



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0134001
(43) 공개일자 2022년10월05일

- | | |
|--|--|
| <p>(51) 국제특허분류(Int. Cl.)
C12N 15/10 (2017.01) A61K 31/713 (2006.01)
C12N 15/113 (2010.01) C12Q 1/6844 (2018.01)</p> <p>(52) CPC특허분류
C12N 15/10 (2022.05)
A61K 31/713 (2013.01)</p> <p>(21) 출원번호 10-2022-7030090</p> <p>(22) 출원일자(국제) 2021년01월29일
심사청구일자 없음</p> <p>(85) 번역문제출일자 2022년08월30일</p> <p>(86) 국제출원번호 PCT/EP2021/052203</p> <p>(87) 국제공개번호 WO 2021/152146
국제공개일자 2021년08월05일</p> <p>(30) 우선권주장
20382064.2 2020년01월31일
유럽특허청(EPO)(EP)</p> | <p>(71) 출원인
타이리스 테라뷰틱스, 에스.엘.
스페인 46113 몬카다 씨. 산타 세실리아 넘버 8
바조 이지크</p> <p>(72) 발명자
오야르자발 산타마리나, 줄렌
스페인 28003 마드리드 7디 씨. 호세 아바스칼 58</p> <p>(74) 대리인
특허법인 무한</p> |
|--|--|

전체 청구항 수 : 총 24 항

(54) 발명의 명칭 **단힌 선형 DNA의 제조방법**

(57) 요약

단힌 선형 DNA의 제조방법

본 발명은 (a) 관심 DNA 서열을 포함하는 DNA 주형을 제공하는 단계; (b) 단계 (a)의 DNA 주형으로부터 DNA를 증폭하는 단계로서, 상기 증폭은 프리마제/폴리머라제(primase/polymerase) 효소로 프라이밍(priming)되는 것인, 단계; (c) 단계 (b)에서 제조된 증폭된 DNA로 단힌 선형 DNA를 생성하는 단계; 및 (d) 단계 (c)에서 제조된 단힌 선형 DNA를 정제하는 단계를 포함하는 단힌 선형 DNA의 제조방법을 제공한다. 본 발명은 또한 본 발명의 방법에 따라 얻을 수 있는 단힌 선형 DNA, 본 발명의 단힌 선형 DNA의 치료적 유효량을 포함하는 약제학적 조성물, 및 관심 DNA 서열의 반복부를 포함하는 콘카테머 DNA를 제공한다.

(52) CPC특허분류

C12N 15/113 (2013.01)

C12Q 1/6844 (2018.05)

C12N 2310/531 (2013.01)

C12N 2310/532 (2013.01)

C12Q 2531/125 (2013.01)

명세서

청구범위

청구항 1

단한 선형 DNA의 제조방법으로서, 상기 제조방법은:

- (a) 관심 DNA 서열을 포함하는 DNA 주형을 제공하는 단계;
- (b) 단계 (a)의 DNA 주형으로부터 DNA를 증폭하는 단계로서, 상기 증폭은 프리마제/폴리머라제 (primase/polymerase) 효소로 프라이밍(priming)되는 것인, 단계;
- (c) 단계 (b)에서 제조된 증폭된 DNA로 단한 선형 DNA를 생성하는 단계; 및
- (d) 단계 (c)에서 제조된 단한 선형 DNA를 정제하는 단계;를 포함하는, 제조방법.

청구항 2

제1항에 있어서, 상기 단계 (b)에서 수행된 증폭은 롤링-서클 증폭(rolling-circle amplification)인 것인, 제조방법.

청구항 3

제1항 또는 제2항에 있어서, 상기 프리마제/폴리머라제 효소는 서열 번호 1의 TthPrimPol, 또는 서열 번호 1에 대해 적어도 80%, 구체적으로는, 서열 번호 1에 대해 적어도 85%의 서열 동일성을 갖는 TthPrimPol의 변이체인 것인, 제조방법.

청구항 4

제1항 내지 제3항 중 어느 한 항에 있어서, 상기 단계 (b)의 증폭은 가닥 변위 DNA 폴리머라제, 보다 구체적으로는, phi29 폴리머라제로 수행되는 것인, 제조방법.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 DNA 주형은 단한 선형 DNA 주형 또는 원형 이중 가닥 DNA 주형 으로부터 선택되는 것인, 제조방법.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 DNA 주형이 단한 선형 DNA 주형인 경우에, 상기 단계 (a)는 상기 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 포함하는 플라스미드 벡터를 적어도 하나의 제한 효소와 접촉시켜 상기 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA를 제조하는 단계, 및 상기 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA의 양 말단에 단일 가닥 DNA 어댑터들(DNA adaptors)을 부착시키는 단계에 의해 수행되는 것인, 제조방법.

청구항 7

제1항 내지 제5항 중 어느 한 항에 있어서, 상기 DNA 주형이 단한 선형 DNA 주형인 경우에, 상기 단계 (a)는 상기 관심 DNA 서열을 플랭킹하는 적어도 2개의 프로텔로머라제 표적 서열들을 포함하는 플라스미드 벡터를 프로

텔로머라제, 보다 구체적으로는, TeIN과 접촉시키는 단계에 의해 수행되는 것인, 제조방법.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 단계 (b)로부터 생성된 증폭된 DNA는 상기 관심 DNA 서열의 반복부를 포함하는 콘카테머 DNA이며, 이때, 반복된 관심 DNA 서열 각각은 제한 자리들 및/또는 프로텔로머라제 표적 서열들에 의해 플랭킹되는 것인, 제조방법.

청구항 9

제8항에 있어서, 상기 콘카테머 DNA가 적어도 제한 자리들에 의해 플랭킹된 상기 관심 DNA 서열의 반복부를 포함하는 경우에, 상기 단계 (c)는: (c.1) 상기 콘카테머 DNA를 적어도 하나의 제한 효소와 접촉시켜 각각 상기 관심 DNA 서열을 함유하는 복수의 열린 이중 가닥 DNA 단편들을 제조하는 단계, 및 (c.2) 상기 열린 이중 가닥 DNA 단편들의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계;에 의해 수행되는 것인, 제조방법.

청구항 10

제8항에 있어서, 상기 콘카테머 DNA가 적어도 프로텔로머라제 표적 서열들에 의해 플랭킹된 상기 관심 DNA 서열의 반복부를 포함하는 경우에, 상기 단계 (c)는 상기 콘카테머 DNA를 프로텔로머라제, 보다 구체적으로는, TeIN과 접촉시키는 단계에 의해 수행되는 것인, 제조방법.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서, 상기 DNA 주형은 프로텔로머라제 표적 부위를 포함하지 않는 닫힌 선형 DNA 주형인 것인, 제조방법.

청구항 12

제11항에 있어서, 상기 단계 (a)는 상기 관심 DNA 서열을 플랭킹 하는 적어도 2개의 제한 자리들을 플랭킹하는 2개의 프로텔로머라제 표적 서열들을 포함하는 플라스미드 벡터를 프로텔로머라제, 예를 들면, TeIN과 접촉시키는 단계에 의해 수행되고; 단계 (c)는 (c.1) 상기 콘카테머 DNA를 적어도 하나의 제한 효소와 접촉시켜 각각 상기 관심 DNA 서열을 함유하는 복수의 열린 이중 가닥 DNA 단편들을 제조하는 단계, 및 (c.2) 상기 열린 이중 가닥 DNA 단편들의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행되는 것인, 제조방법.

청구항 13

제11항에 있어서, 상기 단계 (a)는 상기 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 포함하고 프로텔로머라제 표적 부위를 포함하지 않는 플라스미드 벡터를 적어도 하나의 제한 효소와 접촉시켜 상기 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA을 제조하는 단계, 및 단일 가닥 DNA 어댑터들이 프로텔로머라제 표적 부위를 함유하지 않는 조건 하에서, 상기 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행되고; 상기 단계 (c)는 (c.1) 상기 콘카테머 DNA를 적어도 하나의 제한 효소와 접촉시켜 각각 상기 관심 DNA 서열을 함유하는 복수의 열린 이중 가닥 DNA 단편들을 제조하는 단계, 및 (c.2) 상기 열린 이중 가닥 DNA 단편들의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행되는 것인, 제조방법.

청구항 14

제1항 내지 제13항 중 어느 한 항에 있어서, 단힌 선형 DNA 주형이 프리마제/폴리머라제 프라이밍 부위를 포함하지 않는 조건 하에서, 상기 DNA 주형은 단힌 선형 DNA 주형인 것인, 제조방법.

청구항 15

제14항에 있어서, 상기 단계 (a)는:

- 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 포함하는 플라스미드 벡터를 적어도 하나의 제한 효소와 접촉시켜 상기 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA를 제조하는 단계, 및 단일 가닥 DNA 어댑터들이 프리마제/폴리머라제 프라이밍 부위를 함유하지 않는 조건 하에서, 상기 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계; 또는, 대안적으로는,
- 관심 DNA 서열을 플랭킹하는 적어도 2개의 프로텔로머라제 표적 서열들을 포함하는 플라스미드 벡터를 프로텔로머라제, 보다 구체적으로는, *Te1N*과 접촉시키는 단계; 그 결과, 상기 관심 DNA 서열을 함유하는 단힌 선형 DNA 주형인 DNA 주형을 얻는 단계;에 의해 수행되는 것인, 제조방법.

청구항 16

제1항 내지 제15항 중 어느 한 항에 있어서, 상기 관심 서열은 발현 카세트를 플랭킹하는 역 말단 반복부 (inverted terminal repeats; ITRs)를 포함하는 것인, 제조방법.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, 상기 단힌 선형 DNA가 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA의 양 말단에 단일 가닥 DNA 어댑터들을 포함하는 경우에, 상기 어댑터는 변형된 올리고뉴클레오타이드들을 포함하는 것인, 제조방법.

청구항 18

제1항 내지 제17항 중 어느 한 항에 있어서, 상기 제조방법은 무세포 시험관 내 방법인, 제조방법.

청구항 19

제1항 내지 제18항 중 어느 한 항에 정의된 제조방법에 따라 얻을 수 있는 단힌 선형 DNA.

청구항 20

요법, 보다 구체적으로는, DNA-기반 요법에 사용하기 위한 제19항에 기재된 단힌 선형 DNA.

청구항 21

제20항에 있어서, 상기 DNA-기반 요법은 유전자 요법, 유전자-편집, 세포-요법 (예컨대, CAR-Ts), 백신 및 단일 클론 항체의 발현으로부터 선택되는 것인, 제조방법.

청구항 22

제19항에 기재된 단힌 선형 DNA의 치료적 유효량 및 약제학적으로 허용가능한 담체 또는 부형제를 포함하는 약

제학적 조성물.

청구항 23

관심 DNA 서열의 반복부를 포함하는 콘카테머 DNA로서, 반복된 관심 DNA 서열 각각은 적어도 제한 자리들 및 프로텔로머라제 표적 서열들에 의해 플랭킹되는 것인, 콘카테머 DNA.

청구항 24

제23항에 있어서, 상기 반복된 관심 DNA 서열 각각은 ITRs에 의해 추가로 플랭킹되는 것인, 콘카테머 DNA.

발명의 설명

기술 분야

[0001] 본 출원은 2020년 1월 31일 출원된 유럽 특허 출원 번호 제EP20382064호의 이익을 주장한다.

[0002] 기술분야

[0003] 본 발명은 핵산 및 요법 분야에 속한다. 특히, 본 발명은 닫힌 선형 DNA, 이의 제조방법, 및 이를 포함하는 약제학적 조성물에 관한 것이다. 본 발명의 방법에 의해 얻어진 닫힌 선형 DNA는 치료 목적에 유용하다.

배경 기술

[0004] 유전자 요법은 여러 질병들의 치료에 매우 유용하다. 이는 표적 인간 세포의 핵으로 유전 물질을 성공적으로 전달하는 것을 기초로 한다. 유전자 전달 시스템은 설계에 있어 바이러스성 또는 비-바이러스성일 수 있다. 바이러스 DNA 벡터와 비교시, 비-바이러스성 전이유전자 전달 시스템은 더 안전한 유전자 전달 및 백신 설계 접근들을 제공하며, 숙주에서 염증 및 면역 반응을 일으킬 가능성이 더 적고, 더 큰 전이유전자 용량(transgene capacity)을 가지며, 저장하기에 더 용이하다.

[0005] 하지만, 비-바이러스성 벡터의 효과는 매우 제한적이어서, 임상으로의 도입을 방해하고 있다. 예를 들면, 기존의 유전자 요법용 플라스미드 DNA 벡터들을 사용하는 것은 이들이 함유한 박테리아 서열로 인해 이상 면역 반응(adverse immune response)을 유발할 수 있고, 이들의 큰 분자 크기 때문에 생체 이용률이 저하된다. 그러므로, 새로운 유형의 비-바이러스성 DNA 구조체들이 최근에 개발되고 있다.

[0006] 닫힌 선형 DNA 벡터(closed linear DNA vectors; cIDNA)는 면역원성 박테리아 백분의 벌크 없이 관심 DNA 서열만을 포함하는 덤벨 형태(dumbbell-shaped)의 분자이며, 그 결과 높은 생체 이용률, 높은 형질감염 효능, 연장된 유전자 발현 지속 시간을 보장한다. cIDNA의 선형 특성은 무작위 게놈 통합으로 인한 삽입 돌연변이 유발의 가능성을 최소화한다. cIDNA 벡터는 다양한 치료적 적응증(therapeutic indications)에 성공적으로 사용되고 있고 시험관 내(in vitro) 및 생체 내(in vivo)에서 유망한 결과를 갖는다.

[0007] 요법에 사용하기 위한, cIDNA와 같은 뉴클레오티드 벡터들의 제조는 다양한 과제들을 제시한다. 먼저, 벡터는 환자에 부작용을 일으킬 수 있는 임의의 박테리아 성분 또는 독소(toxin)를 함유하지 않아야 한다. 또한, 생체 내에서 세포로의 벡터의 형질감염 효능이 다소 낮으므로, 생체 내 투여에 필요한 양에 도달하기 위해 매우 높은 생산 수율이 필요하다. 마지막으로, 외래 뉴클레오티드 서열을 환자에 주입하는 것은 위험이 없는 것이 아니며, 예를 들면, 제조 과정 동안 생성되는 돌연변이된 단백질 또는 비특이적 생성물의 발현은 환자를 심각하게 해칠 수 있다.

[0008] 상기의 관점에서, 임상에 사용되는 벡터들은 박테리아성 잔류물(bacterial remains)과 항생제 내성 서열의 부재를 동시에 보장하는 과정에 의해 제조되어야 하고, 감소된 비용으로 대량 생산을 가능하게 해야 하고, 매우 높은 정도의 서열 균질도(즉, 매우 낮은 퍼센트의 변경된 서열)을 또한 보장해야 한다.

[0009] 따라서, 현재까지 이뤄낸 노력에도 불구하고, 치료 목적에 적합한 품질을 갖는 cIDNA의 대량의 안전한 제조를 위한 효율적인 방법이 필요하다.

