ser Pro Gly Gln Gly Asn Tyr Ala Ala Ala Asn Ala Phe
 1925 1930 1935
 Leu Asp Ala Leu Ala Phe Gln Arg Arg Ala Met Gly Leu Pro Ala Val
 1940 1945 1950

Ser Leu Ala Trp Gly Leu Trp Glu Glu Ala Ser Gly Met Thr Gly His
1955 1960 1965
Leu Asp Gln Thr Asp Arg Thr Arg Met Ala Arg Val Gly Leu Arg Pro
1970 1975 1980

Leu Ala Thr Asp Glu Ala Leu Ala Leu Phe Asp Asn Ala Leu Val Asp
1985 1990 1995 2000
Gly Pro Pro Leu Leu Leu Pro Ala Arg Ile Asp Thr Lys Ala Leu Arg
2005 2010 2015
Gly Thr Thr Ala Pro Pro Leu Phe Gln Ser Leu Val Arg Pro Thr Thr
2020 2025 2030
Gly His Arg Pro Arg Pro Ala Thr Pro Asp Gly Arg Ser Ser Leu Arg
2035 2040 2045
Ala Arg Leu Ala Gly Leu Asp Pro Ala Ala Gln His Glu Val Leu Leu

2050 2055 2060
Thr Leu Val Arg Gly His Ala Ala Thr Val Leu Gly His Pro Ser Pro
2065 2070 2075 2080
Asp Ala Ile Ala Arg Glu Ala Ala Phe Arg Asp Leu Gly Phe Asp Ser
2085 2090 2095
Leu Thr Ala Val Glu Leu Arg Asn Arg Leu Lys Glu Ala Thr Gly Leu
2100 2105 2110
Arg Leu Pro Pro Pro Pro Arg Leu Lys Glu Ala Thr Gly Leu Arg Leu
2115 2120 2125

Pro Ala Thr Ile Val Phe Asp His Pro Thr Pro Ala Ala Leu Ala Gln
2130 2135 2140
His Leu Arg Asp Gly Leu Ile Gly Gly Ala Asp Thr Val Thr Leu Ala
2145 2150 2155 2160
Ala Ala Pro Ala Pro Ser Lys Val Ala Met Val Ala Asp Glu Ala Ile
2165 2170 2175
Ala Ile Ile Gly Met Ala Cys Arg Tyr Pro Gly Gly Val Arg Ser Ala
2180 2185 2190
Glu Gly Leu Trp Asp Leu Val Ala Ser Gly Thr Asp Ala Met Ser Gly

2195 2200 2205
 Phe Pro Ser Asp Arg Gly Trp Asp Leu Asp Arg Leu Tyr Ala Pro Gln
 2210 2215 2220
 Asp Gln Asp Val Pro Gly Thr Thr Tyr Thr Arg His Gly Gly Phe Leu
 2225 2230 2235 2240
 His Asp Ala Gly Lys Phe Asp Ala Gly Phe Phe Gly Ile Gly Pro Arg
 2245 2250 2255
 Glu Ala Leu Ala Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Thr Ser
 2260 2265 2270

 Trp Glu Val Phe Glu His Ala Gly Ile Asp Pro Ser Ser Val Arg Arg
 2275 2280 2285
 Ser Arg Thr Gly Val Phe Ala Gly Val Met Pro Thr Asp Tyr Gly Pro
 2290 2295 2300
 Arg Leu Gln Asp Thr Val Ala Glu Val Glu Gly Tyr Val Leu Thr Gly
 2305 2310 2315 2320
 Asn Ser Gly Ser Val Ala Ser Gly Arg Ile Ala Tyr Thr Phe Gly Leu
 2325 2330 2335
 Glu Gly Pro Ala Val Ser Val Asp Thr Ala Cys Ser Ser Ser Leu Val

 2340 2345 2350
 Ala Leu His Leu Ala Cys Gln Ala Leu Arg Ala Gly Glu Cys Ser Met
 2355 2360 2365
 Ala Leu Ala Gly Gly Val Thr Val Met Ala Thr Pro Gly Ala Phe Val
 2370 2375 2380
 Glu Phe Ala Arg Gln Arg Gly Leu Ser Val Asp Gly Arg Cys Lys Ala
 2385 2390 2395 2400
 Phe Gly Val Gly Ala Asp Gly Thr Gly Trp Ala Glu Gly Val Gly Met
 2405 2410 2415

 Leu Leu Val Glu Arg Leu Ser Asp Ala Arg Arg Leu Gly His Arg Val
 2420 2425 2430
 Leu Ala Val Val Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn
 2435 2440 2445
 Gly Leu Thr Ala Pro Asn Gly Pro Ser Gln Gln Arg Val Ile Arg Gln

2450 2455 2460
 Ala Leu Ala Ser Ala Arg Val Gly Gly Ala Asp Val Asp Val Val Glu
 2465 2470 2475 2480
 Gly His Gly Thr Gly Thr Arg Leu Gly Asp Pro Ile Glu Ala Gln Ala

 2485 2490 2495
 Leu Leu Ala Thr Tyr Gly Gln Glu Arg Pro Asp Asp Arg Pro Val Trp
 2500 2505 2510
 Leu Gly Ser Val Lys Ser Asn Ile Gly His Ala Gln Ala Ala Ala Gly
 2515 2520 2525
 Val Ala Gly Val Ile Lys Met Val Met Ala Met Arg Tyr Gly Val Leu
 2530 2535 2540
 Pro Arg Thr Leu His Val Gln Glu Pro Ser Pro His Val Asp Trp Ser
 2545 2550 2555 2560

 Ser Gly Gly Val Arg Leu Leu Thr Glu Ala Val Pro Trp Pro Glu Thr
 2565 2570 2575
 Gly Arg Ala Arg Arg Ala Gly Val Ser Ser Phe Gly Val Ser Gly Thr
 2580 2585 2590
 Asn Ala His Ile Ile Leu Glu Gln Ala Pro Pro Glu Glu His Asp Asp
 2595 2600 2605
 Pro Ala Asp Val Ser Ser Gly Ser Phe Pro Trp Met Val Ser Ala Lys
 2610 2615 2620
 Ser Glu Gln Ala Leu Gln Ala Gln Ala Ala Gln Leu Arg Ala Tyr Leu

 2625 2630 2635 2640
 Ala Ala His Pro Glu Leu Gly Leu Ala Asp Val Gly Tyr Ala Leu Ala
 2645 2650 2655
 Ser Gly Arg Thr Ala Phe Gly His Arg Ala Val Leu Leu Gly Pro Asp
 2660 2665 2670
 Arg Glu Ala Phe Val Glu Glu Leu Gly Ala Leu Glu Ala Gly Glu Glu
 2675 2680 2685
 His Ala Gly Leu Val Arg Gly Val Ala Thr Gly Ala Gly Lys Leu Ala
 2690 2695 2700

Phe Val Cys Ser Gly Gln Gly Thr Gln Arg Pro Arg Met Gly His Gly
 2705 2710 2715 2720
 Leu Tyr Ser Pro Ser Arg Cys Ser Pro Gln Pro Trp Thr Lys Pro Ala
 2725 2730 2735
 His Thr Trp Thr His Thr Ser Thr Ile Pro Cys Gly Met Ser Cys Ser
 2740 2745 2750
 Pro Ser Arg Ala Pro Thr Pro Pro Ser Cys Ser Thr Arg Pro Ala Thr
 2755 2760 2765
 Pro Ser Pro Arg Cys Ser Pro Ser Arg Ser Pro Cys Thr Ala Trp Ser

 2770 2775 2780
 Pro Asn Thr Thr Ala Leu Pro Pro Thr Thr Thr Pro Ala Ile Pro Ser
 2785 2790 2795 2800
 Glu Arg Ser Pro Arg Pro Thr Ser Pro Gly Ser Ser Pro Ser Pro Thr
 2805 2810 2815
 Arg Pro Ala Trp Ser Pro Pro Ala Pro Ala Ser Cys Asn Leu Ser Pro
 2820 2825 2830
 Pro Pro Ala Pro Met Thr Thr Leu Gln Ala Asp Pro Asp Glu Leu His
 2835 2840 2845

 Glu His Leu Thr Arg Cys Glu Gly Arg Val Ser Leu Ala Ala Val Asn
 2850 2855 2860
 Ala Pro Gly Ser Val Val Ile Ser Gly Asp Arg His Asp Val Asp Ala
 2865 2870 2875 2880
 Thr Ala Glu Asn Leu Arg Ala Met Gly Arg Lys Thr Thr Ala Leu Lys
 2885 2890 2895
 Val Ser Gly Ala Phe His Ser His His Ile Asp Pro Leu Leu Asn Glu
 2900 2905 2910
 Leu Arg Asn Thr Ala Glu Thr Leu Thr Tyr His Pro Pro His Thr Pro

 2915 2920 2925
 Leu Ile Thr Thr Asn Pro Thr Asp His Asp Pro Thr Thr Pro His Tyr
 2930 2935 2940
 Trp Val Arg Gln Ala Arg Glu Thr Val His Tyr Ala His Thr Thr Gln
 2945 2950 2955 2960

Gln Leu His Thr His Gly Val Thr Ala Tyr Leu Glu Leu Gly Pro Asp
2965 2970 2975

His Thr Leu Thr Ala Leu Thr His His Asn Leu Pro Asp His Thr Pro
2980 2985 2990

Leu Ala Val Pro Leu Leu His Pro Asp Gln Ser Glu Thr His Thr Thr
2995 3000 3005

His Thr Ala Leu Ala His Leu His Thr His Gly His Pro Thr Thr Trp
3010 3015 3020

His His His His Thr Pro Thr His Tyr His Pro Asn Leu Pro Thr Tyr
3025 3030 3035 3040

Pro Phe Gln His His His Tyr Trp Leu Asn Thr Thr Thr Ala Thr Gly
3045 3050 3055

Asp Met Ser Ala Ala Gly Leu Glu Pro Ala Arg His Pro Leu Leu Gly
3060 3065 3070

Ala Ala Val Gly Leu Ala Asp Gly Glu Gly Leu Leu Phe Thr Gly Arg
3075 3080 3085

Ile Ser Leu Arg Thr His Pro Trp Leu Ala Asp His Ala Val Gly Gly
3090 3095 3100

Ala Val Leu Leu Pro Gly Thr Ala Phe Leu Glu Leu Ala Leu Gln Ala
3105 3110 3115 3120

Ala Ala His Ala Asp Cys Arg Arg Val Glu Glu Leu Thr Leu His Thr
3125 3130 3135

Pro Leu Val Val Pro Asp Ser Ala Gly Val Val Leu Gln Val Thr Val
3140 3145 3150

Ala Ala Pro Asn Glu Ala Gly Asn Arg Ala Val Asp Ile Tyr Ser Arg
3155 3160 3165

Ile Asp Val Gly Gly Leu Thr Ala Asp Ser Ala Gly Glu Pro Trp Thr
3170 3175 3180

Arg His Ala Ala Gly Tyr Leu Ala Asp Lys Pro Asp Pro Asp Cys Gly
3185 3190 3195 3200

Asp Ser Ala Asp Gly Val Met Pro Ala Gly Ala Trp Pro Pro Pro Gly

3205	3210	3215	
Ala Val Ala Val Asp Leu Glu Gly Leu Tyr Glu Gln Leu Ala Glu Gly			
3220	3225	3230	
Gly Phe His Tyr Gly Ala Ala Phe Arg Cys Leu Asp Ala Ala Trp Gln			
3235	3240	3245	
Arg Gly Asp Glu Val Phe Ala Thr Ala Tyr Met Ser Glu Asp Gln Leu			
3250	3255	3260	
Gly Asp Thr Ala Ala Ala Arg Phe Ala Leu His Pro Ala Leu Leu Asp			
3265	3270	3275	3280
Ser Ala Leu His Thr Ile Pro Leu Leu Pro Ser Leu Arg Gly Gln Gln			
3285	3290	3295	
Asp Ser Gly Leu Pro Phe Thr Trp Thr Gly Val Thr Leu Arg Ala Ser			
3300	3305	3310	
Gly Ala Thr Ala Leu Arg Val Arg Leu Arg Pro Asp Gly His Gly Pro			
3315	3320	3325	
Gly Ala Val Ser Val Asp Val Ser Asp Glu Ala Gly Glu Pro Val Ala			
3330	3335	3340	
Ser Val Arg Ser Leu Ala Leu Arg Pro Val Thr Arg Ala Glu Leu His			
3345	3350	3355	3360
Thr Ala Glu Leu Arg Thr Ala Ala Pro Val Ala Pro His Gly Ser Leu			
3365	3370	3375	
Phe Glu Val Arg Trp Glu Pro Val Pro Gln Pro Ser Ala Ala Glu Glu			
3380	3385	3390	
Ala Ala Pro Trp Val Met Ile Gly Thr Gly Pro Thr Leu Arg Pro Val			
3395	3400	3405	
Glu Asp Phe Val Thr Pro Pro Glu Arg Thr Tyr Ala Asp Leu Ala Ala			
3410	3415	3420	
Leu Cys Val Ala Ile Ala Asp Asp Ala Pro Val Pro Arg Thr Val Val			
3425	3430	3435	3440
Ala Trp Ser Pro Ala Gly Ser Glu Asp Glu Ser Ser Glu Ala Leu Arg			
3445	3450	3455	
Gln Ala Thr His His Met Leu Gly Leu Leu Gln Gln Trp Leu Ala Asp			

