(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2017-0065028
(43) 공개일자 2017년06월12일

(51) 국제특허분류(Int. Cl.)
C12N 9/14 (2006.01) *C12Q 1/68* (2006.01)
(52) CPC특허분류
C12N 9/14 (2013.01)
C12Q 1/6844 (2013.01)
(21) 출원번호 10-2017-7011954
(22) 출원일자(국제) 2015년09월29일
심사청구일자 2017년04월28일
(85) 번역문제출일자 2017년04월28일
(86) 국제출원번호 PCT/US2015/053012
(87) 국제공개번호 WO 2016/054088
국제공개일자 2016년04월07일
(30) 우선권주장
62/057,056 2014년09월29일 미국(US)

(71) 출원인
일루미나 케임브리지 리미티드
영국 에섹스 니어 새프런 월든, 리틀 체스터폴드 (우: 씨비10 1엑스엘)
(72) 발명자
보마티 에린
미국 캘리포니아주 92122 샌디에이고 일루미나 웨이 5200
켈링거 매튜 윌리엄
미국 캘리포니아주 92122 샌디에이고 일루미나 웨이 5200
부텔 조나단 마크
영국 니어 새프런 월든 에섹스 씨비10 1엑스엘 리틀 체스터포드
(74) 대리인
특허법인아주김장리

전체 청구항 수 : 총 39 항

(54) 발명의 명칭 **재조합효소 돌연변이체**

(57) 요 약

개선된 클러스터 증폭을 위한 패턴화된 플로우 셀 표면 상에서의 단일-가닥 어댑터 영역을 갖는 PCR-라이브러리와 같은 핵산의 개선된 재조합효소-매개의 증폭을 위한 재조합효소, 및 이를 이용하는 방법 및 키트가 본 명세서에 제시된다.

대 표 도 - 도2

```
RB49    MS-VLEKLKKNSTLKTTAVLSKSSFFNEKTNTRTKIPMLNIAFSGDLKKGFQSGLIFFAGPSKHFK  65
T4      MSDLKSRLIKASTSKLTAELTASKFFNEKDVVRTKIPMMNIALSGEITGGMQSGLLILAGPSKSFK  66

RB49    SNMGLTCVSAYMKQNPDAACLFFDSEFGITSAYLESMGVDPDRVVHVPIKNIEELKFEIM  125
T4      SNFGLTMVSSYMRQYPDAVCLFYDSEFGITPAYLRSMGVDPERVIHTPVQSLEQLRIDMV  126

RB49    NQLEQITREDKVIIFIDSIGNLASKKEVEDAINEKSAQDMTRAKALKGLFRMVTPYLTMN  185
T4      NQLDAIERGEKVVVFIDSLGNLASKKETEDALNEKVVSDMTRAKTMKSLFRIVTPYFSTK  186

RB49    DIPCIAINHTYETQEMFSKTVMSGGTGAMYSANEVFIIGRRQQKEGTEITGYDFILNAEK  245
T4      NIPCIAINHTYETQEMFSKTVMGGGTGPMYSADTVFIIGKRQIKDGSDLQGYQFVLNVEK  246

RB49    SRTVKEKSKFFISVTFSGGIDPYSGLLELAVELGWVVKPSNGWYSRSILNTETGEMETEE  305
T4      SRTVKEKSKFFIDVKFDGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEETGEMIREE  306

RB49    RKFRAKETNSIEFWKPLLTNDKFNEAINDHYKLGQVISDEAVDKEIEDML  355
T4      KSWRAKDTNCTTFWGPLFKHQPFRDAIKRAYQLGAIDSNEIVEAEVDELI  356
```

(52) CPC특허분류
C12Y 306/01008 (2013.01)
C12Q 2521/501 (2013.01)

명 세 서

청구범위

청구항 1

서열번호 1과 적어도 60%, 70%, 80%, 90%, 95%, 99% 동일한 아미노산 서열을 포함하는 재조합 UvsX로서, 상기 재조합 UvsX가 RB49 UvsX 아미노산 서열에서의 Pro256과 기능적으로 동등한 위치에 아미노산 치환 돌연변이를 포함하는, 재조합 UvsX.

청구항 2

제1항에 있어서, 상기 치환 돌연변이가 하전된 잔기로의 돌연변이를 포함하는, 재조합 UvsX.

청구항 3

제1항에 있어서, 상기 치환 돌연변이가 염기성 잔기로의 돌연변이를 포함하는, 재조합 UvsX.

청구항 4

제1항에 있어서, 상기 치환 돌연변이가 RB49 UvsX 아미노산 서열에서의 Pro256Lys과 상동성인 돌연변이를 포함하는, 재조합 UvsX.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 재조합 UvsX가 RB49 UvsX 아미노산 서열에서의 His63와 기능적으로 동등한 위치에 치환 돌연변이를 더 포함하는, 재조합 UvsX.

청구항 6

제5항에 있어서, 상기 재조합 UvsX가 RB49 UvsX 아미노산 서열에서의 His63Ser과 상동성인 치환 돌연변이를 포함하는, 재조합 UvsX.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 상기 재조합 UvsX가 C-말단에 하나 이상의 글루탐산 잔기의 부가; C-말단에 하나 이상의 아스파르트산 잔기의 부가; 및 이들의 조합으로 이루어진 군으로부터 선택되는 돌연변이를 더 포함하는, 재조합 UvsX.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 재조합 UvsX가 T4, T6, Rb69, Aeh1, KVP40, 아시네토박터(Acinetobacter) 파지 133, 아에로모나스(Aeromonas) 파지 65, 시아노파지(cyanophage) P-SSM2, 시아노파지 PSSM4, 시아노파지 S-PM2, Rb32, 비브리오(Vibrio) 파지 nt-1, Rb16, Rb43 및 Rb49로 이루어진 군으로부터 선택되는 미오비리대(myoviridae) 파지로부터 유래되는, 재조합 UvsX.

청구항 9

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 재조합 UvsX가 T2, Rb14, 아에로모나스 파지 25, phi-1, 파지 31, 파지 44RR2.8t, 파지 Rb3 및 파지 LZ2로 이루어진 군으로부터 선택되는 미오비리대 파지로부터 유래되는, 재조합 UvsX.

청구항 10

서열번호 2 및 22 내지 35 중 어느 하나의 아미노산 서열을 포함하는, 재조합 UvsX.

청구항 11

서열번호 3 내지 5 중 임의의 것의 아미노산 서열을 포함하는 반-보존된 도메인에 대한 치환 돌연변이를 포함하

는 재조합 UvsX로서, 상기 치환 돌연변이가 위치 7에서 Phe, Pro, Asp, Glu 또는 Asn 이외의 임의의 잔기로의 치환으로부터 선택되는 돌연변이를 포함하는, 재조합 UvsX.

청구항 12

제11항에 있어서, 상기 돌연변이가 하전된 잔기로의 돌연변이를 포함하는, 재조합 UvsX.

청구항 13

제11항에 있어서, 상기 돌연변이가 위치 7에서 Lys으로의 치환을 포함하는, 재조합 UvsX.

청구항 14

서열번호 6 내지 7 중 임의의 것의 아미노산 서열을 포함하는 반-보존된 도메인에 대한 치환 돌연변이를 포함하는 재조합 UvsX로서, 상기 치환 돌연변이가 위치 12에서 Phe, Pro, Asp, Glu 또는 Asn 이외의 임의의 잔기로의 치환으로부터 선택되는 돌연변이를 포함하는, 재조합 UvsX.

청구항 15

제14항에 있어서, 상기 돌연변이가 하전된 잔기로의 돌연변이를 포함하는, 재조합 UvsX.

청구항 16

제14항에 있어서, 상기 돌연변이가 위치 12에서 Lys으로의 치환을 포함하는, 재조합 UvsX.

청구항 17

제10항 내지 제15항 중 어느 한 항에 있어서, 상기 재조합 UvsX가 RB49 UvsX 아미노산 서열에서의 His63과 기능적으로 동등한 위치에 치환 돌연변이를 더 포함하는, 재조합 UvsX.

청구항 18

제17항에 있어서, 상기 재조합 UvsX가 RB49 UvsX 아미노산 서열에서의 His63Ser과 상동성인 치환 돌연변이를 포함하는, 재조합 UvsX.

청구항 19

제10항 내지 제18항 중 어느 한 항에 있어서, 상기 재조합 UvsX가 C-말단에서의 하나 이상의 글루탐산 잔기의 부가; C-말단에서의 하나 이상의 아스파르트산 잔기의 부가; 및 이들의 조합으로 이루어진 군으로부터 선택되는 돌연변이를 더 포함하는, 재조합 UvsX.

청구항 20

제10항 내지 제19항 중 어느 한 항에 있어서, 상기 재조합 UvsX가 T4, T6, Rb69, Aeh1, KVP40, 아시네토박터 파지 133, 아에로모나스 파지 65, 시아노파지 P-SSM2, 시아노파지 PSSM4, 시아노파지 S-PM2, Rb32, 비브리오 파지 nt-1, Rb16, Rb43 및 Rb49로 이루어진 군으로부터 선택되는 미오비리대 파지로부터 유래되는, 재조합 UvsX.

청구항 21

제10항 내지 제19항 중 어느 한 항에 있어서, 상기 재조합 UvsX가 T2, Rb14, 아에로모나스 파지 25, phi-1, 파지 31, 파지 44RR2.8t, 파지 Rb3 및 파지 LZ2로 이루어진 군으로부터 선택되는 미오비리대 파지로부터 유래되는, 재조합 UvsX.

청구항 22

제1항 내지 제21항 중 어느 한 항에 정의된 바와 같은 재조합 UvsX를 인코딩하는 핵산 분자.

청구항 23

제22항의 핵산 분자를 포함하는 발현 벡터.

청구항 24

제23항의 벡터를 포함하는 숙주 세포.

청구항 25

표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법으로서,

(a) 제1항 내지 제21항 중 어느 한 항의 재조합 UvsX를 제1 및 제2 핵산 프라이머와 접촉시켜, 제1 및 제2 핵단백질 프라이머를 형성하는 단계로서, 상기 핵산 프라이머가 3' 말단에 단일 가닥 영역을 포함하는, 상기 제1 및 제2 핵단백질 프라이머를 형성하는 단계;

(b) 상기 제1 및 제2 핵단백질 프라이머를 상기 표적 핵산 분자에 접촉시킴으로써, 상기 제1 가닥의 제1 부분에 제1 이중-가닥 구조를 형성하고, 상기 제2 가닥의 제2 부분에 제2 이중 가닥 구조를 형성하여, 상기 제1 핵산 프라이머 및 상기 제2 핵산 프라이머의 3' 말단이 동일한 이중-가닥 주형 핵산 분자 상에서 서로를 향해 배향되게 하는 단계;

(c) 상기 제1 및 제2 핵산 프라이머의 3' 말단을 하나 이상의 중합효소 및 dNTP로 연장시켜, 제1 및 제2 이중-가닥 핵산 및 핵산의 제1 및 제2 치환 가닥을 생성하는 단계; 및

(d) 목적으로 하는 증폭 정도가 달성될 때까지 단계 (b) 및 (c)의 반복을 통해 반응을 지속시키는 단계를 포함하는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 26

제25항에 있어서, 상기 표적 핵산 분자가 이중 가닥 핵산을 포함하는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 27

제25항에 있어서, 상기 표적 핵산 분자가 단일 가닥 어댑터 영역을 포함하는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 28

제25항에 있어서, 상기 방법이 재조합효소 로딩 단백질의 존재 하에 수행되는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 29

제28항에 있어서, 상기 재조합효소 로딩 단백질이 T4 UvsY, 에스케리키아 콜라이(*E. coli*) recO, 에스케리키아 콜라이 recR 및 이들의 조합으로 이루어진 군으로부터 선택되는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 30

제25항에 있어서, 상기 방법이 gp32, 에스케리키아 콜라이 SSB 단백질, T4 gp32 단백질 및 그들의 유도체로 이루어진 군으로부터 선택되는 단일 가닥 안정화제의 존재 하에 수행되는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 31

제25항에 있어서, 상기 방법이 폴리에틸렌 글라이콜, 폴리에틸렌 옥사이드, 폴리스타이렌, 피콜(Ficoll), 덱스트란, PVP 및 알부민을 포함하는 군으로부터 선택되는 크라우딩 작용제(crowding agent)의 존재 하에 수행되어, 상기 크라우딩 작용제가 증폭을 자극하게 하는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 32

제25항에 있어서, 상기 방법이 증폭 위치의 어레이에서 수행되는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 33

제32항에 있어서, 각 증폭 위치가 상기 표적 핵산의 증폭을 위한 복수의 증폭 프라이머를 포함하는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 34

제32항에 있어서, 상기 증폭 위치의 어레이가 표면 상의 특징부(feature)의 어레이를 포함하는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 35

제34에 있어서, 상기 특징부는 불연속적이며, 상기 증폭 프라이머가 결여된 상기 표면의 틈새(interstitial) 영역에 의해 분리되는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 36

제32항에 있어서, 상기 증폭 위치의 어레이가 용액 중의 비드 또는 표면 상의 비드를 포함하는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 37

제32항에 있어서, 상기 증폭 위치의 어레이가 에멀전을 포함하는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 38

제25항에 있어서, 상기 방법이 등온으로 발생하는, 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법.

청구항 39

제1항 내지 제21항 중 어느 한 항에 정의된 바와 같은 재조합 UvsX 및 다음 중 하나 이상을 포함하는, 재조합효소 중합효소 반응을 수행하기 위한 키트:

단일 가닥 DNA 결합 단백질;

DNA 중합효소;

dNTP, 또는 dNTP 및 ddNTP의 혼합물;

크라우딩 작용제;

완충제;

환원제;

ATP 또는 ATP 유사체;

재조합효소 로딩 단백질;

제1 프라이머 및 임의로 제2 프라이머.

발명의 설명

기 술 분 야

[0001] 본 발명은 개선된 재조합효소-매개 핵산의 증폭을 위한 재조합효소에 관한 것이다.

배 경 기 술

[0002] 재조합효소는 재조합효소-매개 핵산의 증폭에 유용하다. 예를 들어, 재조합효소는 DNA 표적으로의 올리고뉴클레오타이드의 표적화를 용이하게 하여, 중합효소에 의한 DNA의 복제를 가능하게 할 수 있다. 개선된 특성을 갖는 변형된 재조합효소가 여전히 필요하다.

[0003] **서열 목록**

[0004] 본 출원은 전자 형식의 서열 목록과 함께 출원된다. 서열 목록은 크기가 98 Kb인 2014년 9월 26일에 생성된 파일명이 IP1264.TXT인 파일로서 제공된다. 서열 목록의 전자 형식의 정보는 그의 전문이 본 명세서에 참조로 포함된다.

발명의 내용

[0005] 개선된 재조합효소-매개 핵산의 증폭을 위한 재조합효소가 본 명세서에 제시된다. 본 발명자들은 패턴화된 플로우 셀(patterned flow cell) 표면 상으로의 핵산의 씨딩에서 상당히 개선된 특징을 갖는 특정한 변경된 재조합효소를 놀랍게도 확인하였다. 특정 실시형태에 있어서, 변경된 재조합효소는 개선된 클러스터 증폭을 위해 패턴화된 플로우 셀 표면 상의, 단일-가닥 어댑터 영역을 갖는 PCR-라이브러리와 같은 PCR-프리(free) 라이브러리의 씨딩(seeding)을 개선시킨다.

[0006] 특정 실시형태에 있어서, 재조합효소는 재조합 UvsX이며, RB49 UvsX 아미노산 서열에서의 Pro256과 기능적으로 동등한 위치에 아미노산 치환 돌연변이를 포함한다. 야생형 RB49 UvsX 아미노산 서열은 서열번호 1에 기재되어 있다. 특정 실시형태에 있어서, 재조합 UvsX는 서열번호 1과 적어도 60%, 70%, 80%, 90%, 95%, 99% 동일한 아미노산을 포함하고, RB49 UvsX 아미노산 서열에서의 Pro256과 기능적으로 동등한 위치에 아미노산 치환 돌연변이를 포함하는 아미노산 서열을 포함한다. 특정 실시형태에 있어서, 치환 돌연변이는 하전된 잔기로의 돌연변이를 포함한다. 특정 실시형태에 있어서, 치환 돌연변이는 염기성 잔기로의 돌연변이를 포함한다. 특정 실시형태에 있어서, 치환 돌연변이는 RB49 UvsX 아미노산 서열에서의 Pro256Lys과 상동성인 돌연변이를 포함한다.

[0007] 일부 실시형태에 있어서, 재조합 UvsX는 상기 돌연변이에 더하여, RB49 UvsX 아미노산 서열에서의 His63과 기능적으로 동등한 위치에 치환 돌연변이를 추가로 포함할 수 있다. 예를 들어, 특정 실시형태에 있어서, 재조합 UvsX는 RB49 UvsX 아미노산 서열에서의 His63Ser과 상동성인 치환 돌연변이를 포함한다.

[0008] 일부 실시형태에 있어서, 재조합 UvsX는 상기 돌연변이 중 임의의 것에 더하여, C-말단에 하나 이상의 글루탐산 잔기의 부가; C-말단에 하나 이상의 아스파르트산 잔기의 부가; 및 이들의 조합으로 이루어진 군으로부터 선택되는 돌연변이를 추가로 포함할 수 있다.

[0009] 일부 실시형태에 있어서, 재조합 UvsX는 T4, T6, Rb69, Aeh1, KVP40, 아시네토박터(*Acinetobacter*) 파지 133, 아에로모나스(*Aeromonas*) 파지 65, 시아노파지(cyanophage) P-SSM2, 시아노파지 PSSM4, 시아노파지 S-PM2, Rb32, 비브리오(Vibrio) 파지 nt-1, Rb16, Rb43 및 Rb49로 이루어진 군으로부터 선택되는 미오비리대(myoviridae) 파지로부터 유래된다.

[0010] 일부 실시형태에 있어서, 재조합 UvsX는 T2, Rb14, 아에로모나스 파지 25, phi-1, 파지 31, 파지 44RR2.8t, 파지 Rb3 및 파지 LZ2로 이루어진 군으로부터 선택되는 미오비리대 파지로부터 유래된다.

[0011] 또한, 서열번호 2 및 22 내지 35 중 어느 하나의 아미노산 서열을 포함하는 재조합 UvsX가 본 명세서에 제시된다. 특정 실시형태에 있어서, 재조합 UvsX는 서열번호 2 및 22 내지 35 중 어느 하나와 적어도 60%, 70%, 80%, 90%, 95%, 99% 동일한 아미노산을 포함하며, RB49 UvsX 아미노산 서열에서의 Pro256과 기능적으로 동등한 위치에 아미노산 치환 돌연변이를 포함하는 아미노산 서열을 포함한다.

[0012] 또한, 서열번호 3 내지 5 중 임의의 것의 아미노산 서열을 포함하는 반-보존된 도메인에 대한 치환 돌연변이를 포함하는 재조합 UvsX가 본 명세서에 제시되며, 치환 돌연변이는 위치 7에서 Phe, Pro, Asp, Glu 또는 Asn 이외의 임의의 잔기로의 치환으로부터 선택되는 돌연변이를 포함한다. 특정 실시형태에 있어서, 재조합 UvsX는 서열번호 3 내지 5 중 임의의 것의 아미노산 서열을 포함하는 반-보존된 도메인을 포함하는 재조합효소와 적어도 60%, 70%, 80%, 90%, 95%, 99% 동일한 아미노산을 포함하며, 재조합 UvsX는 위치 7에서 Phe, Pro, Asp, Glu 또는 Asn 이외의 임의의 잔기로의 치환으로부터 선택되는 치환 돌연변이를 포함한다. 특정 실시형태에 있어서, 돌연변이는 하전된 잔기로의 돌연변이를 포함한다. 특정 실시형태에 있어서, 돌연변이는 염기성 잔기로의 돌연변이를 포함한다. 특정 실시형태에 있어서, 돌연변이는 위치 7에서 Lys으로의 치환을 포함한다.

[0013] 또한, 서열번호 6 내지 7 중 임의의 것의 아미노산 서열을 포함하는 반-보존된 도메인에 대한 치환 돌연변이를 포함하는 재조합 UvsX가 본 명세서에 제시되며, 치환 돌연변이는 위치 12에서 Phe, Pro, Asp, Glu 또는 Asn 이외의 임의의 잔기로의 치환으로부터 선택되는 돌연변이를 포함한다. 특정 실시형태에 있어서, 재조합 UvsX는 서열번호 6 내지 7 중 임의의 것의 아미노산 서열을 포함하는 반-보존된 도메인을 포함하는 재조합효소와 적어도 60%, 70%, 80%, 90%, 95%, 99% 동일한 아미노산을 포함하며, 재조합 UvsX는 위치 12에서 Phe, Pro, Asp, Glu

또는 Asn 이외의 임의의 잔기로의 치환으로부터 선택되는 치환 돌연변이를 포함한다. 특정 실시형태에 있어서, 돌연변이는 하전된 잔기로의 돌연변이를 포함한다. 특정 실시형태에 있어서, 돌연변이는 염기성 잔기로의 돌연변이를 포함한다. 특정 실시형태에 있어서, 돌연변이는 위치 12에서 Lys으로의 치환을 포함한다.

[0014] 일부 실시형태에 있어서, 재조합 UvsX는 상기 돌연변이에 더하여, RB49 UvsX 아미노산 서열에서의 His63과 기능적으로 동등한 위치에 치환 돌연변이를 추가로 포함할 수 있다. 예를 들어, 특정 실시형태에 있어서, 재조합 UvsX는 RB49 UvsX 아미노산 서열에서의 His63Ser과 상동성인 치환 돌연변이를 포함한다.

[0015] 일부 실시형태에 있어서, 재조합 UvsX는 상기 돌연변이 중 임의의 것에 더하여, C-말단에서의 하나 이상의 글루탐산 잔기의 부가; C-말단에서의 하나 이상의 아스파르트산 잔기의 부가; 및 이들의 조합으로 이루어진 군으로부터 선택되는 돌연변이를 추가로 포함할 수 있다.

[0016] 일부 실시형태에 있어서, 재조합 UvsX는 T4, T6, Rb69, Aeh1, KVP40, 아시네토박터 파지 133, 아에로모나스 파지 65, 시아노파지 P-SSM2, 시아노파지 PSSM4, 시아노파지 S-PM2, Rb32, 비브리오 파지 nt-1, Rb16, Rb43 및 Rb49로 이루어진 군으로부터 선택되는 미오비리대 파지로부터 유래된다.

[0017] 일부 실시형태에 있어서, 재조합 UvsX는 T2, Rb14, 아에로모나스 파지 25, phi-1, 파지 31, 파지 44RR2.8t, 파지 Rb3 및 파지 LZ2로 이루어진 군으로부터 선택되는 미오비리대 파지로부터 유래된다.

[0018] 또한, 임의의 상기 실시형태에서 정의된 바와 같은 재조합 UvsX를 인코딩하는 핵산 분자가 본 명세서에 제시된다. 또한, 상기 기재된 핵산 분자를 포함하는 발현 벡터가 본 명세서에 제시된다. 또한, 상기 기재된 벡터를 포함하는 숙주 세포가 본 명세서에 제시된다.

