



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2021-0088615
(43) 공개일자 2021년07월14일

<p>(51) 국제특허분류(Int. Cl.) C12N 15/10 (2017.01) C40B 40/06 (2006.01) C40B 50/06 (2017.01)</p> <p>(52) CPC특허분류 C12N 15/1068 (2013.01) C12N 15/1093 (2013.01)</p> <p>(21) 출원번호 10-2021-7016645</p> <p>(22) 출원일자(국제) 2019년10월31일 심사청구일자 없음</p> <p>(85) 번역문제출일자 2021년05월31일</p> <p>(86) 국제출원번호 PCT/US2019/059051</p> <p>(87) 국제공개번호 WO 2020/092704 국제공개일자 2020년05월07일</p> <p>(30) 우선권주장 62/753,254 2018년10월31일 미국(US)</p>	<p>(71) 출원인 지머젠 인코포레이티드 미국 94608 캘리포니아 에머리빌 스위트 105 호턴 스트리트 5980</p> <p>(72) 발명자 딘 에릭 제데디아 미국 94608 캘리포니아 에머리빌 스위트 105 호턴 스트리트 5980 지머젠 인코포레이티드 내 파텔 케다 미국 94608 캘리포니아 에머리빌 스위트 105 호턴 스트리트 5980 지머젠 인코포레이티드 내 (뒷면에 계속)</p> <p>(74) 대리인 특허법인(유한)케이비케이</p>
---	---

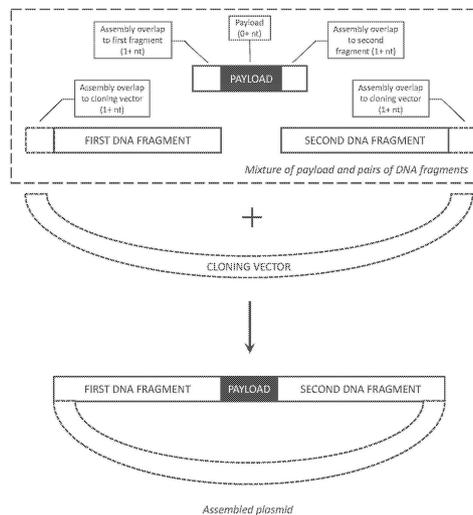
전체 청구항 수 : 총 77 항

(54) 발명의 명칭 DNA 라이브러리의 다중 결정적 어셈블리

(57) 요약

본 발명은 시험관 내 또는 생체 내에서 관심 3개 이상의 이중 가닥(ds) 또는 단일 가닥(ss) DNA 분자를 연결하는 방법에 관한 것이다. 이 방법은 결정적 방식으로 많은 수의 DNA 단편을 연결하게 한다. 이는, 예를 들어, 게놈 편집 및 경로 어셈블리를 포함하는 다양한 응용분야에서 후속적으로 사용될 수 있는 핵산 라이브러리를 빠르게 생성하는 데 사용될 수 있다. 방법을 수행하기 위한 키트도 개시된다.

대표도 - 도1



(52) CPC특허분류

C40B 40/06 (2013.01)

C40B 50/06 (2013.01)

(72) 발명자

필리 애런

미국 94608 캘리포니아 에머리빌 스위트 105 호턴
스트리트 5980 지머젠 인코포레이티드 내

메타 커날

미국 94608 캘리포니아 에머리빌 스위트 105 호턴
스트리트 5980 지머젠 인코포레이티드 내

웨이먼 필립

미국 94608 캘리포니아 에머리빌 스위트 105 호턴
스트리트 5980 지머젠 인코포레이티드 내

명세서

청구범위

청구항 1

폴리뉴클레오타이드의 혼합물을 포함하는 조성물로서, 이 혼합물은 폴리뉴클레오타이드 쌍을 함유하는 제 1 풀을 포함하며, 여기서, 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 함유한다; 및 삽입체 폴리뉴클레오타이드의 제 2 풀을 포함하며, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열을 포함하고 제 1 풀로부터의 한 쌍의 폴리뉴클레오타이드에서 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 이의 반대 3' 말단에 제 2 어셈블리 오버랩 서열을 포함하는 것인 폴리뉴클레오타이드의 혼합물을 포함하는 조성물.

청구항 2

제 1 항에 있어서,

클로닝 벡터를 추가로 포함하고, 여기서 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단은 클로닝 벡터에 상보적인 서열을 포함하는 것인 조성물.

청구항 3

제 2 항에 있어서,

제 1 풀의 각 폴리뉴클레오타이드는 제 1 풀의 폴리뉴클레오타이드가 지정된 역치를 넘어 제 1 풀의 다른 폴리뉴클레오타이드와 공통 서열을 공유하지 않아서, 제 1 풀의 폴리뉴클레오타이드의 쌍과 제 2 풀의 삽입체 폴리뉴클레오타이드 사이 또는 제 1 풀의 폴리뉴클레오타이드의 쌍과 클로닝 벡터 사이의 디자인된 어셈블리 오버랩 서열을 제외하도록 선택되는 것인 조성물.

청구항 4

제 3 항에 있어서,

지정된 역치는 5개 내지 15개의 연속(contiguous) 뉴클레오타이드인 조성물.

청구항 5

제 1 항 내지 제 4 항 중 어느 한 항에 있어서,

중합 효소를 추가로 포함하는 것인 조성물.

청구항 6

제 5 항에 있어서,

중합 효소는 가닥 치환 또는 비 가닥 치환인 조성물.

청구항 7

제 6 항에 있어서,

중합 효소는 비 가닥 치환이고 조성물은 크라우딩제를 추가로 포함하는 것인 조성물.

청구항 8

제 7 항에 있어서,

크라우딩제는 폴리에틸렌 글리콜(PEG)인 조성물.

청구항 9

제 8 항에 있어서,

PEG는 약 3 내지 약 7%(중량/부피)의 농도로 사용되는 것인 조성물.

청구항 10

제 8 항에 있어서,

PEG는 PEG-200, PEG-4000, PEG-6000, PEG-8000 또는 PEG-20,000에서 선택되는 것인 조성물.

청구항 11

제 6 항에 있어서,

중합 효소는 가닥 치환이고 조성물은 단일 가닥 결합 단백질을 추가로 포함하는 것인 조성물.

청구항 12

제 11 항에 있어서,

단일 가닥 DNA 결합 단백질은 극도의 내열성 단일 가닥 DNA 결합 단백질(ET SSB), 대장균 *recA*, T7 유전자 2.5 산물, 파지 람다 *RedB* 또는 *Rac* 프로파지 *RecT*인 조성물.

청구항 13

제 1 항에 있어서,

5'-3' 엑소뉴클레아제를 추가로 포함하는 것인 조성물.

청구항 14

제 1 항에 있어서,

리가아제를 추가로 포함하는 것인 조성물.

청구항 15

제 1 항에 있어서,

제 1 폴의 각 쌍은 이중 가닥 DNA(dsDNA) 또는 단일 가닥 DNA(ssDNA)인 조성물.

청구항 16

제 1 항에 있어서,

제 2 폴의 각 삽입체 폴리뉴클레오타이드는 dsDNA 또는 ssDNA인 조성물.

청구항 17

제 1 항에 있어서,

제 1 폴의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 조성물.

청구항 18

제 1 항에 있어서,

제 1 폴의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 대사 경로의 일부인 유전자에 상응하는 코딩 서열을 포함하는 것인 조성물.

청구항 19

제 1 항에 있어서,

제 1 폴의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 기능적 도메인 또는 하나 이

상의 단백질에 상응하는 코딩 서열을 포함하는 것인 조성물.

청구항 20

제 1 항에 있어서,

제 1 폴의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 단일 구조체에서 함께 연결되며, 여기서 단일 구조체는 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드 사이의 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 것인 조성물.

청구항 21

제 20 항에 있어서,

하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열은 귀소 엔도뉴클레아제 인식 서열을 포함하는 것인 조성물.

청구항 22

제 1 항에 있어서,

제 2 폴의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 폴의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함하는 것인 조성물.

청구항 23

제 1 항에 있어서,

제 2 폴의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 폴의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 약 25개의 뉴클레오타이드를 포함하는 것인 조성물.

청구항 24

제 1 항에 있어서,

제 2 폴의 각 삽입체 폴리뉴클레오타이드는 제 1 어셈블리 오버랩 서열과 제 2 어셈블리 오버랩 서열 사이에 위치한 하나 이상의 페이로드 서열을 포함하는 것인 조성물.

청구항 25

제 24 항에 있어서,

하나 이상의 페이로드 서열은 프로모터, 유전자, 조절 서열, 디그론을 암호화하는 핵산 서열, 용해도 태그를 암호화하는 핵산 서열, 종결자, 고유 식별자 서열 또는 이의 일부로부터 선택되는 것인 조성물.

청구항 26

제 17 항에 있어서,

제 1 폴의 제 1 및 제 2 폴리뉴클레오타이드의 각 쌍은 제 1 폴의 서로의 쌍과 비교하여 숙주 세포의 상이한 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 조성물.

청구항 27

제 17 항에 있어서,

제 1 폴의 제 1 및 제 2 폴리뉴클레오타이드의 각 쌍은 숙주 세포의 동일한 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 조성물.

청구항 28

제 24 항에 있어서,

제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 상이한 것인 조성물.

청구항 29

제 24 항에 있어서,

제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 동일한 것인 조성물.

청구항 30

폴리뉴클레오타이드의 라이브러리를 생성하는 방법으로서, 이 방법은 (a) 폴리뉴클레오타이드의 제 1 풀 및 폴리뉴클레오타이드의 제 2 풀을 조합하는 단계, 여기서 제 1 풀은 폴리뉴클레오타이드의 쌍을 포함하며, 여기서 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 제 2 풀은 삽입체 폴리뉴클레오타이드를 포함하며, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열을 포함하고 제 1 풀로부터의 한 쌍의 폴리뉴클레오타이드에서 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 이의 반대 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다; 및

(b) 제 1 풀 및 제 2 풀을 폴리뉴클레오타이드의 라이브러리로 어셈블리하는 단계를 포함하며, 여기서 라이브러리의 각 폴리뉴클레오타이드는 제 2 풀의 삽입체 폴리뉴클레오타이드 및 제 1 풀의 한 쌍의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 어셈블리는 시험관 내 클로닝 방법 또는 생체 내 클로닝 방법을 통해 수행되는 것인 방법.

청구항 31

제 30 항에 있어서,

제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 풀의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함하는 것인 방법.

청구항 32

제 30 항 또는 제 31 항에 있어서,

제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 풀의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 약 25개의 뉴클레오타이드를 포함하는 것인 방법.

청구항 33

제 30 항에 있어서,

제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 단일 구조체에서 함께 연결되며, 여기서 단일 구조체는 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드 사이의 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 것인 방법.

청구항 34

제 33 항에 있어서,

하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열은 귀소 엔도뉴클레아제 인식 서열을 포함하는 것인 방법.

청구항 35

제 33 항에 있어서,

연결된 단일 구조체는 스플라이싱 및 오버랩 연장 PCR(SOE-PCR), 제한 결찰, 무딘 말단 결찰, 오버랩 기반 어셈블리 방법, 재조합 기반 방법 또는 제 1 및 제 2 폴리뉴클레오타이드를 결합하는 임의의 다른 효소적 또는 화학적 방법을 통해, 또는 단일 구조체를 직접 합성함으로써 개별적인 제 1 및 제 2 폴리뉴클레오타이드를 연결하여 생성되는 것인 방법.

청구항 36

제 30 항에 있어서,

단계(a) 동안 클로닝 벡터를 제 1 풀 및 제 2 풀과 조합하는 단계를 추가로 포함하며, 여기서 클로닝 벡터의 대향 말단은 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 상보적인 서열을 포함하는 것인 방법.

청구항 37

제 30 항에 있어서,

단계(a) 전에 클로닝 벡터를 제 1 풀과 조합하는 단계를 추가로 포함하며, 여기서 클로닝 벡터의 대향 말단은 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 상보적인 서열을 포함하는 것인 방법.

청구항 38

제 36 항 또는 제 37 항에 있어서,

클로닝 벡터 및 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단은 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 것인 방법.

청구항 39

제 38 항에 있어서,

하나 이상의 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제를 첨가함으로써 클로닝 벡터의 대향 말단과 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 사이에 단일 가닥 상보적 오버행을 생성하는 단계를 추가로 포함하는 것인 방법.

청구항 40

제 39 항에 있어서,

클로닝 벡터의 대향 말단과 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 사이에 단일 가닥 상보적 오버행을 결찰하는 단계를 추가로 포함하는 것인 방법.

청구항 41

제 36 항 또는 제 37 항에 있어서,

단계(b)는 제 2 풀의 삽입체 폴리뉴클레오타이드, 제 1 풀의 쌍의 제 1 및 제 2 폴리뉴클레오타이드 및 클로닝 벡터를 포함하는 원형 산물을 생성하는 것인 방법.

청구항 42

제 36 항 또는 제 37 항에 있어서,

제 1 풀은 제 1 풀의 폴리뉴클레오타이드가 지정된 역치를 넘어 제 1 풀의 다른 폴리뉴클레오타이드와 공통 서열을 공유하지 않아서, 제 1 풀의 폴리뉴클레오타이드의 쌍과 제 2 풀의 삽입체 폴리뉴클레오타이드 사이 또는 제 1 풀의 폴리뉴클레오타이드의 쌍과 클로닝 벡터 사이의 디자인된 어셈블리 오버랩 서열을 제외하도록 이란 서열의 더 큰 세트로부터의 폴리뉴클레오타이드 서열의 쌍을 선택함으로써 생성되는 것인 방법.

청구항 43

제 42 항에 있어서,

지정된 역치는 5개 내지 15개의 연속 뉴클레오타이드인 방법.

청구항 44

제 30 항에 있어서,

어셈블리는 시험관 내 클로닝 방법이며, 여기서 제 1 풀과 제 2 풀의 혼합물은 제 1 및 제 2 풀에 존재하는 폴리뉴클레오타이드를 부분적으로 또는 완전히 변성시키기 위해 가열된 다음, 어셈블리 전에 실온으로 냉각되는 것인 방법.

청구항 45

폴리뉴클레오타이드의 라이브러리를 생성하는 방법으로서, 이 방법은 (a) 증합효소 연쇄 반응(PCR)을 통해 폴리뉴클레오타이드의 제 1 풀을 증폭하는 단계, 여기서 제 1 풀은 폴리뉴클레오타이드의 쌍을 포함하며, 여기서 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 한 쌍의 각 제 1 폴리뉴클레오타이드 및 각 제 2 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 포함하고, 여기서 증폭은 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 공통 오버랩 서열을 제 1 풀로부터의 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 도입한다;

(b) 공통 오버랩 서열을 이용하여 제 1 풀로부터의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드의 각 쌍을 단일 핵산 단편으로 어셈블리하는 단계, 여기서 각 쌍에 대한 단일 핵산 단편은 공통 오버랩 서열에 의해 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단으로부터 분리된 제 1 폴리뉴클레오타이드와 제 2 폴리뉴클레오타이드를 포함하며, 각 쌍에 대한 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단은 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 대해 원위에 있는 단일 핵산 단편의 대향 말단 상에 위치된다;

(c) 각 쌍에 대한 단일 핵산 단편을 삽입체 폴리뉴클레오타이드를 포함하는 제 2 풀과 조합하는 단계, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 단일 핵산 단편 내에 존재하는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열 및 단일 핵산 단편 내에 존재하는 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 대향 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다;

(d) 제 1 풀 및 제 2 풀을 원형화된 산물의 제 3 풀로 어셈블리하는 단계, 여기서 어셈블리는 시험관 내 또는 생체 내 오버랩 어셈블리 방법을 통해 수행되고, 여기서 제 3 풀의 각 원형화된 산물은 제 2 풀로부터의 삽입체 서열 및 제 1 풀로부터의 한 쌍의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함한다;

(e) 제 3 풀의 원형화된 산물의 각각의 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치한 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)을 인식하는 하나 이상의 부위 특이적 뉴클레아제(들)에 의한 소화를 통해 제 3 풀의 각 원형화된 산물을 선형화하는 단계; 및

(f) 시험관 내 또는 생체 내 클로닝 방법에 의해 선형화된 산물을 클로닝 벡터로 어셈블리하는 단계를 포함하는 것인 방법.

청구항 46

제 45 항에 있어서,

제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치한 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)은 귀소 뉴클레아제 인식 서열인 방법.

청구항 47

제 45 항 또는 제 46 항에 있어서,

제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 대한 하나 이상의 부위 특이적 뉴클레아제(들)은 귀소 엔도뉴클레아제인 방법.

청구항 48

제 45 항에 있어서,

공통 오버랩 서열은 적어도 1개의 뉴클레오타이드의 어셈블리 오버랩 서열을 포함하고 단계(b)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법에 의해 수행되는 것인 방법.

청구항 49

제 45 항에 있어서,

공통 오버랩 서열은 10-25개 뉴클레오타이드의 어셈블리 오버랩 서열을 포함하고 단계(b)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법에 의해 수행되는 것인 방법.

청구항 50

제 48 항 또는 제 49 항에 있어서,

오버랩 기반 DNA 어셈블리 방법은 SOE-PCR 또는 시험관 내 오버랩 어셈블리 방법에서 선택되는 것인 방법.

청구항 51

제 50 항에 있어서,

제 1 폴리뉴클레오타이드의 5' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)은 각 쌍의 제 2 폴리뉴클레오타이드의 3' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 상보적이며, 여기서 단계(b)에서 각 쌍의 제 1 및 제 2 폴리뉴클레오타이드의 공통 오버랩 서열을 이용하는 것은 SOE-PCR을 수행하는 것을 수반하는 것인 방법.

청구항 52

제 45 항에 있어서,

단계(b)에서 각 쌍의 제 1 및 제 2 폴리뉴클레오타이드의 공통 오버랩 서열을 이용하는 것은 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하여 상보적 서열을 포함하는 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행을 생성하고; 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행 상에 존재하는 상보적 서열을 결합하는 것을 수반하는 것인 방법.

청구항 53

제 45 항에 있어서,

단계(d)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법을 사용하여 수행되는 것인 방법.

청구항 54

제 53 항에 있어서,

오버랩 기반 DNA 어셈블리는 SOE-PCR 및 시험관 내 오버랩 어셈블리 방법에서 선택되는 것인 방법.

청구항 55

제 45 항에 있어서,

각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단은 하나 이상의 부위 특이적 뉴클레아제 인식 서열의 추가 세트를 포함하고 제 2 폴의 각 삽입체 폴리뉴클레오타이드의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 포함하는 것인 방법.

청구항 56

제 55 항에 있어서,

단계(d)에서의 어셈블리는 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단 상에 존재하는 추가의 하나 이상의 부위 특이적 뉴클레아제 인식 서열 및 제 2 플로부터의 각 삽입체 폴리뉴클레오타이드의 제 1 및 제 2 어셈블리 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단 상에 존재하는 추가의 하나 이상의 부위 특이적 뉴클레아제 인식 서열 및 제 2 플로부터의 각 삽입체 폴리뉴클레오타이드의 제 1 및 제 2 어셈블리 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하여 제 2 플로부터의 삽입체 폴리뉴클레오타이드의 제 1 어셈블리 서열의 5' 말단 상의 단일 가닥 오버행 상에 존재하는 서열에 상보적인 서열을 포함하는 제 1 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행 및 제 2 플로부터의 동일한 삽입체 폴리뉴클레오타이드의 제 2 어셈블리 서열의 3' 말단 상의 단일 가닥 오버행 상에 존재하는 서열에 상보적인 서열을 포함하는 제 2 폴리뉴클레오타이드의 5' 말단 상의 단일 가닥 오버행을 생성하고; 단일 가닥 오버행에 존재하는 상보적 서열을 절찰하는 것을 수반하는 것인 방법.

청구항 57

제 45 항에 있어서,

단계(f)의 클로닝 벡터는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 포함하는 것인 방법.

청구항 58

제 57 항에 있어서,

단계(f)에서의 어셈블리는 클로닝 벡터의 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 클로닝 벡터에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하는 것, 여기서 분해는 클로닝 벡터의 대향 말단 상의 단일 가닥 오버행을 생성하고, 클로닝 벡터의 대향 말단 중 하나 상의 단일 가닥 오버행은 단계(e)에서 생성된 선형화된 산물의 말단에 상보적인 서열을 포함하며 클로닝 벡터의 대향 말단 중 다른 하나 상의 단일 가닥 오버행은 단계(e)에서 생성된 선형화된 산물의 대향 말단에 상보적인 서열을 포함한다; 및 클로닝 벡터의 단일 가닥 오버행 상에 존재하는 상보적 서열과 단계(e)로부터의 선형화된 산물을 절찰하는 것을 수반하는 것인 방법.

청구항 59

제 45 항에 있어서,

제 1 풀은 제 1 풀의 폴리뉴클레오타이드가 지정된 역치를 넘어 제 1 풀의 다른 폴리뉴클레오타이드와 공통 서열을 공유하지 않아서, 제 1 풀의 폴리뉴클레오타이드의 쌍과 제 2 풀의 삽입체 폴리뉴클레오타이드 사이 또는 제 1 풀의 폴리뉴클레오타이드의 쌍과 클로닝 벡터 사이의 디자인된 어셈블리 오버랩 서열을 제외하도록 이런 서열의 더 큰 세트로부터의 폴리뉴클레오타이드 서열의 쌍을 선택함으로써 생성되는 것인 방법.

청구항 60

제 59 항에 있어서,

지정된 역치는 5개 내지 15개의 연속 뉴클레오타이드인 방법.

청구항 61

제 45 항에 있어서,

제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 단일 핵산 단편의 대향 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함하는 것인 방법.

청구항 62

제 45 항에 있어서,

제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 단

일 핵산 단편의 대향 말단에 상보적인 약 25개의 뉴클레오타이드를 포함하는 것인 방법.

청구항 63

제 30 항 또는 제 45 항에 있어서,

단계(a) 전에, 폴리뉴클레오타이드의 제 1 풀은 폴리뉴클레오타이드 쌍으로부터 각각의 제 1 폴리뉴클레오타이드를 함유하는 혼합물을 폴리뉴클레오타이드 쌍으로부터 각각의 제 2 폴리뉴클레오타이드를 함유하는 혼합물과 조합함으로써 생성되는 것인 방법.

청구항 64

제 30 항 또는 제 45 항에 있어서,

제 1 풀의 각 쌍은 이중 가닥 DNA(dsDNA) 또는 단일 가닥 DNA(ssDNA)인 방법.

청구항 65

제 30 항 또는 제 45 항에 있어서,

제 2 풀의 각 삽입체 폴리뉴클레오타이드는 dsDNA 또는 ssDNA인 방법.

청구항 66

제 30 항 또는 제 45 항에 있어서,

제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 방법.

청구항 67

제 30 항 또는 제 45 항에 있어서,

제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 대사 경로의 일부인 유전자에 상응하는 코딩 서열을 포함하는 것인 방법.

청구항 68

제 30 항 또는 제 45 항에 있어서,

제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 기능적 도메인 또는 하나 이상의 단백질에 상응하는 코딩 서열을 포함하는 것인 방법.

청구항 69

제 30 항 또는 제 45 항에 있어서,

제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 어셈블리 오버랩 서열과 제 2 어셈블리 오버랩 서열 사이에 위치한 하나 이상의 페이로드 서열을 포함하는 것인 방법.

청구항 70

제 69 항에 있어서,

하나 이상의 페이로드 서열은 프로모터, 유전자, 조절 서열, 디그론을 암호화하는 핵산 서열, 용해도 태그를 암호화하는 핵산 서열, 종결자, 고유 식별자 서열 또는 이의 일부로부터 선택되는 것인 방법.

청구항 71

제 66 항에 있어서,

제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 제 1 풀의 서로의 쌍과 비교하여 숙주 세포의 상이한 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 방법.

청구항 72

제 66 항에 있어서,

제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 동일한 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 방법.

청구항 73

제 69 항에 있어서,

제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 상이한 것인 방법.

청구항 74

제 69 항에 있어서,

제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 동일한 것인 방법.

청구항 75

제 30 항 또는 제 45 항에 있어서,

제 2 풀의 각 삽입체 폴리뉴클레오타이드는

(i) 페이로드 서열, 정방향 프라이머 및 역방향 프라이머를 포함하는 혼합물에서 증합효소 연쇄 반응(PCR)을 수행하는 단계, 여기서 정방향 프라이머는 5'부터 3'로, 페이로드 서열에 상보적인 하나 이상의 뉴클레오타이드의 짧은 스트레치, 제 1 어셈블리 오버랩 서열, 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열, 제 2 어셈블리 오버랩 서열 및 페이로드 서열에 상보적인 하나 이상의 뉴클레오타이드의 제 2 스트레치를 포함하고 역방향 프라이머는 페이로드 서열 또는 페이로드 서열의 다른 하류 서열에 상보적인 서열을 포함하고, 여기서 PCR은 5'부터 3'로, 페이로드 서열에 상보적인 핵산의 짧은 스트레치, 제 1 어셈블리 오버랩 서열, 하나 이상의 부위 특이적 뉴클레아제 인식 서열, 제 2 어셈블리 오버랩 서열 및 페이로드 서열을 포함하는 PCR 산물을 생성한다;

(ii) 스플라이싱 및 오버랩 연장 PCR(SOE-PCR), 제한 결찰, 무딘 말단 결찰, 오버랩 기반 어셈블리 방법, 재조합 기반 방법 또는 두 DNA 분자를 결합하는 임의의 다른 효소적 또는 화학적 방법으로 이루어진 그룹으로부터 선택된 어셈블리 방법을 통해 PCR 산물을 원형화하는 단계; 및

(iii) 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)을 인식하는 하나 이상의 부위 특이적 뉴클레아제(들)로 원형화된 PCR 산물을 선형화하여 폴리뉴클레오타이드의 제 2 풀을 생성하는 단계에 의해 생성되는 것인 방법.

청구항 76

제 20 항에 있어서,

부위 특이적 뉴클레아제(들)은 제한 엔도뉴클레아제(들), 유형 II 엔도뉴클레아제(들), 귀소 엔도뉴클레아제(들), RNA 유도 뉴클레아제(들), DNA 유도 뉴클레아제(들), 징크 핑거 뉴클레아제(들), TALEN(들) 또는 닉킹 효소(들) 중 하나 이상인 조성물.

청구항 77

제 33 항 또는 제 45 항에 있어서,

부위 특이적 뉴클레아제(들)은 제한 엔도뉴클레아제(들), 유형 II 엔도뉴클레아제(들), 귀소 엔도뉴클레아제(들), RNA 유도 뉴클레아제(들), DNA 유도 뉴클레아제(들), 징크 핑거 뉴클레아제(들), TALEN(들) 또는 닉킹 효소(들) 중 하나 이상인 방법.

발명의 설명

기술 분야

[0001] 본 발명은 단일 반응에서 증폭된 말단 서열과 다중 핵산 분자의 시험관 내 또는 생체 내 어셈블리를 허용하는 단일 가닥 및/또는 이중 가닥 핵산 분자를 결합하기 위한 조성물 및 방법에 관한 것이다. 개시된 방법 및 조성물은 핵산 서열 단편의 결정적 어셈블리에 유용할 수 있고, 예를 들어 원하는 숙주 세포 또는 유기체의 게놈에서 플라스미드, 코스미드 또는 특정 유전자와 같은 임의의 DNA 서열을 편집하는 데 사용될 수 있다.

배경 기술

[0002] 전통적으로, 플라스미드 또는 선형 DNA와 같은 핵산 어셈블리는 결정적 방식으로 한 번에 하나씩 생성되므로 느리고 비용이 많이 들고 노동 집약적일 수 있다. 대조적으로, 복잡한 핵산 어셈블리의 라이브러리를 생성하기 위한 현재의 풀링된 접근 방식은 한 번에 많은 어셈블리를 생성할 수 있지만, 어셈블리의 일부 세트 간의 가능한 모든 조합을 나타내는 라이브러리를 생성한다. 이러한 접근 방식은 어셈블리에 대한 비결정적이고 조합적인 접근 방식이며, 특히 서열의 하위 집합이 어셈블리 반응의 원하는 산물인 상황에서 시간과 노동 집약적이며 비용이 많이 들 수 있다.

발명의 내용

해결하려는 과제

[0003] 따라서, 핵산 어셈블리를 생성하는 전통적인 방법에 내재된 상기한 결점을 겪지 않는 복잡한 핵산 어셈블리를 생성하기 위한 새로운 방법이 당업계에 필요하다.

과제의 해결 수단

[0004] 한 양태에서, 폴리뉴클레오타이드의 혼합물을 포함하는 조성물이 본 발명에 제공되며, 상기 혼합물은 폴리뉴클레오타이드 쌍을 함유하는 제 1 풀을 포함하며, 여기서, 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 함유한다; 및 삽입체 폴리뉴클레오타이드의 제 2 풀을 포함하며, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열을 포함하고 제 1 풀로부터의 한 쌍의 폴리뉴클레오타이드에서 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 이의 반대 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다. 일부 경우에, 조성물은 클로닝 벡터를 추가로 포함하고, 여기서 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단은 클로닝 벡터에 상보적인 서열을 포함한다. 일부 경우에, 제 1 풀의 각 폴리뉴클레오타이드는 제 1 풀의 폴리뉴클레오타이드가 지정된 위치를 넘어 제 1 풀의 다른 폴리뉴클레오타이드와 공통 서열을 공유하지 않아서, 제 1 풀의 폴리뉴클레오타이드의 쌍과 제 2 풀의 삽입체 폴리뉴클레오타이드 사이 또는 제 1 풀의 폴리뉴클레오타이드의 쌍과 클로닝 벡터 사이의 디자인된 어셈블리 오버랩 서열을 제외하도록 선택된다. 일부 경우에, 지정된 위치는 5개 내지 15개의 연속 뉴클레오타이드이다. 일부 경우에, 조성물은 중합 효소를 추가로 포함한다. 일부 경우에, 중합 효소는 가닥 치환 또는 비 가닥 치환이다. 일부 경우에, 중합 효소는 비 가닥 치환이고 조성물은 크라운데를 추가로 포함한다. 일부 경우에, 크라운데는 폴리에틸렌 글리콜(PEG)이다. 일부 경우에, PEG는 약 3 내지 약 7%(중량/부피)의 농도로 사용된다. 일부 경우에, PEG는 PEG-200, PEG-4000, PEG-6000, PEG-8000 또는 PEG-20,000에서 선택된다. 일부 경우에, 중합 효소는 가닥 치환이고 조성물은 단일 가닥 결합 단백질(ET SSB), 대장균 recA, T7 유전자 2.5 산물, 파지 람다 RedB 또는 Rac 프로파지 RecT이다. 일부 경우에, 조성물은 5'-3' 엑소뉴클레아제를 추가로 포함한다. 일부 경우에, 조성물은 리가아제를 추가로 포함한다. 일부 경우에, 제 1 풀의 각 쌍은 이중 가닥 DNA(dsDNA) 또는 단일 가닥 DNA(ssDNA)이다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 dsDNA 또는 ssDNA이다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 표적 게놈 유전자좌에 상응하는 서열을 포함한다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 대사 경로의 일부인 유전자에 상응하는 코딩 서열을 포함한다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 기능적 도메인 또는 하나 이상의 단백질에 상응하는 코딩 서열을 포함한다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 단일 구조체에서 함께 연결되며, 여기서 단일 구조체는 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드 사이의 하나 이

상의 부위 특이적 뉴클레아제(들)에 대한 하나 이상의 인식 서열을 포함한다. 일부 경우에서, 하나 이상의 부위 특이적 뉴클레아제(들)에 대한 하나 이상의 인식 서열은 귀소 엔도뉴클레아제 인식 서열을 포함한다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 풀의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함한다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 풀의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 약 25개의 뉴클레오타이드를 포함한다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 어셈블리 오버랩 서열과 제 2 어셈블리 오버랩 서열 사이에 위치한 하나 이상의 페이로드 서열을 포함한다. 일부 경우에, 하나 이상의 페이로드 서열은 프로모터, 유전자, 조절 서열, 디그론을 암호화하는 핵산 서열, 용해도 태그를 암호화하는 핵산 서열, 종결자, 고유 식별자 서열 또는 이의 일부로부터 선택된다. 일부 경우에, 제 1 풀의 제 1 및 제 2 폴리뉴클레오타이드의 각 쌍은 제 1 풀의 서로의 쌍과 비교하여 숙주 세포의 상이한 표적 게놈 유전자좌에 상응하는 서열을 포함한다. 일부 경우에, 제 1 풀의 제 1 및 제 2 폴리뉴클레오타이드의 각 쌍은 숙주 세포의 동일한 표적 게놈 유전자좌에 상응하는 서열을 포함한다. 일부 경우에, 제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 상이하다. 일부 경우에, 제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 동일하다. 일부 경우에, 부위 특이적 뉴클레아제(들)은 제한 엔도뉴클레아제(들), 유형 II 엔도뉴클레아제(들), 귀소 엔도뉴클레아제(들), RNA 유도 뉴클레아제(들), DNA 유도 뉴클레아제(들), 징크 핑거 뉴클레아제(들), TALEN(들) 또는 니킹 효소(들) 중 하나 이상이다.

[0005]

다른 양태에서, 폴리뉴클레오타이드의 라이브러리를 생성하는 방법이 본 발명에 제공되며, 이 방법은 a.) 폴리뉴클레오타이드의 제 1 풀 및 폴리뉴클레오타이드의 제 2 풀을 조합하는 단계, 여기서 제 1 풀은 폴리뉴클레오타이드의 쌍을 포함하며, 여기서 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 제 2 풀은 삽입체 폴리뉴클레오타이드를 포함하며, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열을 포함하고 제 1 풀로부터의 한 쌍의 폴리뉴클레오타이드에서 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 이의 반대 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다; b.) 제 1 풀 및 제 2 풀을 폴리뉴클레오타이드의 라이브러리로 어셈블리하는 단계를 포함하며, 여기서 라이브러리의 각 폴리뉴클레오타이드는 제 2 풀의 삽입체 폴리뉴클레오타이드 및 제 1 풀의 한 쌍의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 어셈블리는 시험관 내 클로닝 방법 또는 생체 내 클로닝 방법을 통해 수행된다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 풀의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함한다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 풀의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 약 25개의 뉴클레오타이드를 포함한다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 단일 구조체에서 함께 연결되며, 여기서 단일 구조체는 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드 사이의 하나 이상의 부위 특이적 뉴클레아제(들)에 대한 하나 이상의 인식 서열을 포함한다. 일부 경우에서, 하나 이상의 부위 특이적 뉴클레아제(들)에 대한 하나 이상의 인식 서열은 귀소 엔도뉴클레아제 인식 서열을 포함한다. 일부 경우에, 연결된 단일 구조체는 스플라이싱 및 오버랩 연장 PCR(SOE-PCR), 제한 결찰, 무딘 말단 결찰, 오버랩 기반 어셈블리 방법, 재조합 기반 방법 또는 제 1 및 제 2 폴리뉴클레오타이드를 결합하는 임의의 다른 효소적 또는 화학적 방법을 통해, 또는 단일 구조체를 직접 합성함으로써 개별적인 제 1 및 제 2 폴리뉴클레오타이드를 연결하여 생성된다. 일부 경우에, 상기 방법은 단계(a) 동안 클로닝 벡터를 제 1 풀 및 제 2 풀과 조합하는 단계를 추가로 포함하며, 여기서 클로닝 벡터의 대향 말단은 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 상보적인 서열을 포함한다. 일부 경우에, 상기 방법은 단계(a) 전에 클로닝 벡터를 제 1 풀과 조합하는 단계를 추가로 포함하며, 여기서 클로닝 벡터의 대향 말단은 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 상보적인 서열을 포함한다. 일부 경우에, 클로닝 벡터 및 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단은 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함한다. 일부 경우에, 상기 방법은 하나 이상의 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제를 첨가함으로써 클로닝 벡터의 대향 말단과 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오

타이드의 3' 말단 사이에 단일 가닥 상보적 오버행을 생성하는 단계를 추가로 포함한다. 일부 경우에, 상기 방법은 클로닝 벡터의 대향 말단과 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 사이에 단일 가닥 상보적 오버행을 결합하는 단계를 추가로 포함한다. 결합은 DNA 리가아제를 사용하여 수행될 수 있다. 일부 경우에, 단계(b)는 제 2 풀의 삽입체 폴리뉴클레오타이드, 제 1 풀의 쌍의 제 1 및 제 2 폴리뉴클레오타이드 및 클로닝 벡터를 포함하는 원형 산물을 생성한다. 일부 경우에, 제 1 풀은 제 1 풀의 폴리뉴클레오타이드가 지정된 위치를 넘어 제 1 풀의 다른 폴리뉴클레오타이드와 공통 서열을 공유하지 않아서, 제 1 풀의 폴리뉴클레오타이드의 쌍과 제 2 풀의 삽입체 폴리뉴클레오타이드 사이 또는 제 1 풀의 폴리뉴클레오타이드의 쌍과 클로닝 벡터 사이의 디자인된 어셈블리 오버랩 서열을 제외하도록 이런 서열의 더 큰 세트로부터의 폴리뉴클레오타이드 서열의 쌍을 선택함으로써 생성된다. 일부 경우에 지정된 위치는 5개 내지 15개의 연속 뉴클레오타이드이다. 일부 경우에, 어셈블리는 시험관 내 클로닝 방법이며, 여기서 제 1 풀과 제 2 풀의 혼합물은 제 1 및 제 2 풀에 존재하는 폴리뉴클레오타이드를 부분적으로 또는 완전히 변성시키기 위해 가열된 다음, 어셈블리 전에 실온으로 냉각된다. 일부 경우에, 단계(a) 전에, 폴리뉴클레오타이드의 제 1 풀은 폴리뉴클레오타이드 쌍으로부터 각각의 제 1 폴리뉴클레오타이드를 함유하는 혼합물을 폴리뉴클레오타이드 쌍으로부터 각각의 제 2 폴리뉴클레오타이드를 함유하는 혼합물과 조합함으로써 생성된다. 일부 경우에, 제 1 풀의 각 쌍은 이중 가닥 DNA(dsDNA) 또는 단일 가닥 DNA(ssDNA)이다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 dsDNA 또는 ssDNA이다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 표적 게놈 유전자좌에 상응하는 서열을 포함한다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 대사 경로의 일부인 유전자에 상응하는 코딩 서열을 포함한다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 기능적 도메인 또는 하나 이상의 단백질에 상응하는 코딩 서열을 포함한다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 어셈블리 오버랩 서열과 제 2 어셈블리 오버랩 서열 사이에 위치한 하나 이상의 페이로드 서열을 포함한다. 일부 경우에, 하나 이상의 페이로드 서열은 프로모터, 유전자, 조절 서열, 디그론을 암호화하는 핵산 서열, 용해도 태그를 암호화하는 핵산 서열, 종결자, 고유 식별자 서열 또는 이의 일부로부터 선택된다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 제 1 풀의 서로의 쌍과 비교하여 숙주 세포의 상이한 표적 게놈 유전자좌에 상응하는 서열을 포함한다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 동일한 표적 게놈 유전자좌에 상응하는 서열을 포함한다. 일부 경우에, 제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 동일하다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 (i) 페이로드 서열, 정방향 프라이머 및 역방향 프라이머를 포함하는 혼합물에서 중합효소 연쇄 반응(PCR)을 수행하는 단계, 여기서 정방향 프라이머는 5'부터 3'로, 페이로드 서열에 상보적인 하나 이상의 뉴클레오타이드의 짧은 스트레치, 제 1 어셈블리 오버랩 서열, 하나 이상의 부위 특이적 뉴클레아제(들)에 대한 하나 이상의 인식 서열, 제 2 어셈블리 오버랩 서열 및 페이로드 서열에 상보적인 하나 이상의 뉴클레오타이드의 제 2 스트레치를 포함하고 역방향 프라이머는 페이로드 서열에 상보적인 서열을 포함하고, 여기서 PCR은 5'부터 3'로, 페이로드 서열에 상보적인 핵산의 짧은 스트레치, 제 1 어셈블리 오버랩 서열, 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들), 제 2 어셈블리 오버랩 서열 및 페이로드 서열을 포함하는 PCR 산물을 생성한다; (ii) 스플라이싱 및 오버랩 연장 PCR(SOE-PCR), 제한 결합, 무딘 말단 결합, 오버랩 기반 어셈블리 방법, 재조합 기반 방법 또는 두 DNA 분자를 결합하는 임의의 다른 효소적 또는 화학적 방법으로 이루어진 그룹으로부터 선택된 어셈블리 방법을 통해 PCR 산물을 원형화하는 단계; 및 (iii) 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)을 인식하는 하나 이상의 부위 특이적 뉴클레아제(들)로 원형화된 PCR 산물을 선형화하여 폴리뉴클레오타이드의 제 2 풀을 생성하는 단계에 의해 생성된다. 일부 경우에, 부위 특이적 뉴클레아제(들)은 제한 엔도뉴클레아제(들), 유형 II 엔도뉴클레아제(들), 귀소 엔도뉴클레아제(들), RNA 유도 뉴클레아제(들), DNA 유도 뉴클레아제(들), 징크 핑거 뉴클레아제(들), TALEN(들) 또는 니킹 효소(들) 중 하나 이상이다.

