(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2017년11월20일
(11) 등록번호 10-1794298
(24) 등록일자 2017년10월31일

(51) 국제특허분류(Int. Cl.)
C12N 15/10 (2017.01) *C12N 15/63* (2006.01)
C12Q 1/68 (2006.01) *C40B 40/06* (2006.01)
(21) 출원번호 10-2011-7014144
(22) 출원일자(국제) 2009년11월19일
심사청구일자 2014년07월30일
(85) 번역문제출일자 2011년06월20일
(65) 공개번호 10-2011-0102351
(43) 공개일자 2011년09월16일
(86) 국제출원번호 PCT/US2009/065048
(87) 국제공개번호 WO 2010/059763
국제공개일자 2010년05월27일
(30) 우선권주장
61/116,109 2008년11월19일 미국(US)
61/162,230 2009년03월20일 미국(US)
(56) 선행기술조사문헌
JP2003533979 A
US05976846 A*
WO2009048885 A2
WO2008045380 A2
*는 심사관에 의하여 인용된 문헌

(73) 특허권자
아미리스 인코퍼레이티드
미국, 캘리포니아주 94608, 에머리빌, 스위트 100, 5885 홀리스 스트리트
(72) 발명자
서버, 자크
미국, 캘리포니아주 94608, 에머리빌, 스위트 100, 5885 홀리스 스트리트
로우, 레이몬드
미국, 캘리포니아주 94608, 에머리빌, 스위트 100, 5885 홀리스 스트리트
(뒷면에 계속)
(74) 대리인
특허법인필앤온지

전체 청구항 수 : 총 35 항 심사관 : 김정희

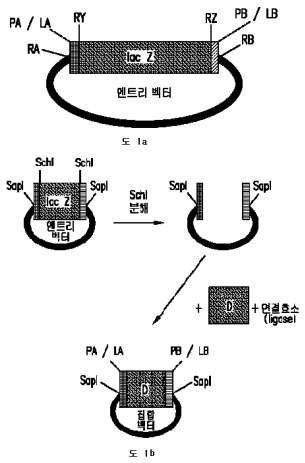
(54) 발명의 명칭 **폴리뉴클레오타이드의 집합체 조성물 및 이에 대한 방법**

(57) 요 약

본 발명은 수많은 폴리뉴클레오타이드 성분들로부터 하나 이상의 집합된(assembled) 폴리뉴클레오타이드의 조성물 및 빠른 집합 방법을 제공한다. 본 발명의 방법은 어닐링될 수 있는 링커 서열, 어닐링될 수 있는 링커 서열 쌍 LA 및 LB, 또는 어닐링될 수 있는 링커 서열/프라이머 결합 절편 쌍 LA과 PB 또는 PA과 LB의 측면에 배치된

(뒷면에 계속)

대 표 도



DNA 절편 D를 포함하는 원형 핵산 벡터를 이용한다. 집합될 DNA 절편(segment)을 포함하는 수많은 벡터의 제한 엔도뉴클레아제 분해는 PA-D-LB, LA-D-LB 및 LA-D-PB 또는 D-LB, LA-D-LB 및 LA-D 성분들을 포함하는 수많은 DNA 절편들을 형성한다. 어닐링될 수 있는 링커 서열 LA 및 LB의 서열은 DNA 절편에 상보적인 말단을 제공하며, 이는 숙주 세포 매개 상동 재조합(homologous recombination)에서 사용되거나 또는, 다양한 DNA 절편(segment)의 하나 이상의 집합된(assembled) 폴리뉴클레오타이드로의 정렬된 집합체에 대한 폴리머라제 사이클링 집합 반응에서 절편 PA와 PB를 결합시키는 프라이머와 함께 사용된다.

(72) 발명자

우버사스, 제프리, 에이.

미국, 캘리포니아주 94608, 에머리빌, 스위트 100, 5885 홀리스 스트리트

찬드란, 선일, 에스.

미국, 캘리포니아주 94608, 에머리빌, 스위트 100, 5885 홀리스 스트리트

딘, 에릭, 제데디아

미국, 캘리포니아주 94608, 에머리빌, 스위트 100, 5885 홀리스 스트리트

플랫, 대런, 엠.

미국, 캘리포니아주 94608, 에머리빌, 스위트 100, 5885 홀리스 스트리트

타케오카, 케네스, 토시키

미국, 캘리포니아주 94608, 에머리빌, 스위트 100, 5885 홀리스 스트리트

명 세 서

청구범위

청구항 1

(a) 각각, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 그룹 D_0로부터 선택된 임의의 DNA 절편, 어닐링될 수 있는 링커 서열 LB_0 및 두번째 제한 부위 RB_0를 포함하며 원형인, 하나 이상의 첫번째 핵산 분자;

(b) 각각, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, 그룹 D_n으로부터 선택된 임의의 DNA 절편, 두번째 어닐링될 수 있는 링커 서열 LB_n 및 두번째 제한 부위 RB_n을 포함하며 원형인, 하나 이상의 중간 핵산 분자(상기 n은 1 내지 중간 핵산 분자의 개수까지의 정수를 나타냄); 및

(c) 각각, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, 그룹 D_m으로부터 선택된 임의의 DNA 절편, 및 두번째 제한 부위 RB_m를 포함하며 원형인, 하나 이상의 마지막 핵산 분자(상기 m은 중간 핵산 분자의 개수보다 큰 정수를 나타냄);를 포함하며,

상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이이며;

여기에서, 제한 부위 RA_0부터 RB_m까지의 절단 및 절단된(excised) 선형 핵산 분자의 변성에 있어서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$는 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 하이브리드할 수 있으며 상기 절단된 선형 핵산 분자에 대한 상보적인 폴리뉴클레오타이드의 합성을 위한 개시 지점으로 기능할 수 있으며, 상기 n은 1부터 (m-1)까지 변하는 정수이며, 상기 p는 1 내지 m의 정수를 나타내며, 및 상기 각각의 그룹 D_0,... D_n,... 및 D_m은, 서로 독립적으로 하나 이상의 DNA 절편(segment)으로 구성된 것을 특징으로 하는 조성물.

청구항 2

제 1항에 있어서, 상기 하나 이상의 첫번째 핵산 분자는 각각, 그룹 D_0로부터 선택된 DNA 절편에 대하여 5'에 위치한 프라이머 결합 절편 PA를 더 포함하며, 하나 이상의 마지막 핵산 분자는 각각, 그룹 D_m으로부터 선택된 DNA 절편에 대하여 3'에 위치한 프라이머 결합 절편 PB를 더 포함하는 것을 특징으로 하는 조성물.

청구항 3

제 1항 또는 제 2항에 있어서, 상기 제한 부위 RA_0부터 RB_m까지의 절단 및 절단된(excised) 선형 핵산 분자의 변성에 있어서, 조성물에서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 선택적으로 하이브리드할 수 있는 것을 특징으로 하는 조성물.

청구항 4

제 1항 또는 제 2항에 있어서, 상기 각각의 어닐링될 수 있는 링커(annealable linker) $LB_{(p-1)}$은 어닐링될 수 있는 링커 서열 LA_p 또는 이의 상보적 서열(complement)과 서열에서 동일한 것을 특징으로 하는 조성물.

청구항 5

제 1항 또는 제 2항에 있어서, 하나의 첫번째 핵산 분자 및 마지막 핵산 분자를 포함하는 것을 특징으로 하는 조성물.

청구항 6

제 1항 또는 제 2항에 있어서, 상기 각각의 제한 부위 RA_0부터 RB_m는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있는 것을 특징으로 하는 조성물.

청구항 7

제 1항 또는 제 2항에 있어서, 상기 제한 부위 RA_0부터 RB_m는 동일한 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있는 것을 특징으로 하는 조성물.

청구항 8

제 1항 또는 제 2항에 있어서, 상기 제한 부위 RA_0부터 RB_m는 SapI 또는 LguI 제한 엔도뉴클레아제에 의해 절단될 수 있는 것을 특징으로 하는 조성물.

청구항 9

제1항에 있어서, 상기 2개 이상의 어닐링될 수 있는 링커 서열은 각각 적어도 60℃의 융점을 갖는 것을 특징으로 하는 조성물.

청구항 10

제 1항 또는 제 2항에 있어서, 상기 2개 이상의 어닐링될 수 있는 링커 서열은 각각, 서로 독립적으로, 적어도 70%의 G-C의 염기쌍의 함량과 적어도 70℃의 융점을 갖는 것을 특징으로 하는 조성물.

청구항 11

제 1항 또는 제 2항에 있어서, 상기 2개 이상의 어닐링될 수 있는 링커 서열은 각각, 서로 독립적으로, 적어도 30%의 A-T의 염기쌍의 함량과 적어도 65℃의 융점을 가지며, 서열 모티프 5'ANNNNNNNANNNAANTANNTTNANA-3'를 포함하며, 상기 A는 아데닌을 나타내며, N은 임의의 뉴클레오타이드를 나타내며, T는 티민을 나타내는 것을 특징으로 하는 조성물.

청구항 12

제 1항 또는 제 2항에 있어서, 상기 2개 이상의 어닐링될 수 있는 링커 서열은 서열번호 1 내지 8 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된 서열을 갖는 것을 특징으로 하는 조성물.

청구항 13

제 1항 또는 제 2항에 있어서, 상기 각각의 어닐링될 수 있는 링커 서열은 서열번호 1 내지 8 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된 서열을 갖는 것을 특징으로 하는 조성물.

청구항 14

제 2항에 있어서, 상기 각각의 프라이머 결합 절편은 서열번호 9 내지 10 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된 서열을 갖는 것을 특징으로 하는 조성물.

청구항 15

제 1항 또는 제 2항에 있어서, 상기 제한 부위 RB_0부터 RBm까지 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제를 더 포함하는 것을 특징으로 하는 조성물.

청구항 16

제 8항에 있어서, SapI 또는 LguI 제한 엔도뉴클레아제를 더 포함하는 것을 특징으로 하는 조성물.

청구항 17

제한 부위 RA_0부터 RB_m까지 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제를 갖는 제 1항 또는 제 2항의 조성물을 분해함으로써 형성된 다수의 선형 핵산 분자를 포함하는 조성물.

청구항 18

제 17항에 있어서, 상기 각각의 선형 핵산 분자는 접착말단(sticky-end)을 포함하는 것을 특징으로 하는

조성물.

청구항 19

제 2항에 있어서, 프라이머 결합 절편 PA 또는 이의 상보적 서열(complement)에 상보적인 첫번째 프라이머, 및 프라이머 결합 절편 PB 또는 이의 상보적 서열(complement)에 상보적인 두번째 프라이머를 더 포함하는 것을 특징으로 하는 조성물.

청구항 20

제 1항 또는 제 2항에 있어서, DNA 폴리머라제를 더 포함하는 것을 특징으로 하는 조성물.

청구항 21

SapI 또는 LguI 제한 엔도뉴클레아제를 갖는 제8항의 조성물을 분해함으로써 형성된 다수의 선형 핵산 분자를 포함하는 조성물.

청구항 22

(a) 5'에서 3'의 방향으로, 제한 부위 RA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

(b) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

(c) 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

(d) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터; 및

(e) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 프라이머 결합 절편 PB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;로부터 선택되며,

상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이이며;

상기 각각의 제한 부위들 RA 및 RB는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있는 것을 특징으로 하는 벡터.

청구항 23

(a) 5'에서 3'의 방향으로, 제한 부위 RA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

(b) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

(c) 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

(d) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터; 및

(e) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 프라이머 결합 절편 PB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;로부터 선택되며,

상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이이며;

상기 각각의 제한 부위들 RA 및 RB는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있는 것을 특징으로 하는 벡터.

청구항 24

(a) 제 23항의 하나 이상의 벡터;

(b) 제한 부위들 RA 및 RB를 절단할 수 있는 하나 이상의 IIS형 제한 엔도뉴클레아제; 및

(c) 제한 부위들 RY 및 RZ를 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제;를 포함하는 키트.

청구항 25

(a) 5'에서 3'의 방향으로, 제한 부위 RA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

(b) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

(c) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터; 중 적어도 하나를 포함하며,

상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이이며;

상기 각각의 제한 부위들 RA 및 RB는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있는 것을 특징으로 하는 핵산 분자의 라이브러리.

청구항 26

(a) 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

(b) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

(c) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 프라이머 결합 절편 PB 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터; 중 적어도 하나를 포함하며,

상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이이며;

상기 각각의 제한 부위들 RA 및 RB는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있는 것을 특징으로 하는 핵산 분자의 라이브러리.

청구항 27

(a) 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 그룹 D_0로부터 선택된 임의의 DNA 절편, 어닐링될 수 있는 링커 서열 LB_0 및 두번째 제한 부위 RB_0를 포함하며 원형인 첫번째 핵산 분자;

(b) 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, 그룹 D_n으로부터 선택된 임의의 DNA 절편, 두번째 어닐링될 수 있는 링커 서열 LB_n 및 두번째 제한 부위 RB_n을 포함하며 원형인 중간 핵산 분자(상기 n은 1 내지 중간 핵산 분자의 개수까지의 정수를 나타냄); 및

(c) 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, 그룹 D_m으로부터 선택된 임의의 DNA 절편, 및 두번째 제한 부위 RB_m를 포함하며 원형인 마지막 핵산 분자(상기 m은 중간 핵산 분자의 개수보다 큰 정수를 나타냄); 중 적어도 하나를 포함하며,

상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이이며;

상기 각각의 제한 부위 RA_0부터 RB_m까지는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있으며,

여기에서, 제한 부위 RA_0부터 RB_m까지의 절단 및 절단된(excised) 선형 핵산 분자의 변성에 있어서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$는 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 하이브리드할 수 있으며 상기 절단된 선형 핵산 분자에 대해 상보적인 폴리뉴클레오타이드의 합성을 위한 개시 지점으로 기능

할 수 있으며, 상기 n은 1부터 (m-1)까지 변하는 정수이며, 상기 p는 1 내지 m의 정수를 나타내며, 및 상기 각각의 그룹 D_0,... D_n,... 및 D_m은 서로 독립적으로, 하나 이상의 DNA 절편(segment)으로 구성된 것을 특징으로 하는 핵산 분자의 라이브러리.

청구항 28

(a) (i), (ii) 및 (iii)의 핵산 분자로부터 절단된(excised) 폴리뉴클레오티드를 형성하기 위하여 하나 이상의 IIS형 제한 엔도뉴클레아제를 사용하여 조성물을 분해하는 단계;

상기 조성물은,

(i) 각각, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 그룹 PA로부터 선택된 임의의 프라이머 결합 절편, 그룹 D_0으로부터 선택된 임의의 DNA 절편, 어닐링될 수 있는 링커 서열 LB_0, 및 두번째 제한 부위 RB_0를 포함하며 원형인, 하나 이상의 첫번째 핵산 분자;

(ii) 각각, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, 그룹 D_n으로부터 선택된 임의의 DNA 절편, 두번째 어닐링될 수 있는 링커 서열 LB_n, 및 두번째 제한 부위 RB_n을 포함하며 원형인, 하나 이상의 중간 핵산 분자(상기 n은 1 내지 중간 핵산 분자의 개수까지의 정수를 나타냄); 및

(iii) 각각, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, 그룹 D_m으로부터 선택된 임의의 DNA 절편, 그룹 PB로부터 선택된 임의의 프라이머 결합 절편, 및 두번째 제한 부위 RB_m을 포함하며 원형인, 하나 이상의 마지막 핵산 분자(상기 m은 중간 핵산 분자의 개수보다 큰 정수를 나타냄);를 포함하며, 여기에서, 제한 부위 RA_0부터 RB_m까지의 절단 및 절단된(excised) 선형 핵산 분자의 변성에 있어서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$는 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 하이브리드할 수 있으며 상기 절단된 선형 핵산 분자에 대한 상보적인 폴리뉴클레오타이드의 합성을 위한 개시 지점으로 기능할 수 있으며, 상기 n은 1부터 (m-1)까지 변하는 정수이며, 상기 p는 1 내지 m의 정수를 나타내며, 및 상기 각각의 그룹 D_0,... D_n,... 및 D_m은 하나 이상의 DNA 절편(segment)으로 구성되며; 및

(b) 상기 조성물을 DNA 폴리머라제, 데옥시리보뉴클레오사이드 트리포스페이트 및 하나 이상의 첫번째 프라이머 및 하나 이상의 두번째 프라이머와 접촉시키는 단계;

상기 각각의 첫번째 프라이머는 그룹 PA로부터 선택된 상기 프라이머 결합 절편 중 하나에 하이브리드 될 수 있으며, 상기 각각의 두번째 프라이머는 그룹 PB로부터 선택된 상기 프라이머 결합 절편 중 하나에 하이브리드 될 수 있으며; 및 상기 성분 조성물을 폴리머라제 사슬 반응을 겪게 하는 단계;를 포함하며,

여기에서, 5'에서 3'의 방향으로, 각각의 그룹 D_0,... D_n,... 및 D_m으로부터 선택된 하나의 DNA 절편을 포함하는 폴리뉴클레오타이드가 상기 어닐링될 수 있는 링커 서열에 의해 집합되며; 상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이인 것을 특징으로 하는

상기 (i), (ii) 및 (iii)의 핵산 분자로부터 절단된(excised) 폴리뉴클레오티드의 다수로부터 어닐링될 수 있는 링커 서열에 의해 집합된 폴리뉴클레오티드를 형성하는 방법.

청구항 29

(a) 제 28항에 따라 어닐링될 수 있는 링커에 의해 집합된 하나 이상의 폴리뉴클레오타이드로 숙주 세포를 형질변환하는 단계; 및

(b) 어닐링될 수 있는 링커에 의해 집합된 하나 이상의 폴리뉴클레오타이드를 포함하는 숙주 세포를 선택하는 단계;를 포함하는 것을 특징으로 하는, 폴리뉴클레오타이드를 포함하는 숙주 세포의 형성 방법.

청구항 30

(a) (i), (ii), 및 (iii) 또는 (i) 및 (iii)을 포함하는 조성물을 사용하여 숙주 세포를 형질변환하는 단계;

(i) 각각, 5'에서 3'의 방향으로, 그룹 D_0으로부터 선택된 임의의 DNA 절편 및 어닐링될 수 있는 링커

서열 LB_0를 포함하는, 하나 이상의 첫번째 선형 핵산 분자;

(ii) 각각, 5'에서 3'의 방향으로, 첫번째 어닐링될 수 있는 링커 서열 LA_n, 그룹 D_n으로부터 선택된 임의의 DNA 절편 및 두번째 어닐링될 수 있는 링커 서열 LB_n을 포함하는, 하나 이상의 중간 선형 핵산 분자(상기 n은 1 내지 중간 핵산 분자의 개수까지의 정수를 나타냄); 및

(iii) 각각, 5'에서 3'의 방향으로, 어닐링될 수 있는 링커 서열 LA_m 및 그룹 D_m으로부터 선택된 임의의 DNA 절편을 포함하는, 하나 이상의 마지막 선형 핵산 분자(상기 m은 중간 핵산 분자의 개수보다 큰 정수를 나타냄);를 포함하며,

여기에서, 상기 n은 1부터 (m-1)까지 변하는 정수이며,

상기 각각의 그룹 D_0,... D_n,... 및 D_m은 하나 이상의 DNA 절편(segment)으로 구성되며,

상기 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은 $LB_{(p-1)}$과 LA_p 사이에서 숙주 세포 매개 상동 재조합(homologous recombination)을 개시하는데 충분한 길이의 어닐링될 수 있는 링커 서열 LA_p를 갖는 상동(homology) 영역을 포함하며, 상기 p는 1 내지 m의 정수를 나타내며, 상기 상동 재조합은 상기 어닐링될 수 있는 링커에 의해 폴리뉴클레오타이드의 집합체를 유발하며; 및

(b) 5'에서 3'의 방향으로, 각각의 그룹 D_0,... D_n,... 및 D_m으로부터 선택된 하나의 DNA 절편을 포함하는 상기 어닐링될 수 있는 링커에 의해 집합된 폴리뉴클레오타이드를 포함하는 숙주 세포를 선택하는 단계;를 포함하며,

상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이인 것을 특징으로 하는 폴리뉴클레오타이드를 포함하는 숙주 세포의 형성 방법.

청구항 31

제 30항의 방법에 의해 형성된 폴리뉴클레오타이드를 포함하는 숙주 세포.

청구항 32

(a) (a1) 하나 이상의 IIS형 제한 엔도뉴클레아제를 사용하여 (i), (ii), 및 (iii)을 포함하는 조성물 내 (i), (ii) 및 (iii)의 핵산 분자를 분해하거나, 또는 (a2) 하나 이상의 IIS형 제한 엔도뉴클레아제를 사용하여 (i) 및 (iii)을 포함하는 조성물 내 (i) 및 (iii)의 핵산 분자를 분해하는 단계:

(i) 각각, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 프라이머 결합 절편 PA, 그룹 D_0로부터 선택된 임의의 DNA 절편, 어닐링될 수 있는 링커 서열 LB_0, 및 두번째 제한 부위 RB_0을 포함하며 원형인, 하나 이상의 첫번째 핵산 분자;

(ii) 각각, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, 그룹 D_n으로부터 선택된 임의의 DNA 절편, 두번째 어닐링될 수 있는 링커 서열 LB_n, 및 두번째 제한 부위 RB_n을 포함하며 원형인, 하나 이상의 중간 핵산 분자(상기 n은 1부터 중간 핵산 분자의 개수까지의 정수를 나타냄); 및

(iii) 각각, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m. 어닐링될 수 있는 링커 서열 LA_m, 그룹 D_m으로부터 선택된 임의의 DNA 절편, 프라이머 결합 절편 PB, 및 두번째 제한 부위 RB_m을 포함하며 원형인, 하나 이상의 마지막 핵산 분자(상기 m은 중간 핵산 분자의 개수를 초과하는 정수임);를 포함하며, 여기에서, 제한 부위 RA_0부터 RB_m까지의 절단 및 절단된(excised) 선형 핵산 분자의 변성에 있어서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$는 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 하이브리드할 수 있으며, 상기 p는 1 내지 m의 정수를 나타내며, 상기 n은 1부터 (m-1)까지 변하는 정수이며, 및 상기 각각의 그룹 D_0,... D_n,... 및 D_m은 하나 이상의 DNA 절편(segment)으로 구성되며;

상기 하나 이상의 IIS형 제한 엔도뉴클레아제는 제한 부위 RA_0부터 RB_m까지 절단할 수 있으며;

상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이이며; 및

(b) 상기 (a) 단계로부터 얻은 분해된 조성물을 사용하여 숙주 세포를 형질변환하는 단계; 상기 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은 $LB_{(p-1)}$과 LA_p 사이에 숙주 세포 매개 상동 재조합(homologous recombination)을 개시하는데 충분한 길이의 어닐링될 수 있는 링커 서열 LA_p를 갖는 상동(homology) 영역을 포함하며, 상기 상동 재조합은 상기 어닐링될 수 있는 링커에 의해 폴리뉴클레오타이드의 집합체를 유발하며, 상기 p는 1 내지 m의 정수를 나타내며; 및

(c) 5'에서 3'의 방향으로, 각각의 그룹 $D_0,...D_n,...D_m$으로부터 선택된 하나의 DNA 절편을 포함하는, 상기 어닐링될 수 있는 링커에 의해 집합된 폴리뉴클레오타이드를 포함하는 숙주 세포를 선택하는 단계;를 포함하는, 폴리뉴클레오타이드를 포함하는 숙주 세포를 형성하는 방법.

청구항 33

제 32항의 방법에 의해 형성된 폴리뉴클레오타이드를 포함하는 숙주 세포.

청구항 34

삭제

청구항 35

삭제

청구항 36

(a) 하기의 벡터에서 선택된 벡터를, 제한 부위들 RY 및 RZ을 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제를 사용하여 분해하는 단계:

(i) 5'에서 3'의 방향으로, 제한 부위 RA, 제한 부위 RY, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는, 원형의 폴리뉴클레오타이드로 구성된 벡터;

(ii) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, 제한 부위 RZ, 및 제한 부위 RB를 포함하는, 원형의 폴리뉴클레오타이드로 구성된 벡터;

(iii) 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA, 제한 부위 RY, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB, 및 제한 부위 RB를 포함하는, 원형의 폴리뉴클레오타이드로 구성된 벡터;

(iv) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는, 원형의 폴리뉴클레오타이드로 구성된 벡터; 및

(v) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, 제한 부위 RZ, 프라이머 결합 절편 PB 및 제한 부위 RB를 포함하는, 원형의 폴리뉴클레오타이드로 구성된 벡터;

상기 분해하는 단계는 선형의 벡터를 형성하며;

상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이이며; 및

(b) 연결효소(ligase)의 존재하에서 선형의 벡터를 선형의 DNA 절편 D와 접촉시키는 단계를 포함하며, 상기 접촉시키는 단계는 DNA 절편 D를 포함하는 원형의 벡터를 형성하며, 상기 제한부위들 RA 및 RB는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있는 것을 특징으로 하는 벡터를 제조하는 방법.

청구항 37

(a) 5'에서 3'의 방향으로, 제한 부위 RA, 제한 부위 RY, 제한 부위 RZ, 및제한 부위 RB를 포함하는, 원형의 폴리뉴클레오타이드로 구성된 벡터를, 제한 부위들 RY 및 RZ를 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제를 사용하여 분해하는 단계;

상기 분해하는 단계는 선형 벡터를 형성하며; 및

(b) 연결효소(ligase)의 존재 하에서, 선형 벡터를 하기의 것들에서 선택된 선형 폴리뉴클레오타이드와 접촉시

키는 단계:

(i) 5'에서 3'의 방향으로, DNA 절편 D, 및 어닐링될 수 있는 링커 서열 LB를 포함하는 선형 폴리뉴클레오타이드;

(ii) 5'에서 3'의 방향으로, 어닐링될 수 있는 링커 서열 LA, 및 DNA 절편을 포함하는 선형 폴리뉴클레오타이드;

(iii) 5'에서 3'의 방향으로, 프라이머 결합 절편 PA, DNA 절편 D, 및 어닐링될 수 있는 링커 서열 LB를 포함하는 선형 폴리뉴클레오타이드;

(iv) 5'에서 3'의 방향으로, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 및 어닐링될 수 있는 링커 서열 LB를 포함하는 선형 폴리뉴클레오타이드; 및

(v) 5'에서 3'의 방향으로, 어닐링될 수 있는 링커(annealable linker) LA, DNA 절편 D, 및 프라이머 결합 절편 PB를 포함하는 선형 폴리뉴클레오타이드;

상기 접촉시키는 단계는 DNA 절편 D를 포함하는 원형의 벡터를 형성하며, 상기 각각의 제한 부위 RA 및 RB는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있으며, 상기 각각의 어닐링될 수 있는 링커 서열은 적어도 24개 뉴클레오티드의 길이인 것을 특징으로 하는 벡터를 제조하는 방법.

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

청구항 50

삭제

청구항 51

삭제

청구항 52

삭제

청구항 53

삭제

청구항 54

삭제

청구항 55

삭제

청구항 56

삭제

청구항 57

삭제

청구항 58

삭제

청구항 59

삭제

청구항 60

삭제

청구항 61

삭제

청구항 62

삭제

청구항 63

삭제

청구항 64

삭제

청구항 65

삭제

청구항 66

삭제

청구항 67

삭제

청구항 68

삭제

청구항 69

삭제

청구항 70

삭제

청구항 71

삭제

청구항 72

삭제

청구항 73

삭제

청구항 74

삭제

청구항 75

삭제

청구항 76

삭제

청구항 77

삭제

청구항 78

삭제

청구항 79

삭제

청구항 80

삭제

청구항 81

삭제

청구항 82

삭제

청구항 83

삭제

청구항 84

삭제

청구항 85

삭제

청구항 86

삭제

청구항 87

삭제

청구항 88

삭제

청구항 89

삭제

청구항 90

삭제

청구항 91

삭제

청구항 92

삭제

청구항 93

삭제

청구항 94

삭제

청구항 95

삭제

청구항 96

삭제

청구항 97

삭제

청구항 98

삭제

청구항 99

삭제

청구항 100

삭제

청구항 101

삭제

청구항 102

삭제

청구항 103

삭제

청구항 104

삭제

청구항 105

삭제

청구항 106

삭제

청구항 107

삭제

청구항 108

삭제

청구항 109

삭제

청구항 110

삭제

발명의 설명

기 술 분 야

[0001] 본 발명은 일반적으로 DNA 재조합 기술 분야에 관한 것이며, 더욱 구체적으로, 다수의 DNA 절편(segment)을 집합된(assembled) 폴리뉴클레오타이드들에 삽입한 정렬된 집합체에 대한 개선된 방법에 관한 것이다.

[0002] 본 출원은 전체 내용이 본원에 인용되어 있는, 2008년 11월 19일에 출원된, 미국 가출원 제 61/116,109호, 2009년 3월 23일에 출원된 미국 가출원 제 61/162,230호의 우선권의 이익을 주장하고 있다.

배 경 기 술

[0003] 폴리뉴클레오타이드의 재조합은 업계에 알려진 수많은 방법을 이용하여 실행될 수 있다. 핵산을 재조합하는 종래의 기술은 새로운 핵산 분자를 만들기 위하여 제한효소와 연결효소(ligating enzyme)를 사용해 왔다. 클로닝 벡터 및 발현 벡터와 같은 재조합 분자는 숙주 세포의 게놈으로 관심있는 핵산 서열을 통합하거나, 및/또는 관심있는 하나 이상의 유전자의 발현을 유도하기 위하여 사용될 수 있다. 세포, 예를 들어, 효모 세포 내의 관심있는 유전자의 발현을 유도하기 위한 벡터의 사용은, 상기 벡터가 관심있는 유전자의 복제 및 발현을 가능하게 하는 필수적 유전적 성분을 포함할 것을 요구한다. 이러한 성분들은, 예를 들어, 관심있는 유전자 또는 유전자들, 프로모터 서열, 종결자(terminator) 서열, 선별표지(selectable marker), 통합 부위(integration loci), 기타 등등을 포함할 수 있다.

[0004] 종래의 제한효소 및 연결효소에 기반한 방법을 사용하는 단일 벡터로 상기 성분들을 집합시키는 것은 시간을 낭비하고 힘이 들 수 있다. 각각의 서브-클로닝 단계, 예를 들어, 존재하는 폴리뉴클레오타이드로 새로운 핵산 절편의 도입은, 추가적인 절편을 도입하기 전에 결과물 클론이 선별되고 특징지어져야 함을 요구할 수 있다. 평활말단(blunt end) 연결에 의해 생성된 클론은, 절편이 적절한 방향으로 도입되었다는 확인을 요구한다. 한편, 접착말단(sticky-end) 연결은, 수용체 절편 상의 접착말단(sticky end)을 형성하기 위하여 사용된 제한 부위가 또한 도너(donor) 절편에 존재해야 하며, 도너(donor) 절편 내 관심있는 서열을 단절시키는 부위에 존재해서는 안된다는 것을 요구한다. 따라서, 실행가능한 제한 부위의 선택은 연결될 부분의 조성에 전적으로 의존하므로 각각의 경우에 있어서 조심스럽게 고려되어야 한다. 또한, 이러한 방법은 원하는 유전자 결과물의 구조 및 기능을 방해할 수 있는 결과물 클론에, 관련없는 핵산 서열을 종종 도입한다. 더욱이, 제한효소에 기반한 클로닝 방법의 효율의 제한은 단일한 반응에서 함께 연결될 수 있는 수많은 핵산 분자 상에서의 고유한 제한이다.

[0005] 중합효소 연쇄 반응 (PCR)은 게놈 DNA, cDNA 및 mRNA을 포함하는, 특정 폴리뉴클레오타이드 서열이 생체 밖에서 증폭되는 강력한 기법이다. PCR은 전형적으로 2개의 가닥 상에서 상보적인 프라이머 신장 생성물의 형성을 허용하는 조건 하에서 2개의 올리고뉴클레오타이드 프라이머를 갖는 타겟 핵산의 분리된 상보적인 가닥들을 접촉시키는 것을 포함한다. 이러한 가닥들은 원하는 핵산 서열의 복제의 합성을 위한 주형으로 작용한다. 자동화 시스템(automated system)에서 분리 과정과 합성 과정을 반복함으로써, 타겟 서열의 기하급수적 복제를 이룰 수 있다.

[0006] PCR의 하나의 방법으로서, "겹침 신장에 의한 스플라이싱(splicing by overlap extension : SOE)" (예를 들어, 미국 특허등록 제 5,023,171호를 참조)로 불리는 방법은, 제한효소 또는 연결효소(ligase)를 사용하지 않고 정확한 접합부에서 DNA 분자의 집합체를 촉진한다. 재조합될 절편 성분은 다른 것에 대해 상보적인 말단을 갖는 앰플리콘을 생산하는, 특이하게 고안된 프라이머를 사용하는 별개의 중합효소 연쇄 반응에서 형성된다. 이러한 앰플리콘의 혼합 및 변성에 있어서, 3' 말단에서 상보적인 서열을 갖는 가닥들이 겹쳐서 서로에 대해 프라이머로서 작용한다. DNA 폴리머라제에 의한 이러한 겹침의 신장(extension)은, 원래의 서열들이 서로 "스플라이싱된(spliced)" 핵산 분자를 생산한다. PCR의 차후의 단계는 결과물인 스플라이싱 된 폴리뉴클레오타이드를 증폭한다.

[0007] 수많은 핵산 절편들을 결합시키는 종래의 연결 효소에 기반한 방법보다 더욱 효율적인, SOE는, 프라이머 서열을 최적화하는 시간과 원하는 생성물을 생산하기 위한 증폭 조건을 요구한다. 서로 스플라이싱 되어야 하는 절편들 사이의 각각의 접합부는 개별적으로 고려되어야 하며, 한 쌍의 프라이머들은 양립될 수 있는 말단들을 형성하도록 각각의 절편에 대하여 고안되어야 한다. PCR 프라이머의 고안을 위한 종래의 고려 사항, 예를 들어, 융점, G-C 함량, 헤어핀(hairpin) 및 다이머 형성의 회피, 잘못된 프라이밍 부위에 대한 엄격함(stringency)은, SOE 반응의 증가에서 스플라이싱 될 수많은 절편으로서 더욱 조심스럽게 고려되어야 한다.

발명의 내용

해결하려는 과제

[0008] 따라서, DNA 재조합 기술의 진보에도 불구하고, 폴리뉴클레오타이드의 빠르고 정렬된 집합체를 제공하는 개선된 방법에 대한 필요성이 존재한다. 특히, 프라이머 호적화 단계에 대한 필요성 없이, 중간체 생성물의 최소한의 조작과 특성화를 갖는 수많은 폴리뉴클레오타이드의 집합체를 촉진할 수 있는 방법에 대한 필요성이 존재한다.

이러한 필요성들은 본 발명의 조성물 및 이에 대한 방법에 의해 충족될 수 있다.

과제의 해결 수단

[0009] 본 발명의 조성물 및 이에 대한 방법은 집합된(assembled) 폴리뉴클레오타이드들에 삽입된 폴리뉴클레오타이드 성분들의 빠르고 정렬된 집합체 또는 "스티칭(stitching)"을 허용한다. 몇몇 실시예에서, 본 발명의 방법은 원형의 핵산 집합 벡터를 사용한다. 몇몇 실시예에서, 집합 벡터는 폴리뉴클레오타이드 성분을 포함하며, 상기 폴리뉴클레오타이드 성분은 (i) 3' 말단의 어닐링될 수 있는 링커(annealable linker); (ii) 5' 말단의 프라이머 결합 절편과 3' 말단의 어닐링될 수 있는 링커(annealable linker); (iii) 3' 말단 및 5' 말단의 어닐링될 수 있는 링커(annealable linker); (vi) 5' 말단의 어닐링될 수 있는 링커(annealable linker) 및 3' 말단의 프라이머 결합 절편; 또는 (v) 5' 말단의 어닐링될 수 있는 링커(annealable linker);의 측면에 위치한 DNA 절편을 포함한다.

[0010] 몇몇 실시예에서, 수많은 폴리뉴클레오타이드 성분들은 단일 반응 베셀 내 수많은 집합 벡터들을 제공함으로써 서로 스티칭될 수 있다. 몇몇 실시예에서, 폴리뉴클레오타이드 성분들은 반응 베셀 내 집합 벡터로부터 잘라낼 수 있다. 몇몇 실시예에서, 그 후, 폴리뉴클레오타이드 성분은 변성될 수 있고, 어닐링될 수 있는 링커 서열은 인접한 폴리뉴클레오타이드 성분 상의 상보적인 가닥들에 어닐링될 수 있으며, 폴리뉴클레오타이드 성분은 겹침 신장에 의한 스플라이싱(SOE) 이후에 PCR을 함으로써 집합된 폴리뉴클레오타이드 내로 함께 스티칭될 수 있다. 다른 실시예에서, 집합 벡터들로부터 잘라진 폴리뉴클레오타이드 성분들은 폴리뉴클레오타이드 성분들로 형질변환된 숙주 세포 내 상동 재조합(homologous recombination)에 의해 생체 내 집합된 폴리뉴클레오타이드로 집합될 수 있다. 집합된 폴리뉴클레오타이드들은 숙주 세포 매개 상동 재조합(homologous recombination)에 의해 생체 내에서 더 결합될 수 있다.

[0011] 폴리뉴클레오타이드 집합체의 효율성은 집합 벡터 내에 사용된 어닐링될 수 있는 링커 서열의 표준 세트의 제공에 의해 증진될 수 있으며, 예를 들어, 본원에 제시된 상기 서열로서 서열번호 1 내지 23을 들 수 있다. 어닐링될 수 있는 링커 서열들은 집합 반응(assembly reaction)에서 인접한 폴리뉴클레오타이드 성분들 간에 서열 겹침을 제공한다. 이상적으로, 어닐링될 수 있는 링커 서열들은 RNA 수준 및 DNA 수준에서 감지할 수 있을 정도의 2차 구조가 부족하며, 원하지 않는 방식으로 서로 교차 반응하지 않으며, 상대적으로 높은 융점(Tm)을 갖는다. 따라서, 많은 폴리뉴클레오타이드 성분들은 각각의 폴리뉴클레오타이드 성분에 대한 특이 프라이머의 고안에 대한 필요성 없이 서로 스티칭할 수 있으며, 이를 통해 시간과 노력을 절감할 수 있다. 본 발명의 조성물 및 이에 대한 방법은, 합성 유전자, 구성, 클로닝 벡터, 발현 벡터, 염색체, 게놈, 펩타이드 라이브러리, 기타 등등을 포함하는, 다양한 타입의 폴리뉴클레오타이드들을 결합하는데 사용될 수 있다.

[0012] 일 측면에서, 본 발명은 벡터, 예를 들어, 수많은 폴리뉴클레오타이드 성분들로부터 집합된 하나 이상의 폴리뉴클레오타이드들의 집합체에 사용될 수 있는, 집합 벡터를 제공한다.

[0013] 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-LA-D-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-D-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D 및 제한 부위 RB (예를 들어, 5'-RA-LA-D-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-PA-D-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 프라이머 결합 절편 PB 및 제한 부위 RB (예를 들어, 5'- RA-LA-D-PB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-LA-D-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 예시적인 집합 벡터들은 도 1b 및 도 2에서 제시되었다.

[0014] 몇몇 실시예에서, 프라이머 결합 절편 (예를 들어, PA 또는 PB)은 집합된 폴리뉴클레오타이드를 제조하는데 사용된 어닐링될 수 있는 링커 서열들 중 어느 것과도 상보적이지 않은 뉴클레오타이드 서열일 수 있다. 몇몇 실시예에서, 프라이머 결합 절편은 제한 엔도뉴클레아제 인식 부위 및/또는 절단 부위를 포함한다. 몇몇 실시예에서, 프라이머 결합 절편은, 특정한 집합 반응(assembly reaction)에 사용되지 않는, 이용가능한 링커 서열들 (예를 들어, 서열번호 1 내지 23 중 하나) 중 어느 하나의 핵산 서열, 또는 이의 상보적 서열(complement)을 포

함한다. 몇몇 실시예에서, 프라이머 결합 절편 PA의 핵산 서열은 서열번호 24, 25 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 프라이머 결합 절편 PB의 핵산 서열은 서열번호 24, 25 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된다. 바람직한 실시예에서, 프라이머 결합 절편 PA 및 프라이머 결합 절편 PB는 서열에서 동일하지 않다.

[0015] 몇몇 실시예에서, 2개 이상의 어닐링될 수 있는 링커 서열들은 적어도 24개의 뉴클레오타이드 길이이며 적어도 60℃의 Tm을 갖는다.

[0016] 몇몇 실시예에서, 2개 이상의 어닐링될 수 있는 링커 서열들은 적어도 70%의 G-C 함량 및 적어도 70℃의 Tm을 가지며, 인지가능한 2차 DNA 구조를 형성하지 않는다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LA의 핵산 서열은 서열번호 1 내지 8 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 8 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LA 및 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 8 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된다.

[0017] 몇몇 실시예에서, 2개 이상의 어닐링될 수 있는 링커 서열들은 적어도 30%의 A-T 함량 및 적어도 65℃의 Tm을 가지며, 인지가능한 2차 DNA 또는 RNA 구조를 형성하지 않는다. 몇몇 실시예에서, 2개 이상의 어닐링될 수 있는 링커 서열들은 낮은 G-C 함량 및 적어도 65℃의 Tm을 가지며, 서열 모티프 5'-ANNNNNNNNANNNAANTANNTTNANA-3'를 포함하며, 상기 A는 아데닌을 나타내며, N은 임의의 뉴클레오타이드를 나타내며, T는 티민을 나타낸다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LA의 핵산 서열은 서열번호 9 내지 23 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 9 내지 23 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LA 및 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 9 내지 23 및 이의 상보적 서열(complement)로 구성된 군으로부터 선택된다.

[0018] 수많은 폴리뉴클레오타이드 성분들의 정렬된 집합체는 집합 벡터 내에 DNA 절편을 측면에 배치한 어닐링될 수 있는 링커 서열들의 선택에 의해 조절될 수 있다. 따라서, 몇몇 실시예에서, 폴리뉴클레오타이드 성분들이 정렬된 방식으로 집합될 수 있는지 확인하기 위하여, 특정한 집합 벡터 내의 어닐링될 수 있는 링커 서열/어닐링될 수 있는 링커 서열 쌍의 서열은 상보적이지 않다. 유사하게, 몇몇 실시예에서, 특정한 집합 벡터 내의 프라이머 결합 절편/어닐링될 수 있는 링커 서열 쌍의 서열들은 상보적이지 않다.

[0019] 특정 실시예에서, 집합 벡터로부터 폴리뉴클레오타이드 성분이 잘려지는 것을 촉진시킬 수 있도록, 제한 부위 RA 및 RB는 동일한 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, 제한 부위 RA 또는 RB는 5' 또는 3' 오버행(overhang)을 남기는 제한 엔도뉴클레아제에 의해 절단될 수 있다. 다른 실시예에서, 제한 부위 RA 또는 RB는 평활 말단(blunt end)을 남기는 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, 제한 부위 RA 및 RB는 동일한 제한 엔도뉴클레아제에 의해 절단될 수 있다. 또다른 실시예에서, 제한 부위 RA 및 RB는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, 제한 부위 RA 및 RB는 동일한 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있다. 특정 실시예에서, 제한 부위 RA 및 RB는 SapI 또는 LguI 제한 엔도뉴클레아제에 의해 절단되 수 있다.

[0020] 다른 측면에서, 본 발명은 DNA 절편을 포함하는 집합 벡터의 제조에 유용한 엔트리 벡터(entry vector)를 제공한다.

[0021] 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-RY-D-RZ-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ 및 제한 부위 RB (예를 들어, 5'-RA-LA-RY-D-RZ-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-LA-RY-D-RZ-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-PA-RY-D-RZ-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 프라이머 결합 절편 PB 및 제한

부위 RB (예를 들어, 5'-RA-LA-RY-D-RZ-PB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 예시적인 엔트리 벡터(entry vector)는 도 1a에 제시하였다.

[0022] RY 및 RZ를 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제를 갖는 엔트리 벡터(entry vector)의 분해(digestion)는 DNA 절편을 수용할 수 있는 선형(linearized) 벡터를 형성할 수 있다. DNA 절편은 본 발명의 집합 벡터를 형성하는 표준 클로닝 기술을 사용하여 RY 및 RZ 부위에 연결될 수 있다. 몇몇 실시예에서, 엔트리 벡터(entry vector)의 제한 부위 RY 및 RZ는 동일한 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, 엔트리 벡터(entry vector)의 제한 부위 RY 및 RZ는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, 엔트리 벡터(entry vector)의 제한 부위 RY 및 RZ는 동일한 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있다. 특정한 실시예에서, IIS형 제한 엔도뉴클레아제는 SchI 또는 MlyI이다.

[0023] 몇몇 실시예에서, 엔트리 벡터(entry vector)의 제한 부위 RA 및 RB는 동일한 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, 엔트리 벡터(entry vector)의 제한 부위 RA 및 RB는 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, 엔트리 벡터(entry vector)의 제한 부위 RA 및 RB는 동일한 IIS형 제한 엔도뉴클레아제에 의해 절단될 수 있다. 특정한 실시예에서, IIS형 제한 엔도뉴클레아제는 SapI 또는 LguI이다.

[0024] 다른 측면에서, 본 발명은 수많은 폴리뉴클레오타이드 성분들로부터 하나 이상의 집합된 폴리뉴클레오타이드들의 집합체에 사용되는 수많은 집합 벡터를 포함하는 집합체 조성물을 제공한다. 몇몇 실시예에서, 집합체 조성물은, 하기의 것들을 포함한다:

[0025] (a) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 그룹 D_0로부터 선택된 임의의 DNA 절편, 어닐링될 수 있는 링커 서열 LBo 및 두번째 제한 부위 RBo를 포함하는, 하나 이상의 첫번째 핵산 분자;

[0026] (b) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, 그룹 D_n으로부터 선택된 임의의 DNA 절편, 두번째 어닐링될 수 있는 링커 서열 LB_n 및 두번째 제한 부위 RB_n을 포함하는, 하나 이상의 중간 핵산 분자(상기 n은 1 내지 중간 핵산 분자의 개수까지의 정수를 나타냄); 및

[0027] (c) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, 그룹 D_m으로부터 선택된 임의의 DNA 절편, 두번째 제한 부위 RB_m를 포함하는, 하나 이상의 마지막 핵산 분자(상기 m은 중간 핵산 분자의 개수보다 큰 정수를 나타냄);

[0028] 여기에서, 제한 부위 RA_0부터 RB_m까지의 절단 및 결과물 선형 핵산 분자의 변성에 있어서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$는 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 하이브리드할 수 있으며, 상기 n은 1부터 (m-1)까지 변하는 정수이며, 상기 p는 1 내지 m의 정수를 나타내며, 및 상기 각각의 그룹 D_0,... D_n,... 및 D_m은 하나 이상의 DNA 절편(segment)으로 구성된다.

[0029] 몇몇 실시예에서, 하나 이상의 첫번째 핵산 분자는 그룹 D_0로부터 선택된 DNA 절편에 대하여 5'에 위치한 프라이머 결합 절편 PA를 더 포함한다. 몇몇 실시예에서, 하나 이상의 마지막 핵산 분자는 그룹 D_m으로부터 선택된 DNA 절편에 대하여 3'에 위치한 프라이머 결합 절편 PB를 더 포함한다.

[0030] 몇몇 실시예에서, 집합체 조성물은 2개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 3개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 4개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 5개 이상의 중간 핵산 분자를 포함한다. 특정 실시예에서, 조성물은 6개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 7개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 8개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 9개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 10개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 15개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 20개 이상의 중간 핵산 분자를 포함한다.

[0031] 몇몇 실시예에서, m은 1이다. 몇몇 실시예에서, m은 2이다. 몇몇 실시예에서, m은 3이다. 몇몇 실시예에서, m은 4이다. 몇몇 실시예에서, m은 5이다. 몇몇 실시예에서, m은 6이다. 몇몇 실시예에서, m은 7이다. 몇몇 실시예에서, m은 8이다. 몇몇 실시예에서, m은 9이다. 몇몇 실시예에서, m은 10이다. 몇몇 실시예에서, m은 10

이상이다.

[0032] 몇몇 실시예에서, 제한 부위 RA_0부터 RB_m까지의 절단 및 결과물 선형 핵산 분자의 변성에서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은, 집합체 조성물에서, 다른 어닐링할 수 있는 링커 서열, 또는 이의 상보적 서열(complement)에 비해, 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 선택적으로 하이브리드할 수 있다. 몇몇 실시예에서, 각각의 어닐링될 수 있는 링커 서열 $L_{(p-1)}$은 어닐링될 수 있는 링커 서열 LA_p과 서열에 있어서 동일하다.

[0033] 특정 실시예에서, 집합 벡터들로부터 폴리뉴클레오타이드 성분들의 잘려짐을 촉진할 수 있도록, 제한 부위들 RA_0부터 RB_m까지는 동일한 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, 제한 부위들 RA_0부터 RB_m까지는 SapI 및/또는 LguI 제한 엔도뉴클레아제들에 의해 절단될 수 있다.

[0034] 다른 측면에서, 본 발명은 수많은 선형 핵산 분자를 포함하는 성분 조성물(component composition)을 제공하며, 상기 선형 핵산 분자는 제한 부위 RA_0부터 RB_m까지 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제로 집합체 조성물을 분해함으로써 형성될 수 있으며, 상기 집합체 조성물은 하기의 것들을 포함한다:

[0035] (a) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 그룹 D_0로부터 선택된 임의의 DNA 절편, 어닐링될 수 있는 링커 서열 LB_0 및 두번째 제한 부위 RB_0를 포함하는, 하나 이상의 첫번째 핵산 분자;

[0036] (b) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, 그룹 D_n으로부터 선택된 임의의 DNA 절편, 두번째 어닐링될 수 있는 링커 서열 LB_n 및 두번째 제한 부위 RB_n를 포함하는, 하나 이상의 중간 핵산 분자(상기 n은 1부터 중간 핵산 분자의 개수까지의 정수를 나타냄); 및

[0037] (c) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, 그룹 D_m으로부터 선택된 임의의 DNA 절편, 두번째 제한 부위 RB_m을 포함하는, 하나 이상의 마지막 핵산 분자(상기 m은 중간 핵산 분자의 개수보다 큰 정수임);

[0038] 여기에서, 제한 부위 RA_0부터 RB_m까지의 절단 및 결과물 선형 핵산 분자의 변성에 있어서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 하이브리드할 수 있으며, 상기 n은 1부터 (m-1)까지 변하는 정수이며, 상기 p는 1부터 m까지의 정수를 나타내며, 상기 각각의 그룹 D_0,... D_n,...및 D_m은 하나 이상의 DNA 절편(segment)으로 구성된다.

[0039] 몇몇 실시예에서, 하나 이상의 첫번째 핵산 분자는 그룹 D_0로부터 선택된 DNA 절편에 대해 5'에 위치한 프라이머 결합 절편 PA를 더 포함한다. 몇몇 실시예에서, 하나 이상의 마지막 핵산 분자는 그룹 D_m으로부터 선택된 DNA 절편에 대해 3'에 위치한 프라이머 결합 절편 PB를 더 포함한다.

[0040] 몇몇 실시예에서, 성분들의 조성물은 2개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 성분들의 조성물은 3개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 성분들의 조성물은 4개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 성분들의 조성물은 5개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 성분들의 조성물은 6개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 성분들의 조성물은 7개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 8개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 9개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 10개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 15개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 20개 이상의 중간 핵산 분자를 포함한다.

[0041] 몇몇 실시예에서, m은 1이다. 몇몇 실시예에서, m은 2이다. 몇몇 실시예에서, m은 3이다. 몇몇 실시예에서, m은 4이다. 몇몇 실시예에서, m은 5이다. 몇몇 실시예에서, m은 6이다. 몇몇 실시예에서, m은 7이다. 몇몇 실시예에서, m은 8이다. 몇몇 실시예에서, m은 9이다. 몇몇 실시예에서, m은 10이다. 몇몇 실시예에서, m은 10 이상이다.

[0042] 다른 측면에서, 본 발명의 방법에 따른 수많은 폴리뉴클레오타이드들을 집합하는데 유용한 키트를 제공한다. 몇몇 실시예에서, 상기 키트는: (a) 본원에 제시된 하나 이상의 엔트리 벡터(entry vector)들; (b) 엔트리 벡터

(entry vector)들의 제한 부위들 RA 및 RB를 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제들; 및 (c) 엔트리 벡터(entry vector)들의 제한 부위들 RY 및 RZ을 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제들을 포함한다.

[0043] 다른 측면에서, 본 발명은 핵산 분자의 라이브러리를 제공한다. 몇몇 실시예에서, 라이브러리의 핵산 분자는 첫번째 제한 부위 RA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 두번째 제한 부위 RB를 포함한다. 몇몇 실시예에서, 라이브러리의 핵산 분자는 첫번째 제한 부위 RA, 프라이머 결합 절편 PA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 두번째 제한 부위 RB를 포함한다. 몇몇 실시예에서, 라이브러리의 핵산 분자는 첫번째 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 두번째 제한 부위 RB를 포함한다. 몇몇 실시예에서, 라이브러리의 핵산 분자는 첫번째 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D 및 두번째 제한 부위 RB를 포함한다. 몇몇 실시예에서, 라이브러리의 핵산 분자는 첫번째 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 프라이머 결합 절편 PB 및 두번째 제한 부위 RB를 포함한다.

[0044] 몇몇 실시예에서, 라이브러리는 하기의 각각의 벡터들 중 적어도 하나를 포함한다:

[0045] (a) 5'에서 3'의 방향으로, 제한 부위 RA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

[0046] (b) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터; 및

[0047] (c) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D 및 제한 부위 RB_0를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터.

[0048] 몇몇 실시예에서, 라이브러리는 하기의 각각의 벡터들 중 적어도 하나를 포함한다:

[0049] (a) 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터;

[0050] (b) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터; 및

[0051] (c) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 프라이머 결합 절편 PB 및 제한 부위 RB_0를 포함하는, 원형 폴리뉴클레오타이드로 구성된 벡터.

[0052] 몇몇 실시예에서, DNA 절편 D는 선별표지(selectable marker), 프로모터, 게놈 타겟팅(targeting) 서열, 에피토프 태그(epitope tag)를 코딩하는 핵산 서열 및 관심 있는 유전자를 코딩하는 핵산 서열, 종결 코돈 및 lacZ를 코딩하는 핵산 서열로 구성된 군으로부터 선택된 핵산 서열을 포함한다.

[0053] 몇몇 실시예에서, 라이브러리는 하기의 각각의 핵산 분자 중 적어도 하나를 포함한다:

[0054] (a) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, DNA 절편 D_0, 어닐링될 수 있는 링커 서열 LB_0 및 두번째 제한 부위 RB_0를 포함하ㄴ는 첫번째 핵산 분자;

[0055] (b) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, DNA 절편 D_n, 두번째 어닐링될 수 있는 링커 서열 LB_n 및 두번째 제한 부위 RB_n를 포함하는 중간 핵산 분자(상기 n은 1부터 중간 핵산 분자의 개수까지의 정수를 나타냄);

[0056] (c) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, DNA 절편 D_m, 두번째 제한 부위 RB_m를 포함하는 마지막 핵산 분자(상기 m은 중간 핵산 분자의 개수를 초과하는 정수임);

[0057] 여기에서, 제한 부위 RA_0부터 RB_n까지의 절단 및 결과물 선형 핵산 분자의 변성에 있어서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 하이브리드할 수 있으며, 상기 p는 1부터 m까지의 정수이다. 몇몇 실시예에서, 첫번째 핵산 분자는 그룹 D_0으로부터 선택된 DNA

절편에 대해 5'에 위치한 프라이머 결합 절편 PA를 더 포함한다. 몇몇 실시예에서, 마지막 핵산 분자는 그룹 D_m으로부터 선택된 DNA 절편에 대해 3'에 위치한 프라이머 결합 절편 PB를 더 포함한다.

[0058] 몇몇 실시예에서, 라이브러리는 2개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 라이브러리는 3개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 라이브러리는 4개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 라이브러리는 5개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 라이브러리는 6개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 라이브러리는 7개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 8개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 9개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 10개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 15개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 20개 이상의 중간 핵산 분자를 포함한다.

[0059] 몇몇 실시예에서, m은 1이다. 몇몇 실시예에서, m은 2이다. 몇몇 실시예에서, m은 3이다. 몇몇 실시예에서, m은 4이다. 몇몇 실시예에서, m은 5이다. 몇몇 실시예에서, m은 6이다. 몇몇 실시예에서, m은 7이다. 몇몇 실시예에서, m은 8이다. 몇몇 실시예에서, m은 9이다. 몇몇 실시예에서, m은 10이다. 몇몇 실시예에서, m은 10 이상이다.

[0060] 다른 측면에서, 본 발명은 수많은 폴리뉴클레오타이드 성분들로부터 하나 이상의 집합된 폴리뉴클레오타이드들을 결합하는 방법을 제공하며, 하기의 단계를 포함한다:

[0061] (a) 성분들 조성물을 형성하기 위하여, 하나 이상의 제한 엔도뉴클레아제들을 사용하여 집합체 조성물을 분해하는 단계로서, 상기 집합체 조성물은 하기의 것들을 포함하며:

[0062] (i) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 그룹 PA로부터 선택된 임의의 프라이머 결합 절편, 그룹 D_0로부터 선택된 임의의 DNA 절편, 어닐링될 수 있는 링커 서열 LB_0 및 두번째 제한 부위 RB_0를 포함하는, 하나 이상의 첫번째 핵산 분자;

[0063] (ii) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, 그룹 D_n으로부터 선택된 임의의 DNA 절편, 두번째 어닐링될 수 있는 링커 서열 LB_n 및 두번째 제한 부위 RB_n를 포함하는, 하나 이상의 중간 핵산 분자(상기 n은 1부터 중간 핵산 분자의 개수까지의 정수를 나타냄); 및

[0064] (iii) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, 그룹 D_m으로부터 선택된 DNA 절편, 그룹 PB로부터 선택된 임의의 프라이머 결합 절편 및 두번째 제한 부위 RB_m를 포함하는, 하나 이상의 마지막 핵산 분자(상기 m은 중간 핵산 분자의 개수보다 큰 정수를 나타냄); 여기에서, 제한 부위 RA_0에서 RB_m까지의 절단 및 결과물 선형 핵산 분자의 변성에 있어서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 하이브리드할 수 있으며, 상기 n은 1부터 (m-1)까지 변하는 정수이며, 상기 P는 1부터 m까지의 정수를 나타내며 및 상기 각각의 그룹 D_0,... D_n,... 및 D_m은 하나 이상의 DNA 절편(segment)으로 구성되며;

[0065] 상기 하나 이상의 제한 엔도뉴클레아제들은 제한 부위 RA_0부터 RB_m까지를 절단할 수 있으며; 및

[0066] (b) 핵산 분자의 변성, 어닐링될 수 있는 링커 서열 LA_p에 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$의 어닐링 및 이들로부터의 신장에 적절한 조건 하에서, 성분 조성물(component composition)을 DNA 폴리머라제, 데옥시리보뉴클레오사이드 트리포스페이트, 하나 이상의 제 1 프라이머들 및 하나 이상의 제 2 프라이머들과 접촉시키는 단계로서; 상기 각각의 제 1 프라이머는 그룹 PA로부터 선택된 상기 프라이머 결합 절편(segment) 중 하나에 하이브리드할 수 있으며, 상기 각각의 제 2 프라이머는 그룹 PB로부터 선택된 상기 프라이머 결합 절편(segment) 중 하나에 하이브리드할 수 있으며; 및 성분 조성물(component composition)을 중합효소 연쇄 반응시키는 단계,

[0067] 여기에서, 5'에서 3'의 방향으로, 각각의 그룹 D_0,... D_n,... 및 D_m으로부터 선택된 하나의 DNA 절편을 포함하는 상기 폴리뉴클레오타이드는 집합된다. 상기 방법에서, p는 1부터 m까지의 정수를 나타낸다.

[0068] 몇몇 실시예에서, 집합체 조성물은 2개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 3개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 4개 이상의 중간 핵산 분자를 포함

한다. 몇몇 실시예에서, 집합체 조성물은 5개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 6개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 7개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 8개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 9개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 10개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 15개 이상의 중간 핵산 분자를 포함한다. 몇몇 실시예에서, 집합체 조성물은 20개 이상의 중간 핵산 분자를 포함한다.

[0069] 몇몇 실시예에서, m은 1이다. 몇몇 실시예에서, m은 2이다. 몇몇 실시예에서, m은 3이다. 몇몇 실시예에서, m은 4이다. 몇몇 실시예에서, m은 5이다. 몇몇 실시예에서, m은 6이다. 몇몇 실시예에서, m은 7이다. 몇몇 실시예에서, m은 8이다. 몇몇 실시예에서, m은 9이다. 몇몇 실시예에서, m은 10이다. 몇몇 실시예에서, m은 10 이상이다.

[0070] 몇몇 실시예에서, 집합체 조성물은 하나의 첫번째 핵산 분자 및 하나의 마지막 핵산 분자를 포함한다. 다른 실시예에서, 집합체 조성물은 하나를 초과하는 첫번째 핵산 분자 및 하나를 초과하는 마지막 핵산 분자를 포함하며, 집합 방법은 조합하는 방식으로 여러 개의 폴리뉴클레오타이드 성분들을 수많은 집합된 폴리뉴클레오타이드들에 제공한다. 몇몇 실시예에서, 집합체 조성물은, 동일한 어닐링될 수 있는 링커 서열 LA 또는 LB, 또는 동일한 프라이머 결합 절편 PA 또는 PB, 또는 동일한 쌍의 어닐링될 수 있는 링커 서열들 LA 및 LB, 또는 동일한 쌍의 어닐링될 수 있는 링커 서열/프라이머 결합 절편들 LA 및 PB, 또는 LB 및 PA를 포함하는 적어도 2개의 핵산 분자들을 포함한다.

[0071] 다른 측면에서, 본 발명은 집합된 폴리뉴클레오타이드들을 포함하는 숙주 세포를 형성하는 방법을 제공한다. 몇몇 실시예에서, 상기 방법은 본원에 제시된 폴리뉴클레오타이드 집합 방법에 의해 형성된 집합된 폴리뉴클레오타이드로 숙주 세포를 형질 변환하는 것을 포함한다. 다른 실시예에서, 상기 방법은 본원에 제시된 폴리뉴클레오타이드 집합 방법에 의해 형성된, 수많은 집합된 폴리뉴클레오타이드로 숙주 세포를 형질 변환하는 것을 포함한다. 특정 실시예에서, 숙주 세포는 2개 이상의 집합된(assembled) 폴리뉴클레오타이드들을 상동 재조합(homologous recombination)에 의해 결합된(combined) 하나 이상의 폴리뉴클레오타이드에 결합시킨다. 또다른 실시예에서, 상기 방법은 수많은 폴리뉴클레오타이드 성분들로 숙주 세포를 형질변환시키고 상기 숙주 세포로 하여금 상동 재조합(homologous recombination)에 의해 집합되거나(assembled) 결합된(combined) 하나 이상의 폴리뉴클레오타이드들을 형성하게 하는 것을 포함한다.

[0072] 다른 측면에서, 본 발명은 수많은 집합된(assembled) 폴리뉴클레오타이드들을 포함하는 수많은 숙주 세포들을 형성하는 방법을 제공한다. 몇몇 실시예에서, 수많은 숙주 세포들은, 폴리뉴클레오타이드 성분들의 조합된 집합체에 의해 형성된 수많은 집합된(assembled) 폴리뉴클레오타이드들을 포함하는 조성물로 숙주세포들을 형질 변환함으로써 형성된다. 다른 실시예에서, 수많은 숙주 세포들은, 수많은 집합된(assembled) 폴리뉴클레오타이드들을 포함하는 조성물로 숙주 세포들을 형질 변환하고 결합된(combined) 폴리뉴클레오타이드를 포함하는 숙주 세포들을 선택함으로써 형성되며, 상기 집합된(assembled) 폴리뉴클레오타이드들 중 적어도 2개의 집합된(assembled) 폴리뉴클레오타이드들은 선별표지(selectable marker)의 비-기능적 절편(segment)을 포함하며, 상기 표지(marker)에서 숙주 세포 매개 상동 재조합(homologous recombination)은 기능적인 선별표지(selectable marker)를 형성한다. 또다른 실시예에서, 수많은 숙주 세포들은 다수의 폴리뉴클레오타이드 성분들을 포함하는 성분 조성물(component composition)로 숙주 세포들을 형질 변환하고, 집합된(assembled) 폴리뉴클레오타이드를 포함하는 숙주 세포들을 선택하는 조합 방법에 의해 형성되며, 상기 다수의 폴리뉴클레오타이드 중 적어도 2개의 폴리뉴클레오타이드 성분들은 동일한 어닐링될 수 있는 링커 서열 LA 또는 LB, 또는 동일한 쌍의 어닐링될 수 있는 링커 서열들 LA 및 LB를 포함한다.

[0073] 다른 측면에서, 본 발명은 서열번호 1 내지 25를 구성하는 군으로부터 선택된 서열을 갖는 폴리뉴클레오타이드를 제공한다.

[0074] 다른 측면에서, 본 발명은 서열번호 1 내지 25를 구성하는 군으로부터 선택된 하나 이상의 서열들을 포함하는 폴리뉴클레오타이드를 제공한다.