[0019] 또한, (a) 상기 실시형태 중 임의의 것의 재조합 UvsX를 제1 및 제2 핵산 프라이머와 접촉시켜, 제1 및 제2 핵단백질 프라이머를 형성하는 단계로서, 상기 핵산 프라이머가 3' 말단에 단일 가닥 영역을 포함하는, 상기 제1 및 제2 핵단백질 프라이머를 형성하는 단계; (b) 제1 및 제2 핵단백질 프라이머를 상기 표적 핵산 분자와 접촉시킴으로써, 상기 제1 가닥의 제1 부분에 제1 이중-가닥 구조를 형성하고, 상기 제2 가닥의 제2 부분에 제2 이중 가닥 구조를 형성하여, 상기 제1 핵산 프라이머 및 상기 제2 핵산 프라이머의 3' 말단이 동일한 이중-가닥 주형 핵산 분자 상에서 서로를 향해 배향되게 하는 단계; (c) 상기 제1 및 제2 핵산 프라이머의 3' 말단을 하나 이상의 중합효소 및 dNTP로 연장시켜, 제1 및 제2 이중-가닥 핵산 및 핵산의 제1 및 제2 치환 가닥을 생성하는 단계; 및 (d) 목적으로 하는 증폭 정도가 달성될 때까지 단계 (b) 및 (c)의 반복을 통해 반응을 지속시키는 단계를 포함하는 표적 핵산 분자의 증폭의 재조합효소 중합효소 증폭 방법이 본 명세서에 제시된다.

[0020] 상기 방법의 특정 실시형태에 있어서, 표적 핵산 분자는 이중 가닥 핵산을 포함한다. 특정 실시형태에 있어서, 표적 핵산 분자는 단일 가닥 핵산을 포함한다. 예를 들어, 일부 실시형태에 있어서, 표적 핵산은 단일 가닥 어댑터 영역을 포함한다. 특정 실시형태에 있어서, 상기 방법은 재조합효소 로딩 단백질의 존재 하에 수행된다. 예를 들어, 재조합효소 로딩 단백질은 T4 UvsY, 에스케리키아 콜라이(*E. coli*) recO, 에스케리키아 콜라이 recR 및 이들의 조합으로 이루어진 군으로부터 선택될 수 있다. 특정 실시형태에 있어서, 상기 방법은 gp32, 에스케리키아 콜라이 SSB 단백질, T4 gp32 단백질 및 그들의 유도체로 이루어진 군으로부터 선택되는 단일 가닥 안정화제의 존재 하에 수행된다. 특정 실시형태에 있어서, 상기 방법은 폴리에틸렌 글라이콜, 폴리에틸렌 옥사이드, 폴리스타이렌, 피콜(Ficoll), 덱스트란, PVP 및 알부민을 포함하는 군으로부터 선택되는 크라우딩 작용제(crowding agent)의 존재 하에 수행되어, 크라우딩 작용제가 증폭을 자극하게 한다.

[0021] 특정 실시형태에 있어서, 상기 방법은 증폭 위치의 어레이에서 수행된다. 특정 실시형태에 있어서, 각각의 증폭 위치는 표적 핵산의 증폭을 위한 복수의 증폭 프라이머를 포함한다. 특정 실시형태에 있어서, 증폭 위치의 어레이는 표면 상의 특징부(feature)의 어레이를 포함한다. 예를 들어, 특징부는 불-연속적일 수 있으며, 증폭 프라이머가 결여된 표면의 틈새(interstitial) 영역에 의해 분리될 수 있다. 특정 실시형태에 있어서, 증폭 위치의 어레이는 용액 중의 비드 또는 표면 상의 비드를 포함한다. 특정 실시형태에 있어서, 증폭 위치의 어레이는 에멀전을 포함한다. 특정 실시형태에 있어서, 상기 방법은 등온으로 발생한다.

[0022] 또한, 재조합효소 중합효소 반응을 수행하기 위한 키트가 본 명세서에 제시된다. 특정 실시형태에 있어서, 키트는 임의의 상기 실시형태에서 정의된 바와 같은 재조합 UvsX, 및 하기의 것들 중 하나 이상을 포함할 수 있다: 단일 가닥 DNA 결합 단백질; DNA 중합효소; dNTP 또는 dNTP 및 ddNTP의 혼합물; 크라우딩 작용제; 완충제; 환원제; ATP 또는 ATP 유사체; 재조합효소 로딩 단백질; 제1 프라이머 및 임의로 제2 프라이머.

[0023] 하나 이상의 실시형태의 상세사항은 하기의 첨부 도면 및 설명에 기재되어 있다. 다른 특징, 목적 및 이점이 설

명 및 도면으로부터, 그리고 청구범위로부터 명백해질 것이다.

도면의 간단한 설명

[0024] 도 1은 US 2009/0029421호의 혼입 물질에도 기재된 바와 같은 엔테로박테리아 파지(*Enterobacteria phage*) T4(T4)(서열번호 8), 엔테로박테리아 파지 T6(T6)(서열번호 9), 아시네토박터 파지 133(Phage133)(서열번호 10), 엔테로박테리아 파지 RB69(Rb69)(서열번호 11), 아에로모나스 파지 Aeh1(Aeh1)(서열번호 12), 아에로모나스 파지 65(Ae65)(서열번호 13), 비브리오 파지 KVP40(Kvp40)(서열번호 14), 엔테로박테리아 파지 RB43(Rb43)(서열번호 15), 프로클로로코커스 파지(*Prochlorococcus phage*) P-SSM2(PSSM2)(서열번호 16) 및 프로클로로코커스 파지 P-SSM4(PSSM4)(서열번호 17)로부터의 UvsX 아미노산 서열의 정렬을 보여주는 개략도이다. RB49 UvsX 아미노산 서열에서의 Pro256과 위치적으로 및/또는 기능적으로 동등한 잔기가 강조표시되어 있으며, 삼각형 기호에 의해 표기되어 있다.

도 2는 엔테로박테리아 파지 RB49(RB49)(서열번호 1) 및 엔테로박테리아 파지 T4(T4)(서열번호 8)로부터의 UvsX 아미노산 서열의 정렬을 보여주는 개략도이다. RB49 UvsX 아미노산 서열에서의 Pro256과 위치적으로 및/또는 기능적으로 동등한 잔기가 강조표시되어 있으며, 삼각형 기호에 의해 표기되어 있다.

도 3a는 T4 UvsX 제형을 사용하여 패턴화된 플로우 셀 상으로 씨딩된 PCR-프리 라이브러리의 클러스터 이미지의 스크린샷을 나타낸다.

도 3b는 RB49 P256K 재조합효소를 포함하는 액체 제형을 사용하여 패턴화된 플로우 셀 상으로 씨딩된 PCR-프리 라이브러리의 클러스터 이미지의 스크린샷을 나타낸다.

도 4a는 T4 UvsX 제형을 사용하여 패턴화된 플로우 셀 상으로 씨딩된 단일 가닥(ssDNA) PCR-프리 라이브러리의 클러스터 이미지의 스크린샷을 나타낸다.

도 4b는 RB49 P256K 재조합효소를 포함하는 액체 제형을 사용하여 패턴화된 플로우 셀 상으로 씨딩된 단일 가닥 PCR-프리 라이브러리의 클러스터 이미지의 스크린샷을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0025] 개선된 재조합효소-매개의 핵산의 증폭을 위한 재조합효소가 본 명세서에 제시된다. 본 발명자들은 패턴화된 플로우 셀 표면 상으로의 핵산의 씨딩에서 실질적으로 개선된 특징을 갖는 특정 변경된 재조합효소를 놀랍게도 확인하였다.

[0026] 하기에 더욱 상세히 기재된 바와 같이, 본 발명자들은 재조합효소에서의 하나 이상의 잔기에 대한 하나 이상의 돌연변이가 패턴화된 플로우 셀 표면 상에, DNA 라이브러리, 예를 들어, 단일-가닥 어댑터 영역을 갖는 PCR-라이브러리를 씨딩하는데 뛰어난 개선을 초래하여, 개선된 클러스터 증폭을 제공하는 것을 놀랍게도 관찰하였다.

[0027] 특정 실시형태에 있어서, 치환 돌연변이는 하전된 측쇄를 갖는 잔기로의 돌연변이를 포함한다. 예를 들어, 일부 실시형태에 있어서, 하전된 아미노산은 양으로 하전된 아미노산 잔기이다. 용어 "양으로 하전된 아미노산"은 7 초과의 측쇄 pKa 값을 갖는 친수성 아미노산, 즉, 염기성 아미노산을 말한다. 염기성 아미노산은 전형적으로 하이드로늄 이온과의 회합으로 인하여 생리학적 pH에서 양으로 하전된 측쇄를 갖는다. 천연 발생(유전자 인코딩된) 염기성 아미노산은 라이신(Lys, K), 아르기닌(Arg, R) 및 히스티딘(His, H)을 포함하는 한편, 비천연(비-유전자 인코딩 또는 비-표준) 염기성 아미노산은 예를 들어, 오르니틴, 2,3-다이아미노프로피온산, 2,4-다이아미노부티르산, 2,5,6-트라이아미노헥산산, 2-아미노-4-구아니디노부탄산 및 호모아르기닌을 포함한다. 용어 "음으로 하전된 아미노산"은 키랄성과 관계 없이, C-말단 카복실기에 더하여, 적어도 하나의 추가의 음으로 하전된 기, 예를 들어, 카복실, 포스페이트, 포스포네이트, 설포네이트 등을 함유하는 천연 또는 비천연 아미노산을 지칭한다.

[0028] 또한, 재조합 UvsX의 반-보존된 도메인에 대한 치환 돌연변이를 포함하는 재조합 UvsX가 본 명세서에 제시된다. 본 명세서에 사용되는 바와 같이, 용어 "반-보존된 도메인"은 다양한 종 간에 완전히 보존되거나, 적어도 부분적으로 보존된 재조합 UvsX의 부분을 지칭한다. 반-보존된 도메인 내의 하나 이상의 잔기의 돌연변이가 특히 단일-가닥 주형 핵산의 존재 하에 재조합효소 활성에 영향을 미쳐, 재조합효소-매개의 증폭 반응에서 씨딩 및/또는 증폭의 증진을 초래하는 것이 놀랍게도 관찰되었다. 이들 돌연변이된 재조합효소는 하기 실시예 부문에 기재된 바와 같이, 패턴화된 플로우 셀 표면 상의 PCR-프리 라이브러리, 예를 들어, 단일-가닥 어댑터 영역을 갖는 PCR-라이브러리의 씨딩에서 개선된 성능을 가져, 개선된 클러스터 증폭을 초래한다.

[0029] 일부 실시형태에 있어서, 반-보존된 도메인은 서열번호 3 내지 7 중 임의의 것에 기재된 서열을 갖는 아미노산을 포함한다. 서열번호 3 내지 7은 다양한 종 간에 반-보존된 도메인 내의 잔기에 상응한다. 서열번호 3은 본 명세서에 서열번호 8로 기재된 T4 UvsX 아미노산 서열의 잔기 251 내지 258에 상응한다. 반-보존된 도메인에서 다양한 종 간의 보존을 보여주는 정렬은 도 1 및 도 2에 나타나 있다. 도 1에 나타낸 UvsX 서열을 젠뱅크(Genbank) 데이터베이스 수탁 번호 NP_049656(T4), YP_004300647(파지 133); NP_861734(RB69); NP_943894.1(Aeh1); YP_004300858(Ae65); NP_899256(KVP40); YP_239013(RB43); YP_214417(P-SSM2); YP_214708(P-SSM4); 및 미국 특허 출원 공개 제2009/0029421호(T6)로부터 수득하였다. 도 2는 엔테로박테리아 파지 RB49(RB49)(서열번호 1) 및 엔테로박테리아 파지 T4(T4)(서열번호 8)로부터의 UvsX 아미노산 서열의 정렬을 보여주는 개략도이다. RB49 UvsX 아미노산 서열에서의 Pro256과 위치적으로 및/또는 기능적으로 동등한 잔기는 강조표시되어 있으며, 삼각형으로 나타나 있다. 도 2에 나타낸 UvsX 서열을 젠뱅크 데이터베이스 수탁 번호 NP_891595(RB49) 및 NP_049656(T4)으로부터 수득하였다.

[0030] 반-보존된 도메인 내의 하나 이상의 잔기에 대한 돌연변이는 특히 단일-가닥 주형 핵산의 존재 하에 재조합효소 활성을 증가시켜, 재조합효소-매개의 증폭 반응에서 씨딩 및/또는 증폭의 증진을 초래하는 것이 놀랍게도 관찰되었다. 이들 돌연변이된 재조합효소는 하기 실시예 부문에 기재된 바와 같이 패턴화된 플로우 셀 표면 상의 PCR-프리 라이브러리, 예를 들어, 단일-가닥 어댑터 영역을 갖는 PCR-라이브러리의 씨딩에서 개선된 성능을 가져, 개선된 클러스터 증폭을 초래한다. 예를 들어, 본 명세서에 제시된 재조합 UvsX의 일부 실시형태에 있어서, 치환 돌연변이는 서열번호 3 내지 5 중 임의의 것의 위치 7에서 Phe, Pro, Asp, Glu 또는 Asn 이외의 임의의 잔기로의 돌연변이를 포함한다. 특정 실시형태에 있어서, 재조합 UvsX는 서열번호 3 내지 5 중 임의의 것의 위치 7에서 Lys으로의 돌연변이를 포함한다. 본 명세서에 제시된 재조합 UvsX의 일부 실시형태에 있어서, 치환 돌연변이는 서열번호 6 내지 7 중 임의의 것의 위치 12에서의 Phe, Pro, Asp, Glu 또는 Asn 이외의 임의의 잔기로의 돌연변이를 포함한다. 특정 실시형태에 있어서, 재조합 UvsX는 서열번호 6 내지 7 중 임의의 것의 위치 12에서의 Lys으로의 돌연변이를 포함한다.

[0031] 일부 실시형태에 있어서, 재조합효소는 UvsX 단백질이다. 예를 들어, 파지 재조합효소, 예를 들어, 미오비리대 파지, 예를 들어, T4, T6, Rb69, Aeh1, KVP40, 아시네토박터 파지 133, 아에로모나스 파지 65, 시아노파지 P-SSM2, 시아노파지 PSSM4, 시아노파지 S-PM2, Rb32, 비브리오 파지 nt-1, Rb16, Rb43 및 Rb49로부터 유래된 UvsX 또는 UvsX-유사 재조합효소를 포함하는 임의의 파지 재조합효소는 본 명세서에 제시된 실시형태에 사용될 수 있다. 특정 실시형태에 있어서, 재조합효소는 미오비리대 파지, 예를 들어, T2, Rb14, 아에로모나스 파지 25, phi-1, 파지 31, 파지 44RR2.8t, 파지 Rb3 및 파지 LZ2로부터 유래된 UvsX 또는 UvsX-유사 재조합효소이다. 다른 재조합효소 단백질이 본 명세서에 제시된 실시형태에서 사용될 수 있는 것이 해당 분야의 숙련자에게 용이하게 명백해질 것이다. 적합한 재조합효소 단백질은 하기에 더욱 상세히 기재된 바와 같이 해당 분야에 알려져 있는 수많은 방법, 예를 들어, BLAST 정렬을 사용하여 UvsX에 대한 상동성에 의해 확인될 수 있다.

[0032] "기능적으로 동등한"은 완전히 상이한 재조합효소를 사용하는 연구의 경우에, 대조군 재조합효소가 효소에서 동일한 기능적 역할을 갖는 다른 재조합효소에서의 아미노산 위치에서 발생하는 것으로 여겨지는 아미노산 치환을 함유할 것을 의미한다. 일 예로서, T4 UvsX 내의 위치 257에서, 페닐알라닌에서 라이신으로의 돌연변이(F257K)는 RB49 UvsX 내의 위치 256에서, 프롤린에서 라이신으로의 치환(P256K)과 기능적으로 동등할 것이다.

[0033] 일반적으로 2개 이상의 상이한 재조합효소에서의 기능적으로 동등한 치환 돌연변이는 재조합효소의 아미노산 서열에서의 상동성 아미노산 위치에서 발생한다. 이런 이유로, 본 명세서에서 용어 "기능적으로 동등한"의 이용은 또한, 돌연변이된 아미노산의 특정 기능이 알려져 있는지의 여부와 관계 없이, 주어진 돌연변이와 "위치적으로 동등"하거나, "상동성"인 돌연변이를 포함한다. 서열 정렬 및/또는 분자 모델링에 기초하여, 둘 이상의 상이한 재조합효소의 아미노산 서열에서 위치적으로 동등한 또는 상동성 아미노산 잔기를 확인할 수 있다. 위치적으로 동등한 및/또는 기능적으로 동등한 잔기를 확인하기 위한 서열 정렬의 일 예는 도 1에 제시되어 있으며, 이는 US 2009/0029421호의 혼입 물질에도 기재된 바와 같은 엔테로박테리아 파지 T4(T4)(서열번호 8), 엔테로박테리아 파지 T6(T6)(서열번호 9), 아시네토박터 파지 133(Phage133)(서열번호 10), 엔테로박테리아 파지 RB69(Rb69)(서열번호 11), 아에로모나스 파지 Aeh1(Aeh1)(서열번호 12), 아에로모나스 파지 65(Ae65)(서열번호 13), 비브리오 파지 KVP40(Kvp40)(서열번호 14), 엔테로박테리아 파지 RB43(Rb43)(서열번호 15), 프로클로로코커스 파지 P-SSM2(PSSM2)(서열번호 16) 및 프로클로로코커스 파지 P-SSM4(PSSM4)(서열번호 17)로부터의 UvsX 아미노산 서열의 정렬을 제시한다. 도 1에 나타낸 UvsX 서열을 젠뱅크 데이터베이스 수탁 번호 NP_049656(T4), YP_004300647(파지 133); NP_861734(RB69); NP_943894.1(Aeh1); YP_004300858(Ae65); NP_899256(KVP40);

YP_239013(RB43); YP_214417(P-SSM2); YP_214708(P-SSM4); 및 미국 공개 제2009/0029421호(T6)로부터 수득하였다.

[0034] 도 2는 엔테로박테리아 파지 RB49(RB49)(서열번호 1) 및 엔테로박테리아 파지 T4(T4)(서열번호 8)로부터의 UvsX 아미노산 서열의 정렬을 보여주는 개략도이다. RB49 UvsX 아미노산 서열에서의 Pro256과 위치적으로 및/또는 기능적으로 동등한 잔기가 강조표시되어 있으며, 삼각형으로 표기되어 있다. 도 2에 나타낸 UvsX 서열을 젠뱅크 데이터베이스 수탁 번호 NP_891595(RB49) 및 NP_049656(T4)으로부터 수득하였다.

[0035] 위치적으로 동등한 및/또는 기능적으로 동등한 잔기는 그들 서열을 참조 서열, 예를 들어, T4 및 RB49의 서열과 비교함으로써 많은 다른 UvsX 서열 중 하나 이상에 대하여 결정될 수 있다. 비제한적인 예로서, 시네코코커스 파지 S-PM2, 엔테로박테리아 파지 RB32, 비브리오 파지 nt-1, 엔테로박테리아 파지 RB16으로부터의 UvsX 서열은 서열번호 18 내지 21에 기재되어 있으며, 이를 젠뱅크 데이터베이스 수탁 번호 YP_195169.1; YP_802982.1; YP_008125207.1; YP_003858336.1로부터 수득하였고, 참조 UvsX 서열, 예를 들어, T4 UvsX(서열번호 8) 및 RB49 UvsX(서열번호 1)와 정렬될 수 있으며, 위치적으로 동등한 및/또는 기능적으로 동등한 잔기가 확인된다. 예를 들어, 하기 표 1에 나타낸 잔기는 RB49 UvsX 아미노산 서열에서의 Pro256과 위치적으로 동등하고/거나 기능적으로 동등한 것으로 확인된다. 다른 UvsX 단백질에 대한 위치적으로 동등한 및/또는 기능적으로 동등한 위치가 유사한 방법에 따라 확인될 수 있는 것이 당업계의 숙련자에 의해 용이하게 인식될 것이다.

표 1

[0036]

파지 종	서열번호	위치적으로/기능적으로 동등한 위치
T4	8	Phe257
T6	9	Phe259
아시네토박터 파지 133	10	Pro257
Rb69	11	Pro258
Aeh1	12	Pro269
아에로모나스 파지 65	13	Asp266
KVP40	14	Pro267
Rb43	15	Pro259
시아노파지 P-SSM2	16	Gln261
시아노파지 PSSM4	17	Glu264
시아노파지 S-PM2	18	Glu264
Rb32	19	Phe259
비브리오 파지 nt-1	20	Pro267
Rb16	21	Pro259

[0037] 상기 기재된 재조합 UvsX 단백질은 재조합효소 활성, 안정성 또는 임의의 다른 바람직한 특성 중 하나 이상의 양상을 증진시키는 것으로 알려져 있는 추가의 치환 돌연변이를 포함할 수 있다. 예를 들어, 일부 실시형태에 있어서, 재조합 UvsX는 상기 돌연변이 중 임의의 것에 더하여, 해당 분야에 알려져 있으며, 전문이 참조로 포함되는 US 2009/0029421호의 명세서에 의해 예시되는 바와 같은 RB49 UvsX 아미노산 서열에서의 His63과 기능적으로 동등한 위치에 치환 돌연변이를 추가로 포함할 수 있다. 예를 들어, 특정 실시형태에 있어서, 재조합 UvsX는 RB49 UvsX 아미노산 서열에서의 His63Ser과 상동성인 치환 돌연변이를 포함한다.

[0038] 일부 실시형태에 있어서, 재조합 UvsX는 상기 돌연변이 중 임의의 것에 더하여, 야생형 재조합효소에 비하여, 추가의 치환, 결실 및/또는 부가 돌연변이를 포함할 수 있다. 하나 이상의 위치에서의 다양한 치환 돌연변이 중 임의의 것은 해당 분야에 알려져 있으며, 2009/0029421호의 혼입 물질에 의해 예시된 바와 같이 이루어질 수 있다. 예를 들어, 일부 실시형태에 있어서, 재조합 UvsX는 상기 돌연변이에 더하여, C-말단에서의 하나 이상의 글루탐산 잔기의 부가; C-말단에서의 하나 이상의 아스파르트산 잔기의 부가; 및 이들의 조합으로 이루어진 군으로부터 선택되는 돌연변이를 추가로 포함할 수 있다.

[0039] 재조합효소의 돌연변이

[0040] 다양한 유형의 돌연변이유발은 예를 들어, 재조합효소 모델 및 모델 예측에 따라, 또는 무작위 또는 준-무작위 돌연변이 방법을 사용하여, 예를 들어, 재조합효소를 변형시켜 변이체를 생성하기 위하여 본 명세서에서 임의로 사용된다. 일반적으로, 임의의 이용 가능한 돌연변이유발 절차가 재조합효소 돌연변이체의 제조를 위해 사용될 수 있다. 그러한 돌연변이유발 절차는 하나 이상의 관심 활성(예를 들어, 고체 지지체 상의 증진된 씨딩 및/또

는 증폭)을 위한 돌연변이 핵산 및 폴리펩타이드의 선택을 임의로 포함한다. 사용될 수 있는 절차는 위치-지정 점 돌연변이유발, 무작위 점 돌연변이유발, 시험관내 또는 생체내 상동성 재조합(DNA 셔플링 및 조합 중첩 PCR), 우라실 함유 주형을 사용한 돌연변이유발, 올리고뉴클레오타이드-지정 돌연변이유발, 포스포로티오에이트-변형 DNA 돌연변이유발, 갭이 있는 듀플렉스 DNA를 사용한 돌연변이유발, 점 미스매치 수복, 수복-결함 숙주 균주를 사용한 돌연변이유발, 제한-선택 및 제한-정제, 결실 돌연변이유발, 전체 유전자 합성에 의한 돌연변이유발, 축퇴성 PCR, 이중-가닥 파단부 수복 및 숙련자에게 알려져 있는 많은 다른 절차를 포함하나 이들에 한정되지 않는다. 돌연변이를 위한 출발 재조합효소는 예를 들어, 전문이 참조로 포함되는 US 2009/0029421호에 확인된 것들과 같은 이용 가능한 재조합효소 돌연변이체를 포함하는 본 명세서에 언급된 것들 중 임의의 것일 수 있다.