3460 3465 3470
 Ser Arg Phe Ala Asp Ser Arg Leu Val Ile Leu Thr Arg Ala Ala Val
 3475 3480 3485
 Ala Thr Ala Pro Asp Glu Glu Val Glu Asp Leu Ala Gly Ala Ala Ala

3490 3495 3500
 Arg Gly Leu Ile Arg Ser Ala Gln Ser Glu His Pro Asp Arg Phe Val
 3505 3510 3515 3520
 Leu Leu Asp Leu Asp Asp Arg Pro Ala Asp Ala Lys Asp His Asp Arg
 3525 3530 3535
 Met Leu Ser Met Ala Leu Ala Cys Gly Glu Pro Glu Val Ala Val Arg
 3540 3545 3550
 Asp Gly Ala Leu Arg Thr Pro Arg Leu Ser Pro Leu Ala Gly Thr Ala
 3555 3560 3565

Thr Glu Ala Met Asp Glu His Pro Trp Asp Gln Asp Gly Thr Val Leu
 3570 3575 3580
 Ile Thr Gly Gly Thr Gly Ser Leu Gly Ala Met Leu Ala Arg His Leu
 3585 3590 3595 3600
 Val Ala Thr His Gly Val Arg His Leu Met Leu Ile Ser Arg Arg Gly
 3605 3610 3615
 Leu Asp Ala Pro Gly Ala Arg Arg Leu Gly Val Glu Leu Ala Glu Leu
 3620 3625 3630
 Gly Ala Gln Val Thr Ile Thr Ala Cys Asp Ala Ala Asp Gln Arg Gln

3635 3640 3645
 Leu Ala Asn Val Leu Ser Glu Ile Ser Val Asp His Pro Leu Thr Ala
 3650 3655 3660
 Val Val His Ala Ala Gly Val Leu Asp Asp Gly Val Ile Thr Ser Leu
 3665 3670 3675 3680
 Thr Pro Glu Gly Leu Thr His Val Leu Arg Ala Lys Val Asp Ser Ala
 3685 3690 3695
 Leu Asn Leu His Gln Leu Thr Arg Asp Leu Pro Leu Ser Ala Phe Val
 3700 3705 3710

Leu Phe Ser Ser Leu Ala Gly Val Met Gly Ser Ala Gly Gln Gly Asn
 3715 3720 3725
 Tyr Ala Ala Ala Asn Ala Ala Leu Asp Ala Leu Ala Ser His Arg Arg
 3730 3735 3740
 Ala Ala Arg Leu Pro Ala Val Ser Leu Ala Trp Gly Val Trp Glu Gln
 3745 3750 3755 3760
 Thr Glu Gly Met Thr Gly Gln Leu Glu Ala Thr Asp His Ala Arg Leu
 3765 3770 3775
 Arg Arg Ser Gly Leu Arg Pro Leu Ala Ile Ser Glu Gly Leu Glu Leu
 3780 3785 3790
 Phe Asp Lys Ala Leu Ser Cys Gly His Ala Leu Val Val Pro Ala Ala
 3795 3800 3805
 Leu Ser Thr Arg Glu Leu Gln Thr Ser Gly Ser Val Pro Pro Phe Leu
 3810 3815 3820
 Arg His Leu Thr Gly Val Ala Pro Ala Arg Pro Ser Arg Thr Arg Asp
 3825 3830 3835 3840
 Ala Ser Ala Gly Glu Pro Thr Ser Leu Arg Arg Arg Leu Thr Gly Leu
 3845 3850 3855
 Gly Pro Glu Glu Arg Leu Arg Glu Val Leu Arg Leu Val Arg Ser Arg
 3860 3865 3870
 Ala Ala Ala Val Leu Gly His Gly Thr Ala Glu Ser Val Pro Ala Asp
 3875 3880 3885
 Ser Ala Phe Arg Asp Leu Gly Phe Asp Ser Leu Ala Ala Val Asp Leu
 3890 3895 3900
 Arg Asn Arg Leu Gln Gln Ala Thr Gly Leu Arg Leu Pro Ala Gly Leu
 3905 3910 3915 3920
 Ile Phe Asp Arg Pro Arg Pro Asp Val Leu Ala Arg Phe Leu Cys Asp
 3925 3930 3935
 Glu Leu Ala Gly Ala Gly Gly Thr Ser Ala Ala Thr Ala Ala Pro Pro
 3940 3945 3950
 Val Ala Ala Gly Gly Gly Gly Gly Arg Gly Gly Ala Gly Gly His Arg
 3955 3960 3965

Arg His Gly Met Pro Val Ser Gly Arg Cys Ala Val Gly Arg Gly Pro
3970 3975 3980
Val Gly Ser Gly Arg Leu Arg Tyr Gly Arg Val Gly Asp Phe Pro Ala
3985 3990 3995 4000
Asp Arg Gly Trp Glu Val Glu Arg Leu Tyr Asp Pro Asp Pro Asp Arg
4005 4010 4015
Thr Gly Thr Ser Tyr Thr Arg Gln Gly Gly Phe Leu Tyr Asp Ala Gly
4020 4025 4030
Glu Phe Asp Ala Ala Phe Phe Gly Ile Gly Pro Arg Glu Ala Val Ala
4035 4040 4045
Met Asp Pro Gln Gln Arg Leu Leu Leu Glu Ile Ser Trp Glu Ala Leu
4050 4055 4060
Glu Arg Ala Gly Ile Asp Pro Ala Ser Leu Arg Gly Ser Ser Thr Gly
4065 4070 4075 4080
Val Phe Ala Gly Val Met Tyr His Asp Tyr Gly Thr Arg Leu Arg Glu
4085 4090 4095
Ile Pro Glu Gly Tyr Glu Gly Tyr Ile Gly Asn Gly Asn Ala Gly Ser
4100 4105 4110
Val Ala Ser Gly Arg Val Ala Tyr Thr Phe Gly Leu Glu Gly Pro Ala
4115 4120 4125
Val Thr Val Asp Thr Ala Cys Ser Ser Ser Leu Val Ala Leu His Leu
4130 4135 4140
Ala Cys Gln Ala Leu Arg Ser Gly Glu Cys Ser Met Ala Leu Ala Gly
4145 4150 4155 4160
Gly Val Thr Val Met Ser Thr Pro Thr Thr Phe Val Glu Phe Ser Arg
4165 4170 4175
Gln Arg Gly Leu Ala Pro Asp Gly Arg Cys Lys Ser Phe Gly Ala Gly
4180 4185 4190
Ala Asp Gly Thr Gly Trp Ala Glu Gly Ala Gly Met Leu Leu Val Glu
4195 4200 4205
Arg Leu Ser Asp Ala Arg Arg Asn Gly His Arg Val Leu Ala Val Val

4210 4215 4220
 Arg Gly Ser Ala Val Asn Gln Asp Gly Ala Ser Asn Gly Leu Thr Ala
 4225 4230 4235 4240
 Pro Asn Gly Pro Ser Gln Glu Arg Val Ile Arg Gln Ala Trp Ala Asn
 4245 4250 4255
 Ala Gly Val Ala Ala Met Asp Ile Asp Ala Val Glu Gly His Gly Thr
 4260 4265 4270
 Gly Thr Thr Leu Gly Asp Pro Ile Glu Ala Gln Ala Leu Leu Gly Thr
 4275 4280 4285

 Tyr Gly Gln Gly Arg Ser Ala Asp Arg Pro Leu Trp Leu Gly Ser Ile
 4290 4295 4300
 Lys Ser Asn Val Gly His Thr Gln Ala Ala Ala Gly Val Gly Gly Val
 4305 4310 4315 4320
 Ile Lys Met Val Met Ala Met Arg His Gly Leu Leu Pro Gln Thr Leu
 4325 4330 4335
 His Ala Glu Glu Pro Ser Pro His Val Asp Trp Ser Gly Gly Thr Val
 4340 4345 4350
 Arg Leu Leu Thr Glu Ser Val Ala Trp Pro Glu Gln Gly Arg Met Arg

 4355 4360 4365
 Arg Ala Gly Val Ser Ser Phe Gly Val Ser Gly Thr Asn Ala His Val
 4370 4375 4380
 Ile Leu Glu Gln Ala Pro Pro Ala Ala Glu Thr His Glu Pro Ala Glu
 4385 4390 4395 4400
 Pro Asn Thr Ala Pro Gly Pro Leu Pro Trp Ala Ile Ser Ala Lys Ser
 4405 4410 4415
 Pro Gln Ala Leu Arg Ala Gln Ala Arg Gln Leu His Thr Tyr Leu Thr
 4420 4425 4430

 Asn Ala Pro Glu Ala Asn Pro Ala Asp Val Gly His Thr Leu Ala Thr
 4435 4440 4445
 Gly Arg Ala Ser Phe Glu His Arg Ala Val Val Ile Gly Ser Asp Arg
 4450 4455 4460
 Ala Glu Phe Leu Gly Gly Leu Asp Ala Leu Ala Ala Asp Glu Ala His

4465 4470 4475 4480
 Thr Ala Val Val Thr Gly Ile Ala Arg Lys Ala Gly Asp Gln Gly Lys
 4485 4490 4495
 Val Val Phe Val Phe Pro Gly Gln Gly Gly Gln Trp Ala Gly Met Gly

 4500 4505 4510
 Leu Arg Leu Leu Lys Thr Ser Pro Val Phe Ala Gln Ser Ile Gln Ala
 4515 4520 4525
 Cys Glu Gln Ala Leu Ala Pro His Thr Asp Trp Thr Leu Thr Asp Ile
 4530 4535 4540
 Leu His Arg Pro His Thr Asp Pro Leu Trp Gln Arg Ala Asp Val Ile
 4545 4550 4555 4560
 Gln Pro Val Leu Phe Ala Leu Met Thr Ser Leu Ala Ala Leu Trp Gln
 4565 4570 4575

 Ser His Gly Leu Asn Pro Asp Ala Val Ile Gly His Ser Gln Gly Glu
 4580 4585 4590
 Ile Thr Ala Ala His Ile Ser Gly Ala Leu Ser Leu Glu Asp Ala Ala
 4595 4600 4605
 Lys Thr Val Ala Leu Arg Ser Arg Ala Leu Gln Thr Leu Arg Gly Ser
 4610 4615 4620
 Gly Gly Met Ala Ser Val Pro Leu Pro Ala Asp Glu Val Thr Gly Leu
 4625 4630 4635 4640
 Leu Arg Thr Gly Leu Ser Gly Gly Gly Ala Pro Pro Pro Ala Thr Val

 4645 4650 4655
 Ile Ser Gly Asn Ala Glu Ala Leu Thr Gln Ala Leu Glu His Tyr Arg
 4660 4665 4670
 Asp Gln Gly Val Asp Ala Lys Arg Ile Pro Val Asp Tyr Ala Ser His
 4675 4680 4685
 Cys Pro His Ile Gln Ala Val Glu Gln Glu Leu Ser Arg Leu Leu Arg
 4690 4695 4700
 Gly Ile Thr Pro Arg Ala Ala Thr Thr Pro Phe Tyr Ser Thr Thr Asp
 4705 4710 4715 4720

Asn Gln Trp Thr Asp Thr Thr Thr Leu Asn Ala His Tyr Trp Tyr Arg
 4725 4730 4735
 Asn Leu Arg Gln Pro Val His Leu Ala Asp Ala Ile Thr Asn Leu Thr
 4740 4745 4750
 His Gln Gly His His Thr Phe Ile Glu Ile Ser Pro His Pro Thr Leu
 4755 4760 4765
 Thr Pro Ala Ile Gln Glu Thr Thr Asp Thr Thr His Thr Pro Thr Thr
 4770 4775 4780
 Val Ile Ser Thr Leu Arg Arg Asn His Asn Asp Thr His Gln Ile Leu

 4785 4790 4795 4800
 His Ala Leu Ala His Ala His Thr Thr Gly His Pro Ile Asn Trp His
 4805 4810 4815
 Thr Thr His Gln His His Thr Pro Thr Pro Gln His Ile Asp Leu Pro
 4820 4825 4830
 Thr Tyr Pro Phe Gln His His His Tyr Trp Leu Asn Thr Pro Thr Gln
 4835 4840 4845
 Thr Gly Asp Ala Ala Ala Val Gly Leu Asp Pro Ala His His Pro Leu
 4850 4855 4860

 Leu Gly Ala Ala Val Ala Val Ala Glu Gly Glu Gly Tyr Leu Leu Thr
 4865 4870 4875 4880
 Gly Arg Leu Ala Leu Ser Thr His Pro Trp Leu Ala Asp His Thr Ile
 4885 4890 4895
 Ala Gly Ala Val Val Leu Pro Gly Thr Ala Leu Leu Glu Ile Ala Leu
 4900 4905 4910
 Gln Ala Gly His Arg Val Asp Cys Trp Arg Ile Glu Glu Leu Thr Leu
 4915 4920 4925
 Gln Ser Pro Leu Phe Ile Pro Glu Glu Gly Ala Val Gln Val Gln Ala

 4930 4935 4940
 Trp Val Ala Ala Pro Asp Glu Asn Gly Cys Arg Ser Leu Thr Val Ser
 4945 4950 4955 4960
 Ser Arg Arg Glu Gly Thr Tyr Glu Asp Ala Thr Trp Val Arg His Ala
 4965 4970 4975