[0075] **실시예의 상세한 설명**

[0076] **1. 정의**

[0077] 본원에서 사용된, 용어 "폴리뉴클레오타이드"는 당업자에 의해 이해되는 뉴클레오타이드 단위들로 구성된 폴리

머를 말한다. 바람직한 뉴클레오타이드 단위들은, 그러나 이에 제한되지 않는, 아데닌 (A), 구아닌 (G), 시토신 (C), 티민 (T) 및 우라실 (U)을 포함하는 뉴클레오타이드 단위들을 포함한다. 유용한 변형된 뉴클레오타이드 단위들은, 그러나 이에 제한되지 않는, 4-아세틸시티딘, 5-(카르복시히드록실메틸)우리딘, 2-O-메틸시티딘, 5-카르복시메틸아미노메틸-2-티오우리딘, 5-카르복시메틸아미노-메틸우리딘, 디히드로우리딘, 2-O-메틸슈도우리딘, 2-O-메틸구아노신, 이노신, N6-이소펜틸아데노신, 1-메틸아데노신, 1-메틸슈도우리딘, 1-메틸구아노신, 1-메틸이노신, 2,2-디메틸구아노신, 2-메틸아데노신, 2-메틸구아노신, 3-메틸시티딘, 5-메틸시티딘, N6-메틸아데노신, 7-메틸구아노신, 5-메틸아미노메틸우리딘, 5-메톡시아미노메틸-2-티오우리딘, 5-메톡시우리딘, 5-메톡시카르보닐메틸-2-티오우리딘, 5-메톡시카르보닐메틸우리딘, 2-메틸티오-N6-이소펜틸아데노신, 우리딘-5-옥시아세트산-메틸에스테르, 우리딘-5-옥시아세트산, 와이부톡소신(wybutoxosine), 와이부토신(wybutosine), 슈도우리딘, 퀘우오신(queuosine), 2-티오시티딘, 5-메틸-2-티오우리딘, 2-티오우리딘, 4-티오우리딘, 5-메틸우리딘, 2-O-메틸-5-메틸우리딘, 2-O-메틸우리딘, 기타 등등을 포함한다. 폴리뉴클레오타이드들은 자연적으로 발생하는 핵산, 예를 들어, 데옥시리보핵산 ("DNA") 및 리보핵산 ("RNA"), 뿐만 아니라 핵산 유사체를 포함한다. 핵산 유사체들은 비-자연적으로 발생하는 염기, 자연적으로 발생하는 포스포디에스테르 결합이 아닌 다른 뉴클레오타이드와의 연결(linkage)에 관여하는 뉴클레오타이드, 또는 포스포디에스테르 결합이 아닌 연결(linkage)을 통하여 부착된 염기들을 포함하는 뉴클레오타이드를 포함하는 핵산 유사체들을 포함한다. 따라서, 뉴클레오타이드 유사체는, 예를 들어, 이에 제한되지 않는, 포스포로티오에이트, 포스포로디티오에이트, 포스포로트리에스테르(phosphorotriester), 포스포라미데이트(phosphoramidate), 보라노포스페이트(boranophosphate), 메틸포스포네이트, 카이랄-메틸 포스포네이트, 2-O-메틸 리보뉴클레오타이드, 펩타이드-핵산 (PNA), 기타 등등을 포함한다.

[0078] 본원에서 사용된, "폴리뉴클레오타이드 성분(component)"은 본원에 제시된 폴리뉴클레오타이드 집합 방법을 사용하여 "집합된(assembled) 폴리뉴클레오타이드"를 형성하기 위하여 서로 집합될 수 있는 폴리뉴클레오타이드 서열을 말한다. 수많은 집합 벡터들이 집합 벡터들로부터 폴리뉴클레오타이드 성분들을 잘라낼 수 있는 하나 이상의 제한 엔도뉴클레아제로 분해될 때, 폴리뉴클레오타이드 성분들의 결과물 집단은 집합된(assembled) 폴리뉴클레오타이드 내로 집합되는 DNA 절편(segment)의 전체를 포함할 수 있다.

[0079] 본원에서 사용된, "집합된(assembled) 폴리뉴클레오타이드"는 본원에 제시된 폴리뉴클레오타이드 집합 방법에 의해 생산된 폴리뉴클레오타이드를 말한다. 집합된(assembled) 폴리뉴클레오타이드는 2개 이상의 폴리뉴클레오타이드 성분들로 구성될 수 있다. 몇몇 실시예에서, 집합된(assembled) 폴리뉴클레오타이드는 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개, 11개, 12개, 13개, 14개, 15개 또는 그 이상의 폴리뉴클레오타이드 성분들을 포함한다. 집합된(assembled) 폴리뉴클레오타이드 길이는 약 100개 내지 약 20,000개의 뉴클레오타이드, 또는 그 이상의 개수의 뉴클레오타이드 길이(length) 일 수 있다. 몇몇 실시예에서, 집합된(assembled) 폴리뉴클레오타이드 길이는 약 200개 내지 약 10,000개, 약 200개 내지 약 8000개, 약 200개 내지 약 5000개, 약 200개 내지 약 3000개, 또는 약 200개 내지 약 1000개 뉴클레오타이드 길이(length) 일 수 있다. 다른 실시예에서, 집합된(assembled) 폴리뉴클레오타이드 길이는 약 200개 내지 약 2000개, 약 2000개 내지 약 5000개, 약 5000개 내지 약 10,000개, 약 10,000개 내지 약 20,000개, 또는 20,000개를 초과하는 뉴클레오타이드 길이(length)일 수 있다.

[0080] 폴리뉴클레오타이드 서열들을 설명하기 위하여 종래의 표기법이 본원에 사용되었다: 단일-가닥 폴리뉴클레오타이드 서열의 왼쪽 말단은 5'-말단이며; 이중-가닥 폴리뉴클레오타이드 서열의 왼쪽 방향은 5'-방향을 말한다.

[0081] 본원에서 사용된, 하기의 실시예에서 "비트(Bit)"로 호환하여 칭하는, 용어 "DNA 절편"은, 분리되거나 분리될 수 있는 DNA 분자를 말한다. 유용한 실시예는, 그러나 이에 제한되지 않는, 단백질-코딩 서열, 리포터 유전자(reporter gene), 형광 표지(marker) 코딩 서열, 프로모터, 증폭자(enhancer), 종결자(terminator), 인트론, 엑손, 폴리-A 꼬리, 다수의 클로닝 부위, 핵 위치 신호(nuclear localization signal), mRNA 안정 신호(mRNA stabilization signal), 선별표지(selectable marker), 통합 부위(integration loci), 에피토프 태그(epitope tag) 코딩 서열, 분해 신호(degradation signal), 기타 자연적으로 발생한 DNA 분자 또는 합성 DNA 분자를 포함한다. 몇몇 실시예에서, DNA 절편은 천연 유래(natural origin)된 것일 수 있다. 선택적으로, DNA 절편은 생체 밖에서 생산된, 완전히 합성 유래(origin)된 것일 수 있다. 더욱이, DNA 절편은 분리되고 자연적으로 발생하는 DNA 분자의 임의의 조합, 또는 분리되고 자연적으로 발생하는 DNA 분자 및 합성 DNA 분자의 임의의 조합을 포함할 수 있다. 예를 들어, DNA 절편은 단백질 코딩서열에 사용할 수 있게 연결된 이형(heterologous) 프로모터, 폴리-A 꼬리에 연결된 단백질 코딩서열, 에피토프 태그(epitope tag) 코딩 서열로 뼈대가 완성되어 연결된 단백질 코딩서열, 기타 등등을 포함할 수 있다.

[0082] 용어 "상보적인(complementary)"은 위상적인(topological) 상보성을 말하거나, 당업자에 의해 이해되는 바와 같

이 2개의 폴리뉴클레오타이드들의 상호작용하는 표면을 서로 짝짓는 것을 말한다. 따라서, 안정한 비-평행의, 이중-가닥 핵산 구조를 형성하기 위하여 서로에게 하이브리드할 수 있다면, 2개의 서열들은 서로에 대해 "상보적"이다. 제 1 폴리뉴클레오타이드의 뉴클레오타이드 서열이 실질적으로 제 2 폴리뉴클레오타이드의 폴리뉴클레오타이드 결합 파트너의 뉴클레오타이드 서열과 동일하다면, 또는 제 1 폴리뉴클레오타이드가 엄격한 하이브리드 조건에서 제 2 폴리뉴클레오타이드와 하이브리드할 수 있다면, 제 1 폴리뉴클레오타이드는 제 2 폴리뉴클레오타이드에 대해 상보적이다. 따라서, 서열이 5'-TATAC-3'의 폴리뉴클레오타이드는, 서열이 5'-GTATA-3'인 폴리뉴클레오타이드에 대해 상보적이다.

[0083] "프라이머"는, 합성에 적절한 조건 하에서, 예를 들어, 뉴클레오타이드 및 합성 반응을 촉진시키는(catalyze) 성분 (예를 들어, DNA 폴리머라제)의 존재 하에서, 폴리뉴클레오타이드 주형(template) 서열, 예를 들어, 프라이머 결합 절편에 특히 하이브리드 할 수 있는 폴리뉴클레오타이드 서열과, 상보적 폴리뉴클레오타이드의 합성을 위한 개시 지점을 제공할 수 있는 폴리뉴클레오타이드를 말한다. 프라이머는 폴리뉴클레오타이드 주형(template) 서열에 대해 상보적이지만, 폴리뉴클레오타이드 주형(template) 서열에 대해 정확한 상보적 서열(complement)일 필요는 없다. 예를 들어, 프라이머는 폴리뉴클레오타이드 주형(template) 서열의 상보적 서열(complement)에 대해 적어도 약 80, 85, 90, 95, 96, 97, 98, 또는 99% 동일할 수 있다. 프라이머는 다양한 길이를 가질 수 있지만, 일반적으로 적어도 15개의 염기 길이를 가진다. 몇몇 실시예에서, 프라이머는 15 내지 35개의 염기 길이를 가진다. 몇몇 실시예에서, 프라이머는 35개를 초과하는 염기 길이를 가진다. 다른 실시예에서, 프라이머는 융점 (Tm)을 가지며, 예를 들어, 적어도 50℃의 상기 융점에서 DNA 이중나선의 1/2은 단일 가닥으로 분리될 것이다. 다른 실시예에서, 프라이머는 약 50℃ 내지 70℃의 Tm을 갖는다. 또다른 실시예에서, 폴리뉴클레오타이드 주형(template) 서열의 하이브리드화(hybridization)의 효율에 영향을 미치지 않도록, 프라이머는 인지가능한 DNA 또는 RNA 2차 구조를 형성하지 않는다.

[0084] 본원에서 사용된, 용어 "프라이머 결합 절편(primer binding segment)"은 합성에 적당한 조건하에서 상보적 폴리뉴클레오타이드의 합성의 개시 지점을 제공하기 위하여 프라이머에 결합하는 폴리뉴클레오타이드 서열이다. 몇몇 실시예에서, 프라이머 결합 서열은 본 발명의 어닐링될 수 있는 링커(annealable linker) 중 하나이다. 서열은, 주어진 세트의 집합체 조성물 내의 상보적 링커 또는 집합체 조성물 내의 폴리뉴클레오타이드 성분들의 부존재 속에서, 어닐링될 수 있는 링커(annealable linker)를 대신하는 프라이머 결합 서열이다. 몇몇 실시예에서, 프라이머 결합 절편은 게놈 타겟팅(targeting) 서열, 예를 들어, 업스트림(upstream) 또는 다운스트림(downstream) 게놈 타겟팅(targeting) 서열로서 기능할 수 있다.

[0085] 본원에서 사용된, 용어 "링커(linker) 서열" 및 "어닐링될 수 있는 링커 서열"은 상호교환적으로 사용되었고, 본 발명에 제시된 엔트리 벡터(entry vector) 및 집합 벡터 내에 함유된 폴리뉴클레오타이드 서열을 말한다. 특히, 어닐링될 수 있는 링커 서열은 엔트리 벡터(entry vector) 또는 집합 벡터 내에 DNA 절편을 측면에 배치한다. 집합 벡터로부터 폴리뉴클레오타이드 성분의 잘라냄과 폴리뉴클레오타이드 성분의 변성에 있어서, 어닐링될 수 있는 링커(annealable linker)는, 본원에 제시된 바와 같이, 폴리뉴클레오타이드 집합 반응(assembly reaction)에서 인접한 폴리뉴클레오타이드 성분의 상보적인 어닐링될 수 있는 링커 서열에 특히 하이브리드할 수 있다. 상보적 링커 가닥과 어닐링함에 있어서, 어닐링될 수 있는 링커(annealable linker)는 상보적 폴리뉴클레오타이드의 합성을 위한 개시 지점을 제공할 수 있다.

[0086] 본원에서 사용된, 용어 "벡터"는 세포에서 복제할 수 있는 염색체 외의(extrachromosomal) 핵산 분자 및 삽입 서열의 복제를 유발하도록 실제로 연결될 수 있는 삽입 서열과 관련하여 사용된다. 유용한 실시예는, 그러나 이에 제한되지 않는, 플라즈미드 구성, 파지(phage) 구성, 코스미드(cosmid) 벡터 등등과 같은 원형 DNA 분자 뿐만 아니라, 선형 핵산 구성 (예를 들어, 람다 파지(phage) 구성, 박테리아의 인공 염색체 (BAC), 효모의 인공 염색체 (YAC), 등)를 포함한다. 벡터는 프로모터 및/또는 종결자(terminator)와 같은 발현 신호, 항생제에 대한 저항성을 부여하는 유전자와 같은 선별표지(selectable marker) 및 삽입 서열들이 복제될 수 있는 하나 이상의 제한 부위를 포함할 수 있다. 벡터들은 (벡터가 수용할 수 있는 DNA 삽입의 크기와 같은) 기타 특이한 특성을 가질 수 있다.

[0087] 본원에서 사용된, 용어 "엔트리 벡터(entry vector)"는 본원에서 제공된 폴리뉴클레오타이드 집합 방법에서 사용되는 집합 벡터의 제조를 위한 원래의(parental) 벡터로 제공할 수 있는 클로닝 벡터 플라즈마를 말한다. 엔트리 벡터(entry vector)는 2개의 어닐링될 수 있는 링커 서열들, 또는 어닐링될 수 있는 링커 서열 및 프라이머 결합 절편을 포함하며, 이들은 집합 벡터를 형성하기 위하여 DNA 절편의 도입에 이용될 수 있는 제한 부위를 측면에 배치한다. 본원에서 사용된, "집합 벡터"는 DNA 절편이 도입된 엔트리 벡터(entry vector)를 말한다. 집합 벡터는 집합된(assembled) 폴리뉴클레오타이드로 집합될 수 있는 폴리뉴클레오타이드 성분을 제공하기 위하

여, 본원에 제시된 폴리뉴클레오타이드 집합 방법에 사용될 수 있다.

[0088] 본원에서 사용된, 용어 "집합 벡터"는 1개의 어닐링될 수 있는 링커 서열, 2개의 어닐링될 수 있는 링커 서열들, 또는 어닐링될 수 있는 링커 서열과 프라이머 결합 절편, 및 DNA 절편을 포함하는 벡터를 말한다.

[0089] 본원에서 사용된, 용어 "제한효소" 또는 "제한 엔도뉴클레아제"는 서열 내 정확한 위치에서 유사한 DNA 서열에 결합하여 DNA 분자를 절단하는 촉매적 분자의 계열의 구성원 또는 구성원들을 말한다. 제한 엔도뉴클레아제는 IIS형 제한 엔도뉴클레아제들을 포함한다. 이러한 계열의 효소는 인식서열(recognition sequence)이 절단 부위로부터 분리된 점에서, 기타 다른 제한 엔도뉴클레아제들과 다르다. IIS형 제한효소의 몇몇 실시예는 AlwI, BsaI, BbsI, BbuI, BsmAI, BsrI, BsmI, BspMI, EarI, Esp3I, FokI, HgaI, HphI, LguI, MboII, MnlI, PleI, SapI, SchI, SfaNi, 기타 등등을 포함한다. 이러한 많은 제한 엔도뉴클레아제들은 상업적으로 구입할 수 있으며 당업자에게 잘 알려져 있다.

[0090] 본원에서 사용된, 용어 "어닐링될 수 있는 링커(annealable linker) 서열 두 가닥의(duplex)"은 비평형 접합(antiparallel association)에서 실질적으로 상보적인 어닐링될 수 있는 링커 서열 가닥과 함께 정렬된 하나의 어닐링될 수 있는 링커 서열 가닥을 말한다. 특정한 실시예에서, 어닐링될 수 있는 링커 서열 두 가닥의(duplex) 서열이 완전한 상보성을 갖는 2개의 어닐링될 수 있는 링커 서열 가닥들을 포함하더라도, 상보성은 완전할 필요는 없으며; 어닐링될 수 있는 링커(annealable linker) 서열 두 가닥(duplex)은 짝짓지 않은 염기쌍 또는 짝짓지 않은 염기들을 포함할 수 있다.

[0091] 본원에서 사용된, 용어 "게놈 타겟팅(targeting) 서열"은 숙주 세포 매개 상동 재조합(homologous recombination)에 의해 삽입되는 부위에서 숙주 세포의 게놈에 존재하는 뉴클레오타이드 서열을 말한다. 용어 "업스트림(upstream) 게놈 타겟팅(targeting) 서열" 및 "다운스트림(downstream) 게놈 타겟팅(targeting) 서열"은 숙주 세포의 게놈에 서로 업스트림 및 다운스트림에 위치한 게놈 타겟팅(targeting) 서열들을 말한다.

[0092] 본원에서 사용된, 용어 "염색체의 타겟팅(targeting) 서열"은 본 발명의 폴리뉴클레오타이드가 숙주 세포 매개 상동 재조합(homologous recombination)에 의해 삽입되는 부위에서 숙주 세포의 염색체에 존재하는 뉴클레오타이드 서열을 말한다. 용어 "업스트림(upstream) 염색체의 타겟팅(targeting) 서열" 및 "다운스트림(downstream) 염색체의 타겟팅(targeting) 서열"은 숙주 세포의 염색체 내에서 서로의 업스트림 및 다운스트림에 위치한 염색체의 타겟팅(targeting) 서열들을 말한다.

[0093] **2. 폴리뉴클레오타이드 집합 방법**

[0094] 일 측면에서, 본 발명은 하나 이상의 집합된(assembled) 폴리뉴클레오타이드들로 삽입되는 수많은 폴리뉴클레오타이드 성분들의 정열된 집합체에 대한 빠르고, 강건하며 높은 생산량의(high-throughput) 방법을 제공한다. 본 발명의 방법들은 집합 벡터들로 불리는 원형 핵산 벡터들을 사용하며, 이들은 각각, 어닐링될 수 있는 링커 서열 (예를 들어, LA 또는 LB)의 측면에 배치된 DNA 절편 D, 한 쌍의 어닐링될 수 있는 링커 서열들 (예를 들어, LA과 LB), 또는 어닐링될 수 있는 링커 서열과 프라이머 결합 절편 (예를 들어, LA과 PB, 또는 LB과 PA) 및 한 쌍의 제한 부위들 RA 및 RB를 포함한다(도 1b). 제한 부위 RA 및 RB에서의 수많은 집합 벡터들의 제한 엔도뉴클레아제 분해는 성분들 5'-LA-D-3', 5'-D-LB-3', 5'-LA-D-LB-3', 5'-LA-D-PB-3', 또는 5'-LB-D-PA-3'을 포함하는 수많은 폴리뉴클레오타이드 성분들을 생성한다 (도 3). 본 발명의 방법에서, 어닐링될 수 있는 링커 서열들 LA 및 LB는 스플라이스 겹침 신장 집합 반응(assembly reaction) 후에, 정렬된 서열과 함께, 폴리뉴클레오타이드 성분들을 집합된(assembled) 폴리뉴클레오타이드로 집합시키는 중합효소 연쇄 반응 (SOE/PCR)에 사용되는 상보적 말단을 갖는 폴리뉴클레오타이드 성분들을 제공한다.

[0095] 특히, 상기 방법들은, 그러나 이에 제한되지 않는, 단백질-코딩 서열, 리포터 유전자, 형광 표지(marker) 코딩 서열, 프로모터, 증폭자(enhancer), 종결자(terminator), 인트론, 엑손, 폴리-A 꼬리, 다수의 클로닝 부위, 핵 위치 신호(nuclear localization signal), mRNA 안정 신호(mRNA stabilization signal), 선별표지(selectable marker), 통합 부위(integration loci), 에피토프 태그(epitope tag) 코딩 서열 및 분해 신호(degradation signal)를 포함하는, 기능적인 수많은 DNA 성분들의 집합된(assembled) 폴리뉴클레오타이드로의 집합을 제공한다. 상기 방법들은, 그러나 이에 제한되지 않는 합성 유전자, 구성, 클로닝 벡터, 발현 벡터, 염색체, 게놈 통합(integration) 구성, 게놈 및 DNA 라이브러리를 포함하는, 집합된(assembled) 폴리뉴클레오타이드의 임의의 형태의 집합을 위하여 사용될 수 있다. 더욱이, 상기 방법들은 중간 생산물의 증폭 및 특성화를 할 필요 없이 단일한 반응에서 DNA 절편을 집합하는데 사용될 수 있다.

[0096] 몇몇 실시예에서, 상기 방법들은, 1개 또는 2개의 어닐링될 수 있는 링커 서열들 LA 및/또는 LB에 의해, 또는 어닐링될 수 있는 링커 서열과 프라이머 결합 절편 (예를 들어, LA과 PB 또는 LB과 PA) 측면에 배치되는, 집합 벡터 (즉, 예를 들어, PCR 증폭, 화학적 합성, 기타 등등과 같이 업계에 알려진 표준 방법에 의해 수득된 DNA 절편(segment))에서 유래되지 않는 수많은 폴리뉴클레오타이드 성분들로부터 집합된(assembled) 폴리뉴클레오타이드의 집합체를 또한 제공한다. 집합 벡터로부터 유래되지 않은 폴리뉴클레오타이드 성분들은 SOE/PCR 반응 또는 집합된(assembled) 폴리뉴클레오타이드로의 집합을 위한 숙주 세포 매개 상동 재조합(homologous recombination) 이전의 단계에서 집합 반응(assembly reaction)으로 첨가될 수 있다. 따라서, 몇몇 실시예에서, 집합 방법은 (1) 1개 또는 2개의 어닐링될 수 있는 링커 서열들, 또는 어닐링될 수 있는 링커 서열과 프라이머 결합 절편을 포함하는 집합 벡터로부터 유도되고 집합 벡터들의 분해에 의해 생성된 폴리뉴클레오타이드 성분들; (2) 1개 또는 2개의 어닐링될 수 있는 링커 서열들에 의해, 또는 어닐링될 수 있는 링커 서열과 프라이머 결합 절편에 의해 측면에 배치된 벡터없는(vectorless) DNA 절편들; 및 (3) 이들의 조합;을 집합하는데 사용될 수 있다.

[0097] 몇몇 실시예에서, 본 발명은 하나 이상의 집합된(assembled) 폴리뉴클레오타이드들로 수많은 폴리뉴클레오타이드 성분들을 집합시키는 방법을 제공하며, 상기 방법은 하기의 단계들을 포함한다:

[0098] (a) 성분 조성물(component composition)을 생성하기 위하여 하나 이상의 제한 엔도뉴클레아제들을 사용하여 집합체 조성물을 분해하는 단계로서, 상기 집합체 조성물은 하기의 것들을 포함하며:

[0099] (i) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 그룹 PA로부터 선택된 임의의 프라이머 결합 절편, 그룹 D_0로부터 선택된 임의의 DNA 절편, 어닐링될 수 있는 링커 서열 LB_0 및 두번째 제한 부위 RB_0을 포함하는, 하나 이상의 첫번째 핵산 분자;

[0100] (ii) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, 그룹 D_n으로부터 선택된 임의의 DNA 절편, 두번째 어닐링될 수 있는 링커 서열 LB_n 및 두번째 제한 부위 RB_n를 포함하는, 하나 이상의 중간 핵산 분자(상기 n은 1부터 중간 핵산 분자의 개수까지의 정수를 나타냄); 및

[0101] (iii) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, 그룹 D_m으로부터 선택된 DNA 절편, 그룹 PB로부터 선택된 임의의 프라이머 결합 절편 및 두번째 제한 부위 RB_m를 포함하는, 하나 이상의 마지막 핵산 분자(상기 m은 중간 핵산 분자의 개수보다 큰 정수를 나타냄); 여기에서, 제한 부위 RA_0에서 RB_m까지의 절단 및 결과물 선형 핵산 분자의 변성에 있어서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 하이브리드할 수 있으며, 상기 n은 1부터 (m-1)까지 변하는 정수이며, 상기 p는 1부터 m까지의 정수를 나타내며, 상기 각각의 그룹 D_0,... D_n,...및 D_m,는 하나 이상의 DNA 절편(segment)으로 구성되며;

[0102] 여기에서, 상기 하나 이상의 제한 엔도뉴클레아제들은 제한 부위 RA_0부터 RB_m까지를 절단할 수 있으며; 및

[0103] (b) 핵산 분자의 변성, 어닐링될 수 있는 링커 서열 LA_p에 대한 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$의 어닐링 및 이들로부터의 신장에 적절한 조건하에서, 성분 조성물(component composition)을 DNA 폴리머라제, 데옥시리보뉴클레오사이드 트리포스페이트, 하나 이상의 제 1 프라이머들 및 하나 이상의 제 2 프라이머들과 접촉시키는 단계; 상기 각각의 제 1 프라이머는 그룹 PA로부터 선택된 상기 프라이머 결합 절편(segment) 중 하나에 하이브리드할 수 있으며, 상기 각각의 제 2 프라이머는 그룹 PB로부터 선택된 상기 프라이머 결합 절편(segment) 중 하나에 하이브리드할 수 있으며; 및 상기 성분 조성물(component composition)을 중합효소 연쇄 반응시키는 단계,

[0104] 상기 폴리뉴클레오타이드는 5'에서 3'의 방향으로, 그룹 D_0,... D_n,... 및 D_m의 각각으로부터 선택된 하나의 DNA 절편을 포함하여 집합된다. 상기 방법에서, p는 1부터 m까지의 정수를 나타낸다.

[0105] 도 3은 예시적인 목적으로, 본 발명의 집합 방법의 일 실시예를 나타낸다. 상기 예시에서, 총 4개의 폴리뉴클레오타이드 성분들은 집합되어(assembled) 집합된(assembled) 폴리뉴클레오타이드를 생산하였다. 그러나, 본 발명의 집합 방법은 임의의 개수의 폴리뉴클레오타이드 성분들을 하나 이상의 집합된(assembled) 폴리뉴클레오타이드로 집합하는데 사용될 수 있다. 몇몇 실시예에서, 본 발명의 상기 방법은 2개, 3개, 4개, 5개, 6개, 7개, 8개, 9개, 10개, 11개, 12개, 13개, 14개, 15개, 또는 그 이상의 폴리뉴클레오타이드 성분들이 하나 이상의 집

합된(assembled) 폴리뉴클레오타이드로 집합된 집합체를 형성한다.

[0106] 도 3에 나타낸 실시예에서, 집합된(assembled) 폴리뉴클레오타이드를 형성하는 집합체 조성물은 "첫번째", "중간 1 (int_1)", "중간 2 (int_2)" 및 "마지막"으로 표시한, 4개의 인풋(input) 집합체들을 포함한다. 각각의 집합 벡터는 어닐링될 수 있는 링커 서열과 프라이머 결합 절편에 의하거나, 2개의 어닐링될 수 있는 링커 서열들 측면에 배치된 DNA 절편을 포함한다. 특히, DNA 절편 D_0는 5' 프라이머 결합 절편 PA와 3' 어닐링될 수 있는 링커 서열 LB_0에 의해 측면에 배치된다. DNA 절편 D_1은 5' 및 3' 어닐링될 수 있는 링커 서열들 LA_1 및 LB_1의 측면에 배치되며, DNA 절편 D_2는 5' 및 3' 어닐링될 수 있는 링커 서열들 LA_2 및 LB_2에 의해 측면에 배치된다. DNA 절편 D_3은 3' 프라이머 결합 절편 PB와 5' 어닐링될 수 있는 링커 서열 LA_3의 측면에 배치된다. 집합 벡터들 내 5'-PA-D-LB-3', 5'-LA-D-LB-3', 또는 5'-LA-D-PB-3' 성분들은 SapI 제한 엔도뉴클레아제 부위의 측면에 더 배치된다.

[0107] 도 3에 나타낸 집합 반응(assembly reaction)의 첫번째 단계에서, 집합체 조성물은 SapI에 의해 분해되어, 집합 벡터 골격으로부터 5'-PA-D-LB-3', 5'-LA-D-LB-3' 또는 5'-LA-D-PB-3'를 포함하는 폴리뉴클레오타이드 성분들을 성분 조성물(component composition)로의 절단을 유발한다. SapI는 IIS형 제한 엔도뉴클레아제이기 때문에, SapI의 인지 부위는 SapI의 절단 부위에서 멀고, 절단은 SapI의 인식서열(recognition sequence)의 바깥쪽에서 일어난다. 제한-부위 자국(scar)을 포함하지 않으며, 만약 그렇지 않으면 비-IIS형 제한 엔도뉴클레아제를 사용하여 제한 부위 RA 및 RB의 절단을 유발할 수 있는, 폴리뉴클레오타이드들이 집합될 수 있기 때문에, 상기 특성은 IIS형 제한 엔도뉴클레아제들을 본 발명의 방법에 따른 폴리뉴클레오타이드의 집합체에 특히 유용하게 만든다. 도 2와 관련하여, IIS형 인지 부위는 절단 부위 RB_0, RA_n 및 RA_m의 각각의 RA_0, RA_n 및 RA_m 및 3'에 대응되는 절단 부위의 5'이다. 따라서, 제한 부위 RA_0부터 RB_m까지는, RA_0부터 RB_m까지 절단할 수 있는 하나 이상의 IIS형 제한 엔도뉴클레아제들에 의한 절단이, D_0로부터의 RA_0 분리, RB_0로부터의 LB_0 분리, LA_n으로부터의 RA_n 분리, RM_n으로부터의 LB_n 분리, LA_m으로부터의 RA_m 분리 및 RB_m으로부터의 D_m 분리를 유발하도록 지향되며, 상기 결과물인 D_0, LB_0, RA_n, LB_n, LA_m 또는 D_m를 포함하는 선형의 핵산 분자는 RA_0부터 RB_m까지의 임의의 것을 포함하지 않는다. 그 결과, 결과물 폴리뉴클레오타이드 성분들은 제한효소의 인지 또는 절단 부위를 적은 양이라도 포함하지 않는다. 그 결과, 독창적인 폴리뉴클레오타이드 집합 방법은 유전적 불안정성을 일으킬 수 있는 서열 반복의 유도 없이 숙주 세포들을 여러번 형질 변환하는데 사용될 수 있다.

[0108] 따라서, 제한 엔도뉴클레아제는 선택적으로 불활성화된다. 만약 불활성화가 요구된다면, 컬럼 또는 겔-기반 정제 방법을 포함하는, 엔도뉴클레아제 효소 활성을 불활성화하는 업계에 알려진 임의의 방법이 사용될 수 있다. 하나의 편리한 방법은 열 불활성화(heat inactivation), 예를 들어, 20분 동안 65℃에서의 열 불활성화이며, 상기 방법은 반응 튜브 바깥의 성분 조성물(component composition)들의 적은 증폭을 필요로 하거나 증폭을 전혀 필요로 하지 않는다.

[0109] 집합된(assembled) 폴리뉴클레오타이드로 삽입된 폴리뉴클레오타이드 성분들의 집합체는 폴리뉴클레오타이드 성분들 중에서 상보적인 말단의 중복되는 가닥들에 의해 형성된 두 가닥의(duplex) 서열들에 의해 이루어질 수 있다. 특히, 어닐링될 수 있는 링커 서열들이 고안되어, 어닐링될 수 있는 링커 서열 LB_0는 어닐링될 수 있는 링커 서열 LA_1의 상보적 서열(complement)에 하이브리드될 수 있으며, 어닐링될 수 있는 링커 서열 LB_1은 어닐링될 수 있는 링커 서열 LA_2의 상보적 서열(complement)에 하이브리드할 수 있으며, 어닐링될 수 있는 링커 서열 LB_2는 어닐링될 수 있는 링커 서열 LA_3의 상보적 서열(complement)에 하이브리드할 수 있다. 따라서, 집합 반응(assembly reaction)의 두번째 단계에서, 폴리뉴클레오타이드 성분들은 변성 조건 (예를 들어, 열)에서 단일 가닥의 폴리뉴클레오타이드 성분들을 형성하며, 상기 집합 반응(assembly reaction)의 변성 단계와 함께, 또는 그 후에 생긴 부수물이 내열성의 DNA 폴리머라제 및 데옥시리보뉴클레오사이드 트리포스페이트와 접촉된다.

[0110] 내열성의 DNA 폴리머라제는 당업자에 의해 적당한 것으로 여겨지는 임의의 내열성의 DNA 폴리머라제일 수 있다. 본 발명의 방법의 사용에 적당한 내열성의 DNA 폴리머라제들은, 그러나 이에 제한되지 않는, *Thermus thermophilus* (Tth) DNA 폴리머라제, *Thermus aquaticus* (Taq) DNA 폴리머라제, *Thermotoga neopolitana* (Tne) DNA 폴리머라제, *Thermotoga maritima* (Tma) DNA 폴리머라제, *Therm° C° Ccus litoralis* (Tli 또는 VENT™) DNA 폴리머라제, *Pyr° C° Ccus furiosus* (Pfu 또는 DEEPVENT™) DNA 폴리머라제, *Pyr℃℃cus woosii* (Pwo) DNA 폴리머라제, *Bacillus sterothermophilus* (Bst) DNA 폴리머라제, *Sulfolobus acid° Caldarius* (SAC) DNA

폴리머라제, *Thermoplasma acidophilum* (Tac) DNA 폴리머라제, *Thermus flavus* (Tfl/Tub) DNA 폴리머라제, *Thermus ruber* (Tru) DNA 폴리머라제, *Thermus br° Ckianus* (DYNAZYME™) DNA 폴리머라제, *Methanobacterium ruminantium* (Mth) DNA 폴리머라제 및 이의 돌연변이, 변종 및 유도체를 포함한다. 높은 복제성 (예를 들어, 프루프리딩 물성)과 낮은 오류율을 갖는 내열성의 DNA 폴리머라제들이 바람직하다. 몇몇 실시예에서, DNA 폴리머라제는 Phusion™ DNA 폴리머라제 (New England Biolabs, Ipswich, MA)이다. 다른 실시예에서, DNA 폴리머라제는 PfuUltra™ II Fusion DNA 폴리머라제 (Strategene / Agilent, La Jolla, CA)이다.

[0111] 그 후, 내열성의 DNA 폴리머라제가, 겹치는 어닐링될 수 있는 링커 서열들 사이의 부분들을 채우는 동안, 집합 반응(assembly reaction)은 겹치는 어닐링될 수 있는 링커 서열들의 3'-히드록시 부분으로부터 가닥의 신장(strand elongation)을 허용하는 조건에 둔다. 상당한 양의 이중-가닥 집합된(assembled) 폴리뉴클레오타이드들이 형성되는 동안, 상기 집합 반응(assembly reaction)은 변성 / 어닐링 / 신장의 제한된 수의 반복 사이클(예를 들어, 5 내지 15 사이클)을 겪게 된다. 상기 사이클이 일어나는 동안, 폴리뉴클레오타이드 성분들은 프라이머와 주형(template)으로 작용하여 집합된(assembled) 폴리뉴클레오타이드에 대하여 완전한 길이의 주형(template)을 형성한다. 몇몇 실시예에서, PCR의 어닐링 및 신장은 모두 72℃에서 실시된다.

[0112] 어닐링될 수 있는 링커 서열들 LA 및 LB과 반대로, 프라이머 결합 절편(segment) PA 및 PB는 서로 겹치지 않거나, 어닐링될 수 있는 링커 서열들 또는 DNA 절편(segment)과 겹치지 않도록 만들어지나, 완전한 길이의 집합된(assembled) 폴리뉴클레오타이드를 증폭하는데 사용된 프라이머들에 대한 결합 부위로 역할을 한다. 따라서, 집합 반응(assembly reaction)의 4 단계 및 5 단계에서, 프라이머 결합 절편(segment)들 PA 및 PB에 상보적인 프라이머들이 첨가되어, 조성물은 종래의 PCR 증폭 조건을 겪게 된다. PCR 증폭 조건은 당업자에 의해 적당한 것으로 여겨지는 임의의 PCR 증폭 조건일 수 있으며, 이러한 조건은 *PCR Technology: Principles and Applications for DNA Amplication*, ed. HA Erlich, Stockton Press, New York, N.Y. (1989); *PCR Prot° Cols: A Guide to Methods and Applications*, eds. Innis, Gelfland, Snisky, and White, Academic Press, San Diego, Calif. (1990); Mattila *et al.* (1991) *Nucleic acids Res.* 19: 4967; Eckert, K. A. and Kunkel, T. A. (1991) PCR Methods and Applications 1 : 17; 및 미국 특허등록 제 4,683,202호 및 제 4,965,188호에 제시된 조건들을 포함하며, 이들은 각각 인용에 의해 본원에 통합된다. 몇몇 실시예에서, 집합 반응(assembly reaction)의 PCR 단계는 결합 절편(segment)들 PA 및 PB에 결합하는 프라이머에 상보적인 프라이머의 존재시에, 약 35 사이클의 변성, 어닐링 및 신장을 포함한다. 몇몇 실시예에서, PCR의 어닐링 및 신장 단계는 모두 72℃에서 실행될 수 있다. 그러나, 성공적인 증폭을 위한 선택적 조건은 내열성의 DNA 폴리머라제와 사용된 어닐링될 수 있는 링커 서열들에 따라 다르며, 따라서 이러한 조건들은 조정될 수 있다는 것을 당업자는 이해할 것이다.

[0113] 선택적으로, 집합된(assembled) 폴리뉴클레오타이드는 당업자에게 명백한 임의의 기술, 예를 들어, 겔 전기영동 정제 방법에 의해 정제될 수 있으며, 다양한 목적으로 사용된다. 예를 들어, 집합된(assembled) 폴리뉴클레오타이드는 서열 검증을 위한 발현 벡터 골격으로 삽입될 수 있다.

[0114] **3 집합된(assembled) 폴리뉴클레오타이드들을 포함하는 숙주 세포들을 제조하는 방법**

[0115] 다른 측면에서, 본 발명은 집합된(assembled) 폴리뉴클레오타이드들을 포함하는 숙주 세포들을 제조하는 방법을 제공한다. 몇몇 실시예에서, 집합된(assembled) 폴리뉴클레오타이드는 적어도 3 kb의 크기이다. 다른 실시예에서, 집합된(assembled) 폴리뉴클레오타이드는 적어도 5 kb의 크기이다. 또다른 실시예에서, 집합된(assembled) 폴리뉴클레오타이드는 적어도 6, 7, 8, 9 또는 10 kb의 크기이다. 또다른 실시예에서, 집합된(assembled) 폴리뉴클레오타이드는 10 kb를 초과하는 크기이다. 또다른 실시예에서, 집합된(assembled) 폴리뉴클레오타이드는 15 kb를 초과하는 크기이다. 또다른 실시예에서, 집합된(assembled) 폴리뉴클레오타이드는 20 kb를 초과하는 크기이다.

[0116] 몇몇 실시예에서, 본 발명은 본원에 제시된 폴리뉴클레오타이드의 집합 방법에 의해 제조된 집합된(assembled) 폴리뉴클레오타이드로 숙주 세포를 형질 변환하는 것을 포함하는 방법을 제공한다. 집합된(assembled) 폴리뉴클레오타이드는 형질 변환 전에 원형으로 만들어질 수 있거나, 선형 분자로서 형질 변환될 수 있다. 집합된(assembled) 폴리뉴클레오타이드는 염색체 외의(extrachromosomal) 폴리뉴클레오타이드로서 숙주 세포에서 유지될 수 있다. 선택적으로, 집합된(assembled) 폴리뉴클레오타이드는, 예를 들어, 숙주 세포 매개 상동 재조합(homologous recombination)에 의해, 숙주 세포의 게놈으로 통합될 수 있다. 상동 재조합(homologous recombination)에 의해, 집합된(assembled) 폴리뉴클레오타이드를 게놈으로 통합하기 위하여, 집합된(assembled) 폴리뉴클레오타이드는, 한쪽 말단에서 업스트림 게놈 타겟팅(targeting) 서열을 포함하는 핵산 서

열을 포함하며, 다른 쪽 말단에서 다운스트림 게놈 타겟팅(targeting) 서열을 포함하는 핵산 서열을 포함한다. 따라서, 숙주 세포의 염색체로 통합되는 집합된(assembled) 폴리뉴클레오타이드는 업스트림 염색체의 타겟팅(targeting) 서열을 포함하는 첫번째 핵산 분자;와 다운스트림 염색체의 타겟팅(targeting) 서열을 포함하는 마지막 핵산 분자;를 포함하는 집합체 조성물로부터 형성되며, 각각의 염색체의 타겟팅(targeting) 서열은 상기 염색체를 갖는 숙주 세포에 의해 상동 재조합(homologous recombination)을 개시할 수 있는 충분한 길이를 갖는다.

[0117] 다른 실시예에서, 본 발명의 방법들은 본원에 제시된 폴리뉴클레오타이드 집합 방법에 의해 제조된 수많은 집합된(assembled) 폴리뉴클레오타이드들로 숙주 세포를 형질 변환하는 것을 포함한다. 특정 실시예에서, 숙주 세포는 2개 이상의 집합된(assembled) 폴리뉴클레오타이드들을 상동 재조합(homologous recombination)에 의해 단일 결합된 폴리뉴클레오타이드와 결합시킨다. 결합된(combined) 폴리뉴클레오타이드들을 포함하는 숙주 세포 형질 변환체들은 집합된(assembled) 폴리뉴클레오타이드들을 결합시키는 과정에서 제조된 선별표지(selectable marker)를 표시함으로써 선택된다. 상동 재조합(homologous recombination)에 의해, 상대적으로 큰 조각의 폴리뉴클레오타이드를 타겟 폴리뉴클레오타이드에 삽입하는데 상기 방법이 특히 유용하다. 염색체의 통합(integration)이 발생하기 위하여, 결합된 폴리뉴클레오타이드는 선별표지(selectable marker)의 코딩 서열의 5' 또는 3'에 위치한 업스트림 게놈 타겟팅(targeting) 서열과 선별표지(selectable marker)의 코딩 서열의 3' 또는 5'에 위치한 다운스트림 게놈 타겟팅(targeting) 서열을, 각각 포함하여야 한다. 본원에서 사용된 게놈 통합(integration)은 염색체의 통합(integration), 예를 들어, 숙주 세포의 염색체로의 폴리뉴클레오타이드의 통합을 포함한다. *Saccharomyces cerevisiae* 내 적당한 염색체의 통합(integration) 부위는, 그러나 이에 제한되지 않는, *NDT80*, *HO*, *GAL2* 및 *GAL1-GAL 10-GAL7* 위치를 포함한다. 상기 방법은, 염색체 외로(extrachromosomally) 유지된 폴리뉴클레오타이드, 예를 들어, 벡터 및 발현 플라즈미드를 포함하는 숙주 세포를 제조하는데 유용할 수 있다. 결합된(combined) 폴리뉴클레오타이드가, 동일한 어닐링될 수 있는 링커 서열들이나, 그 외에는 폴리뉴클레오타이드 성분의 절편(segment)의 절단을 유발하는 추가적인 상동 재조합(homologous recombination) 이벤트를 개시할 수 있는 직접적 반복으로 배열된 DNA 절편(segment)을 포함하지 않을 때, 염색체로 통합되거나 염색체 외로(extrachromosomally) 유지된 결합된(combined) 폴리뉴클레오타이드의 안정성이 상승한다. 따라서, 몇몇 실시예에서, 집합된(assembled) 폴리뉴클레오타이드들은 특유의 어닐링될 수 있는 링커 서열들 및 DNA 절편(segment)을 포함한다. 다른 실시예에서, 집합된(assembled) 폴리뉴클레오타이드들은 하나 이상의 동일한 어닐링될 수 있는 링커 서열 또는 DNA 절편(segment)을 포함하며, 여기에서 집합된(assembled) 폴리뉴클레오타이드들의 조합은 결합된(combined) 폴리뉴클레오타이드에서 역반복(inverted repeat)으로 배열된다.

[0118] 예시적인 결합된(combined) 폴리뉴클레오타이드의 제조와, 상동 재조합(homologous recombination)에 의한, 결합된(combined) 폴리뉴클레오타이드의 숙주 세포의 염색체로의 통합(integration)은 도 8에 나타나 있다. 2개의 집합된(assembled) 폴리뉴클레오타이드들 A 및 B는 상동 재조합(homologous recombination)을 할 수 있는 숙주 세포에 흡수된다. 각각의 집합된(assembled) 폴리뉴클레오타이드는 선별표지(selectable marker)의 절편을 코딩하는 DNA 절편 D_m을 포함하며, 여기에서, 집합된(assembled) 폴리뉴클레오타이드 A의 DNA 절편 D_{m1}은 선별표지(selectable marker)의 첫번째 절편을 코딩하며, 집합된(assembled) 폴리뉴클레오타이드 B의 DNA 절편 D_{m2}는 선별표지(selectable marker)의 두번째 절편을 코딩하며, 여기에서, DNA 절편 D_{m1} 및 DNA 절편 D_{m2}는 숙주 세포 매개 상동 재조합(homologous recombination)을 개시하는데 충분한 상동(homology) 영역을 포함하며, 여기에서, DNA 절편 D_{m1} 또는 DNA 절편 D_{m2} 중 그 어느 것도 기능적인 선별표지(selectable marker)를 생산하지 않지만, 숙주 세포에 의한 상동 재조합(homologous recombination)에서 기능적인 선별표지(selectable marker)가 형성된다. 각각의 집합된(assembled) 폴리뉴클레오타이드는 숙주 매개 상동 재조합(homologous recombination)을 개시하는데 충분한 길이의 염색체의 타겟팅(targeting) 서열을 코딩하는 DNA 절편 D_0를 더 포함하며, 여기에서 집합된(assembled) 폴리뉴클레오타이드 A의 DNA 절편 D_{01}은 업스트림 염색체의 타겟팅(targeting) 서열을 코딩하며, 집합된(assembled) 폴리뉴클레오타이드 B의 DNA 절편 D_{02}는 다운스트림 염색체의 타겟팅(targeting) 서열을 코딩한다. 세포 내에서 한번, 숙주 세포는, 결합된(combined) 폴리뉴클레오타이드를 형성하기 위하여 DNA 절편(segment) D_{m1} 및 D_{m2} 내 상동(homology) 영역에서, 집합된(assembled) 폴리뉴클레오타이드 A와 집합된(assembled) 폴리뉴클레오타이드 B를 재결합시킨다. 또한, 숙주 세포는 상동 재조합(homologous recombination)에 의해 결합된(combined) 폴리뉴클레오타이드를 염색체로 삽입하기 위하여, DNA 절편(segment) D_{01} 및 D_{02}에 의해 코딩된 염색체의 타겟팅(targeting) 서열들을 사용한다. 결합된(combined) 폴리뉴클레오타이

드를 포함하는 숙주 세포들은 형성된 기능적 선별표지(selectable marker)에 기반하여 용이하게 식별될 수 있다.

[0119] 또다른 실시예에서, 본 발명의 방법들은 수많은 폴리뉴클레오타이드 성분들로 숙주 세포를 형질 변환하는 것과, 숙주 세포가 상동 재조합(homologous recombination)에 의해 하나 이상의 집합된(assembled) 폴리뉴클레오타이드들을 형성하는 것을 포함한다. 집합된(assembled) 폴리뉴클레오타이드는 숙주 세포에서 염색체 외로(extrachromosomally) 유지될 수 있거나 숙주 세포의 염색체로 통합될 수 있다. 숙주 세포에서 상동 재조합(homologous recombination)에 의한 예시적인 집합된(assembled) 폴리뉴클레오타이드의 형성과, 집합된(assembled) 폴리뉴클레오타이드를 숙주 세포의 염색체로의 통합(integration)은 도 9에 나타나 있다. 첫번째 단계에서, 집합 벡터들을 포함하는 집합체 조성물은 SapI 또는 LguI와 같은 IIS형 제한 엔도뉴클레아제에 의해 분해되어, 폴리뉴클레오타이드 성분들의 집합 벡터 골격으로부터의 절단을 유발한다. 상기 실시예에서, D_0 및 D_3은 업스트림 및 다운스트림 염색체의 타겟팅(targeting) 서열이 될 수 있으며, 이 경우에 첫번째 및 마지막 집합 벡터 내의 프라이머 결합 절편의 존재는 선택적이다. 선택적으로, 2개의 프라이머 결합 절편(segment)은 업스트림 및 다운스트림 게놈 타겟팅(targeting) 서열들로서 기능할 수 있다.

[0120] 일단 절단되면, 각각의 절단된 폴리뉴클레오타이드 성분은 다른 폴리뉴클레오타이드 성분의 어닐링될 수 있는 링커 서열 LA에 상동이고, 숙주 매개 상동 재조합(homologous recombination)을 개시하는데 충분한 길이의 어닐링될 수 있는 링커 서열 LB를 포함한다. 첫번째 집합 벡터로부터 절단된 폴리뉴클레오타이드 성분은 업스트림 염색체의 타겟팅(targeting) 서열을 더 포함하며, 마지막 집합 벡터로부터 절단된 폴리뉴클레오타이드 성분은 다운스트림 염색체의 타겟팅(targeting) 서열을 더 포함하며, 여기에서 상기 2개의 염색체의 타겟팅(targeting) 서열들 모두 숙주 세포의 염색체를 갖는, 숙주 매개 상동 재조합(homologous recombination)을 개시하는데 충분한 길이이다. 따라서 제한 엔도뉴클레아제는 불활성화될 수 있다. 상기 방법의 두번째 단계에서, 성분들 조성물은 상동 재조합(homologous recombination)을 할 수 있는 숙주 세포로 도입된다. 세포 내에서 한번, 숙주 세포는 집합된(assembled) 폴리뉴클레오타이드를 형성하기 위하여 어닐링될 수 있는 링커 서열들 사이의 상동(homology) 영역에서 폴리뉴클레오타이드 성분들을 재결합시키며, 집합된(assembled) 폴리뉴클레오타이드는 염색체로 통합된다. 집합된(assembled) 폴리뉴클레오타이드를 포함하는 숙주 세포들은 집합된(assembled) 폴리뉴클레오타이드의 DNA 절편에 의해 코딩된 선별표지(selectable marker)에 기반하여 용이하게 식별될 수 있다.

[0121] 임의의 숙주 세포는 본원에 제시된 발법들에 사용될 수 있다. 특정 실시예에서, 적당한 숙주 세포들은, 본원에서 제공된 선별표지(selectable marker) 절편(segment), 게놈 타겟팅(targeting) 서열들 및 어닐링될 수 있는 링커 서열들에 의해 제공된 바와 같은, 상보적인 서열 스트레치(stretch)에 기반한 폴리뉴클레오타이드들을 재결합시킬 수 있는 숙주 세포들이다. 예시적인 실시예의 상기와 같은 숙주 세포들은, 그러나 이에 제한되지 않는, *Saccharomyces cerevisiae*을 포함한다. 그러한 숙주 세포들에 의한 DNA의 흡수에 적당한 조건은 업계에 잘 알려져 있다.

[0122] 집합된(assembled) 또는 결합된(combined) 폴리뉴클레오타이드를 포함하는 숙주 세포 형질변환체는, 세포 성장을 위하거나, 세포 성장에 대항하는 선택을 허용하는, 집합된(assembled) 폴리뉴클레오타이드에 의하거나, 결합된(combined) 폴리뉴클레오타이드에 의해 코딩된 선별표지(selectable marker)의 표현에 의해 용이하게 식별될 수 있다. 선별표지(selectable marker)는 집합체 조성물의 집합 벡터에 존재하는 단일 DNA 절편에 의해 코딩될 수 있다. 선택적으로, 선별표지(selectable marker)의 비기능적 절편(segment)은 집합체 조성물의 다수의 집합 벡터들 내 또는 다수의 집합된(assembled) 폴리뉴클레오타이드들에 존재하는 DNA 절편(segment)에 의해 코딩될 수 있어서, 기능적 선별표지(selectable marker)는 집합된(assembled) 폴리뉴클레오타이드의 형성 또는 결합된(combined) 폴리뉴클레오타이드의 형성시에만 각각 제조된다.

[0123] 다양한 선별표지(selectable marker)가 업계에 알려져 있다 (예를 들어, 하기의 것들을 참조 : Kaufman, *Meth. Enzymol.*, 185:487 (1990); Kaufman, *Meth. Enzymol.*, 185:537 (1990); Srivastava and SchIessinger, *Gene*, 103:53 (1991); Romanos *et al.*, in DNA Cloning 2: Expression Systems, 2nd Edition, pages 123-167 (IRL Press 1995); Markie, *Methods Mol. Biol.*, 54:359 (1996); Pfeifer *et al.*, *Gene*, 188:183 (1997); Tucker and Burke, *Gene*, 199:25 (1997); Hashida-Okado *et al.*, *FEBS Letters*, 425:117 (1998)). 몇몇 실시예에서, 선별표지(selectable marker)는 약물 저항 표지(drug resistant marker)이다. 상기 약물 저항 표지는, 만약 독성을 제거하지 않으면 세포를 죽이는 외인성(exogenous) 약물의 독성을 제거할 수 있게 한다. 약물 저항 표지의 예시적인 실시예는, 그러나 이에 제한되지 않는, 앰피실린, 테트라사이클린, 카나마이신, 블레오마이신, 스트렙토마이신, 하이그로마이신(hygromycin), 네오마이신, Zeocin™, 기타 등등과 같은 항생제에 대한 저항성을 부여

하는 것들을 포함한다. 다른 실시예에서, 선별표지(selectable marker)는 영양요구 표지(auxotrophic marker)이다. 영양요구 표지는 필수 성분(통상적으로 아미노산)이 부족한 배지에서 세포가 성장하는 동안, 상기 필수 성분을 합성할 수 있게 한다. 선택가능한 영양요구 유전자 서열들은, 예를 들어, 히스티디놀(histidinol)의 존재하에 히스티딘이 없는 배지에서 성장할 수 있게 하는, hisD을 포함한다. 기타 선별표지(selectable marker)는 블레오마이신-저항 유전자, 메탈로티오네인(metallothionein) 유전자, 하이그로마이신(hygromycin) B-포스포트랜스퍼라제 유전자, AURI 유전자, 아데노신 디아미나아제 유전자, 아미노글리코시드 포스포트랜스퍼라제 유전자, 디하이드로폴레이트 리덕타아제(dihydrofolate reductase) 유전자, 티미딘 키나아제 유전자, 잔틴-구아닌 포스포리보실트랜스퍼라아제(phosphoribosyltransferase) 유전자, 기타 등등을 포함한다.

[0124] 집합된(assembled) 또는 결합된(combined) 폴리뉴클레오타이드의 통합(integration)이, 숙주 세포가 세포 성장에 필요한 성분을 합성하여, 세포를 영양요구성(auxotrophic)이 되도록 요구하는 유전자의 붕괴를 유발할 때, 영양요구성(auxotrophy)은 염색체로 통합된, 집합되거나(assembled) 또는 결합된(combined) 폴리뉴클레오타이드를 포함하는 숙주 세포 형질변환체를 식별하는데 또한 사용될 수 있다.

[0125] 염색체적으로 통합된, 집합되거나(assembled) 결합된(combined) 폴리뉴클레오타이드를 포함하는 숙주 세포 형질변환체는 각각의 DNA 절편(segment)에 의해 코딩된 기타 소질(trait)을 발휘하는 숙주 세포 형질변환체를 선택하거나, DNA 절편(segment)의 조합, 예를 들어, 빛을 방출하는 펩타이드들의 발현에 의하거나, 각각의 숙주 세포 클로니의 분자 분석에 의하거나, 예를 들어, 제한효소 매핑(mapping), PCR 증폭, 또는 분리된 집합된(assembled) 폴리뉴클레오타이드들 또는 염색체의 통합(integration) 부위의 서열 분석에 의해 또한 식별될 수 있다.

[0126] **4 폴리뉴클레오타이드 집합체의 조합 방법과 숙주 세포 형성**

[0127] 다른 측면에서, 본 발명은 다수의 폴리뉴클레오타이드 성분들을 수많은 집합된(assembled) 폴리뉴클레오타이드들에 삽입한 정렬된 집합체에 대한 빠르고, 강건하며 높은 생산량의(high-throughput) 방법들을 제공한다. 상기 방법들은 집합 벡터들을 포함하는 집합체 조성물의 사용에 의존하며, 각각의 집합 벡터는, 어닐링될 수 있는 링커 서열 LA 또는 LB, 한 쌍의 어닐링될 수 있는 링커 서열들 LA 및 LB의 측면에 배치되거나, 한 쌍의 제한 부위들 RA 및 RB의 측면에 배치된, 어닐링될 수 있는 링커 서열과 프라이머 결합 절편, 예를 들어, LA과 PB, 또는 LB와 PA(도 1b)의 측면에 배치된, DNA 절편 D를 포함한다. 그러나, 본원에 제시된 방법들을 사용하는 다양한 집합된(assembled) 폴리뉴클레오타이들을 형성하기 위하여, 어닐링될 수 있는 링커 서열들과 프라이머 결합 절편(segment)이 선택되어 폴리뉴클레오타이드 성분들의 하나 이상의 조합은 반응에서 집합된(assembled) 폴리뉴클레오타이드로 집합될 수 있다. 따라서, 몇몇 실시예에서, 집합체 조성물은, 동일한 어닐링될 수 있는 링커 서열 LA 또는 LB, 또는 동일한 프라이머 결합 절편 PA 또는 PB를 갖지만, DNA 절편과 다른, 적어도 2개의 집합 벡터들을 포함한다. 또다른 실시예에서, 집합체 조성물은, 동일한 한 쌍의 어닐링될 수 있는 링커 서열들 LA 및 LB, 또는 동일한 어닐링될 수 있는 링커 서열과 프라이머 결합 절편 쌍, 예를 들어, LA 및 PB, 또는 LB 및 PA를 갖지만 DNA 절편과 다른, 적어도 2개의 집합 벡터들을 포함한다.

[0128] 도 10은 동일한 반응에서 7개의 폴리뉴클레오타이드 성분들로부터 수많은 집합된(assembled) 폴리뉴클레오타이드들을 형성하는 예시적인 방법을 나타낸다. 집합된(assembled) DNA 절편(segment)을 포함하는 집합 벡터들을 단일 튜브에 넣고 SapI로 분해하여 집합 벡터 골격들로부터 폴리뉴클레오타이드 절편 성분들을 방출한다. SapI의 열 불활성화 후에, 폴리뉴클레오타이드 성분들을 변성 상태에 둔 후, 상보적인 어닐링될 수 있는 링커(annealable linker) 쌍의 하이브리드화에 충분한 어닐링 조건에 둔다. DNA 폴리머라제 및 dNTP들의 존재하에 프라이머 신장을 한 후, 프라이머 결합 절편들(segment) PA 및 PB에 상보적인 프라이머들을 다양한 가능한 조합으로 집합된 DNA 절편(segment) $D_{01/02}$, $D_{1/2}$, D_3 및 $D_{41/42}$를 포함하는 8개의 서로 다른 완전한 길이의 집합된(assembled) 폴리뉴클레오타이드들을 PCR 증폭하기 위하여 첨가한다. 각각의 집합된(assembled) 폴리뉴클레오타이드들은, 예를 들어, DNA 절편(segment) D_{01}, D_{02}, D_{41} 및 D_{42}의 영역에 상보적인 프라이머들을 이용한 성공적인 PCR 증폭에 의해, 혼합된 집합된(assembled) 폴리뉴클레오타이드들의 조성물로부터 분리될 수 있다. 선택적으로, 한 세트의 집합된(assembled) 폴리뉴클레오타이드들은 프라이머 결합 절편(segment) PA 및/또는 PB의 그룹의 하나를 포함하고, 프라이머 결합 절편(segment) PA 및 PB의 선택적인 서브그룹과 하이브리드하는 PCR 증폭용 프라이머를 사용하는, 첫번째 및 마지막 집합 벡터들에 의해 분리될 수 있다. 분리된 집합된(assembled) 폴리뉴클레오타이드들은, 예를 들어, 집합된(assembled) 폴리뉴클레오타이드들을 포함하는 수많은 숙주 세포들을 형성하기 위하여 숙주 세포들을 형질 변환하는데 사용될 수 있다. 선택적으로, 숙주 세포들은 혼합된 집합된

(assembled) 폴리뉴클레오타이드들과 숙주 세포 형질변환체의 조성물로 직접적으로 형질전환될 수 있으며, 각각의 집합된(assembled) 폴리뉴클레오타이드는, 예를 들어, 각각의 숙주 세포 콜로니의 분자 분석에 의하거나, 선별표지(selectable marker)를 포함하거나 각각의 DNA 절편(segment) 또는 DNA 절편(segment)의 조합에 의해 코딩된 기타 소질(trait)을 나타내는 숙주 세포 형질변환체를 선택함으로써, 분리될 수 있다.

[0129] 다른 실시예에서, 조합 방법에 의해 집합된(assembled) 수많은 폴리뉴클레오타이드들을 포함하는 수많은 숙주 세포들은 다수의 집합된(assembled) 폴리뉴클레오타이드들을 포함하는 조성물로 숙주 세포들을 형질전환하고, 결합된(combined) 폴리뉴클레오타이드를 포함하는 숙주 세포들을 선택함으로써 형성되며, 상기 집합된(assembled) 폴리뉴클레오타이드들의 적어도 2개의 집합된(assembled) 폴리뉴클레오타이드들은 기능적 선별표지(selectable marker)를 형성하는 상동 재조합(homologous recombination)에서 선별표지(selectable marker)의 비-기능적 절편(segment)을 포함한다. 도 11은 결합된(combined) 폴리뉴클레오타이드들을 포함하는 수많은 숙주 세포들을 형성하기 위한 조합적 접근을 나타낸다. 실시예에서, 동일한 업스트림 염색체의 타겟팅(targeting) 서열과 선별표지(selectable marker)의 동일한 첫번째 부분을 포함하는 각각의 집합된(assembled) 폴리뉴클레오타이드들 A1 및 A2, 동일한 다운스트림 염색체의 타겟팅(targeting) 서열과 선별표지(selectable marker)의 동일한 두번째 부분을 포함하는 각각의 집합된(assembled) 폴리뉴클레오타이드들 B1 및 B2는, 4개의 서로 다른 숙주 세포들을 형성하는 염색체로 삽입될 수 있는, 4개의 서로 다른 결합된(combined) 폴리뉴클레오타이드들, A1/B1, A1/B2, A2/B1 및 A2/B2를 형성하는 숙주 세포 매개 상동 재조합(homologous recombination)에 의해 조합적으로 결합된다.

[0130] 또다른 실시예에서, 조합적 방법에 의해 집합되고(assembled) 결합된(combined) 수많은 숙주 세포들은 다수의 폴리뉴클레오타이드 성분들을 포함하는 성분 조성물(component composition)로 숙주 세포들을 형질변환하고, 집합되거나(assembled) 결합된(combined) 폴리뉴클레오타이드를 포함하는 숙주 세포들을 선택함으로써 형성되며, 상기 폴리뉴클레오타이드들 중 적어도 2개의 폴리뉴클레오타이드 성분들은, 숙주 세포 매개 상동 재조합(homologous recombination)이 기능적 선별표지(selectable marker)를 형성하는 선별표지(selectable marker)의 비-기능적 절편(segment)을 포함한다.

[0131] **5 엔트리 벡터(entry vector)**

[0132] 다른 측면에서, 본 발명은, 예를 들어, 집합 벡터를 제조하는데 사용될 수 있는, 엔트리 벡터(entry vector)를 제공한다. 몇몇 실시예에서, 엔트리 벡터(entry vector)는 선별표지(selectable marker), 복제의 원점(origin), 및 본원에 제공된 집합 방법에서 집합되는(assembled) 서로 다른 DNA 절편(segment)의 서브크로닝을 촉진하는 2개의 제한 부위에 의해 즉각적으로 측면에 배치되는 DNA 절편을 포함하는 원형 폴리뉴클레오타이드이다. 엔트리 벡터(entry vector)는, 제한 부위의 측면에 배치되는, 1개 또는 2개의 어닐링될 수 있는 링커 서열들, 또는 어닐링될 수 있는 링커 서열 및 프라이머 결합 절편을 더 포함한다. 엔트리 벡터(entry vector)는 DNA 절편의 측면의 바깥쪽에 위치하는, 예를 들어, 1개 또는 2개의 어닐링될 수 있는 링커 서열들, 또는 어닐링될 수 있는 링커 서열과 프라이머 결합 절편의 측면에 위치하는, 추가적 쌍의 제한 부위를 더 포함한다. 따라서, 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ 및 제한 부위 RB를 포함하는 원형 폴리뉴클레오타이드이다. 다른 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는 원형 폴리뉴클레오타이드이다. 다른 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA 또는 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 프라이머 결합 절편 PB 또는 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는 원형 폴리뉴클레오타이드이다.

[0133] 몇몇 실시예에서, 엔트리 벡터(entry vector)의 DNA 절편 D의 서열은 lac Z 리포터 유전자(reporter gene)이다. 상기 lac Z 리포터 유전자(reporter gene)는, 예를 들어, 본원에 제시된 집합 벡터를 제조하는 동안, lac Z 외의 DNA 절편(segment)을 포함하는 벡터들로 형질변환된 콜로니의 청색/백색 선택(blue/white selection)을 촉진하는데 사용된다.

[0134] 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ 및 제한 부위 RB (예를 들어, 5'-RA-LA-RY-D-RZ-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB 및 제한 부

위 RB (예를 들어, 5'-RA-RY-D-RZ-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-LA-RY-D-RZ-LB-RB-3')을 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-PA-RY-D-RZ-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 엔트리 벡터(entry vector)는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 프라이머 결합 절편 PB 및 제한 부위 RB (예를 들어, 5'-RA-LA-RY-D-RZ-PB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 예시적인 엔트리 벡터(entry vector)는 도 1a에 제시하였다.

[0135] 프라이머 결합 절편은 집합된(assembled) 폴리뉴클레오타이드를 제조하는데 사용된 임의의 어닐링될 수 있는 링커 서열들과 상보적인 임의의 뉴클레오타이드 서열일 수 있다. 몇몇 실시예에서, 2개의 프라이머 결합 절편은 제한 엔도뉴클레아제 인지 및 절단 부위를 포함한다. 몇몇 실시예에서, 프라이머 결합 절편은 특정한 집합 반응(assembly reaction)에서 사용되지 않는, 단순하게 이용가능한 링커 서열 중 하나이다. 몇몇 실시예에서, 프라이머 결합 절편 PA의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 프라이머 결합 절편 PB의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 프라이머 결합 절편 PA 및 프라이머 결합 절편 PB의 핵산 서열들은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 바람직한 실시예에서, PA와 PB는 서열에서 동일하지 않다.

[0136] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LA 또는 LB의 핵산 서열은 적어도 24개의 뉴클레오타이드이며 적어도 60℃의 Tm을 갖는다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LA의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LA 및 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다.

[0137] 제한 부위 RY 및 RZ는 집합 벡터의 형성을 위한 다양한 DNA 절편(segment)을 도입하기 위한 클로닝 부위들로서 사용될 수 있다. 몇몇 실시예에서, RY와 RZ는 서열에서 동일하지 않다. 몇몇 실시예에서, RY와 RZ는 동일한 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, RY와 RZ는 서열에서 동일하다. 몇몇 실시예에서, 제한 부위 RY와 RZ는 엇갈린(staggered) 말단들, 예를 들어 5' 또는 3' 오버행(overhang)을 갖는 말단을 형성하는 제한 엔도뉴클레아제에 의해 절단될 수 있다. 다른 실시예에서, 제한 부위 RY와 RZ는 평활 말단(blunt end)을 형성하는 제한 엔도뉴클레아제에 의해 절단될 수 있다.