[0041] 임의로, 돌연변이유발은 천연 발생 재조합효소 분자로부터의 또는 알려져 있는 변경되거나 돌연변이된 재조합효소(예를 들어, 전술된 참조문헌에 열거된 바와 같은 기존의 돌연변이 재조합효소 사용)의 알려져 있는 정보, 예를 들어, 상기 논의된 바와 같은 서열, 서열 비교, 물리적 특성, 결정 구조 등에 의해 안내될 수 있다. 그러나, 다른 부류의 실시형태에 있어서, 변형은 본질적으로 무작위일 수 있다(예를 들어, 통상의 또는 "패밀리" DNA 셔플링에서와 같이, 예를 들어, 문헌[Crameri et al. (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution" Nature 391:288-291] 참조).

[0042] 돌연변이 형식에 대한 추가의 정보는 문헌[Sambrook et al., MolecularCloning--A Laboratory Manual (3rd Ed.), Vol. 1-3, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 2000 ("Sambrook")]; 문헌[Current Protocols in Molecular Biology, F. M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (supplemented through 2011) ("Ausubel"))] 및 문헌[PCR Protocols A Guide to Methods and Applications (Innis et al. eds) Academic Press Inc. San Diego, Calif. (1990) ("Innis")]에서 관찰된다. 하기의 공개문헌 및 그 안에 열거된 참고문헌은 돌연변이 형식에 대한 추가의 상세사항을 제공한다: 문헌[Arnold, Protein engineering for unusual environments, Current Opinion in Biotechnology 4:450-455 (1993)]; 문헌[Bass et al., Mutant Trp repressors with new DNA-binding specificities, Science 242:240-245 (1988)]; 문헌[Bordo and Argos (1991) Suggestions for "Safe" Residue Substitutions in Site-directed Mutagenesis 217:721-729]; 문헌[Botstein & Shortle, Strategies and applications of in vitro mutagenesis, Science 229:1193-1201 (1985)]; 문헌[Carter et al., Improved oligonucleotide site-directed mutagenesis using M13 vectors, Nucl. Acids Res. 13: 4431-4443 (1985)]; 문헌[Carter, Site-directed mutagenesis, Biochem. J. 237:1-7 (1986)]; 문헌[Carter, Improved oligonucleotide-directed mutagenesis using M13 vectors, Methods in Enzymol. 154: 382-403 (1987)]; 문헌[Dale et al., Oligonucleotide-directed random mutagenesis using the phosphorothioate method, Methods Mol. Biol. 57:369-374 (1996)]; 문헌[Eghtedarzadeh & Henikoff, Use of oligonucleotides to generate large deletions, Nucl. Acids Res. 14: 5115 (1986)]; 문헌[Fritz et al., Oligonucleotide-directed construction of mutations: a gapped duplex DNA procedure without enzymatic reactions in vitro, Nucl. Acids Res. 16: 6987-6999 (1988)]; 문헌[Grundstrom et al., Oligonucleotide-directed mutagenesis by microscale 'shot-gun' gene synthesis, Nucl. Acids Res. 13: 3305-3316 (1985)]; 문헌[Hayes (2002) Combining Computational and Experimental Screening for rapid Optimization of Protein Properties PNAS 99(25) 15926-15931]; 문헌[Kunkel, The efficiency of oligonucleotide directed mutagenesis, in Nucleic Acids & Molecular Biology (Eckstein, F. and Lilley, D. M. J. eds., Springer Verlag, Berlin)) (1987)]; 문헌[Kunkel, Rapid and efficient site-specific mutagenesis without phenotypic selection, Proc. Natl. Acad. Sci. USA 82:488-492 (1985)]; 문헌[Kunkel et al., Rapid and efficient site-specific mutagenesis without phenotypic selection, Methods in Enzymol. 154, 367-382 (1987)]; 문헌[Kramer et al., The gapped duplex DNA approach to oligonucleotide-directed mutation construction, Nucl. Acids Res. 12: 9441-9456 (1984)]; 문헌[Kramer & Fritz Oligonucleotide-directed construction of mutations via gapped duplex DNA, Methods in Enzymol. 154:350-367 (1987)]; 문헌[Kramer et al., Point Mismatch Repair, Cell 38:879-887 (1984)]; 문헌[Kramer et al., Improved enzymatic in vitro reactions in the gapped duplex DNA approach to oligonucleotide-directed construction of mutations, Nucl. Acids Res. 16: 7207 (1988)]; 문헌[Ling et al., Approaches to DNA mutagenesis: an overview, Anal Biochem. 254(2): 157-178 (1997)]; 문헌[Lorimer and Pastan Nucleic Acids Res. 23, 3067-8 (1995)]; 문헌[Mandecki, Oligonucleotide-directed double-strand break repair in plasmids of Escherichia coli: a method for site-specific mutagenesis, Proc. Natl. Acad. Sci. USA, 83:7177-7181(1986)]; 문헌

[Nakamaye & Eckstein, Inhibition of restriction endonuclease Nci I cleavage by phosphorothioate groups and its application to oligonucleotide-directed mutagenesis, Nucl. Acids Res. 14: 9679-9698 (1986)]; 문헌[Nambiar et al., Total synthesis and cloning of a gene coding for the ribonuclease S protein, Science 223: 1299-1301(1984)]; 문헌[Sakamar and Khorana, Total synthesis and expression of a gene for the a-subunit of bovine rod outer segment guanine nucleotide-binding protein (transducin), Nucl. Acids Res. 14: 6361-6372 (1988)]; 문헌[Sayers et al., Y-T Exonucleases in phosphorothioate-based oligonucleotide-directed mutagenesis, Nucl. Acids Res. 16:791-802 (1988)]; 문헌[Sayers et al., Strand specific cleavage of phosphorothioate-containing DNA by reaction with restriction endonucleases in the presence of ethidium bromide, (1988) Nucl. Acids Res. 16: 803-814]; 문헌[Sieber, et al., Nature Biotechnology, 19:456-460 (2001)]; 문헌[Smith, In vitro mutagenesis, Ann. Rev. Genet. 19:423-462 (1985)]; 문헌[Methods in Enzymol. 100: 468-500 (1983)]; 문헌[Methods in Enzymol. 154: 329-350 (1987)]; 문헌[Stemmer, Nature 370, 389-91(1994)]; 문헌[Taylor et al., The use of phosphorothioate-modified DNA in restriction enzyme reactions to prepare nicked DNA, Nucl. Acids Res. 13: 8749-8764 (1985)]; 문헌[Taylor et al., The rapid generation of oligonucleotide-directed mutations at high frequency using phosphorothioate-modified DNA, Nucl. Acids Res. 13: 8765-8787 (1985)]; 문헌[Wells et al., Importance of hydrogen-bond formation in stabilizing the transition state of subtilisin, Phil. Trans. R. Soc. Lond. A 317: 415-423 (1986)]; 문헌[Wells et al., Cassette mutagenesis: an efficient method for generation of multiple mutations at defined sites, Gene 34:315-323 (1985)]; 문헌[Zoller & Smith, Oligonucleotide-directed mutagenesis using M 13-derived vectors: an efficient and general procedure for the production of point mutations in any DNA fragment, Nucleic Acids Res. 10:6487-6500 (1982)]; 문헌[Zoller & Smith, Oligonucleotide-directed mutagenesis of DNA fragments cloned into M13 vectors, Methods in Enzymol. 100:468-500 (1983)]; 문헌[Zoller & Smith, Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template, Methods in Enzymol. 154:329-350 (1987)]; 문헌[Clackson et al. (1991) "Making antibody fragments using phage display libraries" Nature 352:624-628]; 문헌[Gibbs et al. (2001) "Degenerate oligonucleotide gene shuffling (DOGS): a method for enhancing the frequency of recombination with family shuffling" Gene 271:13-20]; 및 문헌[Hiraga and Arnold (2003) "General method for sequence-independent site-directed chimeragenesis: J. Mol. Biol. 330:287-296]. 다수의 상기 방법에 대한 추가의 상세사항은 문헌[Methods in Enzymology Volume 154]에서 찾을 수 있으며, 이는 또한, 다양한 돌연변이유발 방법에서 일어날 수 있는 문제점을 해결하기 위해 유용한 제어 방법을 기재하고 있다.

[0043] 재조합 재조합효소의 제조 및 단리

[0044] 일반적으로, 본 명세서에 제시된 바와 같은 재조합효소를 인코딩하는 핵산은 클로닝, 재조합, 시험관내 합성, 시험관내 증폭 및/또는 다른 이용 가능한 방법에 의해 제조될 수 있다. 본 명세서에 제시된 바와 같은 재조합효소를 인코딩하는 발현 벡터를 발현하기 위하여 다양한 재조합 방법이 사용될 수 있다. 재조합 핵산의 제조, 발현 및 발현된 생성물의 단리 방법이 널리 알려져 있으며, 해당 분야에 기재되어 있다. 수많은 예시적인 돌연변이 및 돌연변이의 조합, 및 바람직한 돌연변이의 설계를 위한 전략이 본 명세서에 기재되어 있다.

[0045] 돌연변이, 재조합 및 시험관내 핵산 조작 방법(클로닝, 발현, PCR 등 포함)을 위한 추가의 유용한 참고문헌은 문헌[Berger and Kimmel, Guide to Molecular Cloning Techniques, Methods in Enzymology volume 152 Academic Press, Inc., San Diego, Calif. (Berger)]; 문헌[Kaufman et al. (2003) Handbook of Molecular and Cellular Methods in Biology and Medicine Second Edition Ceske (ed) CRC Press (Kaufman)]; 및 문헌[The Nucleic Acid Protocols Handbook Ralph Rapley (ed) (2000) Cold Spring Harbor, Humana Press Inc (Rapley)]; 문헌[Chen et al. (ed) PCR Cloning Protocols, Second Edition (Methods in Molecular Biology, volume 192) Humana Press]; 및 문헌[Viljoen et al. (2005)Molecular Diagnostic PCR Handbook Springer, ISBN 1402034032]을 포함한다.

[0046] 또한, 세포로부터의 플라스미드 또는 다른 관련 핵산의 정제를 위한 다양한 키트가 구매가능하다(예를 들어, 파마시아 바이오테크(Pharmacia Biotech)로부터의 EasyPrep.TM., FlexiPrep.TM., 스트라타진(Stratagene)으로부터의 StrataClean.TM.; 및 퀴아젠(Qiagen)으로부터의 QIAprep.TM. 참조). 임의의 단리된 및/또는 정제된 핵산은 다른 핵산을 생성하기 위하여 추가로 조작되고/거나, 세포를 트랜스펙션시키기 위해 사용되고/거나, 발현을 위해 유기체를 감염시키기 위하여 관련 벡터 내로 혼입되는 등일 수 있다. 전형적인 클로닝 벡터는 특정 표적

핵산의 발현의 조절에 유용한 전사 및 번역 종결자, 전사 및 번역 개시 서열 및 프로모터를 함유한다. 벡터는 임의로, 적어도 하나의 독립적인 종결자 서열, 진핵생물 또는 원핵생물 또는 둘 모두에서 카세트의 복제를 허용하는 서열(셔틀 벡터) 및 원핵 시스템과 진핵 시스템 둘 모두에 대한 선택 마커를 함유하는 일반 발현 카세트를 포함한다. 벡터는 원핵생물, 진핵생물 또는 둘 모두에서 복제 및 통합에 적합하다.

[0047] 예를 들어, 세포 단리 및 배양을 위한(예를 들어, 이후의 핵산 단리를 위한) 다른 유용한 참고문헌은 문헌[Freshney (1994) Culture of Animal Cells, a Manual of Basic Technique, third edition, Wiley-Liss, New York] 및 거기에 열거된 참고문헌: 문헌[Payne et al. (1992) Plant Cell and Tissue Culture in Liquid Systems John Wiley & Sons, Inc. New York, N.Y.]; 문헌[Gamborg and Phillips (eds) (1995) Plant Cell, Tissue and Organ Culture]; 문헌[Fundamental Methods Springer Lab Manual, Springer-Verlag (Berlin Heidelberg New York)] 및 문헌[Atlas and Parks (eds) The Handbook of Microbiological Media (1993) CRC Press, Boca Raton, Fla]을 포함한다.

[0048] 본 명세서에 개시된 재조합 재조합효소를 인코딩하는 핵산은 또한, 본 명세서에 제시된 실시형태의 특징이다. 특정 아미노산은 다중 코돈에 의해 인코딩될 수 있으며, 특정 번역 시스템(예를 들어, 원핵 또는 진핵 세포)은 종종 코돈 편향을 나타내며, 예를 들어, 상이한 유기체는 종종 동일한 아미노산을 인코딩하는 몇몇 동의 코돈 중 하나를 선호한다. 이와 같이, 본 명세서에 제시된 핵산은 임의로 "코돈 최적화되며", 이는 재조합효소를 발현하기 위해 사용되는 특정 번역 시스템에 의해 선호되는 코돈을 포함하도록 핵산이 합성되는 것을 의미한다. 예를 들어, 박테리아 세포(또는 심지어 특정 박테리아 균주)에서 재조합효소를 발현하는 것이 바람직한 경우, 핵산은 재조합효소의 효율적인 발현을 위하여 박테리아 세포의 게놈에서 가장 빈번하게 관찰되는 코돈을 포함하도록 합성될 수 있다. 진핵 세포에서 재조합효소를 발현하는 것이 바람직한 경우, 유사한 전략이 사용될 수 있으며, 예를 들어, 핵산은 진핵 세포가 선호하는 코돈을 포함할 수 있다.

[0049] 다양한 단백질 단리 및 검출 방법이 알려져 있으며, 이를 사용하여, 예를 들어, 본 명세서에 제시된 재조합 재조합효소를 발현하는 세포의 재조합 배양물로부터 재조합효소를 단리할 수 있다. 예를 들어, 문헌[R. Scopes, Protein Purification, Springer-Verlag, N.Y. (1982)]; 문헌[Deutscher, Methods in Enzymology Vol. 182: Guide to Protein Purification, Academic Press, Inc. N.Y. (1990)]; 문헌[Sandana (1997) Bioseparation of Proteins, Academic Press, Inc.]; 문헌[Bollag et al. (1996) Protein Methods, 2.sup.nd Edition Wiley-Liss, NY]; 문헌[Walker (1996) The Protein Protocols Handbook Humana Press, NJ], 문헌[Harris and Angal (1990) Protein Purification Applications: A Practical Approach IRL Press at Oxford, Oxford, England]; 문헌[Harris and Angal Protein Purification Methods: A Practical Approach IRL Press at Oxford, Oxford, England]; 문헌[Scopes (1993) Protein Purification: Principles and Practice 3.sup.rd Edition Springer Verlag, NY]; 문헌[Janson and Ryden (1998) Protein Purification: Principles, High Resolution Methods and Applications, Second Edition Wiley-VCH, NY]; 및 문헌[Walker (1998) Protein Protocols on CD-ROM Humana Press, NJ]; 및 거기에 열거된 참고문헌에 기재된 것들을 포함하는 다양한 단백질 단리 및 검출 방법이 해당 분야에 널리 알려져 있다. 단백질 정제 및 검출 방법에 관련된 추가의 상세사항은 문헌[Satinder Ahuja ed., Handbook of Bioseparations, Academic Press (2000)]에서 찾을 수 있다.

[0050] 이용 방법

[0051] 본 명세서에 제시된 변경된 재조합효소는 재조합효소-매개의 증폭 절차, 예를 들어, 재조합효소 중합효소 증폭(RPA) 기술에서 사용될 수 있다. 약술하면, RPA는 표적 핵산을 재조합효소 및 표적 핵산 분자에 특이적인 단일 가닥 핵산 프라이머와 접촉시킴으로써 개시될 수 있다. 그 다음, 혼성화된 프라이머는 dNTP의 존재 하에 중합효소, 예를 들어, 가닥을 치환할 수 있는 중합효소에 의해 연장되어, 이중 가닥 표적 핵산 분자 및 핵산 분자의 치환된 가닥을 생성할 수 있다. 핵산 분자의 치환된 가닥으로의 프라이머의 재조합효소-매개의 표적화 및 프라이머의 연장에 의해 추가의 증폭이 일어나서, 이중 가닥 핵산 분자를 생성할 수 있다. RPA 방법은 상기 기재된 요소와 예를 들어, 재조합효소-로딩 인자, 특정 가닥-치환 중합효소 및 강력한 에너지 재생 시스템의 조합에 의해 조절될 수 있다. 본 명세서의 재조합 UvsX 단백질에서 사용하기 위해 용이하게 조정될 수 있는 예시적인 RPA 절차, 시스템 및 요소는 예를 들어, 각각이 본 명세서에 참조로 포함되는 미국 특허 제8,071,308호; 제7,399,590호, 제7,485,428호, 제7,270,981호, 제8,030,000호, 제7,666,598호, 제7,763,427호, 제8,017,399호, 제8,062,850호 및 제7,435,561호에 기재되어 있다.

[0052] 일부 실시형태에 있어서, 등온 증폭은 배제 증폭(ExAmp)으로도 지칭되는 역학적 배제 증폭(KEA)을 이용하여 수행될 수 있다. 본 명세서의 핵산 라이브러리는 역학적 배제를 이용하는 방법을 사용하여 제조될 수 있다. 역학

적 배제는 다른 사건 또는 과정이 발생하는 것을 효과적으로 배제시키기에 충분하게 신속한 속도로 과정이 발생하는 경우에 발생할 수 있다. 어레이의 위치가 용액으로부터의 표적 핵산으로 무작위로 씨딩되고, 표적 핵산의 카피가 증폭 과정에서 생성되어 씨딩된 위치의 각각을 용량까지 충전시키는 핵산 어레이의 제조가 예시된다. 본 명세서의 역학적 배제 방법에 따라, 씨딩과 증폭 과정은 증폭 속도가 씨딩 속도를 초과하는 조건하에서 동시에 진행될 수 있다. 그와 같이, 제1 표적 핵산이 씨딩된 위치에서 카피가 형성되는 비교적 신속한 속도는 제2 핵산이 증폭을 위한 위치에 씨딩되는 것을 효과적으로 배제시킬 것이다. 역학적 배제 증폭 방법은 전문이 본 명세서에 참조로 포함되는 미국 출원 공개 제2013/0338042호의 명세서에 상세히 기재된 바와 같이 수행될 수 있다.

[0053] 일부 실시형태에 있어서, 증폭되는 표적 핵산은 완전히 이중 가닥이다. 일부 실시형태에 있어서, 증폭되는 표적 핵산은 이중 가닥 핵산의 영역을 포함하며, 단일 가닥 핵산을 갖는 영역도 또한 포함한다. 특정 실시형태에 있어서, 표적 핵산은 라이브러리 단편의 각 말단에서 약 5, 10, 15, 20, 25, 30, 35, 40개 또는 약 40개 초과의 염기의 단일 가닥 서열의 영역을 갖는 하나 이상의 분기형(forked) 어댑터를 포함한다. 분기형 어댑터의 설계 및 이용은 각각의 전문이 본 명세서에 참조로 포함되는 미국 특허 제7,742,463호 및 제8,563,748호의 명세서에 더욱 상세히 기재되어 있다.

[0054] 역학적 배제는 표적 핵산의 제1 카피를 제조하기 위해 상대적으로 느린 속도를 이용하는 데 비하여, 표적 핵산 또는 제1 카피의 후속 카피를 제조하기 위해 상대적으로 빠른 속도를 이용할 수 있다. 이전의 단락의 예에서, 역학적 배제는 표적 핵산 씨딩의 상대적으로 느린 속도(예를 들어, 상대적으로 느린 확산 또는 수송)에 비하여 상기 위치를 핵산 씨드의 카피로 충전시키기 위해 증폭이 발생하는 상대적으로 빠른 속도로 인하여 발생한다. 다른 예시적인 실시형태에 있어서, 역학적 배제는 상기 위치를 씨딩하는 표적 핵산의 제1 카피의 형성의 지연(예를 들어, 지연된 또는 느린 활성화)에 비하여, 상기 위치를 충전시키는 이후의 카피가 형성되는 상대적으로 빠른 속도로 인하여 발생할 수 있다. 이러한 예에서, 개별 위치는 몇몇의 상이한 표적 핵산으로 씨딩될 수 있다(예를 들어, 몇몇의 표적 핵산이 증폭 전에 각각의 위치에 존재할 수 있다). 그러나, 임의의 제공되는 표적 핵산에 대한 제1 카피 형성을 무작위로 활성화시켜, 제1 카피 형성의 평균 속도가 후속 카피가 생성되는 속도에 비해서 상대적으로 느리게 할 수 있다. 이러한 경우에, 비록 개별 위치가 몇몇의 상이한 표적 핵산으로 씨딩될 수 있지만, 역학적 배제는 그들 표적 핵산 중 오직 하나만이 증폭되게 할 것이다. 더욱 구체적으로, 일단 제1 표적 핵산이 증폭을 위해 활성화되면, 위치는 신속하게 그의 카피로 용량까지 충전되어, 상기 위치에서 제2 표적 핵산의 카피가 형성되지 않게 할 것이다.

[0055] 증폭 시약에는 앰플리콘 형성을 촉진시키고, 일부 경우에는 앰플리콘 형성 속도를 증가시키는 추가 성분이 포함될 수 있다. 재조합효소, 예를 들어, UvsX는 반복 침범/연장을 허용하여 앰플리콘 형성을 촉진할 수 있다. 보다 구체적으로, 재조합효소는 앰플리콘 형성을 위한 주형으로서 표적 핵산을 이용하여 중합효소에 의한 표적 핵산의 침범 및 중합효소에 의한 프라이머의 연장을 촉진할 수 있다. 이러한 과정은 각 회차의 침범/연장으로부터 생성된 앰플리콘이 후속 회차에서 주형으로 기능하는 연쇄 반응으로 반복될 수 있다. 이 과정은 변성 사이클(예를 들어, 가열 또는 화학적 변성을 통해)이 필요하지 않으므로 표준 PCR보다 더 빠르게 일어날 수 있다. 그와 같이, 재조합효소-촉진 증폭은 등온으로 수행될 수 있다. 일반적으로 증폭을 촉진하기 위해 재조합효소-촉진 증폭 시약에 ATP, 또는 다른 뉴클레오타이드(또는 일부 경우에는 그의 비-가수분해가능 유사체)가 포함되는 것이 바람직하다. 재조합효소 및 단일 가닥 결합(SSB) 단백질의 혼합물은, SSB가 증폭을 추가로 촉진할 수 있으므로, 특히 유용하다. 재조합효소-촉진 증폭을 위한 예시적인 제형에는 TwistDx(Cambridge, UK)에서 TwistAmp 키트로 판매되는 것들이 포함된다. 재조합효소-촉진 증폭 시약의 유용한 성분 및 반응 조건은 각각이 본 명세서에 참조로 포함되는 US 5,223,414호 및 US 7,399,590호에 나타나 있다.

[0056] 서열 비교, 동일성 및 상동성

[0057] 2개 이상의 핵산 또는 폴리펩타이드 서열과 관련하여 용어 "동일한" 또는 "동일성 백분율"은 하기 기술하는 서열 비교 알고리즘(또는 숙련자가 이용 가능한 다른 알고리즘) 중 하나를 사용하여 측정하거나, 시각적 관찰에 의하여 측정시, 최대로 유사하도록 비교 및 정렬된 경우, 동일한 2개 이상의 서열 또는 하위서열, 또는 동일한 아미노산 잔기 또는 뉴클레오타이드를 특정 백분율로 갖는 2개 이상의 서열 또는 하위서열을 지칭한다.