발명의 내용

- [0010] 본 발명자들은 미생물의 사용을 포함하지 않는 매우 높은 서열 품질을 갖는 cDNA의 대량 제조방법을 개발했다. 본 명세서에서 제공되는 이 방법은 주형 DNA의 증폭을 프라이밍(priming)하기 위한 프리마제/폴리머라제 (primase/polymerase)의 사용 후에, 증폭 생성물을 처리하여 요법에 적합한 cDNA를 생성하는 것을 기초로 한다.
- [0011] 놀랍게도, 하기 실시예에서 나타낸 바와 같이, 본 발명자들은 cDNA 형성 이전에 프리마제/폴리머라제에 의해 프라이밍되는 증폭 단계로 인해, cDNA의 우수한 생산 수율이 얻어질 뿐만 아니라, 생성된 cDNA의 서열 정확성 (sequence fidelity)이 랜덤 프라이밍(random priming)과 같은 다른 증폭 과정을 사용하는 것에 비해 매우 개선된 것을 발견했다 (도 2 참고). 따라서, 본 명세서에서 제공되는 이 결과는 프라이밍이 얻어진 cDNA의 최종 특성에 중요한 것을 나타낸다.
- [0012] 프리마제/폴리머라제에 의해 제공되는 높은 서열 품질은 또한 증폭된 DNA의 특정 서열을 인식하는 효소를 사용하는 것을 필요로 하는 증폭 후 단계들의 효능에 영향을 미친다. 따라서, 프리마제/폴리머라제 프라이밍에 의해 제공된 높은 서열 정확성은 증폭된 DNA 상의 모든 표적 서열들이 고도로 보존되며 제한 효소 또는 프로텔로머라제 (protelomerase)와 같은 처리 효소들에 의해 효율적으로 표적화될 수 있음을 보장한다.
- [0013] 놀랍게도, 미생물을 사용하는 것을 필요로 하는 임의의 단계의 부재는 본 발명의 방법을 스케일-업(scale-up)하기에 매우 용이하고 안전하게 만든다.
- [0014] 선행 기술에 개시된 것과 대조적으로, 본 발명자들은 cDNA 주형의 증폭을 프라이밍하기 위해 프리마제/폴리머라제를 사용하는 것은 주형 DNA의 단일 가닥 루프 내의 프리마제 인식 부위들 (즉, 어댑터들(adaptors))의 존재를 필요로 하지 않음을 추가로 발견했다. 이는 본 발명의 방법에서 사용될 수 있는 DNA 주형의 레퍼토리를 크게 확장시킨다. 예를 들면, 프리마제/폴리머라제는 임의의 서열(즉, 프리마제/폴리머라제 프라이밍 부위를 함유하지 않는 서열들)의 어댑터들의 리게이션(ligation)에 의해 생산된 cDNA 주형들, 또는 최소의 단일 가닥 루프를 함유하는, 프로텔로머라제의 작용에 의해 생성된 cDNA 주형들을 프라이밍하는데 사용될 수 있다 (도 4 참고).
- [0015] 마지막으로, 본 발명에 방법에 의해 얻어진 cDNA는 포유동물 세포에 형질감염되는데 적합하며, 또한 이는 포유동물 세포에 함유된 관심 DNA 서열의 효율적인 발현을 가능하게 한다 (도 5 및 6 참고).
- [0016] 종합하면, 하기 제공된 실시예에서, 본 발명자들은 유전자 요법에 적합한 고품질 cDNA의 대량 제조에 대한 본 발명의 방법의 유용성을 입증했다.
- [0017] 따라서, 제1 측면에서, 본 발명은 (a) 관심 DNA 서열을 포함하는 DNA 주형을 제공하는 단계; (b) 단계 (a)의 DNA 주형으로부터 DNA를 증폭하는 단계로서, 상기 증폭은 프리마제/폴리머라제 효소로 프라이밍되는 것인, 단계; (c) 단계 (b)에서 제조된 증폭된 DNA로 단힌 선형 DNA를 생성하는 단계; 및 (d) 단계 (c)에서 제조된 단힌 선형 DNA를 정제하는 단계를 포함하는 단힌 선형 DNA의 제조방법을 제공한다.
- [0018] 앞서 언급한 바와 같이, 본 발명의 방법은 매우 높은 서열 정확성 (즉, 매우 적은 양의 증폭 아티팩트 (amplification artifact) 또는 돌연변이된 서열들)을 갖는 cDNA의 제조를 가능하게 하며, 이는 높은 서열 품질이 필수 요구 사항인 요법에 사용하기에 특히 적합하게 한다.
- [0019] 따라서, 제2 측면에서, 본 발명은 본 발명의 제1 측면에서 정의된 방법에 따라 얻을 수 있는 단힌 선형 DNA를 제공한다.
- [0020] 제3 측면에서, 본 발명은 요법에 사용하기 위한 제2 측면에 따른 단힌 선형 DNA를 제공한다.
- [0021] 제4 측면에서, 본 발명은 제2 측면에 따른 단힌 선형 DNA의 치료적 유효량 및 약제학적으로 허용가능한 담체 또는 부형제를 포함하는 약제학적 조성물을 제공한다.
- [0022] 제5 측면에서, 본 발명은 관심 DNA 서열의 반복부(repeats)를 포함하는 콘카테머 DNA(concatemeric DNA)를 제공하며, 이때, 반복된 관심 DNA 서열 각각은 적어도 재조합 효소(recombinase) 인식 부위들, 또는 대안적으로는, 적어도 제한 자리들 및 프로텔로머라제 표적 서열들에 의해 프랭킹(flanking)된다.

도면의 간단한 설명

- [0023] 실시예 1과 관련된 도 1은, 증폭이 TthPrimPol 또는 랜덤 프라이머(random primer; RP)에 의해 프라이밍된 DNA 주형의 RCA 증폭으로부터 생성된 DNA 생산량(DNA yield)을 나타낸다. y축은 μg 단위의 DNA 생산량을 나타낸다.

NTC는 DNA 주형이 없는 대조군 반응을 지칭한다; 플라스미드는 플라스미드 주형 DNA를 갖는 증폭 반응을 지칭한다.

실시에 1과 관련된 도 2는, TthPrimPol로 프라이밍된 증폭을 RP로 프라이밍된 증폭과 비교한 일루미나 서열 분석(Illumina sequencing)을 나타낸다.

실시에 2와 관련된 도 3은, 다양한 반응 생성물(product reactions)로 로딩된 아가로스 겔의 사진을 나타낸다: 1) DNA 래더(DNA ladder); 2) pUC57-Kan_TEL0-CMV-EGFP의 TthPrimPol 프라이밍된 증폭; 3) TelN 로 분해된 pUC57-Kan_TEL0-CMV-EGFP의 TthPrimPol 프라이밍된 증폭; 4) DNA 래더; 5) TelN, EcoRI 및 HindIII로 분해된 pUC57-Kan_TEL0-CMV-EGFP의 TthPrimPol 프라이밍된 증폭; 6) TelN, EcoRI, HindIII 및 ExoIII로 분해된 pUC57-Kan_TEL0-CMV-EGFP의 TthPrimPol 프라이밍된 증폭. 화살표는 카세트 (표적 분자)에 상응하는 밴드 크기를 나타낸다.

실시에 3과 관련된 도 4는, TelN 로 생성된 RCA 증폭 c1DNA를 나타낸다. A) TelN에 의해 생성된 플라스미드 주형 또는 c1DNA 주형의 RCA 증폭으로부터 생성된 DNA 생산량로서, 상기 증폭은 TthPrimPol 또는 랜덤 프라이머 (RP)에 의해 프라이밍된다. y축은 μg 단위의 DNA 생산량을 나타낸다. NTC는 주형 DNA가 없는 대조군 반응을 지칭한다. B) 표시된 조건으로 얻어지고 TelN으로 처리되거나 처리되지 않은 DNA 증폭 생성물로 로딩된 아가로스 겔의 사진. 로딩 전에, DNA 생성물은 EcoRI, HindIII 및 ExoIII로 처리되었다. 화살표는 카세트 (표적 분자)에 상응하는 밴드 크기를 나타낸다.

실시에 4와 관련된 도 5는, 표시된 구조체로 형질감염된 후 24h 및 48h에서 HEK293 세포의 형광 강도의 정량화를 나타낸다. N.T.는 처리되지 않은 세포를 의미한다. y축은 형광 강도의 임의 단위(Arbitrary unit)를 나타낸다.

실시에 4와 관련된 도 6은, 표시된 구조체로 형질감염된 후 24h 및 48h에서 HEK293 세포의 대표 이미지를 나타낸다.

도 7은 oDNA 41의 품질 제어 매개변수를 나타낸다. A, 아가로스 겔 전기영동 (M1, 슈퍼코일 DNA 래더 마커 TAKARA: 3585A; M2, 1 kb DNA 래더 TIAGEN MD111; 레인 5, oDNA 41); B, 그레이스케일 분석(Grayscale analysis); D, 생어 서열 분석(Sanger Sequencing).

도 8은 oDNA 21의 품질 제어 매개변수를 나타낸다. A, 아가로스 겔 전기영동 (M1, 슈퍼코일 DNA 래더 마커 TAKARA: 3585A; M2, 1 kb DNA 래더 TIAGEN MD111; 레인 4, oDNA 21); B, 그레이스케일 분석; D, 생어 서열 분석.

도 9는 본 발명의 c1DNA의 제조를 위한 관심 서열을 함유하는 eGFP 플라스미드 (서열 번호 20을 갖는 플라스미드)의 단편의 대표도를 나타낸다. 나타낸 단편은 상응하는 프로모터 및 인핸서(enhancer)와 같은 추가 서열들과 함께 관심 서열(이 경우 GFP를 인코딩하는 서열)을 포함한다. 이 관심 서열은 BsaI 제한 자리 및 프로테로머라제 표적 서열에 의해 플랭킹(flanking)된다.

도 10은 본 발명의 c1DNA의 제조를 위한 관심 서열을 함유하는 Luc-ITR 플라스미드 (서열 번호 22를 갖는 플라스미드)의 단편의 대표도를 나타낸다. 나타낸 단편은 상응하는 프로모터, 인핸서, 및 AVV2-ITR와 같은 추가 서열들과 함께 관심 서열(이 경우 루시페라제를 인코딩하는 서열)을 포함한다. 이 관심 서열은 BsaI 제한 자리 및 프로테로머라제 표적 서열에 의해 플랭킹된다.

도 11은 oDNA 4_{ITR}의 아가로스 겔 전기 영동을 나타낸다 (M, DL3000 래더; 레인 12, oDNA 4_{ITR}).

도 12는 실시에 6 (RCA에 이어서 프로테로머라제 처리)에서와 같이, eGFP 플라스미드(서열 번호 20을 갖는 플라스미드)로부터 얻어진 c1DNA의 아가로스 겔 전기영동을 나타낸다(M1, 슈퍼코일 DNA 래더 마커 TAKARA: 3585A; M2, 1 kb DNA 래더 TIAGEN MD111; 레인 2, 실시에 6의 c1DNA).

발명을 실시하기 위한 구체적인 내용

[0024] 본원에 사용되는 모든 용어들은, 달리 나타내지 않는 한, 당업계에 알려진 통상적인 의미로 이해되어야 한다. 본 출원에서 사용된 특정 용어들의 다른 보다 구체적인 정의들은 하기에 제시되고 달리 명시적으로 제시된 정의가 더 넓은 정의를 제공하지 않는 한, 명세서 및 청구범위 전체에 걸쳐 균일하게 적용되는 것이 의도된다.

[0025] 본 명세서에서 사용되는 바와 같이, 부정 관사 "하나(a)" 및 "하나(an)"는 "적어도 하나" 또는 "하나 이상"과

동의어이다. 달리 나타내지 않는 한, "이(the)"와 같은 정관사는 명사의 복수형도 포함한다.

[0026] 본 발명은 제1 측면에서, (a) 관심 DNA 서열을 포함하는 DNA 주형을 제공하는 단계; (b) 단계 (a)의 DNA 주형으로부터 DNA를 증폭하는 단계로서, 상기 증폭은 프리마제/폴리머라제 효소로 프라이밍되는 것인, 단계; (c) 단계 (b)에서 제조된 증폭된 DNA로 단힌 선형 DNA를 생성하는 단계; 및 (d) 단계 (c)에서 제조된 단힌 선형 DNA를 정제하는 단계를 포함하는 단힌 선형 DNA의 제조방법을 제공한다.

[0027] 프라이밍 효소로서 프리마제/폴리머라제를 사용한 DNA 주형의 증폭은, 매우 높은 효능 및 정확성을 갖는 증폭된 DNA를 생성하였으며, 이는 추후 처리되어 치료 용도에 적합한 단힌 선형 DNA를 생성할 수 있다.

[0028] 본 명세서에서 사용되는 바와 같이, 용어 "단힌 선형 DNA" 또는 "cDNA"는 뉴클레오티드 혼성화를 가능하게 하는 조건 하에서 "덤벨" 또는 "강아지뼈(doggy-bone)" 형태 구조를 형성하는 단일 가닥의 공유결합적 단힌 DNA 분자(covalently closed DNA molecule)를 지칭한다. 그러므로, cDNA가 단힌 단일 가닥 DNA 분자에 의해 형성되더라도, 동일한 분자 내의 2개의 상보적 서열들의 혼성화에 의한 "덤벨" 구조의 형성은 2개의 단일 가닥 루프들에 의해 플랭킹된 이중 가닥 중간 세그먼트로 이루어진 구조를 생성한다. 당업자는 일반적인 분자 생물학 기술들을 사용하여 열린 또는 단힌 이중 가닥 DNA - 예를 들면, 단계 (b)에서 제조된 증폭된 DNA -로부터 cDNA를 생성하는 방법을 알고 있다. 예를 들면, 당업자는 열린 이중 가닥 DNA의 양 말단에 단일 가닥 헤어핀 어댑터를 부착함으로써 - 예를 들면, 리가아제의 작용에 의해 - cDNA가 생성될 수 있음을 알고 있다. 단힌 선형 DNA를 생성하기 위한 당업자에게 공지된 또 다른 방법은 적어도 2개의 프로텔로머라제 표적 서열들을 포함하는 이중 가닥 DNA 상의 프로텔로머라제의 작용을 통한 것이다.

[0029] "관심 서열(sequence of interest)"은 올바른 유전자 발현에 필요한 다른 서열들과 함께, 관심 유전자를 인코딩하는 최소로 필요한 서열들(minimum necessary sequences)을 포함하는 이중 가닥 DNA 단편, 예를 들면, 발현 카세트로 이해된다. 관심 서열은 발현 카세트를 플랭킹하는 다른 서열들, 예컨대, 역 말단 반복부 (ITR)를 추가로 포함할 수 있다.

[0030] 본 명세서에서 사용되는 바와 같이, 용어 "프라이밍"은 효소에 의해 폴리뉴클레오티드 주형 상에 올리고뉴클레오티드 프라이머가 생성되는 것을 지칭한다.