[0138] 제한 부위 RY와 RZ가 업계에 알려진 임의의 제한 부위일 수 있더라도, IIS형 제한 엔도뉴클레아제들에 의해 인지된 제한 부위는 특히 유용하다. IIS형 제한 엔도뉴클레아제들은 절단 영역에서 떨어진 DNA 결합 영역을 갖는다. 따라서, IIS형 제한 엔도뉴클레아제들은 특정 서열을 인지하지만, 어느 정도 떨어진 거리에서 절단한다. 예를 들어, IIS형 제한 엔도뉴클레아제 SchI (MIyI로도 또한 알려져 있음)는 서열 GAGTC를포함하는 인지 부위에 결합하며, 인지 부위로부터 떨어진 4개의 염기 쌍을 절단하여, 평활 말단(blunt end) DNA 분자를 형성한다. IIS형 제한 부위는 엔트리 벡터(entry vector)로부터 집합 벡터를 제조하는데 특히 유용하다. 예를 들어, 서브클로닝 과정에서, 엔트리 벡터(entry vector)의 DNA 절편, 예를 들어 lacZ는, 관심있는 DNA로 대체되며, IIS형 제한 엔도뉴클레아제를 사용한 lacZ의 절단은 제한 부위 인식서열(recognition sequence)의 완전한 제거를 유발할 수 있다. 그 결과, 선형의 엔트리 벡터(entry vector)에 대하여 관심있는 DNA 절편의 연결에서, 어닐링될 수 있는 링커 서열 또는 프라이머 결합 절편 사이의 관련없는(extraneous) 서열과 새롭게 도입되는 DNA 절편은 최소화된다.

[0139] 따라서, 몇몇 실시예에서, 제한 부위 RY 및 RZ는 업계에 알려진 임의의 IIS형 제한 엔도뉴클레아제에 의해 인지가능하고 절단될 수 있는 제한 부위이다. 적당한 IIS형 제한 엔도뉴클레아제들은, 그러나 이에 제한되지 않는, 하기의 엔도뉴클레아제들과, 괄호 안에 표시된 이들의 아이소스키조머(isoschizomer)들을 포함한다: Alw26I (BsmAI), AlwI (AclWI, BinI), AsuHPI (HphI), BbvI (Bst71I), BcefI, BstF5I (BseGI, FokI), FauI, HgaI, SapI (LguI), MboII, PleI, SapI, SchI (MlyI), SfaNI 및 TspRI, AceIII, BbsI (BbvII, Bpil, BpuAI), Bce83I, BciVI, BfiI (BmrI), BpmI (GsuI), BsaI (Eco31I), BseRI, BsgI, BsmBI (Esp3I), BsmFI, BspMI, BsrDI (Bse3DI), Bsu6I (Eaml104I, EarI, Ksp632I), Eco57I, FauI, MmeI, RleAI, TaqII 및 Tth111II. 특정한 실시예에서, 제한 부위 RY 및 RZ은 SchI 제한 엔도뉴클레아제에 의해 인지가능하며 절단될 수 있다.

[0140] 몇몇 실시예에서, RA와 RB는 서열에서 동일하지 않다. 몇몇 실시예에서, RA와 RB는 동일한 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, RA와 RB는 서열에서 동일하다. 몇몇 실시예에서, 제한 부위들 RA 및 RB는 엇갈린(staggered) 말단, 예를 들어, 5' 또는 3' 오버행(overhang)을 갖는 말단을 형성하는 제한 엔도뉴클레아제에 의해 절단될 수 있다. 다른 실시예에서, 제한 부위들 RA 및 RB는 평활 말단(blunt end)을 형성하는 제한 엔도뉴클레아제에 의해 절단될 수 있다.

[0141] 제한 부위들 RA 및 RB는 업계에 알려진 임의의 제한 부위일 수 있지만, 하나 이상의 생물의 DNA (예를 들어, cDNA)에서 다소 드문 제한 부위(예를 들어, 드문 커터(cutter))가 특히 유용하다. 몇몇 실시예에서, 제한 부위들 RA 및 RB는 인간 DNA의 다소 드문 제한 부위를 갖는 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있다. 몇몇 실시예에서, 제한 부위들 RA 및 RB는 마우스 DNA의 다소 드문 제한 부위를 갖는 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있다. 몇몇 실시예에서, 제한 부위들 RA 및 RB는 효모 DNA, 예를 들어, *Saccharomyces cerevisiae*, *Pichia pastoris*, *Kluyveromyces lactis*, *Arxula adeninivorans*, 또는 *Hansenula polymorpha*의 DNA에서 다소 드문 제한 부위를 갖는 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있다. 몇몇 실시예에서, 제한 부위들 RA 및 RB는 박테리아의 DNA, 예를 들어, *Escherichia coli* 또는 *Bacillus subtilis*의 DNA에서 다소 많지 않은 제한 부위를 갖는 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있다.

[0142] 몇몇 실시예에서, 제한 부위들 RA 및 RB는 IIS형 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있으며, 상기 인지 부위는, 예를 들어, PA/LA-D-PB/LB를 포함하는 폴리뉴클레오타이드 서열과 거리가 멀다. 몇몇 실시예에서, 각각의 제한 부위들 RA 및 RB는 MssI, NruI (Bsp68I, MluB2I, Sbo13I, SpoI), SnaBI (BstSNI, Eco105I), SrfI 및 SwaI (BstRZ246I, BstSWI, MspSWI, SmiI), HpaI, HincII, PshAI, OliI, AluI, Alw26I, BalI, DraI, DpnI, EcoR47III, EcoRCRI, EcoRV, FokI, HaeIII, HincII, MboI, MspAlI, NaeI, RsaI, PvuII, ScaI, SmaI, SspI, StuI, XmnI, EcaBC3I, SciI, HincII, DraI, BsaBI, Cac8I, Hpy8I, MlyI, PshAI, SspD51, BfrBI, BsaAI, BsrBI, BtrI, CdiI, CviJI, CviRI, Eco47III, Eco78I, EcoICRI, FnuDII, FspAI, HaeI, LpnI, MlyI, MslI, MstI, NaeI, NlaIV, NruI, NspBII, OliI, PmaCI, PshAI, PsiI, SrfI, StuI, XcaI, XmnI, ZraI 및 이의 아이소스키조머(isoschizomer)로 구성된 군으로부터 선택된 제한 엔도뉴클레아제에 의해 독립적으로 인지 및 절단될 수 있다. 특정 실시예에서, 제한 부위들 RA 및 RB는 SapI 또는 LguI 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있다. LguI는 동일한 인지 및 절단 특이성을 갖는 SapI의 아이소스키조머(isoschizomer)이다.

[0143] 몇몇 실시예에서, 본원에서 제공된 엔트리 벡터(entry vector)는 또한, 벡터 (예를 들어, 복제의 원점(origin))의 복제, 유지 또는 통합의 몇몇 기능을 일반적으로 갖는 하나 이상의 핵산 서열들 뿐만 아니라 선별표지(selectable marker)를 포함한다. 복제 원점(origin)들은 다중 결합의 원점(origin)-결합 단백질에 의해 인지되고 원점(origin) 부위에서 DNA 복제 효소를 결합하는데 중요한 역할을 하는 다수의 짧은 반복 서열들을 포함하는 특유의 폴리뉴클레오타이드들이다. 본원에서 제공된 엔트리 벡터 및 집합 벡터에서의 사용을 위한 복제의 적당한 원점(origin)들은, 그러나 이에 제한되지 않는, *E. coli* oriC, colE1 플라즈미드 원점(origin), 2 μ 및 ARS (둘다 효모 계통에 유용함), sf1, SV40 EBV oriP (포유동물 계통에 유용함), 또는 pSC101에서 발견되는 것들을 포함한다. 선별표지(selectable marker)는, 선별표지(selectable marker)를 포함하는 벡터를 사용하여 성공적으로 형질 변환되고 표지를 발현하는 세포의 성장을 위하거나, 세포의 성장에 대항하기 위하여 선택되는 수단을 제공하기 때문에, 벡터 내에서 유용한 성분이 될 수 있다.

[0144] 몇몇 실시예에서, 임의의 벡터는 본원에 제공된 바와 같은 엔트리 벡터(entry vector)를 구성하는데 사용될 수 있다. 특히, 업계에 알려진 벡터들과 상업적으로 이용가능한 벡터들 (및 이의 변종 또는 유도체)은, 본원에서 제공된 방법에서의 사용을 위한, 제한 부위 RA, 선택적으로 프라이머 결합 절편 PA 또는 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 선택적으로 프라이머 결합 절편 PB 또는 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하기 위하여 조작될 수 있다. 그러한 벡터들은, 예를 들어, Vector Laboratories Inc., InVitrogen, Promega, Novagen, NEB, Clontech, Boehringer Mannheim, Pharmacia, EpiCenter, OriGenes Technologies Inc., Stratagene, Perkin Elmer, Pharmingen, Life Technologies, Inc. 및 Research Genetics로부터 얻을 수 있다. 특별히 관심있는 벡터들의 일반적인 계통은 원핵 및/또는 진핵 클로닝 벡터, 발현 벡터, 융합 벡터, 투-하이브리브(two-hybrid) 또는 역 투-하이브리브(reverse two-hybrid) 벡터, 서로 다른 호스트에서의 사용을 위한 셔틀(shuttle) 벡터, 돌연변이유발(mutagenesis) 벡터, 전사 벡터, 거대한 삽입물을 수용하기 위한 벡터, 기타 등등을 포함한다. 관심있는 기타 벡터들은 바이러스성 오리진 벡터(origin vector) (M 13 벡터, 박테리아 파지(phage) λ 벡터, 아데노바이러스 벡터 및 레트로바이러스 벡터), 많고, 적고, 조정가능한 복제 개수 벡터, 단일 호스트 (PACYC184 및 pBR322)와 진핵 에피솜(episomal) 복제 벡터 (pCDM8)에서의 조합에 사용되는 호환가능한 레플리콘(replicon)을 갖는 벡터를 포함한다.

[0145] 특정한 실시예에서, 본원에서 제공된 방법에 따른 사용을 위한 엔트리 벡터(entry vector)는 서열번호 207 내지 221의 뉴클레오타이드 서열을 갖는 pRYSE 벡터이다. pRYSE 벡터의 도식은 도 4에 나타냈으며 pRYSE 벡터의 제조는 하기 실시예 1에 제시하였다.

[0146] **6 집합 벡터**

[0147] 다른 측면에서, 본 발명은 벡터, 예를 들어, 하나 이상의 집합된(assembled) 폴리뉴클레오타이드들에 수많은 폴리뉴클레오타이드 성분들을 삽입한 집합체에 사용될 수 있는, 집합 벡터를 제공한다. 몇몇 실시예에서, 집합 벡터는 선별표지(selectable marker), 복제의 원점(origin) 및 어닐링될 수 있는 링커 서열, 어닐링될 수 있는 링커 서열 쌍의 측면에 배치된 DNA 절편, 또는 한 쌍의 제한 부위의 측면에 배치된, 어닐링될 수 있는 링커 서열/프라이머 결합 절편 쌍의 측면에 배치된 DNA 절편을 포함하는 원형 폴리뉴클레오타이드이다. 제한 부위는 집합 반응(assembly reaction) 동안, 집합 벡터 골격으로부터 폴리뉴클레오타이드 성분의 절단을 촉진시키는 역할을 할 수 있다. 따라서, 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA 또는 어닐링될 수 있는 링커 서열 LA, DNA 절편 D 및 제한 부위 RB를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, DNA 절편 D, 프라이머 결합 절편 PB 또는 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA 또는 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 프라이머 결합 절편 PB 또는 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB을 포함하는 원형 폴리뉴클레오타이드이다.

[0148] 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D 및 제한 부위 RB (예를 들어, 5'-RA-LA-D-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-D-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-LA-D-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB (예를 들어, 5'-RA-PA-D-LB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 몇몇 실시예에서, 집합 벡터는, 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 프라이머 결합 절편 PB 및 제한 부위 RB (예를 들어, 5'-RA-LA-D-PB-RB-3')를 포함하는 원형 폴리뉴클레오타이드이다. 예시적인 집합 벡터들은 도 1b 및 도 2에 나타냈다.

[0149] 몇몇 실시예에서, 프라이머 결합 절편 PA의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 프라이머 결합 절편 PB의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 프라이머 결합 절편 PA와 프라이머 결합 절편 PB의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 바람직한 실시예에서, 프라이머 결합 절편 PA와 프라이머 결합 절편 PB의 핵산 서열은 동일하지 않다.

[0150] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LA 또는 LB의 핵산 서열은 적어도 24개의 뉴클레오타이드이며 적어도 60℃의 Tm를 갖는다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LA의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열 LA와 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다.

[0151] 몇몇 실시예에서, RA와 RB는 서열에서 동일하지 않다. 몇몇 실시예에서, RA와 RB는 동일한 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, RA와 RB는 서열에서 동일하지 않다. 몇몇 실시예에서, 제한 부위들 RA 및 RB는 엇갈린(staggered) 말단, 예를 들어, 5' 또는 3' 오버행(overhang)을 갖는 말단을 형성하는 제한 엔도뉴클레아제에 의해 절단될 수 있다. 다른 실시예에서, 제한 부위들 RA 및 RB는 평활 말단(blunt end)을 형성하는 제한 엔도뉴클레아제에 의해 절단될 수 있다.

[0152] 제한 부위들 RA 및 RB가 업계에 알려진 임의의 제한 부위가 될 수 있다고 하더라도, 하나 이상의 생물의 DNA (예를 들어, cDNA)에서 다소 드문 제한 부위 (예를 들어, 드문 커터(cutter))는 특히 유용하다. 몇몇 실시예에서, 제한 부위들 RA 및 RB는 인간 DNA의 상대적으로 드문 제한 부위를 갖는 제한 엔도뉴클레아제에 의해 인지

및 절단될 수 있다. 몇몇 실시예에서, 제한 부위들 RA 및 RB는 마우스 DNA의 상대적으로 드문 제한 부위를 갖는 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있다. 몇몇 실시예에서, 제한 부위들 RA 및 RB는 효모 DNA, 예를 들어, *Saccharomyces cerevisiae*, *Pichia pastoris*, *Kluyveromyces lactis*, *Arxula adeninivorans*, 또는 *Hansenula polymorpha*의 DNA에서, 상대적으로 드문 제한 부위를 갖는 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있다. 몇몇 실시예에서, 제한 부위들 RA 및 RB는 박테리아의 DNA, 예를 들어, *Escherichia coli* 또는 *Bacillus subtilis*의 DNA에서 다소 많지 않은 제한 부위를 갖는 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있다.

[0153] 몇몇 실시예에서, 제한 부위들 RA 및 RB는 IIS형 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있다. 적당한 IIS형 제한 엔도뉴클레아제의 예시적인 예는, 그러나 이에 제한되지 않는, MssI, NruI (Bsp68I, MluB2I, Sbo13I, SpoI), SnaBI (BstSNI, Eco105I), SrfI 및 SwaI (BstRZ246I, BstSWI, MspSWI, SmiI), HpaI, HincII, PshAI, OliI, AluI, Alw26I, BalI, DraI, DpnI, EcoR47III, EcoRCRI, EcoRV, FokI, HaeIII, HincII, MboI, MspAlI, NaeI, RsaI, PvuII, ScaI, SmaI, SspI, StuI, XmnI, EcaBC3I, SciI, HincII, DraI, BsaBI, Cac8I, Hpy8I, MlyI, PshAI, SspD51, BfrBI, BsaAI, BsrBI, BtrI, CdiI, CviJI, CviRI, Eco47III, Eco78I, EcoICRI, FnuDII, FspAI, HaeI, LpnI, MlyI, MslI, MstI, NaeI, NlaIV, NruI, NspBII, OliI, PmaCI, PshAI, PsiI, SrfI, StuI, XcaI, XmnI, ZraI, 또는 이의 아이소스키조머(isoschizomer)를 포함한다. 특정 실시예에서, 제한 부위들 RA 및 RB는 SapI 또는 LguI 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있다.

[0154] 바람직하게, 집합 벡터의 DNA 절편은 집합 벡터 내의 임의의 제한 부위 RA 및 RB를 절단할 수 있는 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있는 핵산 서열을 포함하지 않는다. 이것은, 폴리뉴클레오타이드 성분이 집합 벡터 골격으로부터 절단되는 동안, 집합 반응(assembly reaction)의 첫번째 단계 동안 DNA 절편이 손상되지 않는다는 것을 보증한다. 특정한 실시예에서, DNA 절편은 SapI/LguI 부위를 포함하지 않으며 RA와 RB는 SapI 또는 LguI에 의해 절단될 수 있다. 부위-특이적(site-directed) 돌연변이유발(mutagenesis) (참조 : Carter, *Biochem. J.* 237:1-7 (1986); Zoller and Smith, *Methods Enzymol.* 154:329-50 (1987)), 카세트 돌연변이 유발(cassette mutagenesis), 제한 선택돌연변이유발(mutagenesis) (Wells *et al*, *Gene* 34:315-323 (1985)), 올리고뉴클레오타이드-매개 (부위-특이적) 돌연변이유발(mutagenesis), PCR 돌연변이유발(mutagenesis), 기타 알려진 기술들은 엔트리 벡터(entry vector)에 DNA 절편이 연결되기 전 또는 후에 DNA 절편 내에 임의의 상기 서열을 개질하기 위하여 실시될 수 있다.

[0155] 몇몇 실시예에서, 본원에서 제공된 집합 벡터는 벡터의 복제, 유지, 또는 통합의 몇몇 기능 (예를 들어, 복제의 원점(origin)) 뿐만 아니라 하나 이상의 선별표지(selectable marker)를 일반적으로 갖는 하나 이상의 핵산 서열들을 또한 포함한다. 복제 원점(origin)은 다중 결합의 원점(origin)-결합 단백질에 의해 인지되고 원점(origin) 부위에서 DNA 복제 효소를 결합시키는데 중요한 역할을 하는 다수의 짧은 반복 서열들을 포함하는 특이한 폴리뉴클레오타이드이다. 본원에서 제공된 엔트리 벡터 및 집합 벡터에서의 사용을 위한 복제의 적당한 원점(origin)은, 그러나 이에 제한되지 않는, E. coli oriC, colE1 플라즈미드 원점(origin), 2 μ와 ARS (둘 다 효모 계통에서 유용함), sf1, SV40 EBV oriP (포유동물 계통에서 유용함), 또는 pSC101에서 발견되는 것들을 포함한다. 선별표지(selectable marker)를 포함하는 벡터를 사용하여 성공적으로 형질변환되고 상기 표지를 발현하는 세포의 성장을 위하거나, 세포의 성장에 대항하기 위하여 선택되는 수단을 제공하기 때문에, 선별표지(selectable marker)는 벡터에서 유용한 성분이 될 수 있다.

[0156] 몇몇 실시예에서, 임의의 벡터는 본원에서 제공된 바와 같은 집합 벡터를 구성하는데 사용될 수 있다. 특히, 업계에 알려진 벡터들과 상업적으로 이용가능한 벡터들 (및 이의 변종 또는 유도체)은, 본원에서 제공된 방법들에 사용하기 위한, 제한 부위 RA, 프라이머 결합 절편 PA 또는 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 프라이머 결합 절편 PB 또는 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하도록 조작될 수 있다. 그러한 벡터들은, 예를 들어, Vector Laboratories Inc., InVitrogen, Promega, Novagen, NEB, Clontech, Boehringer Mannheim, Pharmacia, EpiCenter, OriGenes Technologies Inc., Stratagene, Perkin Elmer, Pharmingen, Life Technologies, Inc. 및 Research Genetics로부터 얻을 수 있다. 특히 관심있는 벡터들의 일반적인 계열은 원핵 및/또는 진핵 클로닝 벡터, 발현 벡터, 융합 벡터, 투-하이브리브(two-hybrid) 또는 역 투-하이브리브(reverse two-hybrid) 벡터, 서로 다른 호스트에서의 사용을 위한 셔틀 벡터, 돌연변이유발(mutagenesis) 벡터, 전사 벡터, 거대한 삽입물을 수용하기 위한 벡터, 기타 등등을 포함한다. 기타 관심있는 벡터는 바이러스의 오리진 벡터(origin vector) (M13 벡터, 박테리아 파지(phage) λ 벡터, 아데노바이러스 벡터 및 레트로바이러스 벡터), 많고, 적고, 조정가능한 복제 개수 벡터, 단일 호스트 (PACYC 184 및 pBR322)와 진핵 에피솜(episomal) 복제 벡터 (pCDM8)에서 조합에 사용하기 위한 호환가능한 레플리콘(replicon)을 갖는 벡터를 포함한다.

[0157] 집합 벡터는 엔트리 벡터(entry vector)로부터 제조될 수 있다. 엔트리 벡터(entry vector)로부터 집합 벡터를 제조하기 위하여, 엔트리 벡터(entry vector)는 벡터를 선형으로 만들어서 RY와 RZ를 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제를 사용하여 분해될 수 있으므로, 상기 엔트리 벡터는 DNA 절편을 수용할 수 있다. 본 발명의 집합 벡터를 형성하기 위하여 표준 클로닝 기술을 사용하여, DNA 절편은 RY와 RZ 부위에 연결될 수 있다. 예를 들어, DNA 절편은, 화학적 합성에 의하거나, cDNA 클로닝에 의하거나, 또는 원하는 세포로부터 정제된, 게놈 DNA, 또는 이의 절편의 클로닝에 의하거나, PCR 증폭과 클로닝에 의해, 복제된 DNA (예를 들어, DNA "라이브러리")로부터 업계에 알려진 표준 과정에 의해 수득될 수 있다. 예를 들어, Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual*, 3d. ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (2001); Glover, D.M. (ed.), *DNA Cloning: A Practical Approach*, 2d. ed., MRL Press, Ltd., Oxford, U.K. (1995)를 참조한다.

[0158] 집합 벡터는 어닐링될 수 있는 링커 서열, 어닐링될 수 있는 링커 서열 쌍, 또는 DNA 절편의 부위에 측면으로 부착된 어닐링될 수 있는 링커 서열/프라이머 결합 절편 쌍을 포함하지 않는 다른 벡터로부터 또한 제조될 수 있다. 상기 벡터로부터 집합 벡터를 제조하기 위하여, 벡터는 DNA 절편의 삽입에 적당한 부위, 예를 들어, 다수의 클로닝 부위에서 벡터를 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제를 사용하여 분해될 수 있으며, 이로써 벡터를 선형으로 만들어서 상기 벡터는 DNA 절편을 수용할 수 있다. 삽입되어야 하는 DNA 절편은, 예를 들어, 클로닝, 화학적 합성, 또는 PCR 증폭과 같이, 업계에서 알려진 표준 과정에 의해 수득될 수 있다. DNA 절편은 어닐링될 수 있는 링커 서열, 어닐링될 수 있는 링커 서열 쌍 또는 어닐링될 수 있는 링커 서열/프라이머 결합 절편 쌍의 측면에 배치된 DNA 절편을 포함한다. 따라서, 몇몇 실시예에서, DNA 절편은, 5'에서 3' 방향으로, 어닐링될 수 있는 링커 서열 LA 또는 프라이머 결합 절편 PA, DNA 절편 D 및 어닐링될 수 있는 링커 서열 LB 또는 프라이머 결합 절편 PB (예를 들어, 5'-LA-D-LB-3' 또는 5'-PA-D-LB-3' 또는 5'-LA-D-PB-3')를 포함한다. 몇몇 실시예에서, DNA 절편은, 5'에서 3'의 방향으로, DNA 절편 D 및 어닐링될 수 있는 링커 서열 LB 또는 프라이머 결합 절편 PB (예를 들어, 5'-D-LB-3' 또는 5'-D-PB-3')를 포함한다. 몇몇 실시예에서, DNA 절편은, 5'에서 3'의 방향으로, 어닐링될 수 있는 링커 서열 LA 또는 프라이머 결합 절편 PA 및 DNA 절편 D (예를 들어, 5'-LA-D-3' 또는 5'-PA-D-3')를 포함한다. DNA 절편은 어닐링될 수 있는 링커 서열, 어닐링될 수 있는 링커 서열 쌍 또는 어닐링될 수 있는 링커 서열/프라이머 결합 절편 쌍을 측면에 배치하는 한쌍의 제한 부위를 포함하며, 그 부위에서 제한 엔도뉴클레아제에 의한 절단은, DNA 절편이 삽입되어야 하는 벡터를 선형으로 만들어서 형성된 말단과 호환할 수 있는 말단을 형성한다. 선택적으로, DNA 절편이 형성될 수 있어서, 상기 DNA 절편은 상기 호환가능한 말단을 포함하며, 호환가능한 말단을 형성하기 위하여 제한 엔도뉴클레아제를 사용하여 추가적인 분해를 요구하지 않는다. 집합 벡터를 형성하기 위하여 선형으로 만들어진 벡터를 사용하여 DNA 절편의 연결함에 있어서, 호환가능한 말단을 형성하는데 사용된 제한 부위는 집합 벡터의 제한 부위 RA 및 RB로서 역할을 하기 위하여 보존될 수 있다. 선택적으로, 상기 연결은 원래의 제한 부위를 제거할 수 있지만, 추가적인 제한 부위가 집합 벡터의 제한 부위 RA 및 RB로서 역할을 할 수 있는 선형으로 만들어진 벡터에 존재할 수 있다.

[0159] 엔트리 벡터(entry vector) (예를 들어, pRYSE 벡터) 또는 다른 벡터 (예를 들어, pMULE 벡터)로부터 집합 벡터를 형성하는 예시적인 방법은 하기 실시예 6에 제시되었다.

[0160] **7 어닐링될 수 있는 링커 서열**

[0161] 다른 측면에서, 본 발명은 엔트리 벡터(entry vector)와 집합 벡터 내에 위치한 DNA 절편을 측면에 배치하는 어닐링될 수 있는 링커 서열을 제공한다. 어닐링될 수 있는 링커 서열은 집합 반응(assembly reaction)에서 인접한 폴리뉴클레오타이드 성분들 사이에 겹치는 서열을 제공하여, 집합체를 위한 폴리뉴클레오타이드 성분을 집합된(assembled) 폴리뉴클레오타이드로 준비시킨다. 따라서, 바람직한 실시예에서, 엔트리 벡터 및 집합 벡터의 어닐링될 수 있는 링커 서열 LA 및 LB는, 집합 반응(assembly reaction) 동안에 상보적인 어닐링될 수 있는 링커 서열에 효율적이고 정확한 프라이밍을 제공하도록 최적화된다.

[0162] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열의 길이는 상보적인 어닐링될 수 있는 링커 서열과 알맞는 특이성을 제공할 수 있을 정도로 길지만, 집합 반응(assembly reaction)의 어닐링 온도에서 상보적인 어닐링될 수 있는 링커 서열에 용이하게 어닐링 할 수 있을 정도로 짧다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열의 길이는 숙주 세포 매개 상동 재조합(homologous recombination)이 상보적 서열(complement) 어닐링될 수 있는 링커 서열을 갖게 할 정도로 충분히 길다.

[0163] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 약 5개, 10개, 15개, 20개, 25개, 30개, 35개, 40개, 45개,

50개, 55개, 60개, 65개, 70개, 75개, 또는 80개의 뉴클레오타이드 길이다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 적어도 10개, 12개, 14개, 16개, 18개, 20개, 22개, 24개, 26개, 28개, 또는 30개의 뉴클레오타이드 길이다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 30개, 40개, 50개, 60개, 70개, 80개, 90개, 100개, 500개, 1000개, 5000개, 또는 10,000개의 뉴클레오타이드 길이를 초과한다. 몇몇 실시예에서, 어닐링될 수 있는 링커(annealable linker)는 적어도 18개의 뉴클레오타이드 길이이며 3으로 나뉘어질 수 있는 갯수의 뉴클레오타이드 길이라서, DNA 절편을 코딩하기 위하여 연결될 때, 링커의 연속 판독(read-through) 전사를 촉진할 수 있다. 특정한 실시예에서, 어닐링될 수 있는 링커(annealable linker)는 18개, 21개, 24개, 27개, 30개, 33개, 36개, 39개, 42개, 45개, 48개, 51개, 54개, 57개 또는 60개 뉴클레오타이드 길이다.

[0164] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 다소 높은 융점 (Tm), 예를 들어, 어닐링될 수 있는 링커(annealable linker) 두 가닥의(duplex) 서열의 1/2이 단일 가닥으로 분리될 온도를 갖는다. 어닐링될 수 있는 링커의 Tm은 최근접 이웃 알고리즘(nearest neighbor algorithm)을 이용하여 SantaLucia, PNAS, 95:-1460-1465 (1998)에 따라 계산될 수 있다. 다소 높은 Tm은 집합 반응(assembly reaction) 동안 더욱 특이한 프라이밍을 제공할 수 있다. 다소 높은 Tm은 PCR의 어닐링 및 신장 단계의 조합을 허용하거나, PCR의 어닐링 및 신장 단계 사이의 온도를 조절하는데 필요한 양의 시간을 감소시켜서, 본 발명의 집합 방법을 사용하는데 있어서 효율성을 더 높일 수 있다. 따라서, 몇몇 실시예에서, 어닐링될 수 있는 링커(annealable linker) 두 가닥의(duplex) 서열은 약 60℃ 내지 80℃의 Tm를 갖는다. 몇몇 실시예에서, 어닐링될 수 있는 링커(annealable linker) 두 가닥의(duplex) 서열은 약 65℃ 내지 75℃의 Tm를 갖는다. 몇몇 실시예에서, 어닐링될 수 있는 링커(annealable linker) 두 가닥의(duplex) 서열은 50℃, 55℃, 60℃, 65℃, 70℃, 75℃, 80℃, 85℃, 또는 90℃를 초과하는 Tm를 갖는다.

[0165] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열들은, DNA 수준에서, 또는 RNA 수준에서 또는 DNA와 RNA 수준 모두에서, 본원에 제시된 방법의 조건 하에서, 분자내 (예를 들어, 동일한 분자 내) 상호작용을 통하여 형성된 인지가능한 2차 구조 (예를 들어, 헤어핀, 셀프-다이머(self-dimer))를 형성하지 않는다. DNA에서 2차 구조의 존재는 집합 반응(assembly reaction)의 집합된(assembled) 폴리뉴클레오타이드를 적게 생산하거나 전혀 생산하지 않을 수 있다. RNA에서 2차 구조의 존재는 감소된 번역 효율성을 유도할 수 있으며, 이것은, 프로모터와 단백질 코딩서열을 포함하는 폴리뉴클레오타이드 성분들을 집합된(assembled) 폴리뉴클레오타이드에 결합시키기 위하여 어닐링될 수 있는 링커 서열이 사용될 때 특히 중요하며, 상기 어닐링될 수 있는 링커 서열은 프로모터와 단백질 코딩서열 사이에 위치한다. 따라서, 본 발명의 집합 방법에 유용한 어닐링될 수 있는 링커 서열들은 2차 구조의 RNA 및/또는 DNA 구조를 형성하지 않도록 고안된다. 2차 구조의 RNA 또는 DNA 구조를 형성하기 위한 어닐링될 수 있는 링커 서열의 능력은, 예를 들어, IDT Oligo Analyzer (Integrated DNA Technologies, Coralville, IA), mFold (Zuker 2003 *Nucleic acids Res.* 31 (13), 3406-15), 또는 RNAfold (Hofacker & Stadler (2006) *Bioinformatics* 22 (10): 1172-6)와 같은 소프트웨어 도구들을 사용하여 측정될 수 있다. 일반적으로, 이러한 도구들은 선형 상태에서 접힌 상태로의 서열로 전이하는데 필요한 깁스 자유 에너지 (ΔG)를 계산한다. 더 큰 ΔG 일수록, 서열이 2차 구조를 덜 형성할 것이다. 따라서, 몇몇 실시예에서, 어닐링될 수 있는 링커 서열들은 선형 상태에서접혀진 상태로의 전이를 위하여 더 큰 ΔG 값을 갖도록 고안된다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열들은 *Saccharomyces cerevisiae* 게놈에서 많이 발현된 유전자의 코딩 서열의 즉각적인 업스트림에 존재하는 n-염기의 선형 상태에서 접혀진 상태로의 전이에 대한 ΔG 값과 동일하거나 그보다 큰, 선형 상태에서 접혀진 상태로의 전이에 대한 ΔG 값을 갖도록 고안되며, 상기 n은 어닐링될 수 있는 링커 서열에서 수많은 염기에 대응되는 정수를 나타낸다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열을 36개의 염기 길이이며, 선형 상태에서 접혀진 상태로의 전이에 대하여 -1 이상의 ΔG 값을 갖는다.

[0166] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열들은 의도하지 않은 분자간 상호작용 (예를 들어, 서로 다른 분자들 사이)를 피하기 위하여 또한 고안된다. 따라서, 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은, 본원에서 제공된 방법에 의해 폴리뉴클레오타이드 집합체에 요구된, 어닐링될 수 있는 링커 서열 (예를 들어, 벡터 골격 서열)을 포함하는 집합 벡터 내 임의의 기타 서열 및/또는, 상보적인 어닐링될 수 있는 링커 서열들로부터 떨어진 집합체 조성물의 다른 집합 벡터 내의 임의의 기타 서열과 실질적으로 어닐링하지 않는다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 본원에서 제공된 집합체 조성물의 집합 벡터 내 기타 어닐링될 수 있는 링커 서열들과 실질적으로 어닐링하지 않는다.

[0167] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 높은 G-C 함량, 예를 들어, 어닐링될 수 있는 링커 서열 내 총 염기 개수의 백분율로서 어닐링될 수 있는 링커(annealable linker) 내 많은 구아닌과 시토신 뉴클레오타이드를 갖는다. 높은 G-C 함량은 일반적으로 높은 Tm을 제공하기 때문에, 높은 G-C 함량을 갖는 어닐링될 수 있는 링커

서열들은 일반적으로 본 발명의 방법에 유용하며, 그 결과, 이것은 집합 반응(assembly reaction) 동안 더욱 특이한 프라이밍을 제공할 수 있으며, SOE/PCR의 어닐링 및 신장 단계의 조합을 허용함으로써 시간과 공정의 절감을 제공할 수 있다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열의 G-C 함량은 약 20 내지 80%이다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열의 G-C 함량은 약 40 내지 60%이다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열의 G-C 함량은 약 40, 45, 50, 55, 60, 또는 70%이다. 특정한 실시예에서, 어닐링될 수 있는 링커 서열은 70%를 초과하는 G-C 함량을 갖는다. 높은 G-C 함량을 갖는 어닐링될 수 있는 링커 서열들의 예시적인 실시예는, 인지가능한 2차 DNA 구조를 형성하지 않으며 70℃ 이상의 Tm를 갖거나, 서열번호 1 내지 8이다.

[0168] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 높은 A-T 함량, 예를 들어, 어닐링될 수 있는 링커 서열 내 총 염기 개수의 백분율로서 어닐링될 수 있는 링커 서열 내 많은 아데닌 및 티민 뉴클레오타이드를 갖는다. 높은 A-T 함량은 실질적인 2차 구조를 형성하는 어닐링될 수 있는 링커 서열의 감소된 성향을 제공할 수 있으며, 이는, 프로모터와 단백질 코딩서열을 포함하는 폴리뉴클레오타이드를 집합된(assembled) 폴리뉴클레오타이드에 결합시키기 위하여, 프로모터와 단백질 코딩 사이에 위치한, 어닐링될 수 있는 링커 서열이 사용될 때 특히 중요할 수 있다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열의 A-T 함량은 약 20 내지 80%이다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열의 A-T 함량은 약 40 내지 60%이다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열의 A-T 함량은 약 30, 35, 40, 45, 50, 55, 또는 60%이다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 30%를 초과하는 A-T 함량을 갖는다. 몇몇 실시예에서, 3'-어닐링될 수 있는 링커 서열의 최대 26개의 염기의 서열은 공통 서열 5'-ANNNNNNNNANNNAANTANNTTNANA-3'을 가지며, 상기 A는 아데닌을 나타내며, N은 임의의 뉴클레오타이드를 나타내고 T는 티민을 나타낸다. 이러한 공통 서열은 *Saccharomyces cerevisiae*의 게놈에서 높게 발현된 유전자의 시작 코돈의 업스트림에 존재하는 26개의 염기에서 종종 발견된다. 이러한 공통 서열을 포함하는 어닐링될 수 있는 링커 서열의 예시적인 실시예는, 다소 높은 A-T 함량을 가지며, 인지가능한 2차 구조의 RNA 또는 DNA 구조를 형성하지 않으며, 65℃ 이상의 Tm를 가지며, 서열번호 9 내지 23이다.

[0169] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 하나 이상의 제한 부위를 포함한다. 제한 부위의 어닐링될 수 있는 링커 서열로의 통합은, 엔트리 벡터(entry vector) 또는 집합 벡터 내 제한 부위 RA 및 RB를 유지하는 반면, 엔트리 벡터 또는 집합 벡터로부터 DNA 절편을 절단하게 한다. 어닐링될 수 있는 링커 서열 내 제한 부위은 DNA 절편(segment)의 기타 엔트리 벡터 또는 집합 벡터로의 방향성 서브클로닝(directional subcloning)을 또한 촉진시킨다. 이러한 특성은, 예를 들어, 하기에 제시된 서로 다른 어닐링될 수 있는 링커 서열 쌍을 포함하는 집합 벡터의 라이브러리를 형성하기 위하여, 동일한 DNA 절편을 포함하지만 서로 다른 어닐링될 수 있는 링커 서열 쌍 또는 프라이머 결합 절편/어닐링될 수 있는 링커 서열 상을 갖는 집합 벡터의 효율적인 구조를 촉진시킨다. 이러한 특성은 또한, 재-증폭에 대한 필요성과 DNA 절편을 포함하는 추가적 집합 벡터를 형성하기 위한 서열 DNA 절편에 대한 필요성을 제거할 수 있다. 따라서, 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 특유한 제한 부위를 포함한다. 몇몇 실시예에서, 제한 부위는 7개의 염기쌍 제한 부위이며, 예를 들어, 7개의 염기쌍 뉴클레오타이드 서열을 인지하는 제한 엔도뉴클레아제에 의해 절단될 수 있다. 몇몇 실시예에서, 제한 부위는 8개의 염기쌍 제한 부위를 포함한다. 특정한 실시예에서, 어닐링될 수 있는 링커 서열 내 제한 부위는 MreI, FseI, SbfI, AsiSI, NotI, AscI, 또는 BbvCI에 의해 인지 및 절단될 수 있다.

[0170] 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은, 일단 코딩 DNA 절편에 연결된 링커를 연속 판독(read-through) 전사할 수 있게 하는 서열을 포함한다. 몇몇 실시예에서, 어닐링될 수 있는 링커 서열은 5'에서 3'의 방향으로, 그리고 3'에서 5'의 방향으로 연속 판독(read-through) 전사를 하게 한다. 이러한 실시예에서, 어닐링될 수 있는 링커 서열의 길이는, 바람직하게, 3에 의해 나누어질 수 있는 개수의 뉴클레오타이드이다.

[0171] 특정한 실시예에서, 어닐링될 수 있는 링커 서열은 *Escherichia coli* (*E. coli*) 또는 *Saccharomyces cerevisiae* (*S. cerevisiae*)에서 드물게 사용되는 코돈을 포함하지 않는다. *E. coli* 또는 *S. cerevisiae*의 이형(heterologous) 유전자의 효율적인 발현은, 드물게 사용된 코돈의 존재와, 드물게 사용된 코돈이 더욱 자주 사용된 코돈으로 대체될 때 종종 상승하는 이형(heterologous) 단백질의 발현 정도에 의해 역으로 영향을 받을 수 있다. 예를 들어, Williams *et al., Nucleic acids Res.* 16: 10453-10467, 1988 및 Hoog *et al., Gene* 43: 13-21, 1986를 참조한다. 따라서, 연속 판독(read-through) 서열을 포함하는 어닐링될 수 있는 링커 서열은, 바람직하게 *E. coli* 또는 *S. cerevisiae*에서 드물게 사용된 코돈을 포함하지 않아서, 어닐링될 수 있는 링커 서열을 포함하는 집합된(assembled) 폴리뉴클레오타이드에 의해 코딩된 단백질의 효율적인 발현을 가능하게 한다.

[0172] 몇몇 실시예에서, 일련의 어닐링될 수 있는 링커 서열은 의도된 호스트 생물에서 발견되지 않는 특이한 서열이다. 몇몇 실시예에서, 일련의 어닐링될 수 있는 링커 서열은 *E. coli*에서 발견되지 않는 특이한 서열이다. 다른

실시예에서, 일련의 어닐링될 수 있는 링커 서열은 *S. cerevisiase*에서 발견되지 않는 특이한 서열이다.

[0173] 몇몇 실시예에서, 적당한 어닐링될 수 있는 링커 서열은 테스트용 집합된(assembled) 폴리뉴클레오타이드에서 식별된다. 테스트용 집합된(assembled) 폴리뉴클레오타이드는 테스트되어야 하는 어닐링될 수 있는 링커 서열과 어닐링될 수 있는 링커 서열의 테스트를 허용하는 추가적 성분들을 포함한다. 예를 들어, 집합된(assembled) 폴리뉴클레오타이드 내 프로모터의 조절 하에 있도록, 어닐링될 수 있는 링커(annealable linker)가 프로모터 서열을 포함하는 첫번째 폴리뉴클레오타이드 성분;과 단백질 코딩서열을 포함하는 두번째 폴리뉴클레오타이드 성분을 결합시키는데 적당한지 테스트하기 위하여, 테스트용 집합된(assembled) 폴리뉴클레오타이드는, 5'에서 3'의 방향으로, 프라이머 결합 절편 또는 어닐링될 수 있는 링커 서열을 포함하는 첫번째 폴리뉴클레오타이드 성분; 프로모터와 테스트되어야 하는 어닐링될 수 있는 링커 서열을 포함하는 DNA 절편; 및 5'에서 3'의 방향으로, 테스트 되어야 하는 어닐링될 수 있는 링커 서열, 리포터 유전자(reporter gene)를 코딩하는 DNA 절편 (예를 들어, 녹색 형광 단백질 (GFP)) 및 프라이머 결합 절편 또는 어닐링될 수 있는 링커 서열을 포함하는 두번째 폴리뉴클레오타이드 성분;으로부터 집합될 수 있다. 테스트용 집합된(assembled) 폴리뉴클레오타이드는 리포터 유전자(reporter gene)의 발현의 효율성을 위하여 생체 내 또는 생체 밖에서 테스트될 수 있다. 유사한 테스트용 집합된(assembled) 폴리뉴클레오타이드들은, 증폭자(enhancer), 종결자(terminator), 폴리-A 꼬리, 핵 위치 신호(nuclear localization signal), mRNA 안정 신호(mRNA stabilization signal), 선별표지(selectable marker), 에피토프 태그(epitope tag) 코딩 서열, 분해 신호(degradation signal), 기타 등등과 같은 기타 성분들을 포함하는 DNA 절편을 포함하는 폴리뉴클레오타이드 성분들을 결합시키는데 어닐링될 수 있는 링커 서열의 적절성을 테스트하기 위하여 집합될 수 있다. 테스트용 집합된(assembled) 폴리뉴클레오타이드는 테스트를 가능하게 하는 추가적인 폴리뉴클레오타이드 성분들, 예를 들어, 생체 내 테스트를 위하여 숙주 세포로의 테스트용 집합된(assembled) 폴리뉴클레오타이드의 도입과 양성 형질변환체의 선택을 가능하게 하는 게놈 타겟팅(targeting) 서열과 선별표지(selectable marker)를 포함할 수 있다.

[0174] 표 1은 Tm, 제한 부위 및 서열번호(SEQ ID NO) 1 내지 23에 상응되는 예시적인 어닐링될 수 있는 링커 서열의 연속 판독(read-through) 아미노산을 나타낸다.

표 1

[0175]

표 1 - 어닐링될 수 있는 링커 서열의 서열및 특성

어닐링할 수 있는 링커 서열	서열 이름	길이 (염기)	% G-C	% A-T	융점(Tm)	제한 효소 부위	연속판독(read-through) 아미노산	
							Fwd	Rev
SEQ ID NO: 1	RYSE 1	24	79.2	20.8	72.4			
SEQ ID NO: 2	RYSE 2	24	75.0	25.0	71.4	MreI		
SEQ ID NO: 3	RYSE 3	24	75.0	25.0	73.7	FseI		TAGQARGD
SEQ ID NO: 4	RYSE 4	24	70.8	29.2	71.5	SbfI	NLQAASAD	IGARGLQV
SEQ ID NO: 5	RYSE 5	24	70.8	29.2	71.2	AsiSI	NAIADAAD	IGGVGDRV
SEQ ID NO: 6	RYSE 6	24	70.8	29.2	70.9	NotI	KAAAGEGD	ISLASGRL
SEQ ID NO: 7	RYSE 7	24	70.8	29.2	71.5	AscI	KARHGRRD	
SEQ ID NO: 8	RYSE 8	24	75.0	25.0	70.7	BbvCI		
SEQ ID NO: 9	RYSE 9	36	50.0	50.0	67.4			
SEQ ID NO: 10	RYSE 10	36	52.8	47.2	67.7			
SEQ ID NO: 11	RYSE 11	36	58.3	41.7	69.2			
SEQ ID NO: 12	RYSE 12	36	50.0	50.0	67.4			
SEQ ID NO: 13	RYSE 13	36	58.3	41.7	69.4			
SEQ ID NO: 14	RYSE 14	36	52.8	47.2	67.4			
SEQ ID NO: 15	RYSE 15	36	52.8	47.2	67.8			
SEQ ID NO: 16	RYSE 16	36	52.8	47.2	67.8			
SEQ ID NO: 17	RYSE 17	36	52.8	47.2	68.4			
SEQ ID NO: 18	RYSE 18	36	50.0	50.0	67.8			
SEQ ID NO: 19	RYSE 19	36	52.8	47.2	68.1			

SEQ ID NO: 20	RYSE 20	36	55.6	44.4	68.3			
SEQ ID NO: 21	RYSE 21	36	55.6	44.4	67.9			
SEQ ID NO: 22	RYSE 22	36	52.8	47.2	67.4			
SEQ ID NO: 23	RYSE 23	36	55.6	44.4	68.8			

[0176] **8 라이브러리**

[0177] 다른 측면에서, 본 발명은 수많은 집합 벡터들을 포함하는 라이버러리를 제공한다. 라이브러리는 원핵세포 또는 진핵세포에서 기능적인 하나 이상의 집합된(assembled) 폴리뉴클레오타이드로 수많은 폴리뉴클레오타이드 성분들의 효율적인 집합을 촉진시키는 기능을 할 수 있어서, 따라서, 시간을 소모하는 제한 엔도뉴클레아제와 연결효소(ligase)에 기초한 클로닝 기술에 대한 필요성 없이, 특이한 생물, 예를 들면, 박테리아 또는 효모의 재조합 종(strain)의 형성을 촉진한다. 본원에서 제공된 집합 방법 및 조성물은, 발현 구성 내에서, 기능적 DNA 단위, 예를 들어, 프로모터, 증폭자(enhancer), 복제 원점(origin), 등등의 효율적인 대체 또는 도입을 촉진할 수 있어서, 호스트 생물 내 발현 구성의 복제, 및/또는 그로부터의 발현의 효율적인 최적화를 제공할 수 있다.

[0178] [00154] 라이브러리는 단일한 조성물 또는 컨테이너(container), 예를 들어, 본원에서 제공된 집합 방법을 실시하는데 적당한 조성물 또는 컨테이너(container) 내에서 집합된 수많은 집합 벡터들을 포함할 수 있다. 선택적으로, 라이브러리는 동일한 조성물 또는 컨테이너(container) 내에 집합되지 않은 수많은 집합 벡터들을 포함할 수 있다. 몇몇 실시예에서, 라이브러리는 적어도 3개, 적어도 6개, 적어도 10개, 적어도 20개, 적어도 50개, 또는 50개를 초과하는 집합 벡터들을 포함하며, 각각의 벡터는 DNA 절편을 포함한다.

[0179] 몇몇 실시예에서, 라이브러리는 수많은 집합 벡터들을 포함하며, 상기 각각의 집합 벡터들은, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 두번째 제한 부위 RB를 포함한다. 몇몇 실시예에서, 라이브러리는 수많은 집합 벡터들을 포함하며, 각각의 집합 벡터는, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA, 프라이머 결합 절편 PA 또는 첫번째 어닐링될 수 있는 링커 서열 LA, DNA 절편 D 및 두번째 제한 부위 RB를 포함한다. 몇몇 실시예에서, 라이브러리는 수많은 집합 벡터들을 포함하며, 상기 각각의 집합 벡터들은, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA, 첫번째 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 또는 프라이머 결합 절편 PB 및 두번째 제한 부위 RB를 포함한다. 몇몇 실시예에서, 라이브러리의 각각의 집합 벡터 내의 어닐링될 수 있는 링커 서열 쌍 또는 어닐링될 수 있는 링커 서열/1차 결합 절편 쌍은 동일한 서열을 포함하지 않는다. 몇몇 실시예에서, 각각의 집합 벡터 내 어닐링될 수 있는 링커 서열 LA 및/또는 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 각각의 집합 벡터 내 프라이머 결합 절편 PA 또는 PB의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다.

[0180] 몇몇 실시예에서, 라이브러리는 하기의 각각의 벡터들 중 적어도 하나를 포함한다:

[0181] (a) 5'에서 3'의 방향으로, 제한 부위 RA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는 원형 폴리뉴클레오타이드로 구성되는 벡터;

[0182] (b) 5'에서 3' 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는, 원형 폴리뉴클레오타이드로 구성되는 벡터; 및

[0183] (c) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D 및 제한 부위 RB_0를 포함하는 원형 폴리뉴클레오타이드로 구성된 벡터.

[0184] 몇몇 실시예에서, 라이브러리는 하기의 각각의 벡터들 중 적어도 하나를 포함한다:

[0185] (a) 5'에서 3'의 방향으로, 제한 부위 RA, 프라이머 결합 절편 PA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는 원형 폴리뉴클레오타이드로 구성된 벡터;

[0186] (b) 5'에서 3'의 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함하는 원형 폴리뉴클레오타이드로 구성된 벡터; 및

[0187] (c) 5'에서 3' 방향으로, 제한 부위 RA, 어닐링될 수 있는 링커 서열 LA, DNA 절편 D, 프라이머 결합 절편 PB 및 제한 부위 RB_0를 포함하는 원형 폴리뉴클레오타이드로 구성된 벡터.

[0188] 몇몇 실시예에서, 프라이머 결합 절편 PA의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 몇

몇 실시예에서, 프라이머 결합 절편 PB의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 프라이머 결합 절편 PA와 프라이머 결합 절편 PB의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다.

[0189] 몇몇 실시예에서, 라이브러리 내 임의의 어닐링될 수 있는 링커 서열 LA와 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 라이브러리 내 어닐링될 수 있는 링커 서열 LA 중 적어도 하나와 어닐링될 수 있는 링커 서열 LB 중 적어도 하나의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 라이브러리 내 어닐링될 수 있는 링커 서열 LA와 어닐링될 수 있는 링커 서열 LB의 각각의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다.

[0190] 몇몇 실시예에서, DNA 절편 D는 선별표지(selectable marker), 프로모터, 게놈 타겟팅(targeting) 서열, 에피토프 태그(epitope tag)를 코딩하는 핵산 서열, 관심있는 유전자를 코딩하는 핵산 서열, 종결 코돈을 코딩하는 핵산 서열 및 lacZ로 구성된 군으로부터 선택된 핵산 서열을 포함한다.

[0191] 몇몇 실시예에서, 라이브러리는 하기의 각각의 핵산 분자 중 적어도 하나를 포함한다:

[0192] (a) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 그룹 D_0, 어닐링될 수 있는 링커 서열 LB_0 및 두번째 제한 부위 RB_0로부터 선택된 임의의 DNA 절편을 포함하는, 첫번째 핵산 분자;

[0193] (b) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, 그룹 D_n로부터 선택된 임의의 DNA 절편, 두번째 어닐링될 수 있는 링커 서열 LB_n 및 두번째 제한 부위 RB_n을 포함하는, 중간 핵산 분자(n은 1부터 중간 핵산 분자의 개수까지의 정수를 나타냄); 및

[0194] (c) 각각, 원형이며, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, 그룹 D_m으로부터 선택된 임의의 DNA 절편, 두번째 제한 부위 RB_m을 포함하는, 마지막 핵산 분자(m은 중간 핵산 분자의 개수보다 큰 정수를 나타냄);

[0195] 여기에서, 제한 부위 RA_0부터 RB_m까지의 절단과 결과물 선형 핵산 분자의 변성에서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은 어닐링될 수 있는 링커 서열 LA_p의 상보적 서열(complement)에 하이브리드할 수 있으며, 상기 p는 1부터 m까지의 정수를 나타내며, 상기 각각의 그룹 D_0,... D_n,...및 D_m은 하나 이상의 DNA 절편(segment)으로 구성된다. 몇몇 실시예에서, 첫번째 핵산 분자는 그룹 D_0으로부터 선택된 DNA 절편에 대해 5'에 위치한 프라이머 결합 절편 PA를 더 포함한다. 몇몇 실시예에서, 마지막 핵산 분자는 그룹 D_m으로부터 선택된 DNA 절편에 대해 3'에 위치한 프라이머 결합 절편 PB를 더 포함한다.

[0196] 몇몇 실시예에서, 제한 부위 RA_0부터 RB_m까지의 절단과 결과물 선형 핵산 분자의 변성에서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은, 성분 조성물(component composition)에서, 다른 어닐링될 수 있는 링커 서열 대비 어닐링될 수 있는 링커 서열 LA_p 또는 이들의 상보적 서열(complement)의 상보적 서열(complement)에 선택적으로 하이브리드될 수 있다. 몇몇 실시예에서, 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은 어닐링될 수 있는 링커 서열 LA_p과 서열에서 동일하다.

[0197] 특정 실시예에서, 제한 부위 RA_0부터 RB_m까지는 동일한 제한 엔도뉴클레아제에 의해 절단될 수 있어서 집합 벡터들로부터의 폴리뉴클레오타이드 성분들의 절단을 촉진할 수 있다. 몇몇 실시예에서, 제한 부위 RA_0부터 RB_m까지는 SapI 및 LguI 제한 엔도뉴클레아제들에 의해 절단될 수 있다.

[0198] 몇몇 실시예에서, 프라이머 결합 절편 PA의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 프라이머 결합 절편 PB의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 프라이머 결합 절편 PA와 프라이머 결합 절편 PB의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 바람직한 실시예에서, 프라이머 결합 절편 PA와 프라이머 결합 절편 PB의 핵산 서열은 동일하지 않다.

[0199] 몇몇 실시예에서, 라이브러리 내 어닐링될 수 있는 링커 서열 LA와 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 라이브러리 내 적어도 하나의 어닐

링될 수 있는 링커 서열 LA와 적어도 하나의 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 라이브러리 내 각각의 어닐링될 수 있는 링커 서열 LA와 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 조성물 내 각각의 어닐링될 수 있는 링커 서열 LA의 핵산 서열은 다른 핵산 서열과 동일하지 않다. 몇몇 실시예에서, 조성물 내 각각의 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 다른 핵산 서열과 동일하지 않다.

[0200] 특정 실시예에서, 라이브러리는 하기의 핵산 분자를 포함한다:

[0201] (a) 2개의 첫번째 핵산 분자로서, 상기 하나의 첫번째 핵산 분자는, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 프라이머 결합 절편 PA, DNA 절편 D_{01}, 어닐링될 수 있는 링커 서열 LB_0 및 두번째 제한 부위 RB_0를 포함하며, 상기 다른 첫번째 핵산 분자는, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_0, 프라이머 결합 절편 PA, DNA 절편 D_{02}, 어닐링될 수 있는 링커 서열 LB_0 및 두번째 제한 부위 RB_0를 포함하며, 상기 DNA 절편 D_{01}은 첫번째 게놈 타겟팅(targeting) 서열을 코딩하며, 상기 DNA 절편 D_{02}는 타겟 게놈 내 첫번째 게놈 타겟팅(targeting) 서열의 다운스트림에 위치한 두번째 게놈 타겟팅(targeting) 서열을 코딩하며, 상기 DNA 절편 D_{02}는 프라이머 결합 절편 PA와 어닐링될 수 있는 링커 서열 LB_0 대비 DNA 절편 D_{01}에 대해 반대쪽 방향에 위치하며;

[0202] (b) 적어도 하나의 중간 핵산 분자로서, 상기 중간 핵산 분자는 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, DNA 절편 D_n, 두번째 어닐링될 수 있는 링커 서열 LB_n 및 두번째 제한 부위 RB_n를 포함하며, 상기 n은 1부터 중간 핵산 분자의 개수까지의 정수를 나타내며; 및

[0203] (c) 2개의 마지막 핵산 분자로서, 상기 하나의 마지막 핵산 분자는, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, DNA 절편 D_{m1}, 프라이머 결합 절편 PB 및 두번째 제한 부위 RB_m을 포함하며, 상기 다른 마지막 핵산 분자는, 5'에서 3'의 방향으로, 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, DNA 절편 D_{m2}, 프라이머 결합 절편 PB 및 두번째 제한 부위 RB_m를 포함하며, 상기 m은 중간 핵산 분자의 개수보다 큰 정수를 나타내며, 상기 DNA 절편 D_{m1}은 선별표지(selectable marker)의 첫번째 절편을 코딩하며, 상기 DNA 절편 D_{m2}는 선별표지(selectable marker)의 두번째 절편을 코딩하며, 상기 DNA 절편 D_{m2}는 어닐링될 수 있는 링커 서열 LA_m과 프라이머 결합 절편 PB에 대비 DNA 절편 D_{m1}에 대해 반대쪽 방향에 위치하며, 상기 DNA 절편 D_{m1} 또는 DNA 절편 D_{m2}는 기능적 선별표지(selectable marker)를 형성하지 않지만, DNA 절편(segment) D_{m1}과 D_{m2}의 상동 재조합(homologous recombination)에서 기능적 선별표지(selectable marker)가 형성되며;

[0204] 상기 각각의 어닐링될 수 있는 링커 서열 $LB_{(p-1)}$은 어닐링될 수 있는 링커 서열 LA_p와 동일하며, 상기 p는 1부터 m까지의 정수를 나타낸다.

[0205] 몇몇 실시예에서, 라이브러리는 수많은 집합 벡터들을 포함하며, 상기 각각의 집합 벡터는 동일한 어닐링될 수 있는 링커 서열, 어닐링될 수 있는 링커 서열 쌍 또는 어닐링될 수 있는 링커 서열/1차 결합 절편 쌍을 포함하지만, 상기 벡터의 각각의 DNA 절편 D와 서열에서 다르다.

[0206] 다른 실시예에서, 라이브러리는 수많은 집합 벡터들을 포함하며, 상기 각각의 집합 벡터는 특유한 어닐링될 수 있는 링커 서열, 어닐링될 수 있는 링커 서열 쌍 또는 어닐링될 수 있는 링커 서열/프라이머 결합 절편 쌍의 측면에 배치된 동일한 DNA 절편 D를 포함한다. 그러한 라이브러리는 집합된(assembled) 폴리뉴클레오타이드로의 집합된 다른 DNA 절편(segment) 대비 특정한 위치 또는 배향으로의 DNA 절편 D의 빠른 집합을 촉진하는 기능을 할 수 있다.

[0207] 몇몇 실시예에서, 라이브러리의 구성원은 공동의 구조적 또는 기능적 특성을 갖는 DNA 절편(segment)을 포함한다. 예를 들어, 라이브러리는 동일한 기능적 DNA 단위를 포함하는 수많은 집합 벡터들을 포함할 수 있다. 예시적인 기능적 DNA 단위들은, 그러나 이에 제한되지 않는, 단백질-코딩 서열, 리포터 유전자, 형광 표지, 프로모터, 증폭자(enhancer), 종결자(terminator), 인트론, 엑손, 폴리-A 꼬리, 다수의 클로닝 부위, 핵 위치 신호(nuclear localization signal), 핵 배출 신호(nuclear export signal), mRNA 안정 신호(mRNA stabilization

signal), 선별표지(selectable marker), 통합 부위(integration loci), 에피토프 태그(epitope tag) 및 분해 신호(degradation signal)를 포함한다. 몇몇 실시예에서, 라이브러리는 수많은 집합 벡터들을 포함하며, 상기 각각의 집합 벡터는 동일한 프로모터를 포함한다. 집합 벡터는 업계에 알려진 임의의 원핵 또는 진핵 프로모터 서열을 포함할 수 있다. 예시적인 진핵 프로모터는, 그러나 이에 제한되지 않는, 메탈로티오네인(metallothionein) 프로모터, 항시성 아데노바이러스 메이저 레이트 프로모터(constitutive major late promotor), 덱사메타손-유도성 MMTV 프로모터, SV40 프로모터, MRP *pol III* 프로모터, 항시성 MPSV 프로모터, RSV 프로모터, 테트라사이클린-유도성 CMV 프로모터 (예를 들어, 인간 조기발현(immediate-Early) CMV 프로모터) 및 항시성 CMV 프로모터를 포함한다. 특정한 실시예에서, 집합 벡터는 효모 프로모터 서열을 포함한다. 예시적인 효모 프로모터는, 그러나 이에 제한되지 않는, PGAL3, PGAL7, PCTR3, PMET3, PPGK1, PTDH1, PTDH3, PFBA1, PTEF1, PENO1, PENO2, PCYC1, PTDH2, PCUP1, PGAL80, PGAL2, PBNA6, PTMA29, PSBP1, PPUP3, PACS2, PTPO1, PRPT1, PAAT2, PAHP1, PSSE1, PTEF2, PNPL3, PPET9, PTUB2, POLE1, PCPR1, PIPPP1 및 PSOD1를 포함한다.

[0208] 몇몇 실시예에서, 라이브러리는 수많은 집합 벡터들을 포함하며, 상기 각각의 집합 벡터는 동일한 종결자(terminator) 서열을 포함한다. 집합 벡터는 업계에 알려진 임의의 원핵 또는 진핵 종결자(terminator) 서열을 포함할 수 있다. 특정한 실시예에서, 집합 벡터는 효모 종결자(terminator) 서열을 포함한다. 예시적인 효모 종결자(terminator)는, 그러나 이에 제한되지 않는, TADH1, TENO1, TENO2, TCYC1, TNDT80, TTDH3, TTDH1 및 TPGK1를 포함한다.

[0209] 몇몇 실시예에서, 라이브러리는 수많은 집합 벡터들을 포함하며, 상기 각각의 집합 벡터는 동일한 선별표지(selectable marker)를 포함한다. 집합 벡터는 업계에 알려진 임의의 원핵 또는 진핵 선별표지(selectable marker)를 포함할 수 있다. 선별표지(selectable marker)의 예는, 그러나 이에 제한되지 않는, 항생제 내성 표지 (예를 들어, 카나마이신, 앰피실린, 클로람페니콜, 젠타마이신, 또는 트리메토프림에 대한 내성을 코딩하는 유전자) 및 대사성 표지(metabolic marker)(예를 들어, 아미노산 합성 유전자 또는 운반 RNA 유전자)를 포함한다.

[0210] **9 키트(Kit)**

[0211] 다른 측면에서, 본원은 폴리뉴클레오타이드의 집합체 키트를 제공하며, 상기 키트는 2개 이상의 하기의 것들을 포함한다: (a) 본원에서 제시된 하나 이상의 엔트리 벡터(entry vector); (b) 상기 하나 이상의 엔트리 벡터(entry vector)의 제한 부위들 RA 및 RB를 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제; (c) 상기 엔트리 벡터(entry vector)의 제한 부위들 RY 및 RZ를 절단할 수 있는 하나 이상의 제한 엔도뉴클레아제; 및 (d) 상기 하나 이상의 엔트리 벡터(entry vector)의 프라이머 결합 절편(segment)들 PA 및 PB에 어닐링할 수 있는 올리고뉴클레오타이드프라이머.

[0212] 몇몇 실시예에서, 상기 키트의 각각의 엔트리 벡터(entry vector)의 제한 부위 RA 및 RB는 SapI 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있으며, 상기 키트는 SapI 제한 엔도뉴클레아제를 포함한다. 몇몇 실시예에서, 상기 키트의 각각의 엔트리 벡터(entry vector)의 제한 부위 RY 및 RZ는 SchI (또는 MlyI) 제한 엔도뉴클레아제에 의해 인지 및 절단될 수 있으며, 상기 키트는 SchI (또한 MlyI) 제한 엔도뉴클레아제를 포함한다.

[0213] 몇몇 실시예에서, 키트 내 하나 이상의 엔트리 벡터(entry vector)의 프라이머 결합 절편 PA의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 키트 내 하나 이상의 엔트리 벡터(entry vector)의 프라이머 결합 절편 PB의 핵산 서열은 서열번호 24 및 25로 구성된 군으로부터 선택된다. 바람직한 실시예에서, 프라이머 결합 절편 PA와 프라이머 결합 절편 PB의 핵산 서열은 동일하지 않다.

[0214] 몇몇 실시예에서, 키트 내 하나 이상의 엔트리 벡터(entry vector)의 어닐링될 수 있는 링커 서열 LA의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 키트 내 하나 이상의 엔트리 벡터(entry vector)의 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다. 몇몇 실시예에서, 키트 내 모든 엔트리 벡터(entry vector)의 어닐링될 수 있는 링커 서열 LA 및 어닐링될 수 있는 링커 서열 LB의 핵산 서열은 서열번호 1 내지 23으로 구성된 군으로부터 선택된다.

[0215] 몇몇 실시예에서, 키트는 pRYSE 벡터 #1, 서열번호 221와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #2, 서열번호 207와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는

pRYSE 벡터 #3, 서열번호 208와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #4, 서열번호 209와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #5, 서열번호 210와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #6, 서열번호 211와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #7, 서열번호 212와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #8, 서열번호 213와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #9, 서열번호 214와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #10, 서열번호 215와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #11, 서열번호 216와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #12, 서열번호 217와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #13, 서열번호 218와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #14, 서열번호 219와 같이 본원에서 제공된 서열을 포함한다. 몇몇 실시예에서, 키트는 pRYSE 벡터 #15, 서열번호 220와 같이 본원에서 제공된 서열을 포함한다.

[0216] 몇몇 실시예에서, 키트는 본원에서 제시된 폴리뉴클레오타이드 집합 방법을 설명하는 용도에 대한 지침을 더 포함한다. 몇몇 실시예에서, 내열성의 DNA 폴리머라제 (예를 들어, Pfu DNA 폴리머라제) 및 데옥시리보뉴클레오사이드 트리포스페이트 (dNTP)와 같은 폴리뉴클레오타이드 폴리머라제 또한, 키트 내에 존재한다. 몇몇 실시예에서, 집합된(assembled) 폴리뉴클레오타이드로 집합되는 폴리뉴클레오타이드 성분을 각각 포함하는 2개 이상의 집합 벡터는 키트 내에 제공될 수 있다. 예를 들어, 집합 벡터는 교정 및/또는 키트의 올바른 실시를 증명하기 위한 양성 대조군으로서의 사용에 유용한 폴리뉴클레오타이드 성분을 포함하도록 제공될 수 있다. 기타 실시예는, 그러나 이에 제한되지 않는, 집합 벡터는 폴리뉴클레오타이드 성분으로서 단백질-코딩 서열, 리포터 유전자(reporter gene), 형광 표지(marker) 코딩 서열, 프로모터, 증폭자(enhancer), 종결자(terminator), 인트론, 엑손, 폴리-A 꼬리, 다수의 클로닝 부위, 핵 위치 신호(nuclear localization signal), mRNA 안정 신호(mRNA stabilization signal), 선별표지(selectable marker), 통합 부위(integration loci), 에피토프 태그(epitope tag) 코딩 서열 및 분해 신호(degradation signal)를 포함한다.