[0058] 2개의 핵산 또는 폴리펩타이드(예를 들어, 재조합효소를 인코딩하는 DNA 또는 재조합효소의 아미노산 서열)와 관련하여 "실질적으로 동일한"이라는 어구는 서열 비교 알고리즘을 사용하여 측정하거나, 시각적 관찰에 의하여 측정시, 최대로 유사하도록 비교 및 정렬된 경우, 적어도 약 60%, 약 80%, 약 90 내지 95%, 약 98%, 약 99% 이상의 뉴클레오타이드 또는 아미노산 잔기 동일성을 갖는 2개 이상의 서열 또는 하위서열을 지칭한다. 그러한 "실질적으로 동일한" 서열은 전형적으로 실제 유래와는 상관없이 "상동성"인 것으로 간주된다. 바람직하게는, "

실질적인 동일성"은, 적어도 약 50개 잔기 길이인 서열의 영역에 걸쳐, 더욱 바람직하게는 적어도 약 100개 잔기의 영역에 걸쳐 존재하며, 가장 바람직하게는 서열은 비교되는 2개의 서열의 적어도 약 150개 잔기에 걸쳐, 또는 그의 전장에 걸쳐 실질적으로 동일하다.

[0059] 단백질 및/또는 단백질 서열은 그들이 공통 조상 단백질 또는 단백질 서열로부터 천연적으로 또는 인공적으로 유래되는 경우, "상동성"이다. 유사하게, 핵산 및/또는 핵산 서열은 그들이 공통 조상 핵산 또는 핵산 서열로부터 천연적으로 또는 인공적으로 유래되는 경우, 상동성이다. 상동성은 일반적으로 둘 이상의 핵산 또는 단백질(또는 그의 서열) 간의 서열 유사성으로부터 추론된다. 상동성 확립에 유용한 서열 간의 정확한 유사성 백분율은 쟁점이 되는 핵산 및 단백질에 따라 달라지지만, 상동성 확립에는 50, 100, 150개 이상의 잔기에 걸쳐 25% 정도의 낮은 서열 유사성이 통상 이용된다. 상동성 확립을 위해서는 보다 높은 수준의 서열 유사성, 예를 들어, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95% 또는 99% 이상의 서열 유사성이 또한 사용될 수 있다. 서열 유사성 백분율의 결정 방법(예를 들어, 디폴트 파라미터를 사용한 BLASTP 및 BLASTN)은 본 명세서에 기재되어 있으며, 일반적으로 이용 가능하다.

[0060] 서열 비교 및 상동성 결정을 위하여, 전형적으로 하나의 서열은 시험 서열이 비교되는 참조 서열로서의 기능을 한다. 서열 비교 알고리즘을 사용하는 경우, 시험 및 참조 서열은 컴퓨터에 입력되며, 하위서열 좌표는 필요에 따라 지정되고, 서열 알고리즘 프로그램의 파라미터가 지정된다. 이어서, 서열 비교 알고리즘은 참조 서열에 비한 시험 서열(들)의 서열 동일성 백분율을 지정된 프로그램 파라미터에 기초하여 계산한다.

[0061] 비교를 위한 서열의 최적 정렬은 예를 들어, 문헌[Smith & Waterman, Adv. Appl. Math. 2:482 (1981)]의 국소 상동성 알고리즘에 의해, 문헌[Needleman & Wunsch, J. Mol. Biol. 48:443 (1970)]의 상동성 정렬 알고리즘에 의해, 문헌[Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988)]의 유사성 검색 방법에 의해, 이들 알고리즘(미국 위스콘신주 매디슨 사이언스 드라이브 575 제네틱스 컴퓨터 그룹(Genetics Computer Group), 위스콘신 제네틱스 소프트웨어 패키지(Wisconsin Genetics Software Package)에서 GAP, BESTFIT, FASTA 및 TFASTA)의 컴퓨터 구현예에 의해 또는 육안의 검사에 의해(일반적으로 문헌[Current Protocols in Molecular Biology, Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., supplemented through 2004] 참조) 행해질 수 있다.

[0062] 서열 동일성 및 서열 유사성 백분율을 결정하기에 적합한 알고리즘의 하나의 예는 BLAST 알고리즘이며, 이는 문헌[Altschul et al., J. Mol. Biol. 215:403-410 (1990)]에 기술되어 있다. BLAST 분석을 수행하기 위한 소프트웨어는 미국 국립생물공학정보센터(National Center for Biotechnology Information)를 통해 공개적으로 입수 가능하다. 이러한 알고리즘은 먼저 데이터베이스 서열 내의 동일한 길이의 워드와 함께 정렬되는 경우 일부 양의 값 역치 점수 T를 만족시키거나 일치하는 질의 서열 내의 길이 W의 짧은 워드를 확인함으로써 높은 점수 서열 쌍(HSP: high scoring sequence pair)을 확인하는 것을 포함한다. T는 이웃 워드 점수 역치로서 지칭된다(상기 문헌[Altschul et al.]). 이들 초기 이웃 워드 힛트(word hit)는 그들을 함유하는 더 긴 HSP를 찾기 위해 검색을 시작하기 위한 근원으로서 기능한다. 이어서, 워드 힛트는 누적 정렬 점수가 증가될 수 있는 한, 각 서열을 따라 양 방향으로 연장된다. 누적 점수는 뉴클레오타이드 서열에 대해, 파라미터 M(일치하는 잔기의 쌍에 대한 보상 점수; 항상 > 0) 및 N(불일치 잔기에 대한 페널티 점수; 항상 < 0)을 사용하여 계산한다. 아미노산 서열에 있어서, 점수화 매트릭스를 사용하여 누적 점수를 계산한다. 누적 정렬 점수가 그의 최대 달성 값으로부터 X의 양만큼 하락하거나; 하나 이상의 음으로 점수화된 잔기 정렬의 축적으로 인해 누적 점수가 0 이하로 떨어지거나; 어느 한쪽의 서열의 말단에 도달한 경우, 각 방향에서의 워드 힛트의 연장이 중단된다. BLAST 알고리즘 파라미터 W, T 및 X는 정렬의 감도 및 속도를 결정한다. BLASTN 프로그램(뉴클레오타이드 서열의 경우)은 디폴트로서 워드 길이(W) 11, 기대값(E) 10, 컷오프 100, M=5, N=-4 및 양쪽 가닥의 비교를 사용한다. 아미노산 서열의 경우, BLASTP 프로그램은 디폴트로서 워드 길이(W) 3, 기대값(E) 10, 및 BLOSUM62 점수화 매트릭스(문헌[Henikoff & Henikoff (1989) Proc. Natl. Acad. Sci. USA 89:10915] 참조)를 사용한다.

[0063] BLAST 알고리즘은 또한 서열 동일성 백분율을 계산하는데 더하여, 두 서열 간의 유사성의 통계적 분석을 수행한다(예를 들어, 문헌[Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 90:5873-5787 (1993)] 참조). BLAST 알고리즘에 의해 제공되는 유사성의 한 척도는 두 뉴클레오타이드 또는 아미노산 서열 간의 일치가 우연히 발생할 확률의 지표를 제공하는 최소 합계 확률(P(N))이다. 예를 들어, 핵산은 참조 핵산에 대한 시험 핵산의 비교에서 최소 합계 확률이 약 0.1 미만, 더욱 바람직하게는 약 0.01 미만, 가장 바람직하게는 약 0.001 미만인 경우에 참조 서열과 유사한 것으로 간주된다.

[0064] 변경된 재조합효소를 인코딩하는 핵산

[0065] 본 명세서에 제시된 변경된 재조합효소를 인코딩하는 핵산 분자가 본 명세서에 추가로 제시된다. 아미노산 서열 및 바람직하게는 재조합효소를 인코딩하는 야생형 뉴클레오타이드 서열이 알려져 있는 돌연변이 형태의 재조합효소인 임의의 주어진 변경된 재조합효소에 있어서, 분자 생물학의 기본 원리에 따라 돌연변이체를 인코딩하는 뉴클레오타이드 서열을 수득할 수 있다. 예를 들어, RB49 UvsX 재조합효소를 인코딩하는 야생형 뉴클레오타이드 서열이 알려져 있는 것을 고려해볼 때, 표준 유전 코드를 사용하여 하나 이상의 아미노산 치환을 갖는 임의의 제공되는 돌연변이 형태의 RB49 UvsX를 인코딩하는 뉴클레오타이드 서열을 추론할 수 있다. 유사하게, 뉴클레오타이드 서열은 돌연변이 형태의 다른 재조합효소, 예를 들어, T4, T6, Rb69, Aeh1, KVP40, 아시네토박터 파지 133, 아에로모나스 파지 65, 시아노파지 P-SSM2, 시아노파지 PSSM4, 시아노파지 S-PM2, Rb32, 비브리오 파지 nt-1, Rb16, Rb43, T2, Rb14, 아에로모나스 파지 25, phi-1, 파지 31, 파지 44RR2.8t, 파지 Rb3 및 파지 LZ2 등에 대하여 용이하게 유래될 수 있다. 이어서, 요구되는 뉴클레오타이드 서열을 갖는 핵산 분자는 해당 분야에 알려져 있는 표준 분자 생물학 기술을 사용하여 작제될 수 있다.

[0066] 본 명세서에 제시된 실시형태에 따라, 정의된 핵산은 동일한 핵산뿐 아니라 특히 보존적 아미노산 치환에서 축퇴성 코드로 인하여 동의 코돈(동일한 아미노산 잔기를 특정하는 상이한 코돈)을 초래하는 경우의 치환을 포함하는 임의의 미소 염기 변이도 포함한다. 또한, 용어 "핵산 서열"은 염기 변이에 관하여 제공되는 임의의 단일 가닥 서열에 상보적인 서열을 포함한다.

[0067] 본 명세서에 기재된 핵산 분자는 또한, 유리하게는 적합한 발현 벡터에 포함되어, 적합한 숙주에서 그로부터 인코딩되는 재조합 단백질을 발현시킬 수 있다. 이후의 상기 세포의 형질전환 및 이후의 형질전환된 세포의 선택을 위한 적합한 발현 벡터로의 클로닝된 DNA의 혼입은 그의 전문이 참조로 포함되는 문헌[Sambrook et al. (1989), Molecular cloning: A Laboratory Manual, Cold Spring Harbor Laboratory]에 제공되는 바와 같이 해당 분야의 숙련자에게 널리 알려져 있다.

[0068] 그러한 발현 벡터는 상기 DNA 단편의 발현을 가져올 수 있는 조절 서열, 예를 들어, 프로모터 영역에 작동가능하게 연결된 본 명세서에 제시된 실시형태에 따른 핵산을 갖는 벡터를 포함한다. 용어 "작동가능하게 연결된"은 기재된 성분이 그들의 의도된 방식으로 그들이 기능할 수 있게 하는 관계로 존재하는 배치를 의미한다. 그러한 벡터를 적합한 숙주 세포 내로 형질전환시켜, 본 명세서에 제시된 실시형태에 따른 단백질의 발현을 제공할 수 있다.

[0069] 핵산 분자는 성숙 단백질, 또는 숙주 세포에 의해 절단되어 성숙 단백질을 형성하는 프레단백질(preprotein) 상의 리더 서열을 인코딩하는 것을 포함하는 프로서열(prosequence)을 갖는 단백질을 인코딩할 수 있다. 벡터는 예를 들어, 복제 원점, 및 임의로 상기 뉴클레오타이드의 발현을 위한 프로모터 및 임의로 프로모터의 조절자와 함께 제공되는 플라스미드, 바이러스 또는 파지 벡터일 수 있다. 벡터는 하나 이상의 선택가능한 마커, 예를 들어, 항생제 내성 유전자를 함유할 수 있다.

[0070] 발현에 필요한 조절 요소는 RNA 중합효소에 결합하고 적절한 수준의 전사 개시를 유도하기 위한 프로모터 서열 및 또한 리보솜 결합을 위한 번역 개시 서열을 포함한다. 예를 들어, 박테리아 발현 벡터는 프로모터, 예를 들어, lac 프로모터 및 번역 개시를 위해 샤인-달가노(Shine-Dalgarno) 서열 및 출발 코돈 AUG를 포함할 수 있다. 유사하게, 진핵 발현 벡터는 RNA 중합효소 II를 위한 이종성 또는 상동성 프로모터, 하류 폴리아데닐화 신호, 출발 코돈 AUG 및 리보솜의 분리를 위한 종결 코돈을 포함할 수 있다. 그러한 벡터는 구매되거나, 해당 분야에 널리 알려져 있는 방법에 의해 기재된 서열로부터 조립될 수 있다.

[0071] 보다 고등의 진핵생물에 의한 재조합효소를 인코딩하는 DNA의 전사는 벡터에 인핸서 서열을 포함시킴으로써 최적화될 수 있다. 인핸서는 전사 수준을 증가시키기 위해 프로모터에서 작용하는 DNA의 시스-작용 요소이다. 또한, 벡터는 일반적으로 선택가능한 마커에 더하여 복제 원점을 포함할 것이다.

[0072] **실시예 1**

[0073] 이러한 실시예는 개선된 클러스터 증폭을 위한 패턴화된 플로우 셀 표면 상의 PCR-프리 라이브러리의 씨딩 방법을 제공한다. 일 실시형태에 있어서, 본 발명의 방법은 상기 기재된 돌연변이를 포함하는 UvsX, 예를 들어, Pro256Lys을 포함하는 RB49 UvsX 돌연변이체(본 명세서에서 "RB49 P256K"로 지칭되는, 서열번호 2로 본 명세서에 기재된 바와 같음)를 포함하는 씨딩 제형을 사용한다. 상기의 것을 사용한 재조합효소-매개의 증폭이 패턴화된 플로우 셀 표면 상으로의, 단일-가닥 어댑터 영역이 있는 PCR-라이브러리의 씨딩을 상당히 개선시키는 것이 놀랍게도 관찰되었다. 다른 실시형태에 있어서, 본 발명의 방법은 RB49 P256K 재조합효소와 함께 상대적으로 고농도의 DNA 중합효소(예를 들어, eBsu 중합효소)를 포함하는 씨딩 제형을 사용한다.

[0074] 패턴화된 플로우 셀 표면 상으로 PCR-프리 라이브러리를 씨딩함에 있어서, RB49 P256K 제형의 효능을 평가하기 위하여, PCR-프리 라이브러리를 TruSeq® DNA PCR-프리 시료 제조 키트(일루미나, 인코포레이티드(Illumina, Inc.))를 사용하여 생성하였다. TruSeq® 라이브러리 제조 키트를 사용하여 생성된 PCR-프리 라이브러리는 라이브러리 단편의 각 말단에 약 40개 염기의 단일 가닥 서열의 영역이 있는 분기형 어댑터를 갖는다.

[0075] 도 3a는 T4 UvsX 재조합효소(본 명세서에 "T4 UvsX"로 지칭되는, 서열번호 8로 본 명세서에 제시된 바와 같음)를 포함하는 표준 제형을 사용하여 패턴화된 플로우 셀상으로 씨딩된 PCR-프리 라이브러리의 클러스터 이미지의 스크린샷(100)을 보여준다. 도 3b는 RB49 P256K 재조합효소를 포함하는 액체 제형을 사용하여 패턴화된 플로우 셀상으로 씨딩된 PCR-프리 라이브러리의 클러스터 이미지의 스크린샷(150)을 보여준다. 이러한 예에서, 라이브러리를 100 pM의 최종 농도로 T4 UvsX 제형 또는 RB49 P256K 제형과 혼합하고, 플로우 셀상으로 쏟아내고, 38℃에서 cBot에서 인큐베이션시켰다. 1시간 인큐베이션 기간 후에, 온도를 20℃로 낮추고, 플로우 셀을 HT2 세척 완충제(일루미나)로 세척하였다. 클러스터를 0.1M 트리스/0.1M 아스코르브산나트륨 중 SYBR® 그린(Green)(라이프 테크놀로지즈(Life Technologies))의 1:5,000 희석액으로 염색하고, 형광 현미경에서 영상화하였다. 도 3a를 참조하여, 표준 제형(예를 들어, T4 UvsX)을 사용하여, 패턴화된 플로우 셀상으로 PCR-프리 라이브러리 씨딩함으로써 생성된 클러스터 밀도는 상대적으로 희박하다. 도 3b를 참조하여, RB49 P256K 재조합효소를 포함하는 제형을 사용하여, 패턴화된 플로우 셀상으로 PCR-프리 리아브러리를 씨딩함으로써 생성된 클러스터 밀도는 상당히 개선된다.

[0076] 도 4a는 표준 T4 UvsX 제형을 사용하여 패턴화된 플로우 셀상으로 씨딩된 단일 가닥(ssDNA) PCR-프리 라이브러리의 클러스터 이미지의 스크린샷(200)을 보여준다. 도 4b는 RB49 P256K 재조합효소를 포함하는 액체 제형을 사용하여 패턴화된 플로우 셀상으로 씨딩된 단일 가닥 PCR-프리 라이브러리의 클러스터 이미지의 스크린샷(250)을 보여준다. 이러한 예에서, 이중-가닥 PCR-프리 라이브러리를 NaOH를 사용하여 변성시키고, 이후에, 50 pM의 농도로 패턴화된 플로우 셀상으로 씨딩하였다. 도 4a를 참조하여, 표준 제형(예를 들어, T4 UvsX)을 사용하여 패턴화된 플로우 셀상으로 ssDNA, PCR-프리 DNA 라이브러리를 씨딩함으로써 생성된 클러스터 밀도는 상대적으로 희박하다. 도 4b를 참조하여, RB49 P256K 재조합효소를 포함하는 제형을 사용하여 패턴화된 플로우 셀상으로 ssDNA PCR-프리 라이브러리를 씨딩함으로써 생성된 클러스터의 밀도는 상당히 개선된다.

[0077] **실시예 2**

[0078] RB49 P256K 돌연변이체를 사용한 개선된 증폭

[0079] 이러한 실시예는 본 명세서에 기재된 P256K 돌연변이가 있거나, 이것이 없는 재조합효소 간의 증폭 성능의 비교를 설명한다. 이러한 실시예의 목적을 위하여, "대조군" RB49 UvsX(서열번호 1에 기재)는 추가로 H63S 돌연변이를 포함한다. 본 명세서에 서열번호 2로 기재된 바와 같이 위치 256에 Lys 잔기를 갖도록 대조군을 추가로 돌연변이시킴으로써 P256K 돌연변이체가 생성된다.

[0080] 패턴화된 플로우 셀상으로의 PCR-프리 라이브러리의 클러스터링을 대조군 또는 P256K 돌연변이체 중 어느 하나를 사용하여 실시예 1에서 상기 기재된 바와 같이 cBot에서 수행한다. 그 다음, 시퀀싱을 HiSeq 기기(일루미나, 인코포레이티드)에서 수행하고, 시퀀싱 결과를 분석하여, 전형적으로 이전의 시퀀싱 데이터에서 불량하게 나타난 다양한 영역의 호출가능성(callability)을 결정한다.

[0081] 호출가능성은 단일 뉴클레오타이드 다형성(SNP)이 올바르게 호출되는 위치의 분율의 측정치이다. 이상적으로, 이러한 값은 1(100%)이며, 이는 특정 유형의 영역(즉, 고 GC 등) 내의 위치의 100%에서 SNP가 올바르게 호출되는 것을 의미한다. 커버리지는 n 초과의 커버리지를 갖는 위치의 분율의 측정치이며, 여기서, n은 전형적으로 30x(즉, 인간 게놈에 대한 표준 커버리지)이다. 포스미드 프로모터는 이전의 시퀀싱 데이터에서 불량하게 나타나는 것으로 확인된 100개 유전자 프로모터의 세트이다. 프로모터를 포스미드 벡터 내로 클로닝하였다. 고 GC 영역은 GC 함량이 75% 이상인 적어도 100 bp를 갖는 영역으로 정의될 수 있다(N50 (G + C ≥ 0.75) 100 N50). 초고 GC 영역은 GC 함량이 85% 이상인 적어도 100 bp를 갖는 영역으로 정의될 수 있다(N50 (G + C ≥ 0.85) 100 N50). 저 GC 영역은 GC 함량이 40% 이하인 적어도 100 bp를 갖는 영역으로 정의될 수 있다(N50 (G + C ≥ 0.40) 100 N50). 고 AT 영역은 약 50k 영역으로 다운샘플링되는 AT 함량이 75% 이상인 적어도 100 bp를 갖는 영역으로 정의될 수 있다(N50 (A + T ≥ 0.75) 100 N50). 초고 AT 영역은 약 50k 영역으로 다운샘플링되는 AT 함량이 85% 이상인 적어도 100 bp를 갖는 영역으로 정의될 수 있다(N50 (A + T ≥ 0.85) 100 N50). AT 다이뉴클레오타이드 반복 영역은 긴 스트레치의 ATAT 반복부를 포함하는 영역으로 정의될 수 있다.

[0082] 대조군 대 P256K 돌연변이체에 대한 호출가능성 데이터의 비교에 의해, P256K 돌연변이체가 대조군의 것에 비하

여, 포스미드 프로모터 영역 중 하나 이상의 호출가능성의 예상치 않은 유의미한 개선, 고 GC, 초고 GC, 저 GC, 고 AT, 초고 AT 및 AT 다이뉴클레오타이드 반복 영역을 보이는 것이 입증된다.

[0083] **실시예 3**

[0084] P256K와 상동성인 돌연변이를 갖는 돌연변이체를 사용한 개선된 증폭

[0085] 실시예 2에서 상기 기재된 성능 비교를 다른 재조합효소에 대하여 반복한다. 이러한 실시예에서, "대조군" 재조합효소는 야생형 재조합효소를 하기 표 2에서 "대조군" 열에 기재된 바와 같이, RB49에서의 H63S와 상동성인 돌연변이를 포함하도록 변형시킴으로써 생성된다. "P256K 상동체" 돌연변이체는 대조군을 하기 표에서 "P256K 상동체" 열에 기재된 바와 같이, RB49에서의 P256K와 상동성인 돌연변이를 갖도록 추가로 변형시킴으로써 생성된다.

[0086] 예를 들어, T6 UvsX에 있어서, 대조군은 야생형 T6 UvsX(서열번호 9)를 H66S 돌연변이를 지니도록 변형시킴으로써 생성된다. P256K 상동체는 H66S 및 F259K 돌연변이 둘 모두를 지니도록 추가로 변형된다.

표 2

[0087]

야생형 백본	야생형 백본 서열번호	대조군	P256K 상동체
T4	8	64S	64S F257K
T6	9	H66S	H66S F259K
아시네토박터 파지 133	10	H64S	H64S P257K
Rb69	11	H64S	H64S P258K
Aeh1	12	H76S	H76S P269K
아에로모나스 파지 65	13	H73S	H73S D266K
KVP40	14	H64S	H64S P267K
Rb43	15	H66S	H66S P259K
시아노파지 P-SSM2	16	T62S	T62S Q261K
시아노파지 PSSM4	17	T65S	T65S E264K
시아노파지 S-PM2	18	T65S	T65S E264K
Rb32	19	H66S	H66S F259K
비브리오 파지 nt-1	20	H64S	H64S P267K
Rb16	21	H66S	H66S P259K

[0088] 패턴화된 플로우 셀 상에서의 PCR-프리 라이브러리의 클러스터링을 대조군 또는 P256K 돌연변이체를 사용하여 실시예 1에서 상기 기재된 바와 같이 cBot에서 수행한다. 그 다음, 시퀀싱을 HiSeq 기기(일루미나, 인코포레이티드)에서 수행하고, 시퀀싱 결과를 실시예 2에서 상기 기재된 바와 같이 분석하여, 전형적으로 이전의 시퀀싱 데이터에서 불량하게 나타난 다양한 영역의 호출가능성을 결정한다.

[0089] 대조군 대 P256K 상동체 돌연변이체에 대한 호출가능성 데이터의 비교에 의해, P256K 상동체 돌연변이체가 대조군의 것에 비하여, 포스미드 프로모터 영역 중 하나 이상의 호출가능성의 예상치 않은 유의미한 개선, 고 GC, 초고 GC, 저 GC, 고 AT, 초고 AT 및 AT 다이뉴클레오타이드 반복 영역을 보이는 것이 입증된다.