[0031] 용어 "프리마제/폴리머라제 효소"는 고세균-진핵생물 프리마제(archaeo-eukaryotic primase; AEP) 슈퍼패밀리의 효소들과 같은 DNA-지시된(DNA-directed) 프리마제/폴리머라제 효소를 지칭한다. 이 효소들은, 예를 들면, 테르무스 테르모필루스 프리마제/폴리머라제(Thermus thermophilus primase/polymerase; TthPrimPol) 또는 인간 프리마제/폴리머라제(hsPrimPol, CCDC111, FLJ33167, EukPrim2 또는 hPrimPol1)이다. "테르무스 테르모필루스 프리마제/폴리머라제" 또는 "TthPrimPol"는 서열 번호 1의 박테리아 테르무스 테르모필루스의 프리마제/폴리머라제를 지칭한다. 이 뉴클레오티드 및 단백질 서열은 각각 NC_005835 및 WP_01 1173100.1로서 NCBI Entrez 데이터베이스에서 이용가능하다.

표 1

서열 번호	명칭	서열
서열 번호 1	TthPrimPol	MRPIEHALSYAAQGYGLPLRPGGKEPLGKLVPHGLKNASRDPATLEAWWRSCPRCGVGLPGPEVLVLFDFDDPEAWEGLRQEHPALEAAPRQRTPKGGRHVFLRLPEGVRLSASVRAIPGVDLRGMGRAYVVAAPTRLKDGRTYTWEAPLTPPEELPPVPQALLKLLPPPPPPRPSWGAVGTASPKRLQALLQAYAAQVARTPEGQRHLTLIRYAVAAGGLIPHGLDPREAEEVLVAAAMSAGLPEWEARDAVRWGLGVGASRPLVLESSSKPPEPRTYRARVYARMRRWV
서열 번호 2	TeIN	MSKVKIGELINTLVNEVEAIDASDRPQGDKTKRIKAAAARYKNALFNDKRFKRGKGLQKRITANTFNAYMSRARKRFDDKLLHHSFDKNINKLSEKYPLYSEELSSWLSMPTANIRQHMSLQSKLKEIMPLAEELSNVRIGSKGSDAKIARLIKYPDWSFALSDLNSDDWKERRDYLYKLFQQGSALLEELHQLKVNHEVLYHLQLSPAERTSIQQRWADVLRKKRNVVVDYPTYMQSIYDILNPNATLFLSNTRSGMAPLAFALAAVSGRRMIEIMFQGEFAVSGKYTVNFSGQAKKRSEDKSVTRTIYTLCEAKLFVELLTELRSCSAASDFDEVVKGYGKDDTRSENGRINA ILAKAFNPWVKSFSGDDRRVYKDSRAIYARIAYEMFFRVDPRWKNVDEDVFFMELGHDDENTQLHYKQFKLANFSRTWRPEVGDENTRLVALQKLDDEMPGFARGDAGVRLHETVKQLVEQDPSAKITNSTLRAFKFSPTMISRYLEFAADALGQFVGENGQWQLKIETPAIVLPDEESVETIDEPDDESQDD ELDEDEIELDEGGGDEPTEEEGPEEHQPTALKPVFKPAKNNGDGYKIEFEYDGHKHYAWSGPADSPMAAMRSAWETYYS
서열 번호 3	TeIN 표적 서열	TATCAGCACACAATTGCCATTATACGCGGTATAATGGACTATTGTGTGCTGATA
서열 번호 4	AflIII 어댑터	TTAAGTAACATTGTGGCCACTCAGCCAACAATGTTAC

서열 번호 5	NheI-HF 어댑터	CTAGCTAACATTTGTTGGCCACTCAGGCCAACAAATGTTAG
서열 번호 6	EcoRV 어댑터	TCTAACATTTGTTGGCCACTCAGGCCAACAAATGTTAGAT
서열 번호 7	ScaI 어댑터	CTTAACATTTGTTGGCCACTCAGGCCAACAAATGTTAAGT
서열 번호 8	5' ITR 어댑터	CCTGCAGGCAGCTGCGCGCTCGCTCGCTACTGAGGCCGCCCGGGCAAAGCCCGGGCGTCGGGCGACCT TTGGTCGCCCGGCCTCAGTGAGCGAGCGAGCGCGCAGAGAGGGAGTGGCCAACTCCATCACTAGGGGTT CCT
서열 번호 9	3' ITR 서열	AGGAACCCCTAGTGATGGAGTGGCCACTCCCTCTCTGCGGCTCGCTCGCTACTGAGGCCGGGCGAC CAAAGGTCGCCCGACGCCCGGCTTTGCCCGGGCGCCTCAGTGAGCGAGCGAGCGCGCAGCTGCCTGC AGG

- [0033] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 상기 방법은 (a) 관심 DNA 서열을 포함하는 DNA 주형을 제공하는 단계; (b) 단계 (a)의 DNA 주형으로부터 DNA를 증폭하는 단계로서, 상기 증폭은 프라이마제/폴리머라제 효소로 DNA 주형을 프라이밍하는 단계 (b.1), 및 생성 서열을 폴리머라제로 연장하는 단계 (b.2)를 포함하는 것인, 단계; (c) 단계 (b)에서 제조된 증폭된 DNA로 단힌 선형 DNA를 생성하는 단계; 및 (d) 단계 (c)에서 제조된 단힌 선형 DNA를 정제하는 단계를 포함한다.
- [0034] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 프라이마제/폴리머라제 효소는 TthPrimPol 또는 hsPrimPol로부터 선택된다. 특정한 일 구현예에서, 프라이마제/폴리머라제 효소는 TthPrimPol이다. 보다 특정한 일 구현예에서, 프라이마제/폴리머라제 효소는 서열 번호 1의 TthPrimPol 또는 서열 번호 1과 적어도 70%, 적어도 75%, 적어도 80%, 적어도 85%, 적어도 90%, 또는 적어도 95%의 서열 동일성을 갖는 TthPrimPol의 변이체이다. 당업자는 TthPrimPol의 프라이마제 활성을 유지하는 TthPrimPol의 임의의 변이체는 본 발명의 방법에 사용하기에 적합할 것임을 알 것이다.
- [0035] 본 발명에, 용어 "동일성"은 서열들이 최적으로 정렬될 때, 2개의 서열들에서 동일한 잔기들의 백분율을 지칭한다. 최적 정렬에서, 제1 서열의 위치에 제2 서열의 상응하는 위치와 동일한 아미노산이 점유되면, 서열들은 그 위치에 대해 동일성을 나타낸다. 2개의 서열들 사이의 동일성 수준 (또는 "퍼센트 서열 동일성")은 서열 크기에 대한 서열들이 공유하는 동일한 위치의 수의 비로 측정된다 (즉, 퍼센트 서열 동일성 = (동일한 위치들의 수/위치들의 총 수) x 100).
- [0036] 최적 정렬을 빠르게 얻고 둘 이상의 서열들 사이의 동일성을 계산하기 위한 다수의 수학 알고리즘들은 공지되어 있으며 이들은 다수의 이용가능한 소프트웨어 프로그램들에 포함되어 있다. 이러한 프로그램들의 예시들은 아미노산 서열 분석 등을 위한 MATCH-BOX, MULTAIN, GCG, FASTA, 및 ROBUST 프로그램이 있다. 바람직한 소프트웨어 분석 프로그램은 ALIGN, CLUSTAL W, 및 BLAST 프로그램 (예를 들면, BLAST 2.1, BL2SEQ, 및 이들의 이후 버전들)을 포함한다.
- [0037] 아미노산 서열 분석의 경우, BLOSUM 매트릭스들 (예를 들면, BLOSUM45, BLOSUM50, BLOSUM62, 및 BLOSUM80 매트릭스), Gonnet 매트릭스들, 또는 PAM 매트릭스들 (예를 들면, PAM30, PAM70, PAM120, PAM160, PAM250, 및 PAM350 매트릭스)과 같은 가중치 매트릭스(weight matrix)가 동일성을 결정하는데 사용된다.
- [0038] BLAST 프로그램들은, 데이터베이스의 다양한 서열들에 대해 선택된 서열을 정렬하거나(예를 들면, GenSeq), 또는, BL2SEQ로 2개의 선택된 서열들 사이에 정렬함으로써, 적어도 2개의 아미노산 서열들의 분석을 제공한다. BLAST 프로그램들은, 바람직하게는 BLAST 프로그램 작동에 통합된 DUST 또는 SEG 프로그램들과 같은 낮은 복잡성 필터링 프로그램들(low complexity filtering programs)에 의해 바람직하게는 변형된다. 갭 존재 비용(gap existence cost) (또는 갭 점수(gap score))이 사용되는 경우, 갭 존재 비용은 바람직하게는 약 -5 내지 -15로 설정된다. 유사한 갭 매개변수들은 다른 프로그램들과 적절히 사용될 수 있다. BLAST 프로그램들 및 이들의 기반이 되는 원리는 예를 들면, Altschul et al., "Basic local alignment search tool", 1990, J. Mol. Biol, v. 215, pages 403-410에 추가로 개시되어 있다. 동일성의 특정한 퍼센트는 TthPrimPol 효소가 여전히 효과적이어서, 적합한 서열들을 프라이밍할 수 있게 하는 하나 이상의 아미노산의 보존적 돌연변이들에 의한 서열의 변이들을 포함한다. 또한, 단백질 변이들은 하나 이상의 아미노산의 삽입 또는 결실로 인한 것이다.
- [0039] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과

조합하여, 상기 방법은 단힌 선형 DNA의 제조를 위한 시험관 내 무세포(cell-free) 방법이다.

- [0040] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단계 (b)의 증폭은 롤링-서클 증폭(rolling-circle amplification)이다.
- [0041] 용어 "롤링-서클 증폭" 또는 "RCA"는 공유결합적 단힌 DNA 분자들, 예컨대, c1DNA 또는 이중 가닥 원형 DNA의 증폭을 포함하는 핵산 증폭 반응을 지칭하며, 이때, 폴리머라제는 단힌 DNA 분자 주위에서 프라이머의 연장을 수행한다. 폴리머라제는 혼성화된 카피를 변위(displacement)시키고 주형 주위에서 폴리클레오티드 연장을 계속 하여 증폭된 DNA의 탠덤 유닛들(tandem units)를 포함하는 콘카테머 DNA를 생성한다. 이 선형 단일 가닥 생성물들은 다중 혼성화, 프라이머 연장 및 가닥 변위(strand displacement) 사건에 대한 기초로 역할을 하여, 콘카테머 이중 가닥 DNA 생성물들을 생성한다. 따라서, 콘카테머 이중 가닥 DNA 생성물들 내에는 증폭된 단일 유닛 DNA 각각의 다중 카피들이 있다. 당업자들은 이들의 일반 지식 및/또는 제조업체의 지침을 활용하여 효소 및 증폭될 주형의 특성에 따라 증폭 단계의 조건을 어떻게 조정해야 할지 알고 있다. 주형 DNA가 생성되는 방법에 따라, 콘카테머 DNA는 증폭된 관심 DNA 서열 각각을 플랭킹하는 상이한 서열들을 함유할 것이다. 예를 들면, 콘카테머 DNA에서 반복된 관심 DNA 서열은 제한 자리들, 프로텔로머라제 표적 서열들, 재조합 효소 인식 부위들, 또는 이들의 임의의 조합에 의해 플랭킹될 수 있다.
- [0042] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단계 (b)의 증폭은 가닥 변위 DNA 폴리머라제로 수행된다. 용어 "가닥-변위 DNA 폴리머라제"는 주형 DNA의 이중 가닥 부분을 제거하면서 3' 말단 연장 반응을 수행하는 DNA 폴리머라제를 지칭한다. 본 발명에 사용될 수 있는 가닥 변위 DNA 폴리머라제는 가닥-변위 활성을 갖는다면 특별히 제한되지 않으며, 예컨대, phi29 DNA 폴리머라제 및 Bst DNA 폴리머라제가 있다. 따라서, 선택되는 폴리머라제 유형에 따라, 당업자는 3' 말단 연장 반응의 반응 조건은 적절하게 설정될 수 있는 것을 알 것이다. 예를 들면, phi29 DNA 폴리머라제가 사용되는 경우에, 반응은 25℃ 내지 35℃의 반응을 위한 최적 온도에서 수행될 수 있다.
- [0043] 따라서, 특정한 일 구현예에서, 가닥 변위 DNA 폴리머라제는 phi29 DNA 폴리머라제, Bst DNA 폴리머라제, Bca (엑소-) DNA 폴리머라제, 대장균(*Escherichia coli*) DNA 폴리머라제 I의 Klenow 단편, Vent (엑소-) DNA 폴리머라제, DeepVent (엑소-) DNA 폴리머라제, 및 KOD DNA 폴리머라제로부터 선택된다. 보다 특정한 일 구현예에서, 가닥 변위 DNA 폴리머라제는 phi29 DNA 폴리머라제이다. 더욱 특정한 일 구현예에서, 가닥 변위 DNA 폴리머라제는 phi29 DNA 폴리머라제를 포함하는 키메라 단백질이다. 당업자는, 예를 들면, WO2011000997에 개시된 바와 같은 개선된 특성들을 갖는 키메라 DNA 폴리머라제를 얻기 위한 방법을 알고 있다.
- [0044] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, DNA 주형은 단힌 DNA 주형 또는 원형 이중 가닥 DNA 주형으로부터 선택된다.
- [0045] 본 명세서에서 사용되는 바와 같이, 용어 "원형 이중 가닥 DNA"는 공유결합적 단힌 이중 가닥 DNA 분자이다.
- [0046] 본 발명의 단힌 선형 DNA의 제조방법은 또한 랜덤 프라이머로 단계 (b)의 증폭을 프라이밍함으로써 수행될 수 있다.
- [0047] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, DNA 주형이 단힌 선형 DNA 주형인 경우에, 단계 (a)는 하기에 의해 수행된다:
- [0048] - 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 포함하는 플라스미드 벡터를 적어도 하나의 제한 효소와 접촉시켜 상기 관심 DNA 서열을 포함하는 열린 이중 가닥 DNA를 제조하는 단계, 및 단일 가닥 DNA 어댑터들을 상기 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA의 양 말단에 부착시키는 단계; 또는, 대안적으로는, 하기에 의해 수행된다:
- [0049] - 관심 DNA 서열을 플랭킹하는 적어도 2개의 프로텔로머라제들을 포함하는 플라스미드 벡터를 프로텔로머라제, 보다 구체적으로는, Te1N과 접촉시키는 단계; 그 결과, 상기 관심 DNA 서열을 함유하는 단힌 선형 DNA 주형인 DNA 주형을 얻는 단계.
- [0050] 본 발명자들은 놀랍게도, 최신 기술에서 개시된 것과 대조적으로, 프리마제/폴리머라제 효소는 프리마제 인식 부위를 보유한 어댑터들을 함유하지 않는 c1DNA를 프라이밍할 수 있음을 발견했다. 특히, c1DNA 주형이 프로텔로머라제에 의해 생성되는 경우에, 생성된 c1DNA는 말단에 매우 작은 단일 가닥 루프들이 존재하는 구조를 제시하고 상기 단일 가닥 루프들은 프로텔로머라제 표적 서열을 함유하지 않는다. 예기치 않게, 프리마제/폴리머라제는 주형 c1DNA가 변성 조건을 거치지 않는 경우에도, 이 유형의 c1DNA를 프라이밍하여 폴리머라제가 증폭 단

계를 시작하는 것을 가능하게 한다 (도 4 참고).