[0006]

또 다른 양태에서, 폴리뉴클레오타이드의 라이브러리를 생성하는 방법이 본 발명에 제공되며, 이 방법은 (a) 중합효소 연쇄 반응(PCR)을 통해 폴리뉴클레오타이드의 제 1 풀을 증폭하는 단계, 여기서 제 1 풀은 폴리뉴클레오타이드의 쌍을 포함하며, 여기서 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 한 쌍의 각 제 1 폴리뉴클레오타이드 및 각 제 2 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 포함하고, 여기서 증폭은 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 공통 오버랩 서열을 제 1 풀로부터의 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말

단에 도입한다; (b) 공통 오버랩 서열을 이용하여 제 1 풀로부터의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드의 각 쌍을 단일 핵산 단편으로 어셈블리하는 단계, 여기서 각 쌍에 대한 단일 핵산 단편은 공통 오버랩 서열에 의해 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단으로부터 분리된 제 1 폴리뉴클레오타이드와 제 2 폴리뉴클레오타이드를 포함하며, 각 쌍에 대한 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단은 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 대해 원위에 있는 단일 핵산 단편의 대향 말단 상에 위치된다; (c) 각 쌍에 대한 단일 핵산 단편을 삽입체 폴리뉴클레오타이드를 포함하는 제 2 풀과 조합하는 단계, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 단일 핵산 단편 내에 존재하는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열 및 단일 핵산 단편 내에 존재하는 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 대향 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다; (d) 제 1 풀 및 제 2 풀을 원형화된 산물의 제 3 풀로 어셈블리하는 단계, 여기서 어셈블리는 시험관 내 또는 생체 내 오버랩 어셈블리 방법을 통해 수행되고, 여기서 제 3 풀의 각 원형화된 산물은 제 2 풀로부터의 삽입체 서열 및 제 1 풀로부터의 한 쌍의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함한다; (e) 제 3 풀의 원형화된 산물의 각각의 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치한 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)을 인식하는 하나 이상의 부위 특이적 뉴클레아제(들)에 의한 소화를 통해 제 3 풀의 각 원형화된 산물을 선형화하는 단계; 및 (f) 시험관 내 또는 생체 내 클로닝 방법에 의해 선형화된 산물을 클로닝 벡터로 어셈블리하는 단계를 포함한다. 일부 경우에서, 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치한 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)은 귀소 뉴클레아제 인식 서열이다. 일부 경우에, 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 대한 하나 이상의 부위 특이적 뉴클레아제(들)은 귀소 엔도뉴클레아제이다. 일부 경우에, 공통 오버랩 서열은 적어도 1개의 뉴클레오타이드의 어셈블리 오버랩 서열을 포함하고 단계(b)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법에 의해 수행된다. 일부 경우에, 공통 오버랩 서열은 10-25개 뉴클레오타이드의 어셈블리 오버랩 서열을 포함하고 단계(b)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법에 의해 수행된다. 일부 경우에, 오버랩 기반 DNA 어셈블리 방법은 SOE-PCR 또는 시험관 내 오버랩 어셈블리 방법(예를 들어, HiFi 어셈블리)에서 선택된다. 일부 경우에, 제 1 폴리뉴클레오타이드의 5' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)은 각 쌍의 제 2 폴리뉴클레오타이드의 3' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 상보적이며, 여기서 단계(b)에서 각 쌍의 제 1 및 제 2 폴리뉴클레오타이드의 공통 오버랩 서열을 이용하는 것은 SOE-PCR을 수행하는 것을 수반한다. 일부 경우에, 단계(b)에서 각 쌍의 제 1 및 제 2 폴리뉴클레오타이드의 공통 오버랩 서열을 이용하는 것은 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하여 상보적 서열을 포함하는 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행을 생성하고; 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행 상에 존재하는 상보적 서열을 결찰하는 것을 수반한다. 단계(d)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법을 사용하여 수행된다. 일부 경우에, 오버랩 기반 DNA 어셈블리는 SOE-PCR 및 시험관 내 오버랩 어셈블리 방법(예를 들어, HiFi 어셈블리)에서 선택된다. 일부 경우에, 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단은 하나 이상의 부위 특이적 뉴클레아제 인식 서열의 추가 세트를 포함하고 제 2 풀의 각 삽입체 폴리뉴클레오타이드의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 포함한다. 일부 경우에, 단계(d)에서의 어셈블리는 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단 상에 존재하는 추가의 하나 이상의 부위 특이적 뉴클레아제 인식 서열 및 제 2 풀로부터의 각 삽입체 폴리뉴클레오타이드의 제 1 및 제 2 어셈블리 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단 상에 존재하는 추가의 하나 이상의 부위 특이적 뉴클레아제 인식 서열 및 제 2 풀로부터의 각 삽입체 폴리뉴클레오타이드의 제 1 및 제 2 어셈블리 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하여 제 2 풀로부터의 삽입체 폴리뉴클레오타이드의 제 1 어셈블리 서열의 5' 말단 상의 단일 가닥 오버행 상에 존재하는 서열에 상보적인 서열을 포함하는 제 1 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행 및 제 2 풀로부터의 동일한 삽입체 폴리뉴클레오타이드의 제 2 어셈블리 서열의 3' 말단 상의 단일 가닥 오버행 상에 존재하는 서열에 상보적인 서열을 포함하는 제 2 폴리뉴클레오타이드의 5' 말단 상의 단일 가닥 오버행을 생성하고; 단일 가닥 오버행에 존재하는 상보적 서열을 결찰하는 것을 수반한다. 일부 경우에, 단계(f)의 클로닝 벡터는 하나 이상의 부위

특이적 뉴클레아제 인식 서열을 포함한다. 일부 경우에, 단계(f)에서의 어셈블리는 클로닝 벡터의 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 클로닝 벡터에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하는 것, 여기서 분해는 클로닝 벡터의 대향 말단 상의 단일 가닥 오버행을 생성하고, 클로닝 벡터의 대향 말단 중 하나 상의 단일 가닥 오버행은 단계(e)에서 생성된 선형화된 산물의 말단에 상보적인 서열을 포함하며 클로닝 벡터의 대향 말단 중 다른 하나 상의 단일 가닥 오버행은 단계(e)에서 생성된 선형화된 산물의 대향 말단에 상보적인 서열을 포함한다; 및 클로닝 벡터의 단일 가닥 오버행 상에 존재하는 상보적 서열과 단계(e)로부터의 선형화된 산물을 결합하는 것을 수반한다. 일부 경우에, 제 1 풀은 제 1 풀의 폴리뉴클레오타이드가 지정된 위치를 넘어 제 1 풀의 다른 폴리뉴클레오타이드와 공통 서열을 공유하지 않아서, 제 1 풀의 폴리뉴클레오타이드의 쌍과 제 2 풀의 삽입체 폴리뉴클레오타이드 사이 또는 제 1 풀의 폴리뉴클레오타이드의 쌍과 클로닝 벡터 사이의 디자인된 어셈블리 오버랩 서열을 제외하도록 이런 서열의 더 큰 세트로부터의 폴리뉴클레오타이드 서열의 쌍을 선택함으로써 생성된다. 일부 경우에 지정된 위치는 5개 내지 15개의 연속 뉴클레오타이드이다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 단일 핵산 단편의 대향 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함한다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 단일 핵산 단편의 대향 말단에 상보적인 약 25개의 뉴클레오타이드를 포함한다. 일부 경우에, 단계(a) 전에, 폴리뉴클레오타이드의 제 1 풀은 폴리뉴클레오타이드 쌍으로부터 각각의 제 1 폴리뉴클레오타이드를 함유하는 혼합물을 폴리뉴클레오타이드 쌍으로부터 각각의 제 2 폴리뉴클레오타이드를 함유하는 혼합물과 조합함으로써 생성된다. 일부 경우에, 제 1 풀의 각 쌍은 이중 가닥 DNA(dsDNA) 또는 단일 가닥 DNA(ssDNA)이다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 dsDNA 또는 ssDNA이다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 표적 게놈 유전자좌에 상응하는 서열을 포함한다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 대사 경로의 일부인 유전자에 상응하는 코딩 서열을 포함한다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 기능적 도메인 또는 하나 이상의 단백질에 상응하는 코딩 서열을 포함한다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 어셈블리 오버랩 서열과 제 2 어셈블리 오버랩 서열 사이에 위치한 하나 이상의 페이로드 서열을 포함한다. 일부 경우에, 하나 이상의 페이로드 서열은 프로모터, 유전자, 조절 서열, 디그론을 암호화하는 핵산 서열, 용해도 태그를 암호화하는 핵산 서열, 종결자, 고유 식별자 서열 또는 이의 일부로부터 선택된다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 제 1 풀의 서로의 쌍과 비교하여 숙주 세포의 상이한 표적 게놈 유전자좌에 상응하는 서열을 포함한다. 일부 경우에, 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 동일한 표적 게놈 유전자좌에 상응하는 서열을 포함한다. 일부 경우에, 제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 상이하다. 일부 경우에, 제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 동일하다. 일부 경우에, 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 (i) 페이로드 서열, 정방향 프라이머 및 역방향 프라이머를 포함하는 혼합물에서 중합효소 연쇄 반응(PCR)을 수행하는 단계, 여기서 정방향 프라이머는 5'부터 3'로, 페이로드 서열에 상보적인 하나 이상의 뉴클레오타이드의 짧은 스트레치, 제 1 어셈블리 오버랩 서열, 하나 이상의 부위 특이적 뉴클레아제(들)에 대한 하나 이상의 인식 서열, 제 2 어셈블리 오버랩 서열 및 페이로드 서열에 상보적인 하나 이상의 뉴클레오타이드의 제 2 스트레치를 포함하고 역방향 프라이머는 페이로드 서열 또는 페이로드 서열의 다른 서열 하부에 상보적인 서열을 포함하고, 여기서 PCR은 5'부터 3'로, 페이로드 서열에 상보적인 핵산의 짧은 스트레치, 제 1 어셈블리 오버랩 서열, 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들), 제 2 어셈블리 오버랩 서열 및 페이로드 서열을 포함하는 PCR 산물을 생성한다; (ii) 스플라이싱 및 오버랩 연장 PCR(SOE-PCR), 제한 결찰, 무딘 말단 결찰, 오버랩 기반 어셈블리 방법, 재조합 기반 방법 또는 두 DNA 분자를 결합하는 임의의 다른 효소적 또는 화학적 방법으로 이루어진 그룹으로부터 선택된 어셈블리 방법을 통해 PCR 산물을 원형화하는 단계; 및 (iii) 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)을 인식하는 하나 이상의 부위 특이적 뉴클레아제(들)로 원형화된 PCR 산물을 선형화하여 폴리뉴클레오타이드의 제 2 풀을 생성하는 단계에 의해 생성된다. 일부 경우에, 부위 특이적 뉴클레아제(들)은 제한 엔도뉴클레아제(들), 유형 II 엔도뉴클레아제(들), 귀소 엔도뉴클레아제(들), RNA 유도 뉴클레아제(들), DNA 유도 뉴클레아제(들), 징크 핑거 뉴클레아제(들), TALEN(들) 또는 닉킹 효소(들) 중 하나 이상이다.

발명의 효과

[0007] 본 발명의 내용 중에 포함되어 있다.

도면의 간단한 설명

[0008] 도 1은 삽입체 폴리뉴클레오타이드(들) 및 벡터 오버랩 어셈블리 서열을 포함하는 제 1 폴리뉴클레오타이드 및 벡터 오버랩 어셈블리 서열과 선택적 클로닝 벡터를 포함하는 제 2 폴리뉴클레오타이드의 초기 조성물을 보여주는 DNA 라이브러리의 다중화 결정적 어셈블리 방법을 도시한다. 삽입체 폴리뉴클레오타이드는 결실 인 경우 0 뉴클레오타이드 길이, 삽입 또는 대체인 경우 0이 아닌 길이를 갖는 페이로드 서열을 포함할 수 있다.

도 2는 최대 합성 올리고뉴클레오타이드 길이보다 긴 삽입체 폴리뉴클레오타이드(예를 들어, 프로모터)를 어셈블링을 허용하는 도 1의 방법에서 사용하기 위해 제 1 및 제 2 폴리뉴클레오타이드를 사전-결합하기 위한 인사이드-아웃 어셈블리 방법을 예시한다.

도 3은 최대 합성 올리고뉴클레오타이드 길이보다 긴 삽입체 폴리뉴클레오타이드(예를 들어, 프로모터)를 어셈블링하는 것을 허용하는 도 1의 방법의 적용을 예시한다.

도 4는 도 1의 방법을 사용한 결정적 라이브러리의 어셈블리를 예시한다. 각 라이브러리에서 고유한 유전자좌, 페이로드 및 총 가능한 구조체의 수는 표에 제공된다.

도 5는 원형 치환된 페이로드(삽입체)를 포함하는 정밀하게 디자인된 DNA 일부를 포함하는 풀을 사용하는 결정적 라이브러리의 성공적인 시험관 내 어셈블리의 결과를 예시한다. 상단의 긴 막대는 풀에 어셈블링될 플라스미드의 구조를 나타내고 하단의 짧은 막대는 어셈블리 풀에서 분리된 세 개의 샘플에 대한 상응하는 참조 서열에 정렬된 생거 서열을 나타낸다. 리드(reads)의 내부 말단에 있는 희미한 수직선은 생거 리드의 꼬리 말단에서 예상되는 시퀀싱 아티팩트(sequencing artifacts)를 나타낸다.

도 6은 풀링된 어셈블리에 대한 총 성공률을 예시하며 이에 대한 일부를 함유하는 페이로드는 숙주 계능에서 유래된 템플릿으로부터의 어셈블리 오버랩을 추가하는 프라이머를 사용하여 PCR을 통해 생성된다.

발명을 실시하기 위한 구체적인 내용

[0009] 정의

[0010] 다음의 용어는 당해 기술 분야에서 통상의 지식을 가진 자에 의해 잘 이해되는 것으로 생각되지만, 다음 정의는 본 발명에 개시된 주제의 설명을 용이하게 하기 위해 제시된다.

[0011] 본 발명에 사용된 용어 하나("a" 또는 "an")는 그 실체 중 하나 이상을 의미하며, 즉 복수의 지시대상을 의미할 수 있다. 이와 같이, 용어 "하나", "하나 이상" 및 "적어도 하나"라는 본 발명에서 상호 교환적으로 사용될 수 있다. 또한, 부ظ관사에 의한 "한 요소"에 대한 언급은, 내용이 분명하게는 요소의 하나 및 단지 하나가 존재하는 것을 요구하지 않는 한, 하나 이상의 요소가 존재하는 가능성을 배제하지 않는다.

[0012] 문맥이 달리 요구하지 않는 한, 본 명세서 및 청구 범위 전체에 걸쳐, 단어 "포함한다(comprise)" 및 "포함한다(comprises)" 및 "포함하는"(comprising)과 같은 이의 변형은 "포함하나 이에 제한되지 않는다"와 같은 개방적이고 포괄적인 의미로 해석되어야 한다.

[0013] 본 명세서 전체에서 "일 실시태양" 또는 "한 실시태양"에 대한 언급은 실시태양과 관련하여 기술된 특정 특징, 구조 또는 특성이 본 발명의 적어도 하나의 실시태양에 포함될 수 있음을 의미한다. 따라서, 본 명세서 전반에 걸쳐 다양한 장소에서 "일 실시태양에서" 또는 "한 실시태양에서"라는 문구의 출현은 반드시 모두 동일한 실시태양을 지칭하는 것은 아니다. 명확성을 위해 별도의 실시태양의 맥락에서 기술된 본 발명의 특정 특징은 또한 단일 실시태양에서 조합되어 제공될 수 있다는 것이 이해된다. 반대로, 간결함을 위해, 단일 실시태양의 맥락에서 기술된 본 발명의 다양한 특징은 또한 개별적으로 또는 임의의 적절한 하위 조합으로 제공될 수 있다.

[0014] 본 발명에 사용된 용어 "세포 생물", "미세유기체" 또는 "미생물"은 광범위하게 이해되어야 한다. 이 용어들은 상호 교환적으로 사용되나 두 원핵생물 영역인 박테리아와 고세균 뿐만 아니라 진핵 균류와 원생 생물을 포함하나 이에 제한되지 않는다. 일부 실시태양에서, 본 발명은 본 발명에 제공된 목록/표 및 도면의 "미세유기체" 또는 "세포유기체" 또는 "미생물"을 의미한다. 이런 특성화는 표와 도면의 확인된 분류학적 속과 확인된 분류학적 종뿐만 아니라 상기 표 또는 도면의 임의의 유기체의 다양한 신규하고 새로운 확인된 또는 디자인된 균주를 의미할 수 있다. 동일한 특성화는 실시예에서와 같이, 명세서의 다른 부분에서 이런 용어의 인용에 대해서도 마찬가지이다.

- [0015] 본 발명에 사용된 용어 "원핵 생물"은 당업계에 공지되어 있으며 핵 또는 다른 세포 기관을 함유하지 않는 세포를 의미한다. 원핵 생물은 일반적으로 두 영역, 박테리아와 고세균의 하나로 분류된다. 고세균과 박테리아 영역의 유기체 사이의 명확한 차이는 16S 리보솜 RNA에서 뉴클레오타이드 염기 서열의 근본적인 차이에 기초한다.
- [0016] 본 발명에 사용된 용어 "고세균"은 전형적으로 특이한 환경에서 발견되고 세포벽에서 리보솜 단백질의 수 및 유라민산의 부족을 포함하는 몇몇 기준에 의해 나머지 원핵 생물과 구별되는 멘도시쿠테스(Mendosicutes) 문의 유기체의 범주를 의미한다. *ssrRNA* 분석에 기초하여, 고세균은 두 계통발생학적으로 다른 그룹으로 구성된다: 크렌고세균(Crenarchaeota) 및 유리고세균(Euryarchaeota). 생리학을 기초로, 고세균은 세 가지 유형으로 구성될 수 있다: 메테인 생성균(methanogens)(메테인을 생성하는 원핵 생물); 고염성 세균(extreme halophiles)(매우 높은 농도의 염(NaCl)에서 사는 원핵 생물; 및 고온성(초고온성) 세균(extreme(hyper) thermophilus)(초고온에서 사는 원핵 생물). 박테리아와 구별되는 통일된 고세균의 특징(즉, 세포벽, 에스터-연결 막 지질 등에 유레인 없음) 이외에, 이런 원핵 생물은 이들의 특정한 서식 환경에 적응시키는 특이한 구조 또는 생화학적 특성을 나타낸다. 크렌고세균은 주로 초고온성 황 의존성 원핵 생물로 이루어지며 유리고세균은 메테인 생성균과 고염성 세균을 함유한다.
- [0017] 본 발명에 사용된 "박테리아" 또는 "진정세균(eubacteria)"는 원핵 생물의 영역을 의미할 수 있다. 박테리아는 다음과 같이 적어도 11개의 구별된 그룹을 포함한다: (1) 그람 양성 (그람+) 박테리아, 2개위 주요 세부구분이 존재한다: (1) 높은 G+C 그룹(액티노마이세테스, 마이코박테리아, 마이코코쿠스, 기타) (2) 낮은 G+C 그룹(바실러스, 클로스트리디아, 락토바실러스, 스타필로코키, 스트렙토코키, 마이코플라스마스); (2) 프로테오박테리아, 예를 들어, 보라색 광합성 + 비 광합성 그람 음성 박테리아 (대부분의 "일반적인" 그람 음성 박테리아 포함); (3) 사이아노박테리아, 예를 들면, 산소성 광영양생물; (4) 스피로체테스 및 관련 종; (5) 플랑크토마이세스; (6) 박테로이데스, 플라보박테리아; (7) 클라미디아; (8) 녹색 황 박테리아; (9) 녹색 비 황 박테리아(또한 혐기성 광영양생물); (10) 방사성 저항성 마이크로코키 및 동족; (11) 써모토가(Thermotoga) 및 써모시포 써모필레스(Thermosipho thermophiles).
- [0018] 본 발명에 사용된 "진핵 생물"은 세포가 막 내에 둘러싸인 핵 및 다른 세포 기관을 함유하는 임의의 유기체이다. 진핵 생물은 분류군(Eukarya 또는 Eukaryota)에 속한다. 진핵 생물 세포를 원핵 생물(상기 박테리아와 고세균)와 구분시키는 정의하는 특징은 막으로 둘러싸인 유전자 물질, 특히 유전자 물질을 함유하고 핵막에 의해 둘러싸인 핵을 가진다는 것이다.
- [0019] 본 발명에 사용된 용어 "유전자 변형된 숙주 세포", "재조합 숙주 세포" 및 "재조합 균주"는 본 발명에서 상호 교환적으로 사용되고 본 발명의 복제 및 변형 방법에 의해 유전자 변형된 숙주 세포를 의미할 수 있다. 따라서, 이 용어는 유전자 변경, 변형 또는 조작되어, 숙주 세포가 유도된 자연 발생 유기체와 비교하여 변경, 변형 또는 상이한 유전자형 및/또는 표현형을 나타내는(예를 들어, 유전자 변형이 미생물의 핵산 서열 암호화에 영향을 미칠 때) 세포(예를 들어, 박테리아, 효모 세포, 곰팡이 세포, CHO, 인간 세포 등)를 포함한다. 일부 실시태양에서, 이 용어는 문제의 특정 재조합 숙주 세포뿐만 아니라 이런 숙주 세포의 자손 또는 잠재적 자손을 의미하는 것으로 이해된다.
- [0020] 본 발명에 사용된 용어 "야생형 미생물" 또는 "야생형 숙주 세포"는 자연에서 발생하는 세포, 즉 유전자 변형되지 않은 세포를 기술할 수 있다.
- [0021] 본 발명에 사용된 용어 "유전자 조작된"은 (예를 들어, 핵산의 삽입, 결실, 돌연변이 또는 대체에 의한) 숙주 세포 게놈의 임의적 조작을 의미할 수 있다.
- [0022] 본 발명에 사용된 용어 "대조군" 또는 "대조군 숙주 세포"는 유전자 변형 또는 실험적 처리의 효과를 측정하기 위한 적절한 비교기 숙주 세포를 의미한다. 일부 실시태양에서, 대조군 숙주 세포는 야생형 세포이다. 다른 실시태양에서, 대조군 숙주 세포는 처리 숙주 세포를 분화시키는 유전자 변형(들)을 제외하고, 유전자 변형된 숙주 세포와 유전적으로 동일하다. 일부 실시태양에서, 본 발명은 대조군 숙주 세포(예를 들어, 균주 개선 프로그램의 기초로 사용된 S₁ 균주)로서 모 균주의 사용을 고시한다. 다른 실시태양에서, 숙주 세포는 치료 숙주 세포에서 테스트되는 특정 프로모터 또는 SNP가 결여된 유전적으로 동일한 세포일 수 있다.
- [0023] 본 발명에 사용된 용어 "대립 유전자(들)"은 유전자의 하나 이상의 대안 형태의 임의의 것을 의미할 수 있고, 이의 모두 대립 유전자는 적어도 하나의 형질 또는 특성과 관련된다. 2배체 세포에서, 주어진 유전자의 2개의 대립 유전자는 한 쌍의 상동 염색체상의 상응하는 유전자좌를 차지한다.
- [0024] 본 발명에 사용된 용어 "유전자좌(locus)"(복수형 loci)는 천연 게놈 서열에 대한 편집이 요구되는 임의의 위치

를 의미할 수 있다. 한 실시태양에서, 상기 용어는 예를 들어 유전자 또는 유전자 마커가 발견되는 염색체 상의 특정 장소 또는 장소들 또는 위치를 의미할 수 있다.

- [0025] 본 발명에 사용된 용어 "유전적으로 연결된"은 교차를 통해 분리하기가 어려워 변식 동안 높은 비율로 공동유전되는 2개 이상의 형질을 의미할 수 있다.
- [0026] 본 발명에 사용된 바와 같이 "재조합" 또는 "재조합 사건"은 염색체 교차 또는 독립된 분류를 의미할 수 있다.
- [0027] 본 발명에 사용된 용어 "표현형"은 개체의 유전적 구성(즉, 유전자형)과 환경 사이의 상호작용으로부터 기인하는 개별 세포, 세포 배양, 유기체 또는 유기체의 그룹의 관찰 가능한 특성을 의미할 수 있다.
- [0028] 본 발명에 사용된 바와 같이, 핵산 서열 또는 단백질 서열을 기술할 때 용어 "키메라" 또는 "재조합체"는 적어도 2개의 이종 폴리뉴클레오타이드 또는 2개의 이종 폴리펩타이드를 단일 거대분자 속에 연결하거나 적어도 하나의 천연 핵산 또는 단백질 서열의 하나 이상의 요소를 배열하는 핵산 또는 단백질 서열을 의미할 수 있다. 예를 들어, 용어 "재조합체"는 예를 들어, 화학적 합성 또는 유전 공학 기술에 의한 핵산의 분리된 단편의 조작에 의해 서열의 두 개의 분리된 단편의 인공적 조합을 의미할 수 있다.
- [0029] 본 발명에 사용된 바와 같이 "합성 뉴클레오타이드 서열" 또는 "합성 폴리뉴클레오타이드 서열"은 자연에서 발생하는 것으로 알려지지 않았거나 자연 발생적이지 않은 뉴클레오타이드 서열이다. 일반적으로, 이런 합성 뉴클레오타이드 서열은 임의의 다른 자연 발생 뉴클레오타이드 서열과 비교할 때 적어도 하나의 뉴클레오타이드 차이를 포함한다.
- [0030] 본 발명에 사용된 용어 "핵산"은 임의의 길이의 중합체 형태의 리보뉴클레오타이드 또는 데옥시리보뉴클레오타이드 또는 이의 유사체를 의미할 수 있다. 이 용어는 분자의 1차 구조를 의미하고, 따라서 이중 나선 및 단일 가닥의 DNA뿐 아니라 이중 및 단일 가닥의 RNA를 포함한다. 또한, 메틸화 및/또는 캡핑된 핵산과 같은 변형된 핵산, 변형된 염기를 함유하는 핵산, 골격 변형 등과 같은 변형 핵산을 포함한다. 용어 "핵산" 및 "뉴클레오타이드 서열"은 상호 교환적으로 사용된다.
- [0031] 본 발명에 사용된 용어 "유전자"는 생물학적 기능과 관련된 DNA의 임의의 단편을 의미할 수 있다. 따라서, 유전자는 암호화 서열 및/또는 그의 발현에 요구되는 조절 서열을 포함하나 이에 제한되지 않는다. 유전자는 또한, 예를 들어, 다른 단백질에 대한 인식 서열을 형성하는 비 발현 DNA 단편을 포함할 수 있다. 유전자는 관심 공급원으로부터의 클로닝 또는 공지되거나 예측된 서열 정보로부터의 합성을 포함하는 다양한 공급원으로부터 얻을 수 있고, 원하는 파라미터를 갖도록 디자인된 서열을 포함할 수 있다.
- [0032] 본 발명에 사용된 용어 "상동의" 또는 "상동체" 또는 "오솔로그"(ortholog 또는 orthologue)는 당업계에 공지되어 있고 공통 조상 또는 가족 구성원을 공유하고 서열 동일성의 정도에 기초하여 결정되는 관련 서열을 의미할 수 있다.
- [0033] 용어 "상동성", "상동의", "실질적으로 유사" 및 "상응하게 실질적으로"는 본 발명에서 상호 교환적으로 사용된다. 상기 용어는 하나 이상의 뉴클레오타이드 염기의 변화가 유전자 발현을 중재하거나 특정 표현형을 생성시키는 핵산 단편의 능력에 영향을 미치지 않는 핵산 단편을 의미할 수 있다. 이런 용어는 또한 초기의 변형되지 않은 단편에 비해 생성된 핵산 단편의 기능적 특성을 실질적으로 변화시키지 않는 하나 이상의 뉴클레오타이드의 결실 또는 삽입과 같은 본 발명의 핵산 단편의 변형을 의미할 수 있다. 따라서, 당업자가 알 수 있는 바와 같이, 본 발명은 특정 예시적인 서열 이상을 포함하는 것으로 이해된다. 이런 용어는 한 종, 아종, 품종, 품종 또는 균주에서 발견된 유전자 및 다른 종, 아종, 품종, 품종 또는 균주에서 상응하는 또는 동등한 유전자 사이의 관계를 기술한다. 본 발명을 위해서, 상동성 서열이 비교될 수 있다.
- [0034] "상동 서열" 또는 "상동체" 또는 "오솔로그"는 기능적으로 관련이 있다고 생각되고, 믿거나 알려져 있다. 기능적 관계는 (a) 서열 동일성 및/또는 (b) 동일하거나 유사한 생물학적 기능을 포함하나 이에 제한되지 않는 다수의 방식 중 임의의 하나로 표시될 수 있다. 바람직하게는, (a) 및 (b) 모두가 표시된다. 아미노산 또는 핵산 서열 간의 서열 상동성은 공유 조상으로 정의할 수 있다. 핵산의 두 세그먼트는 종 분화 이벤트(오솔로그) 또는 중복 이벤트(파라로그)로 인해 공유된 조상을 가질 수 있다. 아미노산 또는 핵산 서열 간의 상동성은 아미노산 또는 핵산 서열이 상동성이라고 말하고 아미노산 또는 핵산 서열이 상당한 유사성을 공유하도록 서열 유사성으로부터 추론될 수 있다. 중요한 유사성은 두 서열이 공통 조상으로부터 다른 진화에 의해 관련되어 있다는 강력한 증거가 될 수 있다. 여러 서열의 정렬을 사용하여 상동 영역을 발견할 수 있다. 상동성은 Current Protocols in Molecular Biology(F.M. Ausubel et al., eds., 1987) Supplement 30, 섹션 7.718, 표 7.71에서 논의된 바와 같은 당해분야에서 용이하게 이용 가능한 소프트웨어 프로그램을 사용하여 결정될 수 있다. 일부 정렬 프로

그램은 BLAST(NCBI), 맥벡터(MacVector)(Oxford Molecular Ltd, Oxford, U.K.), ALIGN 플러스(Plus)(Scientific and Educational Software, Pennsylvania) 및 AlignX(Vector NTI, Invitrogen, Carlsbad, CA)이다. 다른 정렬 프로그램은 기본 매개변수를 사용하는 시퀀처(Sequencher)(Gene Codes, Ann Arbor, Michigan)이다.

- [0035] 본 발명에 사용된 용어 "내인성" 또는 "내인성 유전자"는 숙주 세포 게놈 내에서 발견되는 위치에서 자연적으로 발생하는 유전자를 의미할 수 있다. 본 발명의 내용에서, 이중 프로모터를 내인성 유전자에 작동 가능하게 연결시키는 것은 그 유전자가 자연적으로 존재하는 위치에서, 현존하는 유전자 앞에 이중 프로모터 서열을 유전적으로 삽입하는 것을 의미한다. 본 발명에 기술된 내인성 유전자는 본 발명의 임의의 방법에 따라 돌연변이된 자연 발생 유전자의 대립 유전자를 포함할 수 있다.
- [0036] 본 발명에 사용된 용어, "외인성"은 "이중성"이라는 용어와 상호 교환적으로 사용되고, 천연 출처 이외의 일부 출처로부터 유래하는 물질을 의미한다. 예를 들어, 용어 "외인성 단백질" 또는 "외인성 유전자"는 비천연 출처 또는 위치로부터이며 인위적으로 생물학적 시스템에 공급된 단백질 또는 유전자를 의미한다.
- [0037] 본 발명에 사용된 용어 "뉴클레오타이드 변화"는 당업계에서 잘 알려진 바와 같이, 예를 들어 뉴클레오타이드 치환, 결실 및/또는 삽입을 의미한다. 예를 들어 돌연변이는 침묵 치환, 추가 또는 결실을 생성하나 암호화된 단백질의 특성 또는 활성 또는 단백질이 어떻게 만들어지는지를 변형하지 않는 변경을 함유한다. 대안적으로, 돌연변이는 암호화된 단백질의 아미노산 서열을 변경시킬 수 있는 비동의적 치환 또는 변화일 수 있고 단백질의 특성 또는 활성의 변경을 초래할 수 있다.
- [0038] 본 발명에서 사용된 용어 "단백질 변형"은 당업계에 잘 알려진 바와 같이, 예를 들어 아미노산 치환, 아미노산 변형, 결실 및 또는 삽입을 의미할 수 있다.
- [0039] 본 발명에 사용된 용어 핵산 또는 폴리펩타이드의 "적어도 일부" 또는 "단편"은 전장 분자를 포함하는 전장 분자의 이런 서열 또는 임의의 더 큰 단편의 최소 크기 특성을 갖는 부분을 의미할 수 있다. 본 발명의 폴리뉴클레오타이드 단편은 유전자 조절 요소의 생물학적 활성 부분을 암호화할 수 있다. 유전자 조절 요소의 생물학적 활성 부분은 유전자 조절 요소를 포함하는 본 발명의 폴리뉴클레오타이드 중 하나의 일부를 분리하고 본 발명에 기재된 바와 같은 활성을 평가함으로써 제조될 수 있다. 유사하게, 폴리펩타이드의 일부는 전장 폴리펩타이드까지 이르는 4개 아미노산, 5개 아미노산, 6개 아미노산, 7개 아미노산 동일 수 있다. 사용될 부분의 길이는 특정 용도에 따라 다를 것이다. 하이브리드화 프로브로서 유용한 핵산의 일부는 12개 뉴클레오타이드 정도로 짧을 수 있으며; 일부 실시태양에서, 이것은 20개 뉴클레오타이드이다. 에피토프로서 유용한 폴리펩타이드의 일부는 4개의 아미노산 정도로 짧을 수 있다. 전장 폴리펩타이드의 기능을 수행하는 폴리펩타이드의 일부는 일반적으로 4개 이상의 아미노산보다 길 수 있다.
- [0040] 변이체 폴리뉴클레오타이드는 또한 DNA 서플링과 같은 돌연변이 및 재조합 절차로부터 유도될 수 있는 서열을 포함할 수 있다. 그러한 DNA 서플링을 위한 전략은 당업계에 공지되어 있다. 예를 들어, Stemmer(1994) PNAS 91:10747-10751; Stemmer(1994) Nature 370:389-391; Cramer *et al.*(1997) Nature Biotech. 15:436-438; Moore *et al.*(1997) J. Mol. Biol. 272:336-347; Zhang *et al.*(1997) PNAS 94:4504-4509; Cramer *et al.*(1998) Nature 391:288-291; 및 미국 특허 제5,605,793호 및 제5,837,458호 참조.
- [0041] 본 발명에 개시된 PCR 증폭을 위해, 임의의 관심 유기체로부터 추출된 cDNA 또는 게놈 DNA로부터의 상응하는 DNA 서열을 증폭시키기 위한 PCR 반응에 사용하기 위해 올리고뉴클레오타이드 프라이머가 디자인될 수 있다. PCR 프라이머 및 PCR 클로닝을 디자인하기 위한 방법은 당업계에 일반적으로 공지되어 있으며, Sambrook *et al.*(2001) Molecular Cloning: A Laboratory Manual (3rd ed., Cold Spring Harbor Laboratory Press, Plainview, New York). See also Innis *et al.*, eds. (1990) PCR Protocols: A Guide to Methods and Applications (Academic Press, New York); Innis and Gelfand, eds. (1995) PCR Strategies (Academic Press, New York); and Innis and Gelfand, eds. (1999) PCR Methods Manual (Academic Press, New York)에 개시된다. PCR의 공지된 방법은 쌍을 이룬 프라이머, 네스티드 프라이머, 단일 특이적 프라이머, 축퇴성 프라이머, 유전자 특이적 프라이머, 벡터 특이적 프라이머, 부분적 불일치 프라이머 등을 사용하는 방법을 포함하나 이에 제한되지 않는다.
- [0042] 본 발명에 사용된 용어 "프라이머"는 DNA 중합효소를 부착시키는 증폭 표적에 대한 어닐링을 행할 수 있어, 프라이머 연장 생성물의 합성이 유도되는 조건하에 놓일 때, 즉, 뉴클레오타이드 및 DNA 중합효소와 같은 중합화제의 존재하에서 및 적절한 온도 및 pH에서 DNA 합성의 개시점으로서 작용하는 올리고뉴클레오타이드를 의미할

수 있다. (증폭) 프라이머는 증폭 효율을 최대화하기 위해 단일 가닥일 수 있다. 프라이머는 올리고데옥시리보뉴클레오타이드일 수 있다. 프라이머는 중합화제의 존재 하에서 증량 생성물의 합성을 시작하기에 충분히 길어야 한다. 프라이머의 정확한 길이는 프라이머의 온도 및 조성(A/T 대 G/C 함량)을 포함하는 많은 요소에 따라 달라질 것이다. 한 쌍의 양방향성 프라이머는 PCR 증폭과 같은 DNA 증폭 기술 분야에서 일반적으로 사용되는 것과 같은 하나의 순방향 및 역방향 프라이머로 구성된다.

[0043] 본 발명에 사용된 용어 "프로모터"는 암호화 서열 또는 기능성 RNA의 발현을 제어할 수 있는 DNA 서열을 의미할 수 있다. 일부 실시태양에서, 프로모터 서열은 근위 및 더 원위 상류 요소로 구성되고, 후자의 요소는 종종 인헨서로 불린다. 따라서, "인헨서"는 프로모터 활성을 자극할 수 있는 DNA 서열이며, 프로모터의 선천적인 요소 또는 프로모터의 수준 또는 조직 특이성을 향상시키기 위해 삽입된 이중성 요소일 수 있다. 프로모터는 천연 유전자로부터 완전히 유도되거나 자연계에서 발견되는 다른 프로모터로부터 유도된 상이한 요소로 구성되거나 심지어 합성 DNA 단편을 포함할 수 있다. 당업자는 상이한 프로모터가 상이한 조직 또는 세포 유형에서, 또는 상이한 발달 단계에서, 또는 상이한 환경 조건에 반응하여 유전자의 발현을 지시할 수 있음을 이해한다. 대부분의 경우 조절 서열의 정확한 경계가 완전히 정의되지 않았기 때문에, 일부 변이의 DNA 단편이 동일한 프로모터 활성을 가질 수 있음이 추가로 인식된다.