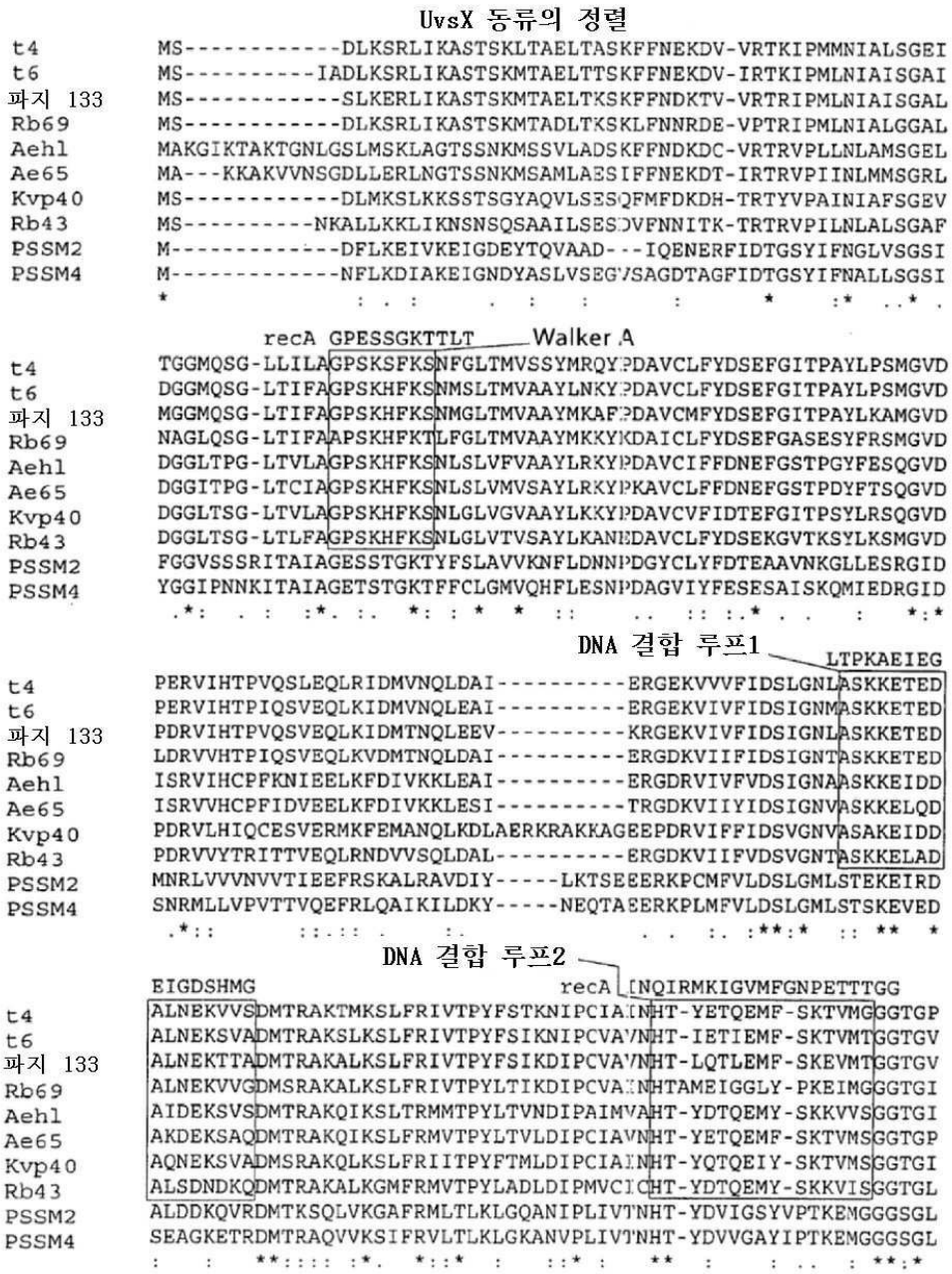
[0090] 본 출원에 걸쳐, 다양한 간행물, 특허 및/또는 특허 출원이 참조되었다. 이들 간행물의 전체의 개시내용은 본 출원에 참조로서 포함된다.

[0091] 용어 "포함하는"은 본 명세서에서 인용된 요소를 포함할 뿐 아니라, 추가의 임의의 요소를 추가로 포함하여, 제한이 없는 것으로 의도된다.

[0092] 많은 실시형태가 기재되어 있다. 그럼에도 불구하고, 다양한 변형이 이루어질 수 있는 것이 이해될 것이다. 따라서, 다른 실시형태는 하기의 청구범위의 범주 이내이다.

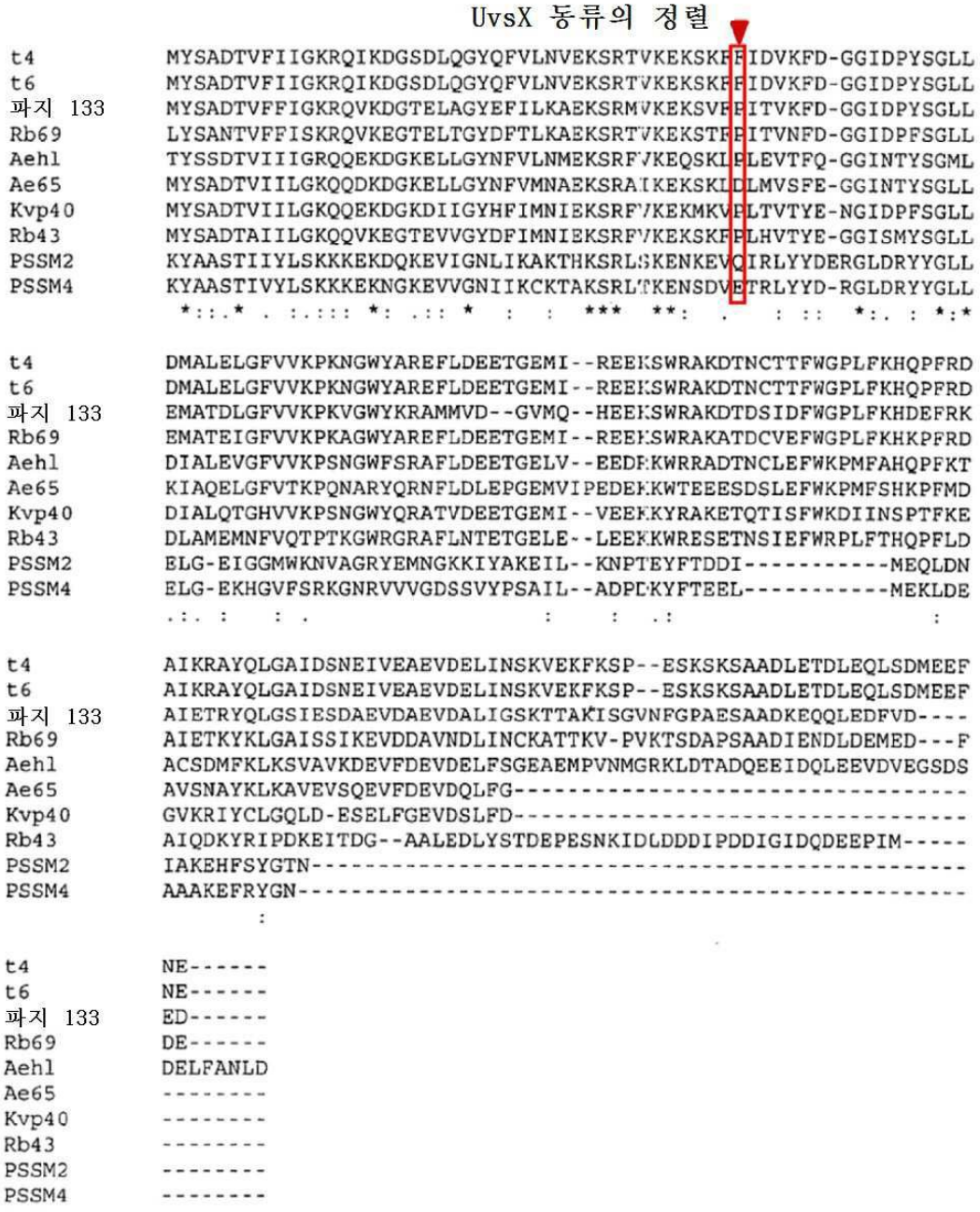
도면

도면1a



도면1b

UvsX 동류의 정렬



도면2

```
RB49    MS-VLEKLKKNSTLKTTAVLSKSSFFNEKTNTRTKIPMLNIAFSGDLKKGFQSGLIFFAGPSKHFK  65
T4      MSDLKSRLIKASTSKLTAELTASKFFNEKDVVRTKIPMMNIALSGEITGGMQSGLLILAGPSKSFK  66

RB49    SNMGLTCVSAYMKQNPDAACLFFDSEFGITSAYLESMGVDPDRVVHVPIKNIEELKFEIM  125
T4      SNFGLTMVSSYMRQYPDAVCLFYDSEFGITPAYLRSMGVDPERVIHTPVQSLEQLRIDMV  126

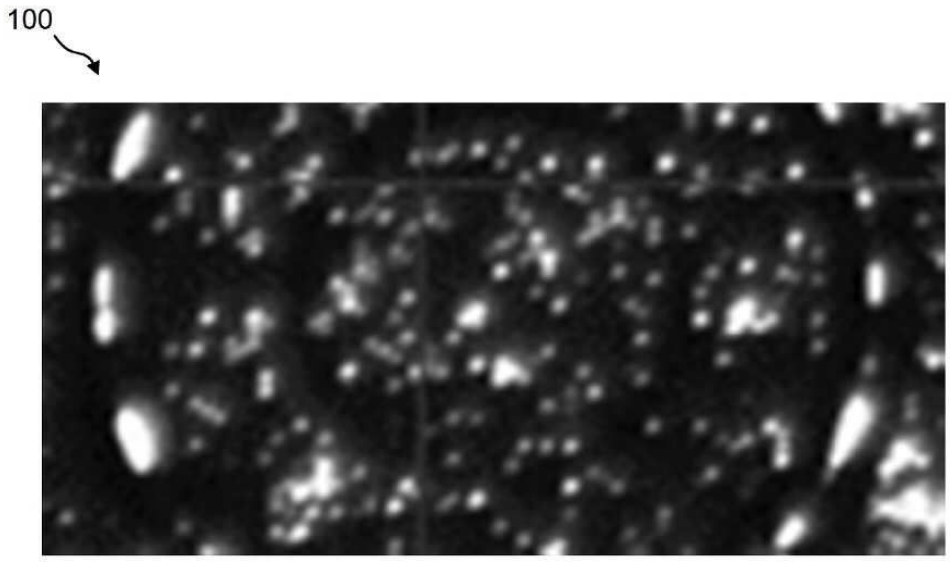
RB49    NQLEQITREDKVIIFIDSIGNLASKKEVEDAINEKSAQDMTRAKALKGLFRMVTPYLTMN  185
T4      NQLDAIERGEKVVVFIDSLGNLASKKETEDALNEKVVSDMTRAKTMKSLFRIVTPYFSTK  186

RB49    DIPCIAINHTYETQEMFSKTVMSGGTGAMYSANEVFIIGRRQQKEGTEITGYDFILNAEK  245
T4      NIPCIAINHTYETQEMFSKTVMGGGTGPMYSADTVFIIGKRQIKDGSDLQGYQFVLNVEK  246

RB49    SRTVKEKSKFPISVTFSGGIDPYSGLLELAVELGWVVKPSNGWYSRSILNTETGEMETEE  305
T4      SRTVKEKSKFFIDVKFDGGIDPYSGLLDMALELGFVVKPKNGWYAREFLDEETGEMIREE  306

RB49    RKFRAKETNSIEFWKPLLTNDKFNEAINDHYKLGQVISDEAVDKEIEDML  355
T4      KSWRAKDTNCTTFWGPLFKHQPFRDAIKRAYQLGAIDSNEIVEAEVDELI  356
```

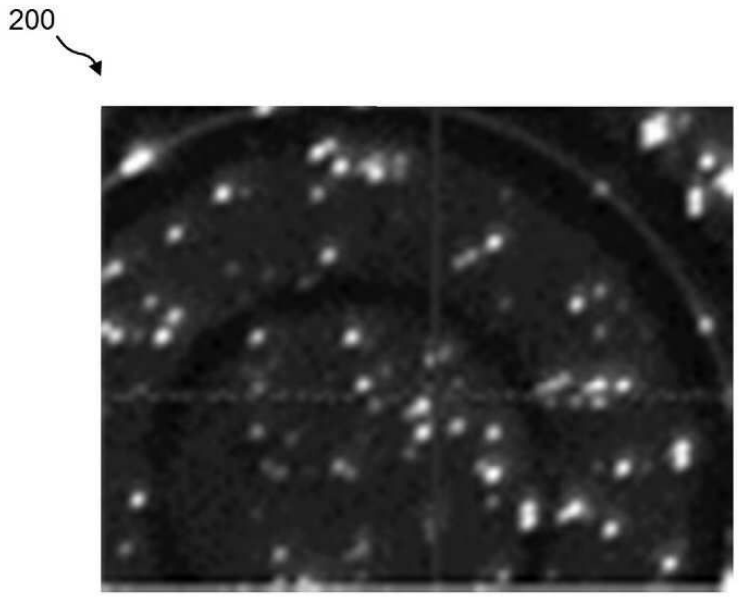
도면3a



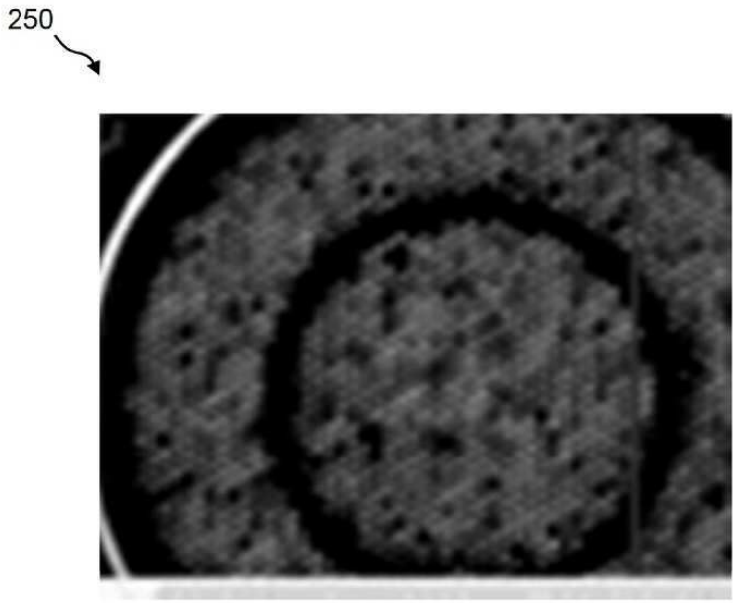
도면3b



도면4a



도면4b



서 열 목 록

SEQUENCE LISTING

<110> ILLUMINA CAMBRIDGE LIMITED

<120> RECOMBINASE MUTANTS

<130> WO 2016/054088

<140> PCT/US2015/053012

<141> 2015-09-29

<150> US 62/057,056

<151> 2014-09-29

<160> 35

<170> PatentIn version 3.5

<210> 1

<211> 356

<212> PRT

<213> Enterobacteria phage RB49

<400> 1

Met Ser Val Leu Glu Lys Leu Lys Lys Asn Ser Thr Leu Lys Thr Thr

1 5 10 15

Ala Val Leu Ser Lys Ser Ser Phe Phe Asn Glu Lys Thr Asn Thr Arg

20 25 30

Thr Lys Ile Pro Met Leu Asn Ile Ala Phe Ser Gly Asp Leu Lys Lys

35 40 45

Gly Phe Gln Ser Gly Leu Ile Phe Phe Ala Gly Pro Ser Lys His Phe

50 55 60

Lys Ser Asn Met Gly Leu Thr Cys Val Ser Ala Tyr Met Lys Gln Asn

65 70 75 80

Pro Asp Ala Ala Cys Leu Phe Phe Asp Ser Glu Phe Gly Ile Thr Ser

85 90 95

Ala Tyr Leu Glu Ser Met Gly Val Asp Pro Asp Arg Val Val His Val

100 105 110

Pro Ile Lys Asn Ile Glu Glu Leu Lys Phe Glu Ile Met Asn Gln Leu

115 120 125

Glu Gln Ile Thr Arg Glu Asp Lys Val Ile Ile Phe Ile Asp Ser Ile

130 135 140

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

145 150 155 160

Ser Ala Gln Asp Met Thr Arg Ala Lys Ala Leu Lys Gly Leu Phe Arg

165 170 175

Met Val Thr Pro Tyr Leu Thr Met Asn Asp Ile Pro Cys Ile Ala Ile

180 185 190

Asn His Thr Tyr Glu Thr Gln Glu Met Phe Ser Lys Thr Val Met Ser

195 200 205

Gly Gly Thr Gly Ala Met Tyr Ser Ala Asn Glu Val Phe Ile Ile Gly

210 215 220

Arg Arg Gln Gln Lys Glu Gly Thr Glu Ile Thr Gly Tyr Asp Phe Ile

225 230 235 240

Leu Asn Ala Glu Lys Ser Arg Thr Val Lys Glu Lys Ser Lys Phe Pro

245 250 255

Ile Ser Val Thr Phe Ser Gly Gly Ile Asp Pro Tyr Ser Gly Leu Leu

260 265 270

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

275 280 285

Tyr Ser Arg Ser Ile Leu Asn Thr Glu Thr Gly Glu Met Glu Thr Glu

290 295 300

Glu Arg Lys Phe Arg Ala Lys Glu Thr Asn Ser Ile Glu Phe Trp Lys

305 310 315 320

Pro Leu Leu Thr Asn Asp Lys Phe Asn Glu Ala Ile Asn Asp His Tyr

325 330 335

Lys Leu Gly Gln Val Ile Ser Asp Glu Ala Val Asp Lys Glu Ile Glu

340 345 350

Asp Met Leu Ala

355

<210> 2

<211> 356

<212> PRT

<213> Enterobacteria phage RB49

<400> 2

Met Ser Val Leu Glu Lys Leu Lys Lys Asn Ser Thr Leu Lys Thr Thr

1 5 10 15

Ala Val Leu Ser Lys Ser Ser Phe Phe Asn Glu Lys Thr Asn Thr Arg

20 25 30

Thr Lys Ile Pro Met Leu Asn Ile Ala Phe Ser Gly Asp Leu Lys Lys

35 40 45

Gly Phe Gln Ser Gly Leu Ile Phe Phe Ala Gly Pro Ser Lys Ser Phe

50 55 60

Lys Ser Asn Met Gly Leu Thr Cys Val Ser Ala Tyr Met Lys Gln Asn

65 70 75 80

Pro Asp Ala Ala Cys Leu Phe Phe Asp Ser Glu Phe Gly Ile Thr Ser

85 90 95

Ala Tyr Leu Glu Ser Met Gly Val Asp Pro Asp Arg Val Val His Val

100 105 110

Pro Ile Lys Asn Ile Glu Glu Leu Lys Phe Glu Ile Met Asn Gln Leu

115 120 125

Glu Gln Ile Thr Arg Glu Asp Lys Val Ile Ile Phe Ile Asp Ser Ile

130 135 140

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

145 150 155 160

Ser Ala Gln Asp Met Thr Arg Ala Lys Ala Leu Lys Gly Leu Phe Arg

165 170 175

Met Val Thr Pro Tyr Leu Thr Met Asn Asp Ile Pro Cys Ile Ala Ile

180 185 190

Asn His Thr Tyr Glu Thr Gln Glu Met Phe Ser Lys Thr Val Met Ser

195 200 205

Gly Gly Thr Gly Ala Met Tyr Ser Ala Asn Glu Val Phe Ile Ile Gly

210 215 220

Arg Arg Gln Gln Lys Glu Gly Thr Glu Ile Thr Gly Tyr Asp Phe Ile

225 230 235 240

Leu Asn Ala Glu Lys Ser Arg Thr Val Lys Glu Lys Ser Lys Phe Lys

245 250 255

Ile Ser Val Thr Phe Ser Gly Gly Ile Asp Pro Tyr Ser Gly Leu Leu

260 265 270

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

275 280 285

Tyr Ser Arg Ser Ile Leu Asn Thr Glu Thr Gly Glu Met Glu Thr Glu

290 295 300

Glu Arg Lys Phe Arg Ala Lys Glu Thr Asn Ser Ile Glu Phe Trp Lys

305 310 315 320

Pro Leu Leu Thr Asn Asp Lys Phe Asn Glu Ala Ile Asn Asp His Tyr

325 330 335

Lys Leu Gly Gln Val Ile Ser Asp Glu Ala Val Asp Lys Glu Ile Glu

340 345 350

Asp Met Leu Ala

355

<210> 3

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> Semi Conserved Domain

<400> 3

Lys Glu Lys Ser Lys Phe Phe Ile

1 5

<210> 4

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> Semi Conserved Domain

<400> 4

Lys Glu Lys Ser Lys Phe Pro Ile

1 5

<210> 5

<211> 8

<212> PRT

<213> Artificial Sequence

<220><223> Semi Conserved Domain

<220><221> VARIANT

<222> (5)..(5)

<223> X = any AA

<220

><221> VARIANT

<222> (7)..(7)

<223> X = any AA

<400> 5

Lys Glu Lys Ser Xaa Phe Xaa Ile

1 5

<210> 6

<211> 13

<212> PRT

<213> Artificial Sequence

<220><223> Semi Conserved Domain

<220><221> VARIANT

<222> (4)..(4)

<223> X = any AA

<220><221> VARIANT

<222> (5)..(5)

<223> X = any AA

<220><221> VARIANT

<222> (8)..(8)

<223> X = any AA

<220><221> VARIANT

<222> (9)..(9)

<223> X = any AA

<220><221> VARIANT

<222> (10)..(10)

<223> X = any AA

<220><221> VARIANT

<222> (11)..(11)

<223> X = any AA

<220><221> VARIANT

<222> (12)..(12)

<223> X = any AA

<220><221> VARIANT

<222> (13)..(13)

<223> X = any AA

<400> 6

Lys Ser Arg Xaa Xaa Lys Glu Xaa Xaa Xaa Xaa Xaa Xaa

1 5 10

<210> 7

<211> 13

<212> PRT

<213> Artificial Sequence

<220><223> Semi Conserved Domain

<400> 7

Lys Ser Arg Thr Val Lys Glu Lys Ser Lys Phe Phe Ile

1 5 10

<210> 8

<211> 391

<212> PRT

<213> Enterobacteria phage T4

<400> 8

Met Ser Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser Lys Leu
1 5 10 15

Thr Ala Glu Leu Thr Ala Ser Lys Phe Phe Asn Glu Lys Asp Val Val
20 25 30

Arg Thr Lys Ile Pro Met Met Asn Ile Ala Leu Ser Gly Glu Ile Thr
35 40 45

Gly Gly Met Gln Ser Gly Leu Leu Ile Leu Ala Gly Pro Ser Lys Ser

50 55 60

Phe Lys Ser Asn Phe Gly Leu Thr Met Val Ser Ser Tyr Met Arg Gln
65 70 75 80

Tyr Pro Asp Ala Val Cys Leu Phe Tyr Asp Ser Glu Phe Gly Ile Thr
85 90 95

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

Thr Pro Val Gln Ser Leu Glu Gln Leu Arg Ile Asp Met Val Asn Gln

115 120 125

Leu Asp Ala Ile Glu Arg Gly Glu Lys Val Val Val Phe Ile Asp Ser
130 135 140

Leu Gly Asn Leu Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu Asn Glu
145 150 155 160