도면의 간단한 설명

[0217] 도 1a는 본 발명의 집합 벡터의 제조에 유용한 엔트리 벡터(entry vector)의 도식을 제공한다. 상기 벡터는 제한 부위 RA_0, 프라이머 결합 절편 PA 또는 어닐링될 수 있는 링커 서열 LA, 제한 부위 RY, DNA 절편 D, 제한 부위 RZ, 프라이머 결합 절편 PB 또는 어닐링될 수 있는 링커 서열 LB 및 제한 부위 RB를 포함한다.

도 1b는 집합 벡터를 형성하기 위하여 DNA 절편의 수용을 위한 엔트리 벡터(entry vector)의 예시적인 제조 방법을 제공한다. 실시예에서, RY=RZ=SchI이다. SchI를 사용한 분해, 즉 평활 말단(blunt end)을 형성할 수 있는 IIS형 제한 엔도뉴클레아제는 DNA 절편 D로 융합되도록, 열려져 있는 링커 부위를 갖는 벡터의 분리를 허용한다. 엔트리 벡터(entry vector)로의 D의 평활 말단(blunt-end) 연결은, 예를 들어, T4 DNA 연결효소(ligase)를 사용하는 종래의 방법에 의해 실시될 수 있다.

도 2는 수많은 집합 벡터 (첫번째, 중간 및 마지막)를 포함하는 집합체 조성물의 도식을 나타내며, 각각의 집합 벡터는 관심있는 DNA 절편 (D_0, D_n, D_m)을 포함한다. 첫번째 핵산 분자는 첫번째 제한 부위 RA_0, 프라이머 결합 절편 PA, DNA 절편 D_0, 어닐링될 수 있는 링커 서열 LB_0 및 두번째 제한 부위 RB_0를 포함한다. 하나 이상의 중간 핵산 분자는 첫번째 제한 부위 RA_n, 첫번째 어닐링될 수 있는 링커 서열 LA_n, DNA 절편 D_n, 두번째 어닐링될 수 있는 링커 서열 LB_n 및 두번째 제한 부위 RB_n을 포함하며, 상기 n은 1부터 중간 핵산 분자의 개수까지의 정수를 나타내며; 마지막 핵산 분자는 첫번째 제한 부위 RA_m, 어닐링될 수 있는 링커 서열 LA_m, DNA 절편 D_m, 프라이머 결합 절편 PB, 두번째 제한 부위 RB_m을 포함하며, 상기 m은 중간 핵산 분자의 개수보다 큰 정수를 나타낸다.

도 3은 집합하는 예시적인 방법, 예를 들어, 4개의 폴리뉴클레오타이드 성분들로부터 집합된(assembled) 폴리뉴클레오타이드를 "스티칭(stitching)"하는 방법을 나타낸다. 집합될 DNA 절편을 포함하는 집합 벡터들은 하나의 싱글 튜브에 넣어진 후에, 집합 벡터 골격으로부터 폴리뉴클레오타이드 절편 성분을 분리하기 위하여 SapI를 사용하여 분해된다. SapI의 열 불활성화 후에, 폴리뉴클레오타이드 절편 성분들은 변성 조건을 겪게 된 후, 상보적인 어닐링될 수 있는 링커(annealable linker) 쌍의 하이브리드에 충분한 어닐링 조건을 겪게 된다. DNA 폴리머라제 및 dNTP의 존재 하에서 프라이머 신장을 한 후, PA 및 PB에 상보적인 프라이머가 첨가된 후, 종래의 PCR 증폭이 실시된다. 집합 반응의 결과에 따라, 5'에서 3'의 방향으로 집합된 폴리뉴클레오타이드 성분들 D_0, D_1,

D_2 및 D_3를 포함하는 집합된(assembled) 폴리뉴클레오타이드가 형성된다.

도 4는 pRYSE 벡터의 맵(map)을 보여준다.

도 5는 SapI 제한 엔도뉴클레아제 (컬럼 정제 또는 열 불활성화)를 제거하는 서로 다른 방법, 기타 모든 절편들의 동일한 몰 농도를 갖는 가장 작은 절편의 서로 다른 집합 벡터 DNA 농도 (5 ng (낮은 DNA 농도) 또는 50 ng (높은 DNA 농도)) 및 PCR 증폭을 위한 서로 다른 어닐링 온도 (54℃ 및 72℃)를 사용하여, 2개 내지 4개의 폴리뉴클레오타이드 성분들 (표 7의 집합체 1 내지 6)을 집합하여 얻은 집합된(assembled) 폴리뉴클레오타이드들을 보여준다.

도 6은 서로 다른 DNA 폴리머라제들 (Phusion (New England Biolabs, Ipswich, MA) 및 PfuUltraII (Stratagene/Agilent, La Jolla, CA))를 사용하여 6개 또는 9개 폴리뉴클레오타이드 성분들 (표 7의 집합체 7 및, 13 내지 16)을 집합하여 얻은 집합된(assembled) 폴리뉴클레오타이드들을 보여준다.

도 7은 pMULE 벡터의 맵(map)을 보여준다. pMULE 엔트리 벡터(entry vector)는, 프라이머 결합 절편(segment) 또는 어닐링될 수 있는 링커 서열이 없다는 점에서 pRYSE 엔트리 벡터(entry vector)와 다르다.

도 8은 집합된(assembled) 폴리뉴클레오타이드들을 숙주 세포 매개 상동 재조합(homologous recombination)에 의해 결합된(combined) 폴리뉴클레오타이드에 결합시키고, 상기 결합된(combined) 폴리뉴클레오타이드를 숙주 세포의 염색체에 통합시키는 예시적인 방법을 나타낸다. 집합된(assembled) 폴리뉴클레오타이드 A는 선별표지(selectable marker)의 비-기능적 첫번째 절편을 코딩하는 DNA 절편 D_{m1} 및 업스트림 게놈 타겟팅(targeting) 서열을 코딩하는 DNA 절편 D_{01}을 포함한다. 집합된(assembled) 폴리뉴클레오타이드 B는 선별표지(selectable marker)의 비-기능적 두번째 절편을 코딩하는 DNA 절편 D_{m2} 및 다운스트림 게놈 타겟팅(targeting) 서열을 코딩하는 DNA 절편 D_{02}를 포함한다. 숙주 세포는, 기능적 선별표지(selectable marker)를 포함하는 결합된(combined) 폴리뉴클레오타이드를 형성하기 위하여 집합된(assembled) 폴리뉴클레오타이드 A와 집합된(assembled) 폴리뉴클레오타이드 B를 DNA 절편(segment) D_{m1} 및 D_{m2}의 상동(homology) 영역에서 재결합하고, 결합된(combined) 폴리뉴클레오타이드를 상동 재조합(homologous recombination)에 의해 숙주 세포 염색체에 삽입하기 위하여 DNA 절편(segment) D_{01} 및 D_{02}에 의해 코딩된 게놈 타겟팅(targeting) 서열을 사용한다.

도 9는 숙주 세포에서 상동 재조합(homologous recombination)에 의해 집합된(assembled) 폴리뉴클레오타이드를 형성하고 집합된(assembled) 폴리뉴클레오타이드를 숙주 세포의 염색체로 통합하는 예시적인 방법을 나타낸다. 첫번째 단계에서, 집합 벡터를 포함하는 집합체 조성물은 제한 엔도뉴클레아제에 의해 분해되어, 집합 벡터 골격으로부터 폴리뉴클레오타이드 성분들의 절단을 유발한다. 두번째 단계에서, 폴리뉴클레오타이드 성분들은 숙주 세포에 도입되며, 여기에서, 집합된(assembled) 폴리뉴클레오타이드를 형성하기 위하여 어닐링될 수 있는 링커 서열의 상동(homology) 영역에서 재결합되고 상기 집합된(assembled) 폴리뉴클레오타이드는 숙주 세포의 염색체로 통합된다.

도 10은 동일한 반응에서 7개의 폴리뉴클레오타이드 성분들로부터 수많은 집합된(assembled) 폴리뉴클레오타이드를 집합하는 예시적인 방법을 나타낸다. 집합될 DNA 절편(segment)을 포함하는 집합 벡터는 하나의 싱글 튜브에 옮겨진 후 집합 벡터 골격으로부터 폴리뉴클레오타이드 성분들을 분리하기 위하여 SapI를 사용하여 분해된다. SapI의 열 불활성화 이후에, 폴리뉴클레오타이드 절편 성분들은 변성 조건을 겪게 되고, 상보적인 어닐링될 수 있는 링커(annealable linker) 쌍의 하이브리드에 충분한 어닐링 조건을 겪게 된다. DNA 폴리머라제 및 dNTP들의 존재 하에서 프라이머 신장 이후에, PA 및 PB에 상보적인 프라이머들이 첨가된 후, 종래의 PCR 증폭이 실시된다. 집합 반응(assembly reaction)은 5'에서 3' 방향으로 집합된 폴리뉴클레오타이드 성분들 $D_{01/02}$, $D_{1/2}$, D_3 및 $D_{41/42}$을 포함하는 집합된(assembled) 폴리뉴클레오타이드의 형성을 유발한다.

도 11은 조합적으로 결합된(combined) 폴리뉴클레오타이드들을 포함하는 수많은 숙주 세포들을 형성하는 예시적인 방법을 나타낸다. 집합된(assembled) 폴리뉴클레오타이드들 A1 및 A2 (이들은 각각, 동일한 업스트림 게놈 타겟팅(targeting) 서열 및 선별표지(selectable marker)의 비기능적인 동일한 첫번째 부분을 포함함) 및 집합된(assembled) 폴리뉴클레오타이드들 B1 및 B2 (이들은 각각 동일한 다운스트림 게놈 타겟팅(targeting) 서열 및 선별표지(selectable marker)의 비기능적인 동일한 두번째 부분을 포함함)는, 서로 다른 4개의 결합된(combined) 폴리뉴클레오타이드들, A1/B1, A1/B2, A2/B1 및 A2/B2를 형성하기 위하여 숙주 세포 매개 상동 재조합(homologous recombination)에 의하여 조합적으로 결합되며, 이들은 각각, 서로 다른 4개의 숙주 세포들을

형성하기 위하여 염색체로 삽입될 수 있는, 기능적 선별표지(selectable marker)를 포함한다.

도 12a는 조합적으로 집합된(assembled) 폴리뉴클레오타이드 및 조합적으로 집합되고(assembled) 결합된(combined) 폴리뉴클레오타이드들 및 예상된 집합되고(assembled) 결합된(combined) 폴리뉴클레오타이드들을 포함하는 효모 세포들의 높은 생산량의(high-throughput)의 형성을 위하여 실시예 10에서 사용된 폴리뉴클레오타이드 성분들을 보여준다. US = 업스트림 게놈 타겟팅(targeting) 서열, DS = 다운스트림 게놈 타겟팅(targeting) 서열, P = 다양한 프로모터 서열들, G = 다양한 단백질 코딩서열들, URA = 선별표지(selectable marker)의 5' 절편, RA3 = 선별표지(selectable marker)의 3' 절편, PA = 프라이머 결합 절편 PmeI-5', PB = 프라이머 결합 절편 PmeI-3', LB_0 = 어닐링될 수 있는 링커 서열 RYSE 2, LA_{n1} = 어닐링될 수 있는 링커 서열 RYSE 2, LB_{n1} = 어닐링될 수 있는 링커 서열 RYSE 15, LA_{n2} = 어닐링될 수 있는 링커 서열 RYSE 3, LB_{n2} = 어닐링될 수 있는 링커 서열 RYSE 16, LA_{n3} = 어닐링될 수 있는 링커 서열 RYSE 15, LB_{n3} = 어닐링될 수 있는 링커 서열 RYSE 3, LA_{n4} = 어닐링될 수 있는 링커 서열 RYSE 16, LB_{n4} = 어닐링될 수 있는 링커 서열 RYSE 4, LA_{m1} = 어닐링될 수 있는 링커 서열 RYSE 3, LA_{m2} = 어닐링될 수 있는 링커 서열 RYSE 4, LA_{m3} = 어닐링될 수 있는 링커 서열 RYSE 3를 나타낸다.

도 12b는 실시예 10에서 제시된 바와 같이 형성되고 1% 아가로스 겔 상에서 분해된 예시적인 집합된(assembled) 폴리뉴클레오타이드 (박스 부분)를 나타낸다.

도 12c는 실시예 10에서 제시된 바에 따라 얻은 예시적인 세포 콜로니에 대한 제한 분석(restriction analysis)을 나타낸다.

도 13a는 실시예 11에서 사용된 집합된(assembled) 폴리뉴클레오타이드 및 폴리뉴클레오타이드 성분과, 집합 및 숙주 세포에 의한 염색체의 통합(integration)으로부터 얻은 예상되는 염색체 위치를 나타낸다.

도 13b는 염색체로 통합된 집합된(assembled) 폴리뉴클레오타이드들을 포함하는, 실시예 11에서 형성된 효모 세포 형질변환체로부터 얻은 cPCR 분석 결과를 보여준다.

도 14는 염색체로 통합된 조합적으로 집합되고(assembled) 조합적으로 결합된(combined) 폴리뉴클레오타이드들 및 집합 및 숙주 세포 매개 상동 재조합(homologous recombination)의 조합에서 얻은 예상되는 결합된(combined) 폴리뉴클레오타이드들을 포함하는 효모 세포의 높은 생산량의(high-throughput) 형성을 위하여, 실시예 12에서 사용된 폴리뉴클레오타이드 성분들을 보여준다. US = 업스트림 게놈 타겟팅(targeting) 서열, DS = 다운스트림 게놈 타겟팅(targeting) 서열, P = 다양한 프로모터 서열, G = 다양한 단백질 코딩서열, URA = 선별표지(selectable marker)의 5' 절편, RA3 = 선별표지(selectable marker)의 3' 절편, PA = 프라이머 결합 절편 PmeI-5', PB = 프라이머 결합 절편 PmeI-3', LB_0 = 어닐링될 수 있는 링커 서열 RYSE 2, LA_{n1} = 어닐링될 수 있는 링커 서열 RYSE 2, LB_{n1} = 어닐링될 수 있는 링커 서열 RYSE 15, LA_{n2} = 어닐링될 수 있는 링커 서열 RYSE 3, LB_{n2} = 어닐링될 수 있는 링커 서열 RYSE 16, LA_{n3} = 어닐링될 수 있는 링커 서열 RYSE 15, LB_{n3} = 어닐링될 수 있는 링커 서열 RYSE 3, LA_{n4} = 어닐링될 수 있는 링커 서열 RYSE 16, LB_{n4} = 어닐링될 수 있는 링커 서열 RYSE 4, LA_{m1} = 어닐링될 수 있는 링커 서열 RYSE 3, LA_{m2} = 어닐링될 수 있는 링커 서열 RYSE 4, LA_{m3} = 어닐링될 수 있는 링커 서열 RYSE 3을 나타낸다.

발명을 실시하기 위한 구체적인 내용

[0218] **실시예**

[0219] 본 발명은 하기의 실시예에 의해 예시적으로 설명되며, 이것은 어떠한 방식으로도 본 발명을 제한하려는 의도가 아니다. 실시예에 제시된 *Saccharyomices cerevisiae* 구성은 *Saccharyomices cerevisiae* 종(strain) CEN.PK2로부터 유래되었다. *Saccharyomices cerevisiae* 종(strain) S288c과 달리, 종(strain) CEN.PK2의 게놈 서열은 대중적으로 이용할 수 없다. 제시된 구성의 몇몇은 서열이 밝혀진 것이어서, 제공된 서열들의 구성은 실제의 CEN.PK2-유래 구성이다. 서열이 밝혀지지 않은 구성에 대하여, 제공된 서열은 종(strain) S288c의 공개된 게놈 서열에 기초하기 때문에, 실제의 CEN.PK2-유래 구성의 서열과 다형성(polymorphic) 차이를 포함할 수 있다.

[0220] 실시예 1

[0221] 본 실시예는 pRYSE 벡터를 만드는 방법을 제시한다. pRYSE 벡터는, 5'에서 3'의 방향으로, 첫번째 SapI 제한효소 인지 부위, 첫번째 어닐링될 수 있는 링커 서열 또는 프라이머 결합 절편, 첫번째 SchI 제한효소 인지 부위, 녹색 형광 단백질 (GFP) 또는 lacZ 표지 유전자, 두번째 SchI 제한효소 인지 부위, 두번째 어닐링될 수 있는 링커 서열 또는 프라이머 결합 절편 및 두번째 SapI 제한효소 인지 부위를 포함한다.

[0222] β-락타마아제를 코딩하는 DNA 절편은, pUC19의 bla 유전자 내 SchI 제한효소 인지 부위가 PCR 프라이머 JCB158-17A (서열번호 227) 및 JCB1 58-17B (서열번호 228)을 사용한 pUC19의 부위-특이적(site-directed) 돌연변이유발(mutagenesis)에 의해 제거된 후에, 프라이머 JCB158-17C (서열번호 229) 및 JCB158-17D (서열번호 230)을 사용한 pUC19 벡터 (유전자 은행의 등록번호 L09137)로부터 PCT 증폭되었다. PCR 생산물은 겔 정제 된 후 TOPO 벡터 (Invitrogen, Carlsbad, CA)로 연결되고, 상기 결과물은 SphI 및 MfeI 제한효소를 사용하여 완전히 분해하여 다시 유리되어, "bla DNA 절편"을 생성하였다.

[0223] DNA 절편들 1040 (서열번호 224), 1041 (서열번호 225) 및 1042 (서열번호 226)은 종합적으로 형성되었다 (Biosearch Technologies, Novato, CA). DNA 절편들 1040 및 1041은 *BstXI* 제한효소를 사용하여 완전히 분해되었고, 각각의 분해된 절편은, 제한효소 *SacI* 및 *KpnI*를 사용하여 pAM1466 (서열번호 223; Biosearch Technologies, Novato, CA에 의해 종합적으로 형성됨)를 완전히 커팅하여 형성된 2.65 kb 벡터 골격과 연결되었다. 1040_pAM1466 DNA 구성은 *BsmBI* 및 *BstXI* 제한효소를 사용하여 완전히 분해되었고, 상기 반응 혼합물은 겔 전기영동에 의해 분해되었고, 1040 DNA 절편을 포함하는 대략 3.5 kb DNA 절편은 겔 정제되었다. 1041_pAM1466 DNA 구성은 *BsaI* 및 *BstXI* 제한효소를 사용하여 완전히 분해되었고, 상기 반응 혼합물은 겔 전기영동에 의해 분해되었고, 및 1041 DNA 절편을 포함하는 대략 0.9 kb 1041 DNA 절편은 겔 정제되었다. 정제된 DNA 절편들이 연결되어, DNA 구성 1040_1041_pAM1466을 형성하였다. 1040_1041 DNA 절편을 형성하기 위한 프라이머들 JO36 (서열번호 69) 및 JO37 (서열번호 70), 1040_1041 DNA 절편의 말단 서열과 겹치는 말단 서열을 갖는 1042 DNA 절편을 형성하기 위한 프라이머들 JO38 (서열번호 71) 및 JO39 (서열번호 72), 및 2개의 PCR 생성물에 결합하기 위한 프라이머들 JO39 (*SphI* 제한효소 인지 부위를 포함함) (서열번호 72) 및 JO36 (*MfeI* 제한효소 인지 부위를 포함함) (서열번호 69)을 사용하는 PCR "스티칭(stitching)" 반응에 의해, DNA 절편 1042은 DNA 구성 1040_1041에 결합되었다. 상기 1040_1041_1042 PCR 생성물은 *SphI* 및 *MfeI* 제한효소를 사용하여 완전해 분해되었고, 상기 반응 혼합물은 겔 전기영동에 의해 분해되었고, 대략 2.4 kb 1040_1041_1042 DNA 절편은 겔 정제되었고, 정제된 DNA 절편은 겔 정제된 *bla* 절편에 연결되어, "1040_1041_1042_bla" DNA 구성을 형성하였다.

[0224] GFP 유전자를 코딩하는 1040_1041_1042_bla DNA 구성의 절편은 PCR 프라이머 1 및 2 (표 2 참조)를 사용하여 PCR 증폭되었다. PCR 반응의 첫번째 단계에서 형성된 겔-추출된 GFP 절편 및 PCR 프라이머 3 및 4를 주형(template)으로 사용하는 PCR 증폭에 의해, *SacI* 및 *XhoI* 제한효소 인지 부위는 상기 증폭된 GFP 절편 말단에 첨가되었다(표 2 참조). 증폭된 PCR 생성물은 겔 추출된 후, *XhoI* 및 *SacI* 제한효소를 사용하여 완전히 분해한 후, 제한효소는 65℃에서 20분 동안 열 불활성화되었고, 분해된 PCR 생성물은 컬럼 정제된 후 겔 정제된, *XhoI* 및 *SacI* 제한효소를 사용하여 1040_1041_1042_bla DNA 구성을 완전히 분해하여 얻은 대략 2.2 kb DNA 절편에 연결되었다. 결과물 벡터는 PCR 프라이머들 5 및 6을 사용하여 PCR 증폭되었고 (표 3 참조), 결과물 혼합물은 겔 전기영동에 의해 분해되었고, 대략 2.2 kb "pRYSE 벡터 골격"이 겔 정제되었다.

표 2

표 2 – 어닐링할 수 있는 링커 쌍 또는 어닐링할 수 있는 링커/프라이머 결합 절편 쌍 및 SacI 및 XhoI 제한효소 부위의 측면에 배치된 GFP 삽입물을 형성하기 위하여 사용된 PCR 프라이머들

GFP 절편	어닐링 할 수 있는 링커 또는 프라이머 결합 절편 1	어닐링 할 수 있는 링커 또는 프라이머 결합 절편 2	프라이머 1	프라이머 2	프라이머 3	프라이머 4
1	Pme1-5'	RYSE 1	J018 (SEQ ID NO: 73)	J073 (SEQ ID NO: 106)	J055 (SEQ ID NO: 88)	J064 (SEQ ID NO: 97)
2	RYSE 1	RYSE 2	J019 (SEQ ID NO: 74)	J074 (SEQ ID NO: 107)	J056 (SEQ ID NO: 89)	J065 (SEQ ID NO: 98)
3	RYSE 2	RYSE 3	J020 (SEQ ID NO: 75)	J029 (SEQ ID NO: 82)	J057 (SEQ ID NO: 90)	J066 (SEQ ID NO: 99)
4	RYSE 3	RYSE 4	J021 (SEQ ID NO: 76)	J030 (SEQ ID NO: 83)	J058 (SEQ ID NO: 91)	J067 (SEQ ID NO: 100)
5	RYSE 4	RYSE 5	J022 (SEQ ID NO: 77)	J031 (SEQ ID NO: 84)	J059 (SEQ ID NO: 92)	J068 (SEQ ID NO: 101)
6	RYSE 5	RYSE 6	J023 (SEQ ID NO: 78)	J032 (SEQ ID NO: 85)	J060 (SEQ ID NO: 93)	J069 (SEQ ID NO: 102)
7	RYSE 6	RYSE 7	J024 (SEQ ID NO: 79)	J033 (SEQ ID NO: 86)	J061 (SEQ ID NO: 94)	J070 (SEQ ID NO: 103)
8	RYSE 7	RYSE 8	J025 (SEQ ID NO: 80)	J034 (SEQ ID NO: 87)	J062 (SEQ ID NO: 95)	J071 (SEQ ID NO: 104)
9	RYSE 2	Pme1-3'	J020 (SEQ ID NO: 75)	J075 (SEQ ID NO: 108)	J057 (SEQ ID NO: 90)	J072 (SEQ ID NO: 105)
10	RYSE 3	Pme1-3'	J021 (SEQ ID NO: 76)	J075 (SEQ ID NO: 108)	J058 (SEQ ID NO: 91)	J072 (SEQ ID NO: 105)
11	RYSE 4	Pme1-3'	J022 (SEQ ID NO: 77)	J075 (SEQ ID NO: 108)	J059 (SEQ ID NO: 92)	J072 (SEQ ID NO: 105)
12	RYSE 5	Pme1-3'	J023 (SEQ ID NO: 78)	J075 (SEQ ID NO: 108)	J060 (SEQ ID NO: 93)	J072 (SEQ ID NO: 105)
13	RYSE 6	Pme1-3'	J024 (SEQ ID NO: 79)	J075 (SEQ ID NO: 108)	J061 (SEQ ID NO: 94)	J072 (SEQ ID NO: 105)
14	RYSE 7	Pme1-3'	J025 (SEQ ID NO: 80)	J075 (SEQ ID NO: 108)	J062 (SEQ ID NO: 95)	J072 (SEQ ID NO: 105)
15	RYSE 8	Pme1-3'	J026 (SEQ ID NO: 81)	J075 (SEQ ID NO: 108)	J063 (SEQ ID NO: 96)	J072 (SEQ ID NO: 105)

[0225]

[0226]

표 3

표 3 - pRYSE 벡터에 존재하는 어닐링할 수 있는 링커 쌍 또는 어닐링할 수 있는 링커 서열/프라이머 결합 절편 쌍 및 pRYSE 벡터 골격을 형성하는데 사용된 PCR 프라이머들

pRYSE 벡터	어닐링 할 수 있는 링커 또는 프라이머 결합 절편 1 (표1참조)	어닐링 할 수 있는 링커 또는 프라이머 결합 절편 2 (표1참조)	프라이머 5	프라이머 6
1	Pme1-5'	RYSE 1	S001 (SEQ ID NO: 46)	S002 (SEQ ID NO: 47)
2	RYSE 1	RYSE 2	S003 (SEQ ID NO: 48)	S004 (SEQ ID NO: 49)
3	RYSE 2	RYSE 3	S005 (SEQ ID NO: 50)	S006 (SEQ ID NO: 51)
4	RYSE 3	RYSE 4	S007 (SEQ ID NO: 52)	S008 (SEQ ID NO: 53)
5	RYSE 4	RYSE 5	S009 (SEQ ID NO: 54)	S010 (SEQ ID NO: 55)
6	RYSE 5	RYSE 6	S011 (SEQ ID NO: 56)	S012 (SEQ ID NO: 57)
7	RYSE 6	RYSE 7	S013 (SEQ ID NO: 58)	S014 (SEQ ID NO: 59)
8	RYSE 7	RYSE 8	S015 (SEQ ID NO: 60)	S016 (SEQ ID NO: 61)
9	RYSE 2	Pme1-3'	S005 (SEQ ID NO: 50)	S018 (SEQ ID NO: 63)
10	RYSE 3	Pme1-3'	S007 (SEQ ID NO: 52)	S018 (SEQ ID NO: 63)
11	RYSE 4	Pme1-3'	S009 (SEQ ID NO: 54)	S018 (SEQ ID NO: 63)
12	RYSE 5	Pme1-3'	S011 (SEQ ID NO: 56)	S018 (SEQ ID NO: 63)
13	RYSE 6	Pme1-3'	S013 (SEQ ID NO: 58)	S018 (SEQ ID NO: 63)
14	RYSE 7	Pme1-3'	S015 (SEQ ID NO: 60)	S018 (SEQ ID NO: 63)
15	RYSE 8	Pme1-3'	S017 (SEQ ID NO: 62)	S018 (SEQ ID NO: 63)

[0227]

[0228] *lacZ* 유전자는 프라이머들 S027 (서열번호 65) 및 S028 (서열번호 66)를 사용하여 pUC 19 벡터로부터 PCR 증폭되었고, 각각의 프라이머는 SchI 제한효소 인지 부위를 포함한다. 상기 반응 혼합물은 겔 전기영동에 의해 분해되었고, 대략 0.5 kb PCR 생성물은 겔 정제되고, 정제된 PCR 생성물은 각각의 pRYSE 벡터 골격과 연결되었다. 부위-특이적(site-directed) 돌연변이유발(mutagenesis)은, 복제의 원점(origin)으로부터 SchI 제한효소 인지 부위를 제거하기 위하여 PCR 프라이머 L012 (서열번호 231) 및 L013 (서열번호 232)를 사용하여 결과물 벡터에서 실시되었다. 마지막으로, 두번째 부위-특이적(site-directed) 돌연변이유발(mutagenesis)은, *lacZ* 절편으로부터 SchI 제한효소 인지 부위를 제거하기 위하여, PCR 프라이머들 S036 (서열번호 67) 및 S037 (서열번호 68)을 사용하여 실시되어, pRYSE 벡터들 1 내지 15를 형성하였다(pRYSE 벡터들 및, pRYSE 벡터들 1 내지 15의 뉴클레오타이드 서열에 대한 서열번호 207의 플라즈미드 맵(map)에 대한 도 4를 참조).

[0229] 실시예 2

[0230] 본 실시예는 pRYSE 벡터들을 만드는 대안적 방법을 제시한다.

[0231] pRYSE 벡터들 1 내지 15는 주형(template)으로서 서열번호 207 내지 221를 사용하여 종합적으로 형성될 수 있다 (예를 들어, Biosearch Technologies, Novato, CA에 의함). 서로 다른 어닐링될 수 있는 링커 서열들을 포함하는 추가적 pRYSE 벡터는 주형(template)으로서 서열번호 221을 사용하여 종합적으로 형성될 수 있으며, 상기 서열번호 221에서 Pmel-5' 프라이머 결합 절편 및/또는 RYSE 1 어닐링될 수 있는 링커 서열들이 다른 적당한 어닐링될 수 있는 링커 서열 또는 프라이머 결합 절편으로 바뀌었다 (표 1 참조).

[0232] 실시예 3

[0233] 본 실시예는, 5'에서 3'의 방향으로, 첫번째 SapI 제한효소 인지 부위, 첫번째 SchI 제한효소 인지 부위, *lacZ* 표지 유전자, 두번째 SchI 제한효소 인지 부위 및 두번째 SapI 제한효소 인지 부위를 포함하는 pMULE 벡터를 만드는 방법을 제시한다. pMULE 벡터는 Mule을 복제하기 위하여 사용될 수 있다.

[0234] pRYSE 벡터 8의 골격은 프라이머들 K162 (서열번호 109) 및 K163 (서열번호 110)을 사용하여 PCR 증폭되었다. 상기 반응 혼합물은 겔 전기영동에 의해 분해되었고, 대략 2.2 kb 벡터 골격은 겔 정제되었다. SchI 제한효소를 사용하여 pRYSE 벡터 8을 완전히 분해하고, 상기 효소를 65℃에서 20분 동안 열 불활성화하고, 겔 전기영동에 의해 반응물 혼합물을 분해하고 및 대략 0.5 kb DNA 절편을 겔 정제함으로써, *lacZ* 유전자를 포함하는 DNA 절편이 형성되었다. *lacZ* 유전자를 포함하는 정제된 DNA 절편은 정제된 벡터 골격과 연결되어, pMULE 벡터를 형성하였다 (플라즈미드 맵(map)에 대한 도 7을 참조).

[0235] 실시예 4

[0236] 본 실시예는 "비트(Bit)"를 만드는 방법을 제시한다. 비트(Bit)는 본원에 제시된 방법들을 사용하여 집합된 (assembled) 폴리뉴클레오타이드로 집합될 수 있는 폴리뉴클레오타이드 성분들을 포함하는 집합 벡터들을 형성하기 위하여, pRYSE 벡터로 삽입될 수 있는 DNA 절편이다. 비트(Bit)는 유전자 또는 관심있는 유전 성분 (예를 들어, 프로모터, 종결자(terminator), 선별표지(selectable marker), 통합 부위(integration loci), 에피토프 태그(epitope tag), 위치 신호(localization signal), 분해 신호(degradation signal), 형광 표지, 다수의 클로닝 부위)을 코딩할 수 있다. 비트(Bit)는 표 4에 제시된 프라이머들을 사용하여 주형으로부터 PCR 증폭되었다.

표 4

표 4 - 증폭된 비트(Bit)

Bit	Type *	프라이머	크기 (bp)	주형
atoB	Gs	L229 (SEQ ID NO: 40) L230 (SEQ ID NO: 41)	1185	Escherichia coli(유전자 은행 등록번호 NC_000913 영역:2324131..2325315) 로부터의 atoB 유전자를 포함하는 플라즈미드 DNA
mvaS	Gs	L235 (SEQ ID NO: 42) L236 (SEQ ID NO: 43)	1152	Saccharomyces cerevisiae에서의 발현에 최적화된 Enter°C°Ccus faecalis (유전자 은행 등록번호 AF290092 영역:142..1293) 코돈으로부터의 mvaS 유전자를 포함하고 효소 활성을 증가시키기 위하여 위치 110에서 알라닌을 글리신으로 변형한 것을 포함하는 합성 DNA 절편(Steussy et al.(2006) Bi°Chemistry 45(48):14407-14414)를 참조)
ERG13-1	GsT	L109 (SEQ ID NO: 235) L110 (SEQ ID NO: 26)	1726	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
3' NDT80	D	L221 (SEQ ID NO: 34) L222 (SEQ ID NO: 35)	516	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
5' NDT80	U	L219 (SEQ ID NO: 32) L220 (SEQ ID NO: 33)	495	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
tP_{FBA1}	P	L225 (SEQ ID NO: 37) L057 (SEQ ID NO: 234)	526	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
tP_{TDH3}	P	L224 (SEQ ID NO: 36) L054 (SEQ ID NO: 233)	559	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
ERG10-1	Gs	L226 (SEQ ID NO: 38) L227 (SEQ ID NO: 39)	1182	Saccharomyces cerevisiae에서의 발현에 최적화된 Ralstonia eutropha (유전자 은행 등록번호 NC_008313 영역:183291..184469) 코돈의 아세틸-CoA 아세틸트랜스퍼라제를 코딩한 후 추가적인 종결 코돈에 의해 합성된 절편
tENO1	T	L248 (SEQ ID NO: 44) L176 (SEQ ID NO: 27)	265	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
tTDH3	T	L185 (SEQ ID NO: 28) L186 (SEQ ID NO: 29)	260	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
HphA	M	TRIX_L_193 (SEQ ID NO: 184) TRIX_L_194 (SEQ ID NO: 185)	1912	TEF1 프로모터 및 Kluyveromyces lactis (각각, 유전자은행 등록번호 CR382122 영역:788874..789380 및 787141..787496)의 종결자 및 Klebsiella pneumonia의 hph 유전자를 포함하는 플라즈미드 DNA
tHMG1	GsT	TRIX_L_232 (SEQ ID NO: 186)	1742	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA

[0237]

		TRIX_L_233 (SEQ ID NO: 187)		
$tP_{GAL1,10}$	P	TRIX_L_266 (SEQ ID NO: 190) TRIX_L_267 (SEQ ID NO: 191)	620	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
ERG10-2	GsT	TRIX_L_106 (SEQ ID NO: 170) TRIX_L_107 (SEQ ID NO: 171)	1467	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
ERG13-2	GsT	TRIX_L_109 (SEQ ID NO: 172) TRIX_L_110 (SEQ ID NO: 173)	1726	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
GAL80US	U	JU-218-168-130-GAL80US-F (SEQ ID NO: 134) JU-219-168-130-GAL80US-R (SEQ ID NO: 135)	500	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
GAL80DS	D	JU-220-168-130-GAL80DS-F (SEQ ID NO: 136) JU-221-168-130-GAL80DS-R (SEQ ID NO: 137)	500	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
P_{TDH3}	P	L224 (SEQ ID NO: 36) TRIX_L_053 (SEQ ID NO: 169)	583	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA

[0238]

표 4 - 증폭된 비트(계속됨)

Bit	Type *	프라이머	크기 (bp)	주형
NatA	M	TRIX_L_193 (SEQ ID NO: 184) TRIX_L_194 (SEQ ID NO: 185)	1456	TEF1 프로모터 및 Kluyveromyces lactis (각각, 유전자은행 등록번호 CR382122 영역:788874..789380 및 787141..787496)의 종결자 및 S. noursei의 natI 유전자를 포함하는 플라즈미드 DNA
ERG12	GsT	TRIX_L_112 (SEQ ID NO: 174) TRIX_L_113 (SEQ ID NO: 175)	1582	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
ERG8	GsT	TRIX_L_118 (SEQ ID NO: 178) TRIX_L_119 (SEQ ID NO: 179)	1616	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
P_{GAL4oc}	P	TRIX_K_131 (SEQ ID NO: 165) PW-91-093-CPK422-G (SEQ ID NO: 162)	270	Saccharomyces cerevisiae 종 CEN.PK2(Griggs & Johnston(1991) PNAS 88(19):8597-8601)의 GAL4 유전자의 프로모터의 "중요한 구성요소의(operative constitutive)" 형태를 포함하는 플라즈미드 DNA
GAL4-1	G	JU-286-275-31-GAL4-F (SEQ ID NO: 140) JU-285-275-31-GAL4-FIX-R2 (SEQ ID NO: 139)	526	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
GAL4-2	G	JU-284-275-31-GAL4-FIX-F2 (SEQ ID NO: 138) JU-287-275-31-GAL4-R (SEQ ID NO: 141)	2414	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
KanA	M	TRIX_L_193 (SEQ ID NO: 184) TRIX_L_194 (SEQ ID NO: 185)	1696	TEF1 프로모터 및 Kluyveromyces lactis (각각, 유전자은행 등록번호 CR382122 영역:788874..789380 및 787141..787496)의 종결자 및 Tn903 트랜스포존의 kanR 유전자를 포함하는 플라즈미드 DNA
ERG19	GsT	TRIX_L_115 (SEQ ID NO: 176) TRIX_L_116 (SEQ ID NO: 177)	1441	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
ERG20	GsT	TRIX_L_124 (SEQ ID NO: 182) TRIX_L_125 (SEQ ID NO: 183)	1319	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
P_{GAL7}	P	TRIX_L_34 (SEQ ID NO:	500	*Saccharomyces cerevisiae* 종

[0239]

		166) TRIX_L_35 (SEQ ID NO: 167)		CEN.PK2 게놈 DNA
tP_{GAL7}	P	TRIX_L_34 (SEQ ID NO: 166) TRIX_L_36 (SEQ ID NO: 168)	476	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
IDI1	GsT	TRIX_L_121 (SEQ ID NO: 180) TRIX_L_122 (SEQ ID NO: 181)	1127	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
tP_{CTR3}	P	TRIX_K_0142 (SEQ ID NO: 163) TRIX_K_0143 (SEQ ID NO: 164)	710	Saccharomyces cerevisiae 종 CEN.PK2의 CTR3 유전자의 프로모터를 포함하는 플라즈미드 DNA
LEU2US	U	JU-164-168-110-LEU2 US-f (SEQ ID NO: 129) JU-165-168-110-LEU2 US-r (SEQ ID NO: 130)	500	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
LEU2DS	D	JU-162-168-110-LEU2 DS-f (SEQ ID NO: 127) JU-163-168-110-LEU2 DS-r (SEQ ID NO: 128)	500	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
ERG9US	U	JU-108-168-110-ERG9 US-f (SEQ ID NO: 126) JU-172-168-110-ERG9 US-r1 (SEQ ID NO: 133)	499	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
ERG9CDS	G	JU-106-168-110-ERG9 CDS-f (SEQ ID NO: 124) JU-107-168-110-ERG9 CDS-r (SEQ ID NO: 125)	501	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
STE5US	U	TRIX_RN017 (SEQ ID NO: 192) TRIX_RN018 (SEQ ID NO: 193)	600	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
STE5DS	D	TRIX_RN019 (SEQ ID NO: 194) TRIX_RN020 (SEQ ID NO: 195)	600	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA

[0240]

표 4 - 증폭된 비트(계속됨)

Bit	Type *	프라이머	크기 (bp)	주형
URA3	M	JU-169-168-110-URA3-f (SEQ ID NO: 131) JU-170-168-110-URA3-r (SEQ ID NO: 132)	1554	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA

* G=유전자; s=종결 코돈; T=종결자; M=표지; D=다운스트림 통합 영역; U=업스트림 통합 영역; P=프로모터.

[0241]

[0242] PCR 증폭은, 제조업자가 제시한 프로토콜에 따라 Phusion DNA 폴리머라제 (New England Biolabs, Ipswich, MA)를 사용하여 이루어졌다. PCR 반응물은 겔 전기영동에 의해 분해되었고, 상기 비트(Bit)는 겔 정제되었고 정제된 비트(Bit)는 제조업자가 제시한 프로토콜에 따라 T4 폴리뉴클레오타이드 키나아제 (PNK) (New England Biolabs, Ipswich, MA)로 처리되었다. PNK는 65℃에서 20분 동안 열 불활성화 되었고, 샘플은 -20℃에서 저장되었다.

[0243] 실시예 5

[0244] 본 실시예는 "MULE"를 만드는 방법을 제시한다. MULE는 본원에 제시된 방법을 사용하여 집합된(assembled) 폴리뉴클레오타이드들에 집합될 수 있는 폴리뉴클레오타이드 성분을 포함하는 집합 벡터를 형성하기 위하여 pMULE 벡터로 삽입될 수 있는 DNA 절편이다. MULE은 유전자, 또는 어닐링될 수 있는 링커 서열 쌍 또는 어닐링될 수 있는 링커 서열 /프라이머 결합 절편 쌍의 측면에 배치된 관심있는 유전 성분 (예를 들어, 프로모터, 종결자(terminator), 선별표지(selectable marker), 통합 부위(integration loci), 에피토프 태그(epitope tag), 위치 신호(localization signal), 분해 신호(degradation signal), 형광 표지, 다수의 클로닝 부위)를 코딩할 수 있다. MULE는, 표 5에 제시되어 있는 바와 같이, 프라이머를 사용하여 주형(template)으로부터 PCR 증폭되었으며, 상기 프라이머의 3' 말단은 타겟 서열에 어닐링되며 5' 말단은 어닐링될 수 있는 링커 서열 또는 프라이머 결합 절편을 포함한다 (적당한 어닐링될 수 있는 링커 서열에 대한 표 1 참조).

표 5

표 5 - 증폭된 MULE

MULE	Type *	프라이머	크기 (bp)	주형
tHMG1-a	G	KMH8-276-1-linker4.tHMG1.fwd (SEQ ID NO: 157) KMH9-276-1-linker9.tHMG1.rev (SEQ ID NO: 160)	1794	RABit 254 프라즈미드 DNA
ERG12	G	KMH46-276-43-ERG12linker4.fwd (SEQ ID NO: 151) KMH14-276-4-linker9.ERG12.rev (SEQ ID NO: 145)	1634	RABit 250 프라즈미드 DNA
ERG19	G	KMH47-276-43-ERG19linker4.fwd (SEQ ID NO: 152) KMH15-276-4-linker9.ERG19.rev (SEQ ID NO: 146)	1493	RABit 241 프라즈미드 DNA
P_{TDH3}-a	P	KMH81-276-116-TDH3.rev.tHMG1 (SEQ ID NO: 155) S004 (SEQ ID NO: 49)	626	RABit 54 프라즈미드 DNA
P_{TDH3}-b	P	KMH91-276-116-TDH3.rev.FS (SEQ ID NO: 158) S004 (SEQ ID NO: 49)	546	RABit 54 프라즈미드 DNA
tHMG1-b	G	KMH82-276-116-tHMG1.fwd.TDH3 (SEQ ID NO: 156) S009 (SEQ ID NO: 54)	1801	RABit 20 프라즈미드 DNA
IME1US	U	KB454-266-53 (SEQ ID NO: 142) KB455-266-53 (SEQ ID NO: 143)	578	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
IME1DS	D	KMH93-276-130-3'IME.linker4.fwd (SEQ ID NO: 161) KB457-266-53 (SEQ ID NO: 144)	554	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
LEU2	M	VH296-235-55-Leu2 12-1 F (SEQ ID NO: 30) VH296-235-55-Leu2 12-1 R (SEQ ID NO: 31)	1795	Saccharomyces cerevisiae 종 CEN.PK2(Sikorski RS, Hieter(1989) Genetics 122(1):19-27)의 LEU2 위치를 포함하는 플라즈미드 DNA

[0245]

FS-a	G	KMH5-276-1-linker3.FS(Kozak).fwd (SEQ ID NO: 153) KMH7-276-1-linker4.TCYC1.rev (SEQ ID NO: 154)	1981	Saccharomyces cerevisiae에서의 발현에 최적화된 Artemisia annua (유전자 은행 등록번호 AY835398) 코돈의 파네센 신타아제의 코딩 서열 및 Saccharomyces cerevisiae 종 CEN.PK2의 CYC1 유전자의 종결자를 포함하는 플라즈미드 DNA
FS-b	G	KMH92-276-116-FS.fwd.TDH3 (SEQ ID NO: 159) KMH7-276-1-linker4.TCYC1.rev (SEQ ID NO: 154)	1976	Saccharomyces cerevisiae에서의 발현에 최적화된 Artemisia annua (유전자 은행 등록번호 AY835398) 코돈의 파네센 신타아제의 코딩 서열 및 Saccharomyces cerevisiae 종 CEN.PK2의 CYC1 유전자의 종결자를 포함하는 플라즈미드 DNA
URA3 블라스터	M	VH228-235-7-URA3LOF3RYSE12-1F (SEQ ID NO: 204) VH229-235-7-URA3LOF3RYSE12-1R (SEQ ID NO: 205)	1565	URA-3 블라스터 주형 **

* G=유전자; s=종결 코돈; T=종결자; M=표지; D=다운스트림 통합 영역; U=업스트림 통합 영역; P=프로모터.

** URA-3 블라스터(blaster) 주형은 서열 A(PCR 프라이머들 TRIX_Z025(서열번호 196) 및 TRIX_Z026(서열번호 197)을 사용하는 서열번호 206을 포함하는 합성 DNA 절편으로부터 형성됨), 서열 B(PCR 프라이머들 TRIX_Z027(서열번호 198) 및 TRIX_Z028(서열번호 199)을 사용하는 서열번호 206을 포함하는 합성 DNA 절편으로부터 형성됨), URA3-c(PCR 프라이머들 TRIX_Z033(서열번호 200) 및 TRIX_Z036(서열번호 203)을 사용하는 Saccharomyces cerevisiae 종 CEN.PK2 게놈 DNA로부터 형성됨), 및 URA3-d(PCR 프라이머들 TRIX_Z034 (서열번호 201) 및 TRIX_Z035(서열번호 202)을 사용하는 Saccharomyces cerevisiae 종 CEN.PK2 게놈 DNA로부터 형성됨)을 측면에 배치하는 DNA 절편을 첫 번째로 형성함으로써 만들어 진다. 그 후, 서열 A, URA3-c 및 URA3-d를 측면에 배치하는 DNA 절편은 PCR 프라이머들 TRIX_Z025 및 TRIX_Z034를 사용하는 DNA 절편으로 함께 스티칭되었고, DNA 절편들 URA3-c, URA3-d 및 서열 B를 측면에 배치하는 DNA 절편들은 PCR 프라이머들 TRIX_Z028 및 TRIX_Z033를 사용하는 DNA 절편으로 함께 스티칭되었다. 마지막으로 DNA 절편들 A 및 B는 PCR 프라이머들 TRIX_Z025 및 TRIX_Z028를 사용하여 함께 스티칭되어, URA-3 블라스터 주형을 형성하였다.

[0246]

[0247] PCR 증폭은, 제조업자가 제시한 프로토콜에 따라, Phusion DNA 폴리머라제 (New England Biolabs, Ipswich, MA)를 사용하여 실시되었다. PCR 반응물은 겔 전기영동에 의해 분해되었고, MULE은 겔 정제되었고, 정제된 MULE은, 제조업자가 제시한 프로토콜에 따라, T4 폴리뉴클레오타이드 키나아제 (PNK) (New England Biolabs, Ipswich, MA)로 처리되었다. PNK는 65℃에서 20분 동안 열 불활성화되었고, 샘플은 -20℃에서 저장되었다.

[0248] 실시예 6

[0249] 본 실시예는 집합 벡터를 형성하기 위하여 비트(Bit)를 pRYSE 벡터로 삽입하거나, MULE를 pMULE 벡터로 삽입하는 방법을 제시한다.

[0250] pRYSE 벡터들 1 내지 8과 pRYSE 벡터 15는 SchI 제한효소를 사용하여 완전히 분해되었고, 분해된 DNA 절편은 Antarctic Phosphatase (New England Biolabs, Ipswich, MA)를 사용하여 처리되었다. 상기 포스파타아제는 65℃에서 20분 동안 열 불활성화되었고, 반응물 혼합물은 겔 전기영동에 의해 분해되었고 대략 2.2 kb pRYSE 벡터 골격(*lacZ* 결핍)은 겔 정제되었다. 정제된 pRYSE 벡터 골격은 표 6에 상세히 나와 있는 바와 같이 비트(Bit)와 연결되어, 집합 벡터를 형성하였다.

[0251] pMULE 벡터는 SchI 제한효소를 사용하여 완전히 분해되었고, 반응물 혼합물은 겔 전기영동에 의해 분해되었고 대략 2.2 kb pMULE 벡터 골격(*lacZ* 결핍)은 겔 정제되었다. 정제된 pMULE 벡터 골격은 포스파타아제 (예를 들어, Antarctic Phosphatase (New England Biolabs, Ipswich, MA), CIAP (New England Biolabs, Ipswich, MA), SAP (New England Biolabs, Ipswich, MA; Fermentas, Glen Burnie, MD), 또는 FastAP (Fermentas, Glen Burnie, MD))로 처리되었고, 상기 포스파타아제는 열 불활성화되었고 (예를 들어, 65℃에서 20분) 및 pMULE 벡터 골격은 MULE과 연결되어 집합 벡터를 형성하였다.

표 6

표 6 - 형성된 집합 벡터

Bit(표 4 참조)	pRYSE 벡터(표 3 참조)	집합 벡터
atoB	4	2
mvaS	7	5
ERG13-1	7	12
3' NDT80	15	29
	10	24
5' NDT80	1	30
	1	97
tP_{FBA1}	6	35
tP_{TDH3}	3	53
ERG10-1	4	60
tENO1	8	62
tTDH3	5	64
GAL80US	1	270
HphA	2	22
tHMG1	3	254
$tP_{GAL1,10}$	4	229
ERG10-2	5	244
ERG13-2	6	253
$tP_{GAL1,10}$	7	228
tHMG1	8	255
GAL80DS	15	271
LEU2US	1	187
NatA	2	262
ERG12	3	250
ERG8	5	252
P_{GAL4oc}	6	268
GAL4 *	7	265
LEU2DS	14	263
ERG9US	1	186
KanA	2	261
ERG19	3	241
ERG20	5	251
tP_{GAL7}	6	249
IDI1	7	237
tP_{CTR3}	8	269
ERG9CDS	15	185
P_{GAL7}	3	44
STE5US	1	567
URA3	2	556(방향 1)
		555(방향 2)
P_{TDH3}	3	54
tHMG1	4	20
STE5DS	11	563

[0252]

연결은, 제조업자가 제시한 프로토콜에 따라, 50 ng 벡터 골격, 3 몰을 초과하는 비트(Bit)및 연결효소 (예를 들어, Quick Ligase (New England Biolabs, Ipswich, MA), T4 DNA 연결효소 (균일하고 높은 농도; 벤더(vendor), Fermentas, Glen Burnie, MD), Fast Ligase (Fermentas, Glen Burnie, MD))를 사용하여 실시되었다.

* 비트(Bit) GAL4는, 프라이머들 JU-286-275-31-GAL4-F (서열번호 140) 및 JU-287-275-31-GAL4-R (서열번호 141)을 사용하여, 비트들 GAL4-1 및 GAL4-2 (표 4 참조)을 함께 스티칭(stitching)함으로써 형성되었다.

[0253]

[0254] 집합 벡터는 형질변환되어 화학 제품 TOP10 *Escherichia coli* 모세포 (Invitrogen, Carlsbad, CA) 성분으로 되었다. 숙주 세포 형질변환체는 100 ug/mL 카르베니실린 및 40 ug/mL X-gal을 포함하는 Luria Bertoni (LB) 아가에서 선택되었다. 단일한 백색 콜로니는 LB 아가로부터, 5 mL의 LB 액체 배지 및 카르베니실린을 포함하는 배

양 튜브로 옮겨진 후, 상기 배양물은 회전 교반기에서 250 rpm에서 37℃에서 밤새도록 배양되었다. 플라즈미드 DNA는 추출된 후, 올바른 방향에서 올바른 서열을 포함하는 클론을 식별하기 위하여 배열되었다. 상기 세포들은 초저온 냉동 용기(cryo-vial)에서 400 uL 멸균 50% 글리세롤 및 600 uL 액체 배양물로 이루어진 1 mL 원액 부분 표본(aliquot)에서 -80℃로 저장되었다.

[0255] 실시예 7

[0256] 본 실시예는, 집합 벡터 및/또는 MULE를 사용하여 폴리뉴클레오타이드 성분들을 집합된(assembled) 폴리뉴클레오타이드로 집합하는 방법을 제시한다.

[0257] 집합 벡터 (표 7 참조)는 하나의 튜브 (333 fmole의 각각의 RABit)에 함께 옮겨진 후, LguI 제한효소(Fermentas, Glen Burnie, MD)를 사용하여 분해되었다. 제한효소는 컬럼 원심분리에 의해 제거되거나 65℃에서 20분 동안 열 불활성화되었다. MULE들 또는 집합된(assembled) 폴리뉴클레오타이드들을 포함하는 집합 반응(assembly reaction)에 대해, 333 fmole의 각각의 MULE 또는 집합된(assembled) 폴리뉴클레오타이드 (표 7 참조)는 함께 하나의 튜브에 옮겨지거나, 분해된 집합 벡터들에 첨가되었다. 샘플은 3개의 30 uL 반응물로 나뉜 후; 물, 버퍼, dNTP 및 DNA 폴리머라제가 각각의 반응물 혼합물에 첨가되고, 첫번째 단계의 PCR 증폭이 개시되었다. 샘플을 얼음 위에 옮긴 후, 0.5 uM의 각각의 말단의 프라이머 (표 7)가 반응물 혼합물에 첨가된 후 및 두 번째 단계의 PCR 증폭이 실시되었다. 3개의 PCR 반응 혼합물이 하나의 튜브에서 결합된 후, 반응물 혼합물은 겔 전기영동에 의해 분해되고, PCR 생성물은 겔 정제되었다.

표 7

표 7 – 집합된 폴리뉴클레오타이드의 집합체에 대한 말단 프라이머

집합체	결합될 집합 벡터(표 6 참조) 또는 MULE(표 5 참조)*	집합될 폴리뉴클레오타이드 크기(kb) (서열)	말단 프라이머 1	말단 프라이머 2
1	30_22_53_60	4.3	S000 (SEQ ID NO: 45)	S009 (SEQ ID NO: 54)
2	30_22_53	3.1	S000 (SEQ ID NO: 45)	S007 (SEQ ID NO: 52)
3	22_53_60	3.7	S002 (SEQ ID NO: 47)	S009 (SEQ ID NO: 54)
4	30_22	2.5	S000 (SEQ ID NO: 45)	S005 (SEQ ID NO: 50)
5	22_53	2.5	S002 (SEQ ID NO: 47)	S007 (SEQ ID NO: 52)
6	53_60	1.8	S004 (SEQ ID NO: 49)	S009 (SEQ ID NO: 54)
7	30_22_53_60_64_35_12_62_29	7.7 (SEQ ID NO: 222)	S000 (SEQ ID NO: 45)	S019 (SEQ ID NO: 64)
8	30_22_53_60_64_35_5_62_29	7.1	S000 (SEQ ID NO: 45)	S019 (SEQ ID NO: 64)
9	30_22_53_2_64_35_5_62_29	7.1	S000 (SEQ ID NO: 45)	S019 (SEQ ID NO: 64)
10	60_64_35_5_62_29	4.1	S006 (SEQ ID NO: 51)	S019 (SEQ ID NO: 64)
11	2_64_35_5_62_29	4.1	S006 (SEQ ID NO: 51)	S019 (SEQ ID NO: 64)
Phase I-A	270_22_254_229_244_253	8.1 (SEQ ID NO: 111)	S000 (SEQ ID NO: 45)	S013 (SEQ ID NO: 58)
Phase I-B	228_255_271	3.0 (SEQ ID NO: 112)	S013 (SEQ ID NO: 58)	S019 (SEQ ID NO: 64)
Phase II complete	187_262_250_229_252_268_265_263	9.7 (SEQ ID NO:113)	S000 (SEQ ID NO: 45)	S019 (SEQ ID NO: 64)

[0258]

Phase III-A	186_261_241_229	4.4 (SEQ ID NO: 114)	S000 (SEQ ID NO: 45)	S008 (SEQ ID NO: 53)
Phase III-B	251_249_237_269_185	4.3 (SEQ ID NO: 115)	S009 (SEQ ID NO: 54)	S018 (SEQ ID NO: 63)
Phase I marker recycling	270_URA3blaster_44_FS-a_tHMG1-a	6.3 (SEQ ID NO: 116)	S000 (SEQ ID NO: 45)	S019 (SEQ ID NO: 64)
Phase II marker recycling	187_ URA3blaster_44_FS-a_ERG12	6.2 (SEQ ID NO: 117)	S000 (SEQ ID NO: 45)	S019 (SEQ ID NO: 64)
Phase III marker recycling	186_URA3blaster_44_FS-a_ERG19	6.0 (SEQ ID NO: 118)	S000 (SEQ ID NO: 45)	S019 (SEQ ID NO: 64)
STE5 knockout	567_556_P_{TDH3}-a_tHMG1-b_563	5.2 (SEQ ID NO: 119)	S000 (SEQ ID NO: 45)	S019 (SEQ ID NO: 64)
IME1 knockout	IME1US_LEU2_P_{TDH3}-b_FS-b_IME1DS	5.4 (SEQ ID NO: 120)	S000 (SEQ ID NO: 45)	S019 (SEQ ID NO: 64)

PCR 증폭의 첫번째 단계는 하기와 같이 실시되었다 : 98℃에서 2분 동안 변성하는 사이클 1회; PCR 생성물 1 킬로베이스 당 98℃에서 30초 동안 변성 및 72℃에서 30초 동안 어닐링/신장하는 사이클 5회. PCR 증폭의 두번째 단계는 하기와 같이 실시되었다 : 98℃에서 2분 동안 변성하는 사이클 1회; PCR 생성물 1 킬로베이스 당 98℃에서 12초 동안 변성 및 72℃에서 20 내지 25초 동안 어닐링/신장하는 단계 35회; 72℃에서 7분 동안 최종 신장하는 사이클 1회; 및 4℃에서 최종 유지. 어닐링 온도가 72℃가 아닐 때(예를 들어, 54℃ 또는 65℃일 때), PCR 증폭의 첫번째 단계에서, 1분간의 어닐링 단계 후, PCR 생성물 1 킬로베이스 당 72℃에서 30초 동안의 신장 단계가 사용되었고, PCR 증폭의 두번째 단계에서, 15초간의 어닐링 단계 후, PCR 생성물 1 킬로베이스 당 72℃에서 20초 동안의 신장 단계가 사용되었다.

* 집합 벡터는 숫자로 표기되었고, MULE는 이름으로 표기되었다.

[0259]

[0260] 도 5 및 6에서 나타난 바와 같이, 2 내지 9 폴리뉴클레오타이드 성분들은 최대 7.7 kb 길이의 집합된(assembled) 폴리뉴클레오타이드들로 정확하게 집합되었다.

[0261] 실시예 8

[0262] 본 실시예는 본원에 제시된 방법에 의해 집합된(assembled) 폴리뉴클레오타이드를 사용하여 유전적으로 변형된 호스트 미생물을 형성하는 방법을 제시한다.

[0263] I-A 형태(Phase) 및 I-B 형태(Phase)의 집합된(assembled) 폴리뉴클레오타이드들은 (표 7 참조) TOPO Zero Blunt II 클로닝 벡터 (Invitrogen, Carlsbad, CA)로 복제되어, 각각, 플라즈미드들 TOPO-I-A 형태(Phase) 및 TOPO-I-B 형태(Phase)를 형성하였다. 50 μg/ml 카나마이신을 포함하는 LB 아가에서 성장한 TOP10 세포 (Invitrogen, Carlsbad, CA)에서 구성물이 증식되었다. 각각의 플라즈미드는 NotI 제한 엔도뉴클레아제를 사용하여 완전히 분해되었고, I-A 형태(Phase) 및 I-B 형태(Phase)의 삽입물들은 겔 정제 키트 (Qiagen, Valencia, CA)를 사용하여 겔 추출되었고, 동일한 몰비율의 정제된 DNA 절편들은 T4 DNA 연결효소(ligase) (New England Biolabs, Ipswich, MA)를 사용하여 연결되어, 완전한 I 형태(Phase I complete)의 집합된(assembled) 폴리뉴클레오타이드를 형성하였다. 상기 완전한 I 형태(Phase I complete)의 집합된(assembled) 폴리뉴클레오타이드는 TOPO Zero Blunt II 클로닝 벡터 (Invitrogen, Carlsbad, CA)로 복제되어, 플라즈미드 TOPO-I 형태(Phase)를 형성하였다. 50 μg/ml 카나마이신을 포함하는 LB 아가에서 성장된 TOP10 세포 (Invitrogen, Carlsbad, CA)에서 구성물이 증식되었다.

[0264] 완전한 II 형태(Phase II complete)의 집합된(assembled) 폴리뉴클레오타이드 (표 7 참조)는 TOPO Zero Blunt II 클로닝 벡터 (Invitrogen, Carlsbad, CA)로 복제되어, 플라즈미드 TOPO-II 형태(Phase)를 형성하였다. 50

μg/ml 카나마이신을 포함하는 LB 아가에서 성장된 TOP10 세포 (Invitrogen, Carlsbad, CA)에서 구성물이 증식되었다.

[0265] III-A 형태(Phase) 및 III-B 형태(Phase)의 집합된(assembled) 폴리뉴클레오타이드들은 (표 7 참조) TOPO Zero Blunt II 클로닝 벡터 (Invitrogen, Carlsbad, CA)로 복제되어, 플라즈미드들 TOPO-III-A 형태(Phase) 및 TOPO-III-B 형태(Phase)를 각각 형성하였다. 50 μg/ml 카나마이신을 포함하는 LB 아가에서 성장된 TOP10 세포 (Invitrogen, Carlsbad, CA)에서 구성물이 증식되었다. 각각의 플라즈미드는 BamHI 및 SbfI 제한 엔도뉴클레아제를 사용하여 완전히 분해되었고, III-A 형태(Phase) 및 III-B 형태(Phase)의 삽입물들은 겔 정제 키트 (Qiagen, Valencia, CA)를 사용하여 겔 추출되었으며 동일한 몰비율의 DNA 절편들은 T4 DNA 연결효소(ligase) (New England Biolabs, Ipswich, MA)를 사용하여 연결되어, 완전한 III 형태(Phase III complete)의 집합된(assembled) 폴리뉴클레오타이드를 형성하였다. 상기 완전한 III 형태(Phase III complete)의 집합된(assembled) 폴리뉴클레오타이드는 TOPO Zero Blunt II 클로닝 벡터 (Invitrogen, Carlsbad, CA)로 복제되어, 플라즈미드 TOPO-III 형태(Phase)를 형성하였다. 50 μg/ml 카나마이신을 포함하는 LB 아가에서 성장된 TOP10 세포 (Invitrogen, Carlsbad, CA)에서 구성물이 증식되었다.

[0266] 효모 세포 형질변환에 있어서, 25 ml의 Yeast Extract Peptone Dextrose (YPD) 배지는 단일 콜로니의 출발 호스트 종(strain)로 접종되었다. 상기 배양물은 200 rpm의 회전 교반기에서 30℃에서 밤새 성장되었다. 배양물의 OD600이 측정된 후, 배양물은 0.15의 OD600가 될 때까지 50 ml의 YPD 배지를 접종하는데 사용되었다. 새롭게 접종된 배양물은 0.7 내지 0.9의 OD600가 될 때까지 200 rpm의 회전 교반기에서 30℃에서 성장되었고, 상기 시점에서 세포들은 1 μg의 DNA를 사용하여 형질변환되었다. 숙주 세포 형질변환체를 식별하기 위한 선택적 성분을 포함하는 아가에 세포들을 옮기기 전에, 4시간 동안 YPD 배지에서 세포들이 회복되도록 하였다.

[0267] 100 ug/mL 카르바마이실린(carbamicillin) 및 50 ug/mL 카나마이신을 포함하는 5 mL의 YPD 배지에서, 활성있는 건식 PE-2 효모 (Santelisa Vale, Sertaozinho, Brazil에서 1994년에 분리됨)를 재분산시킴으로써, 시작자(starter) 호스트 종(strain) Y1198이 형성되었다. 배양물은 200 rpm의 회전 교반기에서 30℃에서 밤새 배양되었다. 그 후, 부분표본(aliquot)의 10 uL의 배양물은 YPD 플레이트 위에 옮겨진 후 건조되도록 하였다. 단일 콜로니에 대하여 상기 세포에 성공적으로 줄무늬를 넣은 후 30℃에서 2일 동안 배양되었다. 12개의 단일 콜로니를 골라, 새로운 YPD 플레이트에서 조금 떼어내어(조금 떼어내어(patched out)) 30℃에서 밤새 성장하도록 하였다. 콜로니의 종(strain) 식별은 제조업자의 설명서에 따라, Bio-Rad CHEF 게놈 DNA Plug Kit (Bio-Rad, Hercules, CA)를 사용하여 Bio-Rad CHEF DR II 시스템 (Bio-Rad, Hercules, CA)에서 염색체 크기를 분석함으로써 확인하였다. 하나의 콜로니를 골라 종(strain) Y1198로 저장하였다.