[0044] 본 발명에서 사용된 용어 "제조합 구조체", "발현 구조체", "키메라 구조체", "구조체" 및 "제조합 DNA 구조체"는 본 발명에서 상호 교환적으로 사용된다. 제조합 구조체는 천연에서 함께 발견되지 않는 조절 및 암호화 서열과 같은 핵산 단편의 인위적인 조합을 포함한다. 예를 들어, 키메라 구조체는 상이한 공급원으로부터 유도된 조절 서열 및 암호화 서열, 또는 동일한 공급원으로부터 유도되지만 자연계에서 발견되는 것과 상이한 방식으로 배열된 조절 서열 및 암호화 서열을 포함할 수 있다. 이러한 구조체는 그 자체로 사용되거나 벡터와 함께 사용될 수 있다. 벡터가 사용되는 경우 벡터의 선택은 당업자에게 주지된 바와 같이 숙주 세포를 형질 전환하는데 사용될 방법에 의존한다. 예를 들어, 플라스미드 벡터가 사용될 수 있다. 당업자는 본 발명의 분리된 핵산 단편을 포함하는 숙주 세포를 성공적으로 형질전환, 선별 및 증식시키기 위해 벡터 상에 존재해야 하는 유전자 요소를 잘 알고 있다. 당업자는 또한 상이한 독립적인 형질전환 사건이 발현의 상이한 수준 및 패턴을 유도한다는 것을 인식할 것이며(Jones *et al.*, (1985) EMBO J. 4 : 2411-2418; De Almeida *et al.*, (1989) Mol. Gen. Genetics 218 : 78-86), 따라서 다수의 사건은 원하는 발현 수준 및 패턴을 나타내는 라인을 수득하기 위해 선별되어야 한다. 이러한 선별은 다른 것들 중에서, 직접 서열화, DNA의 서던 분석, mRNA 발현의 노던 분석, 단백질 발현의 면역 블로팅 분석 또는 표현형 분석에 의해 실행될 수 있다. 벡터는 자율적으로 복제하거나 숙주 세포의 염색체에 통합될 수 있는 플라스미드, 바이러스, 박테리오파지, 프로-바이러스, 파지미드, 트랜스포존, 인공 염색체 등일 수 있다. 벡터는 또한 자율적으로 복제하지 않는 네이키드 RNA 폴리뉴클레오타이드, 네이키드 DNA 폴리뉴클레오타이드, 동일한 가닥 내의 DNA 및 RNA 모두로 구성된 폴리뉴클레오타이드, 폴리-라이신-퀸유게이드된 DNA 또는 RNA, 펩타이드-퀸유게이드된 DNA 또는 RNA, 리포솜-퀸유게이드된 DNA 등일 수 있다. 본 발명에 사용된 용어 "발현"은 기능적 최종 산물, 예를 들어 mRNA 또는 단백질(전구체 또는 성숙)의 생산을 의미한다.

[0045] "작동 가능하게 연결된" 또는 "기능적으로 연결된"은 본 발명에 따른 임의의 기능적 페이로드(예를 들어, 프로모터, 종결자, 테그온, 용해도 태그 등)의 추가 올리고-또는 폴리뉴클레오타이드와의 순차적 배열을 의미할 수 있다. 일부 경우에, 순차적 배열은 상기 추가 폴리뉴클레오타이드의 전사를 초래할 수 있다. 일부 경우에, 순차적 배열은 상기 추가 폴리뉴클레오타이드의 번역을 초래할 수 있다. 기능적 페이로드는 추가 올리고 또는 폴리뉴클레오타이드의 상류 또는 하류에 존재할 수 있다. 한 예에서, "작동 가능하게 연결된" 또는 "기능적으로 연결된"은 프로모터가 상기 프로모터에 인접하거나 하류 또는 3' 유전자의 전사를 제어함을 의미할 수 있다. 또 다른 예에서, "작동 가능하게 연결된" 또는 "기능적으로 연결된"은 종결자가 상기 종결자에 인접하거나 상류 또는 5' 유전자의 전사 종결을 제어함을 의미할 수 있다.

[0046] 본 발명에서 사용된 용어 "관심 생성물" 또는 "생체 분자"는 원료로부터 미생물에 의해 생성된 임의의 생성물을 의미할 수 있다. 일부 경우에, 관심 생성물은 작은 분자, 효소, 펩타이드, 아미노산, 유기산, 합성 화합물, 연료, 알코올 등일 수 있다. 예를 들어, 관심 생성물 또는 생체 분자는 임의의 1차 또는 2차 세포의 대사산물일 수 있다. 1차 대사산물은 특히 에탄올, 시트르산, 락트산, 글루탐산, 글루탐산염, 라이신, 트레오닌, 트립토판 및 다른 아미노산, 비타민, 폴리사카라이드 등일 수 있다. 2차 대사산물은 특히 페니실린과 같은 항생물질 또는 사이클로스포린 A와 같은 면역 억제제, 지베렐린과 같은 식물 호르몬, 로바스타틴과 같은 스타틴 약물, 그리세오폴빈과 같은 살균제 등일 수 있다. 관심 생성물 또는 생체 분자는 또한 카탈라아제, 아밀라아제, 펙티나아제, 글루코오스 아이소머라제, 셀룰라아제, 헤미셀룰라아제, 리파아제, 락타아제, 스트렙토키나아제 및 여러 다른 것을 포함하는 미생물 효소와 같은 미생물에 의해 생성된 임의의 세포내 성분일 수 있다. 세포내 성분은 또한

인슐린, B형 간염 백신, 인터페론, 과립구 콜로니-자극 인자, 스트렙토키나아제 및 기타와 같은 재조합 단백질을 포함할 수 있다.

- [0047] 본 발명에 사용된 용어 "HTP 유전 디자인 라이브러리" 또는 "라이브러리"는 본 발명에 따른 유전자 교란의 집합을 의미한다. 일부 실시태양에서, 본 발명의 라이브러리는 i) 데이터베이스 또는 다른 컴퓨터 파일에서 서열 정보의 집합, ii) 상기한 일련의 유전자 요소를 암호화하는 유전자 구조체의 집합, 또는 iii) 상기 유전자 요소를 포함하는 숙주 세포 균주로서 입증될 수 있다. 일부 실시태양에서, 본 발명의 라이브러리는 개별 요소의 집합 (예를 들어, PRO 스왑 라이브러리를 위한 프로모터 집합 또는 STOP 스왑 라이브러리를 위한 터미네이터 집합, SOLUBILITY TAG 스왑 라이브러리를 위한 단백질 용해도 태그 모음 또는 DEGRADATION TAG 스왑 라이브러리를 위한 단백질 분해 태그 모음)을 의미할 수 있다. 다른 실시태양에서, 본 발명의 라이브러리는 프로모터:유전자, 유전자:터미네이터, 또는 심지어 프로모터:유전자:터미네이터의 조합과 같은 유전자 요소의 조합을 의미할 수 있다. 일부 실시태양에서, 본 발명의 라이브러리는 또한 프로모터, 종결자, 단백질 용해도 태그 및/ 또는 단백질 분해 태그의 조합을 의미할 수 있다. 일부 실시태양에서, 본 발명의 라이브러리는 숙주 유기체에서 라이브러리의 각 구성원을 적용하는 효과와 관련된 메타 데이터를 추가로 포함한다. 예를 들어, 본 발명에서 사용된 라이브러리는 특정 종에서 하나 이상의 표현형에 대한 이들 조합의 얻어진 효과와 함께 프로모터:유전자 서열 조합의 집합을 포함할 수 있으며, 따라서 미래 프로모터 스왑에 상기 조합을 사용하는 미래 예측 가치를 개량시킨다.
- [0048] 본 발명에서 사용된 용어 "SNP"는 작은 핵 다형성(들)을 의미한다. 일부 실시태양에서, 본 발명의 SNP는 광범위하게 해석되어야 하며, 단일 뉴클레오타이드 다형성, 서열 삽입, 결실, 역전 및 다른 서열 치환을 포함한다. 본 발명에서 사용된 용어 "비 동의" 또는 비 동의 SNP "는 숙주 세포 단백질에서 암호화 변화를 유도하는 돌연변이를 의미한다
- [0049] 계능 공학의 "고 처리량(HTP)" 방법은 상기 방법의 적어도 하나의 단계를 수행하기 위한 자동화 장비(예를 들어, 액체 핸들러 또는 플레이트 핸들러 머신)의 적어도 하나의 부품의 이용을 필요로 할 수 있다.
- [0050] 본 발명에 사용된 용어 "폴리뉴클레오타이드"는 올리고뉴클레오타이드를 포함하고 임의의 길이의 핵산을 의미한다. 폴리뉴클레오타이드는 DNA 또는 RNA일 수 있다. 폴리뉴클레오타이드는 달리 명시되지 않는 한 단일 가닥(ss) 또는 이중 가닥(ds)일 수 있다. 폴리뉴클레오타이드는 합성, 예를 들어 DNA 합성기에서 합성되거나, 자연 발생, 예를 들어 천연 공급원에서 추출되거나, 복제 또는 증폭된 물질에서 파생될 수 있다. 본 발명에서 언급된 폴리뉴클레오타이드는 변형된 염기 또는 뉴클레오타이드를 포함할 수 있다.
- [0051] 본 발명에 사용된 용어 "폴"은 2개 이상의 폴리뉴클레오타이드의 집합을 의미할 수 있다. 일부 실시태양에서, 폴리뉴클레오타이드 세트는 적어도 5개, 적어도 10개, 적어도 12개 또는 적어도 15개 이상의 폴리뉴클레오타이드를 포함할 수 있다.
- [0052] 본 발명에 사용된 용어 "오버랩 서열", 또는 "오버랩 어셈블리 서열" 또는 "어셈블리 오버랩 서열"은 2개의 폴리뉴클레오타이드에서 상보적이고 오버랩 서열이 한 폴리뉴클레오타이드 상의 ss인 경우 다른 폴리뉴클레오타이드 상의 또 다른 중첩 상보적 ss 영역에 혼성화될 수 있다. 오버랩 서열은 2개의 별개의 폴리뉴클레오타이드의 말단에 있거나 근접할 수 있다(예를 들어, 약 5, 10, 20개 뉴클레오타이드 내). 예를 들어, 2개의 별개의 폴리뉴클레오타이드가 단일 가닥인 경우, 어셈블리 오버랩 서열은 각 단일 가닥 폴리뉴클레오타이드의 3' 말단 상에 존재할 것이다. 대안적으로, 2개의 별개의 폴리뉴클레오타이드가 이중 가닥인 경우, 폴리뉴클레오타이드 중 하나의 어셈블리 오버랩 서열이 상기 폴리뉴클레오타이드의 3' 말단(즉, ds 폴리뉴클레오타이드의 상단 가닥을 기준으로 3' 말단)에 존재할 수 있는 반면, 다른 폴리뉴클레오타이드 상의 상보적 어셈블리 오버랩 서열은 상기 폴리뉴클레오타이드의 5' 말단(즉, ds 폴리뉴클레오타이드의 상부 가닥을 기준으로 5' 말단)에 존재할 수 있다. 필요에 따라, 임의의 ds 상의 어셈블리 오버랩 서열은 임의의 비 중첩 서열을 제거하여 이용 가능하게 만들 수 있다. 제거는 중합 효소의 3'-5' 엑소뉴클레아제 활성을 사용하는 것과 같이 효소적일 수 있다.
- [0053] 본 발명에 사용된 용어 "어셈블링"은 2개 이상, 4개 이상, 6개 이상, 8개 이상, 10개 이상, 12개 이상, 또는 15개 이상의 폴리뉴클레오타이드, 예를 들어, 4개 이상의 폴리뉴클레오타이드가 다른 폴리뉴클레오타이드에 결합되어 더 긴 폴리뉴클레오타이드를 만든다.
- [0054] 본 발명에 사용된 용어 "적절한 반응 조건 하에서 배양"은 원하는 결과, 즉 폴리뉴클레오타이드 어셈블리를 달성하기 위해 반응을 적합한 온도 및 시간으로 유지하는 것을 의미할 수 있다. 본 방법에서 사용되는 효소 및 시약에 적합한 반응 조건은 공지되어 있으며(예를 들어, 본 발명의 실시예에 기술된 바와 같음), 따라서 본 방법에 적합한 반응 조건을 쉽게 결정할 수 있다. 이러한 반응 조건은 사용되는 효소에 따라 (예를 들어, 최적 온도

등에 따라) 달라질 수 있다.

- [0055] 본 발명에 사용된 용어 "결합"은 두 서열 사이의 공유 결합의 생성을 의미할 수 있다.
- [0056] 본 발명에 사용된 용어 "조성물"은 나열된 것들에 더하여 다른 시약, 예를 들어, 글리세롤, 염, dNTP 등을 함유할 수 있는 시약의 조합을 의미할 수 있다. 조성물은 임의의 형태, 예를 들어 수성 또는 동결 건조일 수 있고, 임의의 상태(예를 들어, 동결 또는 액체 형태)일 수 있다.
- [0057] 본 발명에 사용된 "벡터"는 이 속에 단편 또는 DNA 어셈블리가 통합되어 조작된 벡터가 숙주 세포에서 복제될 수 있는 적합한 DNA이다. 선형화된 벡터는 원형 벡터의 제한 엔도뉴클레아제 분해 또는 PCR에 의해 생성될 수 있다. 단편 및/또는 선형화된 벡터의 농도는 겔 전기 영동 또는 기타 수단에 의해 결정될 수 있다.
- [0058] **개요**
- [0059] 조합 방식이 아니라 결정적 방식으로 단일 반응으로 다수의 어셈블리가 생성되도록 하는 방법 및 조성물이 본 발명에 제공된다. 본 발명에 제공된 방법 및 조성물은 모든 출력 어셈블리가 사전에 결정되는 라이브러리의 생성을 가능하게 하는 동시에 다중 어셈블리의 시간, 비용 및 처리량 이점을 제공한다. 본 발명에 제공된 방법 및 조성물은 단일 조립 반응에서 많은 플라스미드 또는 구조체의 생성을 가능하게 하여 수천 개의 플라스미드 또는 구조체의 라이브러리를 생성하는 데 필요한 총 반응의 수를 줄인다. 본 발명에 제공된 방법 및 조성물은 또한 다수의 가능한 조합의 더 큰 세트로부터 원하는 플라스미드 또는 구조체의 정의된 서브 세트를 어셈블링하는 것을 허용한다. 일부 경우에, 본 발명에 제공된 방법 및 조성물은 임의의 페이로드 또는 삽입 서열 특이적 어셈블리 오버랩을 포함하지 않음으로써 게놈으로부터 증폭(또는 합성)되어야 하는 고유한 부분('상동성 암')의 수를 최소화한다. 따라서 다중 페이로드 또는 삽입체 서열과 결합되도록 디자인된 동일한 상동성 암 쌍의 여러 복사본을 증폭할 필요가 없다. 또한, 페이로드/삽입체 서열 및 상동성 암 쌍의 조합에서 발생하는 다양성은 페이로드/삽입체 서열 자체 상의 서열에 의해 지정된다. 결과적으로 생성되는 다중 페이로드 서열은 종합적이고 저렴하게 생성될 수 있다. 본 발명에 제공된 방법 및 조성물을 사용하여 생성된 라이브러리는, 예를 들어, 임의의 게놈 편집 방법 또는 임의의 풀링된 경로 어셈블리와 같은 임의의 수의 응용분야에 적합할 수 있다. 당업계에 공지된 게놈 편집 방법은 예를 들어, 흉터없는 게놈 편집과 같은 다수의 임의 위치에서 세포의 게놈을 편집하기 위해 RCME(재조합 카세트 매개 교환)를 위한 맞춤형 부위를 필요로 하지 않는 방법일 수 있다.
- [0060] 핵산 구조체 라이브러리의 결정적 방식의 어셈블리를 위한 폴리뉴클레오타이드의 혼합물을 포함하는 조성물이 본 발명에 제공된다. 혼합물은 폴리뉴클레오타이드 부분(예를 들어, 제 1 및 제 2 폴리뉴클레오타이드)의 n개 풀을 포함할 수 있다. n개 풀은 최대, 적어도 또는 정확히 폴리뉴클레오타이드의 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 또는 20개 풀일 수 있다. n개 풀은 각각 동일한 수의 폴리뉴클레오타이드 부분을 포함할 수 있거나, 상이한 수의 폴리뉴클레오타이드 부분(예를 들어, 제 1 및 제 2 폴리뉴클레오타이드)을 포함할 수 있다. 한 실시태양에서, 혼합물은 2개 풀을 포함하여 2개 풀 중 하나가 제 1 폴리뉴클레오타이드를 포함하고 2개 풀 중 다른 하나가 제 2 폴리뉴클레오타이드를 포함한다. 제 1 폴리뉴클레오타이드의 각각의 풀은 별도의 제 2 폴리뉴클레오타이드 풀에 쌍을 이룬 제 2 폴리뉴클레오타이드를 포함할 수 있다. 임의의 상기 실시태양에 추가로, 혼합물은 삽입체 또는 가교 폴리뉴클레오타이드의 n-1 풀을 추가로 포함할 수 있다. 각각의 삽입체 또는 가교 폴리뉴클레오타이드는 5' 말단에서 폴리뉴클레오타이드 부분(예를 들어, 제 1 폴리뉴클레오타이드)의 n개 풀 중 하나의 요소 및 3' 말단에서 폴리뉴클레오타이드 부분(예를 들어, 제 2 폴리뉴클레오타이드)의 다른 풀 중 하나의 요소에 상보적인 서열을 포함할 수 있다. 삽입체 서열은 어셈블리가 폴리뉴클레오타이드의 라이브러리를 생성하도록 디자인될 수 있으며, 여기서 각 폴리뉴클레오타이드는 삽입체 폴리뉴클레오타이드의 n-1 풀의 각각으로부터의 특정 요소가 산재된 폴리뉴클레오타이드 부분의 n개 풀의 각각의 특정 요소를 포함한다.
- [0061] 폴리뉴클레오타이드의 혼합물은 다음을 포함할 수 있다: 폴리뉴클레오타이드의 쌍을 포함하는 제 1 풀을, 여기서, 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 함유한다; 및 삽입체 폴리뉴클레오타이드의 제 2 풀, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열을 포함하고 제 1 풀로부터의 한 쌍의 폴리뉴클레오타이드에서 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 이의 반대 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다. 일부 경우에, 조성물은 클로닝 벡터를 추가로 포함하고, 여기서 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단은 클로닝 벡터에 상보적인 서열을 포함한다. 클로닝 벡터는 예를 들어 *E. coli* 또는 *S. cerevisiae*와 같은 숙주 세포에서의 증식에 적합한 당업계에 공지된 임의의 클로닝 벡터일 수 있다. 다른 실시태양에서, 조성물은 또한 증합 효소, 엑소뉴클레아제, 리가아제 또는

이들의 임의의 조합을 포함한다. 중합 효소는 가닥 치환 또는 비 가닥 치환일 수 있다. 엑소뉴클레아제는 5'-3' 엑소뉴클레아제일 수 있다. 제 1 폴의 폴리뉴클레오타이드 쌍은 이중 가닥, 단일 가닥 또는 이들의 조합일 수 있다. 제 2 폴의 삽입체 폴리뉴클레오타이드는 이중 가닥, 단일 가닥 또는 이들의 조합일 수 있다. 한 실시태양에서, 중합 효소는 비 가닥 치환이고 조성물은 크라우딩제를 추가로 포함한다. 크라우딩제는 폴리에틸렌 글리콜 (PEG), 피콜 또는 텍스트란에서 선택될 수 있다. 한 실시태양에서, 크라우딩제는 폴리에틸렌 글리콜(PEG)이다. PEG는 약 3 내지 약 7%(중량/부피)의 농도로 사용될 수 있다. PEG는 PEG-200, PEG-4000, PEG-6000, PEG-8000 또는 PEG-20,000에서 선택될 수 있다. 다른 실시태양에서, 중합 효소는 가닥 치환이고 조성물은 단일 가닥 결합 단백질(ET SSB), 대장균 recA, T7 유전자 2.5 산물, 파지 램다 RedB 또는 Rac 프로파지 RecT일 수 있다.

[0062] 한 실시태양에서, 본 발명에 제공된 조성물은 다음 폴리뉴클레오타이드: (1) 하나 이상의 제 1 폴리뉴클레오타이드, (2) 하나 이상의 삽입체 폴리뉴클레오타이드, 여기서 삽입체 폴리뉴클레오타이드는 5' 말단에 제 1 어셈블리 오버랩 서열 및 이의 대향 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다, 및 (3) 하나 이상의 제 2 폴리뉴클레오타이드의 혼합물이다. 다른 실시태양에서, 조성물은 다음 폴리뉴클레오타이드: (1) 하나 이상의 제 1 폴리뉴클레오타이드, (2) 하나 이상의 삽입체 폴리뉴클레오타이드, 여기서 삽입체 폴리뉴클레오타이드는 5' 말단에 제 1 어셈블리 오버랩 서열 및 이의 대향 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다, (3) 하나 이상의 제 2 폴리뉴클레오타이드의 혼합물 및 (4) 클로닝 벡터의 혼합물이다. 하나 이상의 제 1 폴리뉴클레오타이드의 각각은 하나 이상의 삽입체 폴리뉴클레오타이드로부터 삽입체 폴리뉴클레오타이드의 5' 또는 근위 말단에 존재하는 제 1 어셈블리 오버랩 서열에 상보적인 3' 또는 원위 말단에 서열을 포함할 수 있다. 하나 이상의 제 2 폴리뉴클레오타이드의 각각은 하나 이상의 삽입체 폴리뉴클레오타이드로부터 삽입체 폴리뉴클레오타이드의 3' 또는 원위 말단에 존재하는 제 2 어셈블리 오버랩 서열에 상보적인 5' 또는 근위 말단에 서열을 포함할 수 있다. 하나 이상의 제 2 폴리뉴클레오타이드 각각은 하나 이상의 제 2 폴리뉴클레오타이드 중 하나 이상과 쌍을 이루어 하나 이상의 제 1 및 제 2 폴리뉴클레오타이드 쌍을 형성할 수 있다. 제 1 및 제 2 폴리뉴클레오타이드의 각각의 쌍은 하나 이상의 삽입체 폴리뉴클레오타이드로부터의 삽입체 폴리뉴클레오타이드의 근위 말단 상의 제 1 어셈블리 오버랩 서열에 상보적인 제 1 폴리뉴클레오타이드의 원위 말단에서의 서열뿐만 아니라 하나 이상의 삽입체 폴리뉴클레오타이드로부터의 삽입체 폴리뉴클레오타이드의 원위 말단에 상보적인 제 2 폴리뉴클레오타이드의 근위 말단에서의 서열을 포함할 수 있다.

[0063] 폴리뉴클레오타이드의 라이브러리를 생성하는 방법이 본 발명에 제공되며, 상기 방법은 a.) 폴리뉴클레오타이드 부분(예를 들어, 제 1 및 제 2 폴리뉴클레오타이드)의 n개 폴 및 삽입체 또는 가교 폴리뉴클레오타이드의 n-1개 폴을 결합하는 단계; 및 b.) 폴리뉴클레오타이드 부분의 n개 폴 및 삽입체 또는 가교 폴리뉴클레오타이드의 n-1개 폴을 폴리뉴클레오타이드의 라이브러리로 어셈블리하는 단계를 포함하며, 여기서 라이브러리의 각 폴리뉴클레오타이드는 폴리뉴클레오타이드 부분과 가교 폴리뉴클레오타이드의 n개 폴의 각각으로부터의 개별 요소의 정의된 조합을 포함한다. 삽입체 또는 가교 폴리뉴클레오타이드의 n-1개 폴의 각 삽입체 또는 가교 폴리뉴클레오타이드는 제 1 및 제 2 폴리뉴클레오타이드의 n개 폴의 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 대향 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다. 어셈블리는 시험관 내 또는 생체내 오버랩 어셈블리 방법을 통해 수행될 수 있다. 일부 경우에, 어셈블리는 시험관 내 클로닝 방법을 통해 수행되며, 여기서 폴리뉴클레오타이드 부분의 n개 폴 및 삽입체 또는 가교 폴리뉴클레오타이드의 n-1개 폴의 혼합물은 가열되어 존재하는 임의의 이중 가닥 폴리뉴클레오타이드 부분을 부분적으로 또는 완전히 변성시키고, 시험관 내 클로닝 방법을 적용하기 전에 느린 속도로 실온으로 냉각된다.

[0064] 폴리뉴클레오타이드의 라이브러리를 생성하는 방법이 본 발명에 제공되며, 이 방법은 (a) 폴리뉴클레오타이드의 제 1 폴 및 폴리뉴클레오타이드의 제 2 폴을 조합하는 단계, 여기서 제 1 폴은 폴리뉴클레오타이드의 쌍을 포함하며, 여기서 제 1 폴의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 제 2 폴은 삽입체 폴리뉴클레오타이드를 포함하며, 여기서 제 2 폴의 각 삽입체 폴리뉴클레오타이드는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열을 포함하고 제 1 폴로부터의 한 쌍의 폴리뉴클레오타이드에서 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 이의 반대 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다; (b) 제 1 폴 및 제 2 폴을 폴리뉴클레오타이드의 라이브러리로 어셈블리하는 단계를 포함하며, 여기서 라이브러리의 각 폴리뉴클레오타이드는 제 2 폴의 삽입체 폴리뉴클레오타이드 및 제 1 폴의 한 쌍의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함한다. 어셈블리는 시험관 내 또는 생체 내 오

버랩 어셈블리 방법을 통해 수행될 수 있다. 일부 경우에, 어셈블리는 시험관 내 클로닝 방법을 통해 수행되며, 여기서 제 1 풀과 제 2 풀의 혼합물은 가열되어 존재하는 임의의 이중 가닥 폴리뉴클레오타이드 부분을 부분적으로 또는 완전히 변성시키고, 시험관 내 클로닝 방법을 적용하기 전에 느린 속도로 실온으로 냉각된다. 일부 경우에, 상기 방법은 단계(a) 동안 클로닝 벡터를 제 1 풀 및 제 2 풀과 조합하는 단계를 추가로 포함하며, 여기서 클로닝 벡터의 대향 말단은 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 상보적인 서열을 포함한다. 일부 경우에, 상기 방법은 단계(a) 전에 클로닝 벡터를 제 1 풀과 조합하는 단계를 추가로 포함하며, 여기서 클로닝 벡터의 대향 말단은 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 상보적인 서열을 포함한다. 일부 경우에, 클로닝 벡터 및 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단은 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함한다. 일부 경우에, 상기 방법은 하나 이상의 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제를 첨가함으로써 클로닝 벡터의 대향 말단과 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 사이에 단일 가닥 상보적 오버행을 생성하는 단계를 추가로 포함한다. 일부 경우에, 상기 방법은 클로닝 벡터의 대향 말단과 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 사이에 단일 가닥 상보적 오버행을 결찰하는 단계를 추가로 포함한다. 결찰은 DNA 리가아제를 사용하여 수행될 수 있다. 일부 경우에, 단계(b)는 제 2 풀의 삽입체 폴리뉴클레오타이드, 제 1 풀의 쌍의 제 1 및 제 2 폴리뉴클레오타이드 및 클로닝 벡터를 포함하는 원형 산물을 생성한다.

[0065]

한 양태에서, 폴리뉴클레오타이드의 라이브러리를 생성하는 방법이 본 발명에 제공되며, 이 방법은 (a) 중합효소 연쇄 반응(PCR)을 통해 폴리뉴클레오타이드의 제 1 풀을 증폭하는 단계, 여기서 제 1 풀은 폴리뉴클레오타이드의 쌍을 포함하며, 여기서 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 한 쌍의 각 제 1 폴리뉴클레오타이드 및 각 제 2 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 포함하고, 여기서 증폭은 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 공통 오버랩 서열을 제 1 풀로부터의 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 도입한다; (b) 공통 오버랩 서열을 이용하여 제 1 풀로부터의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드의 각 쌍을 단일 핵산 단편으로 어셈블리하는 단계, 여기서 각 쌍에 대한 단일 핵산 단편은 공통 오버랩 서열에 의해 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단으로부터 분리된 제 1 폴리뉴클레오타이드와 제 2 폴리뉴클레오타이드를 포함하며, 각 쌍에 대한 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단은 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 대해 원위에 있는 단일 핵산 단편의 대향 말단 상에 위치된다; (c) 각 쌍에 대한 단일 핵산 단편을 삽입체 폴리뉴클레오타이드를 포함하는 제 2 풀과 조합하는 단계, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 단일 핵산 단편 내에 존재하는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열 및 단일 핵산 단편 내에 존재하는 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 대향 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다; (d) 제 1 풀 및 제 2 풀을 원형화된 산물의 제 3 풀로 어셈블리하는 단계, 여기서 어셈블리는 시험관 내 또는 생체 내 오버랩 어셈블리 방법을 통해 수행되고, 여기서 제 3 풀의 각 원형화된 산물은 제 2 풀로부터의 삽입체 서열 및 제 1 풀로부터의 한 쌍의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함한다; (e) 제 3 풀의 원형화된 산물의 각각의 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치한 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)을 인식하는 하나 이상의 부위 특이적 뉴클레아제(들)에 의한 소화를 통해 제 3 풀의 각 원형화된 산물을 선형화하는 단계; 및 (f) 시험관 내 또는 생체 내 클로닝 방법에 의해 선형화된 산물을 클로닝 벡터로 어셈블리하는 단계를 포함한다. 일부 경우에, 공통 오버랩 서열은 적어도 1개의 뉴클레오타이드의 어셈블리 오버랩 서열을 포함하고 단계(b)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법에 의해 수행된다. 일부 경우에, 공통 오버랩 서열은 10-25개 뉴클레오타이드의 어셈블리 오버랩 서열을 포함하고 단계(b)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법에 의해 수행된다. 일부 경우에, 오버랩 기반 DNA 어셈블리 방법은 SOE-PCR 또는 시험관 내 오버랩 어셈블리 방법(예를 들어, NEB® HiFi 빌더를 사용하는 HiFi 어셈블리)에서 선택된다. 일부 경우에, 제 1 폴리뉴클레오타이드의 5' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)은 각 쌍의 제 2 폴리뉴클레오타이드의 3' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 상보적이며, 여기서 단계(b)에서 각 쌍의 제 1 및 제 2 폴리뉴클레오타이드의 공통 오버랩 서열을 이용하는 것은 SOE-PCR을 수행하는 것을 수반한다. 일부 경우에, 단계(b)에서 각 쌍의 제 1 및 제 2 폴리뉴클레오타이드의 공통 오버랩 서열을 이용하는 것은 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 하나 이상의 부위 특

이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하여 상보적 서열을 포함하는 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행을 생성하고; 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행 상에 존재하는 상보적 서열을 결합하는 것을 수반한다. 단계(d)의 어셈블리는 시험관 내 또는 생체 내 오버랩 어셈블리 방법을 사용하여 수행된다. 단계(d)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법을 사용하여 수행된다. 일부 경우에, 오버랩 기반 DNA 어셈블리는 SOE-PCR 및 시험관 내 오버랩 어셈블리 방법(예를 들어, NEB® HiFi 빌더를 사용하는 HiFi 어셈블리)에서 선택된다. 일부 경우에, 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단은 하나 이상의 부위 특이적 뉴클레아제 인식 서열의 추가 세트를 포함하고 제 2 폴의 각 삽입체 폴리뉴클레오타이드의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 포함한다. 일부 경우에, 단계(d)에서의 어셈블리는 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단 상에 존재하는 추가의 하나 이상의 부위 특이적 뉴클레아제 인식 서열 및 제 2 폴리뉴클레오타이드의 제 1 및 제 2 어셈블리 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단 상에 존재하는 추가의 하나 이상의 부위 특이적 뉴클레아제 인식 서열 및 제 2 폴리뉴클레오타이드의 제 1 및 제 2 어셈블리 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하여 제 2 폴리뉴클레오타이드의 삽입체 폴리뉴클레오타이드의 제 1 어셈블리 서열의 5' 말단 상의 단일 가닥 오버행 상에 존재하는 서열에 상보적인 서열을 포함하는 제 1 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행 및 제 2 폴리뉴클레오타이드의 동일한 삽입체 폴리뉴클레오타이드의 제 2 어셈블리 서열의 3' 말단 상의 단일 가닥 오버행 상에 존재하는 서열에 상보적인 서열을 포함하는 제 2 폴리뉴클레오타이드의 5' 말단 상의 단일 가닥 오버행을 생성하고; 단일 가닥 오버행에 존재하는 상보적 서열을 결합하는 것을 수반한다. 일부 경우에, 단계(d)의 어셈블리는 시험관 내 클로닝 방법을 통해 수행되며, 여기서 제 1 폴과 제 2 폴의 혼합물은 가열되어 존재하는 임의의 이중 가닥 폴리뉴클레오타이드 부분을 부분적으로 또는 완전히 변성시키고, 시험관 내 클로닝 방법을 적용하기 전에 느린 속도로 실온으로 냉각된다. 단계(f)의 어셈블리는 시험관 내 클로닝 방법 또는 생체 내 클로닝 방법을 통해 수행된다. 단계(f)의 클로닝 벡터는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 포함한다. 일부 경우에, 단계(f)에서의 어셈블리는 클로닝 벡터의 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 클로닝 벡터에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하는 것, 여기서 분해는 클로닝 벡터의 대향 말단 상의 단일 가닥 오버행을 생성하고, 클로닝 벡터의 대향 말단 중 하나 상의 단일 가닥 오버행은 단계(e)에서 생성된 선형화된 산물의 말단에 상보적인 서열을 포함하며 클로닝 벡터의 대향 말단 중 다른 하나 상의 단일 가닥 오버행은 단계(e)에서 생성된 선형화된 산물의 대향 말단에 상보적인 서열을 포함한다; 및 클로닝 벡터의 단일 가닥 오버행 상에 존재하는 상보적 서열과 단계(e)로부터의 선형화된 산물을 결합하는 것을 수반한다. 본 발명에 제공된 임의의 방법 또는 조성물에 사용하기 위한 부위 특이적 뉴클레아제는 제한 엔도뉴클레아제(들), 유형 II 엔도뉴클레아제(들), 귀소 엔도뉴클레아제(들), RNA 유도 뉴클레아제(들), DNA 유도 뉴클레아제(들), 징크 핑거 뉴클레아제(들), TALEN(들) 또는 니킹 효소(들) 또는 이의 임의의 조합으로부터 선택될 수 있다. 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치한 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)은 하나 이상의 귀소 뉴클레아제 인식 서열(들)일 수 있다. 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 대한 하나 이상의 부위 특이적 뉴클레아제(들)은 귀소 엔도뉴클레아제일 수 있다.

[0066]

다른 양태에서, 폴리뉴클레오타이드의 라이브러리를 생성하는 방법이 본 발명에 제공되며, 이 방법은 (a) 중합 효소 연쇄 반응(PCR)을 통해 폴리뉴클레오타이드의 제 1 폴을 증폭하는 단계, 여기서 제 1 폴은 폴리뉴클레오타이드의 쌍을 포함하며, 여기서 제 1 폴의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 한 쌍의 각 제 1 폴리뉴클레오타이드 및 각 제 2 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 포함하고, 여기서 증폭은 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 공통 오버랩 서열을 제 1 폴리뉴클레오타이드의 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 도입하며, 여기서 제 1 폴리뉴클레오타이드의 제 1 말단 5' 말단상의 하나 이상의 부위 특이적 뉴클레아제(들)에 대한 하나 이상의 인식 서열은 쌍으로 제 2 폴리뉴클레오타이드의 제 1 말단 3' 말단 상의 하나 이상의 부위 특이적 뉴클레아제(들)에 대한 하나 이상의 인식 서열에 상보적이다; (b) 각 쌍 내의 제 1 폴리뉴클레오타이드의 제 1 말단 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 하나 이상의 상보적 부위 특이적 뉴클레아제 인식 서열(들)을 사용하여 스플라이싱 및 오버랩 연장 중합 효소 연쇄 반응(SOE-PCR)을 수행함으로써 제 1

폴로부터의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드의 각 쌍을 단일 핵산 단편으로 어셈블리하는 단계, 여기서 각 쌍에 대한 단일 핵산 단편은 공통 오버랩 서열에 의해 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단으로부터 분리된 제 1 폴리뉴클레오타이드와 제 2 폴리뉴클레오타이드를 포함하며, 각 쌍에 대한 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단은 하나 이상의 부위 특이적 뉴클레아제 인식 서열 (들)에 대해 원위에 있는 단일 핵산 단편의 대향 말단 상에 위치된다; (c) 각 쌍에 대한 단일 핵산 단편을 삽입체 폴리뉴클레오타이드를 포함하는 제 2 풀과 조합하는 단계, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 단일 핵산 단편 내에 존재하는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열 및 단일 핵산 단편 내에 존재하는 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 대향 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다; (d) 제 1 풀 및 제 2 풀을 원형화된 산물의 제 3 풀로 어셈블리하는 단계, 여기서 어셈블리는 시험관 내 또는 생체 내 오버랩 어셈블리 방법을 통해 수행되고, 여기서 제 3 풀의 각 원형화된 산물은 제 2 풀로부터의 삽입체 서열 및 제 1 풀로부터의 한 쌍의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함한다; (e) 제 3 풀의 원형화된 산물의 각각의 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치한 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)을 인식하는 하나 이상의 부위 특이적 뉴클레아제(들)에 의한 소화를 통해 제 3 풀의 각 원형화된 산물을 선형화하는 단계; 및 (f) 시험관 내 또는 생체 내 클로닝 방법에 의해 선형화된 산물을 클로닝 벡터로 어셈블리하는 단계를 포함한다. 단계(d)의 어셈블리는 시험관 내 또는 생체 내 오버랩 어셈블리 방법을 통해 수행될 수 있다. 일부 경우에, 오버랩 기반 DNA 어셈블리는 SOE-PCR 및 시험관 내 오버랩 어셈블리 방법(예를 들어, NEB® HiFi 빌더를 사용하는 HiFi 어셈블리)에서 선택된다. 일부 경우에, 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단은 하나 이상의 부위 특이적 뉴클레아제 인식 서열의 추가 세트를 포함하고 제 2 풀의 각 삽입체 폴리뉴클레오타이드의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 포함한다. 일부 경우에, 단계(d)에서의 어셈블리는 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단 상에 존재하는 추가의 하나 이상의 부위 특이적 뉴클레아제 인식 서열 및 제 2 풀로부터의 각 삽입체 폴리뉴클레오타이드의 제 1 및 제 2 어셈블리 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단 상에 존재하는 추가의 하나 이상의 부위 특이적 뉴클레아제 인식 서열 및 제 2 풀로부터의 각 삽입체 폴리뉴클레오타이드의 제 1 및 제 2 어셈블리 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하여 제 2 풀로부터의 삽입체 폴리뉴클레오타이드의 제 1 어셈블리 서열의 5' 말단 상의 단일 가닥 오버행 상에 존재하는 서열에 상보적인 서열을 포함하는 제 1 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행 및 제 2 풀로부터의 동일한 삽입체 폴리뉴클레오타이드의 제 2 어셈블리 서열의 3' 말단 상의 단일 가닥 오버행 상에 존재하는 서열에 상보적인 서열을 포함하는 제 2 폴리뉴클레오타이드의 5' 말단 상의 단일 가닥 오버행을 생성하고; 단일 가닥 오버행에 존재하는 상보적 서열을 결합하는 것을 수반한다. 일부 경우에, 단계(d)의 어셈블리는 시험관 내 클로닝 방법을 통해 수행되며, 여기서 제 1 풀과 제 2 풀의 혼합물은 가열되어 존재하는 임의의 이중 가닥 폴리뉴클레오타이드 부분을 부분적으로 또는 완전히 변성시키고, 시험관 내 클로닝 방법을 적용하기 전에 느린 속도로 실온으로 냉각된다. 단계(f)의 어셈블리는 시험관 내 클로닝 방법 또는 생체 내 클로닝 방법을 통해 수행된다. 단계(f)의 클로닝 벡터는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 포함한다. 일부 경우에, 단계(f)에서의 어셈블리는 클로닝 벡터의 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 클로닝 벡터에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하는 것, 여기서 분해는 클로닝 벡터의 대향 말단 상의 단일 가닥 오버행을 생성하고, 클로닝 벡터의 대향 말단 중 하나 상의 단일 가닥 오버행은 단계(e)에서 생성된 선형화된 산물의 말단에 상보적인 서열을 포함하며 클로닝 벡터의 대향 말단 중 다른 하나 상의 단일 가닥 오버행은 단계(e)에서 생성된 선형화된 산물의 대향 말단에 상보적인 서열을 포함한다; 및 클로닝 벡터의 단일 가닥 오버행 상에 존재하는 상보적 서열과 단계(e)로부터의 선형화된 산물을 결합하는 것을 수반한다. 본 발명에 제공된 임의의 방법 또는 조성물에 사용하기 위한 부위 특이적 뉴클레아제는 제한 엔도뉴클레아제(들), 유형 II 엔도뉴클레아제(들), 귀소 엔도뉴클레아제(들), RNA 유도 뉴클레아제(들), DNA 유도 뉴클레아제(들), 징크 핑거 뉴클레아제(들), TALEN(들) 또는 니킹 효소(들) 또는 이의 임의의 조합으로부터 선택될 수 있다.