Lys Val Val Ser Asp Met Thr Arg Ala Lys Thr Met Lys Ser Leu Phe
165 170 175

Arg Ile Val Thr Pro Tyr Phe Ser Thr Lys Asn Ile Pro Cys Ile Ala

180 185 190

Ile Asn His Thr Tyr Glu Thr Gln Glu Met Phe Ser Lys Thr Val Met

195 200 205

Gly Gly Gly Thr Gly Pro Met Tyr Ser Ala Asp Thr Val Phe Ile Ile

210 215 220

Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr Gln Phe

225 230 235 240

Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser Lys Phe

245 250 255

Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser Gly Leu

260 265 270

Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys Asn Gly

275 280 285

Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met Ile Arg

290 295 300

Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr Phe Trp

305 310 315 320

Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys Arg Ala

325 330 335

Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala Glu Val

340 345 350

Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro Glu Ser

355 360 365

Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln Leu Ser

370 375 380

Asp Met Glu Glu Phe Asn Glu

385 390

<210> 9

<211> 393

<212> PRT

<213> Enterobacteria phage T6

<400> 9

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp
20 25 30
Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala

35 40 45
Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser
50 55 60
Lys His Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu
65 70 75 80
Asn Lys Tyr Pro Asp Ala Val Cys Leu Phe Tyr Asp Ser Glu Phe Gly
85 90 95
Ile Thr Pro Ala Tyr Leu Arg Ser Met Gly Val Asp Pro Glu Arg Val

100 105 110
Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val
115 120 125
Asn Gln Leu Glu Ala Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile
130 135 140
Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu
145 150 155 160
Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser

165 170 175
Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys
180 185 190
Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
195 200 205
Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe
210 215 220
Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr

225 230 235 240
Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser
245 250 255
Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser

260 265 270

Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys

275 280 285

Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met

290 295 300

Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr

305 310 315 320

Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys

325 330 335

Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala

340 345 350

Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro

355 360 365

Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln

370 375 380

Leu Ser Asp Met Glu Glu Phe Asn Glu

385 390

<210> 10

<211> 387

<212> PRT

<213> Acinetobacter phage 133

<400> 10

Met Ser Ser Leu Lys Glu Arg Leu Ile Lys Ala Ser Thr Ser Lys Met

1 5 10 15

Thr Ala Glu Leu Thr Lys Ser Lys Phe Phe Asn Asp Lys Thr Val Val

20 25 30

Arg Thr Arg Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala Leu Asn

35 40 45

Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser Lys His

50 55 60

Phe Lys Ser Asn Met Gly Leu Thr Met Val Ala Ala Tyr Met Lys Ala

65 70 75 80

Phe Pro Asp Ala Val Cys Met Phe Tyr Asp Ser Glu Phe Gly Ile Thr

85 90 95

Pro Ala Tyr Leu Lys Ala Met Gly Val Asp Pro Asp Arg Val Ile His

100 105 110

Thr Pro Val Gln Ser Val Glu Gln Leu Lys Ile Asp Met Thr Asn Gln

115 120 125

Leu Glu Glu Val Lys Arg Gly Glu Lys Val Ile Val Phe Ile Asp Ser

130 135 140

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

145 150 155 160

Lys Thr Thr Ala Asp Met Thr Arg Ala Lys Ala Leu Lys Ser Leu Phe

165 170 175

Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asp Ile Pro Cys Val Ala

180 185 190

Val Asn His Thr Leu Gln Thr Leu Glu Met Phe Ser Lys Glu Val Met

195 200 205

Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe Phe Ile

210 215 220

Gly Lys Arg Gln Val Lys Asp Gly Thr Glu Leu Ala Gly Tyr Glu Phe

225 230 235 240

Ile Leu Lys Ala Glu Lys Ser Arg Met Val Lys Glu Lys Ser Val Phe

245 250 255

Pro Ile Thr Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser Gly Leu

260 265 270

Leu Glu Met Ala Thr Asp Leu Gly Phe Val Val Lys Pro Lys Val Gly

275 280 285

Trp Tyr Lys Arg Ala Met Met Val Asp Gly Val Met Gln His Glu Glu

290 295 300

Lys Ser Trp Arg Ala Lys Asp Thr Asp Ser Ile Asp Phe Trp Gly Pro

305 310 315 320

Leu Phe Lys His Asp Glu Phe Arg Lys Ala Ile Glu Thr Arg Tyr Gln

325 330 335

Leu Gly Ser Ile Glu Ser Asp Ala Glu Val Asp Ala Glu Val Asp Ala

340 345 350

Leu Ile Gly Ser Lys Thr Thr Ala Lys Ile Ser Gly Val Asn Phe Gly

355 360 365

Pro Ala Glu Ser Ala Ala Asp Lys Glu Gln Gln Leu Glu Asp Phe Val

370 375 380

Asp Glu Asp

385

<210> 11

<211> 390

<212> PRT

<213> Enterobacteria phage RB69

<400> 11

Met Ser Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser Lys Met

1 5 10 15

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

20 25 30

Pro Thr Arg Ile Pro Met Leu Asn Ile Ala Leu Gly Gly Ala Leu Asn

35 40 45

Ala Gly Leu Gln Ser Gly Leu Thr Ile Phe Ala Ala Pro Ser Lys His

50 55 60

Phe Lys Thr Leu Phe Gly Leu Thr Met Val Ala Ala Tyr Met Lys Lys

65 70 75 80

Tyr Lys Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly Ala Ser

85 90 95

Glu Ser Tyr Phe Arg Ser Met Gly Val Asp Leu Asp Arg Val Val His

100 105 110

Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Val Asp Met Thr Asn Gln

115 120 125

Leu Asp Ala Ile Glu Arg Gly Asp Lys Val Ile Ile Phe Ile Asp Ser

130 135 140

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

145 150 155 160

Lys Val Val Gly Asp Met Ser Arg Ala Lys Ala Leu Lys Ser Leu Phe

165 170 175

Arg Ile Val Thr Pro Tyr Leu Thr Ile Lys Asp Ile Pro Cys Val Ala

180 185 190

Ile Asn His Thr Ala Met Glu Ile Gly Gly Leu Tyr Pro Lys Glu Ile

195 200 205

Met Gly Gly Gly Thr Gly Ile Leu Tyr Ser Ala Asn Thr Val Phe Phe

210 215 220

Ile Ser Lys Arg Gln Val Lys Glu Gly Thr Glu Leu Thr Gly Tyr Asp

225 230 235 240

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

245 250 255

Phe Pro Ile Thr Val Asn Phe Asp Gly Gly Ile Asp Pro Phe Ser Gly

260 265 270

Leu Leu Glu Met Ala Thr Glu Ile Gly Phe Val Val Lys Pro Lys Ala

275 280 285

Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met Ile

290 295 300

Arg Glu Glu Lys Ser Trp Arg Ala Lys Ala Thr Asp Cys Val Glu Phe

305 310 315 320

Trp Gly Pro Leu Phe Lys His Lys Pro Phe Arg Asp Ala Ile Glu Thr

325 330 335

Lys Tyr Lys Leu Gly Ala Ile Ser Ser Ile Lys Glu Val Asp Asp Ala

340 345 350

Val Asn Asp Leu Ile Asn Cys Lys Ala Thr Thr Lys Val Pro Val Lys

355 360 365

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

370 375 380

Met Glu Asp Phe Asp Glu

385 390

<210> 12

<211> 411

<212> PRT

<213> Aeromonas phage Aeh1

<400> 12

Met Ala Lys Gly Ile Lys Thr Ala Lys Thr Gly Asn Leu Gly Ser Leu
1 5 10 15

Met Ser Lys Leu Ala Gly Thr Ser Ser Asn Lys Met Ser Ser Val Leu
20 25 30

Ala Asp Ser Lys Phe Phe Asn Asp Lys Asp Cys Val Arg Thr Arg Val
35 40 45

Pro Leu Leu Asn Leu Ala Met Ser Gly Glu Leu Asp Gly Gly Leu Thr

50 55 60

Pro Gly Leu Thr Val Leu Ala Gly Pro Ser Lys His Phe Lys Ser Asn
65 70 75 80

Leu Ser Leu Val Phe Val Ala Ala Tyr Leu Arg Lys Tyr Pro Asp Ala
85 90 95

Val Cys Ile Phe Phe Asp Asn Glu Phe Gly Ser Thr Pro Gly Tyr Phe
100 105 110

Glu Ser Gln Gly Val Asp Ile Ser Arg Val Ile His Cys Pro Phe Lys

115 120 125

Asn Ile Glu Glu Leu Lys Phe Asp Ile Val Lys Lys Leu Glu Ala Ile
130 135 140

Glu Arg Gly Asp Arg Val Ile Val Phe Val Asp Ser Ile Gly Asn Ala
145 150 155 160

Ala Ser Lys Lys Glu Ile Asp Asp Ala Ile Asp Glu Lys Ser Val Ser
165 170 175

Asp Met Thr Arg Ala Lys Gln Ile Lys Ser Leu Thr Arg Met Met Thr

180 185 190

Pro Tyr Leu Thr Val Asn Asp Ile Pro Ala Ile Met Val Ala His Thr

195 200 205

Tyr Asp Thr Gln Glu Met Tyr Ser Lys Lys Val Val Ser Gly Gly Thr

210 215 220

Gly Ile Thr Tyr Ser Ser Asp Thr Val Ile Ile Ile Gly Arg Gln Gln

225 230 235 240

Glu Lys Asp Gly Lys Glu Leu Leu Gly Tyr Asn Phe Val Leu Asn Met

245 250 255

Glu Lys Ser Arg Phe Val Lys Glu Gln Ser Lys Leu Pro Leu Glu Val

260 265 270

Thr Phe Gln Gly Gly Ile Asn Thr Tyr Ser Gly Met Leu Asp Ile Ala

275 280 285

Leu Glu Val Gly Phe Val Val Lys Pro Ser Asn Gly Trp Phe Ser Arg

290 295 300

Ala Phe Leu Asp Glu Glu Thr Gly Glu Leu Val Glu Glu Asp Arg Lys

305 310 315 320

Trp Arg Arg Ala Asp Thr Asn Cys Leu Glu Phe Trp Lys Pro Met Phe

325 330 335

Ala His Gln Pro Phe Lys Thr Ala Cys Ser Asp Met Phe Lys Leu Lys

340 345 350

Ser Val Ala Val Lys Asp Glu Val Phe Asp Glu Val Asp Glu Leu Phe

355 360 365

Ser Gly Glu Ala Glu Met Pro Val Asn Met Gly Arg Lys Leu Asp Thr

370 375 380

Ala Asp Gln Glu Glu Ile Asp Gln Leu Glu Glu Val Asp Val Glu Gly

385 390 395 400

Ser Asp Ser Asp Glu Leu Phe Ala Asn Leu Asp

405 410

<210> 13

<211> 368

<212> PRT

<213> Aeromonas phage 65

<400> 13

Met Ala Lys Lys Ala Lys Val Val Asn Ser Gly Asp Leu Leu Glu Arg
1 5 10 15
Leu Asn Gly Thr Ser Ser Asn Lys Met Ser Ala Met Leu Ala Glu Ser

20 25 30
Ile Phe Phe Asn Glu Lys Asp Thr Ile Arg Thr Arg Val Pro Ile Ile
35 40 45
Asn Leu Met Met Ser Gly Arg Leu Asp Gly Gly Ile Thr Pro Gly Leu
50 55 60
Thr Cys Ile Ala Gly Pro Ser Lys His Phe Lys Ser Asn Leu Ser Leu
65 70 75 80
Val Met Val Ser Ala Tyr Leu Arg Lys Tyr Pro Lys Ala Val Cys Leu

85 90 95
Phe Phe Asp Asn Glu Phe Gly Ser Thr Pro Asp Tyr Phe Thr Ser Gln
100 105 110
Gly Val Asp Ile Ser Arg Val Val His Cys Pro Phe Ile Asp Val Glu
115 120 125
Glu Leu Lys Phe Asp Ile Val Lys Lys Leu Glu Ser Ile Thr Arg Gly
130 135 140
Asp Lys Val Ile Ile Tyr Ile Asp Ser Ile Gly Asn Val Ala Ser Lys

145 150 155 160
Lys Glu Leu Gln Asp Ala Lys Asp Glu Lys Ser Ala Gln Asp Met Thr
165 170 175
Arg Ala Lys Gln Ile Lys Ser Leu Phe Arg Met Val Thr Pro Tyr Leu
180 185 190
Thr Val Leu Asp Ile Pro Cys Ile Ala Val Asn His Thr Tyr Glu Thr
195 200 205
Gln Glu Met Phe Ser Lys Thr Val Met Ser Gly Gly Thr Gly Pro Met

210 215 220
Tyr Ser Ala Asp Thr Val Ile Ile Leu Gly Lys Gln Gln Asp Lys Asp
225 230 235 240
Gly Lys Glu Leu Leu Gly Tyr Asn Phe Val Met Asn Ala Glu Lys Ser

245 250 255

Arg Ala Ile Lys Glu Lys Ser Lys Leu Asp Leu Met Val Ser Phe Glu

260 265 270

Gly Gly Ile Asn Thr Tyr Ser Gly Leu Leu Lys Ile Ala Gln Glu Leu

275 280 285

Gly Phe Val Thr Lys Pro Gln Asn Ala Arg Tyr Gln Arg Asn Phe Leu

290 295 300

Asp Leu Glu Pro Gly Glu Met Val Ile Pro Glu Asp Glu Lys Lys Trp

305 310 315 320

Thr Glu Glu Glu Ser Asp Ser Leu Glu Phe Trp Lys Pro Met Phe Ser

325 330 335

His Lys Pro Phe Met Asp Ala Val Ser Asn Ala Tyr Lys Leu Lys Ala

340 345 350

Val Glu Val Ser Gln Glu Val Phe Asp Glu Val Asp Gln Leu Phe Gly

355 360 365

<210> 14

<211> 366

<212> PRT

<213> Vibrio phage KVP40

<400> 14

Met Ser Asp Leu Met Lys Ser Leu Lys Lys Ser Ser Thr Ser Gly Tyr

1 5 10 15

Ala Gln Val Leu Ser Glu Ser Gln Phe Met Phe Asp Lys Asp His Thr

20 25 30

Arg Thr Tyr Val Pro Ala Ile Asn Ile Ala Phe Ser Gly Glu Val Asp

35 40 45

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

50 55 60

Phe Lys Ser Asn Leu Gly Leu Val Gly Val Ala Ala Tyr Leu Lys Lys

65 70 75 80

Tyr Pro Asp Ala Val Cys Val Phe Ile Asp Thr Glu Phe Gly Ile Thr

85 90 95

Pro Ser Tyr Leu Arg Ser Gln Gly Val Asp Pro Asp Arg Val Leu His
100 105 110
Ile Gln Cys Glu Ser Val Glu Arg Met Lys Phe Glu Met Ala Asn Gln
115 120 125
Leu Lys Asp Leu Ala Glu Arg Lys Arg Ala Lys Lys Ala Gly Glu Glu
130 135 140
Pro Asp Arg Val Ile Phe Phe Ile Asp Ser Val Gly Asn Val Ala Ser
145 150 155 160

Ala Lys Glu Ile Asp Asp Ala Gln Asn Glu Lys Ser Val Ala Asp Met
165 170 175
Ser Arg Ala Lys Gln Leu Lys Ser Leu Phe Arg Ile Ile Thr Pro Tyr
180 185 190
Phe Thr Met Leu Asp Ile Pro Cys Ile Ala Ile Asn His Thr Tyr Gln
195 200 205
Thr Gln Glu Ile Tyr Ser Lys Thr Val Met Ser Gly Gly Thr Gly Ile
210 215 220

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

Thr Gly His Val Val Lys Pro Ser Asn Gly Trp Tyr Gln Arg Ala Thr
290 295 300
Val Asp Glu Glu Thr Gly Glu Met Ile Val Glu Glu Lys Lys Tyr Arg
305 310 315 320
Ala Lys Glu Thr Gln Thr Ile Ser Phe Trp Lys Asp Ile Ile Asn Ser
325 330 335

Pro Thr Phe Lys Glu Gly Val Lys Arg Ile Tyr Cys Leu Gly Gln Leu
340 345 350

Asp Glu Ser Glu Leu Phe Gly Glu Val Asp Ser Leu Phe Asp
355 360 365

<210> 15

<211> 386

<212> PRT

<213> Enterobacteria phage RB43

<400> 15

Met Ser Asn Lys Ala Leu Leu Lys Lys Leu Ile Lys Asn Ser Asn Ser
1 5 10 15

Gln Ser Ala Ala Ile Leu Ser Glu Ser Asp Val Phe Asn Asn Ile Thr
20 25 30

Lys Thr Arg Thr Arg Val Pro Ile Leu Asn Leu Ala Leu Ser Gly Ala
35 40 45

Phe Asp Gly Gly Leu Thr Ser Gly Leu Thr Leu Phe Ala Gly Pro Ser
50 55 60

Lys His Phe Lys Ser Asn Leu Gly Leu Val Thr Val Ser Ala Tyr Leu
65 70 75 80

Lys Ala Asn Glu Asp Ala Val Cys Leu Phe Tyr Asp Ser Glu Lys Gly
85 90 95

Val Thr Lys Ser Tyr Leu Lys Ser Met Gly Val Asp Pro Asp Arg Val
100 105 110

Val Tyr Thr Arg Ile Thr Thr Val Glu Gln Leu Arg Asn Asp Val Val
115 120 125

Ser Gln Leu Asp Ala Leu Glu Arg Gly Asp Lys Val Ile Ile Phe Val
130 135 140

Asp Ser Val Gly Asn Thr Ala Ser Lys Lys Glu Leu Ala Asp Ala Leu
145 150 155 160

Ser Asp Asn Asp Lys Gln Asp Met Thr Arg Ala Lys Ala Leu Lys Gly
165 170 175

Met Phe Arg Met Val Thr Pro Tyr Leu Ala Asp Leu Asp Ile Pro Met
180 185 190
Val Cys Ile Cys His Thr Tyr Asp Thr Gln Glu Met Tyr Ser Lys Lys
195 200 205
Val Ile Ser Gly Gly Thr Gly Leu Met Tyr Ser Ala Asp Thr Ala Ile
210 215 220
Ile Leu Gly Lys Gln Gln Val Lys Glu Gly Thr Glu Val Val Gly Tyr

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

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

355 360 365
Asp Asp Asp Ile Pro Asp Asp Ile Gly Ile Asp Gln Asp Glu Glu Pro
370 375 380
Ile Met
385

<210> 16
<211> 335
<212> PRT
<213> Prochlorococcus phage P-SSM2

<400> 16

Met Asp Phe Leu Lys Glu Ile Val Lys Glu Ile Gly Asp Glu Tyr Thr
1 5 10 15

Gln Val Ala Ala Asp Ile Gln Glu Asn Glu Arg Phe Ile Asp Thr Gly
20 25 30

Ser Tyr Ile Phe Asn Gly Leu Val Ser Gly Ser Ile Phe Gly Gly Val
35 40 45

Ser Ser Ser Arg Ile Thr Ala Ile Ala Gly Glu Ser Ser Thr Gly Lys
50 55 60

Thr Tyr Phe Ser Leu Ala Val Val Lys Asn Phe Leu Asp Asn Asn Pro
65 70 75 80

Asp Gly Tyr Cys Leu Tyr Phe Asp Thr Glu Ala Ala Val Asn Lys Gly
85 90 95

Leu Leu Glu Ser Arg Gly Ile Asp Met Asn Arg Leu Val Val Val Asn
100 105 110

Val Val Thr Ile Glu Glu Phe Arg Ser Lys Ala Leu Arg Ala Val Asp
115 120 125

Ile Tyr Leu Lys Thr Ser Glu Glu Glu Arg Lys Pro Cys Met Phe Val
130 135 140

Leu Asp Ser Leu Gly Met Leu Ser Thr Glu Lys Glu Ile Arg Asp Ala
145 150 155 160

Leu Asp Asp Lys Gln Val Arg Asp Met Thr Lys Ser Gln Leu Val Lys
165 170 175

Gly Ala Phe Arg Met Leu Thr Leu Lys Leu Gly Gln Ala Asn Ile Pro
180 185 190

Leu Ile Val Thr Asn His Thr Tyr Asp Val Ile Gly Ser Tyr Val Pro
195 200 205

Thr Lys Glu Met Gly Gly Gly Ser Gly Leu Lys Tyr Ala Ala Ser Thr
210 215 220

Ile Ile Tyr Leu Ser Lys Lys Lys Glu Lys Asp Gln Lys Glu Val Ile
225 230 235 240

Gly Asn Leu Ile Lys Ala Lys Thr His Lys Ser Arg Leu Ser Lys Glu
245 250 255
Asn Lys Glu Val Gln Ile Arg Leu Tyr Tyr Asp Glu Arg Gly Leu Asp
260 265 270
Arg Tyr Tyr Gly Leu Leu Glu Leu Gly Glu Ile Gly Gly Met Trp Lys
275 280 285

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

<210> 17
<211> 336
<212> PRT
<213> Prochlorococcus phage P-SSM4
<400> 17

Met Asn Phe Leu Lys Asp Ile Ala Lys Glu Ile Gly Asn Asp Tyr Ala

1 5 10 15
Ser Leu Val Ser Glu Gly Val Ser Ala Gly Asp Thr Ala Gly Phe Ile
20 25 30
Asp Thr Gly Ser Tyr Ile Phe Asn Ala Leu Leu Ser Gly Ser Ile Tyr
35 40 45
Gly Gly Ile Pro Asn Asn Lys Ile Thr Ala Ile Ala Gly Glu Thr Ser
50 55 60
Thr Gly Lys Thr Phe Phe Cys Leu Gly Met Val Gln His Phe Leu Glu

65 70 75 80
Ser Asn Pro Asp Ala Gly Val Ile Tyr Phe Glu Ser Glu Ser Ala Ile
85 90 95
Ser Lys Gln Met Ile Glu Asp Arg Gly Ile Asp Ser Asn Arg Met Leu
100 105 110
Leu Val Pro Val Thr Thr Val Gln Glu Phe Arg Leu Gln Ala Ile Lys