[0268] 종(strain) Y1661, Y1662, Y1663 및 Y1664는 종(strain) 홀배수체(haploid)를 제공함으로써 종(strain) Y1198로부터 형성되었다. 종(strain) Y1198는 롤러 드럼(roller drum) 내 유리 튜브에서 30℃에서 5 mL의 YPD 배지에서 밤새 성장되었다. OD600가 측정된 후 세포들은, 2% 포타슘 아세테이트를 포함하는 5 mL의 YP 배지에서 0.2의 OD600가 되도록 희석되었다. 배양물은 롤러 드럼(roller drum) 내 유리 튜브에서 30℃에서 밤새 성장되었다. OD600을 다시 측정한 후 4 OD600*mL의 세포들은 2분 동안 5,000g에서 원심분리함으로써 수집되었다. 세포 펠렛은 멸균수로 한번 수세된 후 0.02% 라피노스를 포함하는 3 mL의 2% 포타슘 아세테이트에서 재분산되었다. 상기 세포는 롤러 드럼(roller drum) 내 유리 튜브에서 30℃에서 3일 동안 성장되었다. 포자 형성은 현미경 검사법으로 확인되었다. 부분표본(aliquot)의 33 uL의 배양물을 1.5 mL의 마이크로퓨지 튜브(microfuge tube)로 옮긴 후 2분 동안 14,000rpm에서 원심분리하였다. 세포 펠렛은 2 uL의 10 mg/mL Zymolyase 100T (MP Biomedicals, Solon, OH)를 포함하는 50 uL의 멸균수에 재분산된 후, 세포는 30℃ 수조에서 10분 동안 배양되었다. 상기 튜브는 얼음으로 옮겨진 후 150 uL의 얼음물이 첨가되었다. 부분표본(aliquot)의 10 uL의 상기 혼합물은 12 mL YPD 플레이트에 첨가된 후 테트라드(tetrad)는 Singer MSM 300 해부 현미경(dissection microscope) (Singer, Somerset, UK)에서 해부되었다. YPD 플레이트는 3일 동안 30℃에서 배양된 후, 포자(spore)를 새로운 YPD 플레이트로 조금 떼어낸(patched out) 후 30℃에서 밤새 성장시켰다. 8개의 4개-포자(spore) 테트라드(tetrad)로부터 각각의 포자의 교배형이 콜로니 PCR에 의해 분석되었다. 2개의 MATA 및 2개의 MATalpha 포자(spore)를 갖는 단일한 4개 포자(spore) 테트라드(tetrad)를 골라, 종(strain) Y1661 (MATA), Y1662 (MATA), Y1663 (MATalpha) 및 Y1664 (MATalpha)로 저장하였다.

[0269] PmeI 제한 엔도뉴클레아제를 사용하여 완전히 분해된 플라즈미드 TOPO-I 형태(Phase)로 종(strain) Y1664를 형질변환함으로써 호스트 종(strain) 1515이 형성되었다. 숙주 세포 형질변환체는 300 ug/mL 하이그로마이신(hygromycin) B를 포함하는 YPD 배지에서 선택되었다.

[0270] PmeI 제한 엔도뉴클레아제를 사용하여 완전히 분해된 플라즈미드 TOPO-II 형태(Phase)로 종(strain) Y1515를 형질변환함으로 써호스트 종(strain) 1762가 형성되었다. 숙주 세포 형질변환체는 100 ug/mL 노르세오트리신(nourseothricin)을 포함하는 YPD 배지에서 선택되었다.

[0271] PmeI 제한 엔도뉴클레아제를 사용하여 완전히 분해된 발현 플라즈미드 pAM404 및 플라즈미드 TOPO-III 형태(Phase)로 2개의 단계에서 종(strain) Y1762를 형질변환함으로써 호스트 종(strain) 1770이 형성되었다. 발현 플라즈미드 pAM404는, β-파네센 신타아제(β-farnesene synthase)를 코딩하는 뉴클레오타이드 서열을 pRS425-Gal1 벡터에 삽입함으로써 형성된, 플라즈미드 pAM353로부터 유래되었다 (Mumberg et. al. (1994) *Nucl. Acids. Res.* 22(25): 5767-5768). *Saccharomyces cerevisiae* (서열번호 121)에서의 발현에 최적화을 위하여 코돈-최적화된 *Artemisia annua* (유전자 은행의 등록번호 AY835398)의 코딩 서열을 주형(template)으로 사용하여, 뉴클레오타이드 서열 삽입물이 종합적으로 형성되었다. 종합적으로 형성된 뉴클레오타이드 서열은 5' *BamHI* 및 3' *XhoI* 제한 부위의 측면에 배치되어, 따라서 표준 pUC 또는 pACYC 오리진 벡터(origin vector)와 같은 클로닝 벡터의 호환가능한 제한 부위로 클론될 수 있다. *BamHI* 및 *XhoI* 제한효소를 사용하여 DNA 합성 구조를 완전히 분해함으로써, 종합적으로 형성된 뉴클레오타이드 서열이 분리되었다. 상기 반응 혼합물은 겔 전기영동에 의해 분해되어, β-파네센 신타아제(β-farnesene synthase) 코딩 서열을 포함하는 대략 1.7 kb DNA 절편이 겔 추출된 후 분리된 DNA 절편이 pRS425-Gal1 벡터의 *BamHI XhoI* 제한 부위로 연결되어 발현 플라즈미드 pAM353을 형성하였다. β-파네센 신타아제(β-farnesene synthase)를 코딩하는 뉴클레오타이드 서열은 프라이머들 GW-52-84 pAM326 BamHI (서열번호 188) 및 GW-52-84 pAM326 NheI (서열번호 189)을 이용하여 pAM353으로부터 PCR 증폭되었다. 결과물 PCR 생성물은 *BamHI* 및 *NheI* 제한효소를 사용하여 완전히 분해되었고, 상기 반응 혼합물은 겔 전기영동에 의해 분해되었고, β-파네센 신타아제(β-farnesene synthase) 코딩 서열을 포함하는 대략 1.7 kb DNA 절편은 겔 추출되고, 분리된 DNA 절편은 벡터 pAM178 (서열번호 122)의 벡터의 *BamHI NheI* 제한 부위로 연결되어, 발현 플라즈미드 pAM404를 형성하였다. pAM404를 갖는 숙주 세포 형질변환체는 pAM404는 메티오닌 및 루신이 결핍된 완전 합성 배지 (CSM)에서 선택되었다. pAM404 및 완전한 III 형태(Phase III complete)의 집합된(assembled) 폴리뉴클레오타이드를 갖는 숙주 세포 형질변환체는 메티오닌 및 루신이 결핍되고 200 ug/mL G418를 포함하는 CSM에서 선택되었다.

[0272] 호스트 종(strain) 1793은 URA3 녹아웃(knockout) 구조 (서열번호 123)을 갖는 종(strain) Y1770을 형질변환함으로써 형성되었다. 녹아웃(knockout) 구조는, DNA 절편들 URA3US (PCR 프라이머들 KMH33-276-21-URA3 5'.fwd (서열번호 147) 및 KMH34-276-21-URA3 5'.rev (서열번호 148)을 사용하여 *Saccharomyces cerevisiae* 종(strain) CEN.PK2 게놈 DNA로부터 형성됨)과 URA3DS (PCR 프라이머들 KMH35-276-21-URA3 3'.fwd (서열번호 149) 및 KMH36-276-21-URA3 3'.rev (서열번호 150)를 사용한 후; PCR 프라이머들 KMH33-276-21-URA3 5'.fwd 및 KMH36-276-21-URA3 3 '.rev를 사용하여 2개의 DNA 절편들을 서로 스티칭(stitching)하여 *Saccharomyces cerevisiae* 종(strain) CEN.PK2 게놈 DNA로부터 형성됨)를 첫번째로 형성함으로써 형성되었다. 숙주 세포 형질변환체는 5-FOA를 포함하는 YPD 배지로부터 선택되었다.

[0273] 호스트 종(strain) YAAA는 I 형태(Phase)의 표지 리사이클링(marker recycling) 집합된(assembled) 폴리뉴클레오타이드 (표 7 참조)를 사용하여 종(strain) Y1793을 형질변환함으로써 형성되었다. 숙주 세포 형질변환체는 메티오닌 및 우라실이 결핍된 CSM에서 선택되었다. 회전 교반기에서 200 rpm에서 30℃에서 YPD 배지에서 밤새 세포를 성장시킨 후 5-FOA를 포함하는 아가로 세포를 옮김으로써 URA3 표지가 절단되었다. 표지 절단은 콜로니 PCR에 의해 확인되었다.

[0274] 호스트 종(strain) YBBB는 II 형태(Phase)의 표지 리사이클링 집합된(assembled) 폴리뉴클레오타이드를 사용하여 종(strain) YAAA를 형질변환함으로써 형성되었다(표 7 참조). 숙주 세포 형질변환체는 메티오닌 및 우라실이 결핍된 CSM에서 선택되었다. 회전 교반기에서 200 rpm에서 30℃에서 YPD 배지에서 밤새 세포를 성장시킨 후 5-FOA를 포함하는 아가로 세포를 옮김으로써 URA3 표지가 절단되었다. 표지 절단은 콜로니 PCR에 의해 확인되었다.

[0275] 호스트 종(strain) Y1912는 III 형태(Phase)의 표지 리사이클링 집합된(assembled) 폴리뉴클레오타이드를 사용하여 종(strain) YBBB를 형질변환함으로써 형성되었다(표 7 참조). 숙주 세포 형질변환체는 메티오닌 및 우라실이 결핍된 CSM에서 선택되었다. 회전 교반기에서 200 rpm에서 30℃에서 YPD 배지에서 밤새 세포를 성장시킨 후 5-FOA를 포함하는 아가로 세포를 옮김으로써 URA3 표지가 절단되었다. 표지 절단은 콜로니 PCR에 의해 확인되었다.

[0276] 호스트 종(strain) Y1913은 STE5 녹아웃(knockout) 집합된(assembled) 폴리뉴클레오타이드를 사용하여 종

(strain) Y1912를 형질변환함으로써 형성되었다(표 7 참조). 숙주 세포 형질변환체는 메티오닌 및 우라실이 결핍된 CSM에서 선택되었다.

[0277] pAM404로부터 종(strain)을 고친 후(curing) IME1 녹아웃(knockout) 집합된(assembled) 폴리뉴클레오타이드를 사용하여 상기 결과물 종(strain)을 형질변환함으로써 호스트 종(strain) Y1915가 형성되었다 (표 7 참조). 종(strain) Y1913은 회전 교반기에서 200 rpm에서 30℃에서 비-선택적 YPD 배지에서 증식되었다. 대략 100개의 세포가 YPD 고체 배지로 옮겨진 후 30℃에서 3일 동안 성장되도록 한 후, 메티오닌 및 루신이 결핍된 CSM 플레이트에서 복제-평판(replica-plated)되었고, 여기에서 세포들은 다시 30℃에서 3일 동안 성장되었다. 고쳐진 세포(cured cell)는, 루신을 포함하는 최소 배지에서 성장하는 능력과 루신이 결핍된 배지에서 성장하는 무능력(inability)에 의해 식별되었다. 콜로니와 같은 싱글(single)이 선택된 후 IME1 녹아웃(knockout) 집합된(assembled) 폴리뉴클레오타이드를 사용하여 형질변환되었다. 숙주 세포 형질변환체는 메티오닌 및 우라실이 결핍된 CSM에서 선택되었다.

[0278] 실시예 9

[0279] 본 실시예는, 본원에 제시된 창조적인 방법에 의해, 프로모터와 단백질 코딩서열을 코딩하는 폴리뉴클레오타이드 성분들을 집합된(assembled) 폴리뉴클레오타이드에 결합하는데 사용되는 어닐링될 수 있는 링커 서열들을 선택하는 방법을 제시한다.

[0280] 2개의 서로 다른 후보인 어닐링될 수 있는 링커 서열들, 어닐링될 수 있는 링커 서열 RYSE 15 (R15; 서열번호 15) 및 어닐링될 수 있는 링커 서열 RYSE 7 (R7; 서열번호 7) 이후에 프로모터들을 코딩하는 MULE 뿐만 아니라, 상기 2개의 어닐링될 수 있는 링커 서열들에 선행하여 GFP를 코딩하는 MULE가, 표 8에 제시된 바와 같이 PCR 증폭되었다.

표 8

[0281]

표 8 - 프로모터 및, 어닐링할 수 있는 링커 서열들 RYSE 15(R15) 또는 어닐링할 수 있는 링커 서열 RYSE 7(R7)을 갖는 GFP를 코딩하는 증폭된 MULE

MULE	TYPE *	프라이머들	크기 (bp)	주형
pGAL1-R15	P	Plan X19(SEQ ID NO:236) Plan X20(SEQ ID NO:237)	698	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
pTDH3-R15	P	Plan X47(SEQ ID NO:238) Plan X48(SEQ ID NO:239)	613	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
pCYC1-R15	P	Plan X11(SEQ ID NO:240) Plan X12(SEQ ID NO:241)	645	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
pGAL1-R7	P	Plan X19(SEQ ID NO:236) Plan X64(SEQ ID NO:242)	692	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
pTDH3-R7	P	Plan X47(SEQ ID NO:238) Plan X71(SEQ ID NO:243)	607	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
pCYC1-R7	P	Plan X11(SEQ ID NO:240) Plan X78(SEQ ID NO:244)	639	*Saccharomyces cerevisiae* 종 CEN.PK2 게놈 DNA
R7-GFP	GsT	Plan X96(SEQ ID NO:247) Plan X88(SEQ ID NO:245)	1378	RABit 634 플라즈미드 DNA **
A-GFP	GsT	Plan X89(SEQ ID NO:246) Plan X88(SEQ ID NO:245)	1385	RABit 634 플라즈미드 DNA **

PCR 반응물은 67 uL ddH2O, 20 uL 5x HF 버퍼, 2 uL의 각각의 프라이머 (10uM), 1 uL dNTP 믹스 (200 uM), 1uL Phusion DNA 폴리머라제(New England Biolabs, Ipswich, MA), 및 9 uL Y002 게놈 DNA 또는 RABit 634 플라즈미드 DNA를 포함하였다.

PCR 증폭은 하기와 같이 실시되었다 : 98℃에서 2분 동안 변성하는 사이클 1회; 98℃에서 15초 동안 변성, 각각의 사이클마다 1℃씩 낮추면서 61℃에서 30초 동안 어닐링 및 72℃에서 1분 동안 신장하는 사이클 9회; 98℃에서 15초 동안 변성, 52℃에서 30초 동안 어닐링, 및 72℃에서 1분 동안 신장하는 단계 26회; 72℃에서 7분 동안 최종 신장하는 사이클 1회; 및 4℃에서 최종 유지.

* G=유전자; s=종결 코돈; T=종결자; P=프로모터.

** RABit 634는 녹색 형광 단백질 (GFP)의 코딩 서열를 뒤잇는 *Saccharomyces cerevisiae*의 ADH1 유전자의 종결자를 포함한다.

[0282] PCR 반응물은 겔 전기영동에 의해 분해되었고, MULE은 겔 정제되었고, 상기 정제된 MULE은 테스트용 집합된 (assembled) 폴리뉴클레오타이드를 결합하는데 사용되었다. 상기 말단에, 집합될 (표 9 참조) MULE 및 집합 벡터 (표 6 참조)는 튜브(333 fmole의 각각의 집합 벡터, 667 fmole의 각각의 MULE)에 함께 넣은 후 LguI 제한효소 (Fermentas, Glen Burnie, MD)를 사용하여 분해되었다. 제한효소는 65℃에서 20분 동안 열 불활성화되었다. 샘플은 3개의 30 uL의 반응물로 나뉘어지고; 물, 버퍼, dNTP 및 DNA 폴리머라제는 각각의 반응물 혼합물에 첨가되고 PCR 증폭의 첫번째 단계가 개시되었다. 그 후 말단의 프라이머는 반응물 혼합물에 첨가되고 PCR 증폭의 두번째 단계가 실시되었다 (표 9 참조).

[0283] 3개의 PCR 반응물 혼합물이 하나의 튜브에서 결합되었고(combined), 상기 반응물 혼합물은 겔 전기영동에 의해 분해되었고, PCR 생성물은 겔 정제되었다.

표 9

[0284]

표 9 - 테스트용 집합된 폴리뉴클레오타이드의 집합체에 대한 말단 프라이머

집합체	**결합될 MULE(표 8 참조) 및 집합 벡터(표 6 참조)***	**집합된 폴리뉴클레오타이드 크기 (kb)**	**말단 프라이머 1**	**말단 프라이머 2**
1	97_555_pGAL1-A_A-GFP_24	4.7	S000 (SEQ ID NO : 45)	S019 (SEQ ID NO : 64)
2	97_555_pTDH3-A_A-GFP_24	4.6		
3	97_555_pCYC1-A_A-GFP_24	4.7		
7	97_555_pGAL1-R7_R7-GFP_24	4.7		
8	97_555_pTDH3-R7_R7-GFP_24	4.6		
9	97_555_pCYC1-R7_R7-GFP_24	4.6		

PCR 반응물은 하기의 것들을 포함한다: 11 uL ddH2O, 20 uL 5x HF 버퍼, 5 uL의 각각의 말단 프라이머(1 uM), 2 uL dNTP 믹스(200 uM), 1.8 uL Phusion DNA 폴리머라아제 및 30 uL MULE 또는 LguI 분해된 결합 벡터.

PCR 증폭의 첫번째 단계는 하기와 같이 실시되었다: 98℃에서 2분 동안 변성하는 사이클 1회; 98℃에서 30초 동안 변성, 60℃에서 30초 동안 어닐링 및 72℃에서 2.5분 동안 신장하는 사이클 5회; 후에 2개의 말단 프라이머들을 첨가하고 4℃에서 유지한다. PCR 증폭의 두번째 단계는 하기와 같이 실시되었다 : 98℃에서 2분 동안 변성하는 사이클 1회; 98℃에서 12초 동안 변성, 60℃에서 30초 동안 어닐링 및 72℃에서 2.5분 동안 신장하는 단계 35회; 72℃에서 7분 동안 최종 신장하는 사이클 1회; 및 4℃에서 최종 유지.

* 결합 벡터는 숫자로 표기되었고, MULE는 이름으로 표기되었다.

[0285] 테스트용 집합된(assembled) 폴리뉴클레오타이드들은, URA3가 부족하고 GAL80 위치에서 HaeI가 삭제된 *Saccharomyces cerevisiae* 호스트 종(strain)을 형질변환하는데 사용되었다. 숙주 세포 형질변환체는 우라실이 결핍된 CSM에서 선택되었고, 집합된(assembled) 폴리뉴클레오타이드의 정확한 게놈 통합(integration)은 콜로니 PCR에 의해 확인되었다. 각각의 형질변환으로부터의 2개의 검증된 콜로니가 선택되어 2% 수크로오스를 포함하는 360 uL Bird Seed Medium (BSM)에 첨가된 후 배양물이 회전 교반기에서 999 rpm에서 30℃에서 48시간 동안 배양되었다. 부분표본(aliquot)의 14.4 uL를 취하여 각각의 웰에 넣고 96-웰 블록 플레이트에서 4% 수크로오스를

포함하는 1.1 mL BSM으로 이동시킨 후 회전 교반기에서 999 rpm에서 30℃에서 6시간 동안 배양하고, 각각의 배양물의 100 uL 지점에서 이들을 GFP 발현의 분석을 위하여 투명한 바닥 96-웰 플레이트의 웰로 옮겼다. 각각의 웰에서의 GFP 발현은 M5 Plate 리더 분광 광도계 (Molecular Devices, Sunnyvale, CA)에서 485 nm 여기(excitation)된 후 515 nm 방출을 측정함으로써 분석되었다. 각각의 배양물에 대하여 세포 배양물 성장을 OD600 리딩(reading)으로 나눔으로써 측정된 GFP 농도가 표준화되었다.

[0286] 표 10에서 나타난 바와 같이, 어닐링될 수 있는 링커 서열 RYSE 7에 비하여, 어닐링될 수 있는 링커 서열 RYSE 15는 테스트용 집합된(assembled) 폴리뉴클레오타이드에서 GFP 리포터 유전자(reporter gene)의 상승된 GAL1, TDH3 및 CYC1 프로모터 유도 발현(driven expression)을 가능하게 하였다.

표 10

[0287]

표 10 - 프로모터와 GFP 리포터 사이의 어닐링할 수 있는 링커 서열 RYSE 15 (R15) 또는 어닐링할 수 있는 링커 서열 RYSE 7 (R7)을 포함하는 테스트용 집합된 폴리뉴클레오타이드를 포함하는 숙주 세포에서의 GFP 발현

테스트용 집합된 폴리뉴클레오타이드에서 프로모터와 GFP 리포터 유전자 사이에 위치한 어닐링할 수 있는 링커 서열	평균 GFP 발현% (3개의 균일한 대조군 구조* 중 하나를 포함하는 숙주 세포를 사용하여 얻은 평균 GFP 발현% 대비; 2개의 독립적인 숙주 세포 분리물)			
	GAL1 프로모터	TDH3 프로모터	CYC1 프로모터	상기 3개의 프로모터의 평균
R15	79.34	91.42	81.92	84.22
R7	27.43	54.68	46.31	42.81

* 균일한(seamless) 대조군 구조는, 프로모터 서열이 GFP 리포터 유전자에 균일하게 연결된 점(예를 들어, 어닐링할 수 있는 링커의 사이에 끼어들지 않음)을 제외하고 테스트용 집합된 폴리뉴클레오타이드와 동일한 구조를 가졌다.

[0288] 실시예 10

[0289] 본 실시예는 폴리뉴클레오타이드의 높은 생산량의(high-throughput) 조합적 집합 방법 및 조합적으로 결합된(combined) 폴리뉴클레오타이드를 포함하는 높은 생산량의(high-throughput) 숙주 세포의 형성 방법을 제시한다.

[0290] 본 실시예에서 사용된 폴리뉴클레오타이드 성분들 및, 상기 폴리뉴클레오타이드 성분들로부터 형성된 예상되는 집합된(assembled) 및 결합된(combined) 폴리뉴클레오타이드는, 도 12a에서 도식적으로 나타냈다. 폴리뉴클레오타이드 성분들은 업스트림 및 다운스트림 염색체의 타겟팅(targeting) 서열 (US 및 DS), 6개의 서로 다른 프로모터 (P), 35개의 서로 다른 단백질 (G), 및 어닐링될 수 있는 링커 서열 쌍 또는 프라이머 결합 절편/어닐링될 수 있는 링커 서열 쌍의 측면에 배치된, URA3 선별표지(selectable marker)의 5' 및 3' 절편 (각각, URA 및 RA3)을 코딩하는 DNA 절편(segment)을 포함하였다.

[0291] LguI 제한 엔도뉴클레아제를 사용하여 RABit 또는 MULE를 분해함으로써 집합 벡터로부터 폴리뉴클레오타이드 성분들이 분리되었다. 이 때문에, 96-웰 플레이트 ("LguI 분해 플레이트(Digestion Plates)")가 하기 표에 나타난 바와 같이 설정되었고 상기 플레이트는 75분 동안 37℃에서 배양되었고, 그 후에 LguI 제한 엔도뉴클레아제는 PCR 기기에서 20분 동안 65℃에서 열 불활성화되었다.

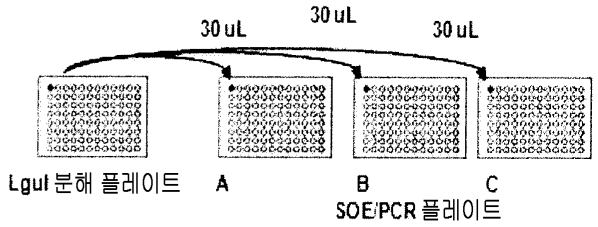
표 11

[0292]

LguI **분해 플레이트**	
성분(웰 당)	**부피(uL)**
667 fMole RABit 또는 MULE	가변적임
10x Tango 버퍼 (Fermentas, Glen Burnie, MD)	10
LguI (Fermentas, Glen Burnie, MD)	2.5
ddH2O	to 100

[0293] 폴리뉴클레오타이드 성분들은 SOE에 의해 집합되었다. 하기 표에 나타난 바와 같이, 각각의 LguI 분해 플레이트에 대하여, 3개의(triplicate) 96-웰 플레이트 ("SOE/PCR 플레이트")가 설정되었고 PCR 기계에서 열화학사이클(thermocycle) 되었다.

표 12

SOE/PCR 플레이트			
성분(웰 당)		부피(uL)	
ddH2O		41	
5x Phusion HF 버퍼 (New England Biolabs, Ipswich, MA)		20	
10 mM dNTP 믹스		2	
Phusion DNA 폴리머라제 (New England Biolabs, Ipswich, MA)		1.8	
집합될 LguI-분해된 RABit 또는 MULE		30 	
		총: 95	
열화학사이클(Thermocycling)조건			
최초 변성		98°C	2 min
7 사이클	변성	98°C	30 sec
	어닐링	67°C	30 sec
	신장	72°C	5 min
유지		4°C	∞

[0294]

[0295] 집합된(assembled) 폴리뉴클레오타이드들은 PCR 증폭되었다. 하기 표에 나타난 바와 같이 각각의 SOE/PCR 플레이트는 추가 성분이 첨가된 후 PCR 기계에서 열화학사이클(thermocycle) 되었다. 대응되는 SOE/PCR 플레이트의 웰은 96-딥 웰 블록(deep well block)으로 넣어진 후, 제조업자가 제시한 프로토콜에 따라(대략 45 uL의 말기(end) 부피) Omega Biotek E-Z 96® Cycle-Pure Kit (Omega Bio-Tek Inc., Norcross, GA)을 사용하여 집합된(assembled) 폴리뉴클레오타이드들이 정제되었다.

표 13

[0296]

SOE/PCR 플레이트			
추가 성분 (웰 당)		부피 (uL)	
말단 프라이머들 S000(SEQ ID NO : 45) 및 S019(SEQ ID NO : 64)의 10 mM 원액		10	
열화학사이클(Thermocycling) 조건			
최초 변성		98℃	2 min
35 사이클	변성	98℃	12 sec
	어닐링	67℃	30 sec
	신장	72℃	4.5 min
최종 신장		72℃	7 min
유지		4℃	∞

[0297] 도 12b는, 예시적인 집합된(assembled) 폴리뉴클레오타이드들 (박스 부분)이 1% 아가로스 겔에서 분해되었음을 보여준다.

[0298] 클로닝을 위한 접착 말단(sticky end)을 형성하기 위하여 LguI 제한 엔도뉴클레아제를 사용하여 정제된 집합된(assembled) 폴리뉴클레오타이드들이 분해되었다. 이를 위하여, 하기 표에 나타난 바와 같이 96-웰 플레이트("플레이트에서의 LguI 집합된 폴리뉴클레오타이드 분해")가 설정되었고 상기 플레이트들은 60분 동안 37℃에서 배양되었고, 그 후 LguI 제한 엔도뉴클레아제는 PCR 기계에서 20 분 동안 65℃에서 열 불활성화되었다. LguI 분해된 집합된(assembled) 폴리뉴클레오타이드들은, 제조업자들이 추천하는 프로토콜에 따라 ZR-96 Zymoclean™ Gel DNA Recovery Kit (Zymo Research Corporation, Orange, CA)를 사용하여 겔 정제되었다.

표 14

[0299]

LguI 집합된 폴리뉴클레오타이드 LguI 분해 플레이트	
성분(웰 당)	**부피(uL)**
정제된 집합된 폴리뉴클레오타이드	43
10x Tango 버퍼	5
LguI	2

[0300] 집합된(assembled) 폴리뉴클레오타이드들은 pUC-19계 벡터 골격으로 연결되었다. 상기 벡터로 어떠한 삽입물도 연결되지 않을 때, pTRC 프로모터 (예를 들어, *Saccharomyces cerevisiae*의 TRC 유전자의 프로모터)는 GluRS의 발현을 유발하고 숙주 세포를 죽인다. 96-웰 플레이트 ("연결 플레이트(Ligation Plates)")는 하기 표에 나타난 바와 같이 설정되었고 상기 플레이트는 15분 동안 24℃에서 배양되 후 밤새 16℃에서 배양되었다. 연결 생성물은, 제조업자가 제안한 프로토콜에 따라, ZR-96 DNA Clean & Concentrator™-5 (Zymo Research Corporation, Orange, CA)를 사용하여 정제되었다.

표 15

[0301]

연결 플레이트	
성분(웰 당)	**부피(uL)**
ddH2O	5
10x T4 DNA Ligase 버퍼	2
벡터 골격	2
정제된 집합된 폴리뉴클레오타이드	10
T4 DNA 리가아제 (NEB, Ipswich, MA)	1

[0302] 연결 생성물은 *E. coli* 성분 세포로 전기천공되었다(electroporate). 하기 표에 나타난 바와 같이, 미리 냉각된 96-웰 전기천공(electroporation) 플레이트가 설정된 후, 전기천공이 실시되었다.

표 16

[0303]

전기천공 플레이트		
성분(웰 당)		**부피(uL)**
정제된 연결 생성물		10
Lucigen 10G 성분 세포(Lucigen Corporation, Middleton, WI)		25
전기천공 설정		
2400V	750Ω	25 uF

[0304] 250 uL의 미리 데워진 SOC를 포함하는 1.1 mL 96-웰 배양 플레이트 ("배양 플레이트(Culture Plate)")가 준비된 후, 각각의 웰로부터 100 uL SOC를 취하여, 전기천공(electroporation) 후에 즉시 전기천공(electroporate)된 세포에 첨가되었다. SOC와 세포들이 혼합되었고, 100 uL의 각각의 혼합물은 배양 플레이트로 다시 옮겨졌다. 배양 플레이트는 Multitron II Incubator Shaker (ATR Biotech, Laurel, MD)에서 1시간 동안 37℃에서 배

양되었다. 세포의 2개의 희석물 (3 ul 및 240 ul)은 50 ug/mL 카나마이신을 포함하는 LB 아가에 옮겨진 후 37℃에서 밤새 배양되었다. 콜로니가 선택된 후, 웰 당 카나마이신을 갖는 1 mL LB 배지를 포함하는 96 딥 웰 플레이트에서 성장되었고, LguI 제한 엔도뉴클레아제를 사용하여 제한 분석(restriction analysis)을 위하여 DNA가 추출되었다. 대략 8 kb 결합된(combined) 폴리뉴클레오타이드를 포함하는 24개의 예시적인 콜로니 중 22개 대한 상기 제한 분석의 결과를 도 12c에 나타냈다.

[0305] 염색체로 통합된 결합된(combined) 폴리뉴클레오타이드들을 포함하는 효모 세포들은 말단 염색체의 타겟팅(targeting) 서열과 집합된(assembled) 폴리뉴클레오타이드들의 선별표지(selectable marker) 절편(segment) 사이의 숙주 세포 매개 상동 재조합(homologous recombination)에 의해 형성되었다. 이를 위하여, 하기 표에 나타낸 바와 같이, 96-웰 PCR 플레이트 ("효모 형질변환 플레이트(Yeast Transformation Plates)")가 준비된 후, PCR 기계에서 열 충격 형질변환(heat shock transformation)이 실시되었다.

표 17

[0306]

효모 형질변환 플레이트	
성분(웰 당)	**부피(uL)**
Miniprep DNA(20 ng/uL)	10
효모 세포 성분 *	40
PEG/SS/LiAc 마스터 믹스 **	200
열 충격	
30℃	30 min
42℃	45 min
24℃(선택적)	30 min
* 100 mL YPD에서 밤새 세포를 성장시키고, 배양물을 희석하고, 밤새 약 0.8의 OD600으로 성장시킨 후, 5분 동안 3,000g에서 배양물을 회전시키고, 1 L ddH2O를 사용하여 세포 펠렛을 수세하고, 1L 100 mM 리튬 아세테이트 (LiAc)로 상기 세포 펠렛을 수세하고, 100 mM LiAc에서 총 부피 18 mL까지 세포 펠렛을 재분산시켜서 제조됨. ** 4개의 PCR 플레이트에 충분한 마스터 믹스는 100 mL 50% PEG, 끓는(적어도 10분 동안 95℃) 4 mL 단일 가닥의 DNA, 15 mL 1M LiAc를 포함한다.	

[0307] 효모 형질변환 플레이트는 2분 동안 2,000g에서 회전되었고, 상층물이 제거된 후 세포 펠렛들은 200 uL ddH2O를 사용하여 3회 수세되었다. 세포 펠렛은, 이전에 제조된, 웰 당 360 uL 차가운 BSM을 포함하는 미리 냉각된 96-웰 배양 플레이트 ("시드 플레이트(Seed Plates)")로부터 취한 100 uL cold Bird Seed Media (BSM)를 사용하여 재분산되었다. 재분산된 세포들은 시드 플레이트로 옮겨진 후 Multitron II Incubator Shaker에서 30℃에서 밤새 성장되었다. 상기 시드 플레이트는 5분 동안 3,000g에서 회전되고, 60 uL의 액체를 제외하고 모두 제거되었고, 세포 펠렛을 재분산하기 위하여 덮개가 덮여진 시드 플레이트는 1,000rpm에서 교반되었다.

[0308] 실시예 11

[0309] 본 실시예는 숙주 세포 매개 상동 재조합(homologous recombination)에 의해 형성된 집합된 폴리뉴클레오타이드들을 포함하는 효모 세포들을 형성하는 방법을 제시한다.

[0310] 본 실시예에 사용된 집합된(assembled) 폴리뉴클레오타이드 및 폴리뉴클레오타이드 성분들 및 집합과 염색체의 통합(integration)에서 얻어진 예상되는 염색체의 위치는, 도 13a에 도식적으로 나타냈다.

[0311] 효모 세포 형질변환은 하기 표에 나타낸 바와 같이 실시되었다. 열 충격(heat shock) 후에, 세포들을 회전시켜 상층물이 제거된 후, 세포들은 400 uL ddH2O에 재분산되었고 숙주 세포 형질변환체는 우라실 결핍 아가로 100 내지 200 uL의 세포 분산물을 옮겨서 선택되었다.

표 18

[0312]

효모 형질변환	
성분	**부피(uL)**
성분 및 집합된 폴리뉴클레오타이드 (각각 300-500 ng)	20

효모 세포 성분 *	세포 펠렛 *
50% PEG 용액	240
1M LiAc pH 8.4-8.9	36
끓는 (95℃에서 5분 동안) 단일 가닥의 DNA (10 mg/mL) (Invitrogen, Carlsbad, CA)	10
ddH2O	54
열 충격	
42℃	40 min
* 밤새 세포를 성장시키고, 배양물을 희석하고, 밤새 30℃에서 25 mL YPD에서 콜로니로부터 0.7 내지 0.9의 OD600으로 세포를 성장시킨 후, 5 내지 40 mL ddH2O로 세포 펠렛을 수세한 후, 1 mL ddH2O로 세포 펠렛을 수세한 후, 1 mL 100 mM 리튬 아세테이트 (LiAc)로 상기 세포 펠렛을 수세하고, 상기 세포를 침전시키기 위하여 30초 동안 마이크로센트리퓨지에서 회전시킨 후, 상층물을 제거시켜서 제조되었다.	

[0313] 집합된(assembled) 폴리뉴클레오타이드들의 성공적인 통합은 cPCR 프라이머 A, B, E 및 F (염색체의 통합(integration) 부위의 5' 접합부) 또는 cPCR 프라이머 C, D, G 및 H (염색체의 통합(integration) 부위의 3' 접합부)를 사용하여 cPCR에 의해 측정되었다 (도 13a). 도 13b에 나타낸 바와 같이, 분석된 모든 8개의 콜로니들은 예상되는 집합된(assembled) 폴리뉴클레오타이드의 양성 염색체의 통합(integration) 이벤트를 나타내는 700 bp PCR 밴드를 생성하였고, 원래의 위치에 생성되었을 950 bp PCR 밴드를 결핍하였다.

[0314] 실시예 12

[0315] 본 실시예는 숙주 세포 매개 상동 재조합(homologous recombination)에 의해 형성된 조합적으로 집합된(assembled) 및 결합된(combined) 폴리뉴클레오타이드들을 포함하는 높은 생산량의(high-throughput)의 효모 세포의 형성 방법을 제시한다.

[0316] 본 실시예에서 사용된 폴리뉴클레오타이드 성분들 및 숙주 세포 매개 상동 재조합(homologous recombination)에 의한 집합 및 결합으로부터 얻은 예상되는 결합된(combined) 폴리뉴클레오타이드들은, 도 14A에 도식적으로 나타냈다. 폴리뉴클레오타이드 성분들은 업스트림 및 다운스트림 염색체의 타겟팅(targeting) 서열 (US 및 DS)을 코딩하는 DNA 절편, 6개의 서로 다른 프로모터 (P), 35개의 서로 다른 단백질 (G) 및 어닐링될 수 있는 링커 서열 쌍 또는 프라이머 결합 절편/어닐링될 수 있는 링커 서열 쌍의 측면에 배치된, URA3 선별표지(selectable marker)의 5' 및 3' 절편 (각각, URA 및 RA3)을 포함하였다.

[0317] LguI 제한 엔도뉴클레아제를 사용하여 RABit 또는 MULE를 분해함으로써, 집합 벡터로부터 폴리뉴클레오타이드 성분들이 분리되었다. 이를 위하여, 96-웰 플레이트 ("LguI 분해 플레이트(LguI Digestion Plate)")는 하기 표에 나타난 바와 같이 준비된 후 상기플레이트는 75분 동안 37℃에서 배양되었고, 그 후에 LguI 제한 엔도뉴클레아제는 PCR 기계에서 20분 동안 65℃에서 열 불활성화되었다.

표 19

[0318]

LguI 분해 플레이트	
성분(웰 당)	**부피(uL)**
667 fMole RABit 또는 MULE	가변적임
10x Tango 버퍼 (Fermentas, Glen Burnie, MD)	5
LguI (Fermentas, Glen Burnie, MD)	2.5
ddH2O	to 50

[0319] 염색체로 통합된(integrated), 조합적으로 집합된(assembled) 및 조합적으로 결합된(combined) 폴리뉴클레오타이드를 포함하는 효모 세포를 형성하기 위하여, 96-웰 PCR 플레이트 ("효모 형질변환 플레이트(Yeast Transformation Plate)")가 준비된 후, 열 충격 형질변환이 하기 표에 나타난 바와 같이 PCR 기계에서 실시되었다.

표 20

[0320]

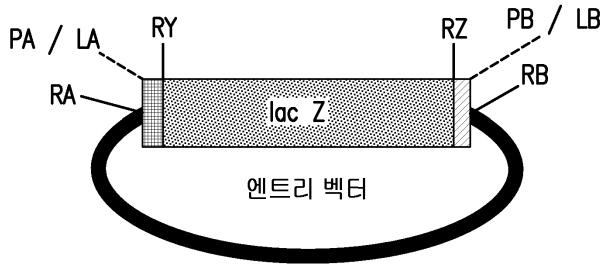
효모 형질변환 플레이트	
성분(웰 당)	**부피(uL)**
폴리뉴클레오타이드 성분	10
효모 세포 성분 *	40
PEG/SS/LiAc 마스터 믹스 **	200
열 충격	
30℃	30 min
42℃	45 min
24℃(선택적)	30 min
*100 mL YPD에서 밤새 세포를 성장시키고, 상기 배양물을 희석시키고 약 0.8의 OD600까지 밤새 성장시킨 후, 5분 동안 3,000g에서 배양물을 회전시킨 후, 1 L ddH2O를 사용하여 세포 펠렛을 수세하고, 1 L 100mM 리튬 아세테이트 (LiAc)를 사용하여 세포 펠렛을 수세한 후, 100 mM LiAc 내 총 18 mL의 부피까지 상기 세포 펠렛을 재분산시켜서 제조됨. **4개의 PCR 플레이트에 대한 마스터 믹스는 contains 100 mL 50% PEG, 4 mL의 끓는 (적어도 10 분 동안의 95℃) 단일 가닥 DNA, 15 mL 1 M LiAc를 포함한다.	

[0321] 효모 형질변환 플레이트가 2분 동안 2,000g에서 회전된 후 상층물은 제거되었으며, 세포 펠렛은 200 uL ddH2O를 사용하여 3회 수세되었다. 웰 당 360 uL의 차가운 BSM을 포함하는, 미리 제조된, 미리 냉각된 96-웰 배양 플레이트 ("시드 플레이트(Seed Plate)")로부터 취한 100 uL 차가운 Bird Seed Media (BSM)를 사용하여 세포 펠렛이 재분산되었다. 분산된 세포는 시드 플레이트로 옮겨진 후 Multitron II Incubator Shaker에서 30℃에서 밤새 성장되었다. 시드 플레이트는 5분 동안 3,000g에서 회전된 후, 60 uL를 제외하고 모든 액체가 제거된 후, 세포 펠렛을 재분산하기 위하여 덮개가 덮여진 시드 플레이트는 1,000rpm에서 교반되었다. 세포의 다양한 희석물은 우라실이 결핍된 아가로 옮겨진 후 37℃에서 밤새 배양되었다. 기능적 URA3 선별표지(selectable marker)를 포함하는 효모 세포 형질변환체의 콜로니를 취하여 분석하였다.

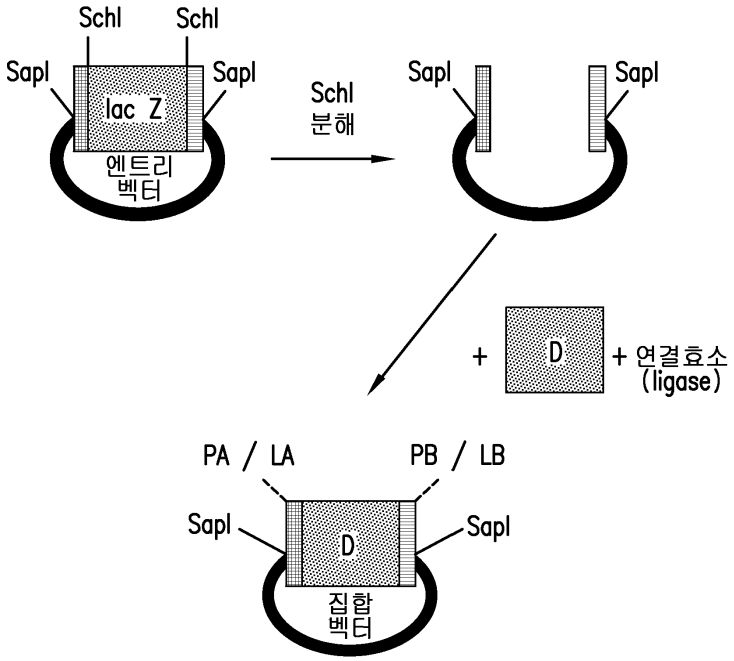
[0322] 각각의 공개 공보 또는 특허 출원이 특히, 그리고 개별적으로 인용에 의해 통합되는 것으로 나타나는 것처럼, 본 명세서에서 인용된 모든 공개 공보, 특허 및 특허 출원은 인용에 의해 본원에 통합된다. 명확히 이해시킬 목적으로 예시 및 실시예의 방법으로 어느 정도 상세하게 발명이 제시되었다고 하더라도, 첨부된 청구항의 사상 또는 범위에서 벗어남이 없이, 어떠한 변화 또는 변형이 본 발명의 내용에 이루어질 수 있다는 것은, 본 발명의 내용의 관점에서 당업자에게 분명할 것이다. 상기에 제시된 본 발명의 실시예는 단지 예시적인 것을 의도할 뿐이며 당업자는 통상적인 경험, 본원에 제시된 특정 공정에 대한 수많은 균등물 이상을 사용하지 않고도 인지하거나 확신할 수 있을 것이다. 그러한 모든 균등물들은 본 발명의 범위 내에 있는 것으로 간주되며 하기의 청구항에 의해 뒷받침된다. 더욱이, 본 명세서 및 청구항에 사용된 바와 같은, 단수 형태들 "하나의(a 또는 an)" 및 "상기(the)"는, 분명하게 달리 표현하지 않는 한, 복수 형태를 포함한다.

도면

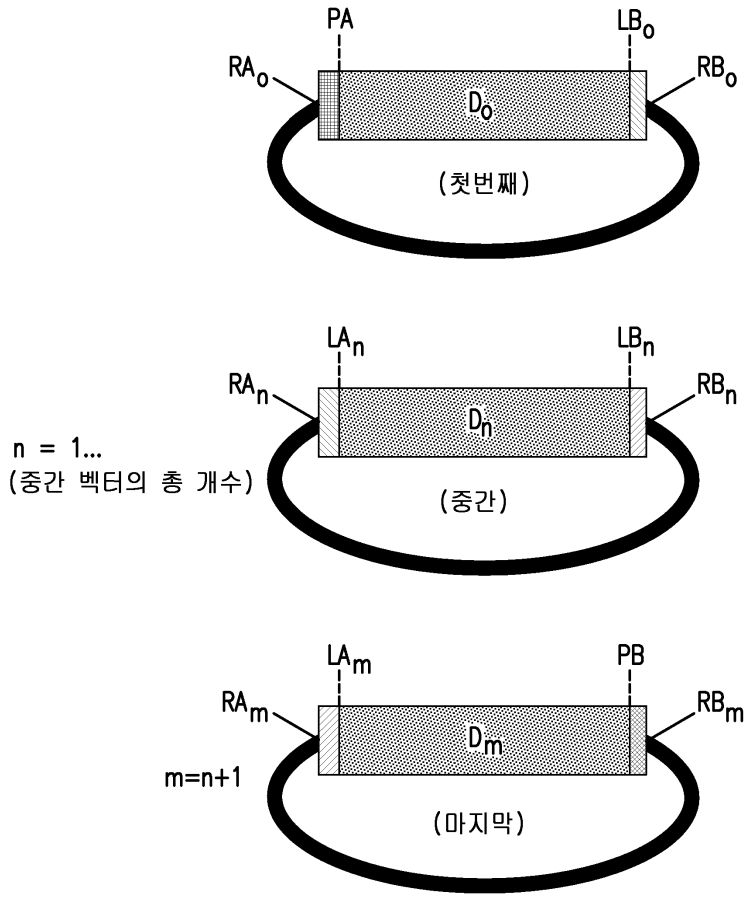
도면1a



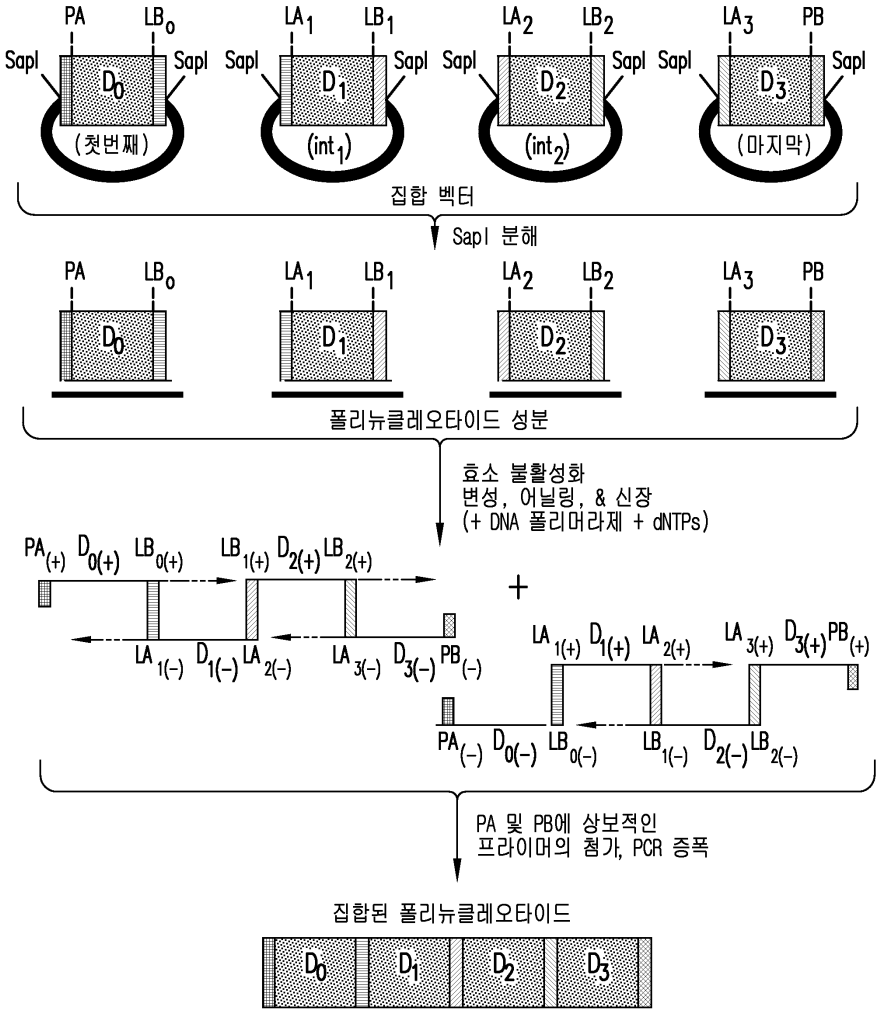
도면1b


SchI
SchI
SapI
SapI
lac Z
엔트리
벡터
SchI
분해
SapI
SapI
+
D
+ 연결효소
(ligase)
PA / LA
PB / LB
SapI
SapI
D
집합
벡터

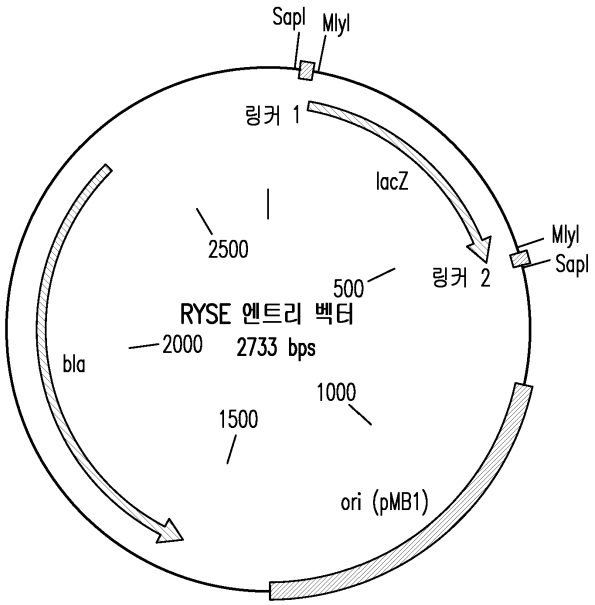
도면2


PA
LBo
RAo
RBo
Do
(첫번째)
LAn
LBn
RAn
RBn
Dn
n = 1...
(중간 벡터의 총 개수)
(중간)
LAm
PB
RAm
RBm
Dm
m=n+1
(마지막)

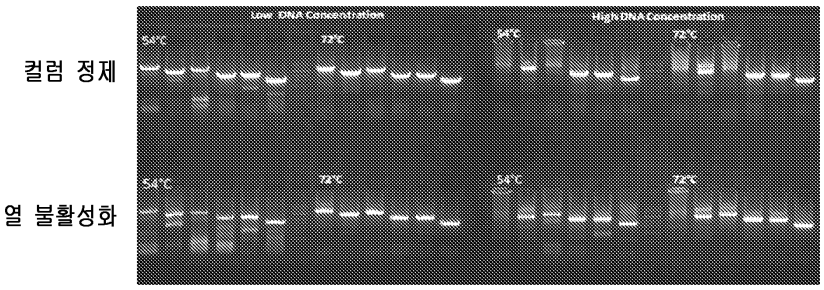
도면3


PA
LB0
LA1
LB1
LA2
LB2
LA3
PB
SapI
D0
D1
D2
D3
(첫번째)
(int1)
(int2)
(마지막)
집합 벡터
SapI 분해
폴리뉴클레오타이드 성분
효소 불활성화
변성, 어닐링, & 신장
(+ DNA 폴리머라제 + dNTPs)
PA(+) D0(+) LB0(+)
LB1(+) D2(+) LB2(+)
LA1(-) D1(-) LA2(-)
LA3(-) D3(-) PB(-)
LA1(+) D1(+) LA2(+)
LA3(+) D3(+) PB(+)
PA(-) D0(-) LB0(-)
LB1(-) D2(-) LB2(-)
PA 및 PB에 상보적인
프라이머의 첨가, PCR 증폭
집합된 폴리뉴클레오타이드

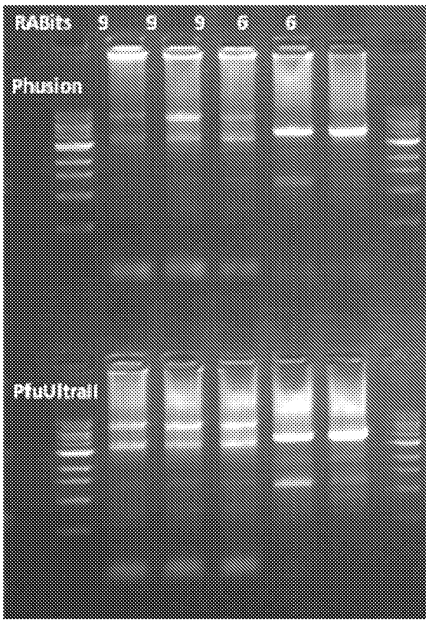
도면4


SapI
MlyI
링커 1
lacZ
2500
500
링커 2
MlyI
SapI
RYSE 엔트리 벡터
2733 bps
2000
bla
1000
1500
ori (pMB1)

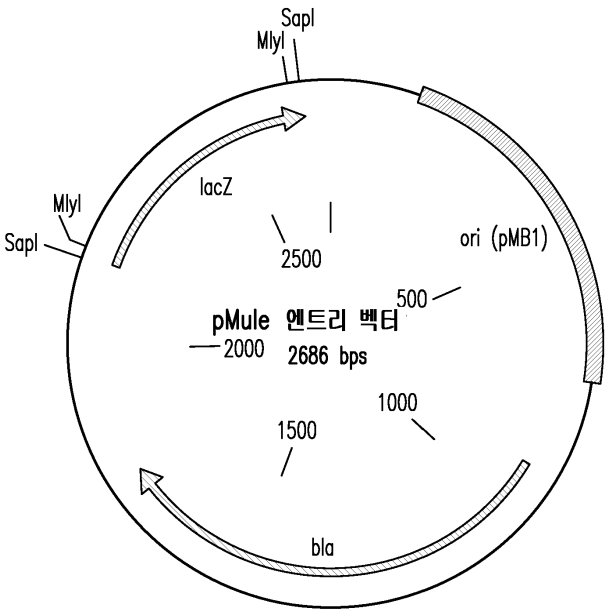
도면5


컬럼 정제
열 불활성화
Low DNA Concentration
High DNA Concentration
54°C
72°C
54°C
72°C
54°C
72°C
54°C
72°C

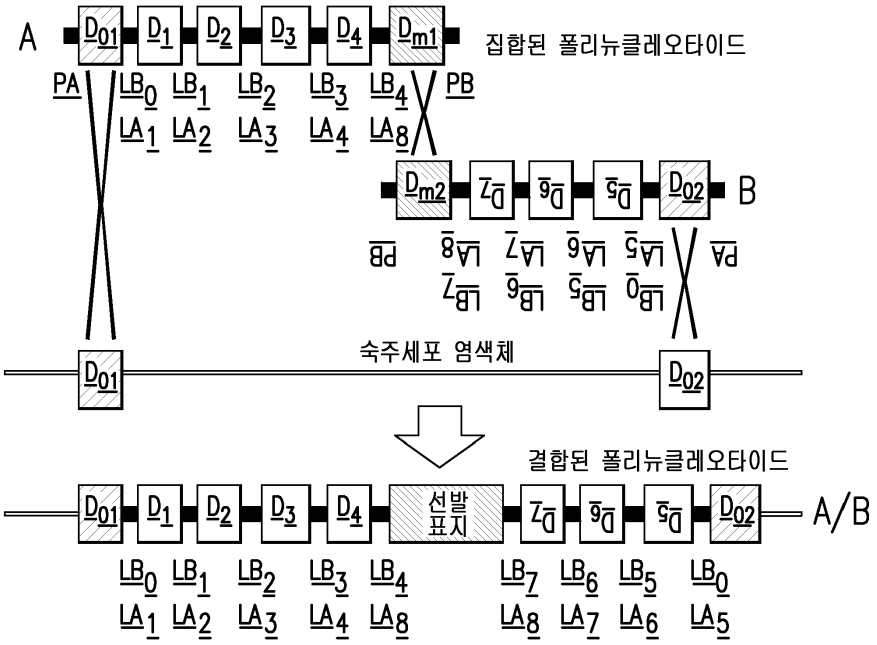
도면6


RABits 9 9 9 6 6
Phusion
PfuUltraII

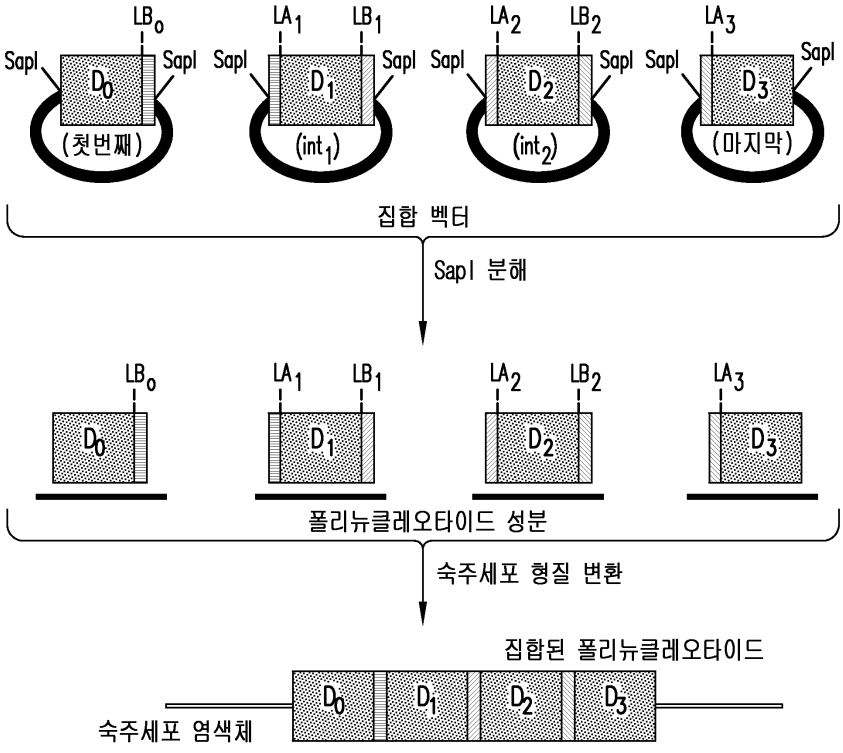
도면7


SapI
MlyI
lacZ
MlyI
SapI
2500
ori (pMB1)
500
pMule 엔트리 벡터
2000
2686 bps
1000
1500
bla

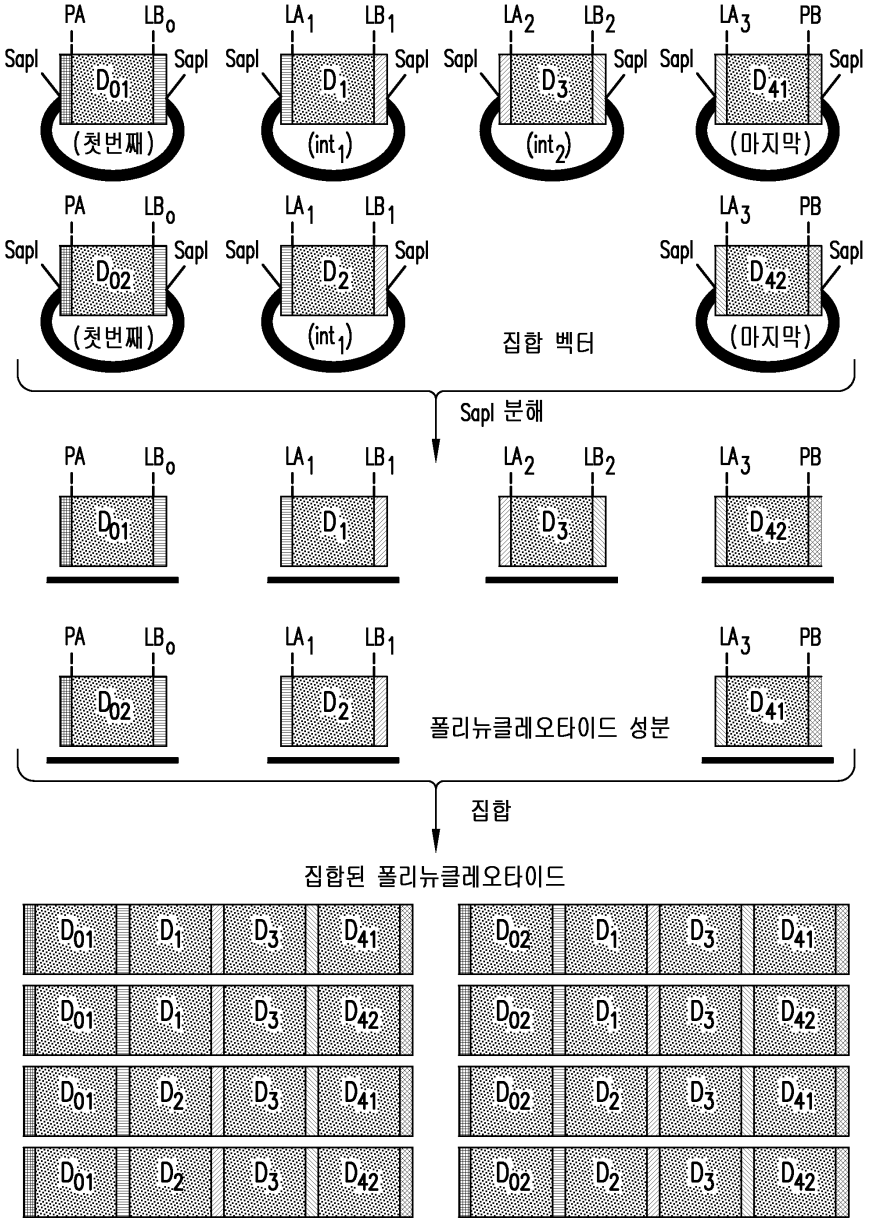
도면8


A
집합된 폴리뉴클레오타이드
PA
PB
B
숙주세포 염색체
결합된 폴리뉴클레오타이드
선발 표지
A/B

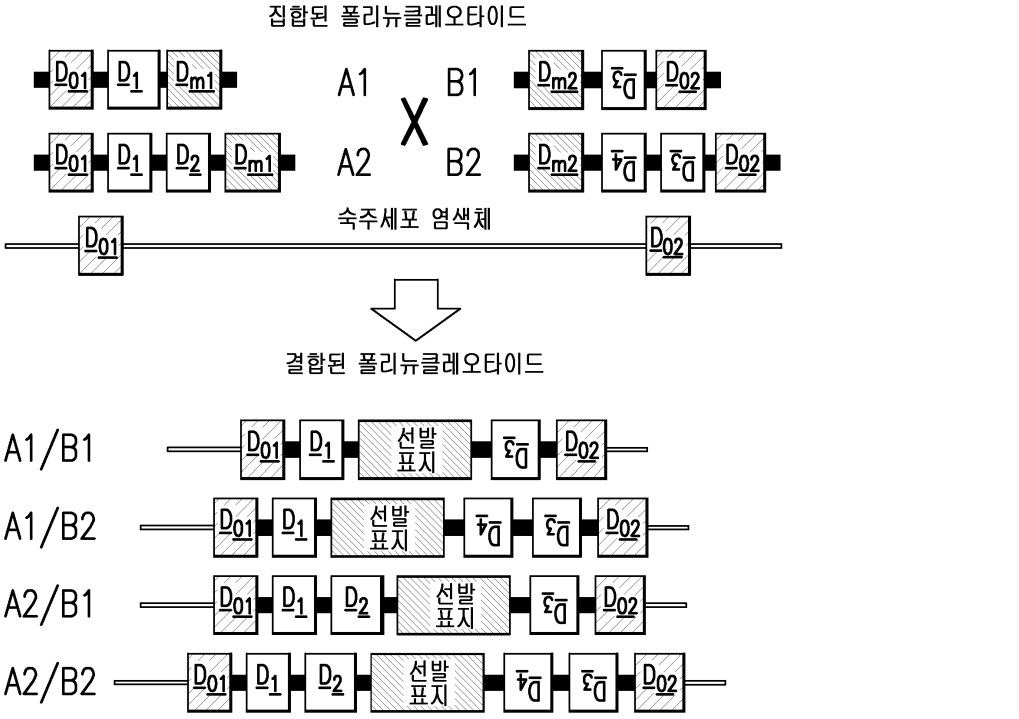
도면9


SapI
(첫번째)
(int1)
(int2)
(마지막)
집합 벡터
SapI 분해
폴리뉴클레오타이드 성분
숙주세포 형질 변환
집합된 폴리뉴클레오타이드
숙주세포 염색체

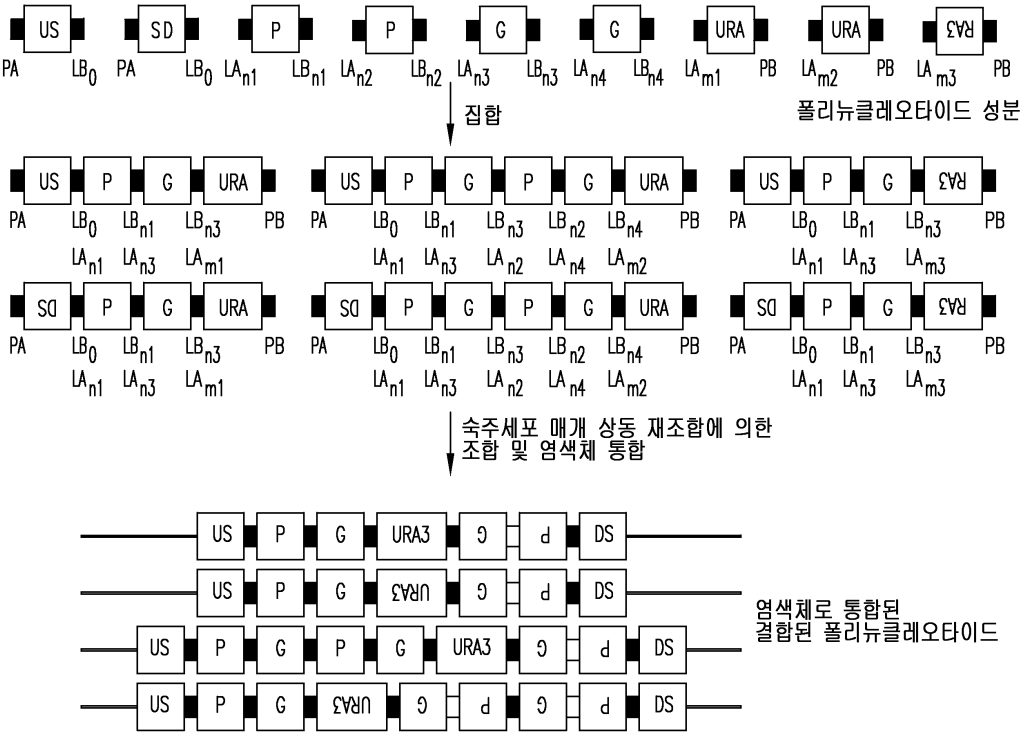
도면10


PA
LB0
SapI
D01
(첫번째)
LA1
LB1
D1
(int1)
LA2
LB2
D3
(int2)
LA3
PB
D41
(마지막)
D02
D2
D42
집합 벡터
SapI 분해
폴리뉴클레오타이드 성분
집합
집합된 폴리뉴클레오타이드

도면11

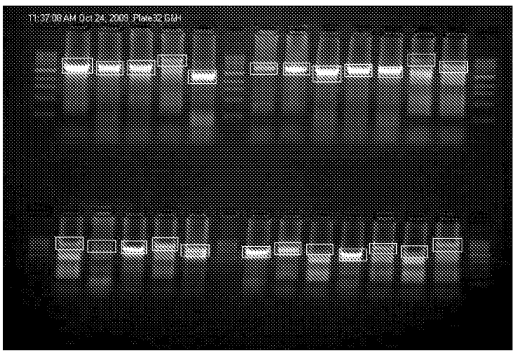

집합된 폴리뉴클레오타이드
A1
B1
A2
B2
숙주세포 염색체
결합된 폴리뉴클레오타이드
A1/B1
A1/B2
A2/B1
A2/B2
선발 표지

도면12a


US
SD
P
G
URA
PA
PB
집합
폴리뉴클레오타이드 성분
숙주세포 매개 상동 재조합에 의한
조합 및 염색체 통합
염색체로 통합된
결합된 폴리뉴클레오타이드

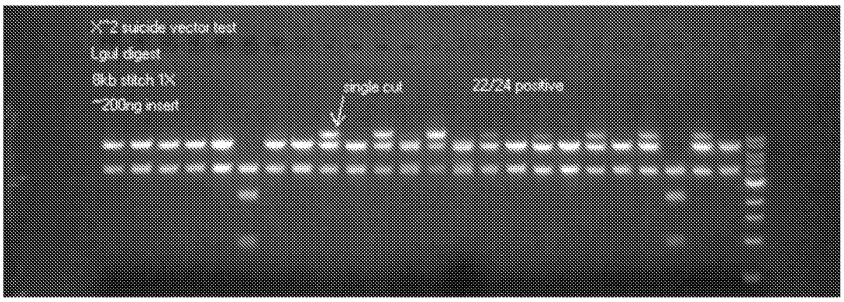
도면12b

상단 패널																				
Ladder	3909	3896	4027	6425	2709	Ladder	5310	5415	3518	3874	3933	8459	5370	Ladder	집합된 폴리뉴클레오타이드 크기(bp)
	4	4	4	6	4		6	6	4	4	4	6	6		집합된 폴리뉴클레오타이드 성분의 개수
	1621	1622	1623	1624	1625		1626	1627	1628	1629	1630	1631	1632		집합 개수

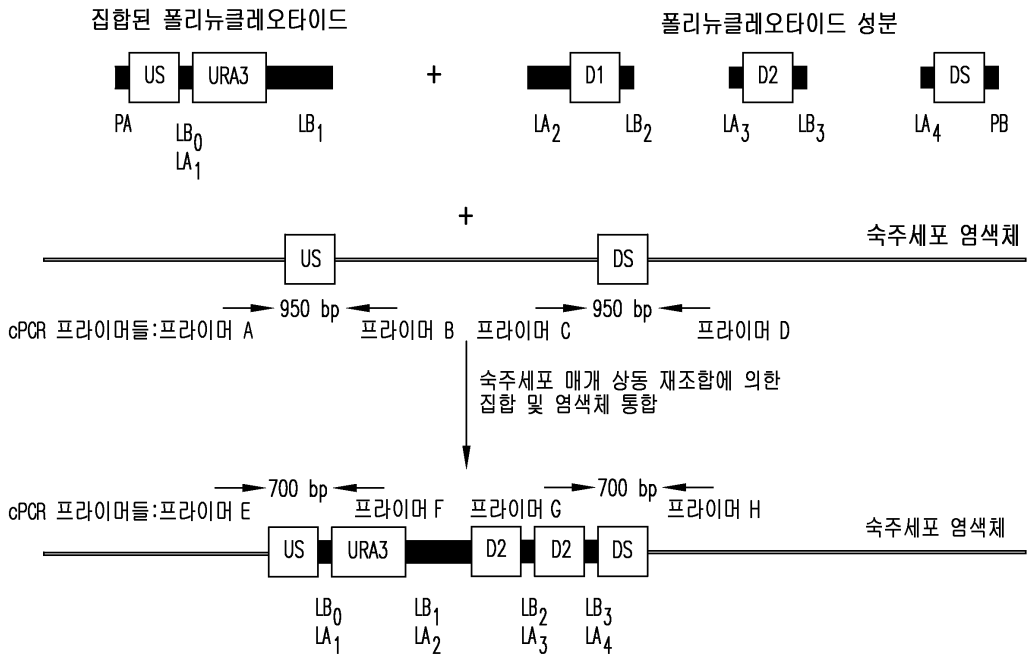


Ladder	1633	1634	1635	1636	1637	Ladder	1638	1639	1640	1641	1642	1643	1644	Ladder	집합 개수
	6	4	4	6	4		4	6	6	4	6	6	6		집합된 폴리뉴클레오타이드 성분의 개수
	5877	3848	4035	5135	3603		3516	6457	5492	3458	5453	4758	6818		집합된 폴리뉴클레오타이드 크기(bp)
하단 패널															

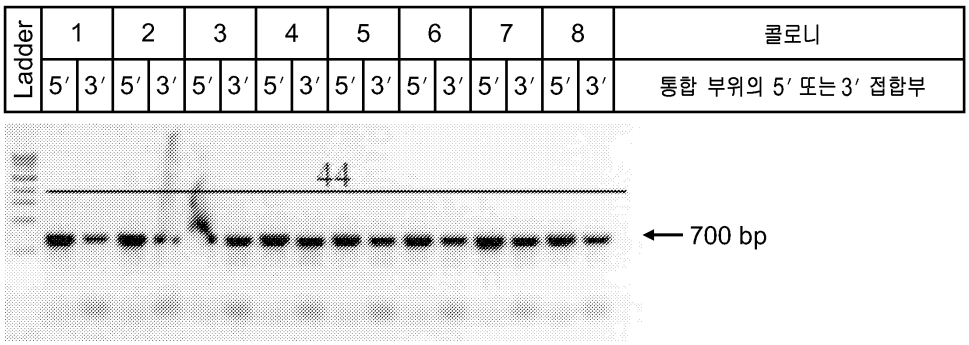
도면12c



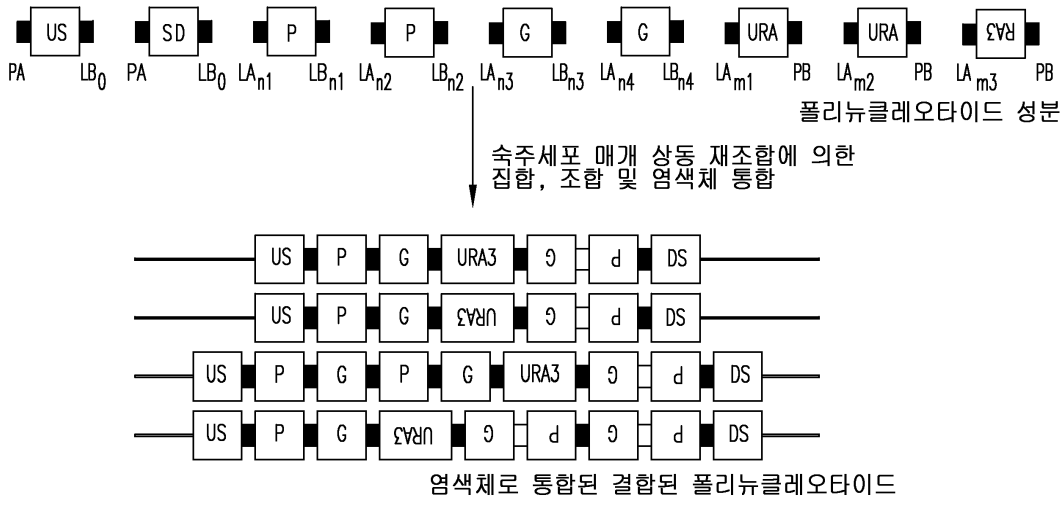
도면13a



도면13b



도면14



서 열 목 록

SEQUENCE LISTING

<110> Amyris Biotechnologies, Inc.