[0067]

한 실시태양에서, 본 발명에 제공된 방법 및 조성물의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 표적 게놈 유전자좌에 상보적이거나 상응하는 서열을 포함한다. 제 1 및 제 2 폴리뉴클레오타이드에 존재하는 표적 게놈 유전자좌에 상보적이거나 상응하는 서열은 삽입체 폴리뉴클레오타이드 상에 존재하는 어셈블리 오버랩 서열에 상보적인 서열을 포함하는 상기 제 1 및 제 2 폴리뉴클레오타이드의 말단에 대향하는 상

기 제 1 및 제 2 폴리뉴클레오타이드의 말단 상에 위치될 수 있다. 숙주 세포의 표적 게놈 유전자좌에 상보적이거나 상응하는 서열을 포함할 때, 제 1 및 제 2 폴리뉴클레오타이드는 상동성 암으로 지칭될 수 있다. 특히, 각각의 제 1 폴리뉴클레오타이드는 좌측 상동성 암으로 지칭될 수 있는 반면, 각각의 제 2 폴리뉴클레오타이드는 우측 상동성 암으로 지칭될 수 있다. 숙주 세포의 표적 게놈 유전자좌에 상보적이거나 상응하는 서열을 포함할 때, 본 발명에 제공된 조성물 및 방법을 사용하여 제 1 및 제 2 폴리뉴클레오타이드 및 삽입체 폴리뉴클레오타이드 쌍의 어셈블리를 통한 핵산 구조체의 라이브러리 생성은 후속적으로 숙주 세포의 게놈을 수정하기 위해 게놈 편집 기술에서 사용될 수 있다. 숙주 세포는 원핵 세포 또는 진핵 숙주 세포일 수 있다.

[0068] 폴리뉴클레오타이드 쌍

[0069] 본 발명에 기술된 바와 같이, 본 발명에 제공된 조성물 및 방법은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하거나 이용할 수 있어서 각각의 제 1 폴리뉴클레오타이드가 제 2 폴리뉴클레오타이드와 쌍을 이룬다. 제 1 및 제 2 폴리뉴클레오타이드는 핵산 합성을 위해 당업계에 공지된 임의의 방법을 사용하여 화학적으로 합성될 수 있다(예를 들어, 어레이 합성 또는 컬럼 합성될 수 있다). 제 1 및 제 2 폴리뉴클레오타이드는, 예를 들어, 게놈 DNA와 같은 기존 DNA로부터 연장 반응(예를 들어, PCR)을 통해 증폭될 수 있다.

[0070] 제 1 및 제 2 폴리뉴클레오타이드의 각각은 기능적 및 비 기능적 서열 또는 이들의 조합을 포함할 수 있다. 기능적 서열은 유전자 또는 이의 일부 또는 도메인, 또는 조절 요소 또는 이의 일부를 나타내는 서열을 의미할 수 있다. 본 발명에 추가로 기술된 바와 같이, 유전자 또는 이의 일부는 대사 또는 생화학적 경로의 일부인 단백질을 암호화할 수 있다. 또한 본 발명에 추가로 기술된 바와 같이, 조절 요소는 프로모터, 종결자, 용해도 태그, 분해 태그 또는 태그론일 수 있다. 비 기능적 서열은 유전자 또는 이의 일부 또는 조절 요소 또는 이의 일부를 나타내지 않는 서열을 의미할 수 있다. 비 기능적 서열은 본 발명에 제공된 바와 같은 삽입체 폴리뉴클레오타이드와의 상기 제 1 및 제 2 폴리뉴클레오타이드의 어셈블리를 돕거나 이를 위해 이용되는 서열일 수 있다. 한 실시태양에서, 각각의 제 1 및 제 2 폴리 뉴클레오타이드는 기능적 및 비 기능적 서열의 혼합물을 포함한다. 다른 실시태양에서, 각각의 제 1 및 제 2 폴리뉴클레오타이드는 기능적 또는 비 기능적 서열 중 하나 또는 다른 하나를 포함한다. 제 1 및 제 2 폴리뉴클레오타이드가 기능적 서열만을 포함하는 실시태양에서, 기능적 서열 또는 기능적 서열의 일부는 본 발명에 제공된 바와 같은 삽입체 폴리뉴클레오타이드와 상기 제 1 및 제 2 폴리뉴클레오타이드의 어셈블리를 위해 이용될 수 있다.

[0071] 제 1 및/또는 제 2 폴리뉴클레오타이드는 각각 길이가 변할 수 있고, 일부 경우에 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 950 또는 1000개의 뉴클레오타이드 염기 길이 및/또는 1kb 또는 2kb 이상의 길이일 수 있다. 대안적으로, 제 1 및/또는 제 2 폴리뉴클레오타이드는 2kb 이상, 또는 1kb 이상 또는 900개 염기, 800개 염기, 700개 염기, 600개 염기, 500개 염기, 400개 염기, 300개 염기, 200개 염기 또는 100개 염기 초과 길이일 수 있다. 제 1 및/또는 제 2 폴리뉴클레오타이드 길이는 100개 뉴클레오타이드-2kb 범위, 예를 들어 최대 100개, 최대 150개, 최대 200개, 최대 250개, 최대 300개, 최대 350개, 최대 400개, 최대 450개, 최대 500개, 최대 550개, 최대 600개, 최대 650개, 최대 700개, 최대 750개 또는 최대 800개, 최대 850개, 최대 900개, 최대 950개, 최대 1000개, 최대 1500개 또는 최대 2000개 뉴클레오타이드일 수 있다. 제 1 및/또는 제 2 폴리뉴클레오타이드의 최소 길이는 경험적으로 결정되는 바람직한 T_m 에 의해 정의될 수 있다.

[0072] 본 발명에 기술된 바와 같이, 각각의 제 1 및 제 2 폴리뉴클레오타이드 서열은 삽입체 폴리뉴클레오타이드와 상기 제 1 및 제 2 폴리뉴클레오타이드의 어셈블리를 돕는 서열을 포함할 수 있다. 상기 어셈블리를 돕기 위해, 상기 서열은 삽입체 폴리뉴클레오타이드 상에 존재하는 어셈블리 오버랩 서열에 상보적일 수 있다. 삽입체 폴리뉴클레오타이드 상에 존재하는 어셈블리 오버랩 서열에 상보적인 서열은 어셈블리 오버랩 서열로도 지칭될 수 있다. 한 실시태양에서, 어셈블리 오버랩 서열은 전체 제 1 및/또는 제 2 폴리 뉴클레오타이드를 나타낸다. 다른 실시태양에서, 어셈블리 오버랩 서열은 제 1 및/또는 제 2 폴리뉴클레오타이드의 일부만을 나타내고, 제 1 및/또는 제 2 폴리뉴클레오타이드는 어셈블리 오버랩 서열을 넘어서는 추가 서열을 추가로 포함한다. 한 실시태양에서, 본 발명에 제공된 바와 같은 한 쌍의 제 1 및 제 2 폴리뉴클레오타이드의 제 1 폴리뉴클레오타이드는 삽입체 폴리뉴클레오타이드의 5' 또는 근위 말단에 존재하는 제 1 어셈블리 오버랩 서열에 상보적인 이의 원위 또는 3' 말단에 어셈블리 오버랩 서열을 포함하는 반면 상기 쌍의 제 2 폴리뉴클레오타이드는 상기 삽입체 폴리뉴클레오타이드의 3' 또는 원위 말단에 존재하는 제 2 어셈블리 오버랩 서열에 상보적인 이의 근위 또는 5' 말단에 오버랩 어셈블리 오버랩 서열을 포함한다. 이 실시태양에 더하여, 제 1 및 제 2 폴리뉴클레오타이드는 각각 어셈블리 오버랩 서열을 넘어서 추가 서열을 포함할 수 있다. 제 1 및/또는 제 2 폴리뉴클레오타이드의 추가

서열은 상기 제 1 및/또는 제 2 폴리뉴클레오타이드를 특정 응용분야에 맞출 수 있다. 특정 응용분야는 당업계에 공지된 핵산 라이브러리를 이용하는 임의의 응용분야, 특히 풀링된 결정적 어셈블리로부터 이익을 얻을 수 있는 응용분야일 수 있다. 예시적인 용도는 게놈 편집 및 경로 어셈블리를 포함할 수 있지만 이에 제한되지는 않는다.

[0073] 제 1 및/또는 제 2 폴리뉴클레오타이드 상에 존재하는 어셈블리 오버랩 서열은 길이가 다양할 수 있으며, 일부 경우에 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30개 뉴클레오타이드 길이 및/또는 최대 100개 뉴클레오타이드 길이(예를 들어, 최대 50개, 최대 30개, 최대 25개, 최대 20개 또는 최대 15개 뉴클레오타이드 길이)일 수 있다. 어셈블리 오버랩 서열 길이는 15개 뉴클레오타이드-100개 뉴클레오타이드 범위, 예를 들어 최대 20개, 최대 25개, 최대 30개, 최대 35개, 최대 40개, 최대 45개, 최대 50개, 최대 55개, 최대 60개, 최대 65개, 최대 70개, 최대 75개, 최대 80개 뉴클레오타이드, 최대 85개 뉴클레오타이드, 최대 90개 뉴클레오타이드, 최대 95개 뉴클레오타이드 또는 최대 100개 뉴클레오타이드일 수 있다. 어셈블리 오버랩 서열은 삽입체 폴리뉴클레오타이드에 존재하는 어셈블리 오버랩 서열과 동일한 길이일 수 있다. 어셈블리 오버랩 서열의 최소 길이는 경험적으로 결정되는 바람직한 T_m 에 의해 정의될 수 있다. 한 실시태양에서, 제 1 및/또는 제 2 폴리뉴클레오타이드 상의 어셈블리 오버랩 서열은 삽입체 폴리뉴클레오타이드의 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함한다. 다른 실시태양에서, 제 1 및/또는 제 2 폴리뉴클레오타이드 상의 어셈블리 오버랩 서열은 삽입체 폴리뉴클레오타이드의 말단에 상보적인 약 25개의 뉴클레오타이드를 포함한다.

[0074] 도 1에 도시된 바와 같이, 제 1 및 제 2 폴리뉴클레오타이드의 각각의 쌍은 클로닝 벡터와 벡터 오버랩 서열을 추가로 포함하여 제 1 폴리뉴클레오타이드(즉, 도 1의 제 1 DNA 단편)가 이의 5' 말단에서 클로닝 벡터에 대한 벡터 오버랩 서열을 포함하는 반면 제 2 폴리뉴클레오타이드(즉, 도 1의 제 2 DNA 단편)는 이의 3' 말단에서 클로닝 벡터에 대한 벡터 오버랩 서열을 포함할 수 있다. 한 쌍의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드 각각이 본 발명에 제공된 바와 같은 제 1 및 제 2 DNA 단편을 추가로 포함하는 실시태양에서, 상기 제 1 및 제 2 DNA 단편은 제 1 폴리뉴클레오타이드의 클로닝 벡터에 대한 벡터 오버랩 서열에 하류 및 인접하게 위치될 수 있고 제 2 폴리뉴클레오타이드의 클로닝 벡터에 대한 벡터 오버랩 서열에 상류 및 인접하게 위치될 수 있다.

[0075] 벡터 오버랩 서열은 길이가 다양할 수 있고, 일부 경우에 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30개 뉴클레오타이드 길이 및/또는 최대 100개 뉴클레오타이드 길이(예를 들어, 최대 50개, 최대 30개, 최대 25개, 최대 20개 또는 최대 15개 뉴클레오타이드 길이)일 수 있다. 대안적으로, 벡터 오버랩 서열은 2kb 이하, 또는 1kb 이하 또는 900개 염기, 800개 염기, 700개 염기, 600개 염기, 500개 염기, 400개 염기, 300개 염기, 200개 염기 또는 100개 염기 미만일 수 있다. 벡터 오버랩 서열 길이는 15개 뉴클레오타이드-80개 뉴클레오타이드 범위, 예를 들어 최대 20개, 최대 25개, 최대 30개, 최대 35개, 최대 40개, 최대 45개, 최대 50개, 최대 55개, 최대 60개, 최대 65개, 최대 70개, 최대 75개 또는 최대 80개 뉴클레오타이드일 수 있다. 벡터 오버랩 서열의 최소 길이는 경험적으로 결정되는 바람직한 T_m 에 의해 정의될 수 있다.

[0076] 한 실시태양에서, 제 1 및 제 2 폴리뉴클레오타이드 쌍을 함유하는 풀은 상기 풀로부터의 폴리뉴클레오타이드가 지정된 역치를 넘어 상기 풀의 다른 폴리뉴클레오타이드와 공통 서열을 공유하지 않아서, 본 발명에 제공된 상기 풀의 폴리뉴클레오타이드의 쌍과 이의 풀의 삽입체 폴리뉴클레오타이드 사이 또는 상기 풀과 클로닝 벡터 사이의 디자인된 어셈블리 오버랩 서열을 제외하도록 이런 서열의 더 큰 세트로부터의 폴리뉴클레오타이드 서열의 쌍을 선택함으로써 생성된다. 지정된 역치는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 또는 20개의 인접 뉴클레오타이드이다. 지정된 역치는 최대 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 또는 20개의 연속 뉴클레오타이드인접 뉴클레오타이드이다. 지정된 역치는 0 내지 2, 1 내지 3, 2 내지 4, 3 내지 5, 4 내지 6, 5 내지 7, 6 내지 8, 7 내지 9, 8 내지 10, 9 내지 11, 10 내지 12, 11 내지 13, 12 내지 14, 13 내지 15, 14 내지 16, 15 내지 17, 16 내지 18, 17 내지 19, 18 내지 20 또는 19 내지 21개 연속 뉴클레오타이드이다. 지정된 역치는 0 내지 5, 0 내지 10, 0 내지 15, 0 내지 20, 5 내지 10, 5 내지 15, 5 내지 20, 10 내지 15 또는 10 내지 20개 연속 뉴클레오타이드이다. 한 실시태양에서, 지정된 역치는 12개의 인접 뉴클레오타이드이다. 지정된 역치를 초과하는 공유 공통 서열의 결정은 BLAST 분석 또는 간단한 하위 문자열 검색을 사용하여 구성 요소가 다른 구성 요소와 서열을 공유하는지 여부를 결정하는 컴퓨터 프로그램을 사용하여 수행될 수 있다. 공유 서열이 지정된 역치를 초과하여 발견되는 경우, 구성 요소가 함께 풀에 배치되지 않는다.

[0077] 한 실시태양에서, 본 발명에 기술된 삽입체 폴리뉴클레오타이드와 원하는 쌍의 제 1 및 제 2 폴리뉴클레오타이드의 페어링은 도 2에 도시된 바와 같은 "인사이드-아웃(inside-out) 어셈블리" 방법을 사용하여 제 1 및 제 2 폴리뉴클레오타이드의 원하는 쌍을 사전어셈블링함으로써 용이해질 수 있다. 이 방법에서, 제 1 및 제 2 폴리뉴클레오타이드는 PCR에 의해 증폭될 수 있어 제 1 폴리뉴클레오타이드의 벡터 근위 말단은 각각 하나 이상의 고유한 부위 특이적 뉴클레아제 부위(들) 또는 인식 서열(들)을 포함한다. 부위 특이적 뉴클레아제 인식 서열은 제한 엔도뉴클레아제, 유형 II 엔도뉴클레아제(들), 귀소 엔도뉴클레아제, RNA 유도 뉴클레아제, DNA 유도 뉴클레아제, 징크 핑거 뉴클레아제, TALEN 및 녹아웃 효소 및 이의 임의의 조합으로부터 선택된 부위 특이적 뉴클레아제에 대한 것일 수 있다. 한 실시태양에서, 제 1 폴리뉴클레오타이드의 벡터 근위 말단은 각각 단일 고유 뉴클레아제 부위 또는 인식 서열을 포함한다. 한 실시태양에서, 고유한 뉴클레아제 인식 서열은 고유한 제한 엔도뉴클레아제 부위로서 상기 제한 엔도뉴클레아제 부위가 본 발명에 제공된 조성물에 존재하는 임의의 폴리뉴클레오타이드에 존재하지 않는다. 한 실시태양에서, 고유한 뉴클레아제 부위는 예를 들어 I-SceI 또는 I-CeuI에 특이적인 귀소 엔도뉴클레아제 서열과 같은 귀소 엔도뉴클레아제 서열이다. 단일 쌍의 제 1 및 제 2 폴리뉴클레오타이드가 결합되고 스플라이싱 및 오버랩 연장 중합 효소 연쇄 반응(SOE-PCR)이 수행되어 추가된 고유한 뉴클레아제 부위(예를 들어, 벡터-근위 말단에서)에서 2개의 폴리뉴클레오타이드를 어셈블하여 삽입체 폴리뉴클레오타이드에 부착하는 말단을 자유롭게 한다. 대안적으로, 연결된 제 1 및 제 2 폴리뉴클레오타이드를 포함하는 전체 서열은 당업계에 공지된 다양한 DNA 합성 방법 중 임의의 것을 사용하여 직접 합성될 수 있다. 부착된 제 1 및 제 2 폴리뉴클레오타이드는 당 업계에 공지되고/되거나 본 발명에 제공된 시험관 내 또는 생체 내 오버랩 어셈블리 방법, 예를 들어, 효모(예를 들어, S 세레비지애) 또는 대장균 상동성 재조합 기반 어셈블리, 김슨 어셈블리 또는 NEB® HiFi 빌더를 사용하여 삽입체 폴리뉴클레오타이드와 어셈블된다. 삽입체 폴리뉴클레오타이드를 갖는 제 1 및 제 2 폴리뉴클레오타이드의 원형화된 산물은 고유한 뉴클레아제 서열(예를 들어, 특정 귀소 엔도뉴클레아제 서열에 대한 귀소 엔도뉴클레아제)에 특이적인 뉴클레아제의 첨가로 선형화될 수 있어 삽입체 폴리뉴클레오타이드가 그 다음 김슨 어셈블리 또는 기타 유사한 방법을 사용하여 벡터로 어셈블될 수 있는 제 1 및 제 2 폴리뉴클레오타이드에 의해 플랭킹되게 한다.

[0078] **삽입체 폴리뉴클레오타이드/페이로드 서열**

[0079] 한 실시태양에서, 본 발명에 제공된 조성물, 키트 또는 방법에서 사용하기 위한 삽입체 폴리뉴클레오타이드는 (1) 상기 삽입체 폴리뉴클레오타이드의 5' 또는 근위 말단 상의 제 1 어셈블리 오버랩 서열, 및 (2) 상기 삽입체 폴리뉴클레오타이드의 대향 3' 또는 원위 말단 상의 제 2 어셈블리 오버랩 서열을 포함한다. 이 실시태양에 추가로, 제 1 어셈블리 오버랩 서열은 제 1 및 제 2 폴리뉴클레오타이드의 한 쌍으로부터의 제 1 폴리뉴클레오타이드의 3' 또는 원위 말단에서 서열에 상보적인 서열(예를 들어, 어셈블리 오버랩 서열)을 포함할 수 있는 반면, 제 2 어셈블리 오버랩 서열은 제 1 및 제 2 폴리뉴클레오타이드의 한 쌍으로부터의 제 2 폴리뉴클레오타이드의 5' 또는 근위 말단에서 서열에 상보적인 서열(예를 들어, 어셈블리 오버랩 서열)을 포함할 수 있다.

[0080] 본 발명에 제공된 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 본 발명에 제공된 폴리뉴클레오타이드의 쌍에서 각각, 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50개 이상의 뉴클레오타이드 길이 및/또는 최대 100개 뉴클레오타이드 길이(예를 들어, 최대 50개, 최대 30개, 최대 25개, 최대 20개 또는 최대 15개 뉴클레오타이드 길이)일 수 있다. 어셈블리 오버랩 서열 길이는 15개 뉴클레오타이드-100개 뉴클레오타이드 범위, 예를 들어 최대 20개, 최대 25개, 최대 30개, 최대 35개, 최대 40개, 최대 45개, 최대 50개, 최대 55개, 최대 60개, 최대 65개, 최대 70개, 최대 75개, 최대 80개 뉴클레오타이드, 최대 85개 뉴클레오타이드, 최대 90개 뉴클레오타이드, 최대 95개 뉴클레오타이드 또는 최대 100개 뉴클레오타이드일 수 있다. 한 실시태양에서, 본 발명에 제공된 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 본 발명에 제공된 폴리뉴클레오타이드의 쌍에서 각각, 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함한다. 다른 실시태양에서, 본 발명에 제공된 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 본 발명에 제공된 폴리뉴클레오타이드의 쌍에서 각각, 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 약 25개의 뉴클레오타이드를 포함한다.

[0081] 다른 실시태양에서, 삽입체 폴리뉴클레오타이드는 하나 이상의 페이로드 서열을 추가로 포함하여 상기 하나 이상의 페이로드 서열이 제 1 및 제 2 어셈블리 오버랩 서열 사이에 위치하도록 한다. 페이로드 서열은 임의의 서열일 수 있다. 페이로드 서열은 마커 서열일 수 있다. 마커 서열은 당업계에 공지된 임의의 마커 서열일 수 있다.

다. 페이로드 서열은 유전자 또는 이의 일부일 수 있다. 유전자 또는 이의 일부는 대사 또는 생화학적 경로의 일부일 수 있다. 유전자 또는 이의 일부는 단백질 또는 이의 도메인을 암호화할 수 있다. 페이로드 서열은 프로모터, 유전자, 조절 서열, 디그론을 암호화하는 핵산 서열, 용해도 태그를 코딩하는 핵산 서열, 분해 태그를 암호화하는 핵산 서열, 종결자, 바코드, 조절 서열 또는 이의 일부로부터 선택될 수 있다. 일부 경우에, 삽입체 폴리뉴클레오타이드의 3 가지 구성 요소(즉, 제 1 어셈블리 오버랩 서열, 제 2 어셈블리 오버랩 서열 및 페이로드 서열)는 본 발명에 제공된 어셈블리 방법에서 사용하기 전에 합성되거나 그렇지 않으면 인접한 DNA 조각으로 결합된다. 한 실시태양에서, 제 1 및 제 2 어셈블리 오버랩은 무작위가 아니며 제 1 및 제 2 폴리뉴클레오타이드의 특정 쌍과 일치하도록 디자인된다.

[0082] 제 1 및 제 2 폴리뉴클레오타이드 쌍이 본 발명에 기술된 바와 같은 표적화 서열을 포함하는 실시태양에서, 삽입체 폴리뉴클레오타이드 내에 존재하는 페이로드 서열은 제 1 및 제 2 폴리뉴클레오타이드의 쌍 상의 표적화 서열에 의해 표적화된 원래 유전자좌에 상대적인 삽입, 제 1 및 제 2 폴리뉴클레오타이드의 쌍 상의 표적화 서열에 의해 표적화된 원래 유전자좌에 대한 서열의 결실, 또는 한 서열의 다른 서열로의 대체를 초래할 수 있다. 삽입 또는 수정의 경우, '페이로드'는 의도한 최종 서열일 수 있다. 결실의 경우 '페이로드'는 마커 서열이거나 서열이 없을 수 있다.

[0083] 한 실시태양에서, 삽입체 폴리뉴클레오타이드는 풀링 방식으로 사용된다. 이 실시태양에 추가로, 삽입체 폴리뉴클레오타이드 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 및 제 2 폴리 뉴클레오타이드의 쌍의 제 1 폴리뉴클레오타이드의 3' 또는 위위 말단에서 서열(예를 들어, 어셈블리 오버랩 서열)에 상보적인 서열을 포함하는 제 1 어셈블리 오버랩 서열 및 제 1 및 제 2 폴리 뉴클레오타이드 쌍으로부터의 제 2 폴리뉴클레오타이드의 5' 또는 근위 말단에서 서열(예를 들어, 어셈블리 오버랩 서열)에 상보적인 서열을 포함하는 제 2 어셈블리 오버랩 서열을 포함할 수 있다.

[0084] 삽입체 폴리뉴클레오타이드 풀은 임의의 수의 고유한 삽입체 폴리뉴클레오타이드 서열을 포함할 수 있다. 삽입체 폴리뉴클레오타이드의 수는 페이로드 서열을 갖거나 갖지 않는 적어도, 최대 또는 약 1, 5, 10, 25, 50, 75, 100, 125, 150, 175, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10,000, 20, 000, 30,000, 40,000, 50,000, 75,000, 100,000, 150,000, 200,000 또는 250,000개의 고유한 삽입체 폴리뉴클레오타이드일 수 있다.

[0085] 페이로드 서열은 최대 또는 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000 또는 10,000개의 뉴클레오타이드 길이일 수 있다. 일부 경우에, 페이로드 서열은 0개 뉴클레오타이드 길이일 수 있다. 페이로드 서열은 삽입체 폴리뉴클레오타이드에 통합될 때, 전체 삽입체 폴리뉴클레오타이드가 화학적으로 합성될 수 있는 길이일 수 있다. 합성은 당업계에 공지된 바와 같은 어레이 기반 또는 컬럼 기반 합성 방법일 수 있다. 한 실시태양에서, 페이로드 서열은 제 1 및 제 2 어셈블리 오버랩 함께 삽입체 올리고뉴클레오타이드에 직접 포함되거나 합성될 수 있는 길이를 갖는다. 합성될 수 있는 삽입체 폴리뉴클레오타이드는 최대 약 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 250, 300, 350, 400개 이상의 뉴클레오타이드 길이일 수 있다.

[0086] 다른 실시태양에서, 삽입체 폴리뉴클레오타이드는 도 3에 기술된 방법을 사용하여 단일 풀에서 생성될 수 있다. 도 3에 도시된 바와 같이, 페이로드 서열(예를 들어, 도 3의 프로모터 서열)은 풀링된 정방향 프라이머, 공통 역방향 프라이머 및 페이로드 템플릿 서열(예를 들어, 도 3의 프로모터)의 세 가지 구성 요소로부터 PCR을 통해 생성될 수 있다. 페이로드 서열 템플릿은 합성 DNA 단편, PCR 산물 또는 기타 단일 또는 이중 가닥 DNA 단편일 수 있다. 전방 프라이머 풀은 당업계에 공지된 어레이 기반 또는 컬럼 기반 합성 방법을 사용하여 합성될 수 있다. 풀의 각 정방향 프라이머는 다음을 포함할 수 있다(5'에서 3'까지): 1) 페이로드 주형 서열의 위위 또는 3' 말단에 상보적인 서열, 2) 제 1 및 제 2 폴리뉴클레오타이드의 쌍의 제 2 폴리뉴클레오타이드에 상보적인 서열을 포함하는 제 2 어셈블리 오버랩 서열, 3) 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열(예를 들어, 귀소 엔도뉴클레아제 부위 또는 인식 서열), 4) 제 1 및 제 2 폴리뉴클레오타이드의 쌍의 제 1 폴리뉴클레오타이드에 상보적인 서열을 포함하는 제 1 어셈블리 오버랩 서열, 및 5) 페이로드 템플릿 서열의 근위 말단 또는 5' 말단에 결합하는 프라이밍 서열. 공통 역방향 프라이머는 페이로드 주형 서열의 말단 또는 3' 말단 또는 페이로드 서열의 다른 서열 또는 하류에 결합될 수 있다. PCR은 풀링된 정방향 프라이머 및 공통 역방향 프라이머를 사용하여 페이로드 템플릿 서열 (예를 들어, 도 3의 프로모터)에서 수행될 수 있다. 증폭 후,

PCR 산물은, 예를 들어, 깃슨 어셈블리, NEB® HIFI 어셈블리 또는 유사한 방법과 같은 당업계에 알려진 오버랩 어셈블리 방법을 사용하여 원형-퍼뮤트 페이로드(삽입체)를 생성하기 위해 원형화될 수 있으며, 그런 다음 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 인식하는 하나 이상의 부위 특이적 뉴클레아제(들)(예를 들어, 도 3의 귀소 엔도뉴클레아제, I-SceI)를 사용하여 선형화될 수 있다. 뉴클레아제 분해는 삽입체 폴리뉴클레오타이드(예를 들어, 도 1에 기술된 "페이로드" 부분)로 사용하기에 적합한 단편을 생성할 수 있으며, 큰 페이로드는 제 1 및 제 2 어셈블리 오버랩 서열(예를 들어, 상동성 암 또는 도 3의 프로모터 서열에 플랭킹하는 영역)에 의해 플랭킹된다. 도 3에 도시된 바와 같이, 말단에서 페이로드 서열은 사용된 오버랩 어셈블리 방법(예를 들어, 깃슨 어셈블리 시약, NEB® HIFI 어셈블리 시약 또는 동등한 혼합물의 3' 및 5' 엑소 뉴클레아제 활성)에 의해 절제될 수 있는 작은 부분 뉴클레아제 인식 서열(예를 들어, 도 3의 I-SceI)일 수 있다. 산물은 원형화 후 및 선형화 전에 선택적으로 증폭될 수 있다(예를 들어, RCA).

[0087] 한 실시태양에서, 각각의 삽입체 폴리뉴클레오타이드는 페이로드 서열을 포함하여 삽입체 폴리뉴클레오타이드의 풀에 있는 각각의 삽입체 폴리뉴클레오타이드가 상기 풀에서 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 상이한 페이로드 서열을 포함하게 한다.

[0088] 다른 실시태양에서, 각각의 삽입체 폴리뉴클레오타이드는 페이로드 서열을 포함하여 삽입체 폴리뉴클레오타이드의 풀에 있는 각각의 삽입체 폴리뉴클레오타이드가 상기 풀에서 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 동일한 페이로드 서열을 포함하게 한다.

[0089] **클로닝 방법**

[0090] 본 발명에 기술된 바와 같이, 삽입체 폴리뉴클레오타이드뿐만 아니라 제 1 및 제 2 폴리뉴클레오타이드의 쌍을 포함하는 조성물은 삽입체 폴리뉴클레오타이드를 사이에 두고 제 1 및 제 2 폴리뉴클레오타이드를 포함하는 핵산 라이브러리로 어셈블링될 수 있다. 본 발명에 제공된 바와 같은 삽입체 폴리뉴클레오타이드와 제 1 및 제 2 폴리뉴클레오타이드 쌍의 어셈블리는 시험관 내 또는 생체 내 클로닝 방법에 의해 수행될 수 있다. 큰 DNA 분자의 어셈블리를 위해, 어셈블리의 최종 단계는 효모 숙주 세포와 같은 생체 내에서 수행될 수 있다. 시험관 내 및 생체 내 어셈블리 단계 사용 간의 균형은 어셈블될 핵산 분자의 특성과 관련하여 방법의 실용성에 의해 결정될 수 있다.

[0091] 한 실시태양에서, 제 1 및 제 2 폴리뉴클레오타이드 쌍과 삽입체 폴리뉴클레오타이드의 어셈블리는 시험관 내 클로닝 방법을 사용하여 수행된다. 시험관 내 클로닝 방법은 당업계에서 공지된 오버랩 어셈블리를 사용하는 임의의 시험관 내 클로닝 방법일 수 있다. 본 발명에 제공된 방법에서 사용되는 시험관 내 클로닝 방법은 주입 클로닝(Clontech®), 골든 게이트 어셈블리, 게이트웨이 어셈블리, 깃슨 어셈블리 및 NEB® HIFI 어셈블리 또는 당 업계에 알려진 임의의 다른 적합한 시험관 내 클로닝 방법으로부터 선택될 수 있다. 주입 클로닝은 본 발명에 제공된 바와 같은 제 1 및 제 2 폴리뉴클레오타이드 쌍의 제 1 풀 및 본 발명에 기술된 바와 같은 삽입체 폴리뉴클레오타이드의 제 2 풀을 주입 클로닝 시약과 혼합한 다음 생성된 어셈블리를 대장균 클로닝 숙주 세포로 형질 전환시키는 것을 수반할 수 있다. 시험관 내 클로닝 방법은 US 8,968,999에 기술된 임의의 오버랩 어셈블리 방법일 수 있으며, 이는 그 전체가 본 발명에 참조로 포함된다. 시험관 내 클로닝 방법은 US20160060671에 기술된 임의의 오버랩 어셈블리 방법일 수 있으며, 이는 그 전체가 본 발명에 참조로 포함된다. 시험관 내 클로닝 방법은 Jun Urano, Ph.D. and Christine Chen, Ph.D., Gibson Assembly® Primer-Bridge End Joining (PBnJ) Cloning, Synthetic Genomics Application Note에 기술된 깃슨 어셈블리 방법일 수 있으며, 이는 그 전체가 본 발명에 참조로 포함된다. 한 실시태양에서, 제 1 및 제 2 폴리뉴클레오타이드의 쌍, 삽입체 폴리뉴클레오타이드 및 클로닝 벡터를 포함하는 조성물은 5'-3' 엑소뉴클레아제; 및 조성물에 또한 존재하는 가닥 치환 중합 효소를 사용하여 연결된다. 조성물은 또한 7mM-150mM, 예를 들어 20mM-50mM 농도 범위의 염화칼륨과 같은 칼륨 염을 함유하는 완충액을 포함할 수 있다. 20mM와 같은 10mM-100mM 범위의 나트륨 염(예를 들어, 염화나트륨)이 칼륨 염에 추가로 사용될 수도 있다. 일부 실시태양에서, 조성물은 폴리에틸렌 글리콜(PEG), 피콜 또는 텍스 트란과 같은 크라우딩제를 함유하지 않는다. 일부 실시태양에서, 조성물은 단일 가닥(ss) 결합 단백질질을 포함한다. 조성물에 사용하기 위한 ss DNA 결합 단백질은 E. coli recA, T7 유전자 2.5 산물, RedB(파지 람다로부터) 또는 RecT(Rac 프로 파지로부터), ET SSB(극단 열 안정성 단일 가닥 DNA 결합 단백질) 또는 임의의 것일 수 있다. 당 업계에 공지된 다른 ss DNA 결합 단백질이 조성물에 사용될 수 있다. ss 결합 단백질의 포함은 특히 콜로니 수에 의해 측정된 바와 같이 ss 결합 단백질의 부재에서 발생하는 것보다 더 긴 오버랩 서열(예를 들어, 적어도 20개의 뉴클레오타이드)을 갖는 핵산 단편에 대한 어셈블리 효율을 개선할 수 있다. 일부 실시태양에서, 조성물은 비 가닥 치환 중합 효소를 함유하지 않는다.

[0092] 다른 실시태양에서, 제 1 및 제 2 폴리뉴클레오타이드의 쌍, 삽입체 폴리뉴클레오타이드 및 클로닝 벡터를 포함하는 조성물은 3' 엑소 뉴클레아제 활성이 결여된 단리된 비 열성 5' 내지 3' 엑소뉴클레아제, 크라우딩제, 3' 엑소뉴클레아제 활성을 갖는 비 가닥 치환 DNA 중합 효소, 또는 상기 DNA 중합 효소와 3' 엑소 뉴클레아제 활성이 결여된 제 2 DNA 중합 효소 및 리가아제의 혼합물을 사용하여 결합된다. 조성물은 폴리뉴클레오타이드와 클로닝 벡터를 결합하는 데 효과적인 조건 하에서 dNTP의 혼합물 및 적합한 완충액을 추가로 포함할 수 있다. 일부 실시태양에서, 조성물은 크라우딩제를 추가로 포함할 수 있다. 크라우딩제는 폴리에틸렌 글리콜(PEG), 텍스트란 또는 피콜에서 선택될 수 있다. 한 실시태양에서, 크라우딩제는 PEG이다. PEG는 약 3 내지 약 7%(중량/부피)의 농도로 사용될 수 있다. PEG는 PEG-200, PEG-4000, PEG-6000, PEG-8000 또는 PEG-20,000 중에서 선택될 수 있다. 일부 실시태양에서, 엑소뉴클레아제는 T5 엑소뉴클레아제이고 접촉은 등은 조건하에 있고/있거나 크라우딩제는 PEG이고/이거나 비 가닥 치환 DNA 중합 효소는 PHUSION® DNA 중합 효소 또는 VENTR® DNA 중합 효소, 및/또는 Taq 리가아제이다.

[0093] 한 실시태양에서, 제 1 및 제 2 폴리뉴클레오타이드 쌍을 삽입체 폴리뉴클레오타이드와 어셈블리는 생체 내 클로닝 방법을 사용하여 수행된다. 생체 내 클로닝 방법은 당업계에서 공지된 임의의 생체 내 클로닝 방법일 수 있다. 생체 내 클로닝 방법은 상동성 재조합 매개 클로닝 방법일 수 있다. 본 발명에 제공된 방법에서 사용되는 생체 내 클로닝 방법은 Tsuge, Kenji et al. "One step assembly of multiple DNA fragments with a designed order and orientation in Bacillus subtilis plasmid." Nucleic acids research vol. 31,21 (2003): e133에 기술된 바와 같이 대장균(RecA 의존성, RecA 의존성 또는 Red/ET 의존성) 상동성 재조합, OEPR(Overlap Extension PCR and Recombination) 클로닝, 효모 상동성 재조합, 및 바실러스에서의 형질 전환 관련 재조합 (TAR) 클로닝 및 유전자 어셈블리로부터 선택될 수 있으며, 이는 본 발명에 참고로 포함된다.

[0094] **응용분야**

[0095] 본 발명에 제공된 조성물 및 어셈블리 방법은 플라스미드, 벡터, 유전자, 대사 경로, 최소 게놈, 부분 게놈, 게놈, 염색체, 염색체 외 핵산, 예를 들어 세포질 세포 기관, 미토콘드리아(동물), 엽록체 및 색소체(식물) 등과 같은 임의의 원하는 조립체를 구축하는 데 사용될 수 있다.