- [0051] 따라서, 본 발명의 제1 측면의 방법의 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, DNA 주형은 프리마제/폴리머라제 프라이밍 부위를 함유하지 않는다. 특정한 일 구현예에서, DNA 주형은 프리마제/폴리머라제 프라이밍 부위를 함유하지 않는 cDNA 주형이다.
- [0052] 본 명세서에서 사용되는 바와 같이, "플라스미드 벡터"는 연결된 또 다른 핵산을 수송할 수 있고 염색체 DNA와 독립적으로 세포 내에서 자율 복제할 수 있는 원형 이중 가닥 핵산 분자를 지칭한다. 그러므로, 플라스미드 벡터는 세포, 특히 박테리아 세포에서 복제에 필요한 모든 인자들을 함유하고 있다.
- [0053] 제한 효소 및 (부착을 위한) 리가아제의 사용은 분자 생물학 분야에서 일반적이므로, 당업자는 사용되는 효소에 따라 반응 조건을 어떻게 조정할 지, 및 표적화될 제한 자리에 따라 어떤 제한 효소를 사용해야 할지를 알 것이다.
- [0054] 당업자는 또한 일부 제한 효소들이 DNA 돌출부(DNA overhangs)를 생성하는 반면에 (점착 말단(sticky ends)), 다른 제한 효소들은 DNA 돌출부들을 생성하지 않는 것 (평활 말단(blunt ends))을 알고 있다. 두 유형의 제한 효소들 모두 본 발명에서 사용될 수 있다. 당업자는 점착 말단을 갖는 어댑터가 점착 말단을 갖는 열린 이중 가닥 DNA에 부착될 수 있는 것(점착-말단 연결)을 알고 있다. 평활 말단을 갖는 열린 이중 가닥 DNA는 또한 예를 들면, Taq 폴리머라제를 사용하여 말단 3' 데옥시 아데노신 뉴클레오티드를 부가하는 방법에 의해 dA-테일링(dA-tailing)된 다음에, 돌출 T를 갖는 어댑터에 연결될 수 있다.
- [0055] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 제한 효소는 평활 말단 또는 점착 말단을 생성한다. 보다 특정한 일 구현예에서, 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 포함하는 플라스미드 벡터를 적어도 하나의 제한 효소와 접촉시키는 것은 점착 말단을 갖는 열린 이중 가닥 DNA 또는 평활 말단을 갖는 열린 이중 가닥 DNA를 생성한다.
- [0056] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단일 가닥 DNA 어댑터들은 헤어핀 구조를 갖는다. 보다 특정한 일 구현예에서, 단일 가닥 DNA 어댑터들은 서열 번호 4, 서열 번호 5, 서열 번호 6, 또는 서열 번호 7의 서열이다. de cDNA를 형성하기 위해 열린 이중 가닥 DNA 양 말단에 부착된 어댑터들은 동일한 어댑터 또는 상이한 어댑터일 수 있다.
- [0057] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단일 가닥 DNA 어댑터들은 하나 이상의 변형된 뉴클레오티드, 예를 들면, 적어도 2개의 변형된 뉴클레오티드들을 함유한다.
- [0058] "변형된 뉴클레오티드"는 화학적으로 변형된 - 염기, 당 또는 인산기의 변형에 의함 -, 또는 비-천연 모이어티가 구조 내에 포함된 임의의 뉴클레오티드 (예를 들면, 아데노신, 구아노신, 시티딘, 및 티미딘)이다. 따라서, 변형된 뉴클레오티드는 변형에 따라 자연 발생이거나 자연 발생이 아닐 수 있다.
- [0059] 본 명세서에 사용되는 변형된 뉴클레오티드는 바람직하게는, 임의의 제한을 암시하지 않고, 임의의 자연 발생된 구아노신, 우리딘, 아데노신, 티미딘 및 시티딘, 또는 예를 들면, 아세틸화, 메틸화, 하이드록실화 등에 의해 화학적으로 변경된 자연 발생되지 않은 구아노신, 우리딘, 아데노신, 티미딘 또는 시티딘의 변이체이며, 5-메틸-데옥시시티딘, 2-아미노-데옥시아데노신, 1-메틸-아데노신, 1-메틸-구아노신, 1-메틸-이노신, 2,2-디메틸-구아노신, 2,6-디아미노퓨린, 2'-아미노-2'-데옥시아데노신, 2 '-아미노-2'-데옥시시티딘, 2'- 아미노-2'-데옥시구아노신, 2 '-아미노-2'-데옥시우리딘, 2-아미노-6-클로로퓨린리보사이드, 2- 아미노퓨린-리보사이드, 2'-아라아데노신, 2'-아라시티딘, 2'-아라우리딘, 2'-아지도-2'- 데옥시아데노신, 2'-아지도-2'-데옥시시티딘, 2'-아지도-2 '-데옥시구아노신, 2'-아지도-2'- 데옥시우리딘, 2-클로로아데노신, 2'-플루오로-2'-데옥시아데노신, 2 '-플루오로-2'-데옥시시티딘, 2'-플루오로-2'-데옥시구아노신, 2'-플루오로-2'-데옥시우리딘, 2'-플루오로티미딘, 2-메틸- 아데노신, 2-메틸-구아노신, 2-메틸-티오-N6-이소펜테닐-아데노신, 2'-0-메틸-2- 아미노아데노신, 2'-0-메틸-2'-데옥시아데노신, 2 '-0-메틸-2'-데옥시시티딘, 2 '-0- 메틸-2'-데옥시구아노신, 2, -0-메틸-2'-데옥시우리딘, 2'-0-메틸-5-메틸우리딘, 2'- 0-메틸이노신, 2'-0-메틸슈도우리딘, 2-티오시티딘, 2-티오-시티딘, 3-메틸- 시티딘, 4-아세틸-시티딘, 4-티오우리딘, 5-(카르복시하이드록시메틸)-우리딘, 5,6- 디하이드로우리딘, 5-아미노알릴시티딘, 5-아미노알릴-데옥시우리딘, 5-브로모우리딘, 5- 카르복시메틸아미노메틸-2-티오-우라실, 5- 카르복시메틸아미노메틸-우라실, 5-클로로- 아라-사이토신, 5-플루오로-우리딘, 5-요오도우리딘, 5-메톡시카르보닐메틸-우리딘, 5-메톡시-우리딘, 5-메틸-2-티오-우리딘, 6-아자시티딘, 6-아자우리딘, 6-클로로-7-테아자-구아노신, 6-클로로퓨린리보사이드, 6-머캅토-구아노신, 6-메틸-머캅토포린-리보사이드, 7-테아자- 2'-데옥시-구

아노신, 7-데아자아데노신, 7-메틸-구아노신, 8-아자아데노신, 8-브로모- 아데노신, 8-브로모-구아노신, 8-머캅토-구아노신, 8-옥소구아노신, 벤즈이미다졸- 리보사이드, 베타-D-만노실-큐오신, 디하이드로-우리딘, 이노신, N1-메틸아데노신, N6-([6-아미노핵실] 카바모일메틸)-아데노신, N6-이소펜테닐-아데노신, N6-메틸-아데노신, N7-메틸-잔토신, N-우라실-5-옥시아세트산 메틸 에스테르, 퓨로마이신, 큐오신, 우라실- 5-옥시아세트산, 우라실-5-옥시아세트산 메틸 에스테르, 와이부톡소신신(wybutoxosine), 잔토신, 및 자일로- 아데노신을 포함한다. 이러한 변이체들의 제조는 당업자에게 공지되어 있고, 예를 들면 US4373071에 공지되어 있다.

[0060] 변형된 뉴클레오티드는, 제한 없이, 피리딘-4-온리보뉴클레오시드, 5-아자-우리딘, 2- 티오-5-아자-우리딘, 2-티오우리딘, 4-티오-슈도우리딘, 2-티오-슈도우리딘, 5- 하이드록시우리딘, 3-메틸우리딘, 5-카르복시메틸-우리딘, 1-카르복시메틸-슈도우리딘, 5-프로피닐-우리딘, 1-프로피닐-슈도우리딘, 5-타우리노메틸우리딘, 1-타우리노메틸-슈도우리딘, 5-타우리노메틸-2-티오-우리딘, 1 -타우리노메틸-4-티오-우리딘, 5-메틸-우리딘, 1-메틸-슈도우리딘, 4-티오-1-메틸-슈도우리딘, 2-티오-1-메틸-슈도우리딘, 1-메틸-1-데아자-슈도우리딘, 2-티오-1 -메틸-1-데아자-슈도우리딘, 디하이드로우리딘, 디하이드로슈도우리딘, 2-티오-디하이드로우리딘, 2-티오-디하이드로슈도우리딘, 2-메톡시우리딘, 2-메톡시-4-티오-우리딘, 4-메톡시-슈도우리딘, 및 4-메톡시-2- 티오-슈도우리딘, 5-아자-시티딘, 슈도이소시티딘, 3-메틸-시티딘, N4-아세틸시티딘, 5-포밀시티딘, N4-메틸시티딘, 5-하이드록시메틸시티딘, 1-메틸-슈도이소시티딘, 피롤로-시티딘, 피롤로-슈도이소시티딘, 2-티오-시티딘, 2-티오-5-메틸-시티딘, 4-티오-슈도이소시티딘, 4-티오-1-메틸-슈도이소시티딘, 4-티오-1-메틸-1-데아자-슈도이소시티딘, 1-메틸-1-데아자-슈도이소시티딘, 제불라린(zebularine), 5-아자-제불라린, 5-메틸-제불라린, 5-아자-2-티오-제불라린, 2-티오-제불라린, 2-메톡시-시티딘, 2-메톡시-5-메틸-시티딘, 4-메톡시-슈도이소시티딘, 및 4-메톡시-1-메틸-슈도이소시티딘을 포함할 수 있다.

[0061] 변형된 뉴클레오티드는, 제한 없이, 2-아미노퓨린, 2,6-디아미노퓨린, 7-데아자- 아데닌, 7-데아자-8-아자-아데닌, 7-데아자-2-아미노퓨린, 7-데아자-8-아자-2-아미노퓨린, 7-데아자-2,6-디아미노퓨린, 7-데아자-8-아자-2,6-디아미노퓨린, 1-메틸아데노신, N6-메틸아데노신, N6-이소펜테닐아데노신, N6-(시스-하이드록시이소펜테닐)아 데노신, 2-메틸티오-N6-(시스-하이드록시이소펜테닐) 아데노신, N6-글리시닐카바모일아데노신, N6-트레오닐카바모일아데노신, 2-메틸티오-N6-트레오닐 카바모일아데노신, N6,N6-디메틸아데노신, 7-메틸아데닌, 2-메틸티오-아 데닌, 및 2-메톡시-아데닌을 포함할 수 있다.

[0062] 변형된 뉴클레오티드는, 제한 없이, 이노신, 1-메틸-이노신, 와이오신(wyosine), 와이부토신(wybutosine), 7-데 아자-구아노신, 7-데아자-8-아자-구아노신, 6-티오-구아노신, 6-티오-7-데아자-구아노신, 6-티오-7-데아자-8-아 자-구아노신, 7-메틸-구아노신, 6-티오-7-메틸-구아노신, 7-메틸이노신, 6-메톡시-구아노신, 1 -메틸구아노신, N2-메틸구아노신, N2,N2-디메틸구아노신, 8-옥소-구아노신, 7-메틸-8-옥소-구아노신, 1-메틸-6-티오-구아노신, N2-메틸-6-티오-구아노신, 및 N2,N2-디메틸-6-티오-구아노신을 포함할 수 있다.

[0063] 변형된 뉴클레오티드는, 제한 없이, 6-아자-시티딘, 2-티오-시티딘, 알파-티오-시티딘, 슈도-이소-시티딘, 5-아 미노알릴-우리딘, 5-요오도-우리딘, N1 -메틸- 슈도우리딘, 5,6-디하이드로우리딘, 알파-티오-우리딘, 4-티오-우리딘, 6-아자-우리딘, 5- 하이드록시-우리딘, 데옥시-티미딘, 5-메틸-우리딘, 피롤로-시티딘, 이노신, 알파- 티오- 구아노신, 6-메틸-구아노신, 5-메틸-시티딘, 8-옥소-구아노신, 7-데아자-구아노신, N1-메틸-아데노신, 2- 아미노-6-클로로-퓨린, N6-메틸-2-아미노-퓨린, 슈도-이소- 시티딘, 6-클로로-퓨린, N6-메틸-아데노신, 알파-티 오-아데노신, 8-아지도-아데노신, 7-데아자-아데노신을 포함할 수 있다.

[0064] 변형된 뉴클레오티드는 2' 위치에서 화학적으로 변형될 수 있다. 바람직하게는, 변형된 뉴클레오티드는 2' 탄소 원자에서 치환체를 포함하는데, 상기 치환체는 할로겐, 알콕시기, 수소, 아릴옥시기, 아미노기 및 아미노알콕시 기로부터 선택되고, 바람직하게는 2'-수소 (2'-데옥시), 2'-O-메틸, 2'-O-메톡시에틸 및 2'-플루오로로부터 선택된다.

[0065] 본 명세서에 기재된 뉴클레오티드의 2' 위치와 관련된 또 다른 화학적 변형은, 잠금 핵산(locked nucleic acid; LNA) 뉴클레오티드, 에틸렌 브릿지 핵산 (ethylene bridged nucleic acid; ENA) 뉴클레시드 및 (S)-구속 에틸 cET 뉴클레오티드이다. 이 백본 변형들은 변형된 뉴클레오티드의 당을 바람직한 노던 구조(northern conformation)로 잠근다.

[0066] 백본의 인산기는 예를 들면, 하나 이상의 산소 원자를 상이한 치환체로 대체함으로써 변형될 수 있다. 또한, 변형된 뉴클레오티드는 변형되지 않은 포스페이트 모이어터를 본 명세서에 기재된 변형된 포스페이트로 완전히 대체하는 것을 포함할 수 있다. 변형된 인산 그룹의 예시는, 제한 없이, 포스포로티오에이트 (트리포스페이트로도 알려져), 포스포셀레네이트, 보라노 포스페이트, 보라노 포스페이트 에스테르, 수소 포스포네이트, 포스포로

아미데이트(phosphoroamidate), 알킬 포스포네이트, 아릴 포스포네이트 및 포스포트리에스테르로 이루어진 그룹을 포함한다. 포스페이트 링커는 연결 산소를 질소(브릿지 포스포아미데이트), 황(브릿지 포스포리티오에이트) 및 탄소(브릿지 메틸렌-포스포네이트)로 대체함으로써 변형될 수 있다.