Serber, Zach

Lowe, Raymond

Ubersax, Jeffrey A.

Chandran, Sunil S.

Dean, Erik Jedediah

Platt, Darren M.

Takeoka, Kenneth Toshiki

<120> COMPOSITIONS AND METHODS FOR THE ASSEMBLY OF POLYNUCLEOTIDES

<130> 11836-047-228

<150> 61/116,109

<151> 2008-11-19

<150> 61/162,230

<151> 2009-03-20

<160> 220

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker Gold (Au)

<400> 1

gctcacacgc ggccaggggg agcc 24

<210> 2

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker Lapis (La)

<400> 2

cgctcgtcca acgccggcgg acct 24

<210> 3

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker Copper (Cu)

<400> 3

atccccgcgt gcttggccgg ccgt 24

<210> 4

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker Quartz (Qz)

<400> 4

aacctgcagg ccgcgagcgc cgat 24

<210> 5

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker Iron (Fe)

<400> 5

aacgcgatcg ccgacgccgc cgat 24

<210> 6

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker Obsidian (Ob)

<400> 6

aaggcggccg ctggcgaggg agat 24

<210> 7

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker Indigo (In)

<400> 7

aaggcgcgcc acggtcgtgc ggat 24

<210> 8

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker Silver (Ag)

<400> 8

agcccctcag ccccctagc gtcg 24

<210> 9

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker Pme1-5prime

<400> 9

gacggcacgg ccacgcgttt aaaccgcc 28

<210> 10

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Linker Pme1-3prime

<400> 10

cggtgtttaa accccagcgc ctggcggg 28

<210> 11

<211> 2737

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 1

<400> 11

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt cagacggcac 60

ggccacgcgt ttaaaccgcc tggcagactc catatgctat gcggcatcag agcagattgt 120

actgagagtg caccatatgc ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg 180

catcaggcgc cattcgccat tcaggctgcg caactgttgg gaagggcgat cggtgcgggc 240

ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat taagttgggt 300

aacgccaggg ttttcccagt cacgacgttg taaaacgacg gccagtgaat tcgagctcgg 360

tacccgggga tcctctagcg tcgacctgca ggcatgcaag cttggcgtaa tcatggtcat 420

agctgtttcc tgtgtgaaat tgttatccgc tcacaattcc acacaacata cgagccggaa 480

gcataaagtg taaagcctgg ggtgcctaat gagtgagcta actcacatta attgcgttgc 540

gcgctagcga gtcatccagc tcacacgcgg ccaggggggag cctgaagagc gagctcccgc 600

tgagcaataa ctagcgtcat agctgtttcc tgggtcgttc ggctgcggcg agcggtatca 660

gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 720

atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 780

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 840

cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 900

tctcctgttc cgaccctgcc gcttacccga tacctgtccg cctttctccc ttcgggaagc 960

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 1020

aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 1080

tatcgtcttg attccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 1140

aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 1200

aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 1260

ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 1320

ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 1380

atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 1440

atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 1500

tcaatctaaa gtatatatga gtaacttggt cgcatgctta ccaatgctta atcagtgagg 1560

cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactg cccgtcgtgt 1620

agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 1680

acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 1740

gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 1800

ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 1860

tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 1920

ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga 1980

tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 2040

attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 2100

agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2160

ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 2220

ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 2280

cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 2340

gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatca 2400

attgcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggttaca 2460

tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 2520

tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 2580

tcacgaggcc ctttcatctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc 2640

agctcccgga gacagtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc 2700

agggcgcgtc agcgggtgtt ggcgggtgtc ggggctg 2737

<210> 12

<211> 7692

<212> DNA

<213> Artificial Sequence

<220>

<223> DNA fragment resulting from Stitch 7

<400> 12

gacggcacgg ccacgcgttt aaaccgccct ccaagctgac ataaatcgca ctttgtatct 60

actttttttt attcgaaaac aaggcacaac aatgaatcta tcgccctgtg agattttcaa 120

tctcaagttt gtgtaataga tagcgttata ttatagaact ataaaggtcc ttgaatatac 180

atagtgtttc attcctatta ctgtatatgt gactttacat tgttacttcc gcggctattt 240

gacgttttct gcttcaggtg cggcttggag ggcaaagtgt cagaaaatcg gccaggccgt 300
atgacacaaa agagtagaaa acgagatctc aaatatctcg aggcctgtcc tctatacaac 360
cgcccagctc tctgacaaag ctccagaacg gttgtctttt gtttcgaaaa gccaaggtcc 420
cttataattg ccctccattt tgtgtcacct atttaagcaa aaaattgaaa gtttactaac 480
ctttcattaa agagaaataa caatattata aaaagcgctt aaagctcaca cgcggccagg 540
gggagccgag ctcctcgaga agttaagatt atatgaataa ctaaatacta aatagaaatg 600
taaatacagt gagaacaaaa caaaaaaaaa cgaacagaga aactaaatcc acattaattg 660
agagttctat ctattagaaa atgcaaactc caactaaatg ggaaaacaga taacctcttt 720

tatttttttt taatgtttga tattcgagtc tttttctttt gttaggttta tattcatcat 780
ttcaatgaat aaaagaagct tcttattttg gttgcaaaga atgaaaaaaa aggatttttt 840
catacttcta aagcttcaat tataaccaaa aattttataa atgaagagaa aaaatctagt 900
agtatcaagt taaacctatt cctttgccct cggacgagtg ctggggcgtc ggtttccact 960
atcggcgagt acttctacac agccatcggt ccagacggcc gcgcttctgc gggcgatttg 1020
tgtacgcccg acagtcccgg ctccggatcg gacgattgcg tcgcatcgac cctgcgccca 1080
agctgcatca tcgaaattgc cgtcaaccaa gctctgatag agttggtcaa gaccaatgcg 1140
gagcatatac gcccggagcc gcggcgatcc tgcaagctcc ggatgcctcc gctcgaagta 1200

gcgcgtctgc tgctccatac aagccaacca cggcctccag aagaagatgt tggcgacctc 1260
gtattgggaa tccccgaaca tcgcctcgct ccagtcaatg accgctgtta tgcggccatt 1320
gtccgtcagg acattgttgg agccgaaatc cgcgtgcacg aggtgccgga cttcggggca 1380
gtcctcggcc caaagcatca gctcatcgag agcctgcgcg acggacgcac tgacggtgtc 1440
gtccatcaca gtttgccagt gatacacatg gggatcagca atcgcgcata tgaaatcacg 1500
ccatgtagtg tattgaccga ttccttgcgg tccgaatggg ccgaacccgc tcgtctggct 1560
aagatcggcc gcagcgatcg catccatggc ctccgcgacc ggctgcagaa cagcgggcag 1620
ttcggtttca ggcaggtctt gcaacgtgac accctgtgca cggcgggaga tgcaataggt 1680

caggctctcg ctgaattccc caatgtcaag cacttccgga atcgggagcg cggccgatgc 1740
aaagtgccga taaacataac gatctttgta gaaaccatcg gcgcagctat ttacccgcag 1800
gacatatcca cgccctccta catcgaagct gaaagcacga gattcttcgc cctccgagag 1860
ctgcatcagg tcggagacgc tgtcgaactt ttcgatcaga aacttctcga cagacgtcgc 1920
ggtgagttca ggctttttca tttttaatgt tacttctctt gcagttaggg aactataatg 1980
taactcaaaa taagattaaa caaactaaaa taaaaagaag ttatacagaa aaacccatat 2040

aaaccagtac taatccataa taataataca caaaaaaact atcaaataaa accagaaaac 2100

agattgaata gaaaaatttt ttcgatctcc ttttatattc aaaattcgat atatgaaaaa 2160

gggaactctc agaaaatcac caaatcaatt taattagatt tttcttttcc ttctagcgtt 2220

ggaaagaaaa atttttcttt tttttttag aaatgaaaaa tttttgccgt aggaatcacc 2280

gtataaaccc tgtataaacg ctactctgtt cacctgtgta ggctatgatt gacccagtgt 2340

tcattgttat tgcgagagag cgggagaaaa gaaccgatac aagagatcca tgctggtata 2400

gttgtctgtc caacactttg atgaacttgt aggacgatga tgtgtattac tagtgtcgac 2460

gctcgtccaa cgccggcgga cctcttttaa ttctgctgta acccgtacat gcccaaaata 2520

gggggcgggt tacacagaat atataacatc gtaggtgtct gggtgaacag tttattcctg 2580

gcatccacta aatataatgg agcccgcttt ttaagctggc atccagaaaa aaaaagaatc 2640

ccagcaccaa aatattgttt tcttcaccaa ccatcagttc ataggtccat tctcttagcg 2700

caactacaga gaacaggggc acaaacaggc aaaaaacggg cacaacctca atggagtgat 2760

gcaacctgcc tggagtaaat gatgacacaa ggcaattgac ccacgcatgt atctatctca 2820

ttttcttaca ccttctatta ccttctgctc tctctgattt ggaaaaagct gaaaaaaaag 2880

gttgaaacca gttccctgaa attattcccc tacttgacta ataagtatat aaagacggta 2940

ggtattgatt gtaattctgt aaatctattt cttaaacttc ttaaattcta cttttatagt 3000

tagtcttttt tttagtttta aaacaccaag aacttagttt cgatccccgc gtgcttggcc 3060

ggccgtatga cagatgtagt aatagtatcc gccgcaagaa cagcagttgg aaagtttgga 3120

ggctctcttg caaagattcc agcccctgaa ttaggagctg ttgttataaa agccgcactt 3180

gaaagggcag gtgtgaagcc tgaacaagtc agtgaagtca taatgggtca agttttaact 3240

gccggctcag gtcaaaaccc agccagacag gctgctatta aagctggttt accggcaatg 3300

gttcccgcga tgactattaa caaagtttgt ggttccggcc ttaaagcagt gatgttagct 3360

gctaacgcaa taatggctgg ggatgctgaa atagtagtcg ccggaggaca agagaatatg 3420

agtgcagccc cacacgtttt accgggctcc agagatggat tccgtatggg tgacgctaag 3480

ttagttgata ctatgatagt agatggacta tgggatgtct ataaccaata tcacatgggt 3540

attacagccg aaaacgtggc gaaagaatat gggattacga gagaagcaca ggatgagttc 3600

gccgtgggta gtcaaaataa ggcggaggcg gctcaaaaag ccggtaaatt tgatgaggaa 3660

atagtacctg tccttatacc acagagaaaa ggagatccgg ttgcctttaa aaccgatgag 3720

tttgtcagac aaggcgccac attagacagc atgtctggtt tgaaacctgc ttttgataag 3780

gccgggaccg tgaccgctgc taatgcgtca ggactaaacg atggagctgc ggcggtggtt 3840

gttatgtctg ctgctaaagc aaaagaatta gggttaactc cattagccac tatcaaatct 3900

tatgctaacg cgggggtgga cccaaaagtg atgggaatgg gacctgttcc agccagtaag 3960

agggcgttat ctagggccga atggactcct caagacttgg atttaatgga aattaatgaa 4020

gcatttgccg cacaggcgtt agctgtccac caacagatgg gttgggatac aagtaaggtc 4080

aatgttaatg gaggtgcaat cgccattggt cacccaattg gtgcgtccgg atgtagaatt 4140

ttagttaccc tactgcatga gatgaagagg cgtgatgcaa agaaaggctt agcttcgttg 4200

tgtatcggtg gtggaatggg tgtggcatta gcagtcgagc gtaaataaaa cctgcaggcc 4260

gcgagcgccg attaagtgaa tttactttaa atcttgcatt taaataaatt ttctttttat 4320

agctttatga cttagtttca atttatatac tattttaatg acattttcga ttcattgatt 4380

gaaagctttg tgtttttttct tgatgcgcta ttgcattgtt cttgtctttt tcgccacatg 4440

taatatctgt agtagatacc tgatacattg tggatgctga gtgaaatttt agttaataat 4500

ggaggcgctc ttaataattt tggggatatt ggcttaacgc gatcgccgac gccgccgatt 4560

gacagcagga ttatcgtaat acgtaatagt tgaaaatctc aaaaatgtgt gggtcattac 4620

gtaaataatg ataggaatgg gattcttcta tttttccttt ttccattcta gcagccgtcg 4680

ggaaaacgtg gcatcctctc tttcgggctc aattggagtc acgctgccgt gagcatcctc 4740

tctttccata tctaacaact gagcacgtaa ccaatggaaa agcatgagct tagcgttgct 4800

ccaaaaaagt attggatggt taataccatt tgtctgttct cttctgactt tgactcctca 4860

aaaaaaaaaa atctacaatc aacagatcgc ttcaattacg ccctcacaaa aacttttttc 4920

cttcttcttc gcccacgtta aattttatcc ctcatgttgt ctaacggatt tctgcacttg 4980

atttattata aaaagacaaa gacataatac ttctctatca atttcagtta ttgttcttcc 5040

ttgcgttatt cttctgttct tctttttctt ttgtcatata taaccaaggc ggccgctggc 5100

gagggagata tgaaactctc aactaaactt tgttggtgtg gtattaaagg aagacttagg 5160

ccgcaaaagc aacaacaatt acacaataca aacttgcaaa tgactgaact aaaaaaacaa 5220

aagaccgctg aacaaaaaac cagacctcaa aatgtcggta ttaaaggtat ccaaatttac 5280

atcccaactc aatgtgtcaa ccaatctgag ctagagaaat ttgatggcgt ttctcaaggt 5340

aaatacacaa ttggtctggg ccaaaccaac atgtctttg tcaatgacag agaagatatc 5400

tactcgatgt ccctaactgt tttgtctaag ttgatcaaga gttacaacat cgacaccaac 5460

aaaattggta gattagaagt cggtactgaa actctgattg acaagtccaa gtctgtcaag 5520

tctgtcttga tgcaattgtt tggtgaaaac actgacgtcg aaggtattga cacgcttaat 5580

gcctgttacg gtggtaccaa cgcgttgttc aactctttga actggattga atctaacgca 5640

tgggatggta gagacgccat tgtagtttgc ggtgatattg ccatctacga taagggtgcc 5700

gcaagaccaa ccggtggtgc cggtactgtt gctatgtgga tcggtcctga tgctccaatt 5760

gtatttgact ctgtaagagc ttcttacatg gaacacgcct acgattttta caagccagat 5820

ttcaccagcg aatatcctta cgtcgatggt cattttttcat taacttgtta cgtcaaggct 5880

cttgatcaag tttacaagag ttattccaag aaggctattt ctaaagggtt ggttagcgat 5940

cccgctggtt cggatgcttt gaacgttttg aaatatttcg actacaacgt tttccatgtt 6000

ccaacctgta aattggtcac aaaatcatac ggtagattac tatataacga tttcagagcc 6060

aatcctcaat tgttcccaga agttgacgcc gaattagcta ctcgcgatta tgacgaatct 6120

ttaaccgata agaacattga aaaaactttt gttaatgttg ctaagccatt ccacaaagag 6180

agagttgccc aatctttgat tgttccaaca aacacaggta acatgtacac cgcatctgtt 6240

tatgccgcct ttgcatctct attaaactat gttggatctg acgacttaca aggcaagcgt 6300

gttggtttat tttcttacgg ttccggttta gctgcatctc tatattcttg caaaattgtt 6360

ggtgacgtcc aacatattat caaggaatta gatattacta acaaattagc caagagaatc 6420

accgaaactc caaaggatta cgaagctgcc atcgaattga gagaaaatgc ccatttgaag 6480

aagaacttca aacctcaagg ttccattgag catttgcaaa gtggtgttta ctacttgacc 6540

aacatcgatg acaaatttag aagatcttac gatgttaaaa aataatcttc ccccatcgat 6600

tgcatcttgc tgaacccccct tcataaatgc tttatttttt tggcagcctg ctttttttag 6660

ctctcattta atagagtagt tttttaatct atatactagg aaaactcttt atttaataac 6720

aatgatatat atatatattt tttttataaa gaattgtata tctatattta taacacaata 6780

aatctaatct caacttttt ctttaaagtt aagcccaacc gatttttttt ctcataaggc 6840

gcgccacggt cgtgcggatt aaagcttttg attaagcctt ctagtccaaa aaacacgttt 6900

ttttgtcatt tatttcattt tcttagaata gtttagttta ttcattttat agtcacgaat 6960

gttttatgat tctatatagg gttgcaaaca agcatttttc attttatgtt aaaacaattt 7020

caggtttacc ttttattctg cttgtggtga cgcgtgtatc cgcccgctct tttggtcacc 7080

catgtattta attgcataaa taattcttaa aagtggagct agtcagcccc tcagcccccc 7140

tagcgtcgat aaactaatga ttttaaatcg ttaaaaaaat atgcgaattc tgtggatcga 7200

acacaggacc tccagataac ttgaccgaag ttttttcttc agtctggcgc tctcccaact 7260

gagctaaatc cgcttactat ttgttatcag ttcccttcat atctacatag aataggttaa 7320

gtattttatt agttgccaga agaactactg atagttggga atatttggtg aataatgaag 7380

attgggtgaa taatttgata attttgagat tcaattgtta atcaatgtta caatattatg 7440

tatacagagt atactagaag ttctcttcgg agatcttgaa gttcacaaaa gggaatcgat 7500

atttctacat aatattatca ttacttcttc cccatcttat atttgtcatt cattattgat 7560

tatgatcaat gcaataatga ttggtagttg ccaaacattt aatacgatcc tctgtaatat 7620

ttctatgaat aattatcaca gcaacgttca attatcttca attccggtgt ttaaacccca 7680

gcgcctggcg gg 7692<210> 13

<211> 3627

<212> DNA

<213> Artificial Sequence

<220>

<223> Vector pAM1466

<400> 13

gcacttttcg gggaaatgtg cgcggaaccc ctatttgttt atttttctaa atacattcaa 60

atatgtatcc gctcatgaga caataaccct gataaatgct tcaataatat tgaaaaagga 120

agagtatgag tattcaacat ttccgtgtcg cccttattcc ctttttgcg gcattttgcc 180

ttcctgtttt tgctcaccca gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg 240

gtgcacgagt gggttacatc gaactggatc tcaacagcgg taagatcctt gagagttttc 300

gccccgaaga acgttttcca atgatgagca cttttaaagt tctgctatgt ggcgcggtat 360

tatcccgtat tgacgccggg caagagcaac tcggtcgccg catacactat tctcagaatg 420

acttggttga gtactcacca gtcacagaaa agcatcttac ggatggcatg acagtaagag 480

aattatgcag tgctgccata accatgagtg ataacactgc ggccaactta cttctgacaa 540

cgatcggagg accgaaggag ctaaccgctt ttttgcacaa catgggggat catgtaactc 600

gccttgatcg ttgggaaccg gagctgaatg aagccatacc aaacgacgag cgtgacacca 660

cgatgcctgt agcaatggca acaacgttgc gcaaactatt aactggcgaa ctacttactc 720

tagcttcccg gcaacaatta atagactgga tggaggcgga taaagttgca ggaccacttc 780

tgcgctcggc ccttccggct ggctggttta ttgctgataa atctggagcc ggtgagcgtg 840

ggtctcgcgg tatcattgca gcactggggc cagatggtaa gccctcccgt atcgtagtta 900

tctacacgac ggggagtcag gcaactatgg atgaacgaaa tagacagatc gctgagatag 960

gtgcctcact gattaagcat tggtaactgt cagaccaagt ttactcatat atactttaga 1020

ttgatttaaa acttcatttt taatttaaaa ggatctaggt gaagatcctt tttgataatc 1080

tcatgaccaa aatcccttaa cgtgagtttt cgttccactg agcgtcagac cccgtagaaa 1140

agatcaaagg atcttcttga gatccttttt ttctgcgcgt aatctgctgc ttgcaaacaa 1200

aaaaaccacc gctaccagcg gtggtttgtt tgccggatca agagctacca actcttttc 1260

cgaaggtaac tggcttcagc agagcgcaga taccaaatac tgtccttcta gtgtagccgt 1320

agttaggcca ccacttcaag aactctgtag caccgcctac atacctcgct ctgctaatcc 1380

tgttaccagt ggctgctgcc agtggcgata agtcgtgtct taccgggttg gactcaagac 1440

gatagttacc ggataaggcg cagcggtcgg gctgaacggg gggttcgtgc acacagccca 1500

gcttggagcg aacgacctac accgaactga gatacctaca gcgtgagcta tgagaaagcg 1560
ccacgcttcc cgaagggaga aaggcggaca ggtatccggt aagcggcagg gtcggaacag 1620
gagagcgcac gagggagctt ccagggggaa acgcctggta tctttatagt cctgtcgggt 1680
ttcgccacct ctgacttgag cgtcgatttt tgtgatgctc gtcaggggggg cggagcctat 1740
ggaaaaacgc cagcaacgcg gccttttttac ggttcctggc cttttgctgg ccttttgctc 1800
acatgttctt tcctgcgtta tcccctgatt ctgtggataa ccgtattacc gcctttgagt 1860
gagctgatac cgctcgccgc agccgaacga ccgagcggcc gccagcgagc tcagtgagcg 1920
aggaagcgga agagcgccca atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt 1980

aatgcagctg gcacgacagg tttcccgact ggaaagcggg cagtgagcgc aacgcaatta 2040
atgtgagtta gctcactcat taggcacccc aggctttaca ctttatgctt ccggctcgta 2100
tgttgtgtgg aattgtgagc ggataacaat ttcacacagg aaacagctat gaccatgagt 2160
ttaaacatga gcaaaggcga agaactgttc acgggcgttg taccgatcct ggtggaactg 2220
gacggggatg tgaatgggca caagttttca gtgagcggcg aaggagaagg cgatgcgacc 2280
tacggcaaac tgaccctgaa attcatttgc accaccggta aacttcccgt gccgtggccc 2340
accctggtga ccacctttgg ctatggcgta cagtgcttcg cgcgttaccc ggatcacatg 2400
aaacgccacg acttctttaa gagcgctatg ccagagggct acgtccagga acgcaccata 2460

ttcttcaaag acgacggcaa ctacaagacg cgcgctgaag tcaagtttga aggggacacg 2520
ctggtgaacc gtattgagct gaagggcatc gacttcaagg aggacgggaa catcctgggc 2580
cataagctgg agtacaatta caacagccac aacgtgtata tcatggcgga caagcagaag 2640
aacggcatca aggtcaactt caagatccgg cacaacatcg aggatggcag cgtgcagctg 2700
gcggatcatt atcaacagaa caccccgatt ggcgatggac cggtgctgct gcccgataat 2760
cattacctga gtacccagag cgccctgagc aaggacccga atgagaagcg tgatcacatg 2820
gtactgctgg aatttgtgac cgcggctggc atcacccacg gcatggatga actgtataaa 2880
taaggtaccg cggccgccgt ctccggggac agacgtctcg ccccttaggg tccatgcagt 2940

tggcttcgat ggtctctttt ttataggtcg agtaccaatt cgccctatag tgagtcgtat 3000
tacaattcac tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg cgttacccaa 3060
cttaatcgcc ttgcagcaca tccccctttc gccagctggc gtaatagcga agaggcccgc 3120
accgatcgcc cttcccaaca gttgcgcagc ctgaatggcg aatgggacgc gccctgtagc 3180
ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac acttgccagc 3240
gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt cgccggcttt 3300

ccccgtcaag ctctaaatcg ggggctccct ttagggttcc gatttagtgc tttacggcac 3360

ctcgacccca aaaaacttga ttagggtgat ggttcacgta gtgggccatc gccctgatag 3420

acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact cttgttccaa 3480

actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg gattttgccg 3540

atttcggcct attggttaaa aaatgagctg atttaacaaa aatttaacgc gaattttaac 3600

aaaatattaa cgcttacaat ttaggtg 3627

<210> 14

<211> 915

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA fragment 1040

<400> 14

taagtagttg accatagcta tggaaggtct cacggaaatg ttgaatactc atcaattgcc 60

tttttcaata ttattgaagc atttatcagg gttattgtct catgagcggt tacatatttg 120

aatgtattta gaaaaataaa caaatagggg ttccgcgcac atttccccga aaagtgccac 180

ctgacgtcta agaaaccatt attatcatga cattaaccta taaaaatagg cgtatcacga 240

ggcccctttca tctcgcgcgt ttcggtgatg acggtgaaaa cctctgacac atgcagctcc 300

cggagacagt cacagcttgt ctgtaagcgg atgccgggag cagacaagcc cgtcagggcg 360

cgtcagcggg tgttggcggg tgtcggggct ggtaaaacga cggccagtat taaccctcac 420

taaagggaac tcgaggctct tcacgctcgt ccaacgccgg cggaccttgg atgactcgct 480

agccgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 540

cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 600

ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 660

attgtgagcg gataacaatt tcacacagga aacagctatg agcaaaggcg aagaactgtt 720

cacgggcgtt gtaccgatcc tggtggaact ggacggggat gtgaatgggc acaagttttc 780

agtgagcggc gaaggagaag gcgatgcgac ctacggcaaa ctgaccctga aattcatttg 840

caccaccggt aaacttcccg tgccgtggcc caccctggtg accacctttg gagacgccac 900

gtacatggct tattg 915

<210> 15

<211> 882

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA fragment 1041

<400> 15

taagtagttg accatagcta tggaaggtct ccctttggct atggcgtaca gtgcttcgcg 60

cgttacccgg atcacatgaa acgccacgac ttctttaaga gcgctatgcc agagggctac 120

gtccaggaac gcaccatatt cttcaaagac gacggcaact acaagacgcg cgctgaagtc 180

aagtttgaag gggacacgct ggtgaaccgt attgagctga agggcatcga cttcaaggag 240

gacgggaaca tcctgggcca taagctggag tacaattaca acagccacaa cgtgtatatc 300

atggcggaca agcagaagaa cggcatcaag gtcaacttca agatccggca caacatcgag 360

gatggcagcg tgcagctggc ggatcattat caacagaaca ccccgattgg cgatggaccg 420

gtgctgctgc ccgataatca ttacctgagt acccagagcg ccctgagcaa ggacccgaat 480

gagaagcgtg atcacatggt actgctggaa tttgtgaccg cggctggcat cacccacggc 540

atggatgaac tgtataaata acatatggag tctgccatcg gtgtttaaac cccagcgcct 600

ggcgggtgaa gagcgagctc ccgctgagca ataactagcg tcatagctgt ttcctgggtc 660

gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa 720

tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt 780

aaaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga gcatcacaaa 840

aatcgacgct caagtcagag acgccacgta catggcttat tg 882

<210> 16

<211> 787

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA fragment 1042

<400> 16

taagtagttg accatagcta tggaaggtct caagtcagag gtggcgaaac ccgacaggac 60

tataaagata ccaggcgttt ccccctggaa gctccctcgt gcgctctcct gttccgaccc 120

tgccgcttac ccgatacctg tccgcctttc tcccttcggg aagcgtggcg ctttctcata 180

gctcacgctg taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc 240

acgaaccccc cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgtgtcca 300

acccggtaag acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag 360

cgaggtatgt aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta 420

gaagaacagt atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg 480

gtagctcttg atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc 540

agcagattac gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt 600

ctgacgctca gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa 660

ggatcttcac ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat 720

atgagtaaac ttggtcgcat gcttaccaat gcttaatcag tggagacgcc acgtacatgg 780

cttattg 787

<210> 17

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JCB158-17A

<400> 17

cgttcatcca tagttgcctg actgcccgtc gtgtag 36

<210> 18

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JCB158-17B

<400> 18

ctacacgacg ggcagtcagg caactatgga tgaacg 36

<210> 19

<211> 59

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JCB158-17C

<400> 19

atatgagtaa acttggtcgc atgcttacca atgcttaatc agtgaggcac ctatctcag 59

<210> 20

<211> 52

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JCB158-17D

<400> 20

ttgaaaaagg caattgatga gtattcaaca tttccgtgtc gcccttattc cc 52

<210> 21

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L012

<400> 21

ccggtaacta tcgtcttgat tccaacccgg taagacacg 39

<210> 22

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L013

<400> 22

cgtgtcttac cgggttggaa tcaagacgat agttaccgg 39

<210> 23

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L054

<400> 23

cgaaactaag ttcttggtgt tttaaaacta aaaaaaag 38

<210> 24

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L057

<400> 24

ggttatatat gacaaaagaa aaagaagaac agaag 35

<210> 25

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L109

<400> 25

atgaaactct ctactaaact ttgttggtg 29

<210> 26

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L110

<400> 26

atgagaaaaa aaatcggttg ggcttaac 28

<210> 27

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L176

<400> 27

gactagctcc acttttaaga attatttatg c 31

<210> 28

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L185

<400> 28

gtgaatttac tttaaatctt gcatttaaat aaattttc 38

<210> 29

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L186

<400> 29

aagccaatat ccccaaaatt attaagagcg 30

<210> 30

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer VH296-235-55-Leu2 12-1 F

<400> 30

gctcacacgc ggccaggggg agcccgttga gccattagta tcaatttgct tacc 54

<210> 31

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer VH296-235-55-Leu2 12-1 R

<400> 31

aggtccgccg gcgttggacg agcgaggcgc ctgattcaag aaatatcttg accgcag 57

<210> 32

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L219

<400> 32

ctccaagctg acataaatcg cactttg 27

<210> 33

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L220

<400> 33

tttaagcgct ttttataata ttgttatttc tctttaatg 39

<210> 34

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L221

<400> 34

ataaactaat gattttaaat cgttaaaaaa atatgcg 37

<210> 35

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L222

<400> 35

gaattgaaga taattgaacg ttgctgtg 28

<210> 36

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L224

<400> 36

cttttaattc tgctgtaacc cgtacatg 28

<210> 37

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L225

<400> 37

tgacagcagg attatcgtaa tacg 24

<210> 38

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L226

<400> 38

atgacagatg tagtaatagt atccgcc 27

<210> 39

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L227

<400> 39

ttatttacgc tcgactgcta atgccac 27

<210> 40

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L229

<400> 40

atgaaaaatt gtgtcatcgt cagtgc 26

<210> 41

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L230

<400> 41

ttaattcaat ctttcaatca ccatcgcaat tc 32

<210> 42

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L235

<400> 42

atgaccatcg gtattgacaa gattag 26

<210> 43

<211> 38

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L236

<400> 43

ttaattacga taagatctta cagtattgtt aatcgcag 38

<210> 44

<211> 35

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer L248

<400> 44

taaagctttt gattaagcct tctagtccaa aaaac 35

<210> 45

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S000

<400> 45

gacggcacgg ccacgcgttt aaaccgcc 28

<210> 46

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S001

<400> 46

ggcggtttaa acgcgtggcc gtgccgtc 28

<210> 47

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S002

<400> 47

gctcacacgc ggccaggggg agcc 24

<210> 48

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S003

<400> 48

ggctcccccct ggccgcgtgt gagc 24

<210> 49

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S004

<400> 49

cgctcgtcca acgccggcgg acct 24

<210> 50

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S005

<400> 50

aggtccgccg gcgttggacg agcg 24

<210> 51

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S006

<400> 51

atccccgcgt gcttggccgg ccgt 24

<210> 52

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S007

<400> 52

acggccggcc aagcacgcgg ggat 24

<210> 53

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S008

<400> 53

aacctgcagg ccgcgagcgc cgat 24

<210> 54

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S009

<400> 54

atcggcgctc gcggcctgca ggtt 24

<210> 55

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S010

<400> 55

aacgcgatcg ccgacgccgc cgat 24

<210> 56

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S011

<400> 56

atcggcggcg tcggcgatcg cgtt 24

<210> 57

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S012

<400> 57

aaggcggccg ctggcgaggg agat 24

<210> 58

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S013

<400> 58

atctccctcg ccagcggccg cctt 24

<210> 59

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S014

<400> 59

aaggcgcgcc acggtcgtgc ggat 24

<210> 60

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S015

<400> 60

atccgcacga ccgtggcgcg cctt 24

<210> 61

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S016

<400> 61

agcccctcag ccccccctagc gtcg 24

<210> 62

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S017

<400> 62

cgacgctagg ggggctgagg ggct 24

<210> 63

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S018

<400> 63

cggtgtttaa accccagcgc ctggcggg 28

<210> 64

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S019

<400> 64

cccgccaggc gctggggttt aaacaccg 28

<210> 65

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S027

<400> 65

tggatgactc gctagcgcgc aacgcaatta atgtgag 37

<210> 66

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S028

<400> 66

tggcagactc catatgctat gcggcatcag agcagattg 39

<210> 67

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S036

<400> 67

tacccgggga tcctctagcg tcgacctgca ggcatgcaag ct 42

<210> 68

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer S037

<400> 68

agcttgcatg cctgcaggtc gacgctagag gatccccggg ta 42

<210> 69

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J036

<400> 69

gccgcaaaaa agggaataag ggcgacacgg aaatgttgaa tactcatcaa ttgc 54

<210> 70

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J037

<400> 70

cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tg 42

<210> 71

<211> 42

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J038

<400> 71

catcacaaaa atcgacgctc aagtcagagg tggcgaaacc cg 42

<210> 72

<211> 55

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J039

<400> 72

ctgagatagg tgcctcactg attaagcatt ggtaagcatg cgaccaagtt tactc 55

<210> 73

<211> 58

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J018

<400> 73

gacggcacgg ccacgcgttt aaaccgcctt ggatggatac gctagccgcc caatacgc 58

<210> 74

<211> 58

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J019

<400> 74

gctcacacgc ggccaggggg agccttggat ggatggatac gctagccgcc caatacgc 58

<210> 75

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J020

<400> 75

cgctcgtcca acgccggcgg accttggatg gatggatacg ctagccgccc aatacgc 57

<210> 76

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J021

<400> 76

atccccgcgt gcttggccgg ccgttggatg gatggatacg ctagccgccc aatacgc 57

<210> 77

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J022

<400> 77

aacctgcagg ccgcgagcgc cgattggatg gatggatacg ctagccgccc aatacgc 57

<210> 78

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J023

<400> 78

aacgcgatcg ccgacgccgc cgattggatg gatggatacg ctagccgccc aatacgc 57

<210> 79

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J024

<400> 79

aaggcggccg ctggcgaggg agattggatg gatggatacg ctagccgccc aatacgc 57

<210> 80

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J025

<400> 80

aaggcgcgcc acggtcgtgc ggatggatgg atggatacgc tagccgccca atacgc 56

<210> 81

<211> 58

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J026

<400> 81

agcccctcag ccccccctagc gtcgttggat ggatggatac gctagccgcc caatacgc 58

<210> 82

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J029

<400> 82

acggccggcc aagcacgcgg ggattggcag gataccatat gttatttata cagttc 56

<210> 83

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J030

<400> 83

atcggcgctc gcggcctgca ggtttggcag gataccatat gttatttata cagttc 56

<210> 84

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J031

<400> 84

atcggcggcg tcggcgatcg cgtttggcag gataccatat gttatttata cagttc 56

<210> 85

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J032

<400> 85

atctccctcg ccagcggccg cctttggcag gataccatat gttatttata cagttc 56

<210> 86

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J033

<400> 86

atccgcacga ccgtggcgcg cctttggcag gataccatat gttatttata cagttc 56

<210> 87

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J034

<400> 87

cgacgctagg ggggctgagg ggcttggcag gataccatat gttatttata cagttc 56

<210> 88

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J055

<400> 88

aaagggaact cgaggctctt cagacggcac ggccacgcgt ttaaaccgcc 50

<210> 89

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J056

<400> 89

aaagggaact cgaggctctt cagctcacac gcggccaggg ggagcc 46

<210> 90

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J057

<400> 90

aaagggaact cgaggctctt cacgctcgtc caacgccggc ggacct 46

<210> 91

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J058

<400> 91

aaagggaact cgaggctctt caatccccgc gtgcttggcc ggccgt 46

<210> 92

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J059

<400> 92

aaagggaact cgaggctctt caaacctgca ggccgcgagc gccgat 46

<210> 93

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J060

<400> 93

aaagggaact cgaggctctt caaacgcgat cgccgacgcc gccgat 46

<210> 94

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J061

<400> 94

aaagggaact cgaggctctt caaaggcggc cgctggcgag ggagat 46

<210> 95

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J062

<400> 95

aaagggaact cgaggctctt caaaggcgcg ccacggtcgt gcggat 46

<210> 96

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J063

<400> 96

aaagggaact cgaggctctt caagcccctc agccccccta gcgtcg 46

<210> 97

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J064

<400> 97

ctcagcggga gctcgctctt caggctcccc ctggccgcgt gtgagc 46

<210> 98

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J065

<400> 98

ctcagcggga gctcgctctt caaggtccgc cggcgttgga cgagcg 46

<210> 99

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J066

<400> 99

ctcagcggga gctcgctctt caacggccgg ccaagcacgc ggggat 46

<210> 100

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J067

<400> 100

ctcagcggga gctcgctctt caatcggcgc tcgcggcctg caggtt 46

<210> 101

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J068

<400> 101

ctcagcggga gctcgctctt caatcggcgg cgtcggcgat cgcgtt 46

<210> 102

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J069

<400> 102

ctcagcggga gctcgctctt caatctccct cgccagcggc cgcctt 46

<210> 103

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J070

<400> 103

ctcagcggga gctcgctctt caatccgcac gaccgtggcg cgcctt 46

<210> 104

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J071

<400> 104

ctcagcggga gctcgctctt cacgacgcta ggggggctga ggggct 46

<210> 105

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J072

<400> 105

ctcagcggga gctcgctctt cacccgccag gcgctggggt ttaaacaccg 50

<210> 106

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J073

<400> 106

ggctcccccct ggccgcgtgt gagcttggca ggataccata tgttatttat acagttc 57

<210> 107

<211> 57

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J074

<400> 107

aggtccgccg gcgttggacg agcgttggca ggataccata tgttatttat acagttc 57

<210> 108

<211> 58

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer J075

<400> 108

cccgccaggc gctggggttt aaacaccgtt ggcaggatac catatgttat ttatacag 58

<210> 109

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer K162

<400> 109

tgaagagcga gctcccgctg 20

<210> 110

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer K163

<400> 110

tgaagagcct cgagttccct ttag 24

<210> 111

<211> 8139

<212> DNA

<213> Artificial Sequence

<220>

<223> Phase I-A stitch product

<400> 111

gacggcacgg ccacgcgttt aaaccgccca gatggaatcc cttccataga gagaaggagc 60

aagcaactga cccaatattg actgccactg gacctgaaga catgcaacaa agtgcaagca 120

tagtggggcc ttcttccaat gctaatccgg tcactgccac tgctgctacg gaaaaccaac 180

ctaaaggtat taacttcttc actataagaa aatcacacga gcgcccggac gatgtctctg 240

tttaaatggc gcaagttttc cgctttgtaa tatatattta tacccctttc ttctctcccc 300

tgcaatataa tagtttaatt ctaatattaa taatatccta tattttcttc atttaccggc 360

gcactctcgc ccgaacgacc tcaaaatgtc tgctacattc ataataacca aaagctcata 420

actttttttt ttgaacctga atatatatac atcacatgtc actgctggtc cttgccgacc 480

agcgtataca atctcgatag ttggtttccc gttctttcca ctcccgtcgc tcacacgcgg 540

ccaggggggag ccgagctcct cgagaagtta agattatatg aataactaaa tactaaatag 600

aaatgtaaat acagtgagaa caaaacaaaa aaaaacgaac agagaaacta aatccacatt 660

aattgagagt tctatctatt agaaaatgca aactccaact aaatgggaaa acagataacc 720

tcttttattt ttttttaatg tttgatattc gagtcttttt cttttgttag gtttatattc 780

atcatttcaa tgaataaaag aagcttctta ttttggttgc aaagaatgaa aaaaaaggat 840

tttttcatac ttctaaagct tcaattataa ccaaaaattt tataaatgaa gagaaaaaat 900

ctagtagtat caagttaaac ctattccttt gccctcggac gagtgctggg gcgtcggttt 960

ccactatcgg cgagtacttc tacacagcca tcggtccaga cggccgcgct tctgcgggcg 1020

atttgtgtac gcccgacagt cccggctccg gatcggacga ttgcgtcgca tcgaccctgc 1080

gcccaagctg catcatcgaa attgccgtca accaagctct gatagagttg gtcaagacca 1140

atgcggagca tatacgcccg gagccgcggc gatcctgcaa gctccggatg cctccgctcg 1200

aagtagcgcg tctgctgctc catacaagcc aaccacggcc tccagaagaa gatgttggcg 1260

acctcgtatt gggaatcccc gaacatcgcc tcgctccagt caatgaccgc tgttatgcgg 1320

ccattgtccg tcaggacatt gttggagccg aaatccgcgt gcacgaggtg ccggacttcg 1380

gggcagtcct cggcccaaag catcagctca tcgagagcct gcgcgacgga cgcactgacg 1440

gtgtcgtcca tcacagtttg ccagtgatac acatggggat cagcaatcgc gcatatgaaa 1500

tcacgccatg tagtgtattg accgattcct tgcggtccga atgggccgaa cccgctcgtc 1560

tggctaagat cggccgcagc gatcgcatcc atggcctccg cgaccggctg cagaacagcg 1620

ggcagttcgg tttcaggcag gtcttgcaac gtgacaccct gtgcacggcg ggagatgcaa 1680

taggtcaggc tctcgctgaa ttccccaatg tcaagcactt ccggaatcgg gagcgcggcc 1740

gatgcaaagt gccgataaac ataacgatct ttgtagaaac catcggcgca gctatttacc 1800

cgcaggacat atccacgccc tcctacatcg aagctgaaag cacgagattc ttcgccctcc 1860

gagagctgca tcaggtcgga gacgctgtcg aacttttcga tcagaaactt ctcgacagac 1920

gtcgcggtga gttcaggctt tttcattttt aatgttactt ctcttgcagt tagggaacta 1980

taatgtaact caaaataaga ttaaacaaac taaaataaaa agaagttata cagaaaaacc 2040

catataaacc agtactaatc cataataata atacacaaaa aaactatcaa ataaaaccag 2100

aaaacagatt gaatagaaaa atttttttcga tctcctttta tattcaaaat tcgatatatg 2160

aaaaagggaa ctctcagaaa atcaccaaat caatttaatt agatttttct tttccttcta 2220

gcgttggaaa gaaaaatttt tctttttttt tttagaaatg aaaaattttt gccgtaggaa 2280

tcaccgtata aaccctgtat aaacgctact ctgttcacct gtgtaggcta tgattgaccc 2340

agtgttcatt gttattgcga gagagcggga gaaaagaacc gatacaagag atccatgctg 2400

gtatagttgt ctgtccaaca ctttgatgaa cttgtaggac gatgatgtgt attactagtg 2460

tcgacgctcg tccaacgccg gcggacctag ttatgacaat tacaacaaca gaattctttc 2520

tatatatgca cgaacttgta atatggaaga aattatgacg tacaaactat aaagtaaata 2580

ttttacgtaa cacatggtgc tgttgtgctt ctttttcaag agaataccaa tgacgtatga 2640

ctaagtttag gatttaatgc aggtgacgga cccatctttc aaacgattta tatcagtggc 2700

gtccaaattg ttaggttttg ttggttcagc aggtttcctg ttgtgggtca tatgactttg 2760

aaccaaatgg ccggctgcta gggcagcaca taaggataat tcacctgcca agacggcaca 2820

ggcaactatt cttgctaatt gacgtgcgtt ggtaccagga gcggtagcat gtgggcctct 2880

tacacctaat aagtccaaca tggcaccttg tggttctaga acagtaccac caccgatggt 2940

acctacttcg atggatggca tggatacgga aattctcaaa tcaccgtcca cttctttcat 3000

caatgttata cagttggaac tttcgacatt ttgtgcagga tcttgtccta atgccaagaa 3060

aacagctgtc actaaattag ctgcatgtgc gttaaatcca ccaacagacc cagccattgc 3120

agatccaacc aaattcttag caatgttcaa ctcaaccaat gcggaaacat cacttttaa 3180

cacttttctg acaacatcac caggaatagt agcttctgcg acgacactct taccacgacc 3240

ttcgatccag ttgatggcag ctggtttttt gtcggtacag tagttaccag aaacggagac 3300

aacctccata tcttcccagc catactcttc taccatttgc tttaatgagt attcgacacc 3360

cttagaaatc atattcatac ccattgcgtc accagtagtt gttctaaatc tcatgaagag 3420

taaatctcct gctagacaag tttgaatatg ttgcagacgt gcaaatcttg atgtagagtt 3480

aaaagctttt ttaattgcgt tttgtccctc ttctgagtct aaccatatct tacaggcacc 3540

agatcttttc aaagttggga aacggactac tgggcctctt gtcataccat ccttagttaa 3600

aacagttgtt gcaccaccgc cagcattgat tgccttacag ccacgcatgg cagaagctac 3660

caaacaaccc tctgtagttg ccattggtat atgataagat gtaccatcga taaccaaggg 3720

gcctataaca ccaacgggca aaggcatgta acctataaca ttttcacaac aagcgccaaa 3780

tacgcggtcg tagtcataat ttttatatgg taaacgatca gatgctaata caggagcttc 3840

tgccaaaatt gaaagagcct tcctacgtac cgcaaccgct ctcgtagtat cacctaattt 3900

tttctccaaa gcgtacaaag gtaacttacc gtgaataacc aaggcagcga cctctttgtt 3960

cttcaattgt tttgtatttc cactacttaa taatgcttct aattcttcta aaggacgtat 4020

tttcttatcc aagctttcaa tatcgcggga atcatcttcc tcactagatg atgaaggtcc 4080

tgatgagctc gattgcgcag atgataaact tttgactttc gatccagaaa tgactgtttt 4140

attggttaaa actggtgtag aagccttttg tacaggagca gtaaaagact tcttggtgac 4200

ttcagtcttc accaattgat ctgcagccat atccccgcgt gcttggccgg ccgttacttt 4260

ttttttggat ggacgcaaag aagtttaata atcatattac atggcaatac caccatatac 4320

atatccatat ctaatcttac ttatatgttg tggaaatgta aagagcccca ttatcttagc 4380

ctaaaaaaac cttctctttg gaactttcag taatacgctt aactgctcat tgctatattg 4440

aagtacggat tagaagccgc cgagcgggcg acagccctcc gacggaagac tctcctccgt 4500

gcgtcctggt cttcaccggt cgcgttcctg aaacgcagat gtgcctcgcg ccgcactgct 4560

ccgaacaata aagattctac aatactagct tttatggtta tgaagaggaa aaattggcag 4620

taacctggcc ccacaaacct tcaaatcaac gaatcaaatt aacaaccata ggataataat 4680

gcgattagtt ttttagcctt atttctgggg taattaatca gcgaagcgat gatttttgat 4740

ctattaacag atatataaat gcaaaagctg cataaccact ttaactaata ctttcaacat 4800

tttcggtttg tattacttct tattcaaatg tcataaaagt atcaacaaaa aattgttaat 4860

atacctctat acttaacctg caggccgcga gcgccgatat gtctcagaac gtttacattg 4920

tatcgactgc cagaacccca attggttcat tccagggttc tctatcctcc aagacagcag 4980

tggaattggg tgctgttgct ttaaaaggcg ccttggctaa ggttccagaa ttggatgcat 5040

ccaaggattt tgacgaaatt atttttggta acgttctttc tgccaatttg ggccaagctc 5100

cggccagaca agttgctttg gctgccggtt tgagtaatca tatcgttgca agcacagtta 5160

acaaggtctg tgcatccgct atgaaggcaa tcattttggg tgctcaatcc atcaaatgtg 5220

gtaatgctga tgttgtcgta gctggtggtt gtgaatctat gactaacgca ccatactaca 5280

tgccagcagc ccgtgcgggt gccaaatttg gccaaactgt tcttgttgat ggtgtcgaaa 5340

gagatgggtt gaacgatgcg tacgatggtc tagccatggg tgtacacgca gaaaagtgtg 5400

cccgtgattg ggatattact agagaacaac aagacaattt tgccatcgaa tcctaccaaa 5460

aatctcaaaa atctcaaaag gaaggtaaat tcgacaatga aattgtacct gttaccatta 5520

agggatttag aggtaagcct gatactcaag tcacgaagga cgaggaacct gctagattac 5580

acgttgaaaa attgagatct gcaaggactg ttttccaaaa agaaaacggt actgttactg 5640

ccgctaacgc ttctccaatc aacgatggtg ctgcagccgt catcttggtt tccgaaaaag 5700

ttttgaagga aaagaatttg aagcctttgg ctattatcaa aggttggggt gaggccgctc 5760

atcaaccagc tgattttaca tgggctccat ctcttgcagt tccaaaggct ttgaaacatg 5820

ctggcatcga agacatcaat tctgttgatt actttgaatt caatgaagcc ttttcggttg 5880

tcggtttggt gaacactaag attttgaagc tagacccatc taaggttaat gtatatggtg 5940

gtgctgttgc tctaggtcac ccattgggtt gttctggtgc tagagtggtt gttacactgc 6000

tatccatctt acagcaagaa ggaggtaaga tcggtgttgc cgccatttgt aatggtggtg 6060

gtggtgcttc ctctattgtc attgaaaaga tatgattacg ttctgcgatt ttctcatgat 6120

ctttttcata aaatacataa atatataaat ggctttatgt ataacaggca taatttaaag 6180

ttttatttgc gattcatcgt ttttcaggta ctcaaacgct gaggtgtgcc ttttgactta 6240

cttttccgcc ttggcaagct ggccgggtga tacttgcaca agttccacta attactgaca 6300

tttgtggtat taactcgttt gactgctcta caattgtagg atgttaatca atgtcttggc 6360

tgcctaacgc gatcgccgac gccgccgata tgagaaaaaa aatcggttgg gcttaacttt 6420

aaagaaaaaa gttgagatta gatttattgt gttataaata tagatataca attctttata 6480

aaaaaaatat atatatatat cattgttatt aaataaagag ttttcctagt atatagatta 6540

aaaaactact ctattaaatg agagctaaaa aaagcaggct gccaaaaaaa taaagcattt 6600

atgaaggggg ttcagcaaga tgcaatcgat gggggaagat tatttttaa catcgtaaga 6660

tcttctaaat ttgtcatcga tgttggtcaa gtagtaaaca ccactttgca aatgctcaat 6720

ggaaccttga ggtttgaagt tcttcttcaa atgggcattt tctctcaatt cgatggcagc 6780

ttcgtaatcc tttggagttt cggtgattct cttggctaat ttgttagtaa tatctaattc 6840

cttgataata tgttggacgt caccaacaat tttgcaagaa tatagagatg cagctaaacc 6900

ggaaccgtaa gaaaataaac caacacgctt gccttgtaag tcgtcagatc caacatagtt 6960

taatagagat gcaaaggcgg cataaacaga tgcggtgtac atgttacctg tgtttgttgg 7020

aacaatcaaa gattgggcaa ctctctcttt gtggaatggc ttagcaacat taacaaaagt 7080

tttttcaatg ttcttatcgg ttaaagattc gtcataatcg cgagtagcta attcggcgtc 7140

aacttctggg aacaattgag gattggctct gaaatcgtta tatagtaatc taccgtatga 7200

ttttgtgacc aatttacagg ttggaacatg gaaaacgttg tagtcgaaat atttcaaaac 7260

gttcaaagca tccgaaccag cgggatcgct aaccaaccct ttagaaatag ccttcttgga 7320

ataactcttg taaacttgat caagagcctt gacgtaacaa gttaatgaaa aatgaccatc 7380

gacgtaagga tattcgctgg tgaaatctgg cttgtaaaaa tcgtaggcgt gttccatgta 7440

agaagctctt acagagtcaa atacaattgg agcatcagga ccgatccaca tagcaacagt 7500

accggcacca ccggttggtc ttgcggcacc cttatcgtag atggcaatat caccgcaaac 7560

tacaatggcg tctctaccat cccatgcgtt agattcaatc cagttcaaag agttgaacaa 7620

cgcgttggta ccaccgtaac aggcattaag cgtgtcaata ccttcgacgt cagtgttttc 7680

accaaacaat tgcatcaaga cagacttgac agacttggac ttgtcaatca gagtttcagt 7740

accgacttct aatctaccaa ttttgttggt gtcgatgttg taactcttga tcaacttaga 7800

caaaacagtt agggacatcg agtagatatc ttctctgtca ttgacaaaag acatgttggt 7860

ttggcccaga ccaattgtgt atttaccttg agaaacgcca tcaaatttct ctagctcaga 7920

ttggttgaca cattgagttg ggatgtaaat ttggatacct ttaataccga cattttgagg 7980

tctggttttt tgttcagcgg tcttttgttt ttttagttca gtcatttgca agtttgtatt 8040

gtgtaattgt tgttgctttt gcggcctaag tcttccttta ataccacacc aacaaagttt 8100

agtagagagt ttcataaggc ggccgctggc gagggagat 8139<210> 112

<211> 2962

<212> DNA

<213> Artificial Sequence

<220>

<223> Phase I-B stitch product

<400> 112

aaggcggccg ctggcgaggg agataagtat agaggtatat taacaatttt ttgttgatac 60

ttttatgaca tttgaataag aagtaataca aaccgaaaat gttgaaagta ttagttaaag 120

tggttatgca gcttttgcat ttatatatct gttaatagat caaaaatcat cgcttcgctg 180

attaattacc ccagaaataa ggctaaaaaa ctaatcgcat tattatccta tggttgttaa 240

tttgattcgt tgatttgaag gtttgtgggg ccaggttact gccaattttt cctcttcata 300

accataaaag ctagtattgt agaatcttta ttgttcggag cagtgcggcg cgaggcacat 360

ctgcgtttca ggaacgcgac cggtgaagac caggacgcac ggaggagagt cttccgtcgg 420

agggctgtcg cccgctcggc ggcttctaat ccgtacttca atatagcaat gagcagttaa 480

gcgtattact gaaagttcca aagagaaggt ttttttaggc taagataatg gggctcttta 540

catttccaca acatataagt aagattagat atggatatgt atatggtggt attgccatgt 600

aatatgatta ttaaacttct ttgcgtccat ccaaaaaaaa agtaaaggcg cgccacggtc 660

gtgcggatat ggctgcagat caattggtga agactgaagt caccaagaag tcttttactg 720

ctcctgtaca aaaggcttct acaccagttt taaccaataa aacagtcatt tctggatcga 780

aagtcaaaag tttatcatct gcgcaatcga gctcatcagg accttcatca tctagtgagg 840

aagatgattc ccgcgatatt gaaagcttgg ataagaaaat acgtccttta gaagaattag 900

aagcattatt aagtagtgga aatacaaaac aattgaagaa caaagaggtc gctgccttgg 960

ttattcacgg taagttacct ttgtacgctt tggagaaaaa attaggtgat actacgagag 1020

cggttgcggt acgtaggaag gctctttcaa ttttggcaga agctcctgta ttagcatctg 1080

atcgtttacc atataaaaat tatgactacg accgcgtatt tggcgcttgt tgtgaaaatg 1140
ttataggtta catgcctttg cccgttggtg ttataggccc cttggttatc gatggtacat 1200
cttatcatat accaatggca actacagagg gttgtttggt agcttctgcc atgcgtggct 1260
gtaaggcaat caatgctggc ggtggtgcaa caactgtttt aactaaggat ggtatgacaa 1320
gaggcccagt agtccgtttc ccaactttga aaagatctgg tgcctgtaag atatggttag 1380
actcagaaga gggacaaaac gcaattaaaa aagcttttaa ctctacatca agatttgcac 1440
gtctgcaaca tattcaaact tgtctagcag gagatttact cttcatgaga tttagaacaa 1500
ctactggtga cgcaatgggt atgaatatga tttctaaggg tgtcgaatac tcattaaagc 1560

aaatggtaga agagtatggc tgggaagata tggaggttgt ctccgtttct ggtaactact 1620
gtaccgacaa aaaaccagct gccatcaact ggatcgaagg tcgtggtaag agtgtcgtcg 1680
cagaagctac tattcctggt gatgttgtca gaaaagtgtt aaaaagtgat gtttccgcat 1740
tggttgagtt gaacattgct aagaatttgg ttggatctgc aatggctggg tctgttggtg 1800
gatttaacgc acatgcagct aatttagtga cagctgtttt cttggcatta ggacaagatc 1860
ctgcacaaaa tgtcgaaagt tccaactgta taacattgat gaaagaagtg gacggtgatt 1920
tgagaatttc cgtatccatg ccatccatcg aagtaggtac catcggtggt ggtactgttc 1980
tagaaccaca aggtgccatg ttggacttat taggtgtaag aggcccacat gctaccgctc 2040

ctggtaccaa cgcacgtcaa ttagcaagaa tagttgcctg tgccgtcttg gcaggtgaat 2100
tatccttatg tgctgcccta gcagccggcc atttggttca aagtcatatg acccacaaca 2160
ggaaacctgc tgaaccaaca aaacctaaca atttggacgc cactgatata aatcgtttga 2220
aagatgggtc cgtcacctgc attaaatcct aaacttagtc atacgtcatt ggtattctct 2280
tgaaaaagaa gcacaacagc accatgtgtt acgtaaaata tttactttat agtttgtacg 2340
tcataatttc ttccatatta caagttcgtg catatataga aagaattctg ttgttgtaat 2400
tgtcataact agcccctcag ccccctagc gtcgaagcat cttgccctgt gcttggcccc 2460
cagtgcagcg aacgttataa aaacgaatac tgagtatata tctatgtaaa acaaccatat 2520

catttcttgt tctgaacttt gtttacctaa ctagttttaa atttcccttt ttcgtgcatg 2580
cgggtgttct tatttattag catactacat ttgaaatatc aaatttcctt agtagaaaag 2640
tgagagaagg tgcactgaca caaaaaataa aatgctacgt ataactgtca aaactttgca 2700
gcagcgggca tccttccatc atagcttcaa acatattagc gttcctgatc ttcatacccg 2760
tgctcaaaat gatcaaacaa actgttattg ccaagaaata aacgcaaggc tgccttcaaa 2820
aactgatcca ttagatcctc atatcaagct tcctcataga acgcccaatt acaataagca 2880