[0096] 본 발명에 제공된 조성물 및 어셈블리 방법은 핵산 분자의 라이브러리를 생성하고 그로부터 생성된 변형된 전체 또는 부분 핵산 분자를 사용하는 방법을 사용하는 데 사용될 수 있다. 라이브러리는 2개 이상의 변이체를 포함할 수 있으며, 상기 다수의 변이체는 관심 원하는 산물의 높은 생산 수준, 관심 산물의 향상된 기능 또는 감소된 기능과 같은 원하는 특성을 가진 구성원에 대해 스크리닝될 수 있다(그것이 유리한 경우). 그러한 스크리닝은 본 발명에 제공된 대로 로봇식/자동화될 수 있는 고 처리량 방법에 의해 수행될 수 있다.

[0097] 본 발명은 또한 본 발명에 제공된 조성물 및 어셈블리 방법에 의해 제조된 산물, 예를 들어 생성된 어셈블된 합성 유전자 또는 게놈(합성 또는 자연 발생) 및 변형된 최적화 유전자 및 게놈, 및 이의 용도(들)를 추가로 포함한다.

[0098] 본 발명에 제공된 조성물 및 어셈블리 방법은, 예를 들어, 원하는 관심 산물의 합성을 위한 경로의 디자인 또는 유전자 산물이 원하는 산물의 합성 또는 발현에서 역할을 하는 하나 이상의 서열의 최적화를 허용하는 매우 다양한 응용분야를 가질 수 있다. 본 발명에 제공된 조성물 및 어셈블리 방법은 또한 유전자 또는 이의 발현의 최적화된 서열을 생성하거나 유전자에 의해 코딩되는 단백질의 하나 이상의 기능적 도메인 또는 모티프를 결합하는 데 사용될 수 있다. 유전자는 생화학 적 또는 대사 경로의 일부일 수 있다. 생화학적 또는 대사 경로는 원하는 관심 산물을 생성할 수 있다.

[0099] 원하는 관심 산물은 세포 배양, 진핵 또는 원핵 발현 시스템 또는 트랜스 제닉 동물 또는 식물에서 어셈블될 수 있는 임의의 분자일 수 있다. 따라서, 본 발명에 제공된 결정적 어셈블리 방법으로부터 생성되는 핵산 분자 또는 이의 라이브러리는 원하는 관심 산물을 생산하기 위해 다양한 맥락에서 사용될 수 있다. 일부 경우에 관심 산물은 소분자, 효소, 펩타이드, 아미노산, 유기산, 합성 화합물, 연료, 알코올 등일 수 있다. 예를 들어, 관심 산물 또는 생체 분자는 1차 또는 2차 세포 외 대사 산물일 수 있다. 1차 대사 산물은 특히 에탄올, 시트르산, 젖산, 글루탐산, 글루타메이트, 라이신, 트레오닌, 트립토판 및 기타 아미노산, 비타민, 다당류 등일 수 있다. 2차 대사 산물은 특히 페니실린, 또는 사이클로스포린 A와 같은 면역 억제제, 지베렐린과 같은 식물 호르몬, 로바스타틴과 같은 스타틴 약물, 그리 세오 폴빈과 같은 살진균제 등과 같은 항생제 화합물일 수 있다. 관심 산물 또는 생체 분자는 또한 카탈라아제, 아밀라아제, 프로테아제, 펙티나아제, 글루코스 이성화효소, 셀룰라아제, 헤미셀룰라아제, 리파아제, 락타아제, 스트렙토키나아제 등을 포함한 미생물 효소와 같은 숙주 세포에 의해 생성된 다음과 같은 세포 내 성분일 수 있다. 세포 내 성분은 또한 인슐린, B 형 간염 백신, 인터페론, 과립구 콜

로니 자극 인자, 스트렙토키나아제 등과 같은 재조합 단백질을 포함할 수 있다. 관심 산물은 관심 단백질을 지칭할 수도 있다.

[0100] 경로 어셈블리

[0101] 한 실시태양에서, 본 발명에 제공된 조성물 및 방법은 유전자 또는 그의 변이체를 어셈블하는 데 사용된다. 유전자 또는 그의 변이체는 대사 또는 생화학적 경로의 일부인 단백질을 암호화할 수 있다. 변이체는 코돈 최적화된 버전 또는 상기 유전자의 돌연변이된 버전일 수 있다. 대사 또는 생화학적 경로는 본 발명에 제공된 관심 산물을 생산할 수 있다. 한 실시태양에서, 유전자 서열 또는 그의 변이체는 본 발명에 제공된 바와 같은 삽입체 폴리뉴클레오타이드 내에 페이로드 서열로서 존재할 수 있다. 제 1 및 제 2 폴리뉴클레오타이드의 쌍은 상기 삽입체 폴리뉴클레오타이드와 어셈블될 때 본 발명에 제공된 유전자 편집 방법을 사용하여 숙주 세포 내의 유전자 요소(예를 들어, 게놈, 플라스미드 등)의 유전자좌의 표적화 및 삽입을 용이하게 하는 역할을 할 수 있는 서열을 포함할 수 있다. 유전자좌는 특정 유전자좌 또는 무작위 유전자좌일 수 있다. 대안적으로, 제 1 및 제 2 폴리뉴클레오타이드의 쌍은 상기 삽입체 폴리뉴클레오타이드와 어셈블될 때 본 발명에 제공된 방법을 사용하여 생성된 다른 어셈블리와 생성된 어셈블리의 추가 어셈블리를 용이하게 하는 역할을 할 수 있는 서열을 포함할 수 있다. 다른 어셈블리는 동일한 대사 또는 생화학적 경로 내에 존재하는 하나 이상의 추가 유전자를 포함할 수 있으며, 이러한 방식으로 상기 대사 또는 생화학적 경로의 어셈블리를 용이하게 한다. 모든 유전자 또는 그의 변이체는 특정 대사 또는 생화학적 경로에 대한 단일 벡터 상의 오버랩 서열의 본 발명에 기재된 기술을 사용하여 어셈블될 수 있거나, 상기 경로의 각 구성원에 대한 독립적인 벡터는 연속적인 변형 혼합물에서 각 구성원에 대한 벡터를 혼합하여 사용될 수 있다. 삽입체 폴리뉴클레오타이드와 제 1 및 제 2 폴리뉴클레오타이드의 어셈블리는 본 발명에 제공된 어셈블리 오버랩 방법을 사용하여 각각의 폴리뉴클레오타이드에 존재하는 어셈블리 오버랩 서열을 통해 달성될 수 있다. 제 1 및 제 2 폴리뉴클레오타이드의 쌍은 적합한 벡터로의 어셈블리를 용이하게 하기 위해 본 발명에 제공된 바와 같은 벡터 오버랩 서열을 추가로 포함할 수 있다. 벡터는 복제 플라스미드일 수 있다. 일부 경우에, 제 1 및/제 2 폴리뉴클레오타이드는 유전자 또는 그의 변이체 또는 그의 전사, 번역, 용해도 또는 분해와 같은 이에 의해 암호화된 단백질의 양태를 지배할 수 있는 조절 또는 제어 요소의 서열을 추가로 포함할 수 있다. 조절 또는 제어 요소는 프로모터, 종결자, 용해도 태데그론해 태그 또는 디그론일 수 있다.

[0102] 다른 실시태양에서, 유전자 서열 또는 그의 변이체는 한 쌍의 제 1 및 제 2 폴리뉴클레오타이드 및 그 사이에 위치하는 삽입체 폴리뉴클레오타이드 또는 제 1 또는 제 2 폴리뉴클레오타이드와 그 사이에 위치한 삽입체 폴리뉴클레오타이드에 걸쳐 퍼져 있다. 각각의 폴리뉴클레오타이드 상의 적절한 어셈블리 오버랩 세그먼트에 의해, 모든 폴리뉴클레오타이드를 포함하는 혼합물은 본 발명에 제공된 오버랩 어셈블리를 사용하여 단일 반응 혼합물에서 정확한 순서로 어셈블될 수 있다. 결과는 유전자 또는 그의 변이체의 전장 코딩 서열이 될 것이다. 제 1 및 제 2 폴리뉴클레오타이드의 쌍은 상기 삽입체 폴리뉴클레오타이드와 어셈블될 때 본 발명에 제공된 유전자 편집 방법을 사용하여 숙주 세포 내의 유전자 요소(예를 들어, 게놈, 플라스미드 등)의 유전자좌의 표적화 및 삽입을 용이하게 하는 역할을 할 수 있는 서열을 추가로 포함할 수 있다. 유전자좌는 특정 유전자좌 또는 무작위 유전자좌일 수 있다. 대안적으로, 제 1 및 제 2 폴리뉴클레오타이드의 쌍은 상기 삽입체 폴리뉴클레오타이드와 어셈블될 때 본 발명에 제공된 방법을 사용하여 생성된 다른 어셈블리와 생성된 어셈블리의 추가 어셈블리를 용이하게 하는 역할을 할 수 있는 서열을 추가로 포함할 수 있다. 다른 어셈블리는 동일한 대사 또는 생화학적 경로 내에 존재하는 하나 이상의 추가 유전자를 포함할 수 있으며, 이러한 방식으로 상기 대사 또는 생화학적 경로의 어셈블리를 용이하게 한다. 모든 유전자 또는 그의 변이체는 특정 대사 또는 생화학적 경로에 대한 단일 벡터 상의 오버랩 서열의 본 발명에 기재된 기술을 사용하여 어셈블될 수 있거나, 상기 경로의 각 구성원에 대한 독립적인 벡터는 연속적인 변형 혼합물에서 각 구성원에 대한 벡터를 혼합하여 사용될 수 있다. 제 1 및 제 2 폴리뉴클레오타이드의 쌍은 적합한 벡터로의 어셈블리를 용이하게 하기 위해 본 발명에 제공된 바와 같은 벡터 오버랩 서열을 추가로 포함할 수 있다. 벡터는 복제 플라스미드일 수 있다. 일부 경우에, 제 1 및/제 2 폴리뉴클레오타이드는 유전자 또는 그의 변이체 또는 그의 전사, 번역, 용해도 또는 분해와 같은 이에 의해 암호화된 단백질의 양태를 지배할 수 있는 조절 또는 제어 요소의 서열을 추가로 포함할 수 있다. 조절 또는 제어 요소는 프로모터, 종결자, 용해도 태데그론해 태그 또는 디그론일 수 있다.

[0103] 다른 실시태양에서, 본 발명에 제공된 조성물 및 방법은 표적 단백질의 모티프 또는 도메인을 암호화하는 핵산 서열을 어셈블하거나 결합하는 데 사용된다. 표적 단백질의 특정 모티프 또는 도메인을 암호화하는 핵산 서열은 한 쌍의 제 1 및 제 2 폴리뉴클레오타이드 및 그 사이에 위치한 삽입체 폴리뉴클레오타이드에 걸쳐 퍼질 수 있거나, 제 1 또는 제 2 폴리뉴클레오타이드 및 그 사이에 위치한 삽입체 폴리뉴클레오타이드에 걸쳐 퍼질 수

다. 표적 단백질의 특정 모티프 또는 도메인을 암호화하는 핵산 서열은 제 1 폴리뉴클레오타이드에 존재할 수 있는 반면, 표적 단백질의 제 2 모티프 또는 도메인은 제 2 폴리뉴클레오타이드에 존재할 수 있고 삽입체 폴리뉴클레오타이드는 각 폴리뉴클레오타이드 상에 존재하는 어셈블리 오버랩 서열과 본 발명에 제공된 오버랩 어셈블리 방법을 사용하여 표적 단백질의 상기 제 1 및 제 2 모티프 또는 도메인을 연결하는 데 사용될 수 있다. 일부 경우에, 삽입체 폴리뉴클레오타이드는 제 1 및/또는 제 2 모티프 또는 도메인의 일부를 포함할 수 있다. 일부 경우에, 삽입체 폴리뉴클레오타이드는 표적 단백질의 제 3 모티프 또는 도메인을 포함할 수 있다. 제 1 및 제 2 폴리뉴클레오타이드의 쌍은 상기 삽입체 폴리뉴클레오타이드와 어셈블될 때 본 발명에 제공된 유전자 편집 방법을 사용하여 숙주 세포 내의 유전자 요소(예를 들어, 게놈, 플라스미드 등)의 유전자좌의 표적화 및 삽입을 용이하게 하는 역할을 할 수 있는 서열을 추가로 포함할 수 있다. 제 1 및 제 2 폴리뉴클레오타이드의 쌍은 적합한 벡터로의 어셈블리를 용이하게 하기 위해 본 발명에 제공된 바와 같은 벡터 오버랩 서열을 추가로 포함할 수 있다. 벡터는 복제 플라스미드일 수 있다.

[0104] 유전자 편집

[0105] 본 발명에 기술된 바와 같이, 제 1 및 제 2 폴리뉴클레오타이드의 쌍뿐만 아니라 삽입체 폴리뉴클레오타이드를 포함하는 조성물은 그 사이에 삽입체 폴리뉴클레오타이드가 있는 제 1 및 제 2 폴리뉴클레오타이드를 포함하는 핵산 라이브러리로 어셈블될 수 있으며, 이는 후속적으로 숙주 세포의 유전적 함량을 변형하는 데 사용될 수 있다. 본 발명에 제공된 바와 같이, 핵산 라이브러리는 제어 요소(예를 들어, 프로모터, 종결자, 용해도 태그, 분해 태그 또는 디그론), 변형된 형태의 유전자(예를 들어, 원하는 SNP를 갖는 유전자), 안티센스 핵산, 및/또는 대사 또는 생화학적 경로의 일부인 하나 이상의 유전자를 포함할 수 있다. 한 실시태양에서, 변형은 숙주 세포의 유전자 편집을 수반한다. 유전자 편집은 숙주 세포의 게놈 및/또는 숙주 세포에 존재하는 별도의 유전 요소, 예를 들어 플라스미드 또는 코스미드를 편집하는 것을 수반할 수 있다. 본 발명에 제공된 방법 및 조성물을 사용하여 생성된 핵산 어셈블리를 이용할 수 있는 유전자 편집 방법은 당업계에 공지된 임의의 유전자 편집 방법 또는 시스템일 수 있으며, 유전자 편집을 원하는 숙주를 기준으로 선택할 수 있다. 유전자 편집의 비 제한적인 예는 상동성 재조합, CRISPR, TALENS, FOK 또는 기타 엔도뉴클레아제를 포함한다.

[0106] 상동성 재조합

[0107] 한 실시태양에서, 유전자 편집 방법은 당업계에 공지된 상동성 재조합 기반 방법이다. 상동성 재조합 기반 방법은 단일 교차 상동성 재조합, 이중 교차 상동성 재조합 또는 람다 레드 재조합 중에서 선택될 수 있다. 이 실시태양에 추가로, 제 1 및 제 2 폴리뉴클레오타이드의 쌍에서 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 각각 숙주의 핵산 요소(예를 들어, 게놈, 플라스미드 또는 코스미드)에서 원하는 유전자좌에 지시되거나 상보적인 서열을 포함하여 이에 의해 그 사이에 위치한 삽입체 폴리뉴클레오타이드를 숙주 세포의 유전자 요소(예를 들어, 게놈, 코스미드 또는 플라스미드)의 원하는 유전자좌로 유도한다. 따라서, 쌍에 존재하는 원하는 유전자좌에 대해 지시되거나 상보적인 서열은 편집을 위해 표적화될 게놈, 코스미드 또는 플라스미드에서의 위치(들)를 결정하기 위해 사용될 수 있다. 도 1에 예시된 바와 같이, 원하는 유전자좌에 대한 또는 상보적인 서열은 제 1 폴리뉴클레오타이드의 근위 또는 5' 말단에 또는 향해 위치될 수 있는 반면, 제 2 폴리뉴클레오타이드에서 원하는 유전자좌에 대한 또는 상보적 서열은 원위 또는 3' 말단에 또는 근처에 위치될 수 있다. 제 1 폴리뉴클레오타이드에서, 원하는 유전자좌에 대해 지시되거나 상보적인 서열은 제 1 폴리뉴클레오타이드에 존재하는 어셈블리 오버랩 서열의 상류 및 존재하는 경우 벡터 오버랩 서열의 하류에 위치될 수 있다. 제 2 폴리뉴클레오타이드에서, 원하는 유전자좌에 대해 지시되거나 상보적인 서열은 제 2 폴리뉴클레오타이드에 존재하는 어셈블리 오버랩 서열의 하류 및 존재하는 경우 벡터 오버랩 서열의 상류에 위치될 수 있다.

[0108] 한 실시태양에서, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드의 쌍을 포함하는 풀의 각 쌍에 대해, 한 쌍의 원하는 유전자좌에 상보적인 서열은 상기 풀의 서로의 쌍에 비해 숙주 세포에서 상이한 표적 유전자좌에 상보적이다.

[0109] 다른 실시태양에서, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드 쌍을 함유하는 풀의 각 쌍에 대해, 한 쌍의 원하는 유전자좌에 상보적인 서열은 상기 풀의 서로의 쌍에 비해 숙주 세포에서 동일한 표적 유전자좌에 상보적이다.

[0110] 루프-인/루프-아웃(Loop-in/Loop-out)

[0111] 일부 실시태양에서, 본 발명은 숙주 유기체로부터 선택된 DNA 영역을 루핑 아웃하는 방법을 교시한다. 루핑 아웃 방법은 Nakashima et al. 2014 "Bacterial Cellular Engineering by Genome Editing and Gene Silencing."

Int. J. Mol. Sci. 15(2), 2773-2793에 기술된 것일 수 있다. 루핑 아웃 결실 방법은 당업계에 공지되어 있고 (Tear et al. 2014 "Excision of Unstable Artificial Gene-Specific inverted Repeats Mediates Scar-Free Gene Deletions in Escherichia coli." Appl. Biochem. Biotech. 175:1858-1867)에 기술되어 있다. 본 발명에 제공된 방법에 사용된 루핑 아웃 방법은 단일 교차 상동성 재조합 또는 이중 교차 상동성 재조합을 사용하여 수행될 수 있다. 한 실시태양에서, 선택된 영역의 루핑 아웃은 단일 교차 상동성 재조합을 사용하는 것을 수반할 수 있다.

[0112] 한 실시태양에서, 본 발명에 제공된 조성물은 제 1 및 제 2 폴리뉴클레오타이드의 쌍(예를 들어, 왼쪽/오른쪽 상동성 압), 삽입체 폴리뉴클레오타이드 및 벡터를 포함하여 본 발명에서 제공되는 시험관 내 또는 생체 내 어셈블리 방법을 사용하는 제 1 및 제 2 폴리뉴클레오타이드 쌍과 삽입체 폴리뉴클레오타이드와 벡터의 어셈블리가 루프 아웃 벡터를 생성한다. 한 실시태양에서, 상기 벡터를 루프-인하기 위해 루프-아웃 벡터와 숙주 세포 계놈 사이에 단일 교차 상동성 재조합이 사용된다. 벡터는 루프-인 단계 후 루프-아웃된 클론의 선택을 용이하게 하는 마커를 포함할 수 있다. 다른 실시태양에서, 상기 벡터를 통합하기 위해 루프-아웃 벡터와 숙주 세포 계놈 사이에 이중 교차 상동성 재조합이 사용된다. 루프-아웃 벡터 내의 삽입 서열은 기존 또는 도입된 인접 숙주 서열의 직접 반복체인 서열로 디자인될 수 있어서, 직접 반복체는 루프 및 결실을 위해 예정된 DNA 영역에 플랭킹한다. 삽입체 서열은 루프 아웃된 클론의 선택을 용이하게 하는 마커를 추가로 포함할 수 있다. 일단 삽입되면, 루프 아웃 플라스미드 또는 벡터를 포함하는 세포는 선택 영역의 삭제에 카운터 선택될 수 있다.

[0113] 본 발명에 제공된 한 양태에서, 본 발명에 제공된 조성물 및/또는 방법을 사용하여 생성된 폴리뉴클레오타이드 또는 폴리뉴클레오타이드 라이브러리는 하나 이상의 재조합 시스템으로부터의 단백질 세트의 사용을 수반할 수 있는 유전자 편집 방법에 사용될 수 있다. 상기 재조합 시스템은 미생물 숙주 세포에 내인성일 수 있거나 이중적으로 도입될 수 있다. 하나 이상의 이중 재조합 시스템의 단백질 세트는 핵산(예를 들어, 플라스미드, 선형 DNA 또는 RNA 또는 인테그론)으로 도입되어 숙주 세포의 계놈에 통합되거나 염색체 외 요소에서 안정적으로 발현될 수 있다. 하나 이상의 이중 재조합 시스템의 단백질 세트는 RNA로 도입되어 숙주 세포에 의해 번역될 수 있다. 하나 이상의 이중 재조합 시스템의 단백질 세트는 숙주 세포에 단백질로 도입될 수 있다. 하나 이상의 재조합 시스템의 단백질 세트는 람다 레드 재조합 시스템, RecET 재조합 시스템, Red/ET 재조합 시스템, 람다 레드 재조합 시스템으로부터의 단백질의 임의의 상동체, 오솔로그 또는 파라로그, RecET 재조합 시스템, 또는 Red/ET 재조합 시스템 또는 이들의 조합에서 유래될 수 있다. RecET 재조합 시스템의 재조합 방법 및/또는 단백질 세트는 Zhang Y., Buchholz F., Muyrers J.P.P. and Stewart A.F. "A new logic for DNA engineering using recombination in E. coli." Nature Genetics 20 (1998) 123-128; Muyrers, J.P.P., Zhang, Y., Testa, G., Stewart, A.F. "Rapid modification of bacterial artificial chromosomes by ET-recombination." Nucleic Acids Res. 27 (1999) 1555-1557; Zhang Y., Muyrers J.P.P., Testa G. and Stewart A.F. "DNA cloning by homologous recombination in E. coli." Nature Biotechnology 18 (2000) 1314-1317 and Muyrers JP et al., "Techniques: Recombinogenic engineering--new options for cloning and manipulating DNA" Trends Biochem Sci. 2001 May;26(5):325-31에 기술된 것 중 임의의 것일 수 있으며, 이는 본 발명에 참조로 포함된다. Red/ET 재조합 시스템의 단백질의 세트는 Rivero-Muller, Adolfo et al. "Assisted large fragment insertion by Red/ET-recombination (ALFIRE)--an alternative and enhanced method for large fragment recombinering" Nucleic acids research vol. 35,10 (2007): e78에 기술된 것 중 임의의 것일 수 있으며, 이는 본 발명에 참조로 포함된다.

[0114] 람다 레드 매개 유전자 편집

[0115] 본 발명에 제공된 바와 같이, 본 발명에 기재된 바와 같은 유전자 편집은 Datsenko and Wanner, PNAS USA 97:6640-6645 (2000)에 기재된 바와 같은 람다 레드-매개 상동성 재조합을 사용하여 수행될 수 있으며, 이의 내용은 전체로 본 발명에 참고로 포함된다.

[0116] 표적 DNA를 변형하기 위해 람다 레드 재조합 시스템을 사용하기 위해, 선형 공여자 DNA 기질(dsDNA 또는 ssDNA)을 람다 레드 재조합 시스템으로부터 단백질 세트를 발현하는 대장균으로 전기 천공할 수 있다. 람다 레드 재조합 시스템으로부터의 단백질 세트는 엑소, 베타 또는 감마 단백질 또는 이들의 임의의 조합을 포함할 수 있다. Gam은 내인성 RecBCD 및 SbcCD 뉴클레아제가 미생물 숙주 세포에 도입된 선형 공여자 DNA(dsDNA 또는 ssDNA)를 분해하는 것을 방지할 수 있는 반면, exo는 5' 말단에서 시작하여 선형 dsDNA를 분해할 수 있고 2개의 가능한 산물(즉, 단일 가닥 3' 오버행이 있는 부분 dsDNA 듀플렉스 또는 전체 상보적 가닥이 분해된 ssDNA)을 생성할 수 있는 5' → 3' dsDNA 의존성 엑소뉴클레아제이며 베타는 Exo에서 생성된 ssDNA를 보호하고 세포에서 상보적 ssDNA 표적에 대한 어닐링을 촉진할 수 있다. 베타 발현은 <https://blog.addgene.org/lambda-red-a->

homologous-recombination-based-technique-for-genetic-engineering에 기술된 대로 ssDNA 올리고 기질과의 람다 레드 기반 재조합에 필요할 수 있으며, 이는 참조로 본 발명에 포함된다.

- [0117] 선형 공여자 DNA 기질(dsDNA 또는 ssDNA)은 본 발명에 제공된 방법 및 조성을 사용하여 생성된 삽입체 폴리뉴클레오타이드가 그 사이에 위치하는 한 쌍의 제 1 및 제 2 폴리뉴클레오타이드를 포함하는 어셈블리일 수 있다. 한 쌍의 제 1 및 제 2 폴리뉴클레오타이드는 숙주 세포의 게놈 내의 특정 유전자좌에 대해 상기 공여자 DNA 기질을 표적화하는 게놈 표적화 서열을 포함할 수 있다. 이어서, 이들 효소는 표적 DNA 서열과 기질의 상동성 재조합을 촉매한다. 이것은 시험관에서 유전적 변화가 발생하는 제한 효소 클로닝과 비교하여 생체 내에서 클로닝이 일어난다는 것을 의미한다. 공여자 DNA 기질은 재조합을 위해 표적 부위에 상동성의 ~50개의 뉴클레오타이드만 필요로 한다. <https://blog.addgene.org/lambda-red-a-homologous-recombination-based-technique-for-genetic-engineering>에 기술된 대로 선형 dsDNA 또는 ssDNA 기질이 사용되는지 여부는 실험의 목표에 따라 달라질 수 있다. dsDNA 기질은 약 20개 뉴클레오타이드보다 큰 삽입 또는 결실에 가장 적합할 수 있는 반면, ssDNA 기질은 점 돌연변이 또는 소수 염기 쌍의 변화에만 가장 적합할 수 있다.
- [0118] dsDNA 기질은 제 1 및 제 2 폴리뉴클레오타이드의 쌍이 대향 말단에서 표적 삽입 부위에 대해 약 50 염기쌍의 상동성을 포함하도록 본 발명에 제공된 조성물 및 방법을 사용하여 제조될 수 있다. 기질 내에 존재하는 dsDNA 삽입체 폴리뉴클레오타이드는 항생제 내성 유전자와 같은 선택 가능한 DNA 단편은 물론 유전자 대체 및 태그와 같은 비 선택적 DNA 단편을 포함하는 큰 삽입 또는 결실을 포함할 수 있다.
- [0119] ssDNA 기질은 또한 dsDNA 기질은 제 1 및 제 2 폴리뉴클레오타이드의 쌍이 대향 말단에서 표적 삽입 부위에 대해 약 50 염기쌍의 상동성을 포함하도록 본 발명에 제공된 조성물 및 방법을 사용하여 제조될 수 있으며 서열의 중앙(즉, 삽입체 폴리뉴클레오타이드 내)에 위치한 원하는 변경(들)을 가질 수 있다.
- [0120] ssDNA 기질은 0.1% 내지 1%의 재조합 빈도로 dsDNA보다 더 효율적일 수 있고, 메틸-지시된 미스매치 복구(MMR) 시스템의 활성화를 피하는 기질을 디자인함으로써 25-50%까지 증가될 수 있다. MMR의 임무는 DNA 복제 중에 발생하는 DNA 미스매치를 수정하는 것이다. MMR의 활성화는 다음에 의해 피할 수 있다: 1) 주요 MMR 단백질이 낙아웃된 박테리아 균주를 사용, 또는 2) MMR을 피하기 위해 ssDNA 기질을 특수하게 디자인함: 1) MMR이 비활성화된 대장균: MMR이 비활성화된 대장균을 사용하는 것이 분명히 두 가지 옵션 중 더 쉬운 방법이지만, 이들 세포는 돌연변이가 발생하기 쉽고 의도하지 않은 게놈에 변화를 줄 수 있다. 2) MMR 활성화를 피하는 ssDNA 기질의 디자인: 일 실시태양에서, 편집 부위의 6개 염기쌍 또는 그 내의 C/C 미스매치가 도입된다. 다른 실시태양에서, 원하는 변화는 위블 코돈에서 4-5 침묵 변화(즉, 번역된 단백질의 아미노산 서열은 아니지만 뉴클레오타이드 서열을 변경하는 인접한 4-5 코돈의 제 3 염기쌍의 변화)에 플랭킹한다. 이러한 변경은 원하는 변경의 5' 또는 3'일 수 있다.
- [0121] 한 실시태양에서, 본 발명에 제공된 조성물 및/또는 방법을 사용하여 생성된 폴리뉴클레오타이드 또는 폴리뉴클레오타이드 라이브러리는 <https://blog.addgene.org/lambda-red-a-homologous-recombination-based-technique-for-genetic-engineering>에 기술된 DY380과 같은 람다 레드 재조합 유전자를 이미 안정적으로 발현하는 미생물 숙주 세포에서 구현되는 유전자 편집 방법에 사용될 수 있으며, 이의 내용은 본 발명에 참조로 포함된다. 람다 레드 재조합 시스템의 성분을 포함하고 본 발명에 제공된 농축 방법(예를 들어, CS-seq 또는 SG-seq)을 사용하여 유전자형화된 유기체를 생성하는 데 사용될 수 있는 다른 박테리아 균주는 Thomason et al (Recombineering: Genetic Engineering in Bacteria Using Homologous Recombination. Current Protocols in Molecular Biology. 106:V:1.16:1.16.1-1.16.39) and Sharan et al (Recombineering: A Homologous Recombination-Based Method of Genetic Engineering. Nature protocols. 2009;4(2):206-223)에서 발견될 수 있고, 이의 각각의 내용은 본 발명에 참조로 포함된다.
- [0122] 본 발명에 제공된 바와 같이, 람다 레드 재조합 시스템의 단백질 세트는 당 업계에 공지되고/되거나 본 발명에 제공된 임의의 편집 방법의 구현 전에 미생물 숙주 세포로 도입될 수 있다. 람다 레드 재조합 시스템의 각 단백질에 대한 유전자는 핵산(예를 들어, 플라스미드, 선형 DNA 또는 RNA, 미니-λ, 람다 레드 프로파지 또는 인테그론)에 도입되어 숙주의 게놈에 통합되거나 세포 또는 염색체 외 요소에서 발현될 수 있다. 일부 경우에, 람다 레드 재조합 시스템의 각 성분(즉, 엑소, 베타, gam 또는 이들의 조합)은 RNA로 도입되어 숙주 세포에 의해 번역될 수 있습니다. 일부 경우에, 람다 레드 재조합 시스템의 각 성분(즉, 엑소, 베타, gam 또는 이들의 조합)은 숙주 세포에 단백질로 도입될 수 있다.
- [0123] 한 실시태양에서, 람다 레드 재조합 시스템의 단백질 세트에 대한 유전자는 플라스미드 상에 도입된다. 플라스미드 상의 람다 레드 재조합 시스템의 단백질 세트는, 예를 들어, 내인성 파지 pL 프로모터와 같은 프로모터의

제어하에 있을 수 있다. 한 실시태양에서, 플라스미드 상의 람다 레드 재조합 시스템의 단백질 세트는 유도성 프로모터의 제어하에 있다. 유도성 프로모터는 시약의 첨가 또는 고갈 또는 온도 변화에 의해 유도될 수 있다. 한 실시태양에서, 플라스미드 상의 람다 레드 재조합 시스템의 단백질 세트는 IPTG-유도성 lac 프로모터 또는 아라비 노스-유도성 pBAD 프로모터와 같은 유도성 프로모터의 제어하에 있다. 람다 레드 재조합 시스템의 단백질 세트에 대한 유전자를 발현하는 플라스미드는 또한 각각, 예를 들어, IPTG-유도성 lac 프로모터, 아라비노스-유도성 pBAD 프로모터 및 내인성 파지 pL 프로모터와 관련된 lacI, araC 또는 cI857 억제자와 같은 특정 프로모터와 관련된 억제자를 발현할 수 있다.

[0124] 한 실시태양에서, 람다 레드 재조합 시스템의 단백질 세트에 대한 유전자는 미니-λ 상에 도입되며, 이는 미생물 숙주 세포에 도입될 때 <https://blog.addgene.org/lambda-red-a-homologous-recombination-based-technique-for-genetic-engineering>에 기술된 바와 같이 계놈 속에 통합되는 결함이 있는 복제되지 않는 원형 파지 DNA 조각이며, 이의 내용은 본 발명에 참조로 포함된다.

[0125] 한 실시태양에서, 람다 레드 재조합 시스템의 단백질 세트에 대한 유전자는 람다 레드 프로 파지에 도입되며, 이는 <https://blog.addgene.org/lambda-red-a-homologous-recombination-based-technique-for-genetic-engineering>에 기술된 바와 같이 람다 레드 재조합 시스템을 미생물 숙주 세포로 안정적으로 통합할 수 있게 하며, 이의 내용은 본 발명에 참조로 포함된다.

[0126] CRISPR 매개 유전자 편집

[0127] 본 발명에 제공된 한 양태에서, 숙주 세포의 유전자 요소(예를 들어, 계놈, 코스미드, 또는 플라스미드)는 CRISPR에 의해 변형될 수 있다.

[0128] CRISPR/Cas 시스템은 플라스미드 및 파지 내에 존재하는 것과 같은 외래 유전자 요소에 대한 내성을 부여하고 획득된 면역의 형태를 제공하는 원핵 면역 시스템이다. CRISPR은 Clustered Regularly Interspaced Short Palindromic Repeat을 나타내며, cas는 CRISPR-관련 시스템(CRISPR-associated system)을 나타내며, CRISPR 복합체와 관련된 작은 cas 유전자를 지칭한다.

[0129] CRISPR-Cas 시스템은 클래스 1 또는 클래스 2 시스템으로 가장 광범위하게 특징지어진다. 이 두 시스템의 주요 특징은 Cas-효과기 모듈의 특성이다. 클래스 1 시스템은 간섭을 중재하기 위해 복합체("캐스케이드 복합체"로 지칭됨)에 다수의 Cas 단백질의 조립을 필요로 하는 반면, 클래스 2 시스템은 간섭을 중재하기 위해 큰 단일 Cas 효소를 사용한다. 클래스 1 및 클래스 2 시스템 각각은 특정 Cas 단백질의 존재에 기초하여 다수의 CRISPR-Cas 유형으로 추가로 분할된다. 예를 들어, 클래스 1 시스템은 다음 세 가지 유형으로 나뉜다: Cas3 단백질을 포함하는 타입 I 시스템; Cas 10 단백질을 함유하는 타입 III 시스템; 및 Cas8-유사 단백질을 포함하는 추정 타입 IV 시스템. 클래스 2 시스템은 일반적으로 클래스 1 시스템보다 덜 일반적이며 다음 세 가지 유형으로 더 나뉜다: Cas9 단백질을 포함하는 타입 II 시스템; Cas 12a 단백질(이전에 Cpf1로 알려짐, 본 발명에서 Cpf1이라고 함), Cas 12b(이전에 C2c1로 알려짐), Cas 12c(이전에 C2c3로 알려짐), Cas 12d(이전에 CasY로 알려짐) 및 Cas12e(이전에 CasX로 알려짐)를 함유하는 타입 V 시스템; 및 Cas13a(이전에 C2c2로 알려짐), Cas13b 및 Cas13c를 함유하는 타입 VI 시스템. Pyzocha *et al.*, ACS Chemical Biology, Vol. 13 (2), pgs. 347-356. 일 실시태양에서, 본 발명에 제공된 방법에 사용하기 위한 CRISPR-Cas 시스템은 클래스 2 시스템이다. 일 실시태양에서, 본 발명에 제공된 방법에 사용하기 위한 CRISPR-Cas 시스템은 타입 II, 타입 V 또는 타입 VI 클래스 2 시스템이다. 일 실시태양에서, 본 발명에 제공된 방법에 사용하기 위한 CRISPR-Cas 시스템은 Cas9, Cas12a, Cas12b, Cas12c, Cas12d, Cas12e, Cas13a, Cas13b, Cas13c 또는 이의 상 동체, 오쏘로그 또는 파라로그로부터 선택된다.

[0130] 본 발명에 개시된 방법에 사용된 CRISPR 시스템은 본 발명에서 Cas 효과기 단백질로 지칭되는 하나 이상의 핵산 유도된 CRISPR-관련(Cas) 뉴클레아제를 포함하는 Cas 효과기 모듈을 포함한다. 일부 실시태양에서, Cas 단백질은 하나 또는 다수의 뉴클레아제 도메인을 포함할 수 있다. Cas 효과기 단백질은 단일 가닥 또는 이중 가닥 핵산 분자(예를 들어, DNA 또는 RNA 핵산)를 표적으로 할 수 있고 이중 가닥 또는 단일 가닥 브레이크를 생성할 수 있다. 일부 실시태양에서, Cas 효과기 단백질은 야생형 또는 자연 발생 Cas 단백질이다. 일부 실시태양에서, Cas 효과기 단백질은 돌연변이체 Cas 단백질이며, 여기서 하나 이상의 돌연변이, 삽입 또는 결실은 WT 또는 자연 발생 Cas 단백질(예를 들어, 부모 Cas 단백질)에서 만들어져 부모 Cas 단백질과 비교하여 하나 이상의 변경된 특성을 갖는 Cas 단백질을 생성한다.

[0131] 일부 경우에, Cas 단백질은 야생형(WT) 뉴클레아제이다. 본 발명에 사용하기에 적합한 Cas 단백질의 비-제한적

인 예는 C2c1, C2c2, C2c3, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (Csn1 및 Csx12로도 알려짐), Cas10, Cpf1, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm1, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx100, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, MAD1-20, SmCsm1, 이의 상동체, 이의 오쏘로그, 이의 변이체, 이의 돌연변이체, 또는 이의 변형된 형태를 포함한다. 적합한 핵산 유도 뉴클레아제(예를 들어, Cas 9)는 다음을 포함하지만 이에 제한되지 않는 속으로부터의 유기체로부터 유래될 수 있다: *티오미크로스피라(Thiomicrospira)*, *숙시니비브리오(Succinivibrio)*, *칸디다투스(Candidatus)*, *포르피로모나스(Porphyrmonas)*, *애시도모노코쿠스(Acidomonococcus)*, *프레보텔라(Prevotella)*, *스미텔라(Smithella)*, *모락셀라(Moraxella)*, *시너지스테스(Synergistes)*, *프란시셀라(Francisella)*, *렙토스피라(Leptospira)*, *카테니박테리움(Catenibacterium)*, *칸드레리아(Kandleria)*, *클로스트리디움(Clostridium)*, *도레아(Dorea)*, *코프로코쿠스(Coproccoccus)*, *엔테로코쿠스(Enterococcus)*, *프루토바실러스(Fructobacillus)*, *웨이셀라(Weissella)*, *페디오코쿠스(Pediococcus)*, *코리네박터(Corynebacter)*, *슈터렐라(Sutterella)*, *레기오넬라(Legionella)*, *트레포네마(Treponema)*, *로세부리아(Roseburia)*, *필리팩터(Filifactor)*, *유박테리움(Eubacterium)*, *스트렙토코쿠스(Streptococcus)*, *락토바실러스(Lactobacillus)*, *마이코플라즈마(Mycoplasma)*, *박테로이데스(Bacteroides)*, *플라비볼라(Flaviivola)*, *플라보박테리움(Flavobacterium)*, *스파에로카에타(Sphaerochaeta)*, *아조스피릴룸(Azospirillum)*, *글루코나세토박터(Gluconacetobacter)*, *나이세리아(Neisseria)*, *로세부리아(Roseburia)*, *파비바쿨럼(Parvibaculum)*, *스타필로코쿠스(Staphylococcus)*, *니트라티프랙터(Nitratifactor)*, *마이코플라즈마(Mycoplasma)*, *알리시클로바실러스(Alicyclobacillus)*, *브레비바실러스(Brevibacillus)*, *바실러스(Bacillus)*, *박테로이데테스(Bacteroidetes)*, *브레비바실러스(Brevibacillus)*, *카르노박테리움(Carnobacterium)*, *클로스트리디아리디움(Clostridiaridium)*, *클로스트리디움(Clostridium)*, *데술포나트로늄(Desulfonatronum)*, *데술포비브리오(Desulfovibrio)*, *헬코쿠스(Helcococcus)*, *렙토티리키아(Leptotrichia)*, *리스테리아(Listeria)*, *메타노메티오피루스(Methanomethyophilus)*, *메틸로박테리움(Methylobacterium)*, *오피튜타세에(Opitutaceae)*, *팔루디박터(Paludibacter)*, *로도박터(Rhodobacter)*, *스파에로카에타(Sphaerochaeta)*, *튜베리바실러스(Tuberibacillus)*, 및 *캠필로박터(Campylobacter)*. 이러한 속의 유기체 종은 본 발명에서 달리 논의된 바와 같을 수 있다.