Thr Gly Arg Val Gly Pro Ala Pro Ala Asp Gln Asp Glu Ala Ile Ala
4980 4985 4990

Arg Leu Thr Asp Pro Gln Gly Asp Gly Ala Ala Ala Val Trp Pro
4995 5000 5005

Pro Gln Gly Ala Val Ala Phe Thr Ala Asp Asp Leu Glu Gly Leu Tyr
5010 5015 5020

Asp Gly Tyr Ala Ala Arg Gly Phe Glu Tyr Gly Pro Val Phe Arg Gly
5025 5030 5035 5040

Leu Arg Ala Ala Trp Arg Arg Gly Glu Asp Ile Phe Ala Glu Val Arg
5045 5050 5055

Leu Pro Asp Thr Ala Asp Gly Asp Ala Ser Gln Phe Ser Val His Pro
5060 5065 5070

Ala Leu Leu Asp Ala Ala Leu His Ala Ala Ala Phe Arg Pro Ala Asp
5075 5080 5085

Lys Leu Pro His Gly Ala Leu Pro Phe Ser Phe Ser Gly Val Arg Leu
5090 5095 5100

His Gly Pro Gly Ala Ser Thr Leu Arg Val Arg Leu Thr Pro Asp Gly
5105 5110 5115 5120

Gln Ala Arg Asp Thr His Ala Trp Ser Val Ala Val Val Asp Gly Glu
5125 5130 5135

Gly Arg Pro Val Ala Ser Ile Ala Ser Leu Ala Val Arg Pro Val Ser
5140 5145 5150

Thr Gln Glu Leu Leu Ala Ala Ser Gly Thr Ala Arg Arg Asp Ser Leu
5155 5160 5165

Phe Ala Val Glu Trp Val Thr Ala Leu Ala Pro Thr Ser Ser Ser Val
5170 5175 5180

Pro Gln Arg Leu Ala Thr Val Gly Pro Ser Asp Arg Leu Pro Ser Ala
5185 5190 5195 5200

Asp Ala Tyr Ala Asn Leu Ala Asp Leu Ala Ala Ala Val Leu Glu Ala
5205 5210 5215

Gly Ala Pro Ala Pro Asp Ala Val Val Val Asp Cys Gly Arg Arg Asp

5220 5225 5230
 Ala Arg Ala Thr Ala Val Pro Glu Asp Val Arg Thr Leu Thr Arg Arg
 5235 5240 5245
 Ile Leu Gly Leu Leu Gln Glu Trp Leu Ala Asp Glu Arg Pro Ala Ser
 5250 5255 5260
 Ser Arg Met Val Val Leu Thr Arg Gly Ala Val Ala Thr Thr Pro Gly
 5265 5270 5275 5280
 Glu Asp Val Ala Asp Leu Ala Gly Ala Ala Val Cys Gly Met Val Arg
 5285 5290 5295

 Ser Ala Gln Ser Glu His Pro Gly Arg Phe Val Leu Leu Asp Leu Asp
 5300 5305 5310
 Pro Asp Pro Asp Leu Asp Gly Gly Glu Val Pro Pro Thr Val Val Pro
 5315 5320 5325
 Ala Ala Leu Ala Cys Gly Glu Pro Gln Ile Ala Val Arg Ala Asn Arg
 5330 5335 5340
 His Leu Val Pro Arg Leu Thr Arg Val Pro Ala Ser Val Pro Val Pro
 5345 5350 5355 5360
 Gly Arg Val Pro Val Pro Ala Ala Glu Ala Ala Asp Pro Asp Thr Thr

 5365 5370 5375
 Pro Thr Ala Phe Asp Pro Asp Gly Thr Val Val Ile Thr Gly Gly Thr
 5380 5385 5390
 Gly Thr Leu Gly Ala Met Leu Ala Arg His Leu Val Ser Arg His Gly
 5395 5400 5405
 Val Arg His Leu Leu Leu Ala Ser Arg Arg Gly Pro Asp Ala Pro Gly
 5410 5415 5420
 Ala Thr Glu Leu Arg Ala Glu Leu Ala Glu Leu Gly Ala Glu Val Thr
 5425 5430 5435 5440

 Val Arg Ala Cys Asp Thr Gly Asp Arg Gly Ala Leu Ala Asp Leu Ile
 5445 5450 5455
 Ala Gly Ile Pro Thr Gly His Pro Leu Thr Gly Val Val His Ala Ala
 5460 5465 5470
 Gly Val Leu Asp Asp Ala Thr Val Ala Ser Leu Thr Pro Arg His Leu

5475 5480 5485
 Asp Thr Ala Leu Thr Pro Lys Ala Asp Ala Ala Phe His Leu His Glu
 5490 5495 5500
 Leu Thr Arg His Ala Arg Pro Arg Ala Phe Val Leu Phe Ser Ser Ala

 5505 5510 5515 5520
 Ala Gly Val Leu Gly Ala Ala Gly Gln Gly Asn Tyr Ala Ala Ala Asn
 5525 5530 5535
 Ala Phe Leu Asp Ala Leu Ala Glu His Arg Arg Ala Gln Gly Leu Pro
 5540 5545 5550
 Ala Leu Ser Leu Ala Trp Gly Leu Trp Glu Gln Gly Ser Gly Met Thr
 5555 5560 5565
 Gly His Leu Asp Arg Thr Asp Arg Ala Arg Ile Asn Arg Ser Gly Leu
 5570 5575 5580

 Ala Pro Leu Ala Thr Glu Asp Ala Leu Ala Leu Phe Asp Ala Ala Leu
 5585 5590 5595 5600
 Ala Gly Asp Arg Pro Phe Leu Val Pro Ala Arg Leu Asp Leu Arg Gly
 5605 5610 5615
 Ser Ser Ala Ala Glu Thr Pro Ala Pro Leu Phe Ser Arg Ile Ala Pro
 5620 5625 5630
 Ala Arg Thr Thr Arg Gly Arg Ser Pro Gly Ala Glu Gly Ala Ala Asp
 5635 5640 5645
 Leu Arg Thr Arg Leu Ala Ala Gln Asp Ala Ala Glu Gln Arg Asp Thr

 5650 5655 5660
 Leu Leu Thr Ile Val Arg Thr His Thr Ala Ala Val Leu Gly His Asp
 5665 5670 5675 5680
 Thr Ala Ala Ala Val Arg Pro Asp Gly Ala Phe Arg Glu Leu Gly Phe
 5685 5690 5695
 Asp Ser Leu Ala Ala Val Glu Leu Arg Asn Arg Leu Gln Thr Thr Thr
 5700 5705 5710
 Ala Leu Thr Leu Pro Ala Thr Thr Val Phe Asp His Pro Thr Pro Ala
 5715 5720 5725

Ala Leu Ala Asp His Leu Arg Thr Gln Leu Cys Gln Asp Ala Gln Ser
 5730 5735 5740
 Ser Ala Ala Ala Thr Ala Met Ala Ala Met Ala Glu Leu Ala Arg Leu
 5745 5750 5755 5760
 Glu Ser Ala Val Ser Asp Ser Val Ala Leu Asp Asp Asp Thr Arg Ser
 5765 5770 5775
 Gly Leu Ala Glu Arg Leu Arg Ser Leu Ala Arg Lys Met Ser Ser Gly
 5780 5785 5790
 Arg Val Val Asp His Asp Gly Gly Gly Ala Ala Asp Leu Asp Leu Gln

5795 5800 5805
 Ser Val Thr Asp Asp Glu Met Phe Glu Leu Ile Asp Lys Glu Val Ser
 5810 5815 5820
 Arg Asp

5825

<210> 79

<211> 15800

<212> DNA

<213> Artificial Sequence

<220><223> insert region of pCR2.1-Adu-milA1-Ac2 (FIG. 6)

<400> 79

tctagagatg acggaatcc cctggctgag ccggcatctg ccggccgagt gagcggcgcg	60
tcgacagcgt tgccgcatca ctccccggg cgccgggcgg tcagcagcac atagccgatg	120
tccggcgctc aggggagcgt cgtgcactgt gcaaccacct tccgcatctc cgcgacgcgc	180
atcccgaagg ccgccgcgta ctctcgagg tgctcgttca cgtggtcacg cagcgcggcg	240
aaagaggggc cgacgacctg atcgccgatg tcggtcagct catgcagtgc cagcccggcg	300
tcggcgatca ttccggcgta ctctcgatg tgcaccaggg ccgggaccgc gagcaactgg	360
gacgtgcact ccccgcgctt catgccggtc cgcccgaagg cgcgcagcgt cacgtccgtg	420
acggccagcc ggccgccggg gcggagcacc cgggcgatct cccggatcac ctgtgcgggg	480
ctgggcatgt gcaggagaca ctccagggcc cagccgcgt cgaaggacc gtcgggaaac	540
ggcagttcca tcgcgtcggc acgggtgaac acgaccggg ccgccacgtg cgactgcttc	600
gcgagagcgg tcgccagccc gacctgaacc tcgctcaccg tcacgccgac gacatcgacg	660
ggcgcgctca gggcgagccg caccgccggc ttccgggaac cgcagccgac gtccaggacc	720

cggcggcccg tgatgcctct cagcttgccg atgaggagat cggtagccg gtcggcgccc	780
ttgccgggtg aactgccgtc ccccggtgc ggccagatc caggtgggt gttccaccc	840
agcgacgat tcatgaggtc ggtcaaacgg tcgtagtagt cccccacttc cagggaagag	900
ggcggggtct gtcggggac ggccatcatg gtcgggaacc tccgaatcc ggccgggacg	960
gcccagctgt cgtggcgatc tactccagga aacgtagacc tttttctgcc acttgtccga	1020
gctatgcaga cacccgatc ccctaagaaa tgaacaccct tgggaacggc acagcccagg	1080
ggtggatagg ggtattcgcc gccgccgccc cgtcattagc tttgaagagt tgaagacgtt	1140
caagacattg atccccggc gtcagcggtt ttctcgctt cctttcattc ttcgacgtg	1200
cattgcagct ctcatcatgt ccgcacggcc gccgagcatt gcctagcggg gaggacacag	1260
ctcagttgcc caaagcccag aacgagttcg cagtggccgg tcatccgtgg atcctctccg	1320
ggcacaccgg aaccgcgtg cgggcccagg cacgccggt ccacgacct gtcgccgacc	1380
accctcggtt ccgtccggaa gacatcgccc acacgtggc gagcagcggc cggcgctca	1440
cccatcgccg ggcggtgatc gcggcgacc gggaaggaca tctccggggg ctcgacgcgg	1500
tggcccgggg tgaggacacc cccggtgtcg tacggggcac ggccggcccg ggcggcgacc	1560
gggtcgctt cgttttccc ggccagggca cccagtggcc cggtagtggc gccgatctgc	1620
tgacggtctc cccgccttc agccgggagg tcgacgcctg cggcgaggcg ttcgaaccgt	1680
atgtctctg gtcaccggag gccgtgtgc gggcgctcc gggcgcccg cccctggagg	1740
ggaccgatgt ggtgcagccg acgtgttctg ccgtcatggt ggggctggc gagctgtggc	1800
ggactcttgg ggtgagcccg acgtcgatcg tgggccactg catcggggag atcgggcag	1860
cccatctctg cggcgccctg tcgtgttccg acgcggcgcg cgtggtgatc gagagcagcc	1920
gggcccaggc gacgtcttcc ggttcgggtg cgtgatcgc ggtcgccgg tccgaggcgc	1980
agctgcttcc gttgtgcgg cggtagccgg gcaggtgac gatcgcccg gtcaacggcc	2040
cgatggccac ggtcgtctcc ggcgatcggc cggcccgca cgagctgtt gcggagttcg	2100
cccgtgccgg tgtccgggccc cgcgaggtgg cgatcgacat ccccgccac tcgccgttca	2160
tggcccccct cagggacggt ctgctcgact cgtgtcacc ggtcacccg ggtgcgtcgc	2220
ggctgccgtt ccactcctcg gtcacgggg ggccgctgga gaccaaggg ctcgacgcgg	2280
cttactggta ccggaacctc gccgacacgg tccgttcca aagcgtctc acggggctgc	2340
tgcggcaggg cacacgtgc ttcgtggagc tgagcccga cccgatgctg accatgtgtg	2400
tgcaggccac cggcgaggag gtggtcggcg gtgagcgcgt cgtgatcctg ccgacgtgc	2460
atcggggca ggccgccgtc gagtccgttc gcaccacgt ggccgagctg tacgtacggg	2520
gcgcactgga tgaccatcgg gcggcgttct cgtgcccgg cggccgcctg atcacctgc	2580