- [0067] 변형된 뉴클레오티드는 무염기 부위일 수 있다. 본 명세서에서 사용되는 바와 같이, "무염기 부위(abasic site)"는 유기 염기가 결핍된 뉴클레오티드이다. 바람직한 구현예에서, 무염기 뉴클레오티드는 리보스의 2' 위치에서 본 명세서에 기재된 바와 같은 화학적 변형을 더 포함할 수 있다. 바람직하게는, 리보스의 2' C 원자는 할로겐, 알콕시기, 수소, 아릴옥시기, 아미노기 및 아미노 알콕시기로 이루어진 군으로부터 선택된 치환체, 바람직하게는 2'-수소 (2'-데옥시), 2'-O-메틸, 2'-O-메톡시메틸 및 2'-플루오로로부터 선택된 치환체로 치환된다.
- [0068] 본 발명의 제1 측면의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 적어도 2개의 변형된 뉴클레오티드들은 2-아미노-데옥시아데노신, 5-메틸-데옥시시티딘, 티오포스페이트 뉴클레오티드, LNA 뉴클레오티드, 이노신, 8-옥소-데옥시아데노신 및 5-플루오로-데옥시우라실 및 L-DNA 뉴클레오티드로 이루어진 군으로부터 독립적으로 선택된다.
- [0069] 본 발명의 제1 측면의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 적어도 2개의 변형된 뉴클레오티드들은 L-DNA 뉴클레오티드, 5-브로모우리딘 또는 5-요오도우리딘이 아니다.
- [0070] 2-아미노-데옥시아데노신 (2-아미노-2'-데옥시아데노신 또는 2-아미노-dA로도 알려짐)은 데옥시아데노신의 유도체이다. 2-아미노-데옥시아데노신은 (2R,3S,5R)-5-(2,6-디아미노퓨린-9-일)-2-(하이드록시메틸)옥솔란-3-올의 IUPAC 명칭을 가지며, CAS 번호 4546-70-7을 갖는다.
- [0071] 5-메틸-데옥시시티딘 (5-메틸-dCTP)은 데옥시시티딘의 유도체이며, IUPAC 명칭은 ([[(2R,3S,5R)-5-(4-아미노-5-메틸-2-옥소피리미딘-1-일)-3-하이드록시옥솔란-2-일]메톡시-하이드록시포스포틸] 포스포노 수소 포스페이트)이고, CAS 번호는 22003-12-9이다.
- [0072] 티오포스페이트 뉴클레오티드는 인산기로서 티오포스페이트(포스포리티오에이트로도 알려짐)를 함유하는 임의의 뉴클레오티드이다. 티오포스페이트는 CAS 번호 15181-41-6을 갖는다.
- [0073] LNA 뉴클레오티드는 리보스 모이어티가 2' 산소 및 4' 탄소를 연결하는 추가 브릿지(extra bridge)로 변형된 것인 변형된 RNA 뉴클레오티드이다.
- [0074] L-DNA 뉴클레오티드는 리보스 또는 데옥시리보스의 L 거울상이성질체를 함유하는 뉴클레오티드를 지칭한다.
- [0075] 본 발명의 제1 측면의 보다 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, cDNA는 티오포스페이트, 잠금 핵산, 2,6-디아미노퓨린, 5-메틸-데옥시시티딘, 이노신, 8-옥소-데옥시아데노신 및 5-플루오로-데옥시우라실 및 L-DNA 뉴클레오티드로 이루어진 군으로부터 독립적으로 선택되는 적어도 3개, 적어도 4개, 또는 적어도 5개의 변형된 뉴클레오티드들을 포함한다.
- [0076] 본 발명의 제1 측면의 보다 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, cDNA는 2개의 LNA 뉴클레오티드들을 포함한다.
- [0077] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단일 가닥 DNA 어댑터들은 적어도 하나의 제한 자리를 포함한다. 보다 특정한 일 구현예에서, 제한 자리는 BsaI 제한 자리, AfIII 제한 자리, HindIII 제한 자리, NheI 제한 자리, 및 EcoRV 제한 자리로 이루어진 군으로부터 선택된다. 더욱 특정한 일 구현예에서, 제한 자리는 BsaI 제한 자리이다.
- [0078] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단일 가닥 DNA 어댑터들은 프리마제 인식 부위를 함유하지 않는다. 보다 특정한 구현예에서, 단일 가닥 DNA 어댑터들은 서열 XTC를 함유하지 않는다.
- [0079] 본 발명의 제1 측면의 방법의 보다 구체적인 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, DNA 주형이 단한 선형 DNA 주형인 경우에, 단계 (a)는 하기에 의해 수행된다:
- [0080] - 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 포함하는 플라스미드 벡터를 적어도 하나의 제한 효소와 접촉시켜 상기 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA를 제조하는 단계, 및 단일 가닥 DNA 어댑터들이 프리마제/폴리머라제 프라이밍 부위를 함유하지 않는 조건 하에서, 관심 DNA 서열을 함유하는 열린 이중

가닥 DNA의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계; 또는, 대안적으로는, 하기에 의해 수행된다:

- [0081] - 관심 DNA 서열을 플랭킹하는 적어도 2개의 프로텔로머라제 표적 서열들을 포함하는 플라스미드 벡터를 프로텔로머라제, 보다 구체적으로는, TelN과 접촉시키는 단계; 그 결과, 상기 관심 DNA 서열을 함유하는 단힌 선형 DNA 주형인 DNA 주형을 얻는 단계. 보다 특정한 일 구현예에서, 단일 가닥 DNA 어댑터들은 서열 XTC를 함유하지 않는다.
- [0082] 보다 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단일 가닥 DNA 어댑터들은 프로텔로머라제 표적 서열을 함유한다. 또 다른 특정한 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단일 가닥 DNA 어댑터들은 프로텔로머라제 표적 서열을 함유하지 않는다. 또 다른 특정한 구현예에서, 단일 가닥 DNA 어댑터들은 프로텔로머라제 표적 서열의 일부를 함유하는데, 상기 프로텔로머라제 표적 서열의 일부는 프로텔로머라제에 의해 인식되지 않는다.
- [0083] 본 명세서에서 사용되는 바와 같이, "프로텔로머라제"는 공유결합적 단힌 선형 DNA 분자를 생성하기 위한, 프로텔로머라제 표적 부위를 포함하는 주형을 절단 및 재결합(rejoining)할 수 있는 임의의 폴리펩티드이다. 따라서, 프로텔로머라제는 DNA 절단 및 리게이션 기능을 가지고 있다. 프로텔로머라제형 활성화(protelomerase-type activity)을 갖는 효소들은 또한 텔로미어 레솔라제(telomere resolvase) (예를 들면, *보렐리아 부르그도르페리(Borrelia burgdorferi)*의 텔로미어 레솔라제)들로 기술되어 왔다. 프로텔로머라제의 일반적인 기질은 원형 이중 가닥 DNA이다. 이 DNA가 프로텔로머라제 표적 부위를 함유하는 경우에, 이 효소는 그 부위에서 DNA를 절단하고 말단들을 리게이팅하여 선형 이중 가닥 공유결합적 단힌 DNA 분자를 생성할 수 있다. 소정의 폴리펩티드가 프로텔로머라제 표적 부위를 포함하는 주형으로부터 단힌 선형 DNA의 제조를 촉매하는 능력은 당업계에서 기술된 임의의 적합한 검정을 사용하여 결정될 수 있다.
- [0084] 본 발명의 방법에 사용하기에 적합한 프로텔로머라제들의 예시는 박테리오파지들, 예컨대, *할로모나스 아쿠아마리나(Halomonas aquamarina)*의 phiHAP-1, *예르시니아 엔테로리티카(Yersinia enterocolitica)*의 PY54, *클렙시엘라 옥시토카(Klebsiella oxytoca)*의 phiK02 및 *비브리오 속(Vibrio sp.)*의 VP882, 및 대장균의 N15의 프로텔로머라제들, 또는 이들의 변이체들을 포함한다.
- [0085] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 프로텔로머라제는 서열 번호 2의 박테리오파지 N15 TelN, 또는 서열 번호 2와 적어도 80% 동일성, 적어도 85% 동일성, 적어도 90% 동일성, 또는 적어도 95% 동일성을 갖는 서열을 포함하는 서열 번호 2의 박테리오파지 N15 TelN의 변이체이다.
- [0086] "프로텔로머라제 표적 서열"은 DNA 주형 내의 존재가 프로텔로머라제의 효소 활성화에 의한 단힌 선형 DNA로의 전환을 가능하게 하는 임의의 DNA 서열이다. 즉, 프로텔로머라제 표적 서열은 공유결합적 단힌 선형 DNA를 형성하기 위해 프로텔로머라제에 의한 이중 가닥 DNA의 절단 및 재-리게이션(re-ligation)에 필요하다. 일반적으로는, 프로텔로머라제 표적 서열은 임의의 완전한 회문 구조 서열(palindromic sequence), 즉, 이중 회전 대칭을 갖는 임의의 이중 가닥 DNA 서열을 포함하며, 본 명세서에서 완전한 역 반복부(perfect inverted repeat)로도 기술된다.
- [0087] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 적어도 2개의 프로텔로머라제 표적 부위들은 완전한 역 반복 DNA 서열들을 포함한다.
- [0088] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 프로텔로머라제 표적 서열은 서열 번호 3에 대해 적어도 80%, 적어도 85%, 적어도 90%, 또는 적어도 95%의 서열 동일성을 갖는 서열을 포함하는, 서열 번호 3의 서열 또는 이의 변이체를 포함한다.
- [0089] 완전한 역 반복부의 길이는 특정 유기체에 따라 달라진다. *보렐리아 부르그도르페리*에서, 완전한 역 반복부는 14개의 염기쌍 길이이다. 다양한 중온성 박테리오파지에서, 완전한 역 반복부는 22개 이상의 염기쌍 길이이다. 또한, 일부 경우에서, 예를 들면, *E. coli* N15에서, 중앙의 완전한 역 회문 구조는 역 반복 서열에 의해 플랭킹되며, 즉, 더 큰 불완전한 역 회문 구조의 일부를 형성한다.
- [0090] 본 발명에 사용된 프로텔로머라제 표적 서열은 바람직하게는 적어도 14개의 염기쌍 길이의 이중 가닥 회문 구조 (완전한 역 반복) 서열을 포함한다.
- [0091] 완전한 역 반복부는 추가의 역 반복 서열들에 의해 플랭킹될 수 있다. 플랭킹하는 역 반복부는 완전한 반복부 또는 불완전한 반복부일 수 있으며, 즉, 완전히 대칭적이거나 부분적으로 대칭적일 수 있다. 플랭킹하는 역 반

복부는 중앙의 회문구조와 인접하거나 인접하지 않을 수 있다. 프로텔로머라제 표적 서열은 적어도 14개의 염기쌍 길이의 완전한 역 반복 서열을 포함하는 불완전한 역 반복 서열일 수 있다.

- [0092] 서열 번호 3의 서열 또는 이의 변이체를 포함하는 프로텔로머라제 표적 서열은 서열 번호 2의 E. coli N15 TeIN 프로텔로머라제 또는 이의 변이체들과 조합하여 사용되는 것이 바람직하다.
- [0093] 상기 기재된 임의의 회문 구조 또는 프로텔로머라제 표적 서열들의 변이체들은 이의 상동체들 또는 돌연변이체들을 포함한다. 돌연변이체들은 천연 서열(native sequence)에 대한 절단, 치환, 또는 결실을 포함한다. 변이체 서열은 DNA 주형 내의 존재가 프로텔로머라제의 효소 활성화에 의한 단편 선형 DNA로의 전환을 가능하게 하는 임의의 DNA 서열이다. 이는 단편 선형 DNA의 형성에 대한 적절한 검정을 사용함으로써 쉽게 결정될 수 있다. 당업계에 기술된 임의의 적합한 검정이 사용될 수 있다. 바람직하게는, 이 변이체는 천연 서열에서 관찰된 것과 필적하는 프로텔로머라제 결합 및 활성을 가능하게 한다. 본 명세서에 기재된 회문 구조 서열들의 바람직한 변이체들의 예시는 완전한 반복 구조를 보존하며, 여전히 단편 선형 DNA를 형성할 수 있는 절단된 회문 구조 서열을 포함한다. 하지만, 변이체 프로텔로머라제 표적 서열들은, 프로텔로머라제 활성화에 대한 기질로서 역할을 할 수 있는 한, 완전한 회문 구조를 보존하지 않도록 변형될 수 있다.
- [0094] 당업자가 상기 설명된 구조적 원리에 기초하여 본 발명에 사용하기에 적합한 프로텔로머라제 표적 서열들을 쉽게 식별할 수 있음을 이해해야 한다. 후보 프로텔로머라제 표적 서열들은 상기 기재된 검정들을 사용하여 단편 선형 DNA의 형성을 촉진하는 능력으로 스크리닝될 수 있다.
- [0095] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, DNA 주형이 관심 DNA 서열을 함유하는 원형 이중 가닥 DNA 주형인 경우에, 단계 (a)는 관심 DNA 서열을 플랭킹하는 적어도 2개의 재조합 효소 인식 부위들을 포함하는 플라스미드 벡터를 부위 특이적 재조합 효소, 보다 구체적으로는, Cre 재조합 효소와 접촉시키는 단계에 의해 수행된다.
- [0096] 플라스미드 벡터 상의 부위 특이적 재조합 효소의 작용은 2개의 재조합 효소 인식 부위들의 재조합을 유발하여 플라스미드 벡터 내 재조합 효소 인식 부위들 사이에 위치했던 관심 DNA 서열을 함유하는 더 작은 원형 이중 가닥 DNA를 생성한다.
- [0097] 본 명세서에서 사용되는 "부위 특이적 재조합 효소"는 재조합 효소 인식 부위들로 알려진, 효소들에 의해 인식되는 특이적 DNA 서열들 사이에서 부위 특이적 재조합을 매개하는 효소들의 패밀리를 지칭한다. 부위 특이적 재조합 효소들의 예시는, 제한 없이, Cre 재조합 효소, F1p 재조합 효소, 람다 인테그라제(lambda integrase), 감마-델타 레졸바제(gamma-delta resolvase), Tn3 레졸바제, Sin 레졸바제, Gin 인버타제(Gin invertase), Hin 인버타제, Tn5044 레졸바제, Tn3 트랜스포사제(Tn3 transposase), 슬리핑 뷰티 트랜스포사제(sleeping beauty transposase), IS607 트랜스포사제, Bxb 1 인테그라제, wBeta 인테그라제, BL3 인테그라제, phiR4 인테그라제, A1 1 8 인테그라제, TGI 인테그라제, MRU 인테그라제, phi370 인테그라제, SPBc 인테그라제, SV1 인테그라제, TP901-1 인테그라제, phiRV 인테그라제, FC1 인테그라제, K38 인테그라제, phiBT1 인테그라제 및 phiC31 인테그라제를 포함한다.
- [0098] "재조합 효소 인식 부위"는 부위 특이적 재조합 효소에 의해 인식되고 재조합 사건에 대한 기질로서 역할을 할 수 있는 뉴클레오티드 서열을 지칭한다. 재조합 효소 인식 부위들의 비제한적인 예시는 FRT, FRT11, FRT71, attP, att, rox, 및 lox 부위들, 예컨대, loxP, lox511, lox2272, lox66, lox71, loxM2, 및 lox5171을 포함한다.
- [0099] 당업자는 일반적인 지식을 사용하여, 각각의 부위 특이적 재조합 효소가 특정한 재조합 효소 인식 부위를 인식하므로, 플라스미드 벡터에 함유된 인식 서열에 따라 플라스미드 벡터로부터 원형 이중 가닥 DNA 주형을 생성하기 위해 상이한 재조합 효소가 사용되어야 함을 알고 있을 것이다.
- [0100] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 부위 특이적 재조합 효소는 Cre 재조합 효소이다. 보다 특정한 일 구현예에서, 재조합 효소 인식 부위는 loxP이다. 더욱 특정한 구현예에서, 부위 특이적 재조합 효소는 Cre 재조합 효소이고 재조합 효소 인식 부위는 loxP이다.
- [0101] 플라스미드 벡터 상의 관심 DNA 서열을 플랭킹하는 서열, 및 단계 (a)에서 DNA 주형을 생성하기 위해 사용되는 방법에 따라, 단계 (b)에서 생성되는 콘카테머 생성물들은 상이한 서열에 의해 플랭킹된 관심 DNA 서열을 포함할 것이다. 예를 들면, 관심 DNA 서열이 오직 플라스미드 벡터 내 프로텔로머라제 표적 서열들에 의해서만 플랭킹되는 경우에, 콘카테머 DNA 내 관심 DNA 서열은 상기 프로텔로머라제 표적 서열들에 의해 플랭킹될 것이다.