[0132] 적합한 핵산 유도된 뉴클레아제(예를 들어, Cas9)는 문(phylum)으로부터의 유기체로부터 유래될 수 있으며, 이는 다음을 포함하지만 이에 제한되지 않는다: 피르미큐트, 악티노박테리아, 박테로이데테스, 프로테오박테리아, 스피로카테스 및 테네리큐테스. 적합한 핵산 유도된 뉴클레아제는 에리시펠로트리키아, 클로스트리디아, 바실리, 악티노박테리아, 박테로이데테스, 플라보박테리아, 알파프로테오박테리아, 베타프로테오박테리아, 감마프로테오박테리아, 델타프로테오박테리아, 입실론프로테오박테리아, 스피로카에테스 및 몰리큐테스. 적합한 핵산 유도된 뉴클레아제는 다음을 포함하지만 이에 제한되지 않는다: 클로스트리디알레스, 락토바실라레스, 악티노마이세탈레스, 박테로이달레스, 플라보박테리알레스, 리조비알레스, 로도스피릴라레스, 버크홀데리알레스, 니세리알레스, 레지오넬라레스, 노틸리알레스, 캄피로박테라레스, 스피로카에탈레스, 마이코플라스마탈레스 및 티오티리칼레스. 적합한 핵산 유도된 뉴클레아제는 다음을 포함하지만 이에 제한되지 않는 패밀리의 유기체로부터 유래될 수 있다: 라크노스피라세에, 엔테로코카세에, 유코노스토카세에, 락토바실라세에, 스트렙토코카세에, 펩토스트렙토코카세에, 스타필로코카세에, 유박테리아세에, 코리네박테리네에, 박테로이다세에, 플라보박테리움, 크리오무어파세에, 로도비아세에, 로도스피릴라세에, 아세토박테라세에, 슈터렐라세에, 나이세리아세에, 레기오넬라세에, 노틸리아세에, 캄필로박테라세에, 스피로카에타세에, 마이코플라스마타세에 및 프란시셀라세에.

[0133] 본 발명의 방법, 시스템 및 조성물에 사용하기에 적합한 다른 핵산 유도 뉴클레아제(예를 들어, Cas9)는 다음과 같은 유기체로부터 유래된 것들을 포함하지만, 이에 제한되지는 않는다: *Thiomicrospira sp. XS5*, *Eubacterium rectale*, *Succinivibrio dextrinosolvens*, *Candidatus Methanoplasma termitum*, *Candidatus Methanomethylophilus alvus*, *Porphyromonas crevioricanis*, *Flavobacterium branchiophilum*, *Acidomonococcus sp.*, *Lachnospiraceae bacterium COE1*, *Prevotella brevis ATCC 19188*, *Smithella sp. SCADC*, *Moraxella bovoculi*, *Synergistes jonesii*, *Bacteroidetes oral taxon 274*, *Francisella tularensis*, *Leptospira inadai serovar Lyme str. 10*, *Acidomonococcus sp. crystal structure (5B43)* *S. mutans*, *S. agalactiae*, *S. equisimilis*, *S. sanguinis*, *S. pneumoniae*; *C. jejuni*, *C. coli*; *N. salsuginis*, *N. tergaricus*; *S. auricularis*, *S. carnosus*; *N. meningitidis*, *N. gonorrhoeae*; *L. monocytogenes*, *L. ivanovii*; *C. botulinum*, *C. difficile*, *C. tetani*, *C. sordellii*; *Francisella tularensis 1*, *Prevotella albensis*, *Lachnospiraceae bacterium MC2017 1*, *Butyrivibrio proteoclasticus*, *Peregrinibacteria bacterium GW2011_GWA2_33_10*, *Parcubacteria bacterium GW2011_GWC2_44_17*, *Smithella sp. SCADC*,

Microgenomates, *Acidaminococcus* sp. BV3L6, *Lachnospiraceae* bacterium MA2020, *Candidatus Methanoplasma* *termitum*, *Eubacterium eligens*, *Moraxella bovoculi* 237, *Leptospira inadai*, *Lachnospiraceae* bacterium ND2006, *Porphyromonas crevioricanis* 3, *Prevotella disiens*, *Porphyromonas macacae*, *Catenibacterium* sp. CAG:290, *Kandleria vitulina*, Clostridiales bacterium KA00274, *Lachnospiraceae* bacterium 3-2, *Dorea longicatena*, *Coprococcus catus* GD/7, *Enterococcus columbae* DSM 7374, *Fructobacillus* sp. EFB-N1, *Weissella halotolerans*, *Pediococcus acidilactici*, *Lactobacillus curvatus*, *Streptococcus pyogenes*, *Lactobacillus versmoldensis*, 및 *Filifactor alocis* ATCC 35896. U.S. Pat. Nos. 8,697,359; 8,771,945; 8,795,965; 8,865,406; 8,871,445; 8,889,356; 8,895,308; 8,906,616; 8,932,814; 8,945,839; 8,993,233; 8,999,641; 9,822,372; 9,840,713; U.S. Pat. App. No. 13/842,859 (US 2014/0068797 A1); 9,260,723; 9,023,649; 9,834,791; 9,637,739; U.S. Pat. App. No. 14/683,443 (US 2015/0240261 A1); U.S. Pat. App. No. 14/743,764 (US 2015/0291961 A1); 9,790,490; 9,688,972; 9,580,701; 9,745,562; 9,816,081; 9,677,090; 9,738,687; U.S. App. No. 15/632,222 (US 2017/0369879 A1); U.S. App. No. 15/631,989; U.S. App. No. 15/632,001; 및 U.S. Pat. No. 9,896,696 참조, 이들 각각은 본 발명에 참조로 포함된다.

- [0134] 일부 실시태양에서, Cas 효과기 단백질은 다음 활성 중 하나 이상을 포함한다:
- [0135] 니카제 활성, 즉 핵산 분자의 단일 가닥을 절단하는 능력;
- [0136] 이중 가닥 뉴클레아제 활성, 즉 이중 가닥 핵산의 두 가닥을 절단하고 이중 가닥 브레이크를 생성하는 능력;
- [0137] 엔도뉴클레아제 활성;
- [0138] 엑소뉴클레아제 활성; 및/또는
- [0139] 헬리카제 활성, 즉 이중 가닥 핵산의 나선 구조를 풀 수 있는 능력.
- [0140] 본 발명의 양태에서, 용어 "가이드 핵산"은 1) 표적 서열(본 발명에서 "표적 세그먼트"로 지칭됨)에 혼성화할 수 있는 가이드 서열 및 2) 본 발명에 기재된 바와 같은 핵산 유도 뉴클레아제(본 발명에서 "스캐폴드 세그먼트"로 지칭됨)와 (단독으로 또는 tracrRNA 분자와의 조합으로) 상호작용 가능한 스캐폴드 서열을 포함하는 폴리뉴클레오타이드를 지칭한다. 가이드 핵산은 DNA일 수 있다. 가이드 핵산은 RNA일 수 있다. 가이드 핵산은 DNA 및 RNA 모두를 포함할 수 있다. 가이드 핵산은 변형된 비-천연 발생 뉴클레오타이드를 포함할 수 있다. 가이드 핵산이 RNA를 포함하는 경우, RNA 가이드 핵산은 본 발명에 제공된 방법과 조성물을 사용하여 생성된 플라즈미드, 선형 구조체 또는 편집 카세트와 같은 폴리뉴클레오타이드 분자상의 DNA 서열에 의해 코딩될 수 있다.
- [0141] 일부 실시태양에서, 본 발명에 기재된 가이드 핵산은 RNA 가이드 핵산("가이드 RNA" 또는 "gRNA")이고 표적 세그먼트 및 스캐폴드 세그먼트를 포함한다. 일부 실시태양에서, gRNA의 스캐폴드 세그먼트는 하나의 RNA 분자에 포함되고 표적 세그먼트는 다른 별도의 RNA 분자에 포함된다. 이러한 실시태양은 본 발명에서 "이중 분자 gRNA" 또는 "2분자 gRNA" 또는 "이중 gRNA"로 지칭된다. 일부 실시태양에서, gRNA는 단일 RNA 분자이고 본 발명에서는 "단일 가이드 RNA" 또는 "sgRNA"로 지칭된다. 용어 "가이드 RNA" 또는 "gRNA"는 2분자 가이드 RNA 및 sgRNA를 모두 지칭한다.
- [0142] 한 실시태양에서, 본 발명에 제공된 방법 및 조성물을 사용하여 생성된 삼입체 폴리뉴클레오타이드가 그 사이에 위치하는 한 쌍의 제 1 및 제 2 폴리뉴클레오타이드를 포함하는 어셈블리는 가이드 RNA(gRNA)이다. 일부 경우에, 본 발명에 제공된 방법은 gRNA의 라이브러리를 생성하는 데 사용된다.
- [0143] gRNA의 DNA-표적 세그먼트는 표적 핵산 서열의 서열에 상보적인 뉴클레오타이드 서열을 포함한다. 이와 같이, gRNA의 표적 세그먼트는 하이브리드화(즉, 염기쌍)를 통해 서열-특이적 방식으로 표적 핵산과 상호 작용하고, 표적 세그먼트의 뉴클레오타이드 서열은 gRNA가 결합할 표적 DNA 내의 위치를 결정한다. 적절한 정렬 알고리즘을 사용하여 최적으로 정렬될 때, 가이드 서열과 이에 상응하는 표적 서열 사이의 상보성 정도는 약 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99% 또는 그 이상이거나 그 보다 크다. 서열 정렬을 위한 임의의 적합한 알고리즘을 사용하여 최적 정렬이 결정될 수 있다. 일부 실시태양에서, 가이드 서열은 약 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 이상 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 45, 50, 75개 또는 그 이상의 뉴클레오타이드 길이이거나 그 보다 크다. 일부 실시태양에서, 가이드 서열은 길이가 약 75, 50, 45, 40, 35, 30, 25, 20개 미만의 뉴클레오타이드이다. 양태들에서, 가이드 서열은 10-30개의 뉴클레오타이드 길이이다. 가이드 서열은 15-20개의 뉴클레오타이드 길이일 수 있다. 가이드 서열은 15개의 뉴클레오타이드 길이일 수 있다. 가이드 서열은 16개의 뉴클레오타이드 길이일 수 있다. 가이드

드 서열은 17개의 뉴클레오타이드 길이일 수 있다. 가이드 서열은 18개의 뉴클레오타이드 길이일 수 있다. 가이드 서열은 19개의 뉴클레오타이드 길이일 수 있다. 가이드 서열은 20개의 뉴클레오타이드 길이일 수 있다.

- [0144] 가이드 RNA의 스캐폴드 세그먼트는 적어도 하나의 Cas 효과기 단백질과 상호 작용하여 리보핵산단백질 복합체 (본 발명에서 CRISPR-RNP 또는 RNP-복합체로 지칭됨)를 형성한다. 가이드 RNA는 상기 기재된 표적 세그먼트를 통해 결합된 폴리펩타이드를 표적 핵산 서열 내의 특정 뉴클레오타이드 서열로 유도한다. 가이드 RNA의 스캐폴드 세그먼트는 서로 상보적이고 이중 가닥 RNA 이중체를 형성하는 2개의 스트레치 뉴클레오타이드를 포함한다. 표적화 뉴클레아제 복합체의 형성을 촉진하기 위해 스캐폴드 서열 내에서 충분한 서열은 스캐폴드 서열 내에서 2개의 서열 영역의 길이를 따라 상보성 정도, 예를 들어 2차 구조를 형성하는 데 관여하는 1개 또는 2개의 서열 영역을 포함할 수 있다. 일부 경우에, 1개 또는 2개의 서열 영역은 동일한 폴리뉴클레오타이드 상에 포함되거나 코딩된다. 일부 경우에, 1개 또는 2개의 서열 영역은 별도의 폴리뉴클레오타이드 상에 포함되거나 코딩된다. 최적의 정렬은 임의의 적합한 정렬 알고리즘에 의해 결정될 수 있고, 1개 또는 2개의 서열 영역 내에서의 자기 상보성과 같은 2차 구조를 추가로 설명할 수 있다. 일부 실시태양에서, 최적으로 정렬될 때 2개 중 짧은 길이를 따라 1개 또는 2개의 서열 영역 사이의 상보성의 정도는 약 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97.5%, 99% 또는 그 이상이거나 그 보다 높다. 일부 실시태양에서, 2개의 서열 영역 중 적어도 하나는 약 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50개 또는 그 이상의 뉴클레오타이드 길이이다.
- [0145] 본 발명의 gRNA의 스캐폴드 서열은 2차 구조를 포함할 수 있다. 2차 구조는 유사매듭(pseudoknot) 영역 또는 스템-루프 구조를 포함할 수 있다. 일부 실시예에서, 가이드 핵산 및 핵산 유도된 뉴클레아제의 상용성은 가이드 RNA의 2차 구조 영역 내에서 또는 가이드 RNA의 인접 서열에 의해 적어도 부분적으로 결정된다. 일부 경우에, 가이드 핵산의 핵산 가이드된 뉴클레아제에 대한 결합 동역학은 스캐폴드 서열 내의 2차 구조에 의해 부분적으로 결정된다. 일부 경우에, 가이드 핵산의 핵산 유도 뉴클레아제에 대한 결합 동역학은 스캐폴드 서열을 갖는 핵산 서열에 의해 부분적으로 결정된다.
- [0146] gRNA-Cas 효과기 단백질 조합에 대한 적합한 스캐폴드 서열은 천연 Cas 뉴클레아제 유전자좌에 인접한 서열을 스캐닝함으로써 발견될 수 있다. 다시 말해, 천연 Cas 뉴클레아제는 상응하는 호환 가능한 가이드 핵산 또는 스캐폴드 서열에 근접한 계놈에서 코딩될 수 있다.
- [0147] 핵산 유도된 뉴클레아제는 뉴클레아제 내인성 숙주 내에서 발견되지 않은 가이드 핵산과 상용성일 수 있다. 이러한 직교 가이드 핵산은 경험적 테스트에 의해 결정될 수 있다. 직교 가이드 핵산은 상이한 박테리아 종으로부터 유래되거나 비-천연적으로 발생하도록 합성되거나 아니면 조작될 수 있다. 공통 핵산 안내 뉴클레아제와 양립할 수 있는 직교 가이드 핵산은 하나 이상의 공통 특징을 포함할 수 있다. 일반적인 특징은 유사매듭 영역 외부의 서열을 포함할 수 있다. 일반적인 특징은 유사매듭 영역을 포함할 수 있다. 일반적인 특징은 1차 서열 또는 2차 구조를 포함할 수 있다.
- [0148] 가이드 서열이 표적 서열에 상보적이도록 가이드 서열을 변경함으로써 원하는 표적 서열을 표적화하도록 가이드 핵산을 조작하여, 가이드 서열과 표적 서열 사이의 혼성화를 가능하게 한다. 조작된 가이드 서열을 갖는 가이드 핵산은 조작된 가이드 핵산으로 지칭될 수 있다. 조작된 가이드 핵산은 종종 자연적으로 발생하지 않으며 자연에서 발견되지 않는다.
- [0149] 일부 실시태양에서, 본 발명은 본 발명에 제공된 조성물과 방법을 사용하여 생성된 gRNA를 암호화하는 폴리뉴클레오타이드를 제공한다. 일부 실시태양에서, 한 쌍의 제 1 및 제 2 폴리뉴클레오타이드 및 삽입체 폴리뉴클레오타이드를 포함하는 조성물은 발현 벡터를 추가로 포함하여 삽입체 폴리뉴클레오타이드 및 발현 벡터와 제 1 및 제 2 폴리뉴클레오타이드 쌍의 어셈블리가 gRNA 암호화 핵산을 포함하는 발현 벡터를 생성하도록 한다.
- [0150] 다른 실시태양에서, 본 발명에 제공된 방법 및 조성물을 사용하여 생성된 삽입체 폴리뉴클레오타이드가 그 사이에 위치하는 한 쌍의 제 1 및 제 2 폴리뉴클레오타이드를 포함하는 어셈블리는 공여자 DNA 서열이다. 일부 경우에, 본 발명에 제공된 방법은 공여자 DNA 서열의 라이브러리를 생성하는 데 사용된다. 공여자 DNA 서열은 상동성 지정 복구(HDR)를 사용하는 유전자 편집의 CRISPR 방법에서 가이드 RNA(gRNA)와 조합하여 사용될 수 있다. CRISPR 복합체는 상동성 지정 복구(HDR)를 사용하여 복구할 수 있는 표적 유전자(들) 내에서 가닥 파손을 초래할 수 있다. HDR 매개 복구는 본 발명에 제공된 방법 및 조성물을 사용하여 생성된 공여자 DNA 서열로 숙주 세포를 공동 형질 전환함으로써 촉진될 수 있다. 공여자 DNA 서열은 원하는 유전자 교란(예를 들어, 결실, 삽입 및/또는 단일 뉴클레오타이드 다형성)뿐만 아니라 제 1 및 제 2 폴리뉴클레오타이드로부터 유래된 표적화 서열을 포함할 수 있다. 이 실시태양에서, CRISPR 복합체는 하나 이상의 gRNA에 의해 지정된 표적 유전자를 절단한

다. 공여자 DNA 서열은 원하는 유전자 삽동을 숙주 세포에 통합하기 위해 상동성 재조합 기구에 대한 주형으로 사용될 수 있다. 공여자 DNA는 단일 가닥, 이중 가닥 또는 이중 가닥 플라스미드일 수 있다. 공여자 DNA는 재절단을 방지하기 위해 PAM 서열이 없거나 스크램블, 변경 또는 비 기능적 PAM을 포함할 수 있다. 일부 경우에, 공여자 DNA는 기능적이거나 변경되지 않은 PAM 부위를 포함할 수 있다. 공여자 DNA의 돌연변이 또는 편집된 서열 (또한 상동성 영역에 플랭킹)은 돌연변이(들)가 게놈에 통합된 후 CRISPR-복합체에 의한 재절단을 방지한다.

[0151] **숙주 세포**

[0152] 본 발명에 제공된 바와 같이, 본 발명에 제공된 조성물 및/또는 방법을 사용하여 생성된 핵산 구조체의 라이브러리를 사용하여 숙주 세포의 유전 요소(예를 들어, 게놈, 코스미드 또는 플라스미드)를 편집 또는 변형하거나 하나 이상의 유전자 요소(들)(예를 들어, 플라스미드 또는 코스미드)를 상기 숙주 세포에 도입(예를 들어, 형질 전환 또는 형질도입)함으로써 숙주 세포를 조작하는 데 사용될 수 있다. 게놈 공학 또는 편집 방법은 유전자 돌연변이 집단에서 원하는 형질을 확인할 수 있는 모든 유기체에 적용할 수 있다. 유기체는 미생물 또는 고등 진핵 유기체일 수 있다.

[0153] 따라서, 본 발명에 사용된 용어 "미생물"은 광범위하게 해석되어야 한다. 미생물은 두 가지 원핵생물 영역인 박테리아와 고세균뿐만 아니라 특정 진핵생물 균류와 원생 생물을 포함한다. 그러나, 특정 양태에서, 곤충, 식물 및 동물과 같은 "보다 높은" 진핵생물이 본 발명에 교시된 방법에서 이용될 수 있다.

[0154] 적합한 숙주 세포는 박테리아 세포, 조류 세포, 식물 세포, 균류 세포, 곤충 세포 및 포유류 세포를 포함하나 이에 제한되지 않는다. 한 예시적인 실시태양에서, 적합한 숙주 세포는 대장균(예를 들어, 매사추세츠주, 입스위치의 뉴잉글랜드 바이오랩스(New England BioLabs)로부터 구입가능한 SHuffle™ competent E. coli)을 포함한다.

[0155] 본 발명의 다른 적합한 숙주 유기체는 코리네박테리움 속의 미생물을 포함한다. 일부 실시태양에서, 바람직한 코리네박테리움 균주/종은 기탁된 유형 균주가 DSM44549인 *C. 에피센스(C. efficiens)*, 기탁된 유형 균주가 ATCC13032인 *C. 글루타미쿰(C. glutamicum)* 및 기탁된 유형 균주가 ATCC6871인 *C. 암모니아게네스(C. ammoniagenes)*를 포함한다. 일부 실시태양에서, 본 발명의 바람직한 숙주는 *C. 글루타미쿰*이다.

[0156] 특히 코리네박테리움 글루타미쿰 종의 코리네박테리움 속의 적합한 숙주 균주는 특히 공지된 야생형 균주이다: 코리네박테리움 글루타미쿰 ATCC13032, 코리네박테리움 아세토글루타미쿰 ATCC15806, 코리네박테리움 아세토아시도필럼 ATCC13870, 코리네박테리움 멜라세콜라 ATCC17965, 코리네박테리움 씨모아미노게네스 FERM BP-1539, 브레비박테리움 플라봄 ATCC14067, 브레비박테리움 락토프레멘툼 ATCC13869 및 브레비박테리움 다이바리카툼 ATCC14020; 및 예를 들어, L-리신 생산 균주로부터 제조된 L-아미노산 생산 돌연변이체 또는 균주: 코리네박테리움 글루타미쿰 FERM-P 1709, 브레비박테리움 플라봄 FERM P-1708, 브레비박테리움 락토프레멘툼 FERM-P 1712, 코리네박테리움 글루타미쿰 FERM-P 6463, 코리네박테리움 글루타미쿰 FERM-P 6464, 코리네박테리움 글루타미쿰 DM58-1, 코리네박테리움 글루타미쿰 DG52-5, 코리네박테리움 글루타미쿰 DSM5714 및 코리네박테리움 글루타미쿰 DSM12866.

[0157] 용어 "마이크로코쿠스 글루타미쿠스"는 *C. 글루타미쿰*에도 사용되어 왔다. *C. 에피센스* 종의 일부 대표는 또한 예를 들어, 균주 FERM BP-1539와 같은 종래 기술에서 *C. 씨모아미노게네스*로 불려왔다.

[0158] 일부 실시태양에서, 본 발명의 숙주 세포는 진핵생물 세포이다. 적합한 진핵생물 숙주 세포는 곰팡이 세포, 조류 세포, 곤충 세포, 동물 세포 및 식물 세포를 포함하나 이에 제한되지 않는다. 적합한 균류 숙주 세포는 아스코코타(*Ascomycota*), 바시디오마이코타(*Basidiomycota*), 테우티로마이코타(*Deuteromycota*), 자이고마이코타(*Zygomycota*), 평기 임퍼펙티(*Fungi imperfecti*)를 포함하나 이에 제한되지 않는다. 특정 바람직한 숙주 세포는 효모 세포 및 사상 균류 세포를 포함한다. 적합한 사상 균류 숙주 세포는 예를 들어 유마이코티나(*Eumycotina*) 및 오마이코타(*Oomycota*) 세부의 임의의 섬유상 형태를 포함한다. (예를 들어, 참조로 본 발명에 포함된 Hawksworth *et al.*, In Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK 참조). 사상 균류는 키틴, 셀룰로오스 및 다른 복합 다당류로 구성된 세포벽을 갖는 식물성 균사체를 특징으로 한다. 사상 균류 숙주 세포는 효모와 형태학적으로 구별된다.

[0159] 특정 예시적이고 비 제한적인 실시태양에서, 사상 균류 숙주 세포는 아칠라(*Achlya*), 아크레모늄(*Acremonium*), 아스퍼길루스(*Aspergillus*), 아우레오바시디움(*Aureobasidium*), 브제르칸데라(*Bjerkandera*), 세리포리오프시스(*Ceriporiopsis*), 세팔로스포리움(*Cephalosporium*), 크리소스포리움(*Chrysosporium*), 코칠리오볼루스

(*Cochliobolus*), 코리나스쿠스(*Corynascus*), 크립토크쿠스(*Cryptococcus*), 코프리누스(*Coprinus*), 코리올루스(*Coriolus*), 디플로디아(*Diplodia*), 엔도디스(*Endothia*), 프사리움(*Fusarium*), 지베렐라(*Gibberella*), 글리오클라디움(*Gliocladium*), 휴미콜라(*Humicola*), 하이포크레아(*Hypocrea*), 마이셀리오프토라(*Myceliophthora*)(예를 들어, *Myceliophthora thermophila*), 무코르(*Mucor*), 네우로스포라(*Neurospora*), 페니실리움(*Penicillium*), 포도스포라(*Podospora*), 필레비아(*Phlebia*), 피로마이세스(*Piromyces*), 피리콜라리아(*Pyricularia*), 리히조무코르(*Rhizomucor*), 리히조푸스(*Rhizopus*), 스킨조필룸(*Schizophyllum*), 스킨탈리둠(*Scytalidium*), 스포로트리쑈(*Sporotrichum*), 탈라로마이세스(*Talaromyces*), 썬모아스쿠스(*Thermoascus*), 티에라비아(*Thielavia*), 트라마테스(*Trametes*), 폴리포클라디움(*Tolypocladium*), 트라이코데마(*Trichoderma*), 베르티실리움(*Verticillium*), 볼바리엘라(*Volvariella*), 또는 텔레오모르프(*teleomorphs*), 또는 아나모르프(*anamorphs*), 이의 동의어 또는 분류학적 동등물의 종의 세포일 수 있다.

[0160] 적합한 효모 숙주 세포는 칸디다(*Candida*), 한세누라(*Hansenula*), 사카로마이세스(*Saccharomyces*), 스킨조사카로마이세스(*Schizosaccharomyces*), 피치아(*Pichia*), 클루베로마이세스(*Kluyveromyces*) 및 야로비아(*Yarrowia*)를 포함하나 이에 제한되지 않는다. 일부 실시태양에서, 효모 세포는 한세누라 폴리모르파(*Hansenula polymorpha*), 사카로마이세스 세레비시애(*Saccharomyces cerevisiae*), 사카로마이세스 칼스베르겐시스(*Saccharomyces carlsbergensis*), 사카로마이세스 다이아스타티쿠스(*Saccharomyces diastaticus*), 사카로마이세스 노르벤시스(*Saccharomyces norbensis*), 사카로마이세스 클루이베리(*Saccharomyces kluyveri*), 스킨조사카로마이세스 폼베(*Schizosaccharomyces pombe*), 피치아 파스토리스(*Pichia pastoris*), 피치아 핀란드리카(*Pichia finlandica*), 피치아 트레할로필라(*Pichia trehalophila*), 피치아 코다메(*Pichia kodamae*), 피치아 멤브라네파시엔스(*Pichia membranaefaciens*), 피치아 오프니티애(*Pichia opuntiae*), 피치아 썬모톨레란스(*Pichia thermotolerans*), 피치아 살릭타리아(*Pichia salictaria*), 피치아 쿠에르쿰(*Pichia quercuum*), 피치아 피제페리(*Pichia pijperi*), 피치아 스티피티스(*Pichia stipitis*), 피치아 메타놀리카(*Pichia methanolica*), 피치아 안구스타(*Pichia angusta*), 클루이베로마이세스 락티스(*Kluyveromyces lactis*), 칸디다 알비칸스(*Candida albicans*), 또는 야로비아 리폴리티카(*Yarrowia lipolytica*)를 포함하나 이에 제한되지 않는다.

[0161] 특정 실시태양에서, 숙주 세포는 클라미도모나스(*Chlamydomonas*)(예를 들어, C. 레인하르티(*C. Reinhardtii*)) 및 포름디움(*Phormidium*)과 같은 조류(*P. sp.* ATCC29409)이다.

[0162] 다른 실시태양에서, 숙주 세포는 원핵생물 세포이다. 적합한 원핵생물 세포는 그람 양성, 그람 음성 및 그람 가변 박테리아 세포를 포함한다. 상기 숙주 세포는 아그로박테리움(*Agrobacterium*), 알리시로바실러스(*Alicyclobacillus*), 아나베아나(*Anabaena*), 아나시스티스(*Anacystis*), 아시네토박터(*Acinetobacter*), 아시도썬무스(*Acidothermus*), 아르쓰로박터(*Arthrobacter*), 아조박터(*Azobacter*), 바실러스(*Bacillus*), 비피도박테리움(*Bifidobacterium*), 브레비박테리움(*Brevibacterium*), 부티리비브리오(*Butyrivibrio*), 부케네라(*Buchnera*), 캠프스트리스(*Campestris*), 캠프로박터(*Campylobacter*), 클로스트리디움(*Clostridium*), 코리네박테리움(*Corynebacterium*), 크로마티움(*Chromatium*), 코프로코쿠스(*Coprococcus*), 에스체리치아(*Escherichia*), 엔테로코쿠스(*Enterococcus*), 엔테로박터(*Enterobacter*), 에르위니아(*Erwinia*), 푸소박테리움(*Fusobacterium*), 파에칼리박테리움(*Faecalibacterium*), 프란시셀라(*Francisella*), 플라보박테리움(*Flavobacterium*), 게오바실루스(*Geobacillus*), 해모필루스(*Haemophilus*), 헬리코박터(*Helicobacter*), 클레브시엘라(*Klebsiella*), 락토바실루스(*Lactobacillus*), 락토코커스(*Lactococcus*), 일로박터(*Ilyobacter*), 마이크로코쿠스(*Micrococcus*), 마이크로박테리움(*Microbacterium*), 메소르히조비움(*Mesorhizobium*), 메틸로박테리움(*Methylobacterium*), 마이코박테리움(*Mycobacterium*), 네이세리아(*Neisseria*), 판도에아(*Pantoea*), 수도모나스(*Pseudomonas*), 프로클로로코쿠스(*Prochlorococcus*), 로도박터(*Rhodobacter*), 로도수도모나스(*Rhodospseudomonas*), 로세부리아(*Roseburia*), 로도스피릴룸(*Rhodospirillum*), 로도코쿠스(*Rhodococcus*), 세네데스무스(*Scenedesmus*), 스트렙토마이세스(*Streptomyces*), 스트렙토코쿠스(*Streptococcus*), 시네코쿠스(*Synecoccus*), 사카로폴리스포라(*Saccharopolyspora*), 스태필로코쿠스(*Staphylococcus*), 세라티아(*Serratia*), 살모넬라(*Salmonella*), 시겔라(*Shigella*), 썬모아나에로박테리움(*Thermoanaerobacterium*), 트로페리마(*Tropheryma*), 툴라렌시스(*Tularensis*), 테메쿨라(*Temecula*), 썬모시네코코쿠스(*Thermosynechococcus*), 썬모코쿠스(*Thermococcus*), 우레아플라즈마(*Ureaplasma*), 잔토모나스(*Xanthomonas*), 자일렐라(*Xylella*), 예르시니아(*Yersinia*) 및 지모모나스(*Zymomonas*)의 종일 수 있으나 이에 제한되지 않는다. 일부 실시태양에서, 숙주 세포는 코리네박테리움 글루타미쿰이다.

[0163] 일부 실시태양에서, 박테리아 숙주 균주는 산업용 균주이다. 수많은 박테리아 산업용 균주가 알려져 있고 본 발명에 기술된 방법 및 조성물에 적합하다.

- [0164] 일부 실시태양에서, 박테리아 숙주 세포는 아그로박테리움 종(예를 들어, *A. radiobacter*, *A. rhizogenes*, *A. rubi*), 아테로박터 종(예를 들어, *A. aurescens*, *A. citreus*, *A. globiformis*, *A. hydrocarboglutamicus*, *A. mysorens*, *A. nicotianae*, *A. paraffineus*, *A. protophonniae*, *A. roseoparaffinus*, *A. sulfureus*, *A. ureafaciens*), 바실루스 종(예를 들어, *B. thuringiensis*, *B. anthracis*, *B. megaterium*, *B. subtilis*, *B. lentus*, *B. circularis*, *B. pumilus*, *B. lautus*, *B. coagulans*, *B. brevis*, *B. firmus*, *B. alkaophilus*, *B. licheniformis*, *B. clausii*, *B. stearothermophilus*, *B. halodurans* 및 *B. amyloliquefaciens*)이다. 특정 실시태양에서, 숙주 세포는 *B. 서브틸리스*, *B. 푸밀루스*, *B. 리체니포르미스*, *B. 메가테륨*, *B. 클라우실*, *B. 스테아로씨도필루스*, 및 *B. 아밀로리퀴에파시엔스*를 포함하나 이에 제한되지 않는 산업용 바실루스일 것이다. 일부 실시태양에서, 숙주 세포는 산업용 클로스트리디움 종(예를 들어, *C. acetobutylicum*, *C. tetani* E88, *C. lituseburense*, *C. saccharobutylicum*, *C. perfringens*, *C. beijerinckii*)일 것이다. 일부 실시태양에서, 숙주 세포는 산업용 코리네박테리움 종(예를 들어, *C. glutamicum*, *C. acetoacidophilum*)일 것이다. 일부 실시태양에서, 숙주 세포는 산업용 에스체리아치아 종(예를 들어, *E. coli*)일 것이다. 일부 실시태양에서, 숙주 세포는 산업용 에르위니아 종(예를 들어, *E. uredovora*, *E. carotovora*, *E. ananas*, *E. herbicola*, *E. punctata*, *E. terreus*)일 것이다. 일부 실시태양에서, 숙주 세포는 산업용 판토에아 종(예를 들어, *P. citrea*, *P. agglomerans*)일 것이다. 일부 실시태양에서, 숙주 세포는 산업용 수도모나스 종(예를 들어, *P. putida*, *P. aeruginosa*, *P. mevalonii*)일 것이다. 일부 실시태양에서, 숙주 세포는 산업용 스트렙토코쿠스 종(예를 들어, *S. equisimiles*, *S. pyogenes*, *S. uberis*)일 것이다. 일부 실시태양에서, 숙주 세포는 산업용 스트렙토마이세스 종(예를 들어, *S. ambofaciens*, *S. achromogenes*, *S. avermitilis*, *S. coelicolor*, *S. aureofaciens*, *S. aureus*, *S. fungicidicus*, *S. griseus*, *S. lividans*)일 것이다. 일부 실시태양에서, 숙주 세포는 자이모모나스 종(예를 들어, *Z. mobilis*, *Z. lipolytica*) 등일 것이다.
- [0165] 일부 실시태양에서, 숙주 세포는 산업용 대장균(예를 들어, *E. coli*)일 것이다.
- [0166] 대장균 종의 적절한 숙주 균주는 장독소성 대장균(ETEC), 장병원성 대장균(EPEC), 장침입성 대장균(EIEC), 장출혈성 대장균(EHEC), 요로병원성 대장균(UPEC), 베로톡신 생성 대장균, 대장균 O157:H7, 대장균 O104:H4, 대장균 O121, 대장균 O104:H21, 대장균 K1 및 대장균 NC101를 포함한다. 일부 실시태양에서, 본 발명은 대장균 K12, 대장균 B 및 대장균 C의 게놈 조각을 교시한다.
- [0167] 일부 실시태양에서, 숙주 세포는 대장균 균주 NCTC 12757, NCTC 12779, NCTC 12790, NCTC 12796, NCTC 12811, ATCC 11229, ATCC 25922, ATCC 8739, DSM 30083, BC 5849, BC 8265, BC 8231, BC 8319, BC 8320, BC 8321, BC 8322, BC 8321, BC 8231, BC 8319, BC 8231, BC 8231, 8326, BC 8327, BC 8331, BC 8335, BC 8338, BC 8341, BC 8344, BC 8345, BC 8346, BC 8347, BC 8348, BC 8863 및 BC 8864일 수 있다.
- [0168] 일부 실시태양에서, 본 발명은 균주 BC 4734 (O26:H11), BC 4735 (O157:H-), BC 4736, BC 4737 (n.d.), BC 4738 (O157:H7), BC 4945 (O26:H-), BC 4946 (O157:H7), BC 4947 (O111:H-), BC 4948 (O157:H), BC 4949 (O5), BC 5579 (O157:H7), BC 5580 (O157:H7), BC 5582 (O3:H), BC 5643 (O2:H5), BC 5644 (O128), BC 5645 (O55:H-), BC 5646 (O69:H-), BC 5647 (O101:H9), BC 5648 (O103:H2), BC 5850 (O22:H8), BC 5851 (O55:H-), BC 5852 (O48:H21), BC 5853 (O26:H11), BC 5854 (O157:H7), BC 5855 (O157:H-), BC 5856 (O26:H-), BC 5857 (O103:H2), BC 5858 (O26:H11), BC 7832, BC 7833 (O 미가공 형태:H-), BC 7834 (ONT:H-), BC 7835 (O103:H2), BC 7836 (O57:H-), BC 7837 (ONT:H-), BC 7838, BC 7839 (O128:H2), BC 7840 (O157:H-), BC 7841 (O23:H-), BC 7842 (O157:H-), BC 7843, BC 7844 (O157:H-), BC 7845 (O103:H2), BC 7846 (O26:H11), BC 7847 (O145:H-), BC 7848 (O157:H-), BC 7849 (O156:H47), BC 7850, BC 7851 (O157:H-), BC 7852 (O157:H-), BC 7853 (O5:H-), BC 7854 (O157:H7), BC 7855 (O157:H7), BC 7856 (O26:H-), BC 7857, BC 7858, BC 7859 (ONT:H-), BC 7860 (O129:H-), BC 7861, BC 7862 (O103:H2), BC 7863, BC 7864 (O 미가공 형태:H-), BC 7865, BC 7866 (O26:H-), BC 7867 (O 미가공 형태:H-), BC 7868, BC 7869 (ONT:H-), BC 7870 (O113:H-), BC 7871 (ONT:H-), BC 7872 (ONT:H-), BC 7873, BC 7874 (O raw form:H-), BC 7875 (O157:H-), BC 7876 (O111:H-), BC 7877 (O146:H21), BC 7878 (O145:H-), BC 7879 (O22:H8), BC 7880 (O 미가공 형태:H-), BC 7881 (O145:H-), BC 8275 (O157:H7), BC 8318 (O55:K-:H-), BC 8325 (O157:H7), 및 BC 8332 (ONT), BC 8333와 같은 베로세포독성 대장균(VTEC)일 수 있는 숙주 세포를 교시한다.
- [0169] 일부 실시태양에서, 본 발명은 균주 BC 8246 (O152 : K- : H-), BC 8247 (O124 : K (72) : H3), BC 8248 (O124), BC 8225 (O112), BC 8250 (O136 : K (78) : H-), BC 8251 (O124 : H-), BC 8252 (O144 : K- : H-), BC 8253 BC 8256 (O112), BC 8256 (O28a.e), BC 8257 (O124 : H-), BC 8258 (O143), BC 8259 (O167 : K- : H5), BC 8262 (O128a.c.:H35), BC 8261 (O164), BC 8262 (O164 : K- : H-), BC 8263 (O164) 및 BC 8264 (O124)과

같은 장침입성 대장균(EIEC)일 수 있는 숙주 세포를 교시한다.