ctctcgagcc gcccgcgagc acgtccgtag agctcgccga cgcgccggac cggcgagg	2640
cctgccggcc ccccttggtg gaggcgcttg cccggctctc caccgaggag cgggaagcggc	2700
ggctgcgcga gctggtgggc gtcgaggcgg ccaaggctct cgaggacgtc gccggggcgg	2760
acgcgcccgg ccacggcatc gcggagcagg agcacttcgt cacttcgggc ttcgactccg	2820
cggcccggt cgcgtcgc aaccgcctga acgacgccac cgtttgtctg ctgcccttca	2880
ccctggcctt cgaccatccg acaccgcgcg ccgtcgccga ccatctgcac tcccggctct	2940
tcgatcacca gggcgggcgg cagccgggcg ccgacggccg gcccgacccc gcggcgggcg	3000
ccggtccggc caggcgccgac gacgagccga tcgccgtcat cggcatggcg ggccgcttcc	3060
ccgggggcgc ccgtaccccg gaggagctgt gggaactggt cgccgaaggc accgacgccc	3120
tctcgccctt cccggagggc cggggtggtg atccgctgcg gctctacgat ccggaccccg	3180
cccggcccgg cactactac cagcggaag cgggattcct ccacgacgc gacaagtctg	3240
acgccgagtt cticggcatc gcgccacgcg aggccaccgc aatggatccc cagcagcggc	3300
tgctcctgga gacctctgg gaggcgctcg aacgggcgcg gatcgaccg accgcgtgc	3360
gcggcagccg caccggggtg ttctcgccg tgccccgct ggactacagc ccccgaatgc	3420
accaggcgtc gccggagctg gagggccatc tgctgaccg caacatcggc gccgcgcct	3480
cggggcggtat ctctacgta ctcggtcttg aggggcccgc ggtgtccgtg gacacggcgt	3540
gctcgtcgtc cctggtcgcc ctgcatctgg cggcccaggc gctcgggcc ggggagtgt	3600
cgctggccct ggtcgggcgg gcgacggtcc tctcgacccc cggcatgttc atcgattct	3660
cgcggcagcg cgtcttggt cggacggcc gctgcaaggc gtacgaggcc gccgcggacg	3720
gcaccggctg gtccgagggt gtgggcatgc tgctcgtcga gcggctgtcc gacgcgcgac	3780
ggctcggaca ccaggtgctt gcggttggtac ggggctccgc cgtcaaccag gacggggcga	3840
gcaacggctt caccgcgcc agcgttccat cacagcaaca ggtcatccg gcggccctgg	3900
ccaatgccgg ggtgtcggt cggagggtcg acgcggtgga ggggcacggc accggcaccc	3960
ggttgggcga tccgatcgag gcgcaggcgc tgctggccgc ctacgggcag gggcgggcgg	4020
ccgaccggcc gctgtggtg gggtcgatca agtcgaacat cggacacacc cagtgggccg	4080
cgggtgtcat cggggtcatc aaaatggtgc tcgcgtcca gcacggtgtg ctgccgcga	4140
cgctgcacgt ggacaagcgc tcggattacg tggactggtc ggccggggcc gtacggtgt	4200
tgacggagcc ggtgccctgg ccggagcggg gccaccgcg ccggcgggg gtgtcgtcct	4260
tcggggtgag cggcaccaac gcccatgtca tctcgagca ggcaacgcca tcgtccacgg	4320

tggtccccgg ggggcataacc gccgaggccg ggcctcccct gccgtgggtg gtctcggcga	4380
agacgcccc ggcactgcgc gaccaggccc gccgcctgca cgaacacctc accgcccagc	4440
cacagctcca accggccgac gtccggccaca cctcgcgcac cggccgcgcc accttcgacc	4500
accgggccgt cctcatcggc tccgaccgcg aacaactcct ccacggcctg gacgcgctcg	4560
ccaccggccg gcccagacca gcggtccacc agacgtcgga ccgtcccgcc accgcccagc	4620
gccgtatcgt cttcgtcttc cccggacaag gcggtcaatg ggccggcatg ggcctacggc	4680
tgctgaacgc ctcacccgtc ttcaccgagc ggatggccgc ctgcgaacag gccctctccc	4740
cctacgtcga ctggtcactc accgacatcc tccaccggcc ggccgacgac gccgtatggc	4800
aacgcgccga catcgtccag cccgccctgt tctcgatcat ggtgtccctg gccgcgctct	4860
ggcgctcttg cggcatcgaa ccggacgccg tctcggcca ctccaaggc gagatcgccg	4920
cggccccagt ctccggcgcc ctgacgctcc acgacgggc caaggtcatc gccctgcga	4980
gccaggccct ccaagccgta cgcggcgccg ggggcatggc ctccgtacc ctgcccgcg	5040
accaggtcac cgaggatctg cgcacccact ggcccagccg gctatgggtg gccgccacca	5100
actccccac ggcaaccgtc atctcgggaa acaccgacgc gcttgacgaa gcgctcgacc	5160
actaccagc ccacgacgta cgggccaagc gcacccgggt cgactacgcc tcccactgcc	5220
cccatatcga cgcggtggcc gagcgactgc ccgacctgt gggcggcac gtcccgcgcg	5280
ccgccgacat ccccttctac tccacggttg acggccgatg ggccgagccg accgagctcg	5340
acgccgacta ctggtaccgc aacctccgca gcccggtacg gttcgccac gccgtccag	5400
ccctcacga gaccgaccac cgcacctttg tcgaagtcag cccacacccc acgctcacc	5460
ccgccatcac ggccaccacc gaaaccaccg accgcaccac caccgtcatc gcctcgtcc	5520
accgcgacca cgacgacacc caccacatcc tcaccaacct cggccaggcc cacatccag	5580
gccacacat cgactggcga caccactacc agactctgcg cccacacca cccatateg	5640
acctccccac ctaccccttc caacaccacc actactggct ccacgactcc accgaggaca	5700
aggcggtggg tacggacctc gcccgggccc gcttctggga ggccgtccac ggcgaggaca	5760
ccaacgccgt cgcgcgctc ctcgacgtcg agccgggcac ctactggac gcgtgctgc	5820
cggccctgtc cgctggcac ggtcggcgtc gcgaccaggc catcaccgac acctggtgtt	5880
accgcgacat ctggaagcgc gccgacctca ccgcgcgcg ccccgccg tccggccgat	5940
ggcttgctgc gatctccga gggcgggccc atcacctcca cgtcagtgcc gtccctggacg	6000
ctctggaacg ccagggtctg cccatcgca cctcgtcct cgacgacacc cacaccgaac	6060

tccccctgct ggagcggcat ctgcacagc cgatcgcgag cgatgggccc gccatcggcg	6120
gcgtgctctc gctgctcgcc ctgcacagc ggccacatcc gcgccaccg gaggtgcccc	6180
tcggcaccgc cctcacctc agcctgatcc aggcgctcat cgcacgcgag gacatggcgc	6240
cccggctgtg gctggccacc caccaggccg tcgccacctc gtccgcggat acgctcgatc	6300
acccctcca ggcatggtc tgggggctgg gacgcaccgc cgactcgaa caccctgatc	6360
tgtggggcgg cctcatcgac ctccggaca ctctaccga acgggtctc caccgctcg	6420
tcacggcgt gaccacctgt caccacgagg acgaactcg gctgcgcgcc accggccac	6480
gcaccggcg cctgatccg acgctgcca ccgccgcgc ggaggacacc ccgctgga	6540
cgccccgtgg caccgtctc atcaccggcg gcaccggggc cctgggtcc cgctcgccc	6600
accgcatcg cgaacgccac ccgactgcc acttgctgt ggtgagccgg cgaggccca	6660
aggcccccgg cgccaccgc ctccgcgacc agctcatga actcggcgc acggtgacc	6720
tcgcccatg tgacaccgc gacccggcg cgtcgcgga tctctcgcc gatgtccct	6780
cggaccgcc cctaccgcg gtctccaca ccgcggcgt cctggacgac agcaccctcg	6840
ccgtacagac ccggaccac ctccgcgcg ttctggggc caagtccat gccgcacacc	6900
atctgcacgc cctcgcccag caccacccc tcgacgcgtt cgtctcttc tcgtccgtcg	6960
cggcgccctt cggcgccgc ggccaggcca actacgcggc cgccaacgc tacctcgacg	7020
ccctcgcca gcaccgcgg gccccagggc tggccgccac ctccatgcc tggggcaact	7080
gggacggcga cgggctcgc agcacacag ccgcgcgac gtacctgcg aaccgcggt	7140
ttctccgat gccgccac ctggcgctgg ccgccctgga gcgagccatc gtctgcccc	7200
acgcccagct cgtctcgcc gacgtcgact ggaagaagc caagccggcg ccgcacacc	7260
gcgacatcc gggaagccg cgcggggccc cggccgccac cgacggcgc gacaggacgg	7320
ccgacgccac cgcgagctc cgtaccgcc tcgcggtca gagccggcc gaacggcacc	7380
agacgtcct cgacctatc agctctcata cagccgcgt cctcgggcac gccacgccc	7440
agacgatcc caccgaccg gccttcgcg acctgggtt cactcgtg acggccatcg	7500
agctcgcaa ccgctcgcg gcggccacc ggctccgct gccgaccacc gtcgcttcg	7560
accgtccgac gccggacaag ctgcggccg acctgctggc gcggtgcgc ccgacaggcc	7620
cggacggcat cggggtgac cccgacgca cggccacgag tggcagttcg cccggtcgcg	7680
cgcatggcg gccggacccc gccgagcca tcgccatgt cggttggcc tgccgtacc	7740
ccggcgggat cggtccccc gaggacctgt gggagtcat caccgcacac cgggacgccg	7800
tcggagactt cccgaccgac cggggtggg acctggcgag gctcttcgac cccgatccg	7860
accggccggg cactctgac agccgacagg gcgcttctt ccgcgacgc ggcgacttcg	7920

acccggagtt cticgggatc agcccacggg aggcgacggc gacggacccc cagcagcgac	7980
tgctcctgga ggcgtcctgg gaagccctcg aacgagccgg gatcaacccc cacgatctcc	8040
acggcagtcg gacggggcgtc ttcaccggca gcaacgcgca ggacttcagc gcgcggctgc	8100
ggcagacgcc gtcggagctg gcggagctgt gcgagggcta tgcgtgact ggcagcaaca	8160
acagcgtcgc ctcggggcgc gtctcgtacg cgtcggcct ggaaggcccg gcggtcagca	8220
tcgacaccgc ctgctcgtcc tcgctcgtgg cgtccatct ggctgccag tcgctccggg	8280
ccggcgaatg ctgctggcc ctggcgggcg gcgtcacggt catgatgacc ccgttcaact	8340
tcgtggagtt ctcccgagc cggggcctgg cggcggacgg ccggtgcaag gcgttctccg	8400
ccacagccga tggcaccggc tggggcgagg gcgtgggcat ggtggtggg gagcggctgt	8460
cggacgcgcg gcgcaacggc catcgtgtgc tggccctcgt ccgcggcagc gccgtcaacc	8520
aggacggtgc cagcaatggg ctgactgccc cgaacggccc ctgcgagcag cgggtcatcc	8580
gcgccgccct ggccgccgcc ggggtcgccg cggcagaagt ggacgcggtc gaggcgcacg	8640
gcacggggac gacgctcggc gatccgatcg aggccaggc cctgctcgcc acctacgggc	8700
aggggcggcc ggccgaccgg gcgctgtggc tcggtacggt caagtccaac atcgacacg	8760
cccagtcggc cgcgggtatc gccggggtca tcaagatggt gctggccctg cggcacggga	8820
tgtgccgcg tacgtgcat gtgtccgagc cgtgccgca tgtggactgg tcggcgggtg	8880
cggtagggct gctgaccgag gaccagccgt ggccggacac cgggcgcccc cggcgggcgg	8940
gggtgtcgtc cticggcgtg agcggcacca atgcccatgt gatcctggag caggcggagc	9000
cggggccgga cccggcaccg acggcctcgc cgcctccct gcccccctgg cccctctcgc	9060
ccaggtcggc ggaggccctg cgggcccagg ccgtaggtt gctggcgtac gtggccgagc	9120
acccgatgt cgacccgcc gacgtgggtt actccctgc gcgcggacgg gccgtgttcg	9180
agcacgggc cgtgctctc ggccaccggc acgacgactt ccggcgcgcc ctggacgccc	9240
tggcgtcggc cgcgcccgc ggccgggtcg tccaggcgc ggcggtggg cggcagggca	9300
aggtcgtctt tgtgtgctc gggcagggca ccagcgcgc cggcatgggc cgcgggctct	9360
accgtcgtc cagggcgtt gccggggcgc tggaggaggt gtgcgccat ctggaccgt	9420
atctggaaca cctctgatg gagtgatgt tcgccatga gaagagcgt acgtcggcgc	9480
tgctgcatct gaccgctac gcccaaccgg cctctctgc cctccagacc gcgctgcacc	9540
gcatggtcac cgaggagtgc gggctcacc ccgactatct ggccggccac tccctgggcg	9600
agctgaccgc cgcctatct gcgggcatcc tcagcctgcc cgacgccgc gcgctggtgg	9660
cggcccgcgc ccgcgccatg cgggaccttc cagcgaccgg agccatggtc gccgtcgagg	9720
ccaccgaggc ggagctgcgg ccccggctcg ccgagttggc ggaccgggtc ggcatcgccc	9780