또한, 플라스미드 벡터 내 관심 DNA 서열은 한 유형의 반응 (예를 들면, TeIN에 의함)에 의해 DNA 주형을 제조한 다음에, 또 다른 반응 (예를 들면, 제한 효소 분해 및 어댑터 리게이션)에 의해 cIDNA를 생성하는 것을 가능하게 하는 상이한 서열들의 조합들에 의해 플랭킹될 수 있다. 이 특정한 경우에서, 관심 DNA 서열은 제한 자리들에 의해 플랭킹되어야 하고, 프로텔로머라제 표적 부위들에 의해 차례로 플랭킹된다.

[0102] 따라서, 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단계 (b)에서 생성된 증폭된 DNA는 관심 DNA 서열의 반복부를 포함하는 콘카테머 DNA이며, 각각의 반복된 관심 DNA 서열은 제한 자리들, 프로텔로머라제 표적 서열들, 및/또는 재조합 효소 인식 부위들에 의해 플랭킹된다.

[0103] 당업자는, 단편 선형 DNA 형태로 관심 DNA 서열의 텐덤 유닛들을 절제(excision)하는 것은 텐덤 유닛들을 절단하고 이 단편들의 열린 말단들을 폐쇄시켜 공유결합적 단편 분자를 형성하도록 하는 상이한 일반적인 분자 생물학 기술들에 의해 수행될 수 있음을 알 것이다. 이 2개의 단계들은, 예를 들면, 제한 효소 분해한 후에 어댑터 리게이션에 의해 수행되거나, 프로텔로머라제의 작용에 의해 동시에 수행될 수 있다.

[0104] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 콘카테머 DNA가 적어도 제한 자리들에 의해 플랭킹된 관심 DNA 서열의 반복부를 포함하는 경우에, 단계 (c)는: (c.1) 콘카테머 DNA를 적어도 하나의 제한 효소와 접촉시켜 각각 관심 DNA 서열을 함유하는 복수의 열린 이중 가닥 DNA 단편들을 제조하는 단계, 및 (c.2) 상기 열린 이중 가닥 DNA 단편들의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행된다. 제한 효소, 제한 자리, 및 단일 가닥 DNA 어댑터와 관련하여 상기에 제공된 모든 구현예들은 또한 이 구현예에 적용됨을 의미한다. 또한, 당업자는 제한 효소가 주형 cIDNA를 제조하기 위해 사용되는 경우에, 동일한 제한 효소가 단계 (b)에서 제조된 증폭된 DNA로부터 cIDNA를 제조하기 위해 추후 사용될 수 있음을 알 것이다. 주형 cIDNA를 제조하기 위해 단계 (a)에서 사용된 단일 가닥 DNA 어댑터들은 단계 (c)에서 사용된 단일 가닥 DNA 어댑터들과 동일하거나 상이할 수 있다.

[0105] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 콘카테머 DNA가 적어도 프로텔로머라제 표적 서열들에 의해 플랭킹된 관심 DNA 서열의 반복부를 포함하는 경우에, 단계 (c)는 콘카테머 DNA를 프로텔로머라제, 보다 구체적으로는, TeIN과 접촉시키는 단계에 의해 수행된다. 프로텔로머라제 및 프로텔로머라제 표적 부위와 관련하여 상기에 제공된 모든 구현예들은 또한 이 구현예에 적용됨을 의미한다. 또한, 당업자는 프로텔로머라제가 단계 (a)에서 주형 cIDNA를 제조하기 위해 사용되는 경우에, 동일한 프로텔로머라제가 증폭된 DNA로부터 cIDNA를 생성하기 위해 단계 (c)에서 추후 사용될 수 있는 것을 알 것이다.

[0106] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 상기 방법은 단편 선형 발현 카세트 DNA를 제조하기 위한 것이다.

[0107] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단계 (a)는 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 포함하는 플라스미드 벡터를 적어도 하나의 제한 효소와 접촉시켜 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA를 제조하는 단계, 및 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행되고; 단계 (c)는 (c.1) 콘카테머 DNA를 적어도 하나의 제한 효소와 접촉시켜 각각 관심 DNA 서열을 함유하는 복수의 열린 이중 가닥 DNA 단편들을 제조하는 단계, 및 (c.2) 열린 이중 가닥 DNA 단편들의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행된다. 보다 특정한 일 구현예에서, 제한 효소는 점착 말단 또는 평활 말단을 생성한다. 제한 효소가 평활 말단을 생성하는 경우에, 생성된 단편은 평활 말단을 함유하는 어댑터들에 부착될 수 있거나, 대안적으로는 앞서 설명한 바와 같이, dA-테일링된 다음에, 돌출 T를 갖는 어댑터에 부착될 수 있다.

[0108] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단계 (a)는 관심 DNA 서열을 플랭킹하는 적어도 2개의 프로텔로머라제 인식 부위들을 플랭킹하는 적어도 2개의 제한자리들을 포함하는 플라스미드 벡터를 적어도 하나의 제한 효소와 접촉시켜 프로텔로머라제 인식 서열들에 의해 플랭킹된 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA를 제조하는 단계, 및 열린 이중 가닥 DNA의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행되고; 단계 (c)는 콘카테머 DNA를 프로텔로머라제, 보다 구체적으로는, TeIN과 접촉시키는 단계에 의해 수행된다.

[0109] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과

조합하여, DNA 주형이 제한 자리들에 의해 플랭킹된 관심 DNA 서열을 함유하는 원형 이중 가닥 DNA 주형인 경우에, 단계 (a)는 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 플랭킹하는 적어도 2개의 재조합 효소 인식 부위들을 포함하는 플라스미드 벡터를 부위 특이적 재조합 효소, 보다 구체적으로는, Cre 재조합 효소와 접촉시키는 단계에 의해 수행되고; 단계 (c)는 (c.1) 콘카테머 DNA를 적어도 하나의 제한 효소와 접촉시켜 각각 관심 DNA 서열을 함유하는 복수의 열린 이중 가닥 DNA 단편들을 제조하는 단계, 및 (c.2) 열린 이중 가닥 DNA 단편들의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행된다.

[0110] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, DNA 주형이 제한 자리들에 의해 플랭킹된 관심 DNA 서열을 함유하는 원형 이중 가닥 DNA 주형인 경우에, 단계 (a)는 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 플랭킹하는 적어도 2개의 재조합 효소 인식 부위들을 포함하는 플라스미드 벡터를 부위 특이적 재조합 효소, 보다 구체적으로는, Cre 재조합 효소와 접촉시키는 단계에 의해 수행되고; 단계 (c)는 콘카테머 DNA를 프로텔로머라제, 보다 구체적으로는 TeIN과 접촉시키는 단계에 의해 수행된다.

[0111] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단계 (a)는 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 플랭킹하는 적어도 2개의 프로텔로머라제 표적 서열들을 포함하는 플라스미드 벡터를 프로텔로머라제, 보다 구체적으로는, TeIN과 접촉시키는 단계에 의해 수행되고; 단계 (c)는 (c.1) 콘카테머 DNA를 적어도 하나의 제한 효소와 접촉시켜 각각 관심 DNA 서열을 함유하는 복수의 열린 이중 가닥 DNA 단편들을 제조하는 단계, 및 (c.2) 열린 이중 가닥 DNA 단편들의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행된다.

[0112] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단계 (a)는 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 플랭킹하는 적어도 2개의 프로텔로머라제 표적 서열들을 포함하는 플라스미드 벡터를 프로텔로머라제, 예를 들면, TeIN과 접촉시키는 단계에 의해 수행되고; 단계 (c)는 (c.1) 콘카테머 DNA를 적어도 하나의 제한 효소와 접촉시켜 각각 관심 DNA 서열을 함유하는 복수의 열린 이중 가닥 DNA 단편들을 제조하는 단계, 및 (c.2) 열린 이중 가닥 DNA 단편들의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행된다.

[0113] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단계 (a)는 관심 DNA 서열을 먼저 (예를 들면, 프로텔로머라제 B에 대한 것) 플랭킹하는 것과 상이한 적어도 2개의 프로텔로머라제 인식 부위들을 플랭킹하는 적어도 2개의 프로텔로머라제 표적 서열들 (예를 들면, 프로텔로머라제 A에 대한 것)을 포함하는 플라스미드 벡터를 상응하는 프로텔로머라제 (예를 들면, 프로텔로머라제 A에 대한 것)와 접촉시키는 단계에 의해 수행되고; 단계 (c)는 콘카테머 DNA를 상응하는 프로텔로머라제 (예를 들면, 프로텔로머라제 B에 대한 것)와 접촉시키는 단계에 의해 수행된다. 특정한 일 구현예에서, 단계 (a) 또는 단계 (c)에서 프로텔로머라제는 TeIN이다.

[0114] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단계 (a)는 관심 DNA 서열을 플랭킹하는 2개의 프로텔로머라제 표적 부위들을 포함하는 플라스미드 벡터를 프로텔로머라제와 접촉시키는 단계에 의해 수행되고; 단계 (c)는 콘카테머 DNA를 프로텔로머라제와 접촉시키는 단계에 의해 수행된다. 특정한 일 구현예에서, 단계 (a) 또는 단계 (c)에서 프로텔로머라제는 TeIN이다.

[0115] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단계 (a)는 관심 DNA 서열을 플랭킹하는 적어도 2개의 제한 자리들을 포함하고 프로텔로머라제 표적 부위를 포함하지 않는 플라스미드 벡터를 적어도 하나의 제한 효소와 접촉시켜 관심 DNA 서열을 함유하는 열린 가닥 DNA를 제조하는 단계; 및 단일 가닥 DNA 어댑터들이 프로텔로머라제 표적 부위를 함유하지 않는 조건 하에서, 상기 관심 DNA 서열을 함유하는 열린 이중 가닥 DNA의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행되고; 단계 (c)는 (c.1) 콘카테머 DNA를 적어도 하나의 제한 효소와 접촉시켜 각각 관심 DNA 서열을 함유하는 복수의 열린 가닥 DNA 단편들을 제조하는 단계, 및 (c.2) 상기 열린 이중 가닥 DNA 단편들의 양 말단에 단일 가닥 DNA 어댑터들을 부착시키는 단계에 의해 수행된다.

[0116] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 관심 DNA 서열은 발현 카세트이다. 보다 특정한 구현예에서, 발현 카세트는 관심 단백질을 인코딩하는 서열에 작동가능하게 연결된 진핵생물 프로모터, 및 선택적으로는 인헨서 및/또는 진핵생물 전사 종결 서열로 이루어진다.

- [0117] 용어 "발현 카세트"는 하나 이상의 프로모터 또는 인핸서 인자 및 mRNA, miRNA, siRNA 또는 관심 단백질을 인코딩하는 유전자 또는 다른 코딩 서열을 포함하는 DNA 서열을 지칭한다. 발현 카세트는 코딩 서열의 발현을 조절하는 다른 인자, 예컨대, 전사 종결 부위를 더 포함할 수 있다.
- [0118] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 관심 DNA 서열은 역 말단 반복부(ITRs)에 의해 플랭킹되는 발현 카세트를 포함한다. ITRs는 발현 카세트와 임의의 적합한 거리에 있을 수 있으며, 예를 들면, ITRs는 발현 카세트에 직접 연결되거나 발현 카세트로부터 1 내지 50 뉴클레오티드, 50 내지 200 뉴클레오티드, 200 내지 1000 뉴클레오티드의 거리에 있을 수 있다. 따라서, 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 관심 DNA는 1 내지 50 뉴클레오티드 거리에 있는 역 말단 반복부(ITRs)를 포함한다.
- [0119] 본 명세서에서 사용되는 바와 같이, 용어 "말단 반복부(terminal repeat)" 또는 "TR"은 적어도 하나의 최소로 요구되는 복제 원점, 및 회문 구조의 헤어핀 구조를 포함하는 영역을 포함하는 임의의 바이러스성 말단 반복부 또는 합성 서열을 포함한다. Rep-결합 서열(Rep-binding sequence; "RBS") (RBE (Rep-binding element)로도 지칭됨) 및 말단 분해 부위(terminal resolution site; "TRS")는 함께 "최소로 요구되는 복제 원점"을 구성하며, 따라서, TR는 적어도 하나의 RBS 및 적어도 하나의 TRS를 포함한다. 폴리뉴클레오티드 서열의 주어진 스트레치(stretch) 내에 서로 역 상보적인 TRs은 일반적으로 각각 "역 말단 반복부" 또는 "ITR"로 지칭된다. 바이러스의 맥락에서, ITRs은 복제, 바이러스 패키징(virus packaging), 통합(integration) 및 프로바이러스 구제(provirus rescue)를 매개한다.
- [0120] 당업자는 복합체 c1DNA 구조에서 2개 초과 ITRs 또는 비대칭 ITR 쌍들이 존재할 수 있음을 이해할 것이다. ITR는 AAV ITR 또는 비-AAV ITR(non-AAV ITR)일 수 있고, AAV ITR 또는 비-AAV ITR로부터 유래할 수 있다. 예를 들면, ITR은 파르보바이러스(parvovirus) 및 데펜도바이러스(dependovirus) (예를 들면, 개 파르보바이러스, 소 파르보바이러스, 마우스 파르보바이러스, 돼지 파르보바이러스, 인간 파르보바이러스 B-19)를 포함하는 파르보비리다에과로부터 유래할 수 있거나, SV40 복제 원점으로 역할을 하는 SV40 헤어핀은 절단, 치환, 결실, 삽입 및/또는 부가에 의해 더 변형될 수 있는 ITR로서 사용될 수 있다. 파르보비리다에과 바이러스들은 하기 2개의 아과(subfamily)들로 이루어진다: 척추동물물을 감염시키는 파르보리나에(Parvovirinae), 및 무척추동물물을 감염시키는 덴소비리나에(Densovirinae). 데펜도파르보바이러스(Dependoparvovirus)들은, 인간, 영장류, 소, 개, 말 및 양의 종들을 포함하나, 이에 제한되지 않는 척추동물 숙주들에서 복제를 할 수 있는 아데노-연관 바이러스(adenovirus-associated viruses; AAV)의 바이러스과를 포함한다. 본 명세서에서 편의를 위해, c1DNA 벡터 내 발현 카세트의 5' (상류)에 위치한 ITR을 "5' ITR" 또는 "좌측 ITR"로 지칭하고, c1DNA 벡터 내 발현 카세트의 3' (하류)에 위치한 ITR을 "3' ITR" 또는 "우측 ITR"로 지칭한다.
- [0121] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 관심 DNA 서열은 서열 서열 번호 8 또는 서열 번호 9의 적어도 하나의 역 말단 반복부에 의해 플랭킹된 발현 카세트를 포함한다.
- [0122] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 관심 DNA 서열은 서열 서열 번호 8의 5' 역 말단 반복부 및/또는 서열 서열 번호 9의 3' 역 말단 반복부에 의해 플랭킹된 발현 카세트를 포함한다.
- [0123] 본 발명의 제1 측면의 방법의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 관심 DNA 서열은 적어도 하나의 DD-ITR에 의해 플랭킹된 발현 카세트를 포함한다. "DD-ITR"는 Xiao X. et al., "A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle", 1997, J Virol., vol. 71(2), pp. 941-948에 개시된 플랭킹하는 D 인자들을 갖는 ITR을 지칭한다.
- [0124] 제조된 c1DNA들이 정제되는 상기 방법의 단계 (d)와 관련하여, 당업자는 핵산, 특히, c1DNA를 정제하기에 적합한 임의의 공지된 방법이 사용될 수 있음을 알고 있다.
- [0125] 앞서 언급한 바와 같이, 제2 측면에서 본 발명은 제1 측면에서 정의된 방법에 따라 얻을 수 있는 단힌 선형 DNA를 제공한다.
- [0126] 본 발명의 목적을 위해, 표현 "얻을 수 있는", "얻어지는" 및 이들과 동등과 표현들은 상호교환적으로 사용되며, 어느 경우에서, 표현 "얻을 수 있는"은 표현 "얻어지는"을 포함한다. 본 발명의 제1 측면 하에서 제공된 모든 구현예들은 또한 본 발명의 제2 측면의 단힌 선형 DNA의 구현예들이다.