- [0170] 일부 실시태양에서, 본 발명은 균주 BC 5581 (078 : H11), BC 5583 (02 : K1), BC 8221 (0118), BC 8222 (0148 : H), BC 8229 (0111), BC 8224 (0110 : H-), BC 8225 (0148), BC 8226 (0 118), BC 8227 (025 : H42), BC 8229 (06), BC 8231 (0153 : H45) BC 8238 (0148), BC 8234 (0128), BC 8235 (0118), BC 8237 (0111), BC 8238 (0 110 : H17), BC 8240 (0148), BC 8241 (0 6H 16) BC 8313 (06 : H6), BC 8315 (0153 : H-), BC 8243 (0153), BC 8244 (015 : H-), BC 8245 (020) , BC 8329, BC 8334 (0118 : H12) 및 BC 8339와 같은 장독성 대장균(ETEC)일 수 있는 숙주 세포를 교시한다.
- [0171] 일부 실시태양에서, 본 발명은 균주 BC 7567 (086), BC 7568 (0128), BC 7571 (0114), BC 7572 (0119), BC 7573 (0125) BC 7575 (0126), BC 7576 (0126), BC 7578 (0142), BC 7579 (026), BC 7580 (OK26), BC 7581 (0142), BC 7582 (055) BC 8535 (0-), BC 7585 (0-), BC 7586 (0-), BC 8330, BC 8550 (026), BC 8551 (055), BC 8552 (0158), BC 8553 (026), BC 8554 (0158), BC 8555 (086), BC 8556 (0128), BC 8557 (OK26), BC 8558 (055), BC 8560 (0158), BC 8561 (0158), BC 8562 BC 8565 (0158), BC 8565 (0158), BC 8565 (0158), BC 8568 (0111), BC 8569 (0128), BC 8570 (0114) BC 8575 (0158), BC 8575 (0158) (0158), BC 8583 (0128), BC 8584 (0158), BC 8585 (0128), BC 8586 (0158), BC 8588 (026), BC 8589 (0 86), BC 8590 (0 127), BC 8591), BC 8592 (0114), BC 8593 (0114), BC 8594 (0114), BC 8595 (0125), BC 8596 (0158), BC 8597 (026), BC 8598 (026), BC 8599 (0158), BC 8605 (0158), BC 8606 (0158), BC 8607 086), BC 8618 (086), BC 8616 (BC 8616), BC 8616 , BC 8621, BC 8622, BC 8623, BC 8624 (0158) 및 BC 8625 (0158)과 같은 장병원성 대장균(EPEC)일 수 있는 숙주 세포를 교시한다.
- [0172] 일부 실시 양태에서, 본 개시 내용은 또한 *Shigella flexneri*, *Shigella dysenteriae*, *Shigella boydii* 및 *Shigella sonnei*를 포함하는 *Shigella* 유기체일 수 있는 숙주 세포를 교시한다.
- [0173] 본 발명은 또한 포유 동물 세포, 예를 들어 인간(293, WI38, PER.C6 및 Bowes 흑색 종 세포 포함), 마우스 (3T3, NS0, NS1, Sp2/0 포함), 햄스터(CHO, BHK), 원숭이(COS, FRhL, Vero) 및 하이브리도마 세포주를 포함하는 다양한 동물 세포 유형과 함께 사용하기에 적합하다.
- [0174] 다양한 실시 양태에서, 원핵 및 진핵 균주 둘 다를 포함하는 본 발명의 실행에 사용될 수 있는 균주는 American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen 및 Zellkulturen GmbH (DSM), Centraalbureau Voor Schimmelcultures (CBS), 농업 연구 서비스 특히 문화 컬렉션, Northern Regional Research Center (NRRL)과 같은 다수의 배양 컬렉션으로부터 대중이 쉽게 접근할 수 있다.
- [0175] 일부 실시태양에서, 본 발명의 방법은 다세포 유기체에도 적용 가능하다. 예를 들어, 플랫폼은 작물의 성능을 개량시키는 데 사용될 수 있다. 유기체는 그라미니아(*Gramineae*), 페투코이데(*Fetucoideae*), 포아코이데(*Poacoideae*), 아그로스티스(*Agrostis*), 필레움(*Phleum*), 닥틸리스(*Dactylis*), 소르검(*Sorgum*), 세타리아(*Setaria*), 제아(*Zea*), 오리자(*Oryza*), 트라이티쿰(*Triticum*), 세칼레(*Secale*), 아베나(*Avena*), 호르데움(*Hordeum*), 삭카룸(*Saccharum*), 포아(*Poa*), 페스투카(*Festuca*), 스테노타프룸(*Stenotaphrum*), 사이노돈(*Cynodon*), 코익스(*Coix*), 올리레아(*Olyrae*), 파레아(*Phareae*), 콤포시태(*Compositae*) 또는 레구미노세(*Leguminosae*)와 같은 복수의 식물을 포함할 수 있다. 예를 들어, 식물은 옥수수, 쌀, 콩, 면화, 밀, 호밀, 귀리, 보리, 완두콩, 콩, 렌즈콩, 땅콩, 참마콩, 동부, 벨벳콩, 클로버, 알팔파, 자주개자리, 루핀, 살갈퀴, 등나무, 완두콩, 사탕수수, 수수, 해바라기, 카놀라 등일 수 있다. 마찬가지로, 유기체는 비-인간 포유류, 어류, 곤충 등과 같은 복수의 동물을 포함할 수 있다.
- [0176] **숙주 세포의 형질전환**
- [0177] 일부 실시태양에서, 본 발명의 방법에 의해 생성된 구조체는 형질 전환, 형질 감염, 형질 도입, 바이러스 감염, 유전자 총 또는 Ti 매개 유전자 전달을 포함하는 다양한 기술 중 임의의 것을 사용하여 숙주 세포에 도입될 수 있다. 특정 방법은 인산 칼슘 형질 감염, DEAE-텍스트란 매개 형질 감염, 형질 주입 또는 전기 천공을 포함한다 (Davis, L., Dibner, M., Battey, I. 1986 "Basic Methods in Molecular Biology"). 형질 전환의 다른 방법은 예를 들어, 아세트산 리튬 형질 전환 및 전기 천공을 포함한다. 예를 들어, Gietz et al., *Nucleic Acids Res.* 27:69-74 (1992); Ito et al., *J. Bacterol.* 153:163-168 (1983); and Becker and Guarente, *Methods in Enzymology* 194:182-187(1991) 참조. 일부 실시태양에서, 형질 전환된 숙주 세포는 재조합 숙주 균주로 불린다.
- [0178] **자동화**

- [0179] 한 실시태양에서, 본 발명에 제공된 조성물 및 방법은 숙주 세포의 유전 공학을 위한 고 처리량(HTP) 방법에 통합된다. 또 다른 실시태양에서, 본 발명에 제공된 방법은 PCT/US18/36360, PCT/US18/36333 또는 WO 2017/100377에 기술된 HTP 분자 도구 세트의 일부인 분자 도구일 수 있으며, 이들 각각은 특히 과학적 통찰력과 반복적 패턴 인식에서 파생된 HTP 유전자 설계 라이브러리를 만드는 데 모든 목적을 위해 본 발명에 참조로 포함된다. 본 발명에 제공된 조성물 및 방법은 PCT/US18/36360, PCT/US18/36333 또는 WO 2017/100377에 기술된 것과 같은 고 처리량 방법에서 사용하기 위한 라이브러리를 생성하는 데 사용될 수 있다. 본 발명에 제공된 방법을 사용하여 생성될 수 있는 라이브러리의 예는 프로모터 래더, 종결자 래더, 용해도 태그 래더 또는 분해 태그 래더를 포함할 수 있지만 이에 제한되지는 않는다. 본 발명에 제공된 조성물 및 방법을 이용할 수 있는 고 처리량 게놈 공학 방법의 예는 PCT/US18/36360, PCT/US18/36333 또는 WO 2017/100377에 기술된 바와 같은 프로모터 스와핑, 종결자(중지) 스와핑, 용해도 태그 스와핑, 분해 태그 스와핑 또는 SNP 스와핑을 포함할 수 있지만 이에 제한되지는 않는다. 고 처리량 방법은 자동화될 수 있고/있거나 로봇 공학 및 액체 취급 플랫폼(예를 들어, 당업계에 공지된 플레이트 로봇 플랫폼 및 액체 취급 기계)를 이용할 수 있다. 고 처리량 방법은, 예를 들어, 미세적정 플레이트와 같은 다중 웰 플레이트를 이용할 수 있다.
- [0180] 일부 실시태양에서, 본 발명의 자동화된 방법은 로봇 시스템을 포함한다. 본 발명에 개괄된 시스템은 일반적으로 96-또는 384-웰 미세적정 플레이트의 사용에 관한 것이지만, 당업자에 의해 이해되는 바와 같이, 임의의 수의 상이한 플레이트 또는 구성이 사용될 수 있다. 또한 본 발명에 개괄된 단계 중 일부 또는 전부는 자동화될 수 있다. 따라서, 예를 들어 시스템은 완전히 또는 부분적으로 자동화될 수 있다. 본 발명에 제공된 방법 및 조성물과 호환되는 로봇 시스템은 PCT/US18/36360, PCT/US18/36333 또는 WO 2017/100377에 기술된 것일 수 있다.
- [0181] **키트**
- [0182] 또한, 본 발명은 상기한 바와 같이 핵산 어셈블리 또는 그로부터 유래된 라이브러리를 생성하는 방법을 실행하기 위한 키트를 제공한다. 키트는 ssDNA 분자(예를 들어, 올리고뉴클레오타이드) 또는 dsDNA 분자를 어셈블링하는 데 필요한 모든 시약을 포함하는 혼합물을 포함할 수 있다. 특정 실시태양에서, 대상 키트는 i. 제 1 및 제 2 폴리뉴클레오타이드 쌍을 함유하는 제 1 폴리뉴클레오타이드 풀, ii. 삽입체 폴리뉴클레오타이드의 제 2 풀, 및 (iii) 선택적으로, 적합한 숙주 세포에서 생성된 어셈블리의 증식을 위한 적합한 클로닝 벡터를 함유할 수 있다. 일부 경우에, 키트에 양성 대조군을 포함한다.
- [0183] 한 실시태양에서, 본 발명에 제공된 키트는 5'-3' 엑소뉴클레아제 및 가닥 치환 중합 효소를 추가로 포함한다. 다른 실시태양에서, 본 발명에 제공된 키트는 5'-3' 엑소뉴클레아제, 리가아제 및 가닥 치환 중합 효소를 추가로 포함한다. 추가 실시태양에서, 본 발명에 제공된 키트는 단일 가닥(ss) 결합 단백질을 포함한다. ss 결합 단백질은 극도의 내열성 단일 가닥 DNA 결합 단백질(ET SSB), 대장균 recA, T7 유전자 2.5 산물, 파지 람다 RedB 또는 Rac 프로파지 RecT일 수 있다.
- [0184] 별도의 실시태양에서, 본 발명에 제공된 키트는 적절한 양으로 3' 엑소뉴클레아제 활성이 결여된 5' 내지 3' 엑소뉴클레아제, 크라우딩제, 3' 엑소뉴클레아제 활성을 갖는 열 안정성 비 가닥 치환 DNA 중합 효소 또는 3' 엑소뉴클레아제 활성이 결여된 제 2 DNA 중합 효소 및 분리된 열 안정성 리가아제와 상기 DNA 중합 효소의 혼합물을 추가로 포함한다. 크라우딩제는 PEG, 텍스트란 또는 피콜일 수 있다. 예를 들어, 키트는 T5 엑소뉴클레아제, PEG, PHUSION®, DNA 중합 효소 및 Taq 리가아제를 포함할 수 있다. 다른 예에서, 키트는 엑소뉴클레아제 III, PEG, AMPLITAQ GOLD® DNA 폴리머라제 및 Taq 리가아제를 포함할 수 있다.
- [0185] 본 발명에 제공된 임의의 키트는 또한 방법이 어떻게 구현될 것인지에 따라 방법에 사용될 수 있는 상기 및 하기 기재된 다른 시약, 예를 들어 mutHLS, cel-1 뉴클레아제, T7 endo 1, uvrD, T4 EndoVII, E. coli EndoV, 버퍼, dNTP, 플라스미드 및/또는 플라스미드, 대조군 등을 수용할 수 있는 실패 및/또는 적격 세포를 삽입할 플라스미드와 같은 미스매치 손상 효소를 포함할 수 있다.
- [0186] 키트의 구성 요소는 하나의 용기에 결합되거나 각 구성 요소가 자체 용기에 있을 수 있다. 예를 들어, 키트의 구성 요소는 단일 반응 튜브 또는 하나 이상의 다른 반응 튜브에서 결합될 수 있다.
- [0187] 상기 언급된 성분에 추가하여, 대상 키트는 대상 방법을 실행하기 위해 키트의 성분을 사용하기 위한 지침을 추가로 포함한다. 주제 방법을 실행하기 위한 지침은 일반적으로 적절한 기록 매체에 기록된다. 예를 들어, 설명서는 종이 또는 플라스틱 등과 같은 기재에 인쇄될 수 있다. 이와 같이 설명서는 키트의 용기 또는 그 구성 요소의 라벨링에 패키지 삽입물로서 키트에 존재할 수 있다(즉, 패키징 또는 서브 패키징과 결합됨). 다른 실시태양에서, 실제 지침은 키트에 존재하지 않지만 원격 소스로부터 지침을 얻기 위한 수단, 예를 들어 인터넷을 통

해 제공된다. 이 실시태양의 예는 지침을 볼 수 있고/있거나 지침을 다운로드할 수 있는 웹 주소를 포함하는 키트이다.

[0188] 본 발명에 기술된 바와 같이 제 1 풀에서 폴리뉴클레오타이드 쌍을 어셈블링하고 제 2 풀에서 폴리뉴클레오타이드를 삽입하기 위한 조성물, 키트 및 방법은 PCR, RCA 또는 유능한 원핵 또는 진핵 숙주 세포의 직접 형질 전환 또는 형질 감염을 포함하는 다양한 다른 분자 생물학 응용분야를 위한 주형 역할을 할 수 있는 dsDNA인 산물을 생성한다.

[0189] **실시예**

[0190] 본 발명은 하기 실시예를 참조하여 추가로 예시된다. 그러나, 전술한 실시 예와 같이 이들 실시예는 예시적인 것이며 어떤 방식으로든 본 발명의 범위를 제한하는 것으로 해석되어서는 안 된다는 점에 유의해야 한다.

[0191] **실시예 1 - DNA 단편을 플라스미드의 결정적 라이브러리로 다중 어셈블리하는 방법의 원리 증명.**

[0192] **목적**

[0193] 이 실시예는 도 1에 개략적으로 도시된 바와 같은 시험관 내 어셈블리 반응의 사용을 설명한다. 원하는 플라스미드의 라이브러리를 생성하기 위해 4개의 구성 요소를 포함하는 정밀하게 디자인된 DNA 부분을 포함하는 풀을 결정적으로 결합한다.

[0194] **방법/결과**

[0195] 삽입 서열(즉, 페이로드)는 컬럼 또는 어레이에서 합성되어 컬럼 합성 페이로드 풀 및 어레이 합성 페이로드 풀을 생성하며, 각각은 아래에 표시된 7개의 페이로드 서열의 혼합물을 포함한다.

[0196] > pMB070_promoter

[0197] ACCGTGCGTGTGACAATTTTACCTCTGGCGGTGATACTGGTTGCATGTACTAAGGAGGTTGT (SEQ ID NO. 1)

[0198] > b2405_promoter

[0199] ATGTCGGATATCTGGTGGTGAATACTTTATGCCATGATAATTTAATACGATGTATTTATTATATGGAGCACTTAATT (SEQ ID NO. 2)

[0200] > b0605_promoter

[0201] TAATGGAAACGCATTAGCCGAATCGGCAAAAATTGGTTACCTTACATCTCATCGAAAACACGGAGGAAGTATAG (SEQ ID NO. 3)

[0202] > pMB043_promoter

[0203] ACCGTGCGTGTGACTATTTTACCTCTGGCGGTAGAGTTAACATCCTACAAGGAGAACAAAAGC (SEQ ID NO. 4)

[0204] > pMB071_promoter

[0205] ACCGTGCGTGTGACTTAAATACCACTGGCGGTGATAATGGTTGCATGTACTAAGGAGGTTGT (SEQ ID NO. 5)

[0206] > b0159_promoter

[0207] CTCTCCCGGTGAGAAATACGCTTCCCCTAAGCGCATGGTAACTATGCCTTCAAATCGGGCTTATCGCGAGTAAATCT (SEQ ID NO. 6)

[0208] > pMB090_promoter

[0209] ACCGTGCGTGTGACAATTTTACCTCTGGCGGTGATAATTAACATCCTACAAGGAGAACAAAAGC (SEQ ID NO. 7)

[0210] 별도로, 왼쪽 및 오른쪽 상동성 암 쌍을 포함하는 6개의 풀이 생성되었다(즉, 도 4에서 참조된 유전자좌 풀 번호). 각 유전자좌 풀은 대장균 숙주 세포의 게놈에서 별도의 유전자좌에 상보적인 서열을 각각 포함하는 다수의 상동성 암을 포함하였다. 고유한 유전자좌의 수(즉, 상동성 암 쌍의 수)는 도 4의 각 풀에 대한 플롯 아래에 제공된다(즉, 그래프 아래 표의 'Loci in Pool'). 각 풀의 각 쌍에서 각 상동성 암의 서열은 SEQ ID NOs 8-179이다.

[0211] 컬럼 합성 페이로드 풀 및 어레이 합성 페이로드 풀은 각각 도 4에 지정된 특정 유전자좌 풀과 개별적으로 혼합되었다. 각 혼합물은 왼쪽 상동성 암:페이로드:오른쪽 상동성 암의 몰비가 대략 1:10:1임을 유의해야 한다. 페이로드 풀에 100-500개의 고유한 왼쪽 및 오른쪽 상동성 암에 해당하는 올리고뉴클레오타이드가 포함되어 있는 반면, 삽입체 올리고뉴클레오타이드의 특정 부분이 반응에서 비활성 상태일 것이라는 점을 고려할 때 주어진 반응에서 10-19개의 상동성 암만 어셈블되었기 때문에 페이로드가 초과되었다. 혼합물은 NEB Hifi DNA 어셈블리

마스터 믹스 및 E. coli 클로닝 균주에서 증식을 위한 클로닝 벡터를 추가로 포함하였다. 각 혼합물은 약 0.05pmol의 각 유전자좌 풀, 0.2-1pmol의 각 페이로드 풀 및 0.0125pmol의 클로닝 벡터를 포함하여 이론적으로 최종 어셈블된 생성물 0.0125pmol을 생성하였다. 일단 어셈블되면, 각각의 혼합물은 시험관 내 오버랩 어셈블리를 위한 NEB Hifi DNA 어셈블리 프로토콜을 적용하고 대장균 클로닝 균주에서 증식시켰다. 각 라이브러리에서 고유한 유전자좌(상동성 암의 쌍), 페이로드 및 총 가능한 구조체의 수는 도 4의 표에 제공된다.

[0212] 증식 후, 어셈블리 당 적어도 100개의 콜로니를 개별적으로 액체 배양으로 골라 밤새 성장시켰다. 액체 배양액은 전체 복제된 플라스미드의 롤링-원 증폭 (RCA)을 위한 주형으로 사용되었다. RCA 생성물은 Tn5 트랜스포사제 단편화 및 어댑터-결찰 키트(Nextera, Illumina)를 사용하여 단편화되었다. 그런 다음 PCR을 통해 샘플 특정 인덱싱 바코드를 추가하고 라이브러리의 플라스미드를 풀링하고 걸림을 정제하였다. 라이브러리 몰 농도는 Tapestation 기기로 평가되었고 라이브러리의 플라스미드는 시퀀싱을 위해 MiSeq 기기(300 사이클 키트)에 로드되었다. 각 어셈블리에 대해 시퀀싱된 플라스미드의 수는 도 4에 도시되어 있다.

[0213] 어셈블리 풀에서 생성된 어셈블리를 결정하기 위해, 어떤 DNA 서열이 어셈블되었는지 확인하기 위해 알고리즘을 사용하여 원시 시퀀싱 판독에서 각 고유 어셈블리의 부분 사이의 각 접합부에 대해 고유한 20-mer 서열을 검색하였다. 그런 다음 생성된 전체 길이 산물을 결정하기 위해 각 샘플에 대한 해당 참조 서열에 판독 값을 매핑하였다. 도 4의 플롯에서, '서열-완벽(sequence-perfect)'은 네 부분(벡터 백본, 왼쪽 및 오른쪽 상동성 암, 페이로드)이 모두 함께 어셈블되었으며 플라스미드에 돌연변이가 없음을 의미한다. '변이가 있는 올바른 어셈블리'는 올바른 배열로 네 부분이 모두 존재하지만 플라스미드에 하나 이상의 점 돌연변이가 있음을 나타낸다. '미스 어셈블리'는 잘못 결합된 상동성 암, 부분(들) 또는 존재하는 않는 부분의 일부를 갖는 플라스미드를 나타낸다.

[0214] **결론**

[0215] 도 4에 도시된 결과는 도 1에 묘사된 프로세스가 DNA 어셈블리의 결정적 라이브러리를 생성하기 위해 성공적으로 활용될 수 있음을 나타낸다.

[0216] **실시예 2 - 순환 치환 페이로드를 사용하는 큰 페이로드가 있는 다중 결정적 어셈블리 방법의 원리 증명.**

[0217] **목적**

[0218] 이 실시예는 도 3에서 제조가 기술된 원형-순환 페이로드(삽입물)를 포함하는 정밀하게 디자인된 DNA 부분을 포함하는 풀을 결정적으로 결합하기 위해 도 1에 개략적으로 도시된 바와 같은 시험관 내 어셈블리 반응의 사용을 기술한다. 한 용도를 기술한다.

[0219] **방법/결과**

[0220] 삽입체는 5'에서 3' 말단까지 페이로드의 오른쪽에 접치는 어셈블리, 53bp의 HOM2, I-SceI 제한 엔도뉴클레아제 인식 부위, 53 bp의 HOM1, 및 페이로드의 좌측에 프라이머 결합 부위를 포함하는 풀링된 정방향 프라이머를 사용하여 페이로드 서열(~ 2670bp)을 포함하는 템플릿을 증폭하여 준비되었다. 증폭 산물은 Qiagen 겔 추출 키트를 사용하여 아가로스 겔에서 절제되었다. 절제된 산물은 NEB HiFi 어셈블리 반응에서 원형화되고 AxyPrep mag 비드 정리를 통해 정제되고 I-Sce ('순환 순열')를 사용하여 선형화되었다.

[0221] 별도로, *Saccharomyces cerevisiae* 숙주 세포의 게놈에서 별도의 유전자좌에 상보적인 서열을 각각 포함하는 왼쪽 및 오른쪽 상동성 암을 게놈 DNA로부터 증폭시켰다.

[0222] 원형 치환 풀링된 페이로드와 왼쪽 및 오른쪽 상동성 암 세트를 클로닝 벡터와 결합하고 NEB HiFi 반응을 사용하여 어셈블하였다. 혼합물은 대략 1:5:1의 왼쪽 상동성 암:풀링된 삽입체:오른쪽 상동성 암의 몰비를 포함하였다. 풀링된 삽입체가 어셈블리에 사용된 10 쌍의 왼쪽/오른쪽 상동성 암과 비교하여 54개의 고유한 시퀀스를 포함하기 때문에 과도한 삽입체가 사용되었다. 혼합물은 약 16fmol의 hom arm 풀, 80fmol의 페이로드 풀 및 2.5fmol의 클로닝 벡터를 포함하여 이론적으로 2.5fmol의 최종 어셈블리 산물을 생성하였다. 일단 어셈블되면, 혼합물을 시험관 내 오버랩 어셈블리를 위해 NEB Hifi DNA 어셈블리 프로토콜을 적용하고 대장균 클로닝 균주에서 증식시켰다.

[0223] 증식 후, 여러 콜로니를 개별적으로 액체 배양으로 골라 밤새 성장시켰다. 액체 배양액은 전체 복제된 플라스미드의 롤링-원 증폭(RCA)을 위한 주형으로 사용되었다. RCA 생성물은 생거 시퀀싱에 의해 두 개의 프라이머(하나는 정방향 및 하나는 역방향)로 시퀀싱되었다. 프라이머는 클로닝 벡터에 결합하고 hom arm 및 페이로드로 판독하도록 디자인되었다.

[0224] 어셈블리 풀에서 생성된 어셈블리를 결정하기 위해, 알고리즘을 사용하여 시퀀싱 판독에서 각 고유 어셈블리에서 고유한 20-mer 서열을 검색하였다. 그런 다음 의도된 접합부가 생성되었는지 확인하기 위해 각 샘플에 대한 해당 참조 서열에 리드를 정렬하여 플라스미드가 올바르게 조립되었음을 나타낸다. 도 5에서, 상단의 긴 막대는 풀에서 어셈블된 플라스미드의 구조를 나타내고, 아래의 짧은 막대는 어셈블리 풀에서 분리된 3개의 샘플에 대한 해당 참조 서열에 정렬된 생거 서열을 나타낸다. 판독의 안쪽 끝에 있는 희미한 수직선은 생거 판독의 꼬리 끝에서 예상되는 시퀀싱 아티팩트를 나타낸다. 데이터는 의도된 모든 접합부가 어셈블되었음을 나타낸다.

[0225] **결론**

[0226] 도 5에 도시된 결과는 도 1 및 도 3에 도시된 프로세스가 긴 페이로드(예를 들어, > 200bp)를 포함하는 DNA 어셈블리의 결정적 라이브러리를 생성하기 위해 성공적으로 사용될 수 있음을 나타낸다.

[0227] **실시예 3 - PCR 증폭 페이로드를 사용하여 큰 페이로드를 사용하는 다중 결정적 어셈블리 방법의 원리 증명.**

[0228] **방법/결과**

[0229] 도 6은 숙주 게놈으로부터 유래된 템플릿으로부터 접치는 어셈블리를 추가하는 프라이머를 사용하여 PCR을 통해 부분을 포함하는 페이로드가 생성된 풀링된 어셈블리에 대한 총 성공률을 보여준다. 증폭된 페이로드는 길이가 182-213bp 범위이었다. 25bp 어셈블리 오버랩이 추가된 PCR 증폭 부분은 마그네틱 비드 기반 프로토콜을 통해 정제한 후 상동성 압을 포함하는 왼쪽 및 오른쪽 부분과 함께 페이로드에 대해 총 .25 피코몰, 왼쪽 및 오른쪽 부분을 별도로 정규화하고 이전에 벡터 백본의 .05 피코몰을 정규화하였다. NEB의 HIFI 어셈블리 마스터 믹스를 사용하여 어셈블리를 수행한 다음 전기 적격 세포로 전기 천공하였다. 성공률은 생성을 시도한 플라스미드에 비해 회수되고 NGS-QC를 통과한 플라스미드의 백분율로 계산되었다. 성공률은 풀 1 : (48/70 플라스미드), 풀 2 : (47/70 플라스미드), 풀 3 : (46/70 플라스미드), 풀 4 : (56/70 플라스미드), 풀 5 : (37/49 플라스미드)에 대한 다음과 같은 고유한 플라스미드 수의 회수 및 시퀀싱을 기반으로 한다. 풀 1 : 4의 경우, 7개의 프로모터 페이로드가 10개의 유전자좌를 대상으로 하고, 풀 5의 경우 7개의 프로모터 페이로드가 7개의 유전자좌를 대상으로 하였다. 총 5개의 풀에서 234/329 또는 71.12 %의 플라스미드가 생성되고 NGS가 확인되었다.

[0230] **결론**

[0231] 도 6에 도시된 결과는 도 1에 도시된 프로세스가 긴 페이로드(예를 들어 > 200bp)를 포함하는 DNA 어셈블리의 결정적 라이브러리를 생성하기 위해 PCR-증폭된 삽입체 단편과 함께 성공적으로 사용될 수 있음을 나타낸다.

[0232] SEQ ID NO 식별자를 가진 본 발명의 서열

표 1

[0233]

명칭	핵산 SEQ ID NO:	설명
pMB070_promoter	1	Payload sequence
b2405_promoter	2	Payload sequence
b0605_promoter	3	Payload sequence
pMB043_promoter	4	Payload sequence
pMB071_promoter	5	Payload sequence
b0159_promoter	6	Payload sequence
pMB090_promoter	7	Payload sequence
pool_1_b3748_left	8	Pool 1-Left Homology Arm
pool_1_b3748_right	9	Pool 1-Right Homology Arm
pool_1_b0388_left	10	Pool 1-Left Homology Arm
pool_1_b0388_right	11	Pool 1-Right Homology Arm
pool_1_b4348_left	12	Pool 1-Left Homology Arm
pool_1_b4348_right	13	Pool 1-Right Homology Arm
pool_1_b1982_left	14	Pool 1-Left Homology Arm
pool_1_b1982_right	15	Pool 1-Right Homology Arm
pool_1_b4367_left	16	Pool 1-Left Homology Arm
pool_1_b4367_right	17	Pool 1-Right Homology Arm
pool_1_b2285_left	18	Pool 1-Left Homology Arm
pool_1_b2285_right	19	Pool 1-Right Homology Arm
pool_1_b2405_left	20	Pool 1-Left Homology Arm

pool_1_b2405_right	21	Pool 1-Right Homology Arm
pool_1_b0495_left	22	Pool 1-Left Homology Arm
pool_1_b0495_right	23	Pool 1-Right Homology Arm
pool_1_b1646_left	24	Pool 1-Left Homology Arm
pool_1_b1646_right	25	Pool 1-Right Homology Arm
pool_1_b3189_left	26	Pool 1-Left Homology Arm
pool_1_b3189_right	27	Pool 1-Right Homology Arm
pool_2_b3125_left	28	Pool 2-Left Homology Arm
pool_2_b3125_right	29	Pool 2-Right Homology Arm
pool_2_b3787_left	30	Pool 2-Left Homology Arm
pool_2_b3787_right	31	Pool 2-Right Homology Arm
pool_2_b1948_left	32	Pool 2-Left Homology Arm
pool_2_b1948_right	33	Pool 2-Right Homology Arm
pool_2_b2790_left	34	Pool 2-Left Homology Arm
pool_2_b2790_right	35	Pool 2-Right Homology Arm
pool_2_b3197_left	36	Pool 2-Left Homology Arm
pool_2_b3197_right	37	Pool 2-Right Homology Arm
pool_2_b3791_left	38	Pool 2-Left Homology Arm
pool_2_b3791_right	39	Pool 2-Right Homology Arm
pool_2_b4260_left	40	Pool 2-Left Homology Arm
pool_2_b4260_right	41	Pool 2-Right Homology Arm
pool_2_b0071_left	42	Pool 2-Left Homology Arm
pool_2_b0071_right	43	Pool 2-Right Homology Arm
pool_2_b1687_left	44	Pool 2-Left Homology Arm
pool_2_b1687_right	45	Pool 2-Right Homology Arm
pool_2_b1006_left	46	Pool 2-Left Homology Arm
pool_2_b1006_right	47	Pool 2-Right Homology Arm
pool_3_b0335_left	48	Pool 3-Left Homology Arm
pool_3_b0335_right	49	Pool 3-Right Homology Arm
pool_3_b1940_left	50	Pool 3-Left Homology Arm
pool_3_b1940_right	51	Pool 3-Right Homology Arm
pool_3_b0109_left	52	Pool 3-Left Homology Arm
pool_3_b0109_right	53	Pool 3-Right Homology Arm
pool_3_b3399_left	54	Pool 3-Left Homology Arm
pool_3_b3399_right	55	Pool 3-Right Homology Arm
pool_3_b2478_left	56	Pool 3-Left Homology Arm
pool_3_b2478_right	57	Pool 3-Right Homology Arm
pool_3_b0320_left	58	Pool 3-Left Homology Arm
pool_3_b0320_right	59	Pool 3-Right Homology Arm
pool_3_b4521_left	60	Pool 3-Left Homology Arm
pool_3_b4521_right	61	Pool 3-Right Homology Arm
pool_3_b2260_left	62	Pool 3-Left Homology Arm
pool_3_b2260_right	63	Pool 3-Right Homology Arm
pool_3_b4169_left	64	Pool 3-Left Homology Arm
pool_3_b4169_right	65	Pool 3-Right Homology Arm
pool_3_b2405_left	66	Pool 3-Left Homology Arm
pool_3_b2405_right	67	Pool 3-Right Homology Arm
pool_4_b3493_left	68	Pool 4-Left Homology Arm
pool_4_b3493_right	69	Pool 4-Right Homology Arm
pool_4_b0479_left	70	Pool 4-Left Homology Arm
pool_4_b0479_right	71	Pool 4-Right Homology Arm
pool_4_b2470_left	72	Pool 4-Left Homology Arm
pool_4_b2470_right	73	Pool 4-Right Homology Arm
pool_4_b1451_left	74	Pool 4-Left Homology Arm
pool_4_b1451_right	75	Pool 4-Right Homology Arm
pool_4_b1981_left	76	Pool 4-Left Homology Arm
pool_4_b1981_right	77	Pool 4-Right Homology Arm
pool_4_b0237_left	78	Pool 4-Left Homology Arm

pool_4_b0237_right	79	Pool 4-Right Homology Arm
pool_4_b2497_left	80	Pool 4-Left Homology Arm
pool_4_b2497_right	81	Pool 4-Right Homology Arm
pool_4_b4260_left	82	Pool 4-Left Homology Arm
pool_4_b4260_right	83	Pool 4-Right Homology Arm
pool_4_b1412_left	84	Pool 4-Left Homology Arm
pool_4_b1412_right	85	Pool 4-Right Homology Arm
pool_4_b4139_left	86	Pool 4-Left Homology Arm
pool_4_b4139_right	87	Pool 4-Right Homology Arm
pool_4_b2039_left	88	Pool 4-Left Homology Arm
pool_4_b2039_right	89	Pool 4-Right Homology Arm
pool_4_b4473_left	90	Pool 4-Left Homology Arm
pool_4_b4473_right	91	Pool 4-Right Homology Arm
pool_4_b3510_left	92	Pool 4-Left Homology Arm
pool_4_b3510_right	93	Pool 4-Right Homology Arm
pool_4_b1007_left	94	Pool 4-Left Homology Arm
pool_4_b1007_right	95	Pool 4-Right Homology Arm
pool_4_b3058_left	96	Pool 4-Left Homology Arm
pool_4_b3058_right	97	Pool 4-Right Homology Arm
pool_4_b2688_left	98	Pool 4-Left Homology Arm
pool_4_b2688_right	99	Pool 4-Right Homology Arm
pool_4_b1716_left	100	Pool 4-Left Homology Arm
pool_4_b1716_right	101	Pool 4-Right Homology Arm
pool_4_b3071_left	102	Pool 4-Left Homology Arm
pool_4_b3071_right	103	Pool 4-Right Homology Arm
pool_4_b2139_left	104	Pool 4-Left Homology Arm
pool_4_b2139_right	105	Pool 4-Right Homology Arm
pool_5_b2434_left	106	Pool 5-Left Homology Arm
pool_5_b2434_right	107	Pool 5-Right Homology Arm
pool_5_b2037_left	108	Pool 5-Left Homology Arm
pool_5_b2037_right	109	Pool 5-Right Homology Arm
pool_5_b2451_left	110	Pool 5-Left Homology Arm
pool_5_b2451_right	111	Pool 5-Right Homology Arm
pool_5_b1902_left	112	Pool 5-Left Homology Arm
pool_5_b1902_right	113	Pool 5-Right Homology Arm
pool_5_b4310_left	114	Pool 5-Left Homology Arm
pool_5_b4310_right	115	Pool 5-Right Homology Arm
pool_5_b0676_left	116	Pool 5-Left Homology Arm
pool_5_b0676_right	117	Pool 5-Right Homology Arm
pool_5_b1497_left	118	Pool 5-Left Homology Arm
pool_5_b1497_right	119	Pool 5-Right Homology Arm
pool_5_b0183_left	120	Pool 5-Left Homology Arm
pool_5_b0183_right	121	Pool 5-Right Homology Arm
pool_5_b3631_left	122	Pool 5-Left Homology Arm
pool_5_b3631_right	123	Pool 5-Right Homology Arm
pool_5_b3791_left	124	Pool 5-Left Homology Arm
pool_5_b3791_right	125	Pool 5-Right Homology Arm
pool_5_b0438_left	126	Pool 5-Left Homology Arm
pool_5_b0438_right	127	Pool 5-Right Homology Arm
pool_5_b1981_left	128	Pool 5-Left Homology Arm
pool_5_b1981_right	129	Pool 5-Right Homology Arm
pool_5_b1709_left	130	Pool 5-Left Homology Arm
pool_5_b1709_right	131	Pool 5-Right Homology Arm
pool_5_b2176_left	132	Pool 5-Left Homology Arm
pool_5_b2176_right	133	Pool 5-Right Homology Arm
pool_5_b2168_left	134	Pool 5-Left Homology Arm
pool_5_b2168_right	135	Pool 5-Right Homology Arm
pool_5_b1872_left	136	Pool 5-Left Homology Arm

pool_5_b1872_right	137	Pool 5-Right Homology Arm
pool_5_b1203_left	138	Pool 5-Left Homology Arm
pool_5_b1203_right	139	Pool 5-Right Homology Arm
pool_5_b2231_left	140	Pool 5-Left Homology Arm
pool_5_b2231_right	141	Pool 5-Right Homology Arm
pool_5_b1622_left	142	Pool 5-Left Homology Arm
pool_5_b1622_right	143	Pool 5-Right Homology Arm
pool_6_b1857_left	144	Pool 6-Left Homology Arm
pool_6_b1857_right	145	Pool 6-Right Homology Arm
pool_6_b4024_left	146	Pool 6-Left Homology Arm
pool_6_b4024_right	147	Pool 6-Right Homology Arm
pool_6_b3942_left	148	Pool 6-Left Homology Arm
pool_6_b3942_right	149	Pool 6-Right Homology Arm
pool_6_b0592_left	150	Pool 6-Left Homology Arm
pool_6_b0592_right	151	Pool 6-Right Homology Arm
pool_6_b1415_left	152	Pool 6-Left Homology Arm
pool_6_b1415_right	153	Pool 6-Right Homology Arm
pool_6_b1762_left	154	Pool 6-Left Homology Arm
pool_6_b1762_right	155	Pool 6-Right Homology Arm
pool_6_b3414_left	156	Pool 6-Left Homology Arm
pool_6_b3414_right	157	Pool 6-Right Homology Arm
pool_6_b4374_left	158	Pool 6-Left Homology Arm
pool_6_b4374_right	159	Pool 6-Right Homology Arm
pool_6_b2917_left	160	Pool 6-Left Homology Arm
pool_6_b2917_right	161	Pool 6-Right Homology Arm
pool_6_b0346_left	162	Pool 6-Left Homology Arm
pool_6_b0346_right	163	Pool 6-Right Homology Arm
pool_6_b3966_left	164	Pool 6-Left Homology Arm
pool_6_b3966_right	165	Pool 6-Right Homology Arm
pool_6_b0406_left	166	Pool 6-Left Homology Arm
pool_6_b0406_right	167	Pool 6-Right Homology Arm
pool_6_b0652_left	168	Pool 6-Left Homology Arm
pool_6_b0652_right	169	Pool 6-Right Homology Arm
pool_6_b1493_left	170	Pool 6-Left Homology Arm
pool_6_b1493_right	171	Pool 6-Right Homology Arm
pool_6_b4159_left	172	Pool 6-Left Homology Arm
pool_6_b4159_right	173	Pool 6-Right Homology Arm
pool_6_b3795_left	174	Pool 6-Left Homology Arm
pool_6_b3795_right	175	Pool 6-Right Homology Arm
pool_6_b4246_left	176	Pool 6-Left Homology Arm
pool_6_b4246_right	177	Pool 6-Right Homology Arm
pool_6_b4440_left	178	Pool 6-Left Homology Arm
pool_6_b4440_right	179	Pool 6-Right Homology Arm

[0234] **본 발명의 번호가 매겨진 실시태양**

[0235] 본 발명에 의해 고려되는 다른 주제는 다음 번호가 매겨진 실시태양에서 설명된다:

[0236] 1. 폴리뉴클레오타이드의 혼합물을 포함하는 조성물로서, 이 혼합물은 폴리뉴클레오타이드 쌍을 함유하는 제 1 풀을 포함하며, 여기서, 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 함유한다; 및

[0237] 삽입체 폴리뉴클레오타이드의 제 2 풀을 포함하며, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열을 포함하고 제 1 풀로부터의 한 쌍의 폴리뉴클레오타이드에서 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 이의 반대 3' 말단에 제 2 어셈블리 오버랩 서열을 포함하는 것인 폴리뉴클레오타이드의 혼합물을 포함하는 조성물.