ccgtcaacgc ccccgcgctcc ctgggtcatca ccggcgacca cgacgccgtg caccagatcg	9840
ccgacgactt ccgcgggcag ggcaggaagg tctctccct ccaggtcagc ggcgccttcc	9900
actctcccca tatggagccc ctgctcgacg agatcgggcg caccgcegaa accctcacct	9960
accaccggcc ccacactccc ctgctcaccg cgtcggccga cggcggcgac gacacgaccg	10020
agcccggggc cgacgacgac ccgggcacgg ccgcgttctg gcctctccag gcccgcgca	10080
ccgtccacta cgcgcgggccc gtggagcggc tgcgcgcccg cggcgtcacc acgttcctgg	10140
aactcggccc cgactccacc ctactacc tctccacca caatctcgcc gcgcacgac	10200
ccgtggccgt ctccctgctc catccggagc ggtgcgagac gcacagcgtc ctcggcgcac	10260
tgcgccgggt ccacgcccac agccgccccg tgcactggac acgccactac accgcacggc	10320
cgcggccgac gccacaccag atcgacgtgc ccacctatgc cttccggcac cggcgctact	10380
ggctccccgc cccggcggcg tgcggcgatg tgacggccgc ggggctcgac gcgcgggagc	10440
acccgctgat cggcgccgcc gtgtggctcg cggagggcga cggctgtctg ctgaccggca	10500
ggatctcgcc gcgtacgcac ccgtggctgg ccgaccatgt catcgccgc actgtgctcc	10560
ttccgggcac cgcgttcgtg gagctggcgc tgcgggccgg ggctacgtg ggctgcgacc	10620
gtgtggagga gctgacctg caccgccac tcccgtgcc cgcgcacggg gaggtggtgc	10680
tgcaggtggc ggtgggggcc gccgacgagt ccggccgccg tgagctgagc atccacgcc	10740
ggccggcgga cgacgttaca tggacacggc acgcatcgg cacgttgga tcggcccgcg	10800
gcgtcggcct cgacgatggc acggggcaca atggccacgc cccggcgggc gacgagccgt	10860
tcgggtcgtg ggccacggcc tggccgccgc ccggtgccga gcccttggac gtcaccgggg	10920
tctacaccg gtttgccgac gccgagttca cgtacgggga ggcatccag gggtggtcg	10980
cggcttggcg gcacggcgac gagacgtgg cggaggtccg cctccccgac cagccggccg	11040
gtgacgccct ccgttcggg ctgcaccccg cgtgctcga cgcggcactg cagaccatgt	11100
ggctcgtgga gcccgcggc acacggccga gcggtggcct gggcggcccc gatcggggcc	11160
tgccttcgc ctggcagggg gtctcgtgc gtacggcggg ccgctcgcc ctgcgggtac	11220
ggctcgcagc gccggcgccg gacaccgtgg ccgtcgccgt ggccgacgc gccggccggc	11280
cggtcgcgtc ggtggagtcg ctgacgtgc gcccggtgcc gcggggcgcc ttgcgcggca	11340
ccgagacggc ggtgcgcacc tcgttgtacg gcctggactg gacggatgtg ccgtgccga	11400
cgccgagac ggccctgccc cgggtgtgcgc tgatcggagc ggacacgctc gacctggtcc	11460
ccgcgtcga ggccgcggcg cccgaccgca tcaccgacgg cgtggagcgc tacgccgacc	11520

tggaggagct ggtgcgctcc gtggcggcgg gcgccccgc cccggacctc gtcacgccc	11580
gctgccacgc agccctgaa gccacggcg cgagcgaaca gccacagccc gagacggtgc	11640
gcacaaggac gggtcagggtg ctggagctgc ttacagcgtg gctcggcgcg gacgggctcg	11700
ccgacgcaca cctgggtctg ttacacctcag gcgcgggtcgc caccggcccg ggcgagccgg	11760
tgcgggacct ggccggggcg gcggtctggg gtctggtgcg ctccggccag tcggagcatc	11820
cggagtgtt caccgtggtg gacatggacg gcgcccagga gtcccgcgcg gcgctgctcg	11880
gcgcgctcgg cctcggcgag ccgcaactgg cgggtgcgcgg cggccgggcg ctggcgccgc	11940
gcctggtgcg cccgggtgac gccgacgacg acagcggcct ggccctgccg caggggcccgg	12000
aaggctggcg gttggagtgt cccggcacgg gcagcctgga cgggttgacc acgaccgagt	12060
ccccggccgc ggccggtgccg ctccggcccgg gcgaggtacg ggtcgcggtg cgggcccgg	12120
ggctgaactt ccgcgatgtg ctgatcgcg gcggcgtggt gcccgggcg acggcgctgg	12180
gcagtgaggg ggccggggatc gtcctcgagg tcggggcgga ggtccgcgat ctcgccccg	12240
gggaccgggt ggtgggtatc ttccccagg cgttcgccc ggtggccgtg gccgagcggg	12300
cgacctggc ggccgtcccc gacggctggt cgttcgccc gcccgcgctg gtccccatcg	12360
tgttcccac cgctaccac ggccgtggtc atctggcgcg cctgcggccg gggaatcgg	12420
tgctgatcca tgcccgggcc ggccgggtgg gcatggcgc cgtgcaactg gcgcgcatc	12480
tgggggcca ggtgtacgc accggccggc ccggcaagtg gcacatcctg cgttcccaag	12540
gcatcgacga cgaccatctg gcctcgtgc gcacgtgga gttcgagcag cgcttcgccg	12600
cgaccgcgg cggcggggg atcgatgtc tcctggactg tctggccat gatttcgtcg	12660
acgcctcgt gcgcctggtg gcgcgtgacg gcggccggtt cctggagatg ggcaagagcg	12720
acatccgtga ccgcggcag gtggcgctgg accatccgg cgtgctctac cggcgcttcg	12780
acctgctgga ggccgggccc gagcgggtcg ggcagatcct gcgcaccgta ctggacctgt	12840
tcgagcgcg tgcctggcg cacctgccga cgacctgtg ggacatccgg caggcggagc	12900
acgccttcg ccatctgcag caggccgtc acatcggaag gaactgtct accgtcccgg	12960
ccggtggaa cgccgagggc accgtactga tcaccggcgg tatgggcacc ctggcgccg	13020
ccctcggcg tcatctggcg ggtaccggc gcgcccga tctgctgtg gccggccgac	13080
gcggccccga cggccgggc gccgaggagc tgcgagagga gctgaccgag ctgggcgcg	13140
gggtcaccat cgccgatgc gatctcgcg accgggcggc ggtcgcccgg ctctggggg	13200
cgatccggc cgagcggccg ctgaccgtg tcattcacgc ggccgggtgtc gtcgacgatg	13260
ccacctcgg gtccctcacc ccccgccacc tggacggcg cctggccgc aaggccgacg	13320
ccgcctggca tctgcacacc ctacccgcc acgccgacgt ggccgcgttc gtctcttct	13380

cctcgggtcgc gggctctgctc ggctcgcccc ggccagggcaa ctacgccgcg gccaacgcct	13440
tcttggagcg gctcgcccc caccggcgcg gctctggcct tccggcggtg tcgctggcgt	13500
gggggctgtg ggagcagacc agcgcatga ccgggcacct ggaccaggcc gaccgcgcc	13560
ggctggcccc gctcgccatc agcccgtca cgaccgggca ggcgctcggc cttttcgacg	13620
ccgccctcgg ccaccaccgc cccgtgctcg tccccgccg cctcgacgtg cccgatccgc	13680
accccggtc gtcgaccgtg ccgccctgt accggggcct ggtcggtacc aggaccggc	13740
ggacaccccc cgcgccgcc gccaccgggc cgttccccct gcataccgc ctcggcggtc	13800
acggccggc cgagcagcac gagatgctgc tctcgctggt ccgctccac gccgccctcg	13860
tgctgggccc cgacgatccg gacacggtcc atccggcgcg gcaattccgc ggcctgggt	13920
tcgactccct gaccgcggtc gagctccga accggctcaa cgccgccacc ggcctccggc	13980
tctccaccac cctcgtcttc gaccaccca cgcccgacga actcgcccg cactccggg	14040
agcaggtgct gggcgacggc gaagcggcg gggtagcccc cgtgttgcg gagctcgaca	14100
aactcgaatc cgccctctcc gccctcgaca aaaccgacag cgccagcgaa agagtcaccc	14160
tgcggtgaa gtcactcatg ttgaggtgga acgcaccca gcatccgaca gccgaaagcg	14220
ctgatgacga cgagaagttc acatcgga cagaggctga gatattcaaa ttcattgaca	14280
acgacctcgg cctgtctga accggacgcc tgccactccg cccgtatccg ctgggccctg	14340
ctaggacgtg aatgcaattg gcgaatgaag cgaagctcct ggaatactc aagcggtca	14400
ctcgggacct ggaccgcact cgccgtcgcc tgtacgaggt ggtcgagcgt gacgaggagc	14460
cgatcgcatg tgtgggcatg gcgtgtctt acccaggcgg ggcgacgtca cccacgcgac	14520
tgtggcatct cgicaagtc cagacggacg ctatcgggga gttcccgacc gaccgtggat	14580
ggaacctgga gcagctctac gaccggacc ccgaccgtc aggaaccagt tacacgcga	14640
gcggagggtt tctctatgac gcggcgact tcgacgccgc gttcttcgag ttgtaccgc	14700
gtgaggcgct ggcaatggac ccgacgacg gctgtgtgt cgaaaccact tgggaaactg	14760
tcgaacaggc cggaatcgac ccgaggtcca tgcgcggaag ccggaccggg gttttcgtg	14820
ggatcaatcc ggaggactac accaccgat acacacatca gccctcaaac gcagtcgagg	14880
gctacctgct cactggcagc gcggcaagca ttgcgtcagg ccgtatctcc tacaacttcg	14940
ggctcgaagg ccttcgcatc actatcgaca ccgcgtgttc ctcctcgtc gtcgccctgc	15000
atctggcctg ccaagcgctc cggtcgggtg aatgcacat ggcgctcgca ggccggcgct	15060
ccgtcatggc cactcccttc gtcttcaccg agttctctcg ccagcggggc ctggccgcag	15120
acggccggtg caaggcgttt tcggcgccgg ccgacgggac cggtcgtcc gaggtgtgg	15180
ggatgtctgt ggtggagcgg ctctccgacg cccgccgcaa cggtcaccgt gtcctggccc	15240

tcgtccgagg cagcgccgtc aaccaggacg gcgcaagcaa cggcctgacc gcaccaacg 15300

gtcgttcaca agtcaaggct atccgccagg ctttggccaa cgcacacctc tcccctgccg 15360

atgtcgatgc ggtggaggcc cagggcacgg ggaccacccct gggcgaccgg atcgaggctc 15420

aagccctcgt cgaagcctac ggtcaggacc gcccacacgg cggccccctc tggctcggaa 15480

ccctcaagtc caacatcggg cactccatgg ccgtgcggg tgtgggcggg gtcataaga 15540

tggatgatgc gctgcggaat ggtctgctgc cgcggacgtt gcatgtgat gagccgtcgc 15600

cgcgtgtgga ctggtccgcg ggtgcgggtc agctgctgac ggagacgggt ccctggcccc 15660

gcggggaggg gcggctacgg cgggcaggag tgtcatcatt cggcgtcagc ggcaccaacg 15720

cccacgtcat cctcgaggaa gcacccgccc acaacatccc gtcagacaca ccgcccagc 15780

acgtcccggg agaatacggc 15800

<210> 80

<211> 22878

<212> DNA

<213> Artificial Sequence

<220><223> nucleotide sequence of pD7K9milA3

<400> 80

ctatgcaggc ccccgaggat cggctcttgc ttgctcgtcg gtgatgtact tcaccagctc 60

cgcgaagtcg ctcttcttga tggagcgcat ggggacgtgc ttggcaatca cgcgcacccc 120

ccggccgttt tagcggctaa aaaagtcatg gctctgcct cgggcggacc acgcccata 180

tgaccttgc aagctcgtcc tgcttctctt cgatcttcgc cagcagggcg aggatcgtgg 240

catcacgaa cgcgcctgt cgcgggtcgt cggtagacca gagtttcagc aggcgcacca 300

ggcggccag gtcgcatatg atgcgggcca gctcgcggac gtgctcatag tccacgacgc 360

ccgtgatatt gtagccctgg ccgacggcca gcaggtaggc cgacaggctc atgcgggcca 420

gcctcgcaga gcaggattcc cgttagcac cgccaggtgc gaataaggga cagtgaagaa 480

ggaacacccg ctgcgggtg ggcctacttc acctatcctg cccggctgac gccgttggat 540

acaccaagga aagtctacac gaaccctttg gcaaaatcct gtatatcgtg cgaaaaagga 600

tggatatacc gaaaaaatcg ctataatgac cccgaagcag gggttatgcag cggaaaagat 660

ccgtcgacct gcatactagc tgctcaaggt cgcgcgcaac cggcgcatca agcccgccgt 720

agctgggtct ccccgcccc tggccgacgc cgcgccccgc gccgcccgc acgacgaacc 780

gtggacctgg accatcagcg cctaggtttc tgcactttct gcactgcca acgcgggcgc 840

gcgtgtgcgc cggccaccgc ccgggtgtg acggaaatcg ccagcagacg tagtccttgg 900

tgagactccg ccccaacgcc ctgatctgcg aagaggaccg ttcgtggaac tgatgagacc	960
gctctacgcg agcgcgcgcg gctgtaacgg atcacgaggt cactccgtcc ataggccgaa	1020
gaggctcacg acgaggaggt cctgacgcac aggaagcccc gccagggccc gggcggggct	1080
tcttcgacc tatcaaatcc caccagcacc gggcgtatcg gacgtctatc ggcgcgtag	1140
gagcgagcag gggcaatgaa cggccacacc aatgccagga tcaacaggac ccagaaatgg	1200
cacgagcccc ggattggcct cccgggctcg ccgccccga accggcgtag catgccgagt	1260
gaggacgttt gctctgtgac gagcgtacga gagaacagcg cgtcgacacc accgcgcccc	1320
gccggaccga agtgccactg cggcaccgcc atcgtccaca tccccgggaa gcgccccaaa	1380
atctactgct cggacgcctg ccgaaagcgc gcaaagcgtc tcattgcgga ccagtacgc	1440
agcgtagcaa gcacgactac agcgggtaat ccggacgcgg gctttagtag taaggaaaa	1500
tctccgtcca ctgcaggta gagcgtgta actggaacgg gttctacccc aaacgagggt	1560
gccgcgacag ccggcgccga ggaccgcgc gacgagcggg tcgcccggcg ggaccggcac	1620
cagacgggtg cgtgaacga ggcggtcaag ggctgcggga cccggctcac tgccgggacg	1680
gccgagctga tgtggaagcc gggcgaggcg acgtggggca acacgtgccg ctgcaacaac	1740
gtccacacct gcccggtgtg catgagccgc atcctggccg tacggggctc gaacgtgcag	1800
ctcgccgccg acggtctcgc ggacgccggg tacgggctgc acctggcac gaacacgttg	1860
cgccacttcg agcgcattgc gttcggcacc gtccgcaagg gaatgcggca cggcctggtc	1920
gcggtcttgc acgacggctg gaagggcgcg tacggctcgt ccggccgccg ctggcgacg	1980
atgcgcgacg acttcgggat catcggctac gagcgggcct tcgaggacac cttcggctgg	2040
ggctccggct ggcacctgca ctggcacacg ctctgggtca cccgcgaggt gctcgccccg	2100
gacgcccagg ccggttccg cgacgcactc gccggagcct gggccgccgg ggtcgaatcc	2160
gcgggcccgt acaccgtcag cgagacgtgc gaccggcccc gttgtctctg tgagggcaag	2220
ggccacggca ccgacgtgcg cccactcaac ggcgcgagcg ccgctgacgg cgatgcgggc	2280
aagcaggccc gctacctcta caaggacggc gacaagacca agggcggcgt cgccaagatc	2340
ggcctggagc tggccggaca gaacttcaag gccggtcgcg gcgacgaccg catgggaccg	2400
ctcgacctgg gcgacgcggc ggccgccgag ctgcagcggc ttcgtcggcc cggcccgttc	2460
gtcgagaagt accgcgagcg cgagttcggc gtcttcagg tccgcaagca ctaccggtcg	2520
cagaacctga accggctgat caaggaactc ggcatccagc aggacgtccg caccgaggag	2580
gagatcaccg acgacaccga gggcctggtc gccatcgcgg tcatcccggc ctacatctgg	2640