115 120 125

Ile Leu Asp Lys Tyr Asn Glu Gln Thr Ala Glu Glu Arg Lys Pro Leu

130 135 140

Met Phe Val Leu Asp Ser Leu Gly Met Leu Ser Thr Ser Lys Glu Val

145 150 155 160

Glu Asp Ser Glu Ala Gly Lys Glu Thr Arg Asp Met Thr Arg Ala Gln

165 170 175

Val Val Lys Ser Ile Phe Arg Val Leu Thr Leu Lys Leu Gly Lys Ala

180 185 190

Asn Val Pro Leu Ile Val Thr Asn His Thr Tyr Asp Val Val Gly Ala

195 200 205

Tyr Ile Pro Thr Lys Glu Met Gly Gly Gly Ser Gly Leu Lys Tyr Ala

210 215 220

Ala Ser Thr Ile Val Tyr Leu Ser Lys Lys Lys Glu Lys Asn Gly Lys

225 230 235 240

Glu Val Val Gly Asn Ile Ile Lys Cys Lys Thr Ala Lys Ser Arg Leu

245 250 255

Thr Lys Glu Asn Ser Asp Val Glu Thr Arg Leu Tyr Tyr Asp Arg Gly

260 265 270

Leu Asp Arg Tyr Tyr Gly Leu Leu Glu Leu Gly Glu Lys His Gly Val

275 280 285

Phe Ser Arg Lys Gly Asn Arg Val Val Val Gly Asp Ser Ser Val Tyr

290 295 300

Pro Ser Ala Ile Leu Ala Asp Pro Asp Lys Tyr Phe Thr Glu Glu Leu

305 310 315 320

Met Glu Lys Leu Asp Glu Ala Ala Ala Lys Glu Phe Arg Tyr Gly Asn

325 330 335

<210> 18

<211> 343

<212> PRT

<213> Synechococcus phage S-PM2

<400> 18

Met Ser Phe Leu Asp Ser Val Ile Lys Asp Ser Lys Asn Glu Tyr Ala
1 5 10 15

Ala Phe Ala Ser Glu Gly Val Ala Ala Gly Asp Val Glu Ser Phe Val
20 25 30

Asp Thr Gly Ser Tyr Ile Phe Asn Ala Leu Val Ser Gly Ser Ile Phe
35 40 45

Gly Gly Ile Pro Ser Asn Lys Ile Thr Ala Leu Ala Gly Glu Ser Gly
50 55 60

Thr Gly Lys Thr Phe Phe Cys Leu Ser Val Val Arg Asn Phe Leu Asn
65 70 75 80

Thr Asp Pro Asp Ala Gly Val Ile Tyr Phe Glu Thr Glu Ser Ala Ile
85 90 95

Ser Lys Gln Met Ile Glu Ser Arg Gly Ile Asp Ser Thr Arg Met Ile
100 105 110

Ile Phe Pro Val Asp Thr Ile Glu Asp Phe Arg Thr Gln Ala Val Arg
115 120 125

Ile Ile Asp Lys Tyr Met Glu Gln Asn Lys Ser Glu Arg Lys Pro Leu
130 135 140

Met Phe Val Leu Asp Ser Leu Gly Met Leu Ala Thr Lys Lys Glu Val
145 150 155 160

Glu Asp Ala Ser Asn Asp Lys Gln Val Arg Asp Met Thr Lys Ala Gln
165 170 175

Ile Val Lys Ser Ala Phe Arg Ile Leu Thr Leu Lys Met Gly Lys Ala
180 185 190

Asn Ile Pro Met Leu Val Thr Asn His Thr Tyr Asp Val Val Gly Ser
195 200 205

Tyr Val Pro Thr Lys Glu Met Gly Gly Gly Ser Gly Leu Lys Tyr Ser
210 215 220

Ala Ser Thr Ile Val Tyr Leu Gly Lys Lys Lys Glu Lys Asp Gly Thr
225 230 235 240

Asp Leu Val Gly Asn Ile Ile Lys Cys Glu Ala Lys Lys Ser Arg Leu
245 250 255

Thr Arg Glu Gly Ser Lys Val Glu Thr Arg Leu Phe Phe Asp Gln Arg
260 265 270

Gly Leu Glu Arg Tyr Tyr Gly Met Leu Glu Leu Gly Glu Arg Ala Gly
275 280 285

Leu Trp Lys Asn Thr Ala Gly Arg Tyr Glu Ile Asn Gly Lys Lys Val
290 295 300

Tyr Gly Lys Gln Ile Leu Ala Asn Pro Asp Glu Phe Phe Thr Glu Glu
305 310 315 320

Ile Leu Gln Glu Leu Asp Lys Gln Ala Gln Arg Glu Phe Leu Tyr Gly
325 330 335

Ala Ser Asp Asp Gly Glu Asp
340

<210> 19

<211> 393

<212> PRT

<213> Enterobacteria phage RB32

<400> 19

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser
1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp
20 25 30

Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala
35 40 45

Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser
50 55 60

Lys His Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu
65 70 75 80

Asn Lys Tyr Pro Asp Ala Val Cys Leu Phe Tyr Asp Ser Glu Phe Gly
85 90 95

Ile Thr Pro Ala Tyr Leu Arg Ser Met Gly Val Asp Pro Glu Arg Val

100 105 110

Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val

115 120 125

Asn Gln Leu Glu Ala Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile

130 135 140

Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

145 150 155 160

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

165 170 175

Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys

180 185 190

Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr

195 200 205

Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe

210 215 220

Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr

225 230 235 240

Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser

245 250 255

Lys Phe Phe Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser

260 265 270

Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys

275 280 285

Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met

290 295 300

Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr

305 310 315 320

Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys

325 330 335

Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala

340 345 350

Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro
355 360 365
Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln
370 375 380
Leu Ser Asp Met Glu Glu Phe Asn Glu
385 390

<210> 20
<211> 366
<212> PRT

<213> Vibrio phage nt-1
<400> 20

Met Ser Asp Leu Leu Lys Ser Leu Lys Lys Ser Ser Thr Ser Gly Tyr
1 5 10 15
Ala His Val Leu Ser Glu Ser Gln Phe Met Phe Glu Lys Asp His Thr
20 25 30
Arg Thr Tyr Val Pro Ala Ile Asn Ile Ala Phe Ser Gly Glu Val Asp
35 40 45
Gly Gly Leu Thr Ser Gly Leu Thr Val Leu Ala Gly Pro Ser Lys His
50 55 60

Phe Lys Ser Asn Leu Gly Leu Val Gly Val Ala Ala Tyr Leu Lys Lys
65 70 75 80
Tyr Pro Glu Ala Ile Cys Val Phe Ile Asp Thr Glu Phe Gly Ile Thr
85 90 95
Pro Ser Tyr Leu Lys Ser Gln Gly Val Asp Pro Glu Arg Val Leu His
100 105 110
Ile Gln Cys Glu Ser Val Glu Arg Met Lys Phe Glu Met Ala Asn Gln
115 120 125

Leu Lys Asp Leu Ala Glu Arg Lys Arg Ala Lys Lys Ala Gly Glu Glu
130 135 140
Pro Asp Arg Val Val Phe Phe Ile Asp Ser Val Gly Asn Val Ala Ser
145 150 155 160
Ala Lys Glu Ile Asp Asp Ala Gln Asn Glu Lys Ser Val Ala Asp Met

165 170 175

Ser Arg Ala Lys Gln Leu Lys Ser Leu Phe Arg Ile Ile Thr Pro Tyr

180 185 190

Phe Thr Met Leu Asp Ile Pro Cys Ile Ala Ile Asn His Thr Tyr Gln

195 200 205

Thr Gln Glu Met Tyr Ser Lys Thr Val Met Ser Gly Gly Thr Gly Ile

210 215 220

Met Tyr Ser Ala Asp Thr Val Ile Ile Leu Gly Lys Gln Gln Glu Lys

225 230 235 240

Asp Gly Lys Glu Ile Ile Gly Tyr His Phe Ile Met Asn Ile Glu Lys

245 250 255

Ser Arg Phe Val Lys Glu Lys Met Lys Val Pro Leu Thr Val Thr Tyr

260 265 270

Glu His Gly Ile Asp Gln Phe Ser Gly Leu Leu Asp Ile Ala Leu Gln

275 280 285

Thr Gly His Val Val Lys Pro Ser Asn Gly Trp Tyr Gln Arg Ala Phe

290 295 300

Ile Asp Glu Glu Thr Gly Glu Ile Glu Ile Glu Glu Lys Lys Tyr Arg

305 310 315 320

Ala Lys Glu Thr Gln Thr Leu Ser Phe Trp Lys Glu Ile Ile Asn Ser

325 330 335

Pro Thr Phe Lys Thr Gly Val Lys Arg Leu Tyr Cys Leu Gly Gln Leu

340 345 350

Asp Glu Ser Glu Leu Leu Asp Glu Val Asp Ser Leu Phe Asp

355 360 365

<210> 21

<211> 381

<212> PRT

<213> Enterobacteria phage RB16

<400> 21

Met Ser Asn Lys Ala Leu Leu Lys Lys Leu Ile Lys Asn Ser Asn Ser

1 5 10 15

Gln Ser Ala Ser Ile Leu Ser Glu Ser Asp Val Phe Asn Asn Ile Thr

20 25 30

Lys Thr Arg Thr Arg Val Pro Ile Leu Asn Leu Val Leu Ser Gly Ala

35 40 45

Phe Asp Gly Gly Leu Thr Ser Gly Leu Thr Leu Ile Ala Gly Pro Ser

50 55 60

Lys His Phe Lys Ser Asn Leu Gly Leu Val Ala Val Ala Ala Tyr Leu

65 70 75 80

Lys Ala Asn Glu Asp Ala Val Cys Leu Phe Tyr Asp Ser Glu Lys Gly

85 90 95

Val Thr Lys Ser Tyr Leu Lys Ser Met Gly Val Asp Pro Asp Arg Val

100 105 110

Val Tyr Thr Arg Ile Thr Thr Val Glu Gln Leu Arg Asn Asp Val Val

115 120 125

Ser Gln Leu Asp Ala Leu Glu Arg Gly Asp Lys Val Ile Ile Phe Val

130 135 140

Asp Ser Val Gly Asn Thr Ala Ser Lys Lys Glu Leu Lys Asp Ala Leu

145 150 155 160

Glu Asp Asn Asp Lys Gln Asp Met Thr Arg Ala Lys Ala Leu Lys Gly

165 170 175

Met Phe Arg Met Val Thr Pro Tyr Leu Ala Asp Ile Asp Ile Pro Met

180 185 190

Val Cys Ile Cys His Thr Tyr Asp Thr Gln Glu Met Tyr Ser Lys Lys

195 200 205

Val Ile Ser Gly Gly Thr Gly Leu Met Tyr Ser Ala Asp Thr Ala Ile

210 215 220

Ile Leu Gly Lys Gln Gln Val Lys Glu Gly Thr Glu Val Val Gly Tyr

225 230 235 240

Asp Phe Ile Met Asn Val Glu Lys Ser Arg Phe Val Lys Glu Lys Ser

245 250 255

Lys Phe Pro Leu His Val Thr Tyr Glu Gly Gly Ile Ser Met Phe Ser

260 265 270

Gly Leu Leu Asp Leu Ala Met Glu Met Asn Phe Val Gln Thr Pro Thr

275 280 285

Lys Gly Trp Arg Gly Arg Ala Phe Leu Asn Thr Glu Thr Gly Glu Leu

290 295 300

Glu Leu Glu Glu Lys Lys Trp Arg Glu Ala Glu Thr Asn Cys Ile Glu

305 310 315 320

Phe Trp Lys Pro Leu Phe Lys His Gln Pro Phe Ile Asp Ala Ile Gln

325 330 335

Asp Lys Tyr Arg Ile Pro Asp Lys Glu Ile Thr Asp Gly Ala Ala Leu

340 345 350

Glu Asp Leu Tyr Ser Asp Asp Val Val Glu Ser Asn Lys Val Asp Phe

355 360 365

Asp Asp Asp Ile Pro Asp Asp Val Asp Leu Met Glu Glu

370 375 380

<210> 22

<211> 391

<212> PRT

<213> Enterobacteria phage T4

<400> 22

Met Ser Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser Lys Leu

1 5 10 15

Thr Ala Glu Leu Thr Ala Ser Lys Phe Phe Asn Glu Lys Asp Val Val

20 25 30

Arg Thr Lys Ile Pro Met Met Asn Ile Ala Leu Ser Gly Glu Ile Thr

35 40 45

Gly Gly Met Gln Ser Gly Leu Leu Ile Leu Ala Gly Pro Ser Lys Ser

50 55 60

Phe Lys Ser Asn Phe Gly Leu Thr Met Val Ser Ser Tyr Met Arg Gln

65 70 75 80

Tyr Pro Asp Ala Val Cys Leu Phe Tyr Asp Ser Glu Phe Gly Ile Thr
85 90 95
Pro Ala Tyr Leu Arg Ser Met Gly Val Asp Pro Glu Arg Val Ile His
100 105 110
Thr Pro Val Gln Ser Leu Glu Gln Leu Arg Ile Asp Met Val Asn Gln
115 120 125

Leu Asp Ala Ile Glu Arg Gly Glu Lys Val Val Val Phe Ile Asp Ser
130 135 140
Leu Gly Asn Leu Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu Asn Glu
145 150 155 160
Lys Val Val Ser Asp Met Thr Arg Ala Lys Thr Met Lys Ser Leu Phe
165 170 175
Arg Ile Val Thr Pro Tyr Phe Ser Thr Lys Asn Ile Pro Cys Ile Ala
180 185 190

Ile Asn His Thr Tyr Glu Thr Gln Glu Met Phe Ser Lys Thr Val Met
195 200 205
Gly Gly Gly Thr Gly Pro Met Tyr Ser Ala Asp Thr Val Phe Ile Ile
210 215 220
Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr Gln Phe
225 230 235 240
Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser Lys Phe
245 250 255

Lys Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser Gly Leu
260 265 270
Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys Asn Gly
275 280 285
Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met Ile Arg
290 295 300
Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr Phe Trp
305 310 315 320

Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys Arg Ala

325 330 335

Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala Glu Val

340 345 350

Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro Glu Ser

355 360 365

Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln Leu Ser

370 375 380

Asp Met Glu Glu Phe Asn Glu

385 390

<210> 23

<211> 393

<212> PRT

<213> Enterobacteria phage T6

<400> 23

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser

1 5 10 15

Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

20 25 30

Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala

35 40 45

Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser

50 55 60

Lys His Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu

65 70 75 80

Asn Lys Tyr Pro Asp Ala Val Cys Leu Phe Tyr Asp Ser Glu Phe Gly

85 90 95

Ile Thr Pro Ala Tyr Leu Arg Ser Met Gly Val Asp Pro Glu Arg Val

100 105 110

Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val

115 120 125

Asn Gln Leu Glu Ala Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile

130 135 140

Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu
145 150 155 160
Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser
165 170 175

Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys
180 185 190
Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
195 200 205
Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe
210 215 220
Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr
225 230 235 240

Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser
245 250 255
Lys Phe Lys Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser
260 265 270
Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys
275 280 285
Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met
290 295 300

Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr
305 310 315 320
Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys
325 330 335
Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala
340 345 350
Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro
355 360 365

Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln
370 375 380
Leu Ser Asp Met Glu Glu Phe Asn Glu

385 390

<210> 24

<211> 387

<212> PRT

<213> Acinetobacter phage 133

<400> 24

Met Ser Ser Leu Lys Glu Arg Leu Ile Lys Ala Ser Thr Ser Lys Met
1 5 10 15

Thr Ala Glu Leu Thr Lys Ser Lys Phe Phe Asn Asp Lys Thr Val Val
20 25 30

Arg Thr Arg Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala Leu Asn
35 40 45

Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser Lys His
50 55 60

Phe Lys Ser Asn Met Gly Leu Thr Met Val Ala Ala Tyr Met Lys Ala
65 70 75 80

Phe Pro Asp Ala Val Cys Met Phe Tyr Asp Ser Glu Phe Gly Ile Thr
85 90 95

Pro Ala Tyr Leu Lys Ala Met Gly Val Asp Pro Asp Arg Val Ile His
100 105 110

Thr Pro Val Gln Ser Val Glu Gln Leu Lys Ile Asp Met Thr Asn Gln
115 120 125

Leu Glu Glu Val Lys Arg Gly Glu Lys Val Ile Val Phe Ile Asp Ser
130 135 140

Ile Gly Asn Leu Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu Asn Glu
145 150 155 160

Lys Thr Thr Ala Asp Met Thr Arg Ala Lys Ala Leu Lys Ser Leu Phe
165 170 175

Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asp Ile Pro Cys Val Ala
180 185 190

Val Asn His Thr Leu Gln Thr Leu Glu Met Phe Ser Lys Glu Val Met
195 200 205

Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe Phe Ile
210 215 220

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

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

Leu Ile Gly Ser Lys Thr Thr Ala Lys Ile Ser Gly Val Asn Phe Gly
355 360 365
Pro Ala Glu Ser Ala Ala Asp Lys Glu Gln Gln Leu Glu Asp Phe Val
370 375 380
Asp Glu Asp
385

<210> 25
<211> 390
<212> PRT
<213> Enterobacteria phage RB69

<400> 25
Met Ser Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser Lys Met
1 5 10 15

Thr Ala Asp Leu Thr Lys Ser Lys Leu Phe Asn Asn Arg Asp Glu Val
20 25 30
Pro Thr Arg Ile Pro Met Leu Asn Ile Ala Leu Gly Gly Ala Leu Asn
35 40 45
Ala Gly Leu Gln Ser Gly Leu Thr Ile Phe Ala Ala Pro Ser Lys His
50 55 60
Phe Lys Thr Leu Phe Gly Leu Thr Met Val Ala Ala Tyr Met Lys Lys
65 70 75 80

Tyr Lys Asp Ala Ile Cys Leu Phe Tyr Asp Ser Glu Phe Gly Ala Ser
85 90 95
Glu Ser Tyr Phe Arg Ser Met Gly Val Asp Leu Asp Arg Val Val His
100 105 110
Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Val Asp Met Thr Asn Gln
115 120 125
Leu Asp Ala Ile Glu Arg Gly Asp Lys Val Ile Ile Phe Ile Asp Ser
130 135 140

Ile Gly Asn Thr Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu Asn Glu
145 150 155 160
Lys Val Val Gly Asp Met Ser Arg Ala Lys Ala Leu Lys Ser Leu Phe
165 170 175
Arg Ile Val Thr Pro Tyr Leu Thr Ile Lys Asp Ile Pro Cys Val Ala
180 185 190
Ile Asn His Thr Ala Met Glu Ile Gly Gly Leu Tyr Pro Lys Glu Ile
195 200 205

Met Gly Gly Gly Thr Gly Ile Leu Tyr Ser Ala Asn Thr Val Phe Phe
210 215 220
Ile Ser Lys Arg Gln Val Lys Glu Gly Thr Glu Leu Thr Gly Tyr Asp
225 230 235 240
Phe Thr Leu Lys Ala Glu Lys Ser Arg Thr Val Lys Glu Lys Ser Thr
245 250 255
Phe Lys Ile Thr Val Asn Phe Asp Gly Gly Ile Asp Pro Phe Ser Gly

260 265 270

Leu Leu Glu Met Ala Thr Glu Ile Gly Phe Val Val Lys Pro Lys Ala
275 280 285
Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met Ile
290 295 300
Arg Glu Glu Lys Ser Trp Arg Ala Lys Ala Thr Asp Cys Val Glu Phe
305 310 315 320
Trp Gly Pro Leu Phe Lys His Lys Pro Phe Arg Asp Ala Ile Glu Thr
325 330 335

Lys Tyr Lys Leu Gly Ala Ile Ser Ser Ile Lys Glu Val Asp Asp Ala
340 345 350
Val Asn Asp Leu Ile Asn Cys Lys Ala Thr Thr Lys Val Pro Val Lys
355 360 365
Thr Ser Asp Ala Pro Ser Ala Ala Asp Ile Glu Asn Asp Leu Asp Glu
370 375 380
Met Glu Asp Phe Asp Glu
385 390

<210> 26
<211> 411
<212> PRT
<213> Aeromonas phage Aeh1

<400> 26
Met Ala Lys Gly Ile Lys Thr Ala Lys Thr Gly Asn Leu Gly Ser Leu
1 5 10 15
Met Ser Lys Leu Ala Gly Thr Ser Ser Asn Lys Met Ser Ser Val Leu
20 25 30
Ala Asp Ser Lys Phe Phe Asn Asp Lys Asp Cys Val Arg Thr Arg Val
35 40 45
Pro Leu Leu Asn Leu Ala Met Ser Gly Glu Leu Asp Gly Gly Leu Thr
50 55 60

Pro Gly Leu Thr Val Leu Ala Gly Pro Ser Lys His Phe Lys Ser Asn

65 70 75 80

Leu Ser Leu Val Phe Val Ala Ala Tyr Leu Arg Lys Tyr Pro Asp Ala

85 90 95

Val Cys Ile Phe Phe Asp Asn Glu Phe Gly Ser Thr Pro Gly Tyr Phe

100 105 110

Glu Ser Gln Gly Val Asp Ile Ser Arg Val Ile His Cys Pro Phe Lys

115 120 125

Asn Ile Glu Glu Leu Lys Phe Asp Ile Val Lys Lys Leu Glu Ala Ile

130 135 140

Glu Arg Gly Asp Arg Val Ile Val Phe Val Asp Ser Ile Gly Asn Ala

145 150 155 160

Ala Ser Lys Lys Glu Ile Asp Asp Ala Ile Asp Glu Lys Ser Val Ser

165 170 175

Asp Met Thr Arg Ala Lys Gln Ile Lys Ser Leu Thr Arg Met Met Thr

180 185 190

Pro Tyr Leu Thr Val Asn Asp Ile Pro Ala Ile Met Val Ala His Thr

195 200 205

Tyr Asp Thr Gln Glu Met Tyr Ser Lys Lys Val Val Ser Gly Gly Thr

210 215 220

Gly Ile Thr Tyr Ser Ser Asp Thr Val Ile Ile Ile Gly Arg Gln Gln

225 230 235 240

Glu Lys Asp Gly Lys Glu Leu Leu Gly Tyr Asn Phe Val Leu Asn Met

245 250 255

Glu Lys Ser Arg Phe Val Lys Glu Gln Ser Lys Leu Lys Leu Glu Val

260 265 270

Thr Phe Gln Gly Gly Ile Asn Thr Tyr Ser Gly Met Leu Asp Ile Ala

275 280 285

Leu Glu Val Gly Phe Val Val Lys Pro Ser Asn Gly Trp Phe Ser Arg

290 295 300

Ala Phe Leu Asp Glu Glu Thr Gly Glu Leu Val Glu Glu Asp Arg Lys

305 310 315 320

Trp Arg Arg Ala Asp Thr Asn Cys Leu Glu Phe Trp Lys Pro Met Phe
325 330 335
Ala His Gln Pro Phe Lys Thr Ala Cys Ser Asp Met Phe Lys Leu Lys
340 345 350
Ser Val Ala Val Lys Asp Glu Val Phe Asp Glu Val Asp Glu Leu Phe
355 360 365
Ser Gly Glu Ala Glu Met Pro Val Asn Met Gly Arg Lys Leu Asp Thr
370 375 380

Ala Asp Gln Glu Glu Ile Asp Gln Leu Glu Glu Val Asp Val Glu Gly
385 390 395 400
Ser Asp Ser Asp Glu Leu Phe Ala Asn Leu Asp
405 410

<210> 27
<211> 368
<212> PRT
<213> Aeromonas phage 65
<400> 27

Met Ala Lys Lys Ala Lys Val Val Asn Ser Gly Asp Leu Leu Glu Arg
1 5 10 15
Leu Asn Gly Thr Ser Ser Asn Lys Met Ser Ala Met Leu Ala Glu Ser
20 25 30

Ile Phe Phe Asn Glu Lys Asp Thr Ile Arg Thr Arg Val Pro Ile Ile
35 40 45
Asn Leu Met Met Ser Gly Arg Leu Asp Gly Gly Ile Thr Pro Gly Leu
50 55 60
Thr Cys Ile Ala Gly Pro Ser Lys His Phe Lys Ser Asn Leu Ser Leu
65 70 75 80
Val Met Val Ser Ala Tyr Leu Arg Lys Tyr Pro Lys Ala Val Cys Leu
85 90 95

Phe Phe Asp Asn Glu Phe Gly Ser Thr Pro Asp Tyr Phe Thr Ser Gln
100 105 110
Gly Val Asp Ile Ser Arg Val Val His Cys Pro Phe Ile Asp Val Glu

```
        115                 120                 125
Glu Leu Lys Phe Asp Ile Val Lys Lys Leu Glu Ser Ile Thr Arg Gly
    130                 135                 140
Asp Lys Val Ile Ile Tyr Ile Asp Ser Ile Gly Asn Val Ala Ser Lys
145                 150                 155                 160

Lys Glu Leu Gln Asp Ala Lys Asp Glu Lys Ser Ala Gln Asp Met Thr
                165                 170                 175
Arg Ala Lys Gln Ile Lys Ser Leu Phe Arg Met Val Thr Pro Tyr Leu
            180                 185                 190
Thr Val Leu Asp Ile Pro Cys Ile Ala Val Asn His Thr Tyr Glu Thr
        195                 200                 205
Gln Glu Met Phe Ser Lys Thr Val Met Ser Gly Gly Thr Gly Pro Met
    210                 215                 220

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

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

Val Glu Val Ser Gln Glu Val Phe Asp Glu Val Asp Gln Leu Phe Gly
        355                 360                 365
```