[0238] 2. 제 1 실시태양에 있어서,

- [0239] 클로닝 벡터를 추가로 포함하고, 여기서 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단은 클로닝 벡터에 상보적인 서열을 포함하는 것인 조성물.
- [0240] 3. 제 2 실시태양에 있어서,
- [0241] 제 1 풀의 각 폴리뉴클레오타이드는 제 1 풀의 폴리뉴클레오타이드가 지정된 역치를 넘어 제 1 풀의 다른 폴리뉴클레오타이드와 공통 서열을 공유하지 않아서, 제 1 풀의 폴리뉴클레오타이드의 쌍과 제 2 풀의 삼입체 폴리뉴클레오타이드 사이 또는 제 1 풀의 폴리뉴클레오타이드의 쌍과 클로닝 벡터 사이의 디자인된 어셈블리 오버랩 서열을 제외하도록 선택되는 것인 조성물.
- [0242] 4. 제 3 실시태양에 있어서,
- [0243] 지정된 역치는 5개 내지 15개의 연속(contiguous) 뉴클레오타이드인 조성물.
- [0244] 5. 제 1 내지 제 4 실시태양 중 어느 하나에 있어서,
- [0245] 중합 효소를 추가로 포함하는 것인 조성물.
- [0246] 6. 제 5 실시태양에 있어서,
- [0247] 중합 효소는 가닥 치환 또는 비 가닥 치환인 조성물.
- [0248] 7. 제 6 실시태양에 있어서,
- [0249] 중합 효소는 비 가닥 치환이고 조성물은 크라우딩제를 추가로 포함하는 것인 조성물.
- [0250] 8. 제 7 실시태양에 있어서,
- [0251] 크라우딩제는 폴리에틸렌 글리콜(PEG)인 조성물.
- [0252] 9. 제 8 실시태양에 있어서,
- [0253] PEG는 약 3 내지 약 7%(중량/부피)의 농도로 사용되는 것인 조성물.
- [0254] 10. 제 8 실시태양 또는 제 9 실시태양에 있어서,
- [0255] PEG는 PEG-200, PEG-4000, PEG-6000, PEG-8000 또는 PEG-20,000에서 선택되는 것인 조성물.
- [0256] 11. 제 6 실시태양에 있어서,
- [0257] 중합 효소는 가닥 치환이고 조성물은 단일 가닥 결합 단백질을 추가로 포함하는 것인 조성물.
- [0258] 12. 제 11 실시태양에 있어서,
- [0259] 단일 가닥 DNA 결합 단백질은 극도의 내열성 단일 가닥 DNA 결합 단백질(ET SSB), 대장균 recA, T7 유전자 2.5 산물, 파지 람다 RedB 또는 Rac 프로파지 RecT인 조성물.
- [0260] 13. 제 1 실시태양 내지 제 12 실시태양 중 어느 하나에 있어서,
- [0261] 5'-3' 엑소뉴클레아제를 추가로 포함하는 것인 조성물.
- [0262] 14. 제 1 실시태양 내지 제 13 실시태양 중 어느 하나에 있어서,
- [0263] 리가아제를 추가로 포함하는 것인 조성물.
- [0264] 15. 제 1 실시태양 내지 제 14 실시태양 중 어느 하나에 있어서,
- [0265] 제 1 풀의 각 쌍은 이중 가닥 DNA(dsDNA) 또는 단일 가닥 DNA(ssDNA)인 조성물.
- [0266] 16. 제 1 실시태양 내지 제 15 실시태양 중 어느 하나에 있어서,
- [0267] 제 2 풀의 각 삼입체 폴리뉴클레오타이드는 dsDNA 또는 ssDNA인 조성물.
- [0268] 17. 제 1 실시태양 내지 제 16 실시태양 중 어느 하나에 있어서,
- [0269] 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 조성물.

- [0270] 18. 제 1 실시태양 내지 제 16 실시태양 중 어느 하나에 있어서,
- [0271] 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 대사 경로의 일부인 유전자에 상응하는 코딩 서열을 포함하는 것인 조성물.
- [0272] 19. 제 1 실시태양 내지 제 18 실시태양 중 어느 하나에 있어서,
- [0273] 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 기능적 도메인 또는 하나 이상의 단백질에 상응하는 코딩 서열을 포함하는 것인 조성물.
- [0274] 20. 제 1 실시태양 내지 제 19 실시태양 중 어느 하나에 있어서,
- [0275] 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 단일 구조체에서 함께 연결되며, 여기서 단일 구조체는 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드 사이의 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 것인 조성물.
- [0276] 21. 제 20 실시태양에 있어서,
- [0277] 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열은 귀소 엔도뉴클레아제 인식 서열을 포함하는 것인 조성물.
- [0278] 22. 제 1 실시태양 내지 제 21 실시태양 중 어느 하나에 있어서,
- [0279] 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 풀의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함하는 것인 조성물.
- [0280] 23. 제 1 실시태양 내지 제 22 실시태양 중 어느 하나에 있어서,
- [0281] 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 풀의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 약 25개의 뉴클레오타이드를 포함하는 것인 조성물.
- [0282] 24. 제 1 실시태양 내지 제 23 실시태양 중 어느 하나에 있어서,
- [0283] 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 어셈블리 오버랩 서열과 제 2 어셈블리 오버랩 서열 사이에 위치한 하나 이상의 페이로드 서열을 포함하는 것인 조성물.
- [0284] 25. 제 24 실시태양에 있어서,
- [0285] 하나 이상의 페이로드 서열은 프로모터, 유전자, 조절 서열, 디그론을 암호화하는 핵산 서열, 용해도 태그를 암호화하는 핵산 서열, 종결자, 고유 식별자 서열 또는 이의 일부로부터 선택되는 것인 조성물.
- [0286] 26. 제 17 실시태양에 있어서,
- [0287] 제 1 풀의 제 1 및 제 2 폴리뉴클레오타이드의 각 쌍은 제 1 풀의 서로의 쌍과 비교하여 숙주 세포의 상이한 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 조성물.
- [0288] 27. 제 17 실시태양에 있어서,
- [0289] 제 1 풀의 제 1 및 제 2 폴리뉴클레오타이드의 각 쌍은 숙주 세포의 동일한 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 조성물.
- [0290] 28. 제 24 실시태양 내지 제 27 실시태양 중 어느 하나에 있어서,
- [0291] 제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 상이한 것인 조성물.
- [0292] 29. 제 24 실시태양 내지 제 27 실시태양 중 어느 하나에 있어서,
- [0293] 제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 동일한 것인 조성물.
- [0294] 30. 폴리뉴클레오타이드의 라이브러리를 생성하는 방법으로서, 이 방법은 (a) 폴리뉴클레오타이드의 제 1 풀 및 폴리뉴클레오타이드의 제 2 풀을 조합하는 단계, 여기서 제 1 풀은 폴리뉴클레오타이드의 쌍을 포함하며, 여기

서 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 제 2 풀은 삽입체 폴리뉴클레오타이드를 포함하며, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열을 포함하고 제 1 풀로부터의 한 쌍의 폴리뉴클레오타이드에서 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 이의 반대 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다; 및

- [0295] (b) 제 1 풀 및 제 2 풀을 폴리뉴클레오타이드의 라이브러리로 어셈블리하는 단계를 포함하며, 여기서 라이브러리의 각 폴리뉴클레오타이드는 제 2 풀의 삽입체 폴리뉴클레오타이드 및 제 1 풀의 한 쌍의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 어셈블리는 시험관 내 클로닝 방법 또는 생체 내 클로닝 방법을 통해 수행되는 것인 방법.
- [0296] 31. 제 30 실시태양에 있어서,
- [0297] 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 풀의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함하는 것인 방법.
- [0298] 32. 제 30 실시태양 또는 제 31 실시태양에 있어서,
- [0299] 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 제 1 풀의 한 쌍의 폴리뉴클레오타이드에서 각각 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 약 25개의 뉴클레오타이드를 포함하는 것인 방법.
- [0300] 33. 제 30 실시태양 내지 제 32 실시태양 중 어느 하나에 있어서,
- [0301] 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 단일 구조체에서 함께 연결되며, 여기서 단일 구조체는 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드 사이의 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 것인 방법.
- [0302] 34. 제 33 실시태양에 있어서,
- [0303] 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열은 귀소 엔도뉴클레아제 인식 서열을 포함하는 것인 방법.
- [0304] 35. 제 33 실시태양에 있어서,
- [0305] 연결된 단일 구조체는 스플라이싱 및 오버랩 연장 PCR(SOE-PCR), 제한 결찰, 무딘 말단 결찰, 오버랩 기반 어셈블리 방법, 재조합 기반 방법 또는 제 1 및 제 2 폴리뉴클레오타이드를 결합하는 임의의 다른 효소적 또는 화학적 방법을 통해, 또는 단일 구조체를 직접 합성함으로써 개별적인 제 1 및 제 2 폴리뉴클레오타이드를 연결하여 생성되는 것인 방법.
- [0306] 36. 제 30 실시태양 내지 제 32 실시태양 중 어느 하나에 있어서,
- [0307] 단계(a) 동안 클로닝 벡터를 제 1 풀 및 제 2 풀과 조합하는 단계를 추가로 포함하며, 여기서 클로닝 벡터의 대향 말단은 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 상보적인 서열을 포함하는 것인 방법.
- [0308] 37. 제 30 실시태양 내지 제 32 실시태양 중 어느 하나에 있어서,
- [0309] 단계(a) 전에 클로닝 벡터를 제 1 풀과 조합하는 단계를 추가로 포함하며, 여기서 클로닝 벡터의 대향 말단은 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 상보적인 서열을 포함하는 것인 방법.
- [0310] 38. 제 36 실시태양 또는 제 37 실시태양에 있어서,
- [0311] 클로닝 벡터 및 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단은 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 것인 방법.
- [0312] 39. 제 38 실시태양에 있어서,
- [0313] 하나 이상의 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제를 첨가함으로써 클로닝 벡터의 대향 말단과 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 사이에 단

일 가닥 상보적 오버행을 생성하는 단계를 추가로 포함하는 것인 방법.

- [0314] 40. 제 39 실시태양에 있어서,
- [0315] 클로닝 벡터의 대향 말단과 제 1 풀의 각 쌍에 대한 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 사이에 단일 가닥 상보적 오버행을 결합하는 단계를 추가로 포함하는 것인 방법.
- [0316] 41. 제 36 실시태양 내지 제 40 실시태양 중 어느 하나에 있어서,
- [0317] 단계(b)는 제 2 풀의 삽입체 폴리뉴클레오타이드, 제 1 풀의 쌍의 제 1 및 제 2 폴리뉴클레오타이드 및 클로닝 벡터를 포함하는 원형 산물을 생성하는 것인 방법.
- [0318] 42. 제 36 실시태양 내지 제 41 실시태양 중 어느 하나에 있어서,
- [0319] 제 1 풀은 제 1 풀의 폴리뉴클레오타이드가 지정된 위치를 넘어 제 1 풀의 다른 폴리뉴클레오타이드와 공통 서열을 공유하지 않아서, 제 1 풀의 폴리뉴클레오타이드의 쌍과 제 2 풀의 삽입체 폴리뉴클레오타이드 사이 또는 제 1 풀의 폴리뉴클레오타이드의 쌍과 클로닝 벡터 사이의 디자인된 어셈블리 오버랩 서열을 제외하도록 이련 서열의 더 큰 세트로부터의 폴리뉴클레오타이드 서열의 쌍을 선택함으로써 생성되는 것인 방법.
- [0320] 43. 제 42 실시태양에 있어서,
- [0321] 지정된 위치는 5개 내지 15개의 연속 뉴클레오타이드인 방법.
- [0322] 44. 제 30 실시태양 내지 제 43 실시태양 중 어느 하나에 있어서,
- [0323] 어셈블리는 시험관 내 클로닝 방법이며, 여기서 제 1 풀과 제 2 풀의 혼합물은 제 1 및 제 2 풀에 존재하는 폴리뉴클레오타이드를 부분적으로 또는 완전히 변성시키기 위해 가열된 다음, 어셈블리 전에 실온으로 냉각되는 것인 방법.
- [0324] 45. 폴리뉴클레오타이드의 라이브러리를 생성하는 방법으로서, 이 방법은 (a) 중합효소 연쇄 반응(PCR)을 통해 폴리뉴클레오타이드의 제 1 풀을 증폭하는 단계, 여기서 제 1 풀은 폴리뉴클레오타이드의 쌍을 포함하며, 여기서 제 1 풀의 각 쌍은 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함하며, 여기서 한 쌍의 각 제 1 폴리뉴클레오타이드 및 각 제 2 폴리뉴클레오타이드는 5' 말단 및 3' 말단을 포함하고, 여기서 증폭은 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열을 포함하는 공통 오버랩 서열을 제 1 풀로부터의 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단에 도입한다;
- [0325] (b) 공통 오버랩 서열을 이용하여 제 1 풀로부터의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드의 각 쌍을 단일 핵산 단편으로 어셈블리하는 단계, 여기서 각 쌍에 대한 단일 핵산 단편은 공통 오버랩 서열에 의해 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단으로부터 분리된 제 1 폴리뉴클레오타이드와 제 2 폴리뉴클레오타이드를 포함하며, 각 쌍에 대한 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단은 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 대해 원위에 있는 단일 핵산 단편의 대향 말단 상에 위치된다;
- [0326] (c) 각 쌍에 대한 단일 핵산 단편을 삽입체 폴리뉴클레오타이드를 포함하는 제 2 풀과 조합하는 단계, 여기서 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 단일 핵산 단편 내에 존재하는 제 1 폴리뉴클레오타이드의 3' 말단에 상보적인 5' 말단에 제 1 어셈블리 오버랩 서열 및 단일 핵산 단편 내에 존재하는 제 2 폴리뉴클레오타이드의 5' 말단에 상보적인 대향 3' 말단에 제 2 어셈블리 오버랩 서열을 포함한다;
- [0327] (d) 제 1 풀 및 제 2 풀을 원형화된 산물의 제 3 풀로 어셈블리하는 단계, 여기서 어셈블리는 시험관 내 또는 생체 내 오버랩 어셈블리 방법을 통해 수행되고, 여기서 제 3 풀의 각 원형화된 산물은 제 2 풀로부터의 삽입체 서열 및 제 1 풀로부터의 한 쌍의 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드를 포함한다;
- [0328] (e) 제 3 풀의 원형화된 산물의 각각의 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치한 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)을 인식하는 하나 이상의 부위 특이적 뉴클레아제(들)에 의한 소화를 통해 제 3 풀의 각 원형화된 산물을 선형화하는 단계; 및
- [0329] (f) 시험관 내 또는 생체 내 클로닝 방법에 의해 선형화된 산물을 클로닝 벡터로 어셈블리하는 단계를 포함하는 것인 방법.
- [0330] 46. 제 45 실시태양에 있어서,

- [0331] 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치한 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)은 귀소 뉴클레아제 인식 서열인 방법.
- [0332] 47. 제 45 실시태양 또는 제 46 실시태양에 있어서,
- [0333] 제 1 폴리뉴클레오타이드 서열과 제 2 폴리뉴클레오타이드 서열 사이에 위치하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 대한 하나 이상의 부위 특이적 뉴클레아제(들)은 귀소 엔도뉴클레아제인 방법.
- [0334] 48. 제 45 실시태양 내지 제 47 실시태양 중 어느 하나에 있어서,
- [0335] 공통 오버랩 서열은 적어도 1개의 뉴클레오타이드의 어셈블리 오버랩 서열을 포함하고 단계(b)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법에 의해 수행되는 것인 방법.
- [0336] 49. 제 45 실시태양 내지 제 47 실시태양 중 어느 하나에 있어서,
- [0337] 공통 오버랩 서열은 10-25개 뉴클레오타이드의 어셈블리 오버랩 서열을 포함하고 단계(b)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법에 의해 수행되는 것인 방법.
- [0338] 50. 제 48 실시태양 또는 제 49 실시태양에 있어서,
- [0339] 오버랩 기반 DNA 어셈블리 방법은 SOE-PCR 또는 시험관 내 오버랩 어셈블리 방법에서 선택되는 것인 방법.
- [0340] 51. 제 50 실시태양에 있어서,
- [0341] 제 1 폴리뉴클레오타이드의 5' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)은 각 쌍의 제 2 폴리뉴클레오타이드의 3' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)에 상보적이며, 여기서 단계(b)에서 각 쌍의 제 1 및 제 2 폴리뉴클레오타이드의 공통 오버랩 서열을 이용하는 것은 SOE-PCR을 수행하는 것을 수반하는 것인 방법.
- [0342] 52. 제 45 실시태양 내지 제 47 실시태양 중 어느 하나에 있어서,
- [0343] 단계(b)에서 각 쌍의 제 1 및 제 2 폴리뉴클레오타이드의 공통 오버랩 서열을 이용하는 것은 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 공통 오버랩 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하여 상보적 서열을 포함하는 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행을 생성하고; 각 쌍의 제 1 폴리뉴클레오타이드의 5' 말단 및 제 2 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행 상에 존재하는 상보적 서열을 결합하는 것을 수반하는 것인 방법.
- [0344] 53. 제 45 실시태양 내지 제 52 실시태양 중 어느 하나에 있어서,
- [0345] 단계(d)의 어셈블리는 오버랩 기반 DNA 어셈블리 방법을 사용하여 수행되는 것인 방법.
- [0346] 54. 제 53 실시태양에 있어서,
- [0347] 오버랩 기반 DNA 어셈블리는 SOE-PCR 및 시험관 내 오버랩 어셈블리 방법에서 선택되는 것인 방법.
- [0348] 55. 제 45 실시태양 내지 제 52 실시태양 중 어느 하나에 있어서,
- [0349] 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단은 하나 이상의 부위 특이적 뉴클레아제 인식 서열의 추가 세트를 포함하고 제 2 폴의 각 삼입체 폴리뉴클레오타이드의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 포함하는 것인 방법.
- [0350] 56. 제 55 실시태양에 있어서,
- [0351] 단계(d)에서의 어셈블리는 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단 상에 존재하는 추가의 하나 이상의 부위 특이적 뉴클레아제 인식 서열 및 제 2 폴로부터의 각 삼입체 폴리뉴클레오타이드의 제 1 및 제 2 어셈블리 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 각 쌍의 단일 핵산 단편의 제 1 폴리뉴클레오타이드의 3' 말단 및 제 2 폴리뉴클레오타이드의 5' 말단 상에 존재하는 추가의 하나 이상의 부위 특이적 뉴클레아제 인식 서열 및 제 2 폴로부터의 각 삼입체 폴리뉴클레오타이드의 제 1 및 제 2 어셈블리 서열에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한

하나 이상의 부위 특이적 뉴클레아제로 분해하여 제 2 폴로부터의 삽입체 폴리뉴클레오타이드의 제 1 어셈블리 서열의 5' 말단 상의 단일 가닥 오버행 상에 존재하는 서열에 상보적인 서열을 포함하는 제 1 폴리뉴클레오타이드의 3' 말단 상의 단일 가닥 오버행 및 제 2 폴로부터의 동일한 삽입체 폴리뉴클레오타이드의 제 2 어셈블리 서열의 3' 말단 상의 단일 가닥 오버행 상에 존재하는 서열에 상보적인 서열을 포함하는 제 2 폴리뉴클레오타이드의 5' 말단 상의 단일 가닥 오버행을 생성하고; 단일 가닥 오버행에 존재하는 상보적 서열을 절찰하는 것을 수반하는 것인 방법.

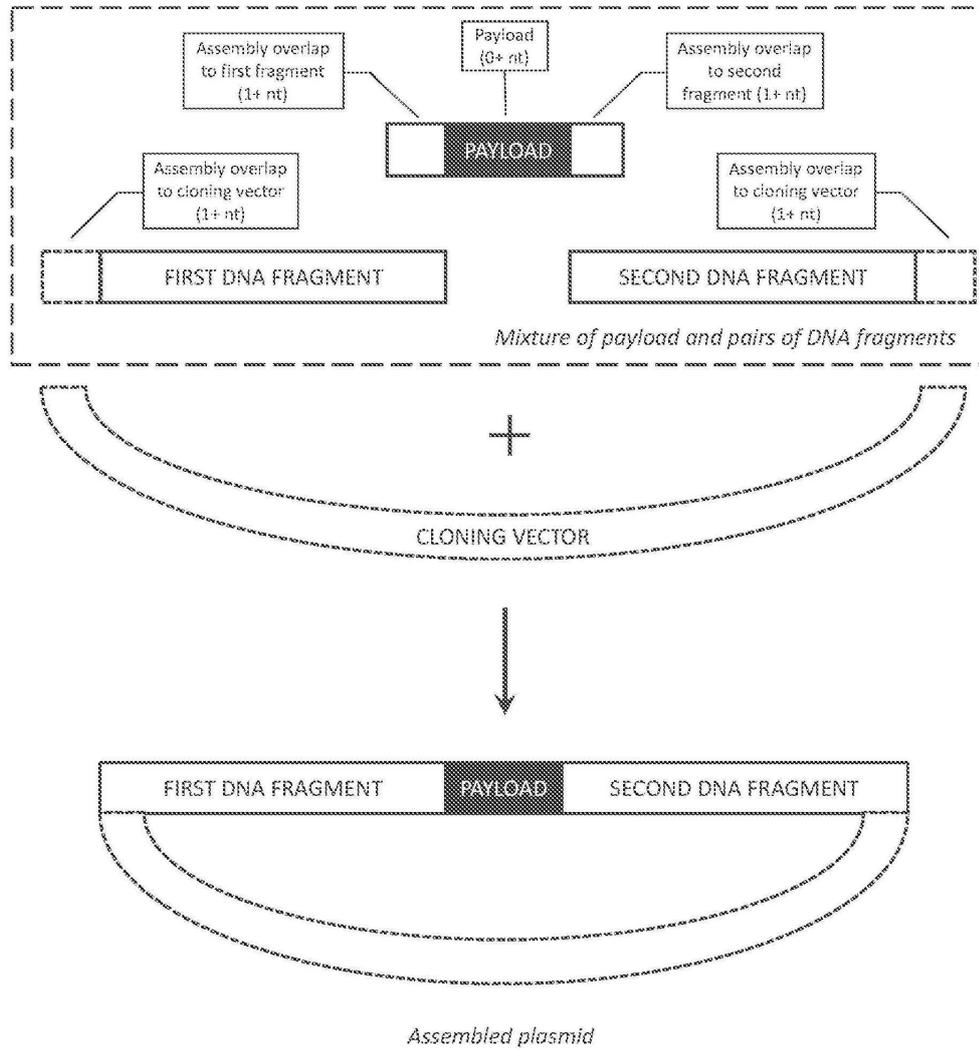
- [0352] 57. 제 45 실시태양 내지 제 56 실시태양 중 어느 하나에 있어서,
- [0353] 단계(f)의 클로닝 벡터는 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 포함하는 것인 방법.
- [0354] 58. 제 57 실시태양에 있어서,
- [0355] 단계(f)에서의 어셈블리는 클로닝 벡터의 하나 이상의 부위 특이적 뉴클레아제 인식 서열을 클로닝 벡터에 존재하는 하나 이상의 부위 특이적 뉴클레아제 인식 서열에 대한 하나 이상의 부위 특이적 뉴클레아제로 분해하는 것, 여기서 분해는 클로닝 벡터의 대향 말단 상의 단일 가닥 오버행을 생성하고, 클로닝 벡터의 대향 말단 중 하나 상의 단일 가닥 오버행은 단계(e)에서 생성된 선형화된 산물의 말단에 상보적인 서열을 포함하며 클로닝 벡터의 대향 말단 중 다른 하나 상의 단일 가닥 오버행은 단계(e)에서 생성된 선형화된 산물의 대향 말단에 상보적인 서열을 포함한다; 및 클로닝 벡터의 단일 가닥 오버행 상에 존재하는 상보적 서열과 단계(e)로부터의 선형화된 산물을 절찰하는 것을 수반하는 것인 방법.
- [0356] 59. 제 45 실시태양 내지 제 48 실시태양 중 어느 하나에 있어서,
- [0357] 제 1 풀은 제 1 풀의 폴리뉴클레오타이드가 지정된 역치를 넘어 제 1 풀의 다른 폴리뉴클레오타이드와 공통 서열을 공유하지 않아서, 제 1 풀의 폴리뉴클레오타이드의 쌍과 제 2 풀의 삽입체 폴리뉴클레오타이드 사이 또는 제 1 풀의 폴리뉴클레오타이드의 쌍과 클로닝 벡터 사이의 디자인된 어셈블리 오버랩 서열을 제외하도록 이런 서열의 더 큰 세트로부터의 폴리뉴클레오타이드 서열의 쌍을 선택함으로써 생성되는 것인 방법.
- [0358] 60. 제 59 실시태양에 있어서,
- [0359] 지정된 역치는 5개 내지 15개의 연속 뉴클레오타이드인 방법.
- [0360] 61. 제 45 내지 제 60 실시태양 중 어느 실시태양에 있어서,
- [0361] 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 단일 핵산 단편의 대향 말단에 상보적인 1개 이상의 뉴클레오타이드를 포함하는 것인 방법.
- [0362] 62. 제 45 실시태양 내지 제 61 실시태양 중 어느 하나에 있어서,
- [0363] 제 2 풀의 각 삽입체 폴리뉴클레오타이드 상의 제 1 어셈블리 오버랩 서열 및 제 2 어셈블리 오버랩 서열은 단일 핵산 단편의 대향 말단에 상보적인 약 25개의 뉴클레오타이드를 포함하는 것인 방법.
- [0364] 63. 제 30 실시태양 내지 제 62 실시태양 중 어느 하나에 있어서,
- [0365] 단계(a) 전에, 폴리뉴클레오타이드의 제 1 풀은 폴리뉴클레오타이드 쌍으로부터 각각의 제 1 폴리뉴클레오타이드를 함유하는 혼합물을 폴리뉴클레오타이드 쌍으로부터 각각의 제 2 폴리뉴클레오타이드를 함유하는 혼합물과 조합함으로써 생성되는 것인 방법.
- [0366] 64. 제 30 실시태양 내지 제 63 실시태양 중 어느 하나에 있어서,
- [0367] 제 1 풀의 각 쌍은 이중 가닥 DNA(dsDNA) 또는 단일 가닥 DNA(ssDNA)인 방법.
- [0368] 65. 제 30 실시태양 내지 제 43 실시태양 중 어느 하나에 있어서,
- [0369] 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 dsDNA 또는 ssDNA인 방법.
- [0370] 66. 제 30 실시태양 내지 제 65 실시태양 중 어느 하나에 있어서,
- [0371] 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 방법.
- [0372] 67. 제 30 실시태양 내지 제 65 실시태양 중 어느 하나에 있어서,
- [0373] 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 대사 경로의 일부인 유전자에

상응하는 코딩 서열을 포함하는 것인 방법.

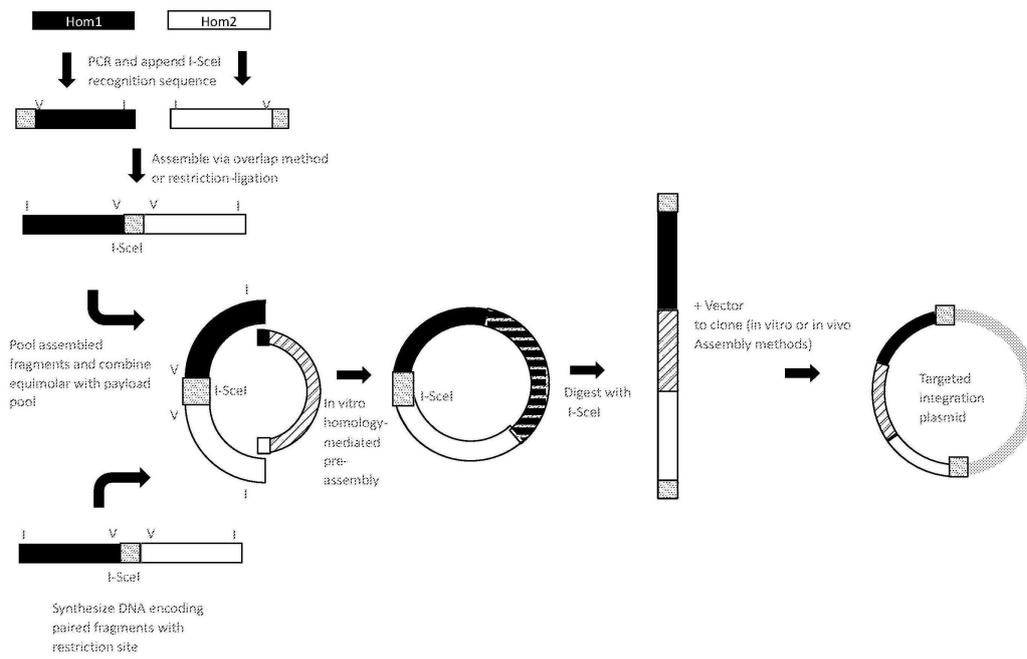
- [0374] 68. 제 30 실시태양 내지 제 65 실시태양 중 어느 하나에 있어서,
- [0375] 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 기능적 도메인 또는 하나 이상의 단백질에 상응하는 코딩 서열을 포함하는 것인 방법.
- [0376] 69. 제 30 실시태양 내지 제 68 실시태양 중 어느 하나에 있어서,
- [0377] 제 2 풀의 각 삽입체 폴리뉴클레오타이드는 제 1 어셈블리 오버랩 서열과 제 2 어셈블리 오버랩 서열 사이에 위치한 하나 이상의 페이로드 서열을 포함하는 것인 방법.
- [0378] 70. 제 69 실시태양에 있어서,
- [0379] 하나 이상의 페이로드 서열은 프로모터, 유전자, 조절 서열, 디그론을 암호화하는 핵산 서열, 용해도 태그를 암호화하는 핵산 서열, 종결자, 고유 식별자 서열 또는 이의 일부로부터 선택되는 것인 방법.
- [0380] 71. 제 66 실시태양에 있어서,
- [0381] 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 제 1 풀의 서로의 쌍과 비교하여 숙주 세포의 상이한 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 방법.
- [0382] 72. 제 66 실시태양에 있어서,
- [0383] 제 1 풀의 각 쌍에 대해, 제 1 폴리뉴클레오타이드 및 제 2 폴리뉴클레오타이드는 숙주 세포의 동일한 표적 게놈 유전자좌에 상응하는 서열을 포함하는 것인 방법.
- [0384] 73. 제 69 실시태양 내지 제 72 실시태양 중 어느 하나에 있어서,
- [0385] 제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 상이한 것인 방법.
- [0386] 74. 제 69 실시태양 내지 제 72 실시태양 중 어느 하나에 있어서,
- [0387] 제 2 풀의 삽입체 폴리뉴클레오타이드의 각 페이로드 서열은 제 2 풀의 서로의 삽입체 폴리뉴클레오타이드의 페이로드 서열과 동일한 것인 방법.
- [0388] 75. 제 30 실시태양 또는 제 69 실시태양 내지 제 74 실시태양 중 어느 하나에 있어서,
- [0389] 제 2 풀의 각 삽입체 폴리뉴클레오타이드는
- [0390] (i) 페이로드 서열, 정방향 프라이머 및 역방향 프라이머를 포함하는 혼합물에서 증합효소 연쇄 반응(PCR)을 수행하는 단계, 여기서 정방향 프라이머는 5'부터 3'로, 페이로드 서열에 상보적인 하나 이상의 뉴클레오타이드의 짧은 스트레치, 제 1 어셈블리 오버랩 서열, 하나 이상의 부위 특이적 뉴클레아제에 대한 하나 이상의 인식 서열, 제 2 어셈블리 오버랩 서열 및 페이로드 서열에 상보적인 하나 이상의 뉴클레오타이드의 제 2 스트레치를 포함하고 역방향 프라이머는 페이로드 서열 또는 페이로드 서열의 다른 하류 서열에 상보적인 서열을 포함하고, 여기서 PCR은 5'부터 3'로, 페이로드 서열에 상보적인 핵산의 짧은 스트레치, 제 1 어셈블리 오버랩 서열, 하나 이상의 부위 특이적 뉴클레아제 인식 서열, 제 2 어셈블리 오버랩 서열 및 페이로드 서열을 포함하는 PCR 산물을 생성한다;
- [0391] (ii) 스플라이싱 및 오버랩 연장 PCR(SOE-PCR), 제한 결찰, 무딘 말단 결찰, 오버랩 기반 어셈블리 방법, 재조합 기반 방법 또는 두 DNA 분자를 결합하는 임의의 다른 효소적 또는 화학적 방법으로 이루어진 그룹으로부터 선택된 어셈블리 방법을 통해 PCR 산물을 원형화하는 단계; 및
- [0392] (iii) 하나 이상의 부위 특이적 뉴클레아제 인식 서열(들)을 인식하는 하나 이상의 부위 특이적 뉴클레아제(들)로 원형화된 PCR 산물을 선형화하여 폴리뉴클레오타이드의 제 2 풀을 생성하는 단계에 의해 생성되는 것인 방법.
- [0393] 76. 제 1 실시태양 내지 제 29 실시태양 중 어느 하나에 있어서,
- [0394] 부위 특이적 뉴클레아제(들)은 제한 엔도뉴클레아제(들), 유형 II 엔도뉴클레아제(들), 귀소 엔도뉴클레아제(들), RNA 유도 뉴클레아제(들), DNA 유도 뉴클레아제(들), 징크 핑거 뉴클레아제(들), TALEN(들) 또는 니킹 효소(들) 중 하나 이상인 조성물.

도면

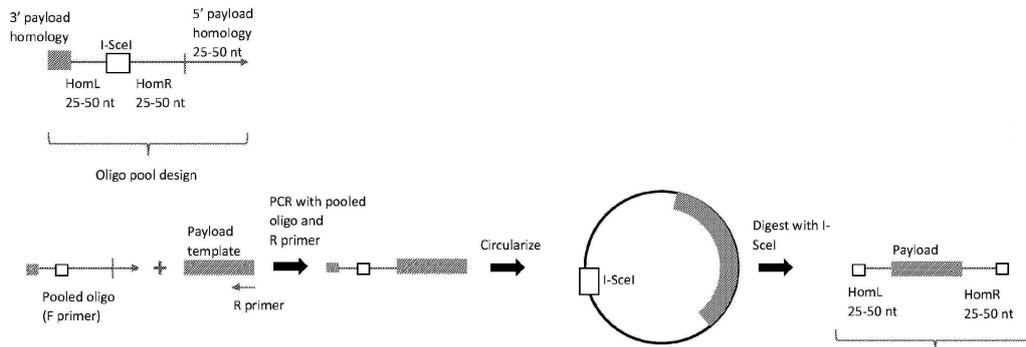
도면1



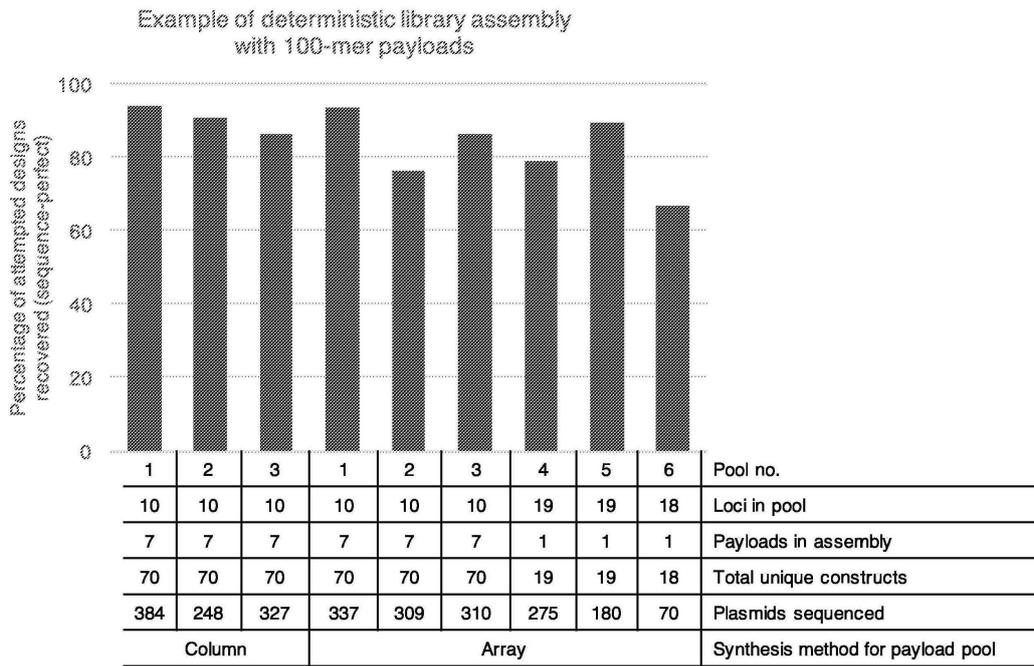
도면2



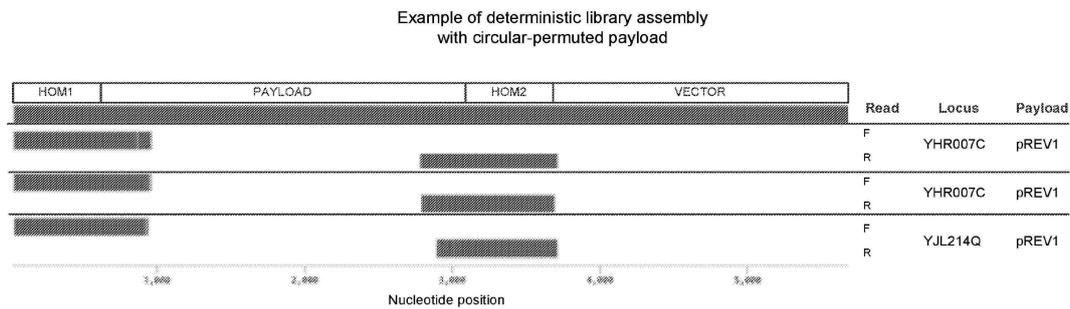
도면3



도면4



도면5



도면6