taccgataca tcgcccgcgt cgccggacgt cgcctcgacc tcatcaaggt cgccgagacg	2700
tacggcctac cgggcgtccg gcggctcgtc gagtccctggg gcctggtgtg gggcaaggac	2760
gtcctggacc cgcccccccc ggaagcgccg gccgcgccgg gcgaccttga cgtgaccag	2820
atgaggttcg aggtcatgtc cgaggaggaa gctgctttcc gcgaggcgcg ccggaaggcc	2880
aacgaggccc gcacggagga gctggccgct tcgctcgacc gagtgcgtca gccgaagaag	2940
gaggcgatcc gcccgacgat cagccttcgg aagcgctga agccgaagcc ggtcacctc	3000
gacgtcaaga ctccccgcc cggcgcccg tcgcccgtgt gccgccgtg caagggaag	3060
ctcgccccgg tactccagcc ctggggccgg ctccccggag actgcctccg cgtcgacaca	3120
gcggtcgctg agggccgctg agggaccgcc atgagaccgc caccgccgac atccgggtgg	3180
cggctctcatg gcggtcccg cgcgaaagcc gagggctctg ctgccgtgag gtggccggcg	3240
gacctgggtt actgctgaa tcccgcccc agtcggagcg tccgccggcg cttttcacct	3300
ggctccgggt gctcgtcgat gagcaccagg ccgtcgacga cgacgcgttc ggccggcagg	3360
agccgcgggg ccgacgtcga caacgacgtc ctggtggtgg tcgacggacc ggtgccgagt	3420
gcgcggcgag ggggtgctcat gccgatgagt ctgaacttct ccggcgaagg gttctaccaa	3480
gtcctgggtt gtgttcgct tggtagcgct cctagtggc gaggtggtc ctttacctg	3540
atgccctcgg gcatcagggt aaaggaccag cctctcaca tgagtaggct gaaacagcag	3600
ttcgttcgc ctcttttcc aaagttgaga agccgcagaa tccctttctg gagggaaacc	3660
agcgtacgta cattgcactt tccctccagg cgagggtttt gcagatgctt ggccggtgtc	3720
actgtctcta gctggcatgg ctttggaat agtccgctc agtccgagg gcttctctgg	3780
ggtggctccc ctgaggaacc tcgacatcga tacgtcgcgg tgagttcagg cttttcata	3840
tctcattgcc cccggacgag cgtctgctcc gccattgcc gtccgccgtg ccaatcggat	3900
cagccgtcca aatgcgggat tttcgttagt cggaggccaa acggcattga gcgtcagcat	3960
atcatcagcg agctgaagaa agacaatccc cgatccgctc cacgtgttg cccagcaatc	4020
agcgcgacct tgcccccca acgtcatctc gttctccgt catgagctca gccaatcgac	4080
tggcgagcgg catcgattc ttgcattccc gcctctggcg gatgcaggaa gatcaacgga	4140
tctcgccca gttgaccag ggctgtgcc acaatgtgc gggagcggat caaccgagca	4200
aaggcatgac cgactggacc ttctttctga aggtcttct ccttgagcca cctgtccgc	4260
aaggcaaagc gtcacagca gtggcattc tcgagataat cgacgcgtac caacttgcca	4320
tcctgaagaa tgggtcagtg tctcggcacc ccataggga ccttgccat caactcgga	4380

agatgcagcg tcgtgttggc atcgtgtccc acgccgagga gaagtacctg cccatcgagt	4440
tcattggacac gggcgaccgg gcttcgaggc gactgagggt gcaggggcaa tggatcagag	4500
atgatctgct ctgcctgtgg ccccgctgcc gcaaaggcaa atggatgggc gctgcgcttt	4560
acatttggca ggcgccagaa tgtgtcagag acaactccaa ggtccggtgt aacgggagac	4620
gtggcaggat cgaacggctc gtcgtccaga cctgaccacg agggcatgac gagcgccct	4680
cccggacceca ggcgagcagc cagggcctcg atcagtccaa gtggcccatc ttcgaggggc	4740
cggacgctac ggaaggagct gtggaccagc agcacaccgc cgggggtaac cccaaggttg	4800
agaagctgac cgatgagctc ggcttttcgc cattcgtatt gcacgacatt gcactccacc	4860
gctgatgaca tcagtcgac atagcacgat caacggcact gttgcaaata gtcgggtggtg	4920
ataaacttat catccccctt tgctgatgga gctgcacatg aaccaaagg atctaggtga	4980
agatcctttt tgataatctc atgacaaaaa tcccttaacg tgagttttcg ttcactgag	5040
cgtcagaccc cgtagaaaaa atcaaaggat cttcttgaga tcctttttt ctgcgcgtaa	5100
tctgctgctt gcaaacaaaa aaaccaccgc taccagcggg ggtttgtttg ccggatcaag	5160
agctaccaac tctttttccg aaggtaactg gcttcagcag agcgagata ccaaatactg	5220
ttcttctagt gtagccgtag ttaggccacc acttcaagaa ctctgtagca ccgcctacat	5280
acctcgtct gctaactctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta	5340
ccgggttggc ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg	5400
gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc	5460
gtgagctatg agaaagcgcc acgcttcccg aaggagaaaa ggccgacagg tatccggtaa	5520
gcccgcagggt cggaacagga gagcgcacga gggagcttcc aggggggaaac gcctggtatc	5580
tttatagtcc tgcgggttt cgcacactct gacttgagcg tcgatttttg tgatgctcgt	5640
caggggggag gagcctatgg aaaaacgcca gcaacgggc ctttttacgg ttcctggcct	5700
tttgctggcc ttttctcac atgttcttc ctgcgttacc cctgattct gtggataacc	5760
gtattaccgc ctttgagtga gctgataccg ctgccgcag ccgaacgacc gagcgagcg	5820
agtcagttag cgaggaagcg gaagagcgcc caatacgcaa accgcctctc cccgcgcgtt	5880
ggccgattca ttaatgcagc tggcacgaca ggtttccga ctggaaagcg ggcagtgagc	5940
gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggcttta cactttatgc	6000
ttccggctcg taigtgtgt ggaattgtga gcggataaca attcacaca ggaaacagct	6060
atgacatgat tacgaattca tacggctacg aggatggggc cgatgaagcg gctcgcgcc	6120
tgtgcgccct tgcctcgggt gaaatcgccc gatccggagg agaccagcat ggtccgcacc	6180
aggctccgggt gcgtggcgag gtaggcggtt ttcggcccga ggccgagctt gagcaggtct	6240

ccgtgatctg ctgcggagcg aaggaactcc aggggctgcc gcatcaacgg cggcacatgg	6300
ccgacgaccg gccaggcgcc gggcgccctcg ggaatgctgc tcgtcgactg ggacatcacg	6360
agtgtctctt tgcggggtga aggggggttg ctgggagggg aacgacagt acgagtgaag	6420
ggggaggtgt gggggttggc gtcggcccgg gggtagagct ggacatggga gtgggagga	6480
gtgaagttag ctcgagtggt tttctgggct tcattgagat tcgaatccga cttccctgtc	6540
gatgagagcg aacatctcct cgtccgatgt ctctgcgagg tcggggcgcg tgtcgtcgcc	6600
gttcaacttc tggcgagggg aatgcagtcg ggaggccagc cgcgtgcgcg gccgtcgtc	6660
cagcggggca gcggaggtat tggtaggga gacactacc gcctccagcc gctcagctc	6720
cgagagcagc gagggcagcc ccgggggtgt cgttgcggcg tccggctcgg ccgcggcgt	6780
gagtccttgc ctgacgagtt gtgtgtggag gtggtgggtg aggggtgggtg ggttggggtg	6840
gtcgaaggcg agggtaggtg ggaggcggag tccggtggtg tgggagagcc ggttgcgtag	6900
ttcgacggcg gtgagggagt cgaagccgag gtcgcggaac gcgcggtcgg ggggatggt	6960
gtcgggggtg gtgtggccca ggacggtggc gatgtgggag cggaccaggg cgaggagggt	7020
ggtgtgctgt tgttcgtgtg tctggccggc cagccggccg tgcagctggg gccgttgtc	7080
cgcaccaccg gtagtggtgg tgcgggtggt gcggcgccgg gtggcgggca ggaggtcctg	7140
cagcaggggc ggcaggggcg gggcgggacg caggtcggcg ggcagcagga ccggccggtc	7200
cagagccagg gccgcacga agagggccag tgcgtccggg gtcgacatgg gatgcaggcc	7260
tgagcggttg atgcgggccc ggtcgtgctg gtcgagatgc ccggtcatgc cgtgcctg	7320
ctccacagg cccacgcga gcgacaaggc cggcaggccc tgcgccctgc ggtgttcggc	7380
gaggcgctcg aggaagcgt tggcgccgc atagttgcc tgcgggctg gccgaggac	7440
accggcgccg gaggagaaca ggacgaaggc gcgcggccgg gcgtggcggg tgagctcgtg	7500
cagatggaag gcggcgtcgg ccttgggtgt cagcgcggtg tccaggtgtc ggggggtgag	7560
cgaggcgacg gtggcgtcgt ccaggacgcc cgcagcgtgg accacaccgg tcaaagggtg	7620
gccggtggga atccccgca tgagatccgc cagcgccct cggtcaccgg tgtcacaagc	7680
gcgcaccgtc acctcggcgc cgagctcggc cagttccgc cgcagctcgg tggcgccggg	7740
tgcgtcgggt ccgcgtcgcg atgccagcag gaggtgtcgt acaccgtgac ggctgaccag	7800
atggcgcgcg agcatcgcg caagggtgcc ggtgcccggt gtgactacta cggtgccgtc	7860
ggggtcgaac gccgtgggcg tgggtgccgg gtcggctgcc tcggcgccgg gaacgggaac	7920
acgccggggg acggggacgg acgccggaac gcgggtcagc cggggcacca ggtgccggtt	7980
cgcacgcacc gcgactcgcg gctcaccaca ggagagagcc gccggaacga cggtcggtg	8040
cacttccccg ccgtcaggtt ccgggtcggg gtcgaggtcc agcaggacga accggccggg	8100