tgttttgctg ttatcaccgg gtgataggtt tgctcaacca tggaaggtag catgcggtgt 2940

ttaaacccca gcgcctggcg gg 2962

<210> 113

<211> 9664

<212> DNA

<213> Artificial Sequence

<220>

<223> Phase II complete stitch product

<400> 113

gacggcacgg ccacgcgttt aaaccgccta ggataattat actctatttc tcaacaagta 60

attggttgtt tggccgagcg gtctaaggcg cctgattcaa gaaatatctt gaccgcagtt 120

aactgtggga atactcaggt atcgtaagat gcaagagttc gaatctctta gcaaccatta 180

ttttttttcct caacataacg agaacacaca ggggcgctat cgcacagaat caaattcgat 240

gactggaaat tttttgttaa tttcagaggt cgcctgacgc atataccttt ttcaactgaa 300

aaattgggag aaaaaggaaa ggtgagagcg ccggaaccgg cttttcatat agaatagaga 360

agcgttcatg actaaatgct tgcatcacaa tacttgaagt tgacaatatt atttaaggac 420

ctattgtttt ttccaatagg tggttagcaa tcgtcttact ttctaacttt tcttaccttt 480

tacatttcag caatatatat atatatattt caaggatata ccattctagc tcacacgcgg 540

ccaggggggag cctcgacact agtaatacac atcatcgtcc tacaagttca tcaaagtgtt 600

ggacagacaa ctataccagc atggatctct tgtatcggtt cttttctccc gctctctcgc 660

aataacaatg aacactgggt caatcatagc ctacacaggt gaacagagta gcgtttatac 720

agggtttata cggtgattcc tacggcaaaa atttttcatt tctaaaaaaa aaaagaaaaa 780

tttttctttc caacgctaga aggaaaagaa aaatctaatt aaattgattt ggtgattttc 840

tgagagttcc ctttttcata tatcgaattt tgaatataaa aggagatcga aaaaattttt 900

ctattcaatc tgttttctgg ttttatttga tagttttttt gtgtattatt attatggatt 960

agtactggtt tatatgggtt tttctgtata acttcttttt attttagttt gtttaatctt 1020

attttgagtt acattatagt tccctaactg caagagaagt aacattaaaa atgaccactc 1080

ttgacgacac ggcttaccgg taccgcacca gtgtcccggg ggacgccgag gccatcgagg 1140

cactggatgg gtccttcacc accgacaccg tcttccgcgt caccgccacc ggggacggct 1200

tcaccctgcg ggaggtgccg gtggacccgc ccctgaccaa ggtgttcccc gacgacgaat 1260

cggacgacga atcggacgcc ggggaggacg gcgacccgga ctcccggacg ttcgtcgcgt 1320

acggggacga cggcgacctg gcgggcttcg tggtcgtctc gtactccggc tggaaccgcc 1380

ggctgaccgt cgaggacatc gaggtcgccc cggagcaccg ggggcacggg gtcgggcgcg 1440

cgttgatggg gctcgcgacg gagttcgccc gcgagcgggg cgccgggcac ctctggctgg 1500

aggtcaccaa cgtcaacgca ccggcgatcc acgcgtaccg gcggatgggg ttcaccctct 1560

gcggcctgga caccgccctg tacgacggca ccgcctcgga cggcgagcag gcgctctaca 1620

tgagcatgcc ctgcccctga gtttaacttg atactactag atttttttctc ttcatttata 1680

aaatttttgg ttataattga agctttagaa gtatgaaaaa atccttttttt ttcattcttt 1740

gcaaccaaaa taagaagctt cttttattca ttgaaatgat gaatataaac ctaacaaaag 1800

aaaaagactc gaatatcaaa cattaaaaaa aaataaaaga ggttatctgt tttcccattt 1860

agttggagtt tgcattttct aatagataga actctcaatt aatgtggatt tagtttctct 1920

gttcgttttt ttttgttttg ttctcactgt atttacattt ctatttagta tttagttatt 1980

catataatct taacttctcg aggagctccg ctcgtccaac gccggcggac ctattcgcgg 2040

gtggaaggac cttgtggagg aatatgaagt tgatagctca aagggattga atggctattt 2100

aactgattat gagtcaatgt atcaaggata ctatggtctg cttaaatttc attctgtctt 2160

cgaaagctga attgatacta cgaaaaattt ttttttgttt ctctttctat ctttattaca 2220

taaaacttca tacacagtta agattaaaaa caactaataa ataatgccta tcgcaaatta 2280

gcttatgaag tccatggtaa attcgtgttt cctggcaata atagatcgtc aatttgttgc 2340

tttgtggtag ttttattttc aaataattgg aatactaggg atttgatttt aagatcttta 2400

ttcaaatttt ttgcgcttaa caaacagcag ccagtcccac ccaagtctgt ttcaaatgtc 2460

tcgtaactaa aatcatcttg caatttcttt ttgaaactgt caatttgctc ttgagtaatg 2520

tctcttcgta acaaagtcaa agagcaaccg ccgccaccag caccggtaag ttttgtggag 2580

ccaattctca aatcatcgct cagattttta ataagttcta atccaggatg agaaacaccg 2640

attgagacaa gcagtccatg atttattctt atcaattcca atagttgttc atacagttca 2700

ttattagttt ctacagcctc gtcatcggtg cctttacatt tacttaactt agtcatgatc 2760

tctaagcctt gtagggcaca ttcacccatg gcatctagaa ttggcttcat aacttcagga 2820

aatttctcgg tgaccaacac acgaacgcga gcaacaagat cttttgtaga ccttggaatt 2880

ctagtatagg ttaggatcat tggaatggct gggaaatcat ctaagaactt aaaattgttt 2940

gtgtttattg ttccattatg tgagtctttt tcaaatagca gggcattacc ataagtggcc 3000

acagcgttat ctattcctga aggggtaccg tgaatacact tttcacctat gaaggcccat 3060

tgattcacta tatgcttatc gttttctgac agcttttcca agtcattaga tcctattaac 3120

ccccccaagt aggccatagc taaggccagt gatacagaaa tagaggcgct tgagcccaac 3180

ccagcaccga tgggtaaagt agactttaaa gaaaacttaa tattcttggc atgggggcat 3240

aggcaaacaa acatatacag gaaacaaaac gctgcatggt agtggaagga ttcggatagt 3300
tgagctaaca acggatccaa aagactaacg agttcctgag acaagccatc ggtggcttgt 3360
tgagccttgg ccaatttttg ggagtttact tgatcctcgg tgatggcatt gaaatcattg 3420
atggaccact tatgattaaa gctaatgtcc gggaagtcca attcaatagt atctggtgca 3480
gatgactcgc ttattagcag gtaggttctc aacgcagaca cactagcagc gacggcaggc 3540
ttgttgtaca cagcagagtg ttcaccaaaa ataataacct ttcccggtgc agaagttaag 3600
aacggtaatg acatatcccc gcgtgcttgg ccggccgtta ctttttttt ggatggacgc 3660
aaagaagttt aataatcata ttacatggca ataccaccat atacatatcc atatctaatc 3720

ttacttatat gttgtggaaa tgtaaagagc cccattatct tagcctaaaa aaaccttctc 3780
tttggaactt tcagtaatac gcttaactgc tcattgctat attgaagtac ggattagaag 3840
ccgccgagcg ggcgacagcc ctccgacgga agactctcct ccgtgcgtcc tggtcttcac 3900
cggtcgcgtt cctgaaacgc agatgtgcct cgcgccgcac tgctccgaac aataaagatt 3960
ctacaatact agcttttatg gttatgaaga ggaaaaattg gcagtaacct ggccccacaa 4020
accttcaaat caacgaatca aattaacaac cataggataa taatgcgatt agttttttag 4080
ccttatttct ggggtaatta atcagcgaag cgatgatttt tgatctatta acagatatat 4140
aaatgcaaaa gctgcataac cactttaact aatactttca acattttcgg tttgtattac 4200

ttcttattca aatgtcataa aagtatcaac aaaaaattgt taatatacct ctatacttaa 4260
cctgcaggcc gcgagcgccg atatgtcaga gttgagagcc ttcagtgccc cagggaaagc 4320
gttactagct ggtggatatt tagttttaga tccgaaatat gaagcatttg tagtcggatt 4380
atcggcaaga atgcatgctg tagcccatcc ttacggttca ttgcaagagt ctgataagtt 4440
tgaagtgcgt gtgaaaagta aacaatttaa agatggggag tggctgtacc atataagtcc 4500
taaaactggc ttcattcctg tttcgatagg cggatctaag aacccttca ttgaaaaagt 4560
tatcgctaac gtatttagct actttaagcc taacatggac gactactgca atagaaactt 4620
gttcgttatt gatattttct ctgatgatgc ctaccattct caggaggaca gcgttaccga 4680

acatcgtggc aacagaagat tgagttttca ttcgcacaga attgaagaag ttcccaaaac 4740
agggctgggc tcctcggcag gtttagtcac agttttaact acagctttgg cctccttttt 4800
tgtatcggac ctggaaaata atgtagacaa atatagagaa gttattcata atttatcaca 4860
agttgctcat tgtcaagctc agggtaaaat tggaagcggg tttgatgtag cggcggcagc 4920
atatggatct atcagatata gaagattccc acccgcatta atctctaatt tgccagatat 4980
tggaagtgct acttacggca gtaaactggc gcatttggtt aatgaagaag actggaatat 5040

aacgattaaa agtaaccatt taccttcggg attaacttta tggatgggcg atattaagaa 5100

tggttcagaa acagtaaaac tggtccagaa ggtaaaaaat tggtatgatt cgcatatgcc 5160

ggaaagcttg aaaatatata cagaactcga tcatgcaaat tctagattta tggatggact 5220

atctaaacta gatcgcttac acgagactca tgacgattac agcgatcaga tatttgagtc 5280

tcttgagagg aatgactgta cctgtcaaaa gtatcctgag atcacagaag ttagagatgc 5340

agttgccaca attagacgtt cctttagaaa aataactaaa gaatctggtg ccgatatcga 5400

acctcccgta caaactagct tattggatga ttgccagacc ttaaaaggag ttcttacttg 5460

cttaatacct ggtgctggtg gttatgacgc cattgcagtg attgctaagc aagatgttga 5520

tcttagggct caaaccgctg atgacaaaag attttctaag gttcaatggc tggatgtaac 5580

tcaggctgac tggggtgtta ggaaagaaaa agatccggaa acttatcttg ataaataact 5640

taaggtagat aatagtggtc catgtgacat ctttataaat gtgaagtttg aagtgaccgc 5700

gcttaacatc taaccattca tcttccgata gtacttgaaa ttgttccttt cggcggcatg 5760

ataaaattcg tttaatgggt acaagctata catactagga tgaggatggt actgagaacg 5820

ataaataaac tttctagata tataacttta tgtcatttaa atataaataa agtgcgtgtt 5880

agcttgaaag tgtgcactaa cgcgatcgcc gacgccgccg atgatgtgta ttactagtgt 5940

cgacgacagc attcgcccag tatttttttt attctacaaa ccttctataa tttcaaagta 6000

tttacataat tctgtatcag tttaatcacc ataatatcgt tttctttgtt tagtgcaatt 6060

aatttttcct attgttactt cgggccttttt tctgttttat gagctatttt ttccgtcatc 6120

cttccggatc cagattttca gcttcatctc cagattgtgt ctacgtaatg cacgccatca 6180

ttttaagaga ggaaggcggc cgctggcgag ggagatatga agctactgtc ttctatcgaa 6240

caagcatgcg atatttgccg acttaaaaag ctcaagtgct ccaaagaaaa accgaagtgc 6300

gccaagtgtc tgaagaacaa ctgggagtgt cgctactctc ccaaaaccaa aaggtctccg 6360

ctgactaggg cacatctgac agaagtggaa tcaaggctag aaagactgga acagctattt 6420

ctactgattt ttcctcgaga agaccttgac atgattttga aaatggattc tttacaggat 6480

ataaaagcat tgttaacagg attatttgta caagataatg tgaataaaga tgccgtcaca 6540

gatagattgg cttcagtgga gactgatatg cctctaacat tgagacagca tagaataagt 6600

gcgacatcat catcggaaga gagtagtaac aaaggtcaaa gacagttgac tgtatcgatt 6660

gactcggcag ctcatcatga taactccaca attccgttgg attttatgcc cagggatgca 6720

cttcatggat ttgattggtc tgaagaggat gacatgtcgg atggcttgcc cttcctgaaa 6780

acggacccca acaataatgg gttctttggc gacggttctc tcttatgtat tcttcgatct 6840

attggcttta aaccggaaaa ttacacgaac tctaacgtta acaggctccc gaccatgatt 6900

acggatagat acacgttggc ttctagatcc acaacatccc gtttacttca aagttatctc 6960

aataattttc acccctactg ccctatcgtg cactcaccga cgctaatgat gttgtataat 7020

aaccagattg aaatcgcgtc gaaggatcaa tggcaaatcc tttttaactg catattagcc 7080

attggagcct ggtgtataga gggggaatct actgatatag atgtttttta ctatcaaaat 7140

gctaaatctc atttgacgag caaggtcttc gagtcaggtt ccataatttt ggtgacagcc 7200

ctacatcttc tgtcgcgata tacacagtgg aggcagaaaa caaatactag ctataatttt 7260

cacagctttt ccataagaat ggccatatca ttgggcttga atagggacct cccctcgtcc 7320

ttcagtgata gcagcattct ggaacaaaga cgccgaattt ggtggtctgt ctactcttgg 7380

gagatccaat tgtccctgct ttatggtcga tccatccagc tttctcagaa tacaatctcc 7440

ttcccttctt ctgtcgacga tgtgcagcgt accacaacag gtcccaccat atatcatggc 7500

atcattgaaa cagcaaggct cttacaagtt ttcacaaaaa tctatgaact agacaaaaca 7560

gtaactgcag aaaaaagtcc tatatgtgca aaaaaatgct tgatgatttg taatgagatt 7620

gaggaggttt cgagacaggc accaaagttt ttacaaatgg atatttccac caccgctcta 7680

accaatttgt tgaaggaaca cccttggcta tcctttacaa gattcgaact gaagtggaaa 7740

cagttgtctc ttatcattta tgtattaaga gatttttca ctaattttac ccagaaaaag 7800

tcacaactag aacaggatca aaatgatcat caaagttatg aagttaaacg atgctccatc 7860

atgttaagcg atgcagcaca aagaactgtt atgtctgtaa gtagctatat ggacaatcat 7920

aatgtcaccc catattttgc ctggaattgt tcttattact tgttcaatgc agtcctagta 7980

cccataaaga ctctactctc aaactcaaaa tcgaatgctg agaataacga gaccgcacaa 8040

ttattacaac aaattaacac tgttctgatg ctattaaaaa aactggccac ttttaaaatc 8100

cagacttgtg aaaaatacat tcaagtactg gaagaggtat gtgcgccgtt tctgttatca 8160

cagtgtgcaa tcccattacc gcatatcagt tataacaata gtaatggtag cgccattaaa 8220

aatattgtcg gttctgcaac tatcgcccaa taccctactc ttccggagga aaatgtcaac 8280

aatatcagtg ttaaatatgt ttctcctggc tcagtagggc cttcacctgt gccattgaaa 8340

tcaggagcaa gtttcagtga tctagtcaag ctgttatcta accgtccacc ctctcgtaac 8400

tctccagtga caataccaag aagcacacct tcgcatcgct cagtcacgcc ttttctaggg 8460

caacagcaac agctgcaatc attagtgcca ctgaccccgt ctgctttgtt tggtggcgcc 8520

aattttaatc aaagtgggaa tattgctgat agctcattgt ccttcacttt cactaacagt 8580

agcaacggtc cgaacctcat aacaactcaa acaaattctc aagcgctttc acaaccaatt 8640

gcctcctcta acgttcatga taacttcatg aataatgaaa tcacggctag taaaattgat 8700

gatggtaata attcaaaacc actgtcacct ggttggacgg accaaactgc gtataacgcg 8760

tttggaatca ctacagggat gtttaatacc actacaatgg atgatgtata taactatcta 8820

ttcgatgatg aagatacccc accaaaccca aaaaaagagt aaaatgaatc gtagatactg 8880

aaaaaccccg caagttcact tcaactgtgc atcgtgcacc atctcaattt ctttcattta 8940

tacatcgttt tgccttcttt tatgtaacta tactcctcta agtttcaatc ttggccatgt 9000

aacctctgat ctatagaatt ttttaaatga ctagaattaa tgcccatctt ttttttggac 9060

ctaaattctt catgaaaata tattacgagg gcttattcag aagcttcgct caaaggcgcg 9120

ccacggtcgt gcggataaag attctctttt tttatgatat ttgtacataa actttataaa 9180

tgaaattcat aatagaaacg acacgaaatt acaaaatgga atatgttcat agggtagacg 9240

aaactatata cgcaatctac atacatttat caagaaggag aaaaaggagg atgtaaagga 9300

atacaggtaa gcaaattgat actaatggct caacgtgata aggaaaaaga attgcacttt 9360

aacattaata ttgacaagga ggagggcacc acacaaaaag ttaggtgtaa cagaaaatca 9420

tgaaactatg attcctaatt tatatattgg aggattttct ctaaaaaaaa aaaaatacaa 9480

caaataaaaa acactcaatg acctgaccat ttgatggagt ttaagtcaat accttcttga 9540

accatttccc ataatggtga aagttccctc aagaatttta ctctgtcaga aacggcctta 9600

acgacgtagt cgacctcctc ttcagtacta aatctacggt gtttaaaccc cagcgcctgg 9660

cggg 9664<210> 114

<211> 4380

<212> DNA

<213> Artificial Sequence

<220>

<223> Phase III-A stitch product

<400> 114

gacggcacgg ccacgcgttt aaaccgccaa aagtgcagct cagagccccc agcaccagta 60

ttagaggtca taatgggctg cgaagcctgc taaaatgcag tggaggccgt gtacccttg 120

ccaaattggc tattggaatc ggcagagaac ctgggtcccg ttctagagac cctgcgagcg 180

tgtcccggtg ggttctggga gctctaactc cgcaggaact acaaaccttg cttacacaga 240

gtgaacctgc tgcctggcgt gctctgactc agtacatttc atagcccatc ttcaacaaca 300

ataccgactt accatcctat ttgctttgcc ctttttcttt tccactgcac tttgcatcgg 360

aaggcgttat cggttttggg tttagtgcct aaacgagcag cgagaacacg accacgggct 420

atataaatgg aaagttagga caggggcaaa gaataagagc acagaagaag agaaaagacg 480

aaaagcagaa gcggaaaacg tatacacgtc acatatcaca cacacacgct cacacgcggc 540

caggggggagc ctcgacacta gtaatacaca tcatcgtcct acaagttcat caaagtgttg 600

gacagacaac tataccagca tggatctctt gtatcggttc ttttctcccg ctctctcgca 660
ataacaatga acactgggtc aatcatagcc tacacaggtg aacagagtag cgtttataca 720
gggtttatac ggtgattcct acggcaaaaa tttttcattt ctaaaaaaaa aaagaaaaat 780
ttttctttcc aacgctagaa ggaaaagaaa aatctaatta aattgatttg gtgattttct 840
gagagttccc tttttcatat atcgaatttt gaatataaaa ggagatcgaa aaaatttttc 900
tattcaatct gttttctggt tttatttgat agtttttttg tgtattatta ttatggatta 960
gtactggttt atatgggttt ttctgtataa cttcttttta ttttagtttg tttaatctta 1020
ttttgagtta cattatagtt ccctaactgc aagagaagta acattaaaaa tgggtaagga 1080

aaagactcac gtttcgaggc cgcgattaaa ttccaacatg gatgctgatt tatatgggta 1140
taaatgggct cgcgataatg tcgggcaatc aggtgcgaca atctatcgat tgtatgggaa 1200
gcccgatgcg ccagagttgt ttctgaaaca tggcaaaggt agcgttgcca atgatgttac 1260
agatgagatg gtcagactaa actggctgac ggaatttatg cctcttccga ccatcaagca 1320
ttttatccgt actcctgatg atgcatggtt actcaccact gcgatccccg gcaaaacagc 1380
attccaggta ttagaagaat atcctgattc aggtgaaaat attgttgatg cgctggcagt 1440
gttcctgcgc cggttgcatt cgattcctgt ttgtaattgt ccttttaaca gcgatcgcgt 1500
atttcgtctc gctcaggcgc aatcacgaat gaataacggt ttggttgatg cgagtgattt 1560

tgatgacgag cgtaatggct ggcctgttga acaagtctgg aaagaaatgc ataagctttt 1620
gccattctca ccggattcag tcgtcactca tggtgatttc tcacttgata accttatttt 1680
tgacgagggg aaattaatag gttgtattga tgttggacga gtcggaatcg cagaccgata 1740
ccaggatctt gccatcctat ggaactgcct cggtgagttt tctccttcat tacagaaacg 1800
gctttttcaa aaatatggta ttgataatcc tgatatgaat aaattgcagt ttcatttgat 1860
gctcgatgag tttttctaag tttaacttga tactactaga ttttttctct tcatttataa 1920
aatttttggt tataattgaa gctttagaag tatgaaaaaa tccttttttt tcattctttg 1980
caaccaaaat aagaagcttc ttttattcat tgaaatgatg aatataaacc taacaaaaga 2040

aaaagactcg aatatcaaac attaaaaaaa aataaaagag gttatctgtt ttcccattta 2100
gttggagttt gcattttcta atagatagaa ctctcaatta atgtggattt agtttctctg 2160
ttcgtttttt tttgttttgt tctcactgta tttacatttc tatttagtat ttagttattc 2220
atataatctt aacttctcga ggagctccgc tcgtccaacg ccggcggacc tcttgtgcta 2280
agtggtgctg ttagacagct acgaataagg aaattccgaa gcatgtaggg aggtcatgat 2340
atgaaaaagc aaaagagtag gcatcaaaaa gtttctcatt caagtggtaa ctgctgttaa 2400

aattaagata tttataaatt gaagcttggt cgttccgacc aataccgtag ggaaacgtaa 2460

attagctatt gtaaaaaaag gaaaagaaaa gaaaagaaaa atgttacata tcgaattgat 2520

cttattcctt tggtagacca gtctttgcgt caatcaaaga ttcgtttgtt tcttgtgggc 2580

ctgaaccgac ttgagttaaa atcactctgg caacatcctt ttgcaactca agatccaatt 2640

cacgtgcagt aaagttagat gattcaaatt gatggttgaa agcctcaagc tgctcagtag 2700

taaatttctt gtcccatcca ggaacagagc caaacaattt atagataaat gcaaagagtt 2760

tcgactcatt ttcagctaag tagtacaaca cagcatttgg acctgcatca aacgtgtatg 2820

caacgattgt ttctccgtaa aactgattaa tggtgtggca ccaactgatg atacgcttgg 2880

aagtgtcatt catgtagaat attggaggga aagagtccaa acatgtggca tggaaagagt 2940

tggaatccat cattgtttcc tttgcaaagg tggcgaaatc tttttcaaca atggctttac 3000

gcatgacttc aaatctcttt ggtacgacat gttcaattct ttctttaaat agttcggagg 3060

ttgccacggt caattgcata ccctgagtgg aactcacatc ctttttaata tcgctgacaa 3120

ctaggacaca agctttcatc tgaggccagt cagagctgtc tgcgatttgt actgccatgg 3180

aatcatgacc atcttcagct tttcccattt cccaggccac gtatccgcca aacaacgatc 3240

tacaagctga accagacccc tttcttgcta ttctagatat ttctgaagtt gactgtggta 3300

attggtataa cttagcaatt gcagagacca atgcagcaaa gccagcagcg gaggaagcta 3360

aaccagctgc tgtaggaaag ttattttcgg agacaatgtg gagtttccat tgagataatg 3420

tgggcaatga ggcgtccttc gattccattt cctttcttaa ttggcgtagg tcgcgcagac 3480

aattttgagt tctttcattg tcgatgctgt gtggttctcc atttaaccac aaagtgtcgc 3540

gttcaaactc aggtgcagta gccgcagagg tcaacgttct gaggtcatct tgcgataaag 3600

tcactgatat ggacgaattg gtgggcagat tcaacttcgt gtccctttc ccccaatact 3660

taagggttgc gatgttgacg ggtgcggtaa cggatgctgt gtaaacggtc atatccccgc 3720

gtgcttggcc ggccgttact tttttttgg atggacgcaa agaagtttaa taatcatatt 3780

acatggcaat accaccatat acatatccat atctaatctt acttatatgt tgtggaaatg 3840

taaagagccc cattatctta gcctaaaaaa accttctctt tggaactttc agtaatacgc 3900

ttaactgctc attgctatat tgaagtacgg attagaagcc gccgagcggg cgacagccct 3960

ccgacggaag actctcctcc gtgcgtcctg gtcttcaccg gtcgcgttcc tgaaacgcag 4020

atgtgcctcg cgccgcactg ctccgaacaa taaagattct acaatactag cttttatggt 4080

tatgaagagg aaaaattggc agtaacctgg ccccacaaac cttcaaatca acgaatcaaa 4140

ttaacaacca taggataata atgcgattag ttttttagcc ttatttctgg ggtaattaat 4200

cagcgaagcg atgatttttg atctattaac agatatataa atgcaaaagc tgcataacca 4260

ctttaactaa tactttcaac attttcggtt tgtattactt cttattcaaa tgtcataaaa 4320

gtatcaacaa aaaattgtta atatacctct atacttaacc tgcaggccgc gagcgccgat 4380

<210

> 115

<211> 4284

<212> DNA

<213> Artificial Sequence

<220>

<223> Phase III-B stitch product

<400> 115

aacctgcagg ccgcgagcgc cgatatggct tcagaaaaag aaattaggag agagagattc 60

ttgaacgttt tccctaaatt agtagaggaa ttgaacgcat cgcttttggc ttacggtatg 120

cctaaggaag catgtgactg gtatgcccac tcattgaact acaacactcc aggcggtaag 180

ttaaatagag gtttgtccgt tgtggacacg tatgctattc tctccaacaa gaccgttgaa 240

caattggggc aagaagaata cgaaaaggtt gctattctag gttggtgcat tgagttgttg 300

caggcttact tcttggtcgc cgatgatatg atggacaagt ccattaccag aagaggccaa 360

ccatgttggt acaaggttcc tgaagttggg gaaattgcca tcaatgacgc attcatgtta 420

gaggctgcta tctacaagct tttgaaatct cacttcagaa acgaaaaata ctacatagat 480

atcaccgaat tgttccatga agtcaccttc caaaccgaat tgggccaatt gatggactta 540

atcactgcac ctgaagacaa agtcgacttg agtaagttct ccctaaagaa gcactccttc 600

atagttactt tcaagactgc ttactattct ttctacttgc ctgtcgcatt ggctatgtac 660

gttgccggta tcacagatga aaaggatttg aaacaagcca gagatgtctt gattccattg 720

ggtgaatatt tccaaattca agatgactac ttagactgct tcggtacccc agaacagatc 780

ggtaagatcg gtacagatat ccaagataac aaatgttctt gggtaatcaa caaggcatta 840

gaacttgctt ccgcagaaca aagaaagact ttagacgaaa attacggtaa gaaggactca 900

gtcgcagaag ccaaatgcaa aaagattttc aatgacttga aaatcgacca gttataccac 960

gaatatgaag agtctgttgc caaggatttg aaggccaaga tctcccaagt cgacgagtct 1020

cgtggcttca aagccgacgt cttaactgcg tttttgaaca aggtttacaa gagaagtaaa 1080

taagactgga agcgttcaat cgataaaaat tggaatacag attagataag gaccatgtat 1140

aagaaatata tacttccact ataatatagt ataagcttac agataattat ctcttgattt 1200

accgttacac gtgactaaag gacgcttttt ctcagccaat gaaagtgaag aaaaacttga 1260

tcggaaatta cgggtagtac gtaaaaggaa cttgagccac cccccaaaaa tttacccata 1320

taataatagg aaaagcaacg acctcaaacg cgatcgccga cgccgccgat gacggtagca 1380

acaagaatat agcacgagcc gcgaagttca tttcgttact tttgatatcg ctcacaacta 1440

ttgcgaagcg cttcagtgaa aaaatcataa ggaaaagttg taaatattat tggtagtatt 1500

cgtttggtaa agtagagggg gtaatttttc ccctttattt tgttcataca ttcttaaatt 1560

gctttgcctc tccttttgga aagctatact tcggagcact gttgagcgaa ggctcattag 1620

atatattttc tgtcattttc cttaacccaa aaataaggga aagggtccaa aaagcgctcg 1680

gacaactgtt gaccgtgatc cgaaggactg gctatacagt gttcacaaaa tagccaagct 1740

gaaaataatg tgtagctatg ttcagttagt ttggctagca aagatataaa agcaggtcgg 1800

aaatatttat gggcattatt atgcagagca tcaacatgat aaaaaaaagg cggccgctgg 1860

cgagggagat atgactgccg acaacaatag tatgccccat ggtgcagtat ctagttacgc 1920

caaattagtg caaaaccaaa cacctgaaga cattttggaa gagtttcctg aaattattcc 1980

attacaacaa agacctaata cccgatctag tgagacgtca aatgacgaaa gcggagaaac 2040

atgtttttct ggtcatgatg aggagcaaat taagttaatg aatgaaaatt gtattgtttt 2100

ggattgggac gataatgcta ttggtgccgg taccaagaaa gtttgtcatt taatggaaaa 2160

tattgaaaag ggtttactac atcgtgcatt ctccgtcttt attttcaatg aacaaggtga 2220

attactttta caacaaagag ccactgaaaa aataactttc cctgatcttt ggactaacac 2280

atgctgctct catccactat gtattgatga cgaattaggt ttgaagggta agctagacga 2340

taagattaag ggcgctatta ctgcggcggt gagaaaacta gatcatgaat taggtattcc 2400

agaagatgaa actaagacaa ggggtaagtt tcacttttta aacagaatcc attacatggc 2460

accaagcaat gaaccatggg gtgaacatga aattgattac atcctatttt ataagatcaa 2520

cgctaaagaa aacttgactg tcaacccaaa cgtcaatgaa gttagagact tcaaatgggt 2580

ttcaccaaat gatttgaaaa ctatgtttgc tgacccaagt tacaagttta cgccttggtt 2640

taagattatt tgcgagaatt acttattcaa ctggtgggag caattagatg acctttctga 2700

agtggaaaat gacaggcaaa ttcatagaat gctataacaa cgcgtcaata atataggcta 2760

cataaaaatc ataataactt tgttatcata gcaaaatgtg atataaaacg tttcatttca 2820

cctgaaaaat agtaaaaata ggcgacaaaa atccttagta atatgtaaac tttatttttct 2880

ttatttattt acagaactct gaatatacat tgattgttca catttttttt ttctcttctc 2940

aatttccctt gattatattc aaaaggttat tggcctcttg aatgtttccc actgatgaag 3000

gcgcgccacg gtcgtgcgga tgtattccaa tgagaatcgc tagaaatgct ttaccagaac 3060

tagactactt gtcgcagatc acttttgaac tgtatgagag tacggatgct tctggtcaaa 3120

aatcgcattc cattagactg aaaatgtctc ctgggtgtca tactcgagat ccgttagatg 3180

ttcaattaga tgacaggcat tatattagtt gtattccaaa gatttccctg acgaagcatt 3240

tggatatgga ctacgttcaa cagaaattga gaaacaaatt taccagggtc attatgcctc 3300
cgaaatttac accagtaaac attacgagcc ccaacttgag tttccagaaa cgcaaaacca 3360
gaagaaagtc ggtatctgtt gagaagttga agcttcctgc ctcgtccgga tcttcatcat 3420
ctacctccgt taacaagaca ttagattagt gatcacaccc aatttttaat ttagcaaccc 3480
aaaataaata agtatttact caacttttt ttaataaaaa aaaacttaat tgaattttgc 3540
tcgcgatctt taggtccggg gttttcgttg aacccttaga cgagcaaatt agcgccataa 3600
ggatatacgt cagagcacat taattagtga catataccta tataaagagc aaccttctcc 3660
gatagacttg taatttatct tatttcattt cctaacactt tggtcgaaga agagggataa 3720

aaacagacga aagcccctca gccccctag cgtcgatggg aaagctatta caattggcat 3780
tgcatccggt cgagatgaag gcagctttga agctgaagtt ttgcagaaca ccgctattct 3840
ccatctatga tcagtccacg tctccatatc tcttgcactg tttcgaactg ttgaacttga 3900
cctccagatc gtttgctgct gtgatcagag agctgcatcc agaattgaga aactgtgtta 3960
ctctctttta tttgatttta agggctttgg ataccatcga agacgatatg tccatcgaac 4020
acgatttgaa aattgacttg ttgcgtcact tccacgagaa attgttgtta actaaatgga 4080
gtttcgacgg aaatgccccc gatgtgaagg acagagccgt tttgacagat ttcgaatcga 4140
ttcttattga attccacaaa ttgaaaccag aatatcaaga agtcatcaag gagatcaccg 4200

agaaaatggg taatggtatg gccgactaca tcttagatga aaattacaac ttgaatcggt 4260
gtttaaaccc cagcgcctgg cggg 4284

<210> 116
<211> 6344
<212> DNA
<213> Artificial Sequence
<220>
<223> Phase I marker recycling stitch product
<400> 116

gacggcacgg ccacgcgttt aaaccgccca gatggaatcc cttccataga gagaaggagc 60
aagcaactga cccaatattg actgccactg gacctgaaga catgcaacaa agtgcaagca 120
tagtggggcc ttcttccaat gctaatccgg tcactgccac tgctgctacg gaaaaccaac 180
ctaaaggtat taacttcttc actataagaa aatcacacga gcgcccggac gatgtctctg 240

tttaaatggc gcaagttttc cgctttgtaa tatatattta tacccctttc ttctctcccc 300

tgcaatataa tagtttaatt ctaatattaa taatatccta tattttcttc atttaccggc 360

gcactctcgc ccgaacgacc tcaaaatgtc tgctacattc ataataacca aaagctcata 420

actttttttt ttgaacctga atatatatac atcacatgtc actgctggtc cttgccgacc 480

agcgtataca atctcgatag ttggtttccc gttctttcca ctcccgtcgc tcacacgcgg 540

ccaggggggag cctcactatt attccataag atgatcatta gcattacgtt caaaacgagt 600

acaaataact taagtaataa cacgagccat atgaccatta gcaagatgac aagcaagtta 660

agaccaatca gcttccatca tagcatcagc ttaacgttca ccattaataa gagtagaaat 720

ttcaccttca agacaataac gattttcgtg gtaataactg atataattaa attgaagctc 780

taatttgtga gtttagtata catgcattta cttataatac agtttttag ttttgctggc 840

cgcatcttct caaatatgct tcccagcctg cttttctgta acgttcaccc tctaccttag 900

catcccttcc ctttgcaaat agtcctcttc caacaataat aatgtcagat cctgtagaga 960

ccacatcatc cacggttcta tactgttgac ccaatgcgtc tcccttgtca tctaaaccca 1020

caccgggtgt cataatcaac caatcgtaac cttcatctct tccacccatg tctctttgag 1080

caataaagcc gataacaaaa tctttgtcac tcttcgcaat gtcaacagta cccttagtat 1140

attctccagt agctagggag cccttgcatg acaattctgc taacatcaaa aggcctctag 1200

gttcctttgt tacttcttcc gccgcctgct tcaaaccgct aacaatacct gggcccacca 1260

caccgtgtgc attcgtaatg tctgcccatt ctgctattct gtatacaccc gcagagtact 1320

gcaatttgac tgtattacca atgtcagcaa attttctgtc ttcgaagagt aaaaaattgt 1380

acttggcgga taatgccttt agcggcttaa ctgtgccctc catggaaaaa tcagtcaaga 1440

tatccacatg tgtttttagt aaacaaattt tgggacctaa tgcttcaact aactccagta 1500

attccttggt ggtacgaaca tccaatgaag cacacaagtt tgtttgcttt tcgtgcatga 1560

tattaaatag cttggcagca acaggactag gatgagtagc agcacgttcc ttatatgtag 1620

ctttcgacat gatttatctt cgtttcctgc aggtttttgt tctgtgcagt tgggttaaga 1680

atactgggca atttcatgtt tcttcaacac cacatatgcg tatatatacc aatctaagtc 1740

tgtgctcctt ccttcgttct tccttctgct cggagattac cgaatcaaaa aaatttcaaa 1800

gaaaccggaa tcaaaaaaaa gaacaaaaaa aaaaaagatg aattgaaaag ctttatggac 1860

cctgaaacca ctcactatta ttccataaga tgatcattag cattacgttc aaaacgagta 1920

caaataactt aagtaataac acgagccata tgaccattag caagatgaca agcaagttaa 1980

gaccaatcag cttccatcat agcatcagct taacgttcac cattaataag agtagaaatt 2040

tcaccttcaa gacaataacg attttcgtgc gctcgtccaa cgccggcgga cctgacggta 2100

gcaacaagaa tatagcacga gccgcggagt tcatttcgtt acttttgata tcactcacaa 2160

ctattgcgaa gcgcttcagt gaaaaaatca taaggaaaag ttgtaaatat tattggtagt 2220

attcgtttgg taaagtagag ggggtaattt ttccccttta ttttgttcat acattcttaa 2280

attgctttgc ctctcctttt ggaaagctat acttcggagc actgttgagc gaaggctcat 2340

tagatatatt ttctgtcatt ttccttaacc caaaaataag ggaaagggtc caaaaagcgc 2400

tcggacaact gttgaccgtg atccgaagga ctggctatac agtgttcaca aaatagccaa 2460

gctgaaaata atgtgtagct atgttcagtt agtttggcta gcaaagatat aaaagcaggt 2520

cggaaatatt tatgggcatt attatgcaga gcatcaacat gataaaaaaa aacagttgaa 2580

tattccctca aaaatccccg cgtgcttggc cggccgtaat taataatgtc aactttgcct 2640

atttcttctg tgtcattttc ctcttctaca tcaccattag tcgtggacga caaagtctca 2700

accaagcccg acgttatcag acatacaatg aatttcaatg cttctatttg gggagatcaa 2760

ttcttgacct atgatgagcc tgaagattta gttatgaaga aacaattagt ggaggaatta 2820

aaagaggaag ttaagaagga attgataact atcaaaggtt caaatgagcc catgcagcat 2880

gtgaaattga ttgaattaat tgatgctgtt caacgtttag gtatagctta ccattttgaa 2940

gaagagatcg aggaagcttt gcaacatata catgttacct atggtgaaca gtgggtggat 3000

aaggaaaatt tacagagtat ttcattgtgg ttcaggttgt tgcgtcaaca gggctttaac 3060

gtctcctctg gcgttttcaa agactttatg gacgaaaaag gtaaattcaa agagtcttta 3120

tgcaatgatg cacaaggaat attagcctta tatgaagctg catttatgag ggttgaagat 3180

gaaaccatct tagacaatgc tttggaattc acaaaagttc atttagatat catagcaaaa 3240

gacccatctt gcgattcttc attgcgtaca caaatccatc aagccttaaa acaaccttta 3300

agaaggagat tagcaaggat tgaagcatta cattacatgc caatctacca acaggaaaca 3360

tctcatgatg aagtattgtt gaaattagcc aagttggatt tcagtgtttt gcagtctatg 3420

cataaaaagg aattgtcaca tatctgtaag tggtggaaag atttagattt acaaaataag 3480

ttaccttatg tacgtgatcg tgttgtcgaa ggctacttct ggatattgtc catatactat 3540

gagccacaac acgctagaac aagaatgttt ttgatgaaaa catgcatgtg gttagtagtt 3600

ttggacgata cttttgataa ttatggaaca tacgaagaat tggagatttt tactcaagcc 3660

gtcgagagat ggtctatctc atgcttagat atgttgcccg aatatatgaa attaatctac 3720

caagaattag tcaatttgca tgtggaaatg gaagaatctt tggaaaagga gggaaagacc 3780

tatcagattc attacgttaa ggagatggct aaagaattag ttcgtaatta cttagtagaa 3840

gcaagatggt tgaaggaagg ttatatgcct actttagaag aatacatgtc tgtttctatg 3900

gttactggta cttatggttt gatgattgca aggtcctatg ttggcagagg agacattgtt 3960

actgaagaca cattcaaatg ggtttctagt tacccaccta ttattaaagc ttcctgtgta 4020

atagtaagat taatggacga tattgtatct cacaaggaag aacaagaaag aggacatgtg 4080

gcttcatcta tagaatgtta ctctaaagaa tcaggtgctt ctgaagagga agcatgtgaa 4140
tatattagta ggaaagttga ggatgcctgg aaagtaatca atagagaatc tttgcgtcca 4200
acagccgttc ccttcccttt gttaatgcca gcaataaact tagctagaat gtgtgaggtc 4260
ttgtactctg ttaatgatgg ttttactcat gctgagggtg acatgaaatc ttatatgaag 4320
tccttcttcg ttcatcctat ggtcgtttga gctagctaag atccgctcta accgaaaagg 4380
aaggagttag acaacctgaa gtctaggtcc ctatttattt ttttatagtt atgttagtat 4440
taagaacgtt atttatattt caaatttttc ttttttttct gtacagacgc gtgtacgcat 4500
gtaacattat actgaaaacc ttgcttgaga aggttttggg acgctcgaag aacctgcagg 4560

ccgcgagcgc cgatagttat gacaattaca acaacagaat tctttctata tatgcacgaa 4620
cttgtaatat ggaagaaatt atgacgtaca aactataaag taaatatttt acgtaacaca 4680
tggtgctgtt gtgcttcttt ttcaagagaa taccaatgac gtatgactaa gtttaggatt 4740
taatgcaggt gacggaccca tctttcaaac gatttatatc agtggcgtcc aaattgttag 4800
gttttgttgg ttcagcaggt ttcctgttgt gggtcatatg actttgaacc aaatggccgg 4860
ctgctagggc agcacataag gataattcac ctgccaagac ggcacaggca actattcttg 4920
ctaattgacg tgcgttggta ccaggagcgg tagcatgtgg gcctcttaca cctaataagt 4980
ccaacatggc accttgtggt tctagaacag taccaccacc gatggtacct acttcgatgg 5040

atggcatgga tacggaaatt ctcaaatcac cgtccacttc tttcatcaat gttatacagt 5100
tggaactttc gacattttgt gcaggatctt gtcctaatgc caagaaaaca gctgtcacta 5160
aattagctgc atgtgcgtta aatccaccaa cagacccagc cattgcagat ccaaccaaat 5220
tcttagcaat gttcaactca accaatgcgg aaacatcact ttttaacact tttctgacaa 5280
catcaccagg aatagtagct tctgcgacga cactcttacc acgaccttcg atccagttga 5340
tggcagctgg ttttttgtcg gtacagtagt taccagaaac ggagacaacc tccatatctt 5400
cccagccata ctcttctacc atttgcttta atgagtattc gacaccctta gaaatcatat 5460
tcatacccat tgcgtcacca gtagttgttc taaatctcat gaagagtaaa tctcctgcta 5520

gacaagtttg aatatgttgc agacgtgcaa atcttgatgt agagttaaaa gctttttaa 5580
ttgcgttttg tccctcttct gagtctaacc atatcttaca ggcaccagat cttttcaaag 5640
ttgggaaacg gactactggg cctcttgtca taccatcctt agttaaaaca gttgttgcac 5700
caccgccagc attgattgcc ttacagccac gcatggcaga agctaccaaa caaccctctg 5760
tagttgccat tggtatatga taagatgtac catcgataac caaggggcct ataacaccaa 5820
cgggcaaagg catgtaacct ataacatttt cacaacaagc gccaaatacg cggtcgtagt 5880