<210> 28

<211> 366

<212> PRT

<213> Vibrio phage KVP40

<400> 28

Met Ser Asp Leu Met Lys Ser Leu Lys Lys Ser Ser Thr Ser Gly Tyr
1 5 10 15

Ala Gln Val Leu Ser Glu Ser Gln Phe Met Phe Asp Lys Asp His Thr
20 25 30

Arg Thr Tyr Val Pro Ala Ile Asn Ile Ala Phe Ser Gly Glu Val Asp
35 40 45

Gly Gly Leu Thr Ser Gly Leu Thr Val Leu Ala Gly Pro Ser Lys His
50 55 60

Phe Lys Ser Asn Leu Gly Leu Val Gly Val Ala Ala Tyr Leu Lys Lys
65 70 75 80

Tyr Pro Asp Ala Val Cys Val Phe Ile Asp Thr Glu Phe Gly Ile Thr
85 90 95

Pro Ser Tyr Leu Arg Ser Gln Gly Val Asp Pro Asp Arg Val Leu His
100 105 110

Ile Gln Cys Glu Ser Val Glu Arg Met Lys Phe Glu Met Ala Asn Gln
115 120 125

Leu Lys Asp Leu Ala Glu Arg Lys Arg Ala Lys Lys Ala Gly Glu Glu
130 135 140

Pro Asp Arg Val Ile Phe Phe Ile Asp Ser Val Gly Asn Val Ala Ser
145 150 155 160

Ala Lys Glu Ile Asp Asp Ala Gln Asn Glu Lys Ser Val Ala Asp Met
165 170 175

Ser Arg Ala Lys Gln Leu Lys Ser Leu Phe Arg Ile Ile Thr Pro Tyr
180 185 190

Phe Thr Met Leu Asp Ile Pro Cys Ile Ala Ile Asn His Thr Tyr Gln
195 200 205

Thr Gln Glu Ile Tyr Ser Lys Thr Val Met Ser Gly Gly Thr Gly Ile

210 215 220

Met Tyr Ser Ala Asp Thr Val Ile Ile Leu Gly Lys Gln Gln Glu Lys

225 230 235 240

Asp Gly Lys Asp Ile Ile Gly Tyr His Phe Ile Met Asn Ile Glu Lys

245 250 255

Ser Arg Phe Val Lys Glu Lys Met Lys Val Lys Leu Thr Val Thr Tyr

260 265 270

Glu Asn Gly Ile Asp Pro Phe Ser Gly Leu Leu Asp Ile Ala Leu Gln

275 280 285

Thr Gly His Val Val Lys Pro Ser Asn Gly Trp Tyr Gln Arg Ala Thr

290 295 300

Val Asp Glu Glu Thr Gly Glu Met Ile Val Glu Glu Lys Lys Tyr Arg

305 310 315 320

Ala Lys Glu Thr Gln Thr Ile Ser Phe Trp Lys Asp Ile Ile Asn Ser

325 330 335

Pro Thr Phe Lys Glu Gly Val Lys Arg Ile Tyr Cys Leu Gly Gln Leu

340 345 350

Asp Glu Ser Glu Leu Phe Gly Glu Val Asp Ser Leu Phe Asp

355 360 365

<210> 29

<211> 386

<212> PRT

<213> Enterobacteria phage RB43

<400> 29

Met Ser Asn Lys Ala Leu Leu Lys Lys Leu Ile Lys Asn Ser Asn Ser

1 5 10 15

Gln Ser Ala Ala Ile Leu Ser Glu Ser Asp Val Phe Asn Asn Ile Thr

20 25 30

Lys Thr Arg Thr Arg Val Pro Ile Leu Asn Leu Ala Leu Ser Gly Ala

35 40 45

Phe Asp Gly Gly Leu Thr Ser Gly Leu Thr Leu Phe Ala Gly Pro Ser

50 55 60

Lys His Phe Lys Ser Asn Leu Gly Leu Val Thr Val Ser Ala Tyr Leu
65 70 75 80
Lys Ala Asn Glu Asp Ala Val Cys Leu Phe Tyr Asp Ser Glu Lys Gly
85 90 95
Val Thr Lys Ser Tyr Leu Lys Ser Met Gly Val Asp Pro Asp Arg Val
100 105 110

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

Met Phe Arg Met Val Thr Pro Tyr Leu Ala Asp Leu Asp Ile Pro Met
180 185 190
Val Cys Ile Cys His Thr Tyr Asp Thr Gln Glu Met Tyr Ser Lys Lys
195 200 205
Val Ile Ser Gly Gly Thr Gly Leu Met Tyr Ser Ala Asp Thr Ala Ile
210 215 220
Ile Leu Gly Lys Gln Gln Val Lys Glu Gly Thr Glu Val Val Gly Tyr
225 230 235 240

Asp Phe Ile Met Asn Ile Glu Lys Ser Arg Phe Val Lys Glu Lys Ser
245 250 255
Lys Phe Lys Leu His Val Thr Tyr Glu Gly Gly Ile Ser Met Tyr Ser
260 265 270
Gly Leu Leu Asp Leu Ala Met Glu Met Asn Phe Val Gln Thr Pro Thr
275 280 285
Lys Gly Trp Arg Gly Arg Ala Phe Leu Asn Thr Glu Thr Gly Glu Leu
290 295 300

Glu Leu Glu Glu Lys Lys Trp Arg Glu Ser Glu Thr Asn Ser Ile Glu
305 310 315 320
Phe Trp Arg Pro Leu Phe Thr His Gln Pro Phe Leu Asp Ala Ile Gln
325 330 335
Asp Lys Tyr Arg Ile Pro Asp Lys Glu Ile Thr Asp Gly Ala Ala Leu
340 345 350
Glu Asp Leu Tyr Ser Thr Asp Glu Pro Glu Ser Asn Lys Ile Asp Leu
355 360 365

Asp Asp Asp Ile Pro Asp Asp Ile Gly Ile Asp Gln Asp Glu Glu Pro
370 375 380
Ile Met
385

<210> 30
<211> 335
<212> PRT
<213> Prochlorococcus phage P-SSM2
<400> 30
Met Asp Phe Leu Lys Glu Ile Val Lys Glu Ile Gly Asp Glu Tyr Thr
1 5 10 15
Gln Val Ala Ala Asp Ile Gln Glu Asn Glu Arg Phe Ile Asp Thr Gly
20 25 30
Ser Tyr Ile Phe Asn Gly Leu Val Ser Gly Ser Ile Phe Gly Gly Val

35 40 45
Ser Ser Ser Arg Ile Thr Ala Ile Ala Gly Glu Ser Ser Thr Gly Lys
50 55 60
Thr Tyr Phe Ser Leu Ala Val Val Lys Asn Phe Leu Asp Asn Asn Pro
65 70 75 80
Asp Gly Tyr Cys Leu Tyr Phe Asp Thr Glu Ala Ala Val Asn Lys Gly
85 90 95
Leu Leu Glu Ser Arg Gly Ile Asp Met Asn Arg Leu Val Val Val Asn

100 105 110
Val Val Thr Ile Glu Glu Phe Arg Ser Lys Ala Leu Arg Ala Val Asp

115 120 125

Ile Tyr Leu Lys Thr Ser Glu Glu Glu Arg Lys Pro Cys Met Phe Val

130 135 140

Leu Asp Ser Leu Gly Met Leu Ser Thr Glu Lys Glu Ile Arg Asp Ala

145 150 155 160

Leu Asp Asp Lys Gln Val Arg Asp Met Thr Lys Ser Gln Leu Val Lys

165 170 175

Gly Ala Phe Arg Met Leu Thr Leu Lys Leu Gly Gln Ala Asn Ile Pro

180 185 190

Leu Ile Val Thr Asn His Thr Tyr Asp Val Ile Gly Ser Tyr Val Pro

195 200 205

Thr Lys Glu Met Gly Gly Gly Ser Gly Leu Lys Tyr Ala Ala Ser Thr

210 215 220

Ile Ile Tyr Leu Ser Lys Lys Lys Glu Lys Asp Gln Lys Glu Val Ile

225 230 235 240

Gly Asn Leu Ile Lys Ala Lys Thr His Lys Ser Arg Leu Ser Lys Glu

245 250 255

Asn Lys Glu Val Lys Ile Arg Leu Tyr Tyr Asp Glu Arg Gly Leu Asp

260 265 270

Arg Tyr Tyr Gly Leu Leu Glu Leu Gly Glu Ile Gly Gly Met Trp Lys

275 280 285

Asn Val Ala Gly Arg Tyr Glu Met Asn Gly Lys Lys Ile Tyr Ala Lys

290 295 300

Glu Ile Leu Lys Asn Pro Thr Glu Tyr Phe Thr Asp Asp Ile Met Glu

305 310 315 320

Gln Leu Asp Asn Ile Ala Lys Glu His Phe Ser Tyr Gly Thr Asn

325 330 335

<210> 31

<211> 336

<212> PRT

<213> Prochlorococcus phage P-SSM4

<400> 31

Met Asn Phe Leu Lys Asp Ile Ala Lys Glu Ile Gly Asn Asp Tyr Ala
1 5 10 15

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

Asp Thr Gly Ser Tyr Ile Phe Asn Ala Leu Leu Ser Gly Ser Ile Tyr
35 40 45

Gly Gly Ile Pro Asn Asn Lys Ile Thr Ala Ile Ala Gly Glu Thr Ser
50 55 60

Thr Gly Lys Thr Phe Phe Cys Leu Gly Met Val Gln His Phe Leu Glu
65 70 75 80

Ser Asn Pro Asp Ala Gly Val Ile Tyr Phe Glu Ser Glu Ser Ala Ile
85 90 95

Ser Lys Gln Met Ile Glu Asp Arg Gly Ile Asp Ser Asn Arg Met Leu
100 105 110

Leu Val Pro Val Thr Thr Val Gln Glu Phe Arg Leu Gln Ala Ile Lys
115 120 125

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

Met Phe Val Leu Asp Ser Leu Gly Met Leu Ser Thr Ser Lys Glu Val
145 150 155 160

Glu Asp Ser Glu Ala Gly Lys Glu Thr Arg Asp Met Thr Arg Ala Gln
165 170 175

Val Val Lys Ser Ile Phe Arg Val Leu Thr Leu Lys Leu Gly Lys Ala
180 185 190

Asn Val Pro Leu Ile Val Thr Asn His Thr Tyr Asp Val Val Gly Ala
195 200 205

Tyr Ile Pro Thr Lys Glu Met Gly Gly Gly Ser Gly Leu Lys Tyr Ala
210 215 220

Ala Ser Thr Ile Val Tyr Leu Ser Lys Lys Lys Glu Lys Asn Gly Lys
225 230 235 240

Glu Val Val Gly Asn Ile Ile Lys Cys Lys Thr Ala Lys Ser Arg Leu

245 250 255

Thr Lys Glu Asn Ser Asp Val Lys Thr Arg Leu Tyr Tyr Asp Arg Gly

260 265 270

Leu Asp Arg Tyr Tyr Gly Leu Leu Glu Leu Gly Glu Lys His Gly Val

275 280 285

Phe Ser Arg Lys Gly Asn Arg Val Val Val Gly Asp Ser Ser Val Tyr

290 295 300

Pro Ser Ala Ile Leu Ala Asp Pro Asp Lys Tyr Phe Thr Glu Glu Leu

305 310 315 320

Met Glu Lys Leu Asp Glu Ala Ala Ala Lys Glu Phe Arg Tyr Gly Asn

325 330 335

<210> 32

<211> 343

<212> PRT

<213> Synechococcus phage S-PM2

<400> 32

Met Ser Phe Leu Asp Ser Val Ile Lys Asp Ser Lys Asn Glu Tyr Ala

1 5 10 15

Ala Phe Ala Ser Glu Gly Val Ala Ala Gly Asp Val Glu Ser Phe Val

20 25 30

Asp Thr Gly Ser Tyr Ile Phe Asn Ala Leu Val Ser Gly Ser Ile Phe

35 40 45

Gly Gly Ile Pro Ser Asn Lys Ile Thr Ala Leu Ala Gly Glu Ser Gly

50 55 60

Thr Gly Lys Thr Phe Phe Cys Leu Ser Val Val Arg Asn Phe Leu Asn

65 70 75 80

Thr Asp Pro Asp Ala Gly Val Ile Tyr Phe Glu Thr Glu Ser Ala Ile

85 90 95

Ser Lys Gln Met Ile Glu Ser Arg Gly Ile Asp Ser Thr Arg Met Ile

100 105 110

Ile Phe Pro Val Asp Thr Ile Glu Asp Phe Arg Thr Gln Ala Val Arg

115 120 125

Ile Ile Asp Lys Tyr Met Glu Gln Asn Lys Ser Glu Arg Lys Pro Leu

130 135 140

Met Phe Val Leu Asp Ser Leu Gly Met Leu Ala Thr Lys Lys Glu Val

145 150 155 160

Glu Asp Ala Ser Asn Asp Lys Gln Val Arg Asp Met Thr Lys Ala Gln

165 170 175

Ile Val Lys Ser Ala Phe Arg Ile Leu Thr Leu Lys Met Gly Lys Ala

180 185 190

Asn Ile Pro Met Leu Val Thr Asn His Thr Tyr Asp Val Val Gly Ser

195 200 205

Tyr Val Pro Thr Lys Glu Met Gly Gly Gly Ser Gly Leu Lys Tyr Ser

210 215 220

Ala Ser Thr Ile Val Tyr Leu Gly Lys Lys Lys Glu Lys Asp Gly Thr

225 230 235 240

Asp Leu Val Gly Asn Ile Ile Lys Cys Glu Ala Lys Lys Ser Arg Leu

245 250 255

Thr Arg Glu Gly Ser Lys Val Lys Thr Arg Leu Phe Phe Asp Gln Arg

260 265 270

Gly Leu Glu Arg Tyr Tyr Gly Met Leu Glu Leu Gly Glu Arg Ala Gly

275 280 285

Leu Trp Lys Asn Thr Ala Gly Arg Tyr Glu Ile Asn Gly Lys Lys Val

290 295 300

Tyr Gly Lys Gln Ile Leu Ala Asn Pro Asp Glu Phe Phe Thr Glu Glu

305 310 315 320

Ile Leu Gln Glu Leu Asp Lys Gln Ala Gln Arg Glu Phe Leu Tyr Gly

325 330 335

Ala Ser Asp Asp Gly Glu Asp

340

<210> 33

<211> 393

<212> PRT

<213> Enterobacteria phage RB32

<400> 33

Met Ser Ile Ala Asp Leu Lys Ser Arg Leu Ile Lys Ala Ser Thr Ser
1 5 10 15
Lys Met Thr Ala Glu Leu Thr Thr Ser Lys Phe Phe Asn Glu Lys Asp

20 25 30
Val Ile Arg Thr Lys Ile Pro Met Leu Asn Ile Ala Ile Ser Gly Ala
35 40 45
Ile Asp Gly Gly Met Gln Ser Gly Leu Thr Ile Phe Ala Gly Pro Ser
50 55 60
Lys His Phe Lys Ser Asn Met Ser Leu Thr Met Val Ala Ala Tyr Leu
65 70 75 80
Asn Lys Tyr Pro Asp Ala Val Cys Leu Phe Tyr Asp Ser Glu Phe Gly

85 90 95
Ile Thr Pro Ala Tyr Leu Arg Ser Met Gly Val Asp Pro Glu Arg Val
100 105 110
Ile His Thr Pro Ile Gln Ser Val Glu Gln Leu Lys Ile Asp Met Val
115 120 125
Asn Gln Leu Glu Ala Ile Glu Arg Gly Glu Lys Val Ile Val Phe Ile
130 135 140
Asp Ser Ile Gly Asn Met Ala Ser Lys Lys Glu Thr Glu Asp Ala Leu

145 150 155 160
Asn Glu Lys Ser Val Ala Asp Met Thr Arg Ala Lys Ser Leu Lys Ser
165 170 175
Leu Phe Arg Ile Val Thr Pro Tyr Phe Ser Ile Lys Asn Ile Pro Cys
180 185 190
Val Ala Val Asn His Thr Ile Glu Thr Ile Glu Met Phe Ser Lys Thr
195 200 205
Val Met Thr Gly Gly Thr Gly Val Met Tyr Ser Ala Asp Thr Val Phe

210 215 220
Ile Ile Gly Lys Arg Gln Ile Lys Asp Gly Ser Asp Leu Gln Gly Tyr

225 230 235 240

Gln Phe Val Leu Asn Val Glu Lys Ser Arg Thr Val Lys Glu Lys Ser

245 250 255

Lys Phe Lys Ile Asp Val Lys Phe Asp Gly Gly Ile Asp Pro Tyr Ser

260 265 270

Gly Leu Leu Asp Met Ala Leu Glu Leu Gly Phe Val Val Lys Pro Lys

275 280 285

Asn Gly Trp Tyr Ala Arg Glu Phe Leu Asp Glu Glu Thr Gly Glu Met

290 295 300

Ile Arg Glu Glu Lys Ser Trp Arg Ala Lys Asp Thr Asn Cys Thr Thr

305 310 315 320

Phe Trp Gly Pro Leu Phe Lys His Gln Pro Phe Arg Asp Ala Ile Lys

325 330 335

Arg Ala Tyr Gln Leu Gly Ala Ile Asp Ser Asn Glu Ile Val Glu Ala

340 345 350

Glu Val Asp Glu Leu Ile Asn Ser Lys Val Glu Lys Phe Lys Ser Pro

355 360 365

Glu Ser Lys Ser Lys Ser Ala Ala Asp Leu Glu Thr Asp Leu Glu Gln

370 375 380

Leu Ser Asp Met Glu Glu Phe Asn Glu

385 390

<210> 34

<211> 366

<212> PRT

<213> Vibrio phage nt-1

<400> 34

Met Ser Asp Leu Leu Lys Ser Leu Lys Lys Ser Ser Thr Ser Gly Tyr

1 5 10 15

Ala His Val Leu Ser Glu Ser Gln Phe Met Phe Glu Lys Asp His Thr

20 25 30

Arg Thr Tyr Val Pro Ala Ile Asn Ile Ala Phe Ser Gly Glu Val Asp

35 40 45

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

50 55 60

Phe Lys Ser Asn Leu Gly Leu Val Gly Val Ala Ala Tyr Leu Lys Lys

65 70 75 80

Tyr Pro Glu Ala Ile Cys Val Phe Ile Asp Thr Glu Phe Gly Ile Thr

85 90 95

Pro Ser Tyr Leu Lys Ser Gln Gly Val Asp Pro Glu Arg Val Leu His

100 105 110

Ile Gln Cys Glu Ser Val Glu Arg Met Lys Phe Glu Met Ala Asn Gln

115 120 125

Leu Lys Asp Leu Ala Glu Arg Lys Arg Ala Lys Lys Ala Gly Glu Glu

130 135 140

Pro Asp Arg Val Val Phe Phe Ile Asp Ser Val Gly Asn Val Ala Ser

145 150 155 160

Ala Lys Glu Ile Asp Asp Ala Gln Asn Glu Lys Ser Val Ala Asp Met

165 170 175

Ser Arg Ala Lys Gln Leu Lys Ser Leu Phe Arg Ile Ile Thr Pro Tyr

180 185 190

Phe Thr Met Leu Asp Ile Pro Cys Ile Ala Ile Asn His Thr Tyr Gln

195 200 205

Thr Gln Glu Met Tyr Ser Lys Thr Val Met Ser Gly Gly Thr Gly Ile

210 215 220

Met Tyr Ser Ala Asp Thr Val Ile Ile Leu Gly Lys Gln Gln Glu Lys

225 230 235 240

Asp Gly Lys Glu Ile Ile Gly Tyr His Phe Ile Met Asn Ile Glu Lys

245 250 255

Ser Arg Phe Val Lys Glu Lys Met Lys Val Lys Leu Thr Val Thr Tyr

260 265 270

Glu His Gly Ile Asp Gln Phe Ser Gly Leu Leu Asp Ile Ala Leu Gln

275 280 285

Thr Gly His Val Val Lys Pro Ser Asn Gly Trp Tyr Gln Arg Ala Phe

290 295 300

Ile Asp Glu Glu Thr Gly Glu Ile Glu Ile Glu Glu Lys Lys Tyr Arg

305 310 315 320

Ala Lys Glu Thr Gln Thr Leu Ser Phe Trp Lys Glu Ile Ile Asn Ser

325 330 335

Pro Thr Phe Lys Thr Gly Val Lys Arg Leu Tyr Cys Leu Gly Gln Leu

340 345 350

Asp Glu Ser Glu Leu Leu Asp Glu Val Asp Ser Leu Phe Asp

355 360 365

<210> 35

<211> 381

<212> PRT

<213> Enterobacteria phage RB16

<400> 35

Met Ser Asn Lys Ala Leu Leu Lys Lys Leu Ile Lys Asn Ser Asn Ser

1 5 10 15

Gln Ser Ala Ser Ile Leu Ser Glu Ser Asp Val Phe Asn Asn Ile Thr

20 25 30

Lys Thr Arg Thr Arg Val Pro Ile Leu Asn Leu Val Leu Ser Gly Ala

35 40 45

Phe Asp Gly Gly Leu Thr Ser Gly Leu Thr Leu Ile Ala Gly Pro Ser

50 55 60

Lys His Phe Lys Ser Asn Leu Gly Leu Val Ala Val Ala Ala Tyr Leu

65 70 75 80

Lys Ala Asn Glu Asp Ala Val Cys Leu Phe Tyr Asp Ser Glu Lys Gly

85 90 95

Val Thr Lys Ser Tyr Leu Lys Ser Met Gly Val Asp Pro Asp Arg Val

100 105 110

Val Tyr Thr Arg Ile Thr Thr Val Glu Gln Leu Arg Asn Asp Val Val

115 120 125

Ser Gln Leu Asp Ala Leu Glu Arg Gly Asp Lys Val Ile Ile Phe Val

130 135 140

Asp Ser Val Gly Asn Thr Ala Ser Lys Lys Glu Leu Lys Asp Ala Leu
145 150 155 160
Glu Asp Asn Asp Lys Gln Asp Met Thr Arg Ala Lys Ala Leu Lys Gly
165 170 175
Met Phe Arg Met Val Thr Pro Tyr Leu Ala Asp Ile Asp Ile Pro Met
180 185 190
Val Cys Ile Cys His Thr Tyr Asp Thr Gln Glu Met Tyr Ser Lys Lys
195 200 205

Val Ile Ser Gly Gly Thr Gly Leu Met Tyr Ser Ala Asp Thr Ala Ile
210 215 220
Ile Leu Gly Lys Gln Gln Val Lys Glu Gly Thr Glu Val Val Gly Tyr
225 230 235 240
Asp Phe Ile Met Asn Val Glu Lys Ser Arg Phe Val Lys Glu Lys Ser
245 250 255
Lys Phe Lys Leu His Val Thr Tyr Glu Gly Gly Ile Ser Met Phe Ser
260 265 270

Gly Leu Leu Asp Leu Ala Met Glu Met Asn Phe Val Gln Thr Pro Thr
275 280 285
Lys Gly Trp Arg Gly Arg Ala Phe Leu Asn Thr Glu Thr Gly Glu Leu
290 295 300
Glu Leu Glu Glu Lys Lys Trp Arg Glu Ala Glu Thr Asn Cys Ile Glu
305 310 315 320
Phe Trp Lys Pro Leu Phe Lys His Gln Pro Phe Ile Asp Ala Ile Gln
325 330 335

Asp Lys Tyr Arg Ile Pro Asp Lys Glu Ile Thr Asp Gly Ala Ala Leu
340 345 350
Glu Asp Leu Tyr Ser Asp Asp Val Val Glu Ser Asn Lys Val Asp Phe
355 360 365
Asp Asp Asp Ile Pro Asp Asp Val Asp Leu Met Glu Glu
370 375 380