atgttccgac tgcgcggagc gcacatgcc gcacaccgcc gcgcccgcc ggtccgccac	8160
gtcctcccc ggagtgggtg ccaccgcacc acgggtcagt acgaccatcc ggctcgaggc	8220
cggcctctcg tccgccagcc actcctgcag cagaccagg atgcgcggg tgagggtcct	8280
tacgtcctcc ggcacggcgg tggcgcgcgc atcgcggcgg ccgcagtcga ccacgaccgc	8340
atcgggcgcc ggggcccccg cctccagcac tgcggcgccc aggtcggcga ggttcgcgta	8400
cgcgtctgcc gaggggaggc ggtcgttggg cccacacctg gccaggcgtt gcggaacgga	8460
cgacgaggtc ggcgccaggc cggtcacca ctgaccgcg aagagcgagt cccgccgcgc	8520
cgtaccggag gccgccagca actcctgcgt cgacaccggg cggaccgcga gcgatgcgat	8580
cgaggccacc ggcgccccct cgccgtcgac caccgcgacc gacatgcgt gcgtgtccc	8640
cgcctggccg tccgggtga ggcgcaccgc cagggtcgac gtcggggcc cgtgcagcct	8700
caccccgctg aaggagaacg gcagggcgcc gtgcgggagt ttgtcgccg ggcggaaggc	8760
cgcggcgtgc agtgcggcgt ccagcagggc ggggtgtacg gagaactggg aggcgtcgcc	8820
gtccgccgtg tcgggaaggc gcacctcggc gaagatgtcc tcgccacgtc gccaggccgc	8880
ccgcaggcct cggaacaccg ggccgtactc gaatccccgc gccgcgtacc cgtcgtacag	8940
gccctccaga tcgtctgcgg tgaacgcgac agcgccctgc ggtggccaga ccgccccgc	9000
cgctccgtcg ccttgtgggt cggtagccg tgcgatggct tcatcctggt cggccggtgc	9060
ggggccgacc cggcccggtg catggcgcac ccacgtggcg tcctcgtacg taccctcgcg	9120
tcgggaggac accgtcaggc ttccgcacc gtctcatcc ggtgccgcca cccatgcctg	9180
cacctgtact gtcctctctt ccgggatgaa cagcggtgat tggagggtga gtctctcgat	9240
gcgccagcag tccacacgat ggcccgcctg aaggcgcatc tcgagaaggc cagttccagg	9300
gaggacaacc gccccgcga tgggtgtgat gcgcagccag ggggtgggtg acaggcgag	9360
ccgaccggtg agcagatagc cctccccctc ggccaccgcg accgccgcgc ccagcaacgg	9420
gtgatgtgcc gggccaggc cgacggccgc cgcacccct gtctgggtgg ggggtgttag	9480
ccagtagtgg tgggtgttga aggggtaggt gggtaggtcg atgtgttggg ggggtgggt	9540
gtggtgttgg tgggtgtgt gccagttagt ggggtggccg gtggtgtggg cgtgggcgag	9600
ggcgtggagg atttgggtgg gtctgttgt gtgcgcgg agtgtgtga tgacgttgt	9660
gggggtgtgg gtggtgtcgg tggtttcttg gatggcggg gtgagggtgg ggtggggct	9720
gatttcgatg aaggtgttgt ggccttggtg ggtgaggttg gtgatggcgt cggcgagggtg	9780
gacgggttgg cggaggtttc ggtaccagta gtggcggtt aggggtgtgg tgtcgtcca	9840

ttggttgtcg gtggtggagt agaagggggt ggtggcggcc cgtggggtga tgccccgcaa	9900
cagccgtgac agttcctgtt ccacggcctg gatgtggggg cagtgggagg cgtagtcgac	9960
cgggatccgt ttgcgctga cgccttggtc ccggtagtgt tccagcgctt gtgtgagagc	10020
ttccgcgttg ccggagatca ccgtggcgggt aggggcgttg acggcggcca cccacagccg	10080
gtccggccac atggtctgca gcagcccgtt gacctggtcc gccggcagtg gtacggaggc	10140
catgcgccc gaaccgcga gggctcgcag ggcccggctg cgcagcgca cggttttcgc	10200
ggcgtcctcc aggcctcagc ctccgtgat gtgggcggcg gtgatttcgc cttgggagtg	10260
gccgatgacg gcgtcggggt taaggccgtg ggattgccag agggcggcga gggaggtcat	10320
gagggcgaag aggacgggct ggatgacgtc ggcgcgctgc cacagggggt cgggtgtgggg	10380
ccggtgcagg atgtcggta gggctcagtc ggtgtggggg gcgagggtt gtgcgaggc	10440
ctggatcgat tgggcgaaga cgggtgaggt ctaagcagc cgcagtccta tcccggccca	10500
ctgaccgccc tgcccgggga acacgaacac caccttccc ttggtcaccg ccttcctcgc	10560
gatccccgtg acgacggcgg tgtgggcctc gtcggccgcc agagcatcca ggccaccag	10620
gaactccgt cggtcggagc cgatgaccac agcacgatgc tcgaaagagg cgcggcccgt	10680
cgcgagggtg tggccgacgt cggcggggtt cgcctcgggg gcgttggta ggtacgtgtg	10740
cagttggcgc gcctgggcac gtagcgcttg cgggctcttc gcggagatcg cccagggcag	10800
tgggcctggc gcgggtgttg gctctgccgg ttctgtgggtc tccgcggcag gcggtgcttg	10860
ttccaggatg acgtgggcgt tggtagcgtt gacaccgaaa gaggagacgc ccgcacggcg	10920
catccgccc tgctcgggcc aggccaccga ctcggtcagc aaccgcaccg tcccggccga	10980
ccagtccaca tgaggtgagg gctcctcggc gtgcagggtc tgcgggagca gcccgtggcg	11040
catggccatc accatcttga tgacggccgc caccgccgcg gcggcctggg tgtgtccgac	11100
gttggacttg atcgatccca accacaacgg ccgatcggcc gaccgtccct gtccgtacgt	11160
ccccagcagc gcctgggcct cgatgggggt accgagcgtc gtccccgtgc cgtgtccctc	11220
caccgcgtcg atgtccatcg cggccacacc cgcgtttgcc cagccttgc ggatcaccg	11280
ctcttcgac gggccgttcg gcgcgctcag ccattgtct gcccgtcct ggttgaccgc	11340
actccccgt accaccgcca ggaccgggtg gccgttgcgc cgggcgtccg aaagccgttc	11400
caccaggagc atccccgcc cctccgcca gcctgttcg tccgcgccg cccgaagga	11460
cttgaccgc ccgtccgggg ccagtccccg ctggcgcgag aactcgaaa aagtgtggg	11520
ggtggacatg acggtgacct cgccggccag cgccatggag cactccccctg accgcagcgc	11580
ctggcaggcc agatgcaggc cgaccaggga cgacgaacac gccgtgtcca cggtgaccgc	11640
cggccctcc aggccgaagg ttagggcgac acgtcccgc gcgacgtgc ccgcgtttcc	11700

attgccgata tagccctcgt agccctctgg gatctcgcgc aggcgggtgc cgtagtcgtg	11760
gtacatcacc ccagcgaaca ccccggtcga actccccgc agcgacgccg ggtcgatccc	11820
cgcacgttcc agcgccctccc aggagatctc cagcagcagc cgctgctgtg gatccatggc	11880
taccgcctca cgcggggccga tcccgaagaa tgccgcgtcg aactcacccg cgtcgtaaag	11940
gaaccgcct tgcgggtgtg aggaggtgcc ggttcggtcc gggtcgggt cgtagagccg	12000
ttccacctcc cagcctcgtt ctgcggggaa gtcaccacc gcgtccatac cggaggcgac	12060
cagatccac aggcctcgg cgcaccgcac acctcccgga aaccggcatg ccatgccgac	12120
gatggccacc ggctcggcg ctgccccgcc gacggccgca acgggtgggg cggccgtggc	12180
cgccgacgta ccgccggcg cggccaactc gtcacacagg aaacgggcga gcacgtccgg	12240
acggcgccgg tcgaagatca agccggccgg caggcgagc ccggtggcct gctgcaaccg	12300
gttccgcagg tccaccgagg cgaggagtc gaacccagg tcgcggaacg ccgagtcgc	12360
cgggaccgat tcggccgtgc cgtgccccag caccgcagcc gcccgaggc gcaccagccg	12420
cagcacctcg ctagccgtt ctccggccc gaggccgtc aaccgccgc gcaggaggt	12480
cggctcaccg gccgaggcgt cgcgggtccg ggacggccgg gccggagcga caccgtcag	12540
gtggcgagg aatggcgga cggatccgga tgtctgaagc tccctcgtgc tgagtgcggc	12600
gggcaccacc agggcgtgtc cacagctcag ggccttgtcg aagagctcca ggcctcgt	12660
gatggccagc ggctcaggc ccgagcgccg gagccgcgcg tggtcgtgg cctccaactg	12720
cccgtcatg cctcgtct gtcctcaaac tcccaggcc aaggacaccg ccggcagccg	12780
agcgccctc cgtgactcg ccagcgctc cagggtgctg ttggcgccgg ctagttgcc	12840
ctgcctgcc gaacccatca ccccgccag cgaggagaag agcacaacg cggacagcgg	12900
caggtcgcgt gtgagctggt ggagattgag cgccgaatc acctggccc gcaggacatg	12960
ggtcaggccc tccggtgtga gggatgtgat cccccgtcg tccagtacgc ctgcccatg	13020
caccacagcg gtcagcggt ggtcgacgga gatctccgac aatacgttcg caagttgcct	13080
ttggtctcgc gcatcgacg cggatgacgt cacctgcgc ccgagctccg caagttcgac	13140
ccccagtcgc ctggccccg gggcgtcgag gccacgtcgg ctgacagca tcagatgcgc	13200
tacgcatggt gtggccacca agtgccgggc aagcatggcg ccgaggctgc cggtgccgc	13260
ggtgatgagt acggtgccgt cctgatccca gggatgctcg tccatggcct cggtgccgt	13320
gccggccagc gggctcagc ggggtgtgcg cagggtcca tcggtacgg ccacttcgg	13380
ttccccgag gccaggcca tcgacagcat tcggtcgtgg tctttcgtc cagcgggacg	13440
gtcgtccagg tcgagcagga cgaatcggtc aggggtgtcc gactgggcgg agcggatcag	13500
accccgccg gccgctccc ccaggtcttc tacctctcg tccggcgag tggccaccgc	13560

ggctcgggtg aggatcacca ggcgactgtc ggccaaccgg ctgtccgcca accactgctg	13620
cagtagggccc agcatgtggt gtgtggcctg ggcagcgccc tcactcgact cgtcttcgct	13680
cccggctggg gaccaggcca cgaccgtccg gggaacgggc gcgtcatcgg cgattgccac	13740
gcacagcgcg gccaggtcgg cgtacgtccg ctccggcgga gtgacgaagt cctcgaccgg	13800
gcgcagcgtc ggcccgtcc cgatcatcac ccatggggcg gcttcttcgg ccgctgaagg	13860
ctgggggacg ggttcccatc gcacctcgaa gagcgagcca tggggggcaa ccggggcggc	13920
tgtgcgaac tcggccgtat gcaactcggc cctggtcacc ggccgcaggg ccaacgaccg	13980
gaccgatgct acgggctcac ccgcctcgtc ggacacgtcg accgacaccg cccccgggcc	14040
atggccgtcc ggcctcagcc ggacgcgcag agccgtcgcc ccgcatgcac gcagggtgac	14100
tcctgtccac gtgaacggca gcccgtgtc ctgttgtccc cgtaggaggagg gcaaaagtgg	14160
aatgtgtgac agtgcggaat ccagcagcgc ggggtgcagc gcgaaccgag ccgcagccgt	14220
gtcggccagc tgatcctctg acatatacgc ggtcgcgaag acctcgtccc cgcgttgcca	14280
ggcggcgctc aggcaacgga aggccgcacc gtagtggaac cccccctcgg ccagttgtct	14340
gtacagtccc tcagatcca cggcgaccgc acccgccggc ggccatgcgc ccgcgggcat	14400
gacaccatcc gccgagtcac cgcagtcctg gtcaggcttg tcggcaaggt acccgccggc	14460
atggcgctc caccgctgc cagccgaatc ggcggtgagg ccgccgacat cgattcgca	14520
gtagatatcc accgcccgtt ttctgtctt gttcggcgcg gccacagtga cctgcagcac	14580
tacgccggcg ctatccgcta cgacgagcgg ggtgtggagc gtaagtcct cgaccgacg	14640
gcagtcgca tggcgggcgg cttggagggc gatttcgaga aaggccgtac cggggagcaa	14700
cacggcgccg ccgacggcgt ggtcggccag ccagggatgc gtacggagag aaatccgcc	14760
agtgaacagc aaccctcac catcgccaa cccgaccgcc gcgccaaca ggggatgccg	14820
cgccggctca aggcctgcag ccgacatatc accggtggca gtggtggtgt tgagccagta	14880
gtggtggtgt tggaaggggt aggtggggag gtttgggtgg tagtgggtgg ggtgtgatg	14940
gtggtgccag gtggtgggtt ggccgtgggt gtggaggtgg gcgagggcgg tgtgggtggt	15000
gtgggtctcg gattggtcgg ggtggagaag cgggacggct agcggggtgt ggtcggggag	15060
gttgtggtgg gtgagggcgg tgagggtgtg gtcggggccg agttcgaggt aggcggtgac	15120
gccgtgggtg tggagtgtt ggggtggtgt ggcgtagtgg accgtctgc gcgtttgccg	15180
gaccagtag tggggtgtgg tggggtcgtg gtcggtgggg ttggtggtga tgaggggggt	15240
gtggggtggg tggtaggtga gggtttctgc cgtgttcgg agttcgttga ggagtgggtc	15300