cataattttt atatggtaaa cgatcagatg ctaatacagg agcttctgcc aaaattgaaa 5940

gagccttcct acgtaccgca accgctctcg tagtatcacc taatttttc tccaaagcgt 6000

acaaaggtaa cttaccgtga ataaccaagg cagcgacctc tttgttcttc aattgttttg 6060

tatttccact acttaataat gcttctaatt cttctaaagg acgtattttc ttatccaagc 6120

tttcaatatc gcgggaatca tcttcctcac tagatgatga aggtcctgat gagctcgatt 6180

gcgcagatga taaacttttg actttcgatc cagaaatgac tgttttattg gttaaaactg 6240

gtgtagaagc cttttgtaca ggagcagtaa aagacttctt ggtgacttca gtcttcacca 6300

attggtctgc agccatcggt gtttaaaccc cagcgcctgg cggg 6344<210> 117

<211> 6184

<212> DNA

<213> Artificial Sequence

<220>

<223

> Phase II marker recycling stitch product

<400> 117

gacggcacgg ccacgcgttt aaaccgccta ggataattat actctatttc tcaacaagta 60

attggttgtt tggccgagcg gtctaaggcg cctgattcaa gaaatatctt gaccgcagtt 120

aactgtggga atactcaggt atcgtaagat gcaagagttc gaatctctta gcaaccatta 180

ttttttcct caacataacg agaacacaca ggggcgctat cgcacagaat caaattcgat 240

gactggaaat tttttgttaa tttcagaggt cgcctgacgc atataccttt ttcaactgaa 300

aaattgggag aaaaaggaaa ggtgagagcg ccggaaccgg cttttcatat agaatagaga 360

agcgttcatg actaaatgct tgcatcacaa tacttgaagt tgacaatatt atttaaggac 420

ctattgtttt ttccaatagg tggttagcaa tcgtcttact ttctaacttt tcttaccttt 480

tacatttcag caatatatat atatatattt caaggatata ccattctagc tcacacgcgg 540

ccaggggggag cctcactatt attccataag atgatcatta gcattacgtt caaaacgagt 600

acaaataact taagtaataa cacgagccat atgaccatta gcaagatgac aagcaagtta 660

agaccaatca gcttccatca tagcatcagc ttaacgttca ccattaataa gagtagaaat 720

ttcaccttca agacaataac gattttcgtg gtaataactg atataattaa attgaagctc 780

taatttgtga gtttagtata catgcattta cttataatac agtttttag ttttgctggc 840

cgcatcttct caaatatgct tcccagcctg cttttctgta acgttcaccc tctaccttag 900

catcccttcc ctttgcaaat agtcctcttc caacaataat aatgtcagat cctgtagaga 960

ccacatcatc cacggttcta tactgttgac ccaatgcgtc tcccttgtca tctaaaccca 1020

caccgggtgt cataatcaac caatcgtaac cttcatctct tccacccatg tctctttgag 1080

caataaagcc gataacaaaa tctttgtcac tcttcgcaat gtcaacagta cccttagtat 1140

attctccagt agctagggag cccttgcatg acaattctgc taacatcaaa aggcctctag 1200

gttccttgt tacttcttcc gccgcctgct tcaaaccgct aacaatacct gggcccacca 1260

caccgtgtgc attcgtaatg tctgcccatt ctgctattct gtatacaccc gcagagtact 1320

gcaatttgac tgtattacca atgtcagcaa attttctgtc ttcgaagagt aaaaaattgt 1380

acttggcgga taatgccttt agcggcttaa ctgtgccctc catggaaaaa tcagtcaaga 1440

tatccacatg tgtttttagt aaacaaattt tgggacctaa tgcttcaact aactccagta 1500

attccttggt ggtacgaaca tccaatgaag cacacaagtt tgtttgcttt tcgtgcatga 1560

tattaaatag cttggcagca acaggactag gatgagtagc agcacgttcc ttatatgtag 1620

ctttcgacat gatttatctt cgtttcctgc aggtttttgt tctgtgcagt tgggttaaga 1680

atactgggca atttcatgtt tcttcaacac cacatatgcg tatatatacc aatctaagtc 1740

tgtgctcctt ccttcgttct tccttctgct cggagattac cgaatcaaaa aaatttcaaa 1800

gaaaccggaa tcaaaaaaaa gaacaaaaaa aaaaaagatg aattgaaaag ctttatggac 1860

cctgaaacca ctcactatta ttccataaga tgatcattag cattacgttc aaaacgagta 1920

caaataactt aagtaataac acgagccata tgaccattag caagatgaca agcaagttaa 1980

gaccaatcag cttccatcat agcatcagct taacgttcac cattaataag agtagaaatt 2040

tcaccttcaa gacaataacg attttcgtgc gctcgtccaa cgccggcgga cctgacggta 2100

gcaacaagaa tatagcacga gccgcggagt tcatttcgtt acttttgata tcactcacaa 2160

ctattgcgaa gcgcttcagt gaaaaaatca taaggaaaag ttgtaaatat tattggtagt 2220

attcgtttgg taaagtagag ggggtaattt ttccccttta ttttgttcat acattcttaa 2280

attgctttgc ctctcctttt ggaaagctat acttcggagc actgttgagc gaaggctcat 2340

tagatatatt ttctgtcatt ttccttaacc caaaaataag ggaaagggtc caaaaagcgc 2400

tcggacaact gttgaccgtg atccgaagga ctggctatac agtgttcaca aaatagccaa 2460

gctgaaaata atgtgtagct atgttcagtt agtttggcta gcaaagatat aaaagcaggt 2520

cggaaatatt tatgggcatt attatgcaga gcatcaacat gataaaaaaa aacagttgaa 2580

tattccctca aaaatccccg cgtgcttggc cggccgtaat taataatgtc aactttgcct 2640

atttcttctg tgtcattttc ctcttctaca tcaccattag tcgtggacga caaagtctca 2700

accaagcccg acgttatcag acatacaatg aatttcaatg cttctatttg gggagatcaa 2760

ttcttgacct atgatgagcc tgaagattta gttatgaaga aacaattagt ggaggaatta 2820

aaagaggaag ttaagaagga attgataact atcaaaggtt caaatgagcc catgcagcat 2880

gtgaaattga ttgaattaat tgatgctgtt caacgtttag gtatagctta ccattttgaa 2940

gaagagatcg aggaagcttt gcaacatata catgttacct atggtgaaca gtgggtggat 3000

aaggaaaatt tacagagtat ttcattgtgg ttcaggttgt tgcgtcaaca gggctttaac 3060

gtctcctctg gcgttttcaa agactttatg gacgaaaaag gtaaattcaa agagtcttta 3120

tgcaatgatg cacaaggaat attagcctta tatgaagctg catttatgag ggttgaagat 3180

gaaaccatct tagacaatgc tttggaattc acaaaagttc atttagatat catagcaaaa 3240

gacccatctt gcgattcttc attgcgtaca caaatccatc aagccttaaa acaaccttta 3300

agaaggagat tagcaaggat tgaagcatta cattacatgc caatctacca acaggaaaca 3360

tctcatgatg aagtattgtt gaaattagcc aagttggatt tcagtgtttt gcagtctatg 3420

cataaaaagg aattgtcaca tatctgtaag tggtggaaag atttagattt acaaaataag 3480

ttaccttatg tacgtgatcg tgttgtcgaa ggctacttct ggatattgtc catatactat 3540

gagccacaac acgctagaac aagaatgttt ttgatgaaaa catgcatgtg gttagtagtt 3600

ttggacgata cttttgataa ttatggaaca tacgaagaat tggagatttt tactcaagcc 3660

gtcgagagat ggtctatctc atgcttagat atgttgcccg aatatatgaa attaatctac 3720

caagaattag tcaatttgca tgtggaaatg gaagaatctt tggaaaagga gggaaagacc 3780

tatcagattc attacgttaa ggagatggct aaagaattag ttcgtaatta cttagtagaa 3840

gcaagatggt tgaaggaagg ttatatgcct actttagaag aatacatgtc tgtttctatg 3900

gttactggta cttatggttt gatgattgca aggtcctatg ttggcagagg agacattgtt 3960

actgaagaca cattcaaatg ggtttctagt tacccaccta ttattaaagc ttcctgtgta 4020

atagtaagat taatggacga tattgtatct cacaaggaag aacaagaaag aggacatgtg 4080

gcttcatcta tagaatgtta ctctaaagaa tcaggtgctt ctgaagagga agcatgtgaa 4140

tatattagta ggaaagttga ggatgcctgg aaagtaatca atagagaatc tttgcgtcca 4200

acagccgttc ccttcccttt gttaatgcca gcaataaact tagctagaat gtgtgaggtc 4260

ttgtactctg ttaatgatgg ttttactcat gctgagggtg acatgaaatc ttatatgaag 4320

tccttcttcg ttcatcctat ggtcgtttga gctagctaag atccgctcta accgaaaagg 4380

aaggagttag acaacctgaa gtctaggtcc ctatttattt ttttatagtt atgttagtat 4440

taagaacgtt atttatattt caaatttttc tttttttttct gtacagacgc gtgtacgcat 4500

gtaacattat actgaaaacc ttgcttgaga aggttttggg acgctcgaag aacctgcagg 4560

ccgcgagcgc cgtaattcgc gggtggaagg accttgtgga ggaatatgaa gttgatagct 4620

caaagggatt gaatggctat ttaactgatt atgagtcaat gtatcaagga tactatggtc 4680

tgcttaaatt tcattctgtc ttcgaaagct gaattgatac tacgaaaaat tttttttttgt 4740

ttctctttct atctttatta cataaaactt catacacagt taagattaaa aacaactaat 4800

aaataatgcc tatcgcaaat tagcttatga agtccatggt aaattcgtgt ttcctggcaa 4860

taatagatcg tcaatttgtt gctttgtggt agttttattt tcaaataatt ggaatactag 4920

ggatttgatt ttaagatctt tattcaaatt ttttgcgctt aacaaacagc agccagtccc 4980

acccaagtct gtttcaaatg tctcgtaact aaaatcatct tgcaatttct ttttgaaact 5040

gtcaatttgc tcttgagtaa tgtctcttcg taacaaagtc aaagagcaac cgccgccacc 5100

agcaccggta agttttgtgg agccaattct caaatcatcg ctcagatttt taataagttc 5160

taatccagga tgagaaacac cgattgagac aagcagtcca tgatttattc ttatcaattc 5220

caatagttgt tcatacagtt cattattagt ttctacagcc tcgtcatcgg tgcctttaca 5280

tttacttaac ttagtcatga tctctaagcc ttgtagggca cattcaccca tggcatctag 5340

aattggcttc ataacttcag gaaatttctc ggtgaccaac acacgaacgc gagcaacaag 5400

atcttttgta gaccttggaa ttctagtata ggttaggatc attggaatgg ctgggaaatc 5460

atctaagaac ttaaaattgt ttgtgtttat tgttccatta tgtgagtctt tttcaaatag 5520

cagggcatta ccataagtgg ccacagcgtt atctattcct gaaggggtac cgtgaataca 5580

cttttcacct atgaaggccc attgattcac tatatgctta tcgttttctg acagcttttc 5640

caagtcatta gatcctatta acccccccaa gtaggccata gctaaggcca gtgatacaga 5700

aatagaggcg cttgagccca acccagcacc gatgggtaaa gtagacttta aagaaaactt 5760

aatattcttg gcatgggggc ataggcaaac aaacatatac aggaaacaaa acgctgcatg 5820

gtagtggaag gattcggata gttgagctaa caacggatcc aaaagactaa cgagttcctg 5880

agacaagcca tcggtggctt gttgagcctt ggccaatttt tgggagttta cttgatcctc 5940

ggtgatggca ttgaaatcat tgatggacca cttatgatta aagctaatgt ccgggaagtc 6000

caattcaata gtatctggtg cagatgactc gcttattagc aggtaggttc tcaacgcaga 6060

cacactagca gcgacggcag gcttgttgta cacagcagag tgttcaccaa aaataataac 6120

ctttcccggt gcagaagtta agaacggtaa tgacatcggt gtttaaaccc cagcgcctgg 6180

cggg 6184<210> 118

<211> 6042

<212> DNA

<213> Artificial Sequence

<220>

<223> Phase III marker recycling stitch product

<400> 118

gacggcacgg ccacgcgttt aaaccgccaa aagtgcagct cagagccccc agcaccagta 60

ttagaggtca taatgggctg cgaagcctgc taaaatgcag tggaggccgt gtacccttg 120

ccaaattggc tattggaatc ggcagagaac ctgggtcccg ttctagagac cctgcgagcg 180

tgtcccggtg ggttctggga gctctaactc cgcaggaact acaaaccttg cttacacaga 240

gtgaacctgc tgcctggcgt gctctgactc agtacatttc atagcccatc ttcaacaaca 300

ataccgactt accatcctat ttgctttgcc ctttttcttt tccactgcac tttgcatcgg 360

aaggcgttat cggttttggg tttagtgcct aaacgagcag cgagaacacg accacgggct 420

atataaatgg aaagttagga caggggcaaa gaataagagc acagaagaag agaaaagacg 480

aaaagcagaa gcggaaaacg tatacacgtc acatatcaca cacacacgct cacacgcggc 540

caggggagc ctcactatta ttccataaga tgatcattag cattacgttc aaaacgagta 600

caaataactt aagtaataac acgagccata tgaccattag caagatgaca agcaagttaa 660

gaccaatcag cttccatcat agcatcagct taacgttcac cattaataag agtagaaatt 720

tcaccttcaa gacaataacg attttcgtgg taataactga tataattaaa ttgaagctct 780

aatttgtgag tttagtatac atgcatttac ttataataca gttttttagt tttgctggcc 840

gcatcttctc aaatatgctt cccagcctgc ttttctgtaa cgttcaccct ctaccttagc 900

atcccttccc tttgcaaata gtcctcttcc aacaataata atgtcagatc ctgtagagac 960

cacatcatcc acggttctat actgttgacc caatgcgtct cccttgtcat ctaaacccac 1020

accgggtgtc ataatcaacc aatcgtaacc ttcatctctt ccacccatgt ctctttgagc 1080

aataaagccg ataacaaaat ctttgtcact cttcgcaatg tcaacagtac ccttagtata 1140

ttctccagta gctagggagc ccttgcatga caattctgct aacatcaaaa ggcctctagg 1200

ttcctttgtt acttcttccg ccgcctgctt caaaccgcta acaatacctg ggcccaccac 1260

accgtgtgca ttcgtaatgt ctgcccattc tgctattctg tatacacccg cagagtactg 1320

caatttgact gtattaccaa tgtcagcaaa ttttctgtct tcgaagagta aaaaattgta 1380

cttggcggat aatgccttta gcggcttaac tgtgccctcc atggaaaaat cagtcaagat 1440

atccacatgt gtttttagta aacaaatttt gggacctaat gcttcaacta actccagtaa 1500

ttccttggtg gtacgaacat ccaatgaagc acacaagttt gtttgctttt cgtgcatgat 1560

attaaatagc ttggcagcaa caggactagg atgagtagca gcacgttcct tatatgtagc 1620

tttcgacatg atttatcttc gtttcctgca ggtttttgtt ctgtgcagtt gggttaagaa 1680

tactgggcaa tttcatgttt cttcaacacc acatatgcgt atatatacca atctaagtct 1740

gtgctccttc cttcgttctt ccttctgctc ggagattacc gaatcaaaaa aatttcaaag 1800

aaaccggaat caaaaaaaag aacaaaaaaa aaaaagatga attgaaaagc tttatggacc 1860

ctgaaaccac tcactattat tccataagat gatcattagc attacgttca aaacgagtac 1920

aaataactta agtaataaca cgagccatat gaccattagc aagatgacaa gcaagttaag 1980

accaatcagc ttccatcata gcatcagctt aacgttcacc attaataaga gtagaaattt 2040

caccttcaag acaataacga ttttcgtgcg ctcgtccaac gccggcggac ctgacggtag 2100

caacaagaat atagcacgag ccgcggagtt catttcgtta cttttgatat cactcacaac 2160

tattgcgaag cgcttcagtg aaaaaatcat aaggaaaagt tgtaaatatt attggtagta 2220

ttcgtttggt aaagtagagg gggtaatttt tccccttttat tttgttcata cattcttaaa 2280

ttgctttgcc tctccttttg gaaagctata cttcggagca ctgttgagcg aaggctcatt 2340

agatatattt tctgtcattt tccttaaccc aaaaataagg gaaagggtcc aaaaagcgct 2400

cggacaactg ttgaccgtga tccgaaggac tggctataca gtgttcacaa aatagccaag 2460

ctgaaaataa tgtgtagcta tgttcagtta gtttggctag caaagatata aaagcaggtc 2520

ggaaatattt atgggcatta ttatgcagag catcaacatg ataaaaaaaa acagttgaat 2580

attccctcaa aaatccccgc gtgcttggcc ggccgtaatt aataatgtca actttgccta 2640

tttcttctgt gtcattttcc tcttctacat caccattagt cgtggacgac aaagtctcaa 2700

ccaagcccga cgttatcaga catacaatga atttcaatgc ttctatttgg ggagatcaat 2760

tcttgaccta tgatgagcct gaagatttag ttatgaagaa acaattagtg gaggaattaa 2820

aagaggaagt taagaaggaa ttgataacta tcaaaggttc aaatgagccc atgcagcatg 2880

tgaaattgat tgaattaatt gatgctgttc aacgtttagg tatagcttac cattttgaag 2940

aagagatcga ggaagctttg caacatatac atgttaccta tggtgaacag tgggtggata 3000

aggaaaattt acagagtatt tcattgtggt tcaggttgtt gcgtcaacag ggctttaacg 3060

tctcctctgg cgttttcaaa gactttatgg acgaaaaagg taaattcaaa gagtctttat 3120

gcaatgatgc acaaggaata ttagccttat atgaagctgc atttatgagg gttgaagatg 3180

aaaccatctt agacaatgct ttggaattca caaaagttca tttagatatc atagcaaaag 3240

acccatcttg cgattcttca ttgcgtacac aaatccatca agccttaaaa caacctttaa 3300

gaaggagatt agcaaggatt gaagcattac attacatgcc aatctaccaa caggaaacat 3360

ctcatgatga agtattgttg aaattagcca agttggattt cagtgttttg cagtctatgc 3420

ataaaaagga attgtcacat atctgtaagt ggtggaaaga tttagattta caaaataagt 3480

taccttatgt acgtgatcgt gttgtcgaag gctacttctg gatattgtcc atatactatg 3540

agccacaaca cgctagaaca agaatgtttt tgatgaaaac atgcatgtgg ttagtagttt 3600

tggacgatac ttttgataat tatggaacat acgaagaatt ggagattttt actcaagccg 3660

tcgagagatg gtctatctca tgcttagata tgttgcccga atatatgaaa ttaatctacc 3720

aagaattagt caatttgcat gtggaaatgg aagaatcttt ggaaaaggag ggaaagacct 3780

atcagattca ttacgttaag gagatggcta aagaattagt tcgtaattac ttagtagaag 3840

caagatggtt gaaggaaggt tatatgccta ctttagaaga atacatgtct gtttctatgg 3900

ttactggtac ttatggtttg atgattgcaa ggtcctatgt tggcagagga gacattgtta 3960

ctgaagacac attcaaatgg gtttctagtt acccacctat tattaaagct tcctgtgtaa 4020

tagtaagatt aatggacgat attgtatctc acaaggaaga acaagaaaga ggacatgtgg 4080

cttcatctat agaatgttac tctaaagaat caggtgcttc tgaagaggaa gcatgtgaat 4140

atattagtag gaaagttgag gatgcctgga aagtaatcaa tagagaatct ttgcgtccaa 4200

cagccgttcc cttccctttg ttaatgccag caataaactt agctagaatg tgtgaggtct 4260

tgtactctgt taatgatggt tttactcatg ctgagggtga catgaaatct tatatgaagt 4320

ccttcttcgt tcatcctatg gtcgtttgag ctagctaaga tccgctctaa ccgaaaagga 4380

aggagttaga caacctgaag tctaggtccc tatttatttt tttatagtta tgttagtatt 4440

aagaacgtta tttatatttc aaatttttct tttttttctg tacagacgcg tgtacgcatg 4500

taacattata ctgaaaacct tgcttgagaa ggttttggga cgctcgaaga acctgcaggc 4560

cgcgagcgcc gatcttgtgc taagtggtgc tgttagacag ctacgaataa ggaaattccg 4620

aagcatgtag ggaggtcatg atatgaaaaa gcaaaagagt aggcatcaaa aagtttctca 4680

ttcaagtggt aactgctgtt aaaattaaga tatttataaa ttgaagcttg gtcgttccga 4740

ccaataccgt agggaaacgt aaattagcta ttgtaaaaaa aggaaaagaa aagaaaagaa 4800

aaatgttaca tatcgaattg atcttattcc tttggtagac cagtctttgc gtcaatcaaa 4860

gattcgtttg tttcttgtgg gcctgaaccg acttgagtta aaatcactct ggcaacatcc 4920

ttttgcaact caagatccaa ttcacgtgca gtaaagttag atgattcaaa ttgatggttg 4980

aaagcctcaa gctgctcagt agtaaatttc ttgtcccatc caggaacaga gccaaacaat 5040

ttatagataa atgcaaagag tttcgactca ttttcagcta agtagtacaa cacagcattt 5100

ggacctgcat caaacgtgta tgcaacgatt gtttctccgt aaaactgatt aatggtgtgg 5160

caccaactga tgatacgctt ggaagtgtca ttcatgtaga atattggagg gaaagagtcc 5220

aaacatgtgg catggaaaga gttggaatcc atcattgttt cctttgcaaa ggtggcgaaa 5280

tctttttcaa caatggcttt acgcatgact tcaaatctct ttggtacgac atgttcaatt 5340

ctttctttaa atagttcgga ggttgccacg gtcaattgca taccctgagt ggaactcaca 5400

tccttttaa tatcgctgac aactaggaca caagctttca tctgaggcca gtcagagctg 5460

tctgcgattt gtactgccat ggaatcatga ccatcttcag cttttcccat ttcccaggcc 5520

acgtatccgc caaacaacga tctacaagct gaaccagacc cctttcttgc tattctagat 5580

atttctgaag ttgactgtgg taattggtat aacttagcaa ttgcagagac caatgcagca 5640

aagccagcag cggaggaagc taaaccagct gctgtaggaa agttattttc ggagacaatg 5700

tggagtttcc attgagataa tgtgggcaat gaggcgtcct tcgattccat ttcctttctt 5760

aattggcgta ggtcgcgcag acaattttga gttctttcat tgtcgatgct gtgtggttct 5820

ccatttaacc acaaagtgtc gcgttcaaac tcaggtgcag tagccgcaga ggtcaacgtt 5880

ctgaggtcat cttgcgataa agtcactgat atggacgaat tggtgggcag attcaacttc 5940

gtgtcccttt tcccccaata cttaagggtt gcgatgttga cgggtgcggt aacggatgct 6000

gtgtaaacgg

tcatcggtgt ttaaacccca gcgcctggcg gg 6042<210> 119

<211> 5213

<212> DNA

<213> Artificial Sequence

<220>

<223> STE5 knockout stitch product

<400> 119

gacggcacgg ccacgcgttt aaaccgccac gaagtgactg acagaatact gacatcagct 60

gatttctcat agagctgttt ctctgataac acgttgtttg aacatcgaca agatgaaaat 120

ctagaagtat caagtttcct ttaaagggat atataacaga ttctaaaact gacagaaata 180

tttcgagtga agaagaagcg ttaaatattg gatctttccg cagttctact ctgatacatt 240

tttgaagtag gagagtcatt tagaaggcgt attgctcaat agtagaaagc aggcctgtgc 300

acatgaatta attaaaaaat ataaaggtag tgattagacg acacatgtcc ataggtaacc 360

tgtcataatt ttgaacaatt tcccttcttt tctttttttt ttttgggtgc ggcgatatgt 420

agcttgttaa tttacacatc atgtactttt ctgcatcaaa atatgaaagg cgatagtagc 480

taaagaaaat accgagaatt tcctcgaaaa gttgacgaca aaagaaaggc ataaaaaagt 540

aatttgaaaa tattttaaaa ctgttttaac ccatctagca tccgcgctaa aaaaggaaga 600

tacaggatac agcggaaaca acttttaagc tcacacgcgg ccaggggggag ccatgcgtcc 660

atctttacag tcctgtctta ttgttcttga tttgtgcccc gtaaaatact gttacttggt 720

tctggcgagg tattggatag ttccttttta taaaggccat gaagcttttt ctttccaatt 780

ttttttttttt cgtcattata gaaatcatta cgaccgagat tcccgggtaa taactgatat 840

aattaaattg aagctctaat ttgtgagttt agtatacatg catttactta taatacagtt 900

ttttagtttt gctggccgca tcttctcaaa tatgcttccc agcctgcttt tctgtaacgt 960

tcaccctcta ccttagcatc ccttcccttt gcaaatagtc ctcttccaac aataataatg 1020

tcagatcctg tagagaccac atcatccacg gttctatact gttgacccaa tgcgtctccc 1080

ttgtcatcta aacccacacc gggtgtcata atcaaccaat cgtaaccttc atctcttcca 1140

cccatgtctc tttgagcaat aaagccgata acaaaatctt tgtcgctctt cgcaatgtca 1200

acagtaccct tagtatattc tccagtagct agggagccct tgcatgacaa ttctgctaac 1260

atcaaaaggc ctctaggttc ctttgttact tcttccgccg cctgcttcaa accgctaaca 1320

atacctgggc ccaccacacc gtgtgcattc gtaatgtctg cccattctgc tattctgtat 1380

acacccgcag agtactgcaa tttgactgta ttaccaatgt cagcaaattt tctgtcttcg 1440

aagagtaaaa aattgtactt ggcggataat gcctttagcg gcttaactgt gccctccatg 1500

gaaaaatcag tcaagatatc cacatgtgtt tttagtaaac aaattttggg acctaatgct 1560

tcaactaact ccagtaattc cttggtggta cgaacatcca atgaagcaca caagtttgtt 1620

tgcttttcgt gcatgatatt aaatagcttg gcagcaacag gactaggatg agtagcagca 1680

cgttccttat atgtagcttt cgacatgatt tatcttcgtt tcctgcaggt ttttgttctg 1740

tgcagttggg ttaagaatac tgggcaattt catgtttctt caacaccaca tatgcgtata 1800

tataccaatc taagtctgtg ctccttcctt cgttcttcct tctgctcgga gattaccgaa 1860

tcaaaaaaat ttcaaagaaa ccggaatcaa aaaaaagaac aaaaaaaaaa aagatgaatt 1920

gaaaagcttt atggaccctg aaaccacagc cacattaacc ttctttgatg gtcaaaactt 1980

atccttcacc ataaatatgc ctcgcaaaaa aggtaattaa catatataga attacattat 2040

ttatgaaata tcatcactat ctcttagcat ctttaatcct tttctacatc agataacttc 2100

ggtttgttat catcgtctgt attgtcatca attggcgcag tagcctcaat ttcaacgtcg 2160

tttgactctg gtgtttgttc atgtgcagat ccatgagatg atgaaccgct cgtccaacgc 2220

cggcggacct cttttaattc tgctgtaacc cgtacatgcc caaaataggg ggcgggttac 2280

acagaatata taacatcgta ggtgtctggg tgaacagttt attcctggca tccactaaat 2340

ataatggagc ccgctttta agctggcatc cagaaaaaaa aagaatccca gcaccaaaat 2400

attgttttct tcaccaacca tcagttcata ggtccattct cttagcgcaa ctacagagaa 2460

caggggcaca aacaggcaaa aaacgggcac aacctcaatg gagtgatgca acctgcctgg 2520

agtaaatgat gacacaaggc aattgaccca cgcatgtatc tatctcattt tcttacacct 2580

tctattacct tctgctctct ctgatttgga aaaagctgaa aaaaaaggtt gaaaccagtt 2640

ccctgaaatt attcccctac ttgactaata agtatataaa gacggtaggt attgattgta 2700

attctgtaaa tctatttctt aaacttctta aattctactt ttatagttag tctttttttt 2760

agttttaaaa caccaagaac ttagtttcga ataaacacac ataaacaaac aaattaataa 2820

tggctgcaga ccaattggtg aagactgaag tcaccaagaa gtcttttact gctcctgtac 2880

aaaaggcttc tacaccagtt ttaaccaata aaacagtcat ttctggatcg aaagtcaaaa 2940

gtttatcatc tgcgcaatcg agctcatcag gaccttcatc atctagtgag gaagatgatt 3000

cccgcgatat tgaaagcttg gataagaaaa tacgtccttt agaagaatta gaagcattat 3060

taagtagtgg aaatacaaaa caattgaaga acaaagaggt cgctgccttg gttattcacg 3120

gtaagttacc tttgtacgct ttggagaaaa aattaggtga tactacgaga gcggttgcgg 3180

tacgtaggaa ggctctttca attttggcag aagctcctgt attagcatct gatcgtttac 3240

catataaaaa ttatgactac gaccgcgtat ttggcgcttg ttgtgaaaat gttataggtt 3300

acatgccttt gcccgttggt gttataggcc ccttggttat cgatggtaca tcttatcata 3360

taccaatggc aactacagag ggttgtttgg tagcttctgc catgcgtggc tgtaaggcaa 3420

tcaatgctgg cggtggtgca acaactgttt taactaagga tggtatgaca agaggcccag 3480

tagtccgttt cccaactttg aaaagatctg gtgcctgtaa gatatggtta gactcagaag 3540

agggacaaaa cgcaattaaa aaagctttta actctacatc aagatttgca cgtctgcaac 3600

atattcaaac ttgtctagca ggagatttac tcttcatgag atttagaaca actactggtg 3660

acgcaatggg tatgaatatg atttctaagg gtgtcgaata ctcattaaag caaatggtag 3720

aagagtatgg ctgggaagat atggaggttg tctccgtttc tggtaactac tgtaccgaca 3780

aaaaaccagc tgccatcaac tggatcgaag gtcgtggtaa gagtgtcgtc gcagaagcta 3840

ctattcctgg tgatgttgtc agaaaagtgt taaaaagtga tgtttccgca ttggttgagt 3900

tgaacattgc taagaatttg gttggatctg caatggctgg gtctgttggt ggatttaacg 3960

cacatgcagc taatttagtg acagctgttt tcttggcatt aggacaagat cctgcacaaa 4020

atgtcgaaag ttccaactgt ataacattga tgaaagaagt ggacggtgat ttgagaattt 4080

ccgtatccat gccatccatc gaagtaggta ccatcggtgg tggtactgtt ctagaaccac 4140

aaggtgccat gttggactta ttaggtgtaa gaggcccaca tgctaccgct cctggtacca 4200

acgcacgtca attagcaaga atagttgcct gtgccgtctt ggcaggtgaa ttatccttat 4260

gtgctgccct agcagccggc catttggttc aaagtcatat gacccacaac aggaaacctg 4320

ctgaaccaac aaaacctaac aatttggacg ccactgatat aaatcgtttg aaagatgggt 4380

ccgtcacctg cattaaatcc taaacttagt catacgtcat tggtattctc ttgaaaaaga 4440

agcacaacag caccatgtgt tacgtaaaat atttacttta tagtttgtac gtcataattt 4500

cttccatatt acaagttcgt gcatatatag aaagaattct gttgttgtaa ttgtcataac 4560

taacctgcag gccgcgagcg ccgatagtat acactaaatt ttatgcaata ataaaaagaa 4620

agcatcccgc caaacgtttc gtaactacat attgttacat agtttgattc cgtgaatttg 4680

aagtggacgc agttcttctt agtctttcaa gttcatagta aatagacatt acccaaagat 4740

ctctttcctg tcttgatctt gccatgaaaa ccattgactt gccgctcaca cctaactttt 4800

tagtgaaatg gatcttatct gtgttacttg gaacattaga gttattcaag tcagcttccg 4860

gagcagaagg tgggtcaata ttattatctt gtctaccata atcttgcgta tattgctttt 4920

cattcccttt cttctgtaag cgattgctag aaagggcacg ccttgttcca aaccaaagcg 4980

taaagcatcg cgaactctcg tcttctacag atcgccatcc gtccccataa actcgtggga 5040

gagcgtggga tccatagttt atttcgtcaa aagtacgatc tctctgcaga aggtctagtt 5100

ctgttgtggt cccagagtat aggtagcaat cgtcaatcgg aatggttaca taatgagcat 5160

actccaaaac ttccttagca aaacccggtg tttaaacccc agcgcctggc ggg 5213<210> 120

<211> 5417

<212> DNA

<213> Artificial Sequence

<220>

<223> IME1 knockout stitch product

<400> 120

gacggcacgg ccacgcgttt aaaccgccat ccaattcctc tattatatgc atacattttt 60

ttgtttgatt atctatgctg tttataatat cgtatatagt taatgaataa ccctacaaca 120

cgaagggcag taatatattc tgattctcag ttgaaatttc aaatttttca agctggtctt 180

actcggcagg taggaacttc ccagtgggtc tgagttttct ctccggaagg tatttatcat 240

tatgcacgac aagtgcgatt caatcgaaag attataaatt tcgcgatgaa caacatctga 300

taaaaaaaaa attaaaaaac aataatctaa atgatatgta tgggttaaaa aggatgtatg 360

gatggagggt tggcataaaa atgaaaggca gaagatgggc ggctaaaaag gtatagagat 420

cgtacgtcac cgtcacaaaa atcactcata gaatgcacta tcatccttac aataccatac 480

cagccgcaag aaaaaaagaa aaaaaaatca attcatatca tatattatct atatcatgct 540

gttctttccg ccacgctcac acgcggccag ggggagcccg ttgagccatt agtatcaatt 600

tgcttacctg tattccttta ctatcctcct ttttctcctt cttgataaat gtatgtagat 660

tgcgtatata gtttcgtcta ccctatgaac atattccatt ttgtaatttc gtgtcgtttc 720

tattatgaat ttcatttata aagtttatgt acaaatatca taaaaaaaga gaatcttttt 780

aagcaaggat tttcttaact tcttcggcga cagcatcacc gacttcggtg gtactgttgg 840

aaccacctaa atcaccagtt ctgatacctg catccaaaac ctttttaact gcatcttcaa 900

tggccttacc ttcttcaggc aagttcaatg acaatttcaa catcattgca gcagacaaga 960

tagtggcgat agggtcaacc ttattctttg gcaaatctgg agcagaaccg tggcatggtt 1020

cgtacaaacc aaatgcggtg ttcttgtctg gcaaagaggc caaggacgca gatggcaaca 1080

aacccaagga acctgggata acggaggctt catcggagat gatatcacca aacatgttgc 1140

tggtgattat aataccattt aggtgggttg ggttcttaac taggatcatg gcggcagaat 1200

caatcaattg atgttgaacc ttcaatgtag ggaattcgtt cttgatggtt tcctccacag 1260
tttttctcca taatcttgaa gaggccaaaa gattagcttt atccaaggac caaataggca 1320
atggtggctc atgttgtagg gccatgaaag cggccattct tgtgattctt tgcacttctg 1380
gaacggtgta ttgttcacta tcccaagcga caccatcacc atcgtcttcc tttctcttac 1440
caaagtaaat acctcccact aattctctga caacaacgaa gtcagtacct ttagcaaatt 1500
gtggcttgat tggagataag tctaaaagag agtcggatgc aaagttacat ggtcttaagt 1560
tggcgtacaa ttgaagttct ttacggattt ttagtaaacc ttgttcaggt ctaacactac 1620
cggtacccca tttaggacca gccacagcac ctaacaaaac ggcatcaacc ttcttggagg 1680

cttccagcgc ctcatctgga agtgggacac ctgtagcatc gatagcagca ccaccaatta 1740
aatgattttc gaaatcgaac ttgacattgg aacgaacatc agaaatagct ttaagaacct 1800
taatggcttc ggctgtgatt tcttgaccaa cgtggtcacc tggcaaaacg acgatcttct 1860
taggggcaga cataggggca gacattagaa tggtatatcc ttgaaatata tatatatatt 1920
gctgaaatgt aaaaggtaag aaaagttaga aagtaagacg attgctaacc acctattgga 1980
aaaaacaata ggtccttaaa taatattgtc aacttcaagt attgtgatgc aagcatttag 2040
tcatgaacgc ttctctattc tatatgaaaa gccggttccg gcctctcacc tttccttttt 2100
ctcccaattt ttcagttgaa aaaggtatat gcgtcaggcg acctctgaaa ttaacaaaaa 2160

atttccagtc atcgaatttg attctgtgcg atagcgcccc tgtgtgttct cgttatgttg 2220
aggaaaaaaa taatggttgc taagagattc gaactcttgc atcttacgat acctgagtat 2280
tcccacagtt aactgcggtc aagatatttc ttgaatcagg cgcctcgctc gtccaacgcc 2340
ggcggacctc ttttaattct gctgtaaccc gtacatgccc aaaatagggg gcgggttaca 2400
cagaatatat aacatcgtag gtgtctgggt gaacagttta ttcctggcat ccactaaata 2460
taatggagcc cgctttttaa gctggcatcc agaaaaaaaa agaatcccag caccaaaata 2520
ttgttttctt caccaaccat cagttcatag gtccattctc ttagcgcaac tacagagaac 2580
aggggcacaa acaggcaaaa aacgggcaca acctcaatgg agtgatgcaa cctgcctgga 2640

gtaaatgatg acacaaggca attgacccac gcatgtatct atctcatttt cttacacctt 2700
ctattacctt ctgctctctc tgatttggaa aaagctgaaa aaaaaggttg aaaccagttc 2760
cctgaaatta ttcccctact tgactaataa gtatataaag acggtaggta ttgattgtaa 2820
ttctgtaaat ctatttctta aacttcttaa attctacttt tatagttagt ctttttttta 2880
gttttaaaac accaagaact tagtttcgaa taaacacaca taaacaaaca aattaataat 2940
gtcaactttg cctatttctt ctgtgtcatt ttcctcttct acatcaccat tagtcgtgga 3000

cgacaaagtc tcaaccaagc ccgacgttat cagacataca atgaatttca atgcttctat 3060

ttggggagat caattcttga cctatgatga gcctgaagat ttagttatga agaaacaatt 3120

agtggaggaa ttaaaagagg aagttaagaa ggaattgata actatcaaag gttcaaatga 3180

gcccatgcag catgtgaaat tgattgaatt aattgatgct gttcaacgtt taggtatagc 3240

ttaccatttt gaagaagaga tcgaggaagc tttgcaacat atacatgtta cctatggtga 3300

acagtgggtg gataaggaaa atttacagag tatttcattg tggttcaggt tgttgcgtca 3360

acagggcttt aacgtctcct ctggcgtttt caaagacttt atggacgaaa aaggtaaatt 3420

caaagagtct ttatgcaatg atgcacaagg aatattagcc ttatatgaag ctgcatttat 3480

gagggttgaa gatgaaacca tcttagacaa tgctttggaa ttcacaaaag ttcatttaga 3540

tatcatagca aaagacccat cttgcgattc ttcattgcgt acacaaatcc atcaagcctt 3600

aaaacaacct ttaagaagga gattagcaag gattgaagca ttacattaca tgccaatcta 3660

ccaacaggaa acatctcatg atgaagtatt gttgaaatta gccaagttgg atttcagtgt 3720

tttgcagtct atgcataaaa aggaattgtc acatatctgt aagtggtgga aagatttaga 3780

tttacaaaat aagttacctt atgtacgtga tcgtgttgtc gaaggctact tctggatatt 3840

gtccatatac tatgagccac aacacgctag aacaagaatg tttttgatga aaacatgcat 3900

gtggttagta gttttggacg atacttttga taattatgga acatacgaag aattggagat 3960

ttttactcaa gccgtcgaga gatggtctat ctcatgctta gatatgttgc ccgaatatat 4020

gaaattaatc taccaagaat tagtcaattt gcatgtggaa atggaagaat ctttggaaaa 4080

ggagggaaag acctatcaga ttcattacgt taaggagatg gctaaagaat tagttcgtaa 4140

ttacttagta gaagcaagat ggttgaagga aggttatatg cctactttag aagaatacat 4200

gtctgtttct atggttactg gtacttatgg tttgatgatt gcaaggtcct atgttggcag 4260

aggagacatt gttactgaag acacattcaa atgggtttct agttacccac ctattattaa 4320

agcttcctgt gtaatagtaa gattaatgga cgatattgta tctcacaagg aagaacaaga 4380

aagaggacat gtggcttcat ctatagaatg ttactctaaa gaatcaggtg cttctgaaga 4440

ggaagcatgt gaatatatta gtaggaaagt tgaggatgcc tggaaagtaa tcaatagaga 4500

atctttgcgt ccaacagccg ttcccttccc tttgttaatg ccagcaataa acttagctag 4560

aatgtgtgag gtcttgtact ctgttaatga tggttttact catgctgagg gtgacatgaa 4620

atcttatatg aagtccttct tcgttcatcc tatggtcgtt tgagctagct aagatccgct 4680

ctaaccgaaa aggaaggagt tagacaacct gaagtctagg tccctattta ttttttttata 4740

gttatgttag tattaagaac gttatttata tttcaaattt ttctttttt tctgtacaga 4800

cgcgtgtacg catgtaacat tatactgaaa accttgcttg agaaggtttt gggacgctcg 4860

aagaacctgc aggccgcgag cgccgatctc gaaaagtact acaatcttcc cccttccctc 4920

aaaatatatc cattcacact catttcttta tttccattgt ctcactcaaa ttgctaagaa 4980

tttgtgtatt tgcatatata tatattatat ataggtatat atatgtattc aatgtctcaa 5040

gctccatgac ataataccgt agcgttatta gccttatcgt atgtcgcgat gggaaaggag 5100

attcgtttta atcttgaaaa accttcgtag cgaataatgc gacataaatc ttgagagagt 5160

acatcaccaa attcactttg ttaaaccgca ccatcgtgct ttgcattctt attccttttg 5220

cctacactaa aattattagc atttctctaa atgagctcca gtgaagacga agacgacaag 5280

ttcttgtatg gttccgactc cgaattagca ctaccttcat ctaaacgatc aagagatgat 5340

gaagcagacg caggtgcgtc cagtaatcct gatatagtta aaaggcaaac ggtgtttaaa 5400

ccccagcgcc tggcggg 5417<210>

121

<211> 1737

<212> DNA

<213> Artificial Sequence

<220>

<223> beta-farnesene synthase of Artemisia annua codon-optimized
for expression in Saccharomyces cerevisiae

<400> 121

ggatccatgt caactttgcc tatttcttct gtgtcatttt cctcttctac atcaccatta 60

gtcgtggacg acaaagtctc aaccaagccc gacgttatca gacatacaat gaatttcaat 120

gcttctattt ggggagatca attcttgacc tatgatgagc ctgaagattt agttatgaag 180

aaacaattag tggaggaatt aaaagaggaa gttaagaagg aattgataac tatcaaaggt 240

tcaaatgagc ccatgcagca tgtgaaattg attgaattaa ttgatgctgt tcaacgttta 300

ggtatagctt accattttga agaagagatc gaggaagctt tgcaacatat acatgttacc 360

tatggtgaac agtgggtgga taaggaaaat ttacagagta tttcattgtg gttcaggttg 420

ttgcgtcaac agggctttaa cgtctcctct ggcgttttca aagactttat ggacgaaaaa 480

ggtaaattca aagagtcttt atgcaatgat gcacaaggaa tattagcctt atatgaagct 540

gcatttatga gggttgaaga tgaaaccatc ttagacaatg ctttggaatt cacaaaagtt 600

catttagata tcatagcaaa agacccatct tgcgattctt cattgcgtac acaaatccat 660

caagccttaa aacaaccttt aagaaggaga ttagcaagga ttgaagcatt acattacatg 720

ccaatctacc aacaggaaac atctcatgat gaagtattgt tgaaattagc caagttggat 780

ttcagtgttt tgcagtctat gcataaaaag gaattgtcac atatctgtaa gtggtggaaa 840

gatttagatt tacaaaataa gttaccttat gtacgtgatc gtgttgtcga aggctacttc 900

tggatattgt ccatatacta tgagccacaa cacgctagaa caagaatgtt tttgatgaaa 960

acatgcatgt ggttagtagt tttggacgat acttttgata attatggaac atacgaagaa 1020

ttggagattt ttactcaagc cgtcgagaga tggtctatct catgcttaga tatgttgccc 1080

gaatatatga aattaatcta ccaagaatta gtcaatttgc atgtggaaat ggaagaatct 1140

ttggaaaagg agggaaagac ctatcagatt cattacgtta aggagatggc taaagaatta 1200

gttcgtaatt acttagtaga agcaagatgg ttgaaggaag gttatatgcc tactttagaa 1260

gaatacatgt ctgtttctat ggttactggt acttatggtt tgatgattgc aaggtcctat 1320

gttggcagag gagacattgt tactgaagac acattcaaat gggtttctag ttacccacct 1380

attattaaag cttcctgtgt aatagtaaga ttaatggacg atattgtatc tcacaaggaa 1440

gaacaagaaa gaggacatgt ggcttcatct atagaatgtt actctaaaga atcaggtgct 1500

tctgaagagg aagcatgtga atatattagt aggaaagttg aggatgcctg gaaagtaatc 1560

aatagagaat ctttgcgtcc aacagccgtt cccttccctt tgttaatgcc agcaataaac 1620

ttagctagaa tgtgtgaggt cttgtactct gttaatgatg gttttactca tgctgagggt 1680

gacatgaaat cttatatgaa gtccttcttc gttcatccta tggtcgtttg actcgag 1737

<210> 122

<211> 7348

<212> DNA

<213> Artificial Sequence

<220>

<223> Vector pAM178

<400> 122

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatatcga ctacgtcgta aggccgtttc tgacagagta aaattcttga gggaactttc 240

accattatgg gaaatgcttc aagaaggtat tgacttaaac tccatcaaat ggtcaggtca 300

ttgagtgttt tttatttgtt gtatttttt ttttttagag aaaatcctcc aatatcaaat 360

taggaatcgt agtttcatga ttttctgtta cacctaactt tttgtgtggt gccctcctcc 420

ttgtcaatat taatgttaaa gtgcaattct ttttccttat cacgttgagc cattagtatc 480

aatttgctta cctgtattcc tttactatcc tccttttttct ccttcttgat aaatgtatgt 540

agattgcgta tatagtttcg tctaccctat gaacatattc cattttgtaa tttcgtgtcg 600

tttctattat gaatttcatt tataaagttt atgtacaaat atcataaaaa aagagaatct 660

ttttaagcaa ggattttctt aacttcttcg gcgacagcat caccgacttc ggtggtactg 720

ttggaaccac ctaaatcacc agttctgata cctgcatcca aaacctttt aactgcatct 780

tcaatggcct taccttcttc aggcaagttc aatgacaatt tcaacatcat tgcagcagac 840

aagatagtgg cgatagggtc aaccttattc tttggcaaat ctggagcaga accgtggcat 900

ggttcgtaca aaccaaatgc ggtgttcttg tctggcaaag aggccaagga cgcagatggc 960

aacaaaccca aggaacctgg gataacggag gcttcatcgg agatgatatc accaaacatg 1020

ttgctggtga ttataatacc atttaggtgg gttgggttct taactaggat catggcggca 1080

gaatcaatca attgatgttg aaccttcaat gtagggaatt cgttcttgat ggtttcctcc 1140

acagtttttc tccataatct tgaagaggcc aaaagattag ctttatccaa ggaccaaata 1200

ggcaatggtg gctcatgttg tagggccatg aaagcggcca ttcttgtgat tctttgcact 1260

tctggaacgg tgtattgttc actatcccaa gcgacaccat caccatcgtc ttcctttctc 1320

ttaccaaagt aaatacctcc cactaattct ctgacaacaa cgaagtcagt acctttagca 1380

aattgtggct tgattggaga taagtctaaa agagagtcgg atgcaaagtt acatggtctt 1440

aagttggcgt acaattgaag ttctttacgg atttttagta aaccttgttc aggtctaaca 1500

ctaccggtac cccatttagg accagccaca gcacctaaca aaacggcatc aaccttcttg 1560

gaggcttcca gcgcctcatc tggaagtgga acacctgtag catcgatagc agcaccacca 1620

attaaatgat tttcgaaatc gaacttgaca ttggaacgaa catcagaaat agctttaaga 1680

accttaatgg cttcggctgt gatttcttga ccaacgtggt cacctggcaa aacgacgatc 1740

ttcttagggg cagacattac aatggtatat ccttgaaata tataaaaaa aaggcgcctt 1800

agaccgctcg gccaaacaac caattacttg ttgagaaata gagtataatt atcctataaa 1860

tataacgttt ttgaacacac atgaacaagg aagtacagga caattgattt tgaagagaat 1920

gtggattttg atgtaattgt tgggattcca tttttaataa ggcaataata ttaggtatgt 1980

ggatatacta gaagttctcc tcgaccgtcg atatgcggtg tgaaataccg cacagatgcg 2040

taaggagaaa ataccgcatc aggaaattgt aaacgttaat attttgttaa aattcgcgtt 2100

aaatttttgt taaatcagct catttttaa ccaataggcc gaaatcggca aaatccctta 2160

taaatcaaaa gaatagaccg agatagggtt gagtgttgtt ccagtttgga acaagagtcc 2220

actattaaag aacgtggact ccaacgtcaa agggcgaaaa accgtctatc agggcgatgg 2280

cccactacgt gaaccatcac cctaatcaag ttttttgggg tcgaggtgcc gtaaagcact 2340

aaatcggaac cctaaaggga gcccccgatt tagagcttga cggggaaagc cggcgaacgt 2400

ggcgagaaag gaagggaaga aagcgaaagg agcgggcgct agggcgctgg caagtgtagc 2460

ggtcacgctg cgcgtaacca ccacacccgc cgcgcttaat gcgccgctac agggcgcgtc 2520

gcgccattcg ccattcaggc tgcgcaactg ttgggaaggg cgatcggtgc gggcctcttc 2580

gctattacgc cagctgaatt ggagcgacct catgctatac ctgagaaagc aacctgacct 2640

acaggaaaga gttactcaag aataagaatt ttcgttttaa aacctaagag tcactttaaa 2700

atttgtatac acttattttt tttataactt atttaataat aaaaatcata aatcataaga 2760

aattcgctta tttagaagtg tcaacaacgt atctaccaac gatttgaccc ttttccatct 2820

tttcgtaaat ttctggcaag gtagacaagc cgacaacctt gattggagac ttgaccaaac 2880

ctctggcgaa gaattgttaa ttaagagctc agatcttatc gtcgtcatcc ttgtaatcca 2940

tcgatactag tgcggccgcc ctttagtgag ggttgaattc gaattttcaa aaattcttac 3000

ttttttttg gatggacgca aagaagttta ataatcatat tacatggcat taccaccata 3060

tacatatcca tatacatatc catatctaat cttacttata tgttgtggaa atgtaaagag 3120

ccccattatc ttagcctaaa aaaaccttct ctttggaact ttcagtaata cgcttaactg 3180

ctcattgcta tattgaagta cggattagaa gccgccgagc gggtgacagc cctccgaagg 3240

aagactctcc tccgtgcgtc ctcgtcttca ccggtcgcgt tcctgaaacg cagatgtgcc 3300

tcgcgccgca ctgctccgaa caataaagat tctacaatac tagcttttat ggttatgaag 3360

aggaaaaatt ggcagtaacc tggccccaca aaccttcaaa tgaacgaatc aaattaacaa 3420

ccataggatg ataatgcgat tagtttttta gccttatttc tggggtaatt aatcagcgaa 3480

gcgatgattt ttgatctatt aacagatata taaatgcaaa aactgcataa ccactttaac 3540

taatactttc aacattttcg gtttgtatta cttcttattc aaatgtaata aaagtatcaa 3600

caaaaaattg ttaatatacc tctatacttt aacgtcaagg agaaaaaacc ccggatccgt 3660

aatacgactc actatagggc ccgggcgtcg acatggaaca gaagttgatt tccgaagaag 3720

acctcgagta agcttggtac cgcggctagc taagatccgc tctaaccgaa aaggaaggag 3780

ttagacaacc tgaagtctag gtccctattt atttttttat agttatgtta gtattaagaa 3840

cgttatttat atttcaaatt tttctttttt ttctgtacag acgcgtgtac gcatgtaaca 3900

ttatactgaa aaccttgctt gagaaggttt tgggacgctc gaagatccag ctgcattaat 3960

gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc 4020

tcactgactc gctgcgctcg gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg 4080

cggtaatacg gttatccaca gaatcagggg ataacgcagg aaagaacatg tgagcaaaag 4140

gccagcaaaa ggccaggaac cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcc 4200

gccccccctga cgagcatcac aaaaatcgac gctcaagtca gaggtggcga aacccgacag 4260

gactataaag ataccaggcg tttcccccctg gaagctccct cgtgcgctct cctgttccga 4320

ccctgccgct taccggatac ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc 4380

atagctcacg ctgtaggtat ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg 4440

tgcacgaacc ccccgttcag cccgaccgct gcgccttatc cggtaactat cgtcttgagt 4500

ccaacccggt aagacacgac ttatcgccac tggcagcagc cactggtaac aggattagca 4560

gagcgaggta tgtaggcggt gctacagagt tcttgaagtg gtggcctaac tacggctaca 4620

ctagaaggac agtatttggt atctgcgctc tgctgaagcc agttaccttc ggaaaaagag 4680

ttggtagctc ttgatccggc aaacaaacca ccgctggtag cggtggtttt tttgtttgca 4740

agcagcagat tacgcgcaga aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg 4800

ggtctgacgc tcagtggaac gaaaactcac gttaagggat tttggtcatg agattatcaa 4860

aaaggatctt cacctagatc cttttaaatt aaaaatgaag ttttaaatca atctaaagta 4920

tatatgagta aacttggtct gacagttacc aatgcttaat cagtgaggca cctatctcag 4980

cgatctgtct atttcgttca tccatagttg cctgactccc cgtcgtgtag ataactacga 5040

tacgggaggg cttaccatct ggccccagtg ctgcaatgat accgcgagac ccacgctcac 5100

cggctccaga tttatcagca ataaaccagc cagccggaag ggccgagcgc agaagtggtc 5160

ctgcaacttt atccgcctcc atccagtcta ttaattgttg ccgggaagct agagtaagta 5220

gttcgccagt taatagtttg cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac 5280

gctcgtcgtt tggtatggct tcattcagct ccggttccca acgatcaagg cgagttacat 5340

gatcccccat gttgtgcaaa aaagcggtta gctccttcgg tcctccgatc gttgtcagaa 5400

gtaagttggc cgcagtgtta tcactcatgg ttatggcagc actgcataat tctcttactg 5460

tcatgccatc cgtaagatgc ttttctgtga ctggtgagta ctcaaccaag tcattctgag 5520

aatagtgtat gcggcgaccg agttgctctt gcccggcgtc aatacgggat aataccgcgc 5580

cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg cgaaaactct 5640

caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca cccaactgat 5700

cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga aggcaaaatg 5760

ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc ttccttttc 5820

aatattattg aagcatttat cagggttatt gtctcatgag cggatacata tttgaatgta 5880

tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg ccacctgaac 5940

gaagcatctg tgcttcattt tgtagaacaa aaatgcaacg cgagagcgct aatttttcaa 6000

acaaagaatc tgagctgcat ttttacagaa cagaaatgca acgcgaaagc gctattttac 6060

caacgaagaa tctgtgcttc atttttgtaa aacaaaaatg caacgcgaga gcgctaattt 6120

ttcaaacaaa gaatctgagc tgcattttta cagaacagaa atgcaacgcg agagcgctat 6180

tttaccaaca aagaatctat acttcttttt tgttctacaa aaatgcatcc cgagagcgct 6240

atttttctaa caaagcatct tagattactt tttttctcct ttgtgcgctc tataatgcag 6300

tctcttgata actttttgca ctgtaggtcc gttaaggtta gaagaaggct actttggtgt 6360

ctattttctc ttccataaaa aaagcctgac tccacttccc gcgtttactg attactagcg 6420

aagctgcggg tgcatttttt caagataaag gcatccccga ttatattcta taccgatgtg 6480

gattgcgcat actttgtgaa cagaaagtga tagcgttgat gattcttcat tggtcagaaa 6540

attatgaacg gtttcttcta ttttgtctct atatactacg tataggaaat gtttacattt 6600

tcgtattgtt ttcgattcac tctatgaata gttcttacta caatttttt gtctaaagag 6660

taatactaga gataaacata aaaaatgtag aggtcgagtt tagatgcaag ttcaaggagc 6720

gaaaggtgga tgggtaggtt atatagggat atagcacaga gatatatagc aaagagatac 6780

ttttgagcaa tgtttgtgga agcggtattc gcaatatttt agtagctcgt tacagtccgg 6840

tgcgtttttg gttttttgaa agtgcgtctt cagagcgctt ttggttttca aaagcgctct 6900

gaagttccta tactttctag agaataggaa cttcggaata ggaacttcaa agcgtttccg 6960

aaaacgagcg cttccgaaaa tgcaacgcga gctgcgcaca tacagctcac tgttcacgtc 7020

gcacctatat ctgcgtgttg cctgtatata tatatacatg agaagaacgg catagtgcgt 7080

gtttatgctt aaatgcgtac ttatatgcgt ctatttatgt aggatgaaag gtagtctagt 7140

acctcctgtg atattatccc attccatgcg gggtatcgta tgcttccttc agcactaccc 7200

tttagctgtt ctatatgctg ccactcctca attggattag tctcatcctt caatgctatc 7260

atttcctttg atattggatc atactaagaa accattatta tcatgacatt aacctataaa 7320

aataggcgta tcacgaggcc ctttcgtc 7348<210> 123

<211> 746

<212> DNA

<213> Artificial Sequence

<220>

<223> URA3 knockout construct

<400> 123

tgcgaggcat atttatggtg aaggataagt tttgaccatc aaagaaggtt aatgtggctg 60

tggtttcagg gtccataaag cttttcaatt catctttttt ttttttgttc tttttttga 120

ttccggtttc tttgaaattt ttttgattcg gtaatctccg agcagaagga agaacgaagg 180

aaggagcaca gacttagatt ggtatatata cgcatatgtg gtgttgaaga aacatgaaat 240

tgcccagtat tcttaaccca actgcacaga acaaaaacct gcaggaaacg gctcacacgc 300

ggccaggggg agccctgtat tataagtaaa tgcatgtata ctaaactcac aaattagagc 360

ttcaatttaa ttatatcagt tattacccgg gaatctcggt cgtaatgatt tctataatga 420

cgaaaaaaaa aaaattggaa agaaaaagct tcatggcctt tataaaaagg aactatccaa 480

tacctcgcca gaaccaagta acagtatttt acggggcaca aatcaagaac aataagacag 540

gactgtaaag atggacgcat tgaactccaa agaacaacaa gagttccaaa aagtagtgga 600

acaaaagcaa atgaaggatt tcatgcgttt gtactctaat ctggtagaaa gatgtttcac 660

agactgtgtc aatgacttca caacatcaaa gctaaccaat aaggaacaaa catgcatcat 720

gaagtgctca gaaaagttct tgaagc 746

<210> 124

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-106-168-110-ERG9 CDS-f

<400> 124

atgggaaagc tattacaatt ggcattg 27

<210> 125

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-107-168-110-ERG9 CDS-r

<400> 125

attcaagttg taattttcat ctaagatgta gtcg 34

<210> 126

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-108-168-110-ERG9 US-f

<400> 126

aaaagtgcag ctcagagccc 20

<210> 127

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-162-168-110-LEU2 DS-f

<400> 127

aaagattctc ttttttttatg atatttgtac ataaact 37

<210> 128

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-163-168-110-LEU2 DS-r

<400> 128

tagatttagt actgaagagg aggtcgac 28

<210> 129

<211> 37

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-164-168-110-LEU2 US-f

<400> 129

taggataatt atactctatt tctcaacaag taattgg 37

<210> 130

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-165-168-110-LEU2 US-r

<400> 130

tagaatggta tatccttgaa atatatatat atatattgct g 41

<210> 131

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-169-168-110-URA3-f

<400> 131

gttcatcatc tcatggatct gcaca 25

<210> 132

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-170-168-110-URA3-r

<400> 132

atgcgtccat ctttacagtc ctg 23

<210> 133

<211> 60

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-172-168-110-ERG9 US-r1

<400> 133

gtgtgtgtgt gatatgtgac gtgtatacgt tttccgcttc tgcttttcgt cttttctctt 60

<210> 134

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-218-168-130-GAL80US-F

<400> 134

cagatggaat cccttccata gagag 25

<210> 135

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-219-168-130-GAL80US-R

<400> 135

gacgggagtg gaaagaacgg 20

<210> 136

<211> 21

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-220-168-130-GAL80DS-F

<400> 136

aagcatcttg ccctgtgctt g 21

<210> 137

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-221-168-130-GAL80DS-R

<400> 137

catgctacct tccatggttg agc 23

<210> 138

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-284-275-31-GAL4-FIX-F2

<400> 138

ggattttatg cccagggatg cacttcatgg atttgattgg tctg 44

<210> 139

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-285-275-31-GAL4-FIX-R2

<400> 139

cagaccaatc aaatccatga agtgcatccc tgggcataaa atcc 44

<210> 140

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-286-275-31-GAL4-F

<400> 140

atgaagctac tgtcttctat cgaacaagc 29

<210> 141

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer JU-287-275-31-GAL4-R

<400> 141

tgagcgaagc ttctgaataa gccc 24

<210> 142

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KB454-266-53

<400> 142

gacggcacgg ccacgcgttt aaaccgccat ccaattcctc tattatatgc 50

<210> 143

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KB455-266-53

<400> 143

ggctcccct ggccgcgtgt gagcgtggcg gaaagaacag c 41

<210> 144

<211> 50

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KB457-266-53

<400> 144

cccgccaggc gctggggttt aaacaccgtt tgccttttaa ctatatcagg 50

<210> 145

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH14-276-4-linker9.ERG12.rev

<400> 145

cccgccaggc gctggggttt aaacaccgat gtcattaccg ttcttaactt ctgc 54

<210> 146

<211> 49

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH15-276-4-linker9.ERG19.rev

<400> 146

cccgccaggc gctggggttt aaacaccgat gaccgtttac acagcatcc 49

<210> 147

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH33-276-21-URA3 5 prime fwd

<400> 147

tgcgaggcat atttatggtg aaggataagt tttgaccatc 40

<210> 148

<211> 64

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH34-276-21-URA3 5 prime rev

<400> 148

ggctcccccct ggccgcgtgt gagccgtttc ctgcaggttt ttgttctgtg cagttgggtt 60

aaga 64

<210> 149

<211> 73

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH35-276-21-URA3 3 prime fwd

<400> 149

gctcacacgc ggccaggggg agccctgtat tataagtaaa tgcatgtata ctaaactcac 60

aaattagagc ttc 73

<210> 150

<211> 40

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH35-276-21-URA3 3 prime rev

<400> 150

gcttcaagaa cttttctgag cacttcatga tgcatgtttg 40

<210> 151

<211> 44

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH46-276-43-ERG12linker4.fwd

<400> 151

aacctgcagg ccgcgagcgc cgatattcgc gggtggaagg acct 44

<210> 152

<211> 49

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH47-276-43-ERG19linker4.fwd

<400> 152

aacctgcagg ccgcgagcgc cgatcttgtg ctaagtggtg ctgttagac 49

<210> 153

<211> 59

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH5-276-1-linker3.FS(Kozak).fwd

<400> 153

atccccgcgt gcttggccgg ccgtaattaa taatgtcaac tttgcctatt tcttctgtg 59

<210> 154

<211> 46

<212

> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH7-276-1-linker4.TCYC1.rev

<400> 154

tacggcgctc gcggcctgca ggttcttcga gcgtcccaaa accttc 46

<210> 155

<211> 53

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH81-276-116-TDH3.rev.tHMG1

<400> 155

ggtctgcagc cattattaat ttgtttgttt atgtgtgttt attcgaaact aag 53

<210> 156

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH82-276-116-tHMG1.fwd.TDH3

<400> 156

cgaataaaca cacataaaca aacaaattaa taatggctgc agaccaattg gtgaag 56

<210> 157

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH8-276-1-linker4.tHMG1.fwd

<400> 157

aacctgcagg ccgcgagcgc cgatagttat gacaattaca acaacagaat tctttc 56

<210> 158

<211> 56

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH91-276-116-TDH3.rev.FS

<400> 158

taggcaaagt tgacattatt aatttgtttg tttatgtgtg tttattcgaa actaag 56

<210> 159

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH92-276-116-FS.fwd.TDH3

<400>

159

aaacacacat aaacaaacaa attaataatg tcaactttgc ctatttcttc tgtg 54

<210> 160

<211> 48

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH9-276-1-linker9.tHMG1.rev

<400> 160

cccgccaggc gctggggttt aaacaccgat ggctgcagac caattggt 48

<210> 161

<211> 47

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer KMH93-276-130-3 prime IME.linker4.fwd

<400> 161

aacctgcagg ccgcgagcgc cgatctcgaa aagtactaca atcttcc 47

<210> 162

<211> 44

<212>

DNA

<213> Artificial Sequence

<220>

<223> Primer PW-91-093-CPK422-G

<400> 162

gatgtgtatt actagtgtcg acgacagcat tcgcccagta tttt 44

<210> 163

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_K_0142

<400> 163

gtattccaat gagaatcgct agaa 24

<210> 164

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_K_0143

<400> 164

ttcgtctgtt tttatccctc ttc 23

<210> 165

<211> 20

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_K_131

<400> 165

cctctcttaa aatgatggcg 20

<210> 166

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_034

<400> 166

gacggtagca acaagaatat agcacgagcc 30

<210> 167

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_035

<400> 167

ttttgaggga atattcaact gttttttttt atcatg 36

<210> 168

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_036

<400> 168

tttttttatca tgttgatgct ctgcataata atgc 34

<210> 169

<211> 34

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_053

<400> 169

tttgtttgtt tatgtgtgtt tattcgaaac taag 34

<210> 170

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_106

<400> 170

atgtctcaga acgtttacat tgtatcg 27

<210> 171

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_107

<400> 171

aggcagccaa gacattgatt aacatcc 27

<210> 172

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_109

<400> 172

atgaaactct ctactaaact ttgttggtg 29

<210> 173

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_110

<400> 173

atgagaaaaa aaatcggttg ggcttaac 28

<210> 174

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_112

<400> 174

atgtcattac cgttcttaac ttctgc 26

<210> 175

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_113

<400> 175

attcgcgggt ggaaggacct tgtgg 25

<210> 176

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_115

<400> 176

atgaccgttt acacagcatc cgttacc 27

<210> 177

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_116

<400> 177

cttgtgctaa gtggtgctgt tagac 25

<210> 178

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_118

<400> 178

atgtcagagt tgagagcctt cagtg 25

<210> 179

<211> 23

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_119

<400> 179

agtgcacact ttcaagctaa cac 23

<210> 180

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_121

<400> 180

atgactgccg acaacaatag tatgccc 27

<210> 181

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_122

<400> 181

catcagtggg aaacattcaa gaggcc 26

<210> 182

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_124

<400> 182

atggcttcag aaaaagaaat taggagagag 30

<210> 183

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_125

<400> 183

tgaggtcgtt gcttttccta ttattatatg 30

<210> 184

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_193

<400> 184

tcgacactag taatacacat catcgtcc 28

<210> 185

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_194

<400> 185

gagctcctcg agaagttaag attatatg 28

<210> 186

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_194

<400> 186

atggctgcag atcaattggt gaagac 26

<210> 187

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_233

<400> 187

agttatgaca attacaacaa cagaattctt tc 32

<210> 188

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer GW-52-84 pAM326 BamHI

<400> 188

taataaggat ccatgtcaac tttgcctatt tc 32

<210> 189

<211> 32

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer GW-52-84 pAM326 NheI

<400> 189

ttatagctag ctcaaacgac cataggatga ac 32

<210> 190

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_266

<400> 190

tacttttttt ttggatggac gcaaag 26

<210> 191

<211> 29

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_L_267

<400> 191

aagtatagag gtatattaac aatttttg 29

<210> 192

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_RN017

<400> 192

acgaagtgac tgacagaata ctgacatcag 30

<210> 193

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_RN018

<400> 193

ttaaaagttg tttccgctgt atcctgtatc 30

<210> 194

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_RN019

<400> 194

agtatacact aaattttatg caataataaa 30

<210> 195

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_RN020

<400> 195

ggttttgcta aggaagtttt ggagtatgct 30

<210> 196

<211> 26

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_Z025

<400> 196

cacgaaaatc gttattgtct tgaagg 26

<210> 197

<211> 52

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_Z026

<400> 197

gctttatgga ccctgaaacc actcactatt attccataag atgatcatta gc 52

<210> 198

<211> 54

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_Z027

<400> 198

gcttcaattt aattatatca gttattacca cgaaaatcgt tattgtcttg aagg 54

<210> 199

<211> 30

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_Z028

<400> 199

tcactattat tccataagat gatcattagc 30

<210> 200

<211> 22

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_Z033

<400> 200

gtggtttcag ggtccataaa gc 22

<210> 201

<211> 28

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_Z034

<400> 201

gtaataactg atataattaa attgaagc 28

<210> 202

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_Z035

<400> 202

ctgttgacat tgcgaagagt gacaaagatt ttgttatcg 39

<210> 203

<211> 39

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer TRIX_Z036

<400> 203

cgataacaaa atctttgtca ctcttcgcaa tgtcaacag 39

<210> 204

<211> 45

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer VH228-235-7-URA3LOF3RYSE12-1F

<400> 204

gctcacacgc ggccaggggg agcctcacta ttattccata agatg 45

<210> 205

<211> 46

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer VH229-235-7-URA3LOF3RYSE12-1R

<400> 205

aggtccgccg gcgttggacg agcgcacgaa aatcgttatt gtcttg 46

<210> 206

<211> 198

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic DNA fragment

<400> 206

cacgaaaatc gttattgtct tgaaggtgaa atttctactc ttattaatgg tgaacgttaa 60

gctgatgcta tgatggaagc tgattggtct taacttgctt gtcatcttgc taatggtcat 120

atggctcgtg ttattactta agttatttgt actcgttttg aacgtaatgc taatgatcat 180

cttatggaat aatagtga 198

<210> 207

<211> 2733

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 2

<400> 207

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt cagctcacac 60

gcggccaggg ggagcctggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccacgctcg tccaacgccg gcggaccttg aagagcgagc tcccgctgag 600

caataactag cgtcatagct gtttcctggg tcgttcggct gcggcgagcg gtatcagctc 660

actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 720

gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 780

ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 840

acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 900

ctgttccgac cctgccgctt acccgatacc tgtccgcctt tctcccttcg ggaagcgtgg 960

cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 1020

tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 1080

gtcttgattc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 1140

ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 1200

acggctacac tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 1260

gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 1320

ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 1380

tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 1440

gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 1500

tctaaagtat atatgagtaa cttggtcgca tgcttaccaa tgcttaatca gtgaggcacc 1560

tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactgcccg tcgtgtagat 1620

aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc 1680

acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 1740

aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag 1800

agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt 1860

ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 1920

agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt 1980

tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc 2040

tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 2100

attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 2160

taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 2220

aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc 2280

caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 2340

gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatcaattg 2400

ccttttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gttacatatt 2460

tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 2520

acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 2580

gaggcccttt catctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 2640

cccggagaca gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 2700

cgcgtcagcg ggtgttggcg ggtgtcgggg ctg 2733

<210> 208

<211> 2733

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 3

<400> 208

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt cacgctcgtc 60

caacgccggc ggaccttggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccaatcccc gcgtgcttgg ccggccgttg aagagcgagc tcccgctgag 600

caataactag cgtcatagct gtttcctggg tcgttcggct gcggcgagcg gtatcagctc 660

actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 720

gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 780

ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 840

acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 900

ctgttccgac cctgccgctt acccgatacc tgtccgcctt tctcccttcg ggaagcgtgg 960

cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 1020

tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 1080

gtcttgattc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 1140

ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 1200

acggctacac tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 1260

gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 1320

ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 1380

tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 1440

gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 1500

tctaaagtat atatgagtaa cttggtcgca tgcttaccaa tgcttaatca gtgaggcacc 1560

tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactgcccg tcgtgtagat 1620

aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc 1680

acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 1740

aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag 1800

agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt 1860

ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 1920

agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt 1980

tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc 2040

tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 2100

attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 2160

taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 2220

aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc 2280

caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 2340

gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatcaattg 2400

ccttttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gttacatatt 2460

tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 2520

acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 2580

gaggcccttt catctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 2640

cccggagaca gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 2700

cgcgtcagcg ggtgttggcg ggtgtcgggg ctg 2733

<210> 209

<211> 2733

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 4

<400> 209

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caatccccgc 60

gtgcttggcc ggccgttggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccaaacctg caggccgcga gcgccgattg aagagcgagc tcccgctgag 600

caataactag cgtcatagct gtttcctggg tcgttcggct gcggcgagcg gtatcagctc 660

actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 720

gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 780

ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 840

acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 900

ctgttccgac cctgccgctt acccgatacc tgtccgcctt tctcccttcg ggaagcgtgg 960

cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 1020

tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 1080

gtcttgattc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 1140

ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 1200

acggctacac tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 1260

gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 1320

ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 1380

tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 1440

gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 1500

tctaaagtat atatgagtaa cttggtcgca tgcttaccaa tgcttaatca gtgaggcacc 1560

tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactgcccg tcgtgtagat 1620

aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc 1680

acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 1740

aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag 1800

agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt 1860

ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 1920

agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt 1980

tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc 2040

tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 2100

attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 2160

taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 2220

aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc 2280

caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 2340

gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatcaattg 2400

ccttttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gttacatatt 2460

tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 2520

acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 2580

gaggcccttt catctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 2640

cccggagaca gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 2700

cgcgtcagcg ggtgttggcg ggtgtcgggg ctg 2733

<210> 210

<211> 2733

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 5

<400> 210

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caaacctgca 60

ggccgcgagc gccgattggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccaaacgcg atcgccgacg ccgccgattg aagagcgagc tcccgctgag 600

caataactag cgtcatagct gtttcctggg tcgttcggct gcggcgagcg gtatcagctc 660

actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 720

gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 780

ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 840

acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 900

ctgttccgac cctgccgctt acccgatacc tgtccgcctt tctcccttcg ggaagcgtgg 960

cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 1020

tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 1080

gtcttgattc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 1140
ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 1200
acggctacac tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 1260
gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 1320
ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 1380
tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 1440
gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 1500
tctaaagtat atatgagtaa cttggtcgca tgcttaccaa tgcttaatca gtgaggcacc 1560

tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactgcccg tcgtgtagat 1620
aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc 1680
acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 1740
aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag 1800
agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt 1860
ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 1920
agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt 1980
tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc 2040

tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 2100
attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 2160
taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 2220
aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc 2280
caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 2340
gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatcaattg 2400
ccttttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gttacatatt 2460
tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 2520

acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 2580
gaggcccttt catctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 2640
cccggagaca gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 2700
cgcgtcagcg ggtgttggcg ggtgtcgggg ctg 2733

<210> 211

<211> 2733

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 6

<400> 211

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caaacgcgat 60

cgccgacgcc gccgattggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccaaaggcg gccgctggcg agggagattg aagagcgagc tcccgctgag 600

caataactag cgtcatagct gtttcctggg tcgttcggct gcggcgagcg gtatcagctc 660

actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 720

gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 780

ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 840

acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 900

ctgttccgac cctgccgctt acccgatacc tgtccgcctt tctcccttcg ggaagcgtgg 960

cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 1020

tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 1080

gtcttgattc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 1140

ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 1200

acggctacac tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 1260

gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 1320

ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 1380

tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 1440

gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 1500

tctaaagtat atatgagtaa cttggtcgca tgcttaccaa tgcttaatca gtgaggcacc 1560

tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactgcccg tcgtgtagat 1620

aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc 1680

acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 1740

aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag 1800

agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt 1860

ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 1920

agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt 1980

tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc 2040

tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 2100

attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 2160

taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 2220

aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc 2280

caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 2340

gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatcaattg 2400

ccttttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gttacatatt 2460

tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 2520

acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 2580

gaggcccttt catctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 2640

cccggagaca gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 2700

cgcgtcagcg ggtgttggcg ggtgtcgggg ctg 2733

<210> 212

<211> 2733

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 7

<400> 212

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caaaggcggc 60

cgctggcgag ggagattggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccaaaggcg cgccacggtc gtgcggattg aagagcgagc tcccgctgag 600

caataactag cgtcatagct gtttcctggg tcgttcggct gcggcgagcg gtatcagctc 660

actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 720

gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 780

ataggctccg ccccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 840

acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 900

ctgttccgac cctgccgctt acccgatacc tgtccgcctt tctcccttcg ggaagcgtgg 960

cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 1020

tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 1080

gtcttgattc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 1140

ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 1200

acggctacac tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 1260

gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 1320

ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 1380

tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 1440

gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 1500

tctaaagtat atatgagtaa cttggtcgca tgcttaccaa tgcttaatca gtgaggcacc 1560

tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactgcccg tcgtgtagat 1620

aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc 1680

acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 1740

aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag 1800

agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt 1860

ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 1920

agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt 1980

tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc 2040

tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 2100

attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 2160

taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 2220

aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc 2280

caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 2340

gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatcaattg 2400

ccttttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gttacatatt 2460

tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 2520

acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 2580

gaggcccttt catctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 2640

cccggagaca gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 2700

cgcgtcagcg ggtgttggcg ggtgtcgggg ctg 2733

<210> 213

<211> 2733

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 8

<400> 213

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caaaggcgcg 60

ccacggtcgt gcggattggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccaagcccc tcagcccccc tagcgtcgtg aagagcgagc tcccgctgag 600

caataactag cgtcatagct gtttcctggg tcgttcggct gcggcgagcg gtatcagctc 660

actcaaaggc ggtaatacgg ttatccacag aatcagggga taacgcagga aagaacatgt 720

gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc 780

ataggctccg cccccctgac gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa 840

acccgacagg actataaaga taccaggcgt ttcccccctgg aagctccctc gtgcgctctc 900

ctgttccgac cctgccgctt acccgatacc tgtccgcctt tctcccttcg ggaagcgtgg 960

cgctttctca tagctcacgc tgtaggtatc tcagttcggt gtaggtcgtt cgctccaagc 1020

tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg cgccttatcc ggtaactatc 1080

gtcttgattc caacccggta agacacgact tatcgccact ggcagcagcc actggtaaca 1140

ggattagcag agcgaggtat gtaggcggtg ctacagagtt cttgaagtgg tggcctaact 1200

acggctacac tagaagaaca gtatttggta tctgcgctct gctgaagcca gttaccttcg 1260

gaaaaagagt tggtagctct tgatccggca aacaaaccac cgctggtagc ggtggttttt 1320

ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc tcaagaagat cctttgatct 1380

tttctacggg gtctgacgct cagtggaacg aaaactcacg ttaagggatt ttggtcatga 1440

gattatcaaa aaggatcttc acctagatcc ttttaaatta aaaatgaagt tttaaatcaa 1500

tctaaagtat atatgagtaa cttggtcgca tgcttaccaa tgcttaatca gtgaggcacc 1560

tatctcagcg atctgtctat ttcgttcatc catagttgcc tgactgcccg tcgtgtagat 1620

aactacgata cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc 1680

acgctcaccg gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag 1740

aagtggtcct gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag 1800

agtaagtagt tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt 1860

ggtgtcacgc tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg 1920

agttacatga tcccccatgt tgtgcaaaaa agcggttagc tccttcggtc ctccgatcgt 1980

tgtcagaagt aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc 2040

tcttactgtc atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc 2100

attctgagaa tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa 2160

taccgcgcca catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg 2220

aaaactctca aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc 2280

caactgatct tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag 2340

gcaaaatgcc gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatcaattg 2400

ccttttttcaa tattattgaa gcatttatca gggttattgt ctcatgagcg gttacatatt 2460

tgaatgtatt tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc 2520

acctgacgtc taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac 2580

gaggcccttt catctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct 2640

cccggagaca gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg 2700

cgcgtcagcg ggtgttggcg ggtgtcgggg ctg 2733

<210> 214

<211> 2737

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 9

<400> 214

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt cacgctcgtc 60

caacgccggc ggaccttggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccacggtgt ttaaacccca gcgcctggcg ggtgaagagc gagctcccgc 600

tgagcaataa ctagcgtcat agctgtttcc tgggtcgttc ggctgcggcg agcggtatca 660

gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 720

atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 780

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 840

cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 900

tctcctgttc cgaccctgcc gcttacccga tacctgtccg cctttctccc ttcgggaagc 960

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 1020

aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 1080

tatcgtcttg attccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 1140

aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 1200

aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 1260

ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 1320

ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 1380

atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 1440

atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 1500

tcaatctaaa gtatatatga gtaacttggt cgcatgctta ccaatgctta atcagtgagg 1560

cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactg cccgtcgtgt 1620
agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 1680
acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 1740
gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 1800
ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 1860
tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 1920
ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga 1980
tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 2040

attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 2100
agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2160
ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 2220
ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 2280
cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 2340
gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatca 2400
attgcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggttaca 2460
tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 2520

tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 2580
tcacgaggcc ctttcatctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc 2640
agctcccgga gacagtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc 2700
agggcgcgtc agcgggtgtt ggcgggtgtc ggggctg 2737

<210> 215
<211> 2737
<212> DNA
<213> Artificial Sequence
<220>
<223> pRYSE Entry vector 10
<400> 215
gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caatccccgc 60
gtgcttggcc ggccgttggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccacggtgt ttaaacccca gcgcctggcg ggtgaagagc gagctcccgc 600

tgagcaataa ctagcgtcat agctgtttcc tgggtcgttc ggctgcggcg agcggtatca 660

gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 720

atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 780

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 840

cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 900

tctcctgttc cgaccctgcc gcttacccga tacctgtccg cctttctccc ttcgggaagc 960

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 1020

aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 1080

tatcgtcttg attccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 1140

aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 1200

aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 1260

ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 1320

ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 1380

atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 1440

atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 1500

tcaatctaaa gtatatatga gtaacttggt cgcatgctta ccaatgctta atcagtgagg 1560

cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactg cccgtcgtgt 1620

agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 1680

acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 1740

gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 1800

ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 1860

tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 1920

ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga 1980

tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 2040

attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 2100

agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2160

ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 2220

ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 2280

cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 2340

gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatca 2400

attgcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggttaca 2460

tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 2520

tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 2580

tcacgaggcc ctttcatctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc 2640

agctcccgga gacagtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc 2700

agggcgcgtc agcgggtgtt ggcgggtgtc ggggctg 2737

<210> 216

<211> 2737

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 11

<400> 216

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caaacctgca 60

ggccgcgagc gccgattggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccacggtgt ttaaacccca gcgcctggcg ggtgaagagc gagctcccgc 600

tgagcaataa ctagcgtcat agctgtttcc tgggtcgttc ggctgcggcg agcggtatca 660

gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 720

atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 780

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 840

cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 900

tctcctgttc cgaccctgcc gcttacccga tacctgtccg cctttctccc ttcgggaagc 960

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 1020

aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 1080

tatcgtcttg attccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 1140

aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 1200

aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 1260

ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 1320

ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 1380

atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 1440

atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 1500

tcaatctaaa gtatatatga gtaacttggt cgcatgctta ccaatgctta atcagtgagg 1560

cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactg cccgtcgtgt 1620

agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 1680

acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 1740

gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 1800

ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 1860

tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 1920

ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga 1980

tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 2040

attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 2100

agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2160

ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 2220

ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 2280

cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 2340

gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatca 2400

attgcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggttaca 2460

tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 2520

tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 2580

tcacgaggcc ctttcatctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc 2640

agctcccgga gacagtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc 2700

agggcgcgtc agcgggtgtt ggcgggtgtc ggggctg 2737

<210> 217

<211> 2737

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 12

<400> 217

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caaacgcgat 60

cgccgacgcc gccgattggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccacggtgt ttaaacccca gcgcctggcg ggtgaagagc gagctcccgc 600

tgagcaataa ctagcgtcat agctgtttcc tgggtcgttc ggctgcggcg agcggtatca 660

gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 720

atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 780

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 840

cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 900

tctcctgttc cgaccctgcc gcttacccga tacctgtccg cctttctccc ttcgggaagc 960

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 1020

aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 1080

tatcgtcttg attccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 1140

aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 1200

aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 1260

ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 1320

ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 1380

atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 1440

atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 1500

tcaatctaaa gtatatatga gtaacttggt cgcatgctta ccaatgctta atcagtgagg 1560

cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactg cccgtcgtgt 1620

agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 1680

acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 1740

gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 1800

ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 1860

tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 1920

ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga 1980

tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 2040

attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 2100

agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2160

ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 2220

ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 2280

cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 2340

gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatca 2400

attgcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggttaca 2460

tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 2520

tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 2580

tcacgaggcc ctttcatctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc 2640

agctcccgga gacagtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc 2700

agggcgcgtc agcgggtgtt ggcgggtgtc ggggctg 2737

<210> 218

<211> 2737

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 13

<400> 218

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caaaggcggc 60

cgctggcgag ggagattggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180
aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240
tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300
ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360
cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420
gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480
aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540
tagcgagtca tccacggtgt ttaaacccca gcgcctggcg ggtgaagagc gagctcccgc 600

tgagcaataa ctagcgtcat agctgtttcc tgggtcgttc ggctgcggcg agcggtatca 660
gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 720
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 780
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 840
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 900
tctcctgttc cgaccctgcc gcttacccga tacctgtccg cctttctccc ttcgggaagc 960
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 1020
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 1080

tatcgtcttg attccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 1140
aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 1200
aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 1260
ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 1320
ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 1380
atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 1440
atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 1500
tcaatctaaa gtatatatga gtaacttggt cgcatgctta ccaatgctta atcagtgagg 1560

cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactg cccgtcgtgt 1620
agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 1680
acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 1740
gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 1800
ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 1860
tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 1920

ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga 1980

tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 2040

attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 2100

agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2160

ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 2220

ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 2280

cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 2340

gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatca 2400

attgcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggttaca 2460

tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 2520

tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 2580

tcacgaggcc ctttcatctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc 2640

agctcccgga gacagtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc 2700

agggcgcgtc agcgggtgtt ggcgggtgtc ggggctg 2737

<210> 219

<211> 2737

<212> DNA

<213> Artificial Sequence

<220>

<223> pRYSE Entry vector 14

<400> 219

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caaaggcgcg 60

ccacggtcgt gcggattggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180

aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240

tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300

ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360

cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420

gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480

aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540

tagcgagtca tccacggtgt ttaaacccca gcgcctggcg ggtgaagagc gagctcccgc 600

tgagcaataa ctagcgtcat agctgtttcc tgggtcgttc ggctgcggcg agcggtatca 660

gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 720

atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 780

ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 840

cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 900

tctcctgttc cgaccctgcc gcttacccga tacctgtccg cctttctccc ttcgggaagc 960

gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 1020

aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 1080

tatcgtcttg attccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 1140

aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 1200

aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 1260

ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 1320

ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 1380

atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 1440

atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 1500

tcaatctaaa gtatatatga gtaacttggt cgcatgctta ccaatgctta atcagtgagg 1560

cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactg cccgtcgtgt 1620

agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 1680

acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 1740

gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 1800

ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 1860

tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 1920

ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga 1980

tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 2040

attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 2100

agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2160

ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 2220

ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 2280

cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 2340

gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatca 2400

attgcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggttaca 2460

tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 2520

tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 2580
tcacgaggcc ctttcatctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc 2640
agctcccgga gacagtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc 2700
agggcgcgtc agcgggtgtt ggcgggtgtc ggggctg 2737

<210> 220
<211> 2737
<212> DNA
<213> Artificial Sequence
<220>
<223> pRYSE Entry vector 15
<400> 220

gtaaaacgac ggccagtatt aaccctcact aaagggaact cgaggctctt caagcccctc 60
agccccccta gcgtcgtggc agactccata tgctatgcgg catcagagca gattgtactg 120

agagtgcacc atatgcggtg tgaaataccg cacagatgcg taaggagaaa ataccgcatc 180
aggcgccatt cgccattcag gctgcgcaac tgttgggaag ggcgatcggt gcgggcctct 240
tcgctattac gccagctggc gaaaggggga tgtgctgcaa ggcgattaag ttgggtaacg 300
ccagggtttt cccagtcacg acgttgtaaa acgacggcca gtgaattcga gctcggtacc 360
cggggatcct ctagcgtcga cctgcaggca tgcaagcttg gcgtaatcat ggtcatagct 420
gtttcctgtg tgaaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 480
aaagtgtaaa gcctggggtg cctaatgagt gagctaactc acattaattg cgttgcgcgc 540
tagcgagtca tccacggtgt ttaaacccca gcgcctggcg ggtgaagagc gagctcccgc 600

tgagcaataa ctagcgtcat agctgtttcc tgggtcgttc ggctgcggcg agcggtatca 660
gctcactcaa aggcggtaat acggttatcc acagaatcag gggataacgc aggaaagaac 720
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggccgcgtt gctggcgttt 780
ttccataggc tccgcccccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 840
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 900
tctcctgttc cgaccctgcc gcttacccga tacctgtccg cctttctccc ttcgggaagc 960
gtggcgcttt ctcatagctc acgctgtagg tatctcagtt cggtgtaggt cgttcgctcc 1020
aagctgggct gtgtgcacga acccccccgtt cagcccgacc gctgcgcctt atccggtaac 1080

tatcgtcttg attccaaccc ggtaagacac gacttatcgc cactggcagc agccactggt 1140

aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttcttgaa gtggtggcct 1200

aactacggct acactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 1260

ttcggaaaaa gagttggtag ctcttgatcc ggcaaacaaa ccaccgctgg tagcggtggt 1320

ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 1380

atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gattttggtc 1440

atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaaa 1500

tcaatctaaa gtatatatga gtaacttggt cgcatgctta ccaatgctta atcagtgagg 1560

cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactg cccgtcgtgt 1620

agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 1680

acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 1740

gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 1800

ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 1860

tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 1920

ggcgagttac atgatccccc atgttgtgca aaaaagcggt tagctccttc ggtcctccga 1980

tcgttgtcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 2040

attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 2100

agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 2160

ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 2220

ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 2280

cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 2340

gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatca 2400

attgcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggttaca 2460

tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 2520

tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 2580

tcacgaggcc ctttcatctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc 2640

agctcccgga gacagtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc 2700

agggcgcgtc agcgggtgtt ggcgggtgtc ggggctg 2737

【심사관 직권보정사항】

【직권보정 1】

【보정항목】 청구범위

【보정세부항목】 청구항 1

【변경전】

(c) ~ 선택된 임의의 DNA 절편, 두번째 제한 부위 RBm를 포함하며

【변경후】

(c) ~ 선택된 임의의 DNA 절편, 및 두번째 제한 부위 RBm를 포함하며

【직권보정 2】

【보정항목】 청구범위

【보정세부항목】 청구항 32

【변경전】

(iii) ~, 프라이머 결합 절편 PB, 두번째 제한 부위 RBm을 포함하며

【변경후】

(iii) ~, 프라이머 결합 절편 PB, 및 두번째 제한 부위 RBm을 포함하며

【직권보정 3】

【보정항목】 청구범위

【보정세부항목】 청구항 27

【변경전】

(c) ~ 선택된 임의의 DNA 절편, 두번째 제한 부위 RBm를 포함하며

【변경후】

(c) ~ 선택된 임의의 DNA 절편, 및 두번째 제한 부위 RBm를 포함하며