- [0127] 본 발명의 제2 측면의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 단편 선형 DNA는 하나 이상의 발현 카세트를 포함한다.
- [0128] 본 발명의 제2 측면의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 발현 카세트는 mRNA, miRNA, siRNA, 또는 단백질을 인코딩하는 서열에 작동가능하게 연결된 진핵생물 프로모터를 포함한다.
- [0129] 본 발명의 제2 측면의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 발현 카세트는 진핵생물 전사 종결 서열을 더 포함한다.
- [0130] 본 발명의 제2 측면의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 발현 카세트는 하기로 이루어진 군으로부터 선택된 하나 이상의 박테리아 서열 또는 벡터 서열이 결합되어 있다:
- [0131] (i) 박테리아성 복제 원점;
- [0132] (ii) 박테리아성 선택 마커; 및
- [0133] (iii) 메틸화되지 않은 CpG 모티프.
- [0134] 본 발명의 제2 측면의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 관심 DNA 서열은 역 말단 반복부(ITRs)에 의해 플랭킹된 발현 카세트를 포함한다.
- [0135] 이전에 언급한 바와 같이, 본 발명은 또한 요법에 사용하기 위한 제1 측면에 따른 단편 선형 DNA를 제공한다.
- [0136] 본 발명의 cDNA는, 특히, DNA 백신 또는 유전자 요법에서, 숙주 세포에서 시험관 내 발현에 사용될 수 있다. DNA 백신은 일반적으로 감염성 유기체 DNA의 변형된 형태를 인코딩한다. DNA 백신은 대상체에 투여되어 감염성 유기체의 선택된 단백질을 발현하고, 일반적으로 보호되는 이 단백질에 대한 면역 반응을 개시한다. DNA 백신은 또한 암 면역요법 접근에서 종양 항원을 인코딩할 수 있다.
- [0137] DNA 백신은 암, 알레르기, 독성, 및 감염을 포함하나, 이에 제한되지 않는 수많은 병태(condition)들의 치료 또는 예방을 위한, 항원을 인코딩하는 핵산 서열을 포함할 수 있는데, 상기 감염은 제한 없이, 예컨대, 진균, 인간 유두종 바이러스(HPV), HIV, HSV2/HSV1, 인플루엔자 바이러스 (A형, B형 및 C형), 폴리오 바이러스(Polio virus), RSV 바이러스, 리노바이러스(Rhinoviruses), 로타바이러스(Rotaviruses), A형 간염 바이러스 (Hepatitis A virus), 노르워크 바이러스 그룹(Norwalk Virus Group), 엔테로바이러스, 아스트로바이러스 (Astroviruses), 홍역 바이러스, 파라인플루엔자 바이러스(Parainfluenza virus), 뎀프스 바이러스(Mumps virus), 수두-대상포진 바이러스(Varicella-Zoster virus), 사이토메갈로바이러스(Cytomegalovirus), 엡스타인-바 바이러스(Epstein-Barr virus), 아데노바이러스, 풍진 바이러스(Rubella virus), 인간T-세포 림프종 I형 바이러스(Human T-cell Lymphoma type I virus; HTLV-I), B형 간염 바이러스(HBV), C형 간염 바이러스(HCV), D형 간염 바이러스, 폭스 바이러스(Pox virus), 마버그(Marburg) 바이러스 및 에볼라 바이러스를 포함하는 바이러스; 마이코박테리움 튜베르쿨로시스(Mycobacterium tuberculosis), 클라미디아(Chlamydia), 나이세리아 고노로이애(Neisseria gonorrhoeae), 쉬겔라(Shigella), 살모넬라, 비브리오 콜레라(Vibrio cholerae), 트레포네마 팔리둠(Treponema pallidum), 슈도모나스, 보르데텔라 페르투스시스(Bordetella pertussis), 브루셀라 (Brucella), 프랜시셀라 툴라렌시스(Franciscella tularensis), 헬리코박터 파이로리, 렙토스피라 인테로간스 (Leptospira interrogans), 레지오넬라 뉴모필라(Legionella pneumophila), 예르시니아 페스티스(Yersinia pestis), 스트렙토코커스 (A형 및 B형), 뉴모코커스(Pneumococcus), 메닝고코커스(Meningococcus), 헤모필루스 인플루엔자 (b형), 톡소플라즈마 곤디(Toxoplasma gondii), 캄필로박테리오시스(Campylobacteriosis), 모락셀라 카타랄리스(Moraxella catarrhalis), 도노바노시스(Donovanosis), 및 액티노마이코시스(Actinomycosis)를 포함하는 박테리아; 칸디디아시스(Candidiasis) 및 아스페르길로시스(Aspergillosis)를 포함하는 진균 병원체; 테니아(Taenia), 흡충(Flukes), 회충, 아메바증(Amoebiasis), 지아디아증(Giardiasis), 크립토스포리디움 (Cryptosporidium), 주혈흡충(Schistosoma), 뉴모시스티스 카리니(Pneumocystis carinii), 트리코모나스증 (Trichomoniasis) 및 선모충증(Trichinosis)을 포함하는 기생충 병원체에 의한 감염이 있다.
- [0138] DNA 백신은 아데노비리다에 (예를 들면, 인간 아데노바이러스를 포함), 헤르페스비리다에 (예를 들면, HSV-1, HSV-2, EBV, CMV 및 VZV 포함), 파포바비리다에 (예를 들면, HPV를 포함), 폭스비리다에 (예를 들면, 천연두 및 백시니아 포함), 파르보비리다에 (예를 들면, 파르보바이러스 B19 포함), 레오비리다에 (예를 들면, 로타바이러스 포함), 코로나비리다에 (예를 들면, SARS 포함), 플라비비리다에 (예를 들면, 황열(yellow fever), 웨스트

나일 바이러스(West Nile virus), 땡기열 바이러스, C형 간염 바이러스 및 진드기 매개 뇌염(tick-borne encephalitis) 바이러스 포함), 피코르나비리다에 (예를 들면, 폴리오 바이러스, 리노바이러스, 및 A형 간염 바이러스 포함), 토가비리다에 (예를 들면, 풍진 바이러스(rubella virus) 포함), 필로비리다에 (예를 들면, 마버그 바이러스 및 에볼라 바이러스 포함), 파라믹소비리다에 (예를 들면, 파라인플루엔자 바이러스, 호흡기 세포융합 바이러스(respiratory syncytial virus), 머프스 바이러스 및 홍역 바이러스 포함), 랩토비리다에 (예를 들면, 광견 바이러스 포함), 분야비리다에 (예를 들면, 한탄 바이러스(Hantaan virus) 포함), 오르토믹소비리다에 (예를 들면, A형, B형 및 C형 인플루엔자 바이러스), 레트로비리다에 (예를 들면, HIV 및 HTLV 포함), 및 헤파드나비리다에 (예를 들면, B형 간염 바이러스 포함)의 구성원으로부터 항원을 인코딩하는 핵산 서열을 포함할 수 있다.

[0139] 이 항원은 수의학적 질병을 일으키는 병원체로부터 유래할 수 있으며, 특히, 예를 들면, 레오바이러스 (예컨대, 아프리카 마역(African Horse sickness) 바이러스, 블루팅(Bluetongue) 바이러스) 및 헤르페스 바이러스 (말 헤르페스 바이러스 포함)을 포함하는 바이러스 병원체로부터 유래할 수 있다. 이 항원은 구제역 바이러스, 진드기 매개 뇌염 바이러스, 땡기열 바이러스, SARS, 웨스트 나일 바이러스 및 한탄 바이러스로부터 유래한 것일 수 있다. 이 항원은 면역결핍 바이러스로부터 유래할 수 있으며, 예를 들면, SIV 또는 고양이 면역결핍 바이러스 (feline immunodeficiency virus)로부터 유래할 수 있다.

[0140] 본 발명의 방법에 의해 제조된 cDNA들은 또한 종양 항원을 인코딩하는 핵산 서열을 포함할 수 있다. 종양 관련 항원들의 예시는, 암-고환 항원, 예컨대, MAGE 패밀리의 구성원들 (MAGE 1, 2, 3 등), NY-ESO-1 및 SSX-2, 분화 항원, 예컨대, 티로시나제, gp100, PSA, Her-2 및 CEA, 돌연변이된 자기 항원, 및 바이러스성 종양 항원, 예컨대, 발암성 HPV 유형들의 E6 및/또는 E7을 포함하나, 이에 제한되지 않는다. 특정 종양 항원들의 추가 예시는 MART-1, Melan-A, p97, 베타-HCG, GaINac, MAGE-1, MAGE-2, MAGE-4, MAGE-12, MUC1, MUC2, MUC3, MUC4, MUC18, CEA, DDC, P1A, EpCam, 흑색종 항원 gp75, Hker 8, 고분자량 흑색종 항원, K19, Tyr1, Tyr2, pMel 17 유전자 패밀리의 구성원들, c-Met, PSM (prostate mucin antigen), PSMA (prostate specific membrane antigen), 전립선 분비 단백질, 알파-태아단백질, CA125, CA19.9, TAG-72, BRCA-1 및 BRCA-2 항원을 포함한다.

[0141] 또한, 본 발명의 방법은 다른 유형의 치료적 cDNA, 예를 들면, 유전자 요법에 사용되는 치료적 cDNA를 제조할 수 있다. 예를 들면, 이러한 DNA 분자들은, 대상체가 유전자의 기능장애 버전에 의해 유발된 유전성 질환을 갖는 경우에 기능적 유전자를 발현하는데 사용될 수 있다. 이러한 질환들의 예시는 뒤센 근이영양증, 낭포성 섬유증, 고셔병(Gaucher's Disease), 및 아데노신 탈아미노효소(adenosine deaminase; ADA) 결핍증을 포함한다. 유전자 요법이 유용할 수 있는 다른 질병들은, 염증성 질환, 자가면역 질환, 예컨대, AIDS, 암, 신경계 질환, 심혈관 질환, 고콜레스테롤혈증, 다양한 빈혈, 지중해빈혈, 및 혈우병을 포함한 다양한 혈액 질환, 및 폐기종을 포함하는 만성 감염 질환(chronic and infectious diseases)을 포함한다. 고품종 종양의 치료를 위해, 독성 펩티드 (즉, 리신(ricin), 디프테리아 독소(diphtheria toxin) 및 코브라 베놈 인자(cobra venom factor)와 같은 화학요법제)를 인코딩하는 유전자, p53과 같은 종양 억제 유전자, 형질전환 암유전자(transforming oncogene)에 대해 안티센스인 mRNA 서열을 코딩하는 유전자, 종양 괴사 인자(tumor necrosis factor; TNF)와 같은 항신생물 펩티드 및 기타 사이토카인, 또는 형질전환 암유전자의 전이우성 음성 돌연변이체(transdominant negative mutant)가 발현될 수 있다.

[0142] 다른 유형의 치료적 cDNA 또한 본 발명의 방법에 의한 제조에 고려된다. 예를 들면, 활성 RNA 형태, 예를 들어, siRNA(small interfering RNA)로 전사되는 cDNA들은 본 발명의 방법에 따라 제조될 수 있다.

[0143] 앞서 언급한 바와 같이, 본 발명은 또한 본 발명의 제2 측면에 따른 단편 선형 DNA의 치료적 유효량 및 약제학적으로 허용가능한 담체 또는 부형제를 포함하는 약제학적 조성물을 제공한다.

[0144] 본 명세서에서 사용되는 표현 "치료적 유효량"은 투여시 다루지는 질병의 하나 이상의 증상들의 발생을 예방하거나, 어느 정도 완화하기에 충분한 cDNA의 양을 지칭한다. 물론, 본 발명에 따라 투여되는 제제의 구체적인 용량은 투여되는 cDNA, 투여 경로, 치료되는 특정 상태, 및 유사한 고려사항들을 포함하는 경우를 둘러싼 특정한 상황들에 의해 결정될 것이다.

[0145] 표현 "약제학적 조성물"은 인간이 아닌 동물 (즉, 수의학적 조성물) 뿐만 아니라 인간을 위한 조성물 둘 모두를 포함한다.

[0146] 표현 "약제학적으로 허용가능한 담체 또는 부형제"는 약제학적으로 허용가능한 물질, 조성물 또는 비히클(vehicle)을 지칭한다. 각 구성성분은 약제학적 조성물의 다른 성분들과 양립할 수 있다는 의미에서 약제학적으

로 허용가능해야 한다. 또한, 과도한 독성, 자극, 알레르기 반응, 면역원성 또는 다른 문제 또는 합병증 없이 인간 및 인간이 아닌 동물의 조직 또는 기관과 접촉하여 사용하기에 적합하고 합리적인 위험/이익 비에 상응해야 한다.

- [0147] 적합한 약제학적으로 허용가능한 부형제들의 예시는 용매, 분산 매질, 희석제, 또는 다른 액체 비히클, 분산 또는 현탁 보조제, 표면 활성제, 등장화제, 증점제 또는 유화제, 보존제, 고체 바인더, 운할제 등이다. 임의의 기존 부형제 매질이 예컨대, 임의의 바람직하지 않은 생물학적 효과를 생성하거나 약제학적 조성물의 임의의 다른 구성성분(들)과 유해한 방식으로 상호작용함으로써 물질 또는 이의 유도체들과 양립할 수 없는 경우를 제외하고, 이의 사용은 본 발명의 범위 내인 것으로 고려된다.
- [0148] 본 발명의 약제학적 조성물에서 단힌 선형 DNA, 약제학적으로 허용가능한 부형제, 및/또는 임의의 추가 성분의 상대량은 치료받는 대상체의 신원, 크기, 및/또는 상태에 따라, 추가되는 조성물이 투여될 경로에 따라 달라질 수 있다.
- [0149] 약제학적 조성물의 제조에 사용되는 약제학적으로 허용가능한 부형제는 불활성 희석제, 분산 및/또는 과립화제, 표면 활성제 및/또는 유화제, 봉해제, 결합제, 보존제, 완충제, 운할제, 및/또는 오일을 포함하나, 이에 제한되지 않는다. 착색제, 코팅제, 감미제, 및 향미제와 같은 부형제는 조제자의 판단에 따라 조성물에 존재할 수 있다.
- [0150] 본 발명의 방법에 따라 제조된 단힌 선형 DNA를 함유하는 약제학적 조성물들은 임의의 투여 형태, 예를 들면, 고체 또는 액체로 존재할 수 있고, 적합한 투여 경로, 예를 들면, 경구, 비경구, 직장, 국소, 비강내 또는 설하 경로로 투여될 수 있으며, 바람직한 투여 형태, 예를 들면, 국소 제형(연고, 크림, 리포겔(lipogel), 하이드로겔 등), 점안액, 에어로졸 스프레이(aerosol sprays), 주사가능한 용액, 삼투 펌프(osmotic pumps) 등의 제형화에 필요한 약제학적으로 허용가능한 부형제를 포함할 것이다.
- [0151] 예시적인 희석제는 탄산 칼슘, 탄산 나트륨, 인산 칼슘, 인산 이칼슘(dicalcium phosphate), 황산 칼슘, 인산 수소 칼슘, 인산 나트륨 락토스, 수크로스, 셀룰로스, 미세결정질 셀룰로스, 카올린(kaolin), 만니톨, 소르비톨, 이노시톨, 염화 나트륨, 건조 전분, 옥수수 전분, 분당(powdered sugar), 및 이들의 조합들을 포함하나, 이에 제한되지 않는다.
- [0152] 예시적인 과립화 및/또는 분산제는 감자 전분, 옥수수 전분, 사피오카 전분, 전분 글리콜산 나트륨, 알긴산, 구아 검(guar gum), 시트루스 펄프(citrus pulp), 아가(agar), 벤토나이트, 셀룰로스, 목제품(wood products), 천연 스펀지, 양이온 교환 수지, 탄산 칼슘, 실리케이이트, 탄산 나트륨, 가교 폴리비닐피롤리돈(크로스포비돈(crospovidone)), 카르복시메틸 전분 나트륨(전분 글리콜산 나트륨), 카르복시 메틸 셀룰로스, 가교 카르복시메틸 셀룰로스(크로스카멜로스(croscarmellose)) 나트륨, 메틸셀룰로스, 전호화 전분(starch 1500), 미세결정질 전분, 수불용성 전분, 카르복시메틸 셀룰로스 칼슘, 마그네슘 알루미늄 실리케이이트(Veegum), 라우릴 황산 나트륨, 4차 암모늄 화합물, 및 이들의 조합물을 포함하나, 이에 제한되지 않는다.
- [0153] 예시적인 결합 부형제는 전분(예를 들면, 옥수수 전분 및 전분 페이스트); 당(예를 들면, 수크로스, 글루코스, 텍스트로스, 텍스트린, 당밀(molasses), 락토스, 락티톨(lactitol), 만니톨); 천연 검 및 합성 검(예를 들면, 아카시아, 아긴산 나트륨, 아이리쉬 모스(irish moss)의 추출물, 판와 검(panwar gum), 가티 검(ghatti gum), 이사폴 허스크(isapol husks)의 점액(mucilage), 카르복시메틸 셀룰로스, 메틸 셀룰로스, 에틸 셀룰로스, 하이드록시에틸셀룰로스, 하이드록시프로필 셀룰로스, 하이드록시프로필 메틸셀룰로스, 미세결정질 셀룰로스, 셀룰로스 아세트산, 폴리비닐피롤리돈), 마그네슘 알루미늄 실리케이이트(Veegum), 및 라치 아라보갈락탄(larch arabogalactan)); 알기네이트; 폴리에틸렌 옥사이드; 폴리에틸렌 글리콜; 무기 칼슘 염; 규산; 폴리메타크릴레이트; 왁스; 물; 알코올; 및 이들의 조합들을 포함하나, 이에 제한되지 않는다.
- [0154] 예시적인 보존제는 항산화제, 킬레이팅제, 항바이러스성 보존제, 항균성 보존제, 알코올 보존제, 산성 보존제, 및 기타 보존제를 포함할 수 있다. 예시적인 항산화제는, 알파 토크오페롤, 아스코르브산, 아스코르빌 팔미테이트, 아스코르빌 스테아레이트, 아스코르빌 올레레이트, 부틸화 하이드록시 아니솔, 부틸화 하이드록시톨루엔, 모노티오글리세롤(monothioglycerol), 메타중아황산 칼륨(potassium metabisulfite), 프로피온산, 프로필 갈레이트(propyl gallate), 아스코르브산 나트륨, 중아황산 나트륨, 메타중아황산 나트륨, 아황산 나트륨을 포함하나, 이에 제한되지 않는다. 예시적인 킬레이팅제는 에틸렌디아민테트라아세트산(EDTA), 시트르산 일수화물, 에데트산 이나트륨(disodium edetate), 에데트산 이칼륨, 푸마르산, 말산, 인산, 에데트산 나트륨, 타르타르산, 및 에데트산 삼나트륨을 포함한다.

- [0155] 예시적인 완충제는 예를 들면, 시트레이트 완충 용액, 아세테이트 완충 용액, 포스페이트 완충 용액, 염화 암모늄, 탄산 칼슘, 염화 칼슘, 시트르산 칼슘, 글루비온산 칼슘, 글루세프산 칼슘, 글루콘산 칼슘, D-글루콘산, 글리세로인산 칼슘, 락트산 칼슘, 프로판산, 레볼린산 칼슘, 펜탄산, 이염기성 인산 칼슘, 인산(phosphoric acid), 삼염기성 인산 칼슘, 인산 수산화 칼슘(calcium hydroxide phosphate), 아세트산 칼슘, 염화 칼슘, 글루콘산 칼슘, 칼슘 혼합물, 이염기성 인산 칼슘, 일염기성 인산 칼슘, 인산 칼슘 혼합물, 아세트산 나트륨, 이탄산 나트륨, 염화 나트륨, 시트르산 나트륨, 락트산 나트륨, 이염기성 인산 나트륨, 일염기성 인산 나트륨, 인산 나트륨 혼합물, 트로메타민(tromethamine), 수산화 마그네슘, 수산화 알루미늄, 알긴산, 발열물질 무함유 물(pyrogen-free water), 등장성 식염수, 링거 용액(Ringer's solution), 에틸 알코올, 및 이들의 조합들을 포함하나, 이에 제한되지 않는다.
- [0156] 예시적인 윤활제는 스테아르산 마그네슘, 스테아르산 칼슘, 스테아르산, 실리카, 탈크, 몰트(malt), 그리세릴 베헤네이트(glyceryl behenate), 수소화된 식물성 오일, 폴리에틸렌 글리콜, 벤조산 나트륨, 아세트산 나트륨, 염화 나트륨, 류신, 라우릴 황산 나트륨, 및 이들의 조합들을 포함하나, 이에 제한되지 않는다.
- [0157] 상기 개시된 바와 같이, 본 발명은 제5 측면에서 관심 DNA 서열의 반복부를 포함하는 콘카테머 DNA를 제공하며, 이때, 반복된 관심 DNA 서열 각각은 적어도 제조할 효소 인식 부위들, 또는 대안적으로는, 적어도 제한 자리들 및 프로토펠모라제 표적 서열들에 의해 플랭킹된다.
- [0158] 본 발명의 제5 측면의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 반복된 관심 DNA 서열 각각은 ITRs에 의해 추가로 플랭킹된다.
- [0159] 본 발명의 제5 측면의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 콘카테머 DNA는 관심 DNA 서열의 10 이상의 반복부를 포함한다.
- [0160] 본 발명의 제5 측면의 특정한 일 구현예에서, 선택적으로는 상기 또는 하기에 제공된 임의의 구현예들과 조합하여, 콘카테머 DNA는 적어도 5kb 크기이다.
- [0161] 발명의 설명 및 청구범위에 걸쳐, 단어 "포함하다" 및 상기 단어의 변형어는 다른 기술적 특징, 첨가물, 구성성분, 또는 단계를 배제하는 것을 의도하지 않는다. 또한, 단어 "포함하다"는 "로 이루어지는"의 경우를 포함한다. 본 발명의 추가 목적, 이점 및 특징은 본 설명을 검토할 때 당업자에게 명백해지거나, 본 발명의 실시예에 의해 학습될 수 있다. 하기 실시예 및 도면은 예시로서 제공되며, 본 발명의 제한하는 것을 의도하지 않는다. 도면과 관련되고 청구범위의 괄호 안에 있는 참조 기호는 청구범위의 명료함을 높이기 위한 것일 뿐, 청구범위를 제한하는 것으로 해석되어서는 안 된다. 또한, 본 발명은 본 명세서에 기재된 특정 바람직한 구현예들의 모든 가능한 조합들을 포함한다.
- [0162] **실시예**
- [0163] **실시예 1: TthPrimPol 기반 DNA 주형 증폭은 랜덤 프라이머들보다 더 높은 서열 정확성을 제공한다**
- [0164] 관심 DNA 서열을 함유하는 플라스미드 벡터(서열 번호 23을 갖는 pUC57-Kan_Telo-CMVEGFP) 10 ng의 RCA 증폭을 TthPrimPol 또는 랜덤 프라이머들(RPs)로 수행했다. 반응 조건은 총 반응 부피 100 μ l에서 30 $^{\circ}$ C에서 6 h 및 65 $^{\circ}$ C에서 10 min이었다.
- [0165] 도 1에 나타난 바와 같이, TthPrimPol로 프라이밍된 증폭은 DNA 주형의 부재 하에 어떠한 증폭 생성물도 생성하지 않았다(좌측 컬럼, NTC). 반면에, RPs로 프라이밍된 증폭은 DNA 주형의 부재하에 높은 DNA 생산량을 생산했는데, 이는 TthPrimPol 프라이밍이 고도로 특이적인 증폭 반응을 제공하는 것을 시사한다. 또한, TthPrimPol 프라이밍에 의해 생산된 DNA 생산량은 DNA 주형의 존재 하에서 RPs에 의해 생산된 DNA 생산량과 동일한 정도의 규모였다(우측 컬럼, 플라스미드).
- [0166] 마지막으로, 프라이머 TthPrimPol 또는 RPs에 의한 증폭 생성물을 정제하고 표준 프로토콜을 사용한 일루미나 기술(Illumina technology) (5백만개 판독 쌍 2 x 150 bps)을 사용하여 서열 분석했다. 도 2에 나타난 바와 같이, 서열 분석 결과의 생물정보학적 분석은 TthPrimPol 기반 증폭은 80%의 사용가능한 판독을 생산하는 반면에, 랜덤 합성 프라이머들로 증폭하는 것은 오직 66%의 사용가능한 판독을 생산하여, 프라이밍 방법인 TthPrimPol의 사용을 기반으로 할 때, 더 적은 DNA 산물(DNA artifact)를 생성하는 것을 확인했다.
- [0167] 결론적으로, 상기 결과는 치료적 폴리뉴클레오티드의 생산을 위한 방법에서 TthPrimPol의 사용은 증폭 단계에서 부여하는 더 높은 정확성으로 인해 더 유리할 수 있음을 시사한다.

- [0168] **실시예 2: TthPrimPol 프라이밍된 증폭에 기초한 cDNA의 제조**
- [0169] 상기 실시예에서 기술한 바와 같이, 플라스미드 pUC57-Kan_TEL0-CMVEGFP를 RCA에 의해 증폭시켰다. 이어서, 증폭 생성물을 DNA 정제의 표준 프로토콜을 사용하여 정제했다.
- [0170] 이어서, 증폭 생성물 (DNA 콘카테머)을 특히 하기 조건에서 제조업체의 프로토콜에 따라 프로텔로머라제 (Tel N)로 처리했다:
- [0171] - 반응 부피: 1008 μ l
- [0172] -DNA 입력: 증폭 DNA 350 μ g
- [0173] - TelN 입력: 125 μ l (625 유닛)
- [0174] - 반응 시간: 30 $^{\circ}$ C에서 30 min, 이어서 75 $^{\circ}$ C에서 5min
- [0175] 마지막으로, TelN과 반응의 생성물을 제한 효소로 분해하고 엑소뉴클레아제로 처리하여 프로텔로머라제 반응으로부터 생성된 원치 않는 DNA 단편들을 제거했다.
- [0176] 제조업체의 지침에 따라 HindIII 및 EcoRI 분해를 수행했다. 간략하게는 하기와 같다:
- [0177] -반응 부피: 1453 μ l
- [0178] -DNA 입력: TelN로 분해된 증폭 DNA 350 μ g
- [0179] - EcoRI 입력: 150 μ l (1500 유닛)
- [0180] - HindIII 입력: 150 μ l (1500 유닛)
- [0181] - 반응 시간: 37 $^{\circ}$ C에서 60 min, 65 $^{\circ}$ C에서 15 min
- [0182] 제조업체의 지침에 따라 엑소뉴클레아제 III 분해를 수행했다. 간략하게는 하기와 같다:
- [0183] -반응 부피: 1628 μ l
- [0184] -DNA 입력: TelN, HindIII 및 EcoRI로 분해된 증폭 DNA 350 μ g
- [0185] - ExoIII 입력: 6 μ l (600 유닛)
- [0186] - 반응 시간: 37 $^{\circ}$ C에서 45 min 및 80 $^{\circ}$ C에서 20 min
- [0187] 도 3에 나타난 바와 같이, TthPrimPol로 프라이밍된 증폭은 TelN의 작용에 의해 cDNA로 성공적으로 변환될 수 있는 증폭 생성물을 생성했다.
- [0188] 마지막으로, 생성된 DNA를 하기 표준 절차에 따라 정제하고 80,5 μ g의 정제 생산량을 제공했다(106 ng/ μ l에서 760 μ l). 생성된 cDNA 생성물의 품질을 생어 서열 분석으로 분석했다.
- [0189] **실시예 3: TelN-생성 cDNA 주형으로부터 cDNA 제조**
- [0190] TthPrimPol이 표적 서열 "XTC"를 함유하지 않는 최소의 단일 가닥 루프들로 cDNA들을 프라이밍할 수 있는지 테스트하기 위해, 발현 카세트를 플랭킹하는 2개의 프로텔로머라제 인식 서열들을 함유하는 플라스미드 pUC57-Kan_TEL0-CMVEGFP를 실시예 2에 기재된 바와 같이 프로텔로머라제로 처리하여, 발현 카세트를 포함하는 cDNA들을 생성했다.
- [0191] 생성된 cDNA 