gatgtggtgt gagtggaaag cgccgctgac cttcagcgca gtggtcttgc gtcccatggc	15360
gcggaggttt tcggccctag cgtctacgtc gtggcgatca ccgctgatga ccacggaccc	15420
gggcgcgttc acggccgcga gtgagaccgc tccttcgcat cgtgtgaggt gttcgtggag	15480
ttcgtcgggg tcigtcttga ggggtggatc ggcgccggtg gcggggagag attgcatgag	15540
gcgggcgcgg gtggtgacca ggcgggccgc gtcggggagg gtgaggatcc cggcgaggtg	15600
ggccgcggtg atctctccga gggaatggcc ggcgtagtag tggggggtaa ggccgtggtg	15660
ttcggtgacc aggcggtgca gggcgatctg gagggcgaac agggcgggct gggcgtagcg	15720
ggctctggtg agcagctggg cgggtgctgg gcccggctcg gcgaacatga catcccgag	15780
gggatggtcg aggtgtgggt ccaggtgtgc gcaggcttcg tccatggctg cggcgaacag	15840
cgggaaggcg tagtacagcc cgtgtcccat acggggacgt tgcgttcctt gcccggaaca	15900
cacaaacgcc agcttccccg caccgtcgc cagcccccgt accagcccgg cgtgttcctc	15960
accggcctcc agagctccca gctcttcgac gaaggcttcg cggtcgggac ccaggagcac	16020
ggcacggtg ccgaaggccg tgcggccgga ggccagcgca taccgacat cagccagccc	16080
cagctcagga tgtgccgcca gatacgcgcg caactgtgct gcctgcgcct gtagtgctg	16140
ttcggacttg gccgacacca tccacggaaa cgacccggac gagacgtccg ccggtatctc	16200
gtgtctctca ggccggcgct gttcgaggat gatgtgcgcg ttggtgccac tgaccccgaa	16260
cgacgacacc cccgcacgcc gcgcacgccc cgtctccggc cagggcaccc cctccgtcag	16320
cagccgcacc ccgcccagg accagtccac atgcggcgac ggctcctgca catgcaacgt	16380
ccgcggaac accccatacc gcatcgccat caccatcttg atgacaccg caacccccgc	16440
cgcggcctgc gcatgccga tattcgactt caccgacccc aaccacaacg ggccgtcccc	16500
caccgcctcc tgaccgtagg tcgccagcaa cgctgcgcc tcgatcggt caccagccg	16560
cgtecccgta ccgtgcccct ccaccacatc cacatccgcc ccgccaaccc gcgcaactggc	16620
caacgcctgc cggatcacc gctgtgcga cggaccattc ggccgctca aaccattcga	16680
cgcaccatcc tgattcaccg cagaaccccg caccaccgcc aacacccgat gcccacccg	16740
ccgcgcatca gacaaccgt ccaccaacag catccccacc cctccgccc acccgtacc	16800
atccgcacc accccaaacg ccttgaccg ccatccacc gacaacccc gctgccgcgc	16860
aaactccag aaggcaccg gcgtcgccat caccgtcacc ccgccgcca gcgcatgga	16920
gcaactcccc gcacgcagcg cctgacacgc cagatgcaac gccaccagag acgacgaaca	16980
cgccgtgtcc accgacaccg cggggccttc caggccgaag gtgtaggcga tacggcccga	17040
ggcgacgtg ccggagtctt cggtagggac atagccctcg acctcgcca cgggtgtctt	17100
cagccggggg ccgtagtccg tcggcatcac accggcgaag actccggtcc ggctccgccc	17160

taccgacgag gggtcgattc ccgcgtgttc gaaaacctcc caggaggtct ccagcagcag	17220
ccgctgctgc ggatccatcg ccagcgcctc acgtgggccg atgccgaaga atcccgcgtc	17280
gaacttgccc gcgtcgtgga ggaagcccc gtggcgggtg tatgtggtgc ccggcacgtc	17340
ctggtcctgg ggggcgtaga ggcggtcgag gtcccagccg cggtcgctgg ggaatccgct	17400
catggcgtcg gtcccgagg cgaccagatc ccacagcccc tcggccgacc gcacgcccc	17460
cggataccgg caggccatgc cgatgatcgc gatggcctca tccgccacca tcgccacctt	17520
gctcggagca ggagccgcag ccagggtgac cgtatcggcg ccgccgatga ggccgtccc	17580
cagggtgctgg gcgagagcgg caggagtggg atggtcgaag acgatggtgg cggggagccg	17640
caggccggtt gcctccttga ggcggttcg gagctccacg gcggtgaggg agtcgaagcc	17700
gaggtcacgg aaggccgcct cgcgggcgat ggcgtcgggg ctcggtggc cgaggaccgt	17760
ggcggcgtgg ccgcggacga gggtagacag gacctcgtgc tgtcggcgcg ggtcgagccc	17820
ggcgagccgg gctcggaggg aggagcggcc gtcgggtgtc gcggggcgtg gccggtggcc	17880
ggtggtgggg cgtacgaggc tctggaacag gggcgggtgcg gtgtgtcccc gtagcgcctt	17940
ggtgtcgata cgggcccggga gcagcagcgg tgggccgtcg acgagagcgt tgtcgaacag	18000
cgccagggcc tcgtccgtgg ccagtggccg gaggccgacg cgggccatgc gggtagcggtc	18060
ggtctggtcg aggtggccgg tcattccgct ggcctcctcc cacagtcccc atgccaggga	18120
cacggcgggg agtcccatcg cccgtcgtg gaaggcgagc gcgtcgagaa aggcggtggc	18180
ggcggcgtag ttgccctgcc ccggcgagcc caggacaccc gcggcggacg agaacagcac	18240
gaacgcggcc agatccagcc ggcaggtcag gtcgtgcaga tgccaggcgg catcggcctt	18300
cgcgggcaga acggtgtcga ggtggcgcgc ggtgaggag ccgagcgtgg cgtcgtcgag	18360
ggtgcccgcg gcgtggacga cggcggtcag cgggtgctcg gcggggatcc gcccagcag	18420
ttcggcgagc gcggcacggt cgcgcagctc gcaggtggcg acggtcgccg tggcgccaa	18480
ctcgtcaatc tcagcgatga gttcgggagc gcctggcgca tcgaggccac ggcggctgat	18540
cagtagcaga tgccgtgctc ggccggtggc ggcgagatgg cggcgagcgc cggcgccgag	18600
ggtaccggtg ccaccggtga tgaggacggt gccgtcgggg tcgggggtgc ggtccagcgt	18660
gagcacattc tttccgacgg tgcggccctg ctggagtgg cggaaggcat ccctggcctg	18720
ccggtatgtec cagcaggtag gcggcagcgg ggtgagcacg ccctcgtcga acagccgcaa	18780
cacggtgcgc aggatctccc cgatcgcctc gggctccggc tccatcaggt cgtacgccc	18840
gtaggcgaca ccgggttacg cctgcgccac ctctccggg tcgaggaggt cggctttgcc	18900
catctccagg aaccggccgc tgtcgccgga caccaggcgc agcgaggcgt cgacgtactc	18960
atgggcccagg gatttagga cgacgtcgt gttccggccg gtgaaccggc tctcgaagtc	19020

gaggggtgcgg gacgaggcga tgcgtccgg ggggatgccg tgggcgcgta ggggtgccca	19080
tttgccgggg ctgccgtgg cgtagacctc agcgcgccagg tggcgggcca gttggacagc	19140
ggccagtccc actccgccgg ccgcagcgtg caccagcacc gactcaccg ggcgcagtec	19200
tgccaggtcc accagcccgt aataggcggf ggcgaacacg acggggacgg cggcggcctg	19260
gccgaaggtc cagtggtcgg ggacgcggat caccgtggcc cggtcggcga tcgcgaacgg	19320
gccgaacgca tccgcgaaca cgcccatcac ccggtcccg ggccgccagg cggccacttc	19380
agggccggct tcgagcacca ccccgccgcc ctgcgtgcc agcccccgct gcccaggac	19440
catccccagg gccaccagga catcgcgga gttcagcccg gccgcacgca cctcgatccg	19500
tacctggtgc gggcccagcg ggtcgcggtc ggccgaggac gccacaggcg ccagcccctg	19560
gaccgtgcc gcgtcggcgc agtcgagccg ccagccgcc ctcccctgtg gcggcgtgag	19620
cagggcgccc cctcgggtgt cggctcgaac cagccgggcc atgtcggct ggccggcccc	19680
gatggccatc tgcggttcgc ccagcgcgac cggcgtgggt agcgcgccccc acgactcctc	19740
ggccgcgtcc aggtccacca ggagatacgt gtcgggctcc tcggcctgcg cggatcgac	19800
cagtcccgag atcgccgcgc ccgccaggtc catcactccc tccgccggcc cgggtggccac	19860
cgcgcgcgag gtgaccagca ccagcctgc gccggccagc cgcgcttcgc cgagccagcc	19920
ctgaagcagg tccagcgtt gtccgggtacg gacacgggta cgggcgggtg ggacgtcggc	19980
gaccgtcgtg tcgtcgggcc cggcggcagc cggaggcag cccacgaaca ccaggtccgg	20040
aggcggatcg cccgcgtcca gggccgtcac cagtccgca aggtccgggt accgccgtac	20100
gggccccgct gccgcgctt ccagcgcggc agcccatcca ggaatggctt ggccaatcac	20160
agcgcaccgc ccgagggccg tcgccggggg ctggggcgcg gcgtccact ccagggcgta	20220
gaggggtgctg cccgcctggc gcccgaggc ccgcaggtcc tccagggcca ccggccgag	20280
cgtgagcgat tcggccgagg ccaccggcgc gcccgctctg tcggccacgt cgacggcaat	20340
cgagtccggg ccttcgggcc gcagccgtac gcgcagcgtc ggtgtgcccg gggagcgag	20400
gctgacgcc cgccaggcga acggcagcag gccatgtccc gctggattct cctcgcttcg	20460
gccgtgtcg agtgccagcg ggtggagcgc ggcgtcagc agcgcggggt ggaggcagaa	20520
gcgtggggcg tcggcgcca cctggtcggg gagccgacc tcggcgtagg tctcctgcc	20580
ttgccccag acggcgcgca gcccttgaa ggcggggccg tagccgaagt ctgcggcggc	20640
cagcgtgtcg tagaggccgc tgatgtccag gggctccgc cccggcgag gccacgcccc	20700
gccagggtc ccggcaccgt cggcaccgt gatgccatca gtcggggcct cggccgccag	20760

tgtgccggtc gcatgccgtg tccacgcgcc ccgggcccgc ccctcccacg cggggtcgtc	20820
gtcgcgctgt gcgtagacgc tcatgtcgcg gcggccggtc tcgtccggtg ccccgacggc	20880
cacctggagc gccacgccgc cgggtgcggg gaggacgagc ggggcgtgca gggtagctc	20940
ctccagccgg gggcagccga ggtgtccgcc gacgtgcacg gccagtcca cgtatgcggt	21000
gcccggcagc aggacgggtc cggcgaccgc gtggtcggtg agccaggggt ggggtgtcgg	21060
ggcgagccgg ccggtgagca gatatccgtc gccctcggcg agtccgaccg ccgcgccgag	21120
cagtgggtgt ccgaccgggg ccagccccgc ggccgtcacg tcccgggtgg tcgccggggc	21180
gtccagccag tagcggcggc gctggaaggc gtaggtgggc agatctaggt ggtggttgtg	21240
ggggtgggtg tgggtgggg tgtagtgtg gtgccagggt gtgggtgtgc cgatgccgt	21300
ggtgtgggtg tgggcgaggc cggtaggaa gcggcgggtg tcgttgtcgc cgcggcggag	21360
gctgccgacg gcggtgacgt cttcggcgggt gtcttcggtg gtgtcttcga tggcggggac	21420
gagggtgggg tgggggctga cttcgacgaa gacgcggtgt ccgtcatccg ccagggcctg	21480
gacggcatcg ctgaaacgga caggctggtg caggttgcgg taccagtagg cggcgtccag	21540
ggttgtggtg tccagccagg tgccctccac cgtggagaag aacggcacgc cggacggctg	21600
cgggctgatg tccccagca gctccagca ctcctccgc aggggctgca catgggggca	21660
gtgcgaggca tagtcgaccg ggatccgccg ggcccgcacc ccggtgccgg cacagtacgc	21720
cagcacctcg tcacccgct cggcatcccc cgagacggcg gtggagcggg ggccgttgac	21780
cgccgccacc cacaaccgcc ccgcccaccg ctcaccaatg agtgcctcca ctcctgggc	21840
aggcagcggc actgaggcca tgccgccccg gcccgcaca gcggccagcg cccggtgcg	21900
cagcgcaaca gtcttcgccg cgtcttcag gtcagcgcc ccacacacat gcgcggccgc	21960
gatctcgccc tggaatggc caaggaccgc gtcgggttcg ataccgtagg aacgccacag	22020
agcagccaaa gacacatga cgctgaacag cacaggtgg accacatcg cccgtccca	22080
caccgcatcc cccgcgtccc ggcgcaggat gtccaccaca gaccagtcca cccacggcgc	22140
cagagcctcc tcgcacgct ccatccgccg ggcgaaacacc ggagaggagg cgagcagacg	22200
cacaccatc cgggcccact gcccacctg tccgggaaac acgaagacga caccgccccg	22260
gtcaccacc ggcgcatgac ccgtggtcac ccgccgatcc ggctcaccg ccgccagcgc	22320
ccccaacct tgcaccagt cctcacggtc cgcggccagg acgaccgcac gatgctccag	22380
cacagccgc ccacaggcca gaccgcacc cacatcggca agcgaaacgt aagcttgga	22440
ctggccgtcg ttttacaac tcgtgactgg gaaaaccctg gcgttaccca acttaatcg	22500
cttcgagcac atccccctt ccgcagctgg cgtaatagcg aagaggcccg caccgatcg	22560
ccttccaac agttgcgac cctgaatggc gaatggcgcc tgatgcggtg ttttctctt	22620

acgcatctgt gcggtatttc acaccgata aattcccaa tgtcaagcac ttccggaatc	22680
gggagcgcg ccatgcaaa gtgccgatca acataacgat cttttagaa accatcggcg	22740
cagctattta cccgcaggac atatccacgc cctctacat cgaagctgaa agcacgagat	22800
tcttcgccct gcgagagctg catcaggctg gagacgctgt cgaacttttc gatcagaaac	22860
 ttctcgacag acgtagat	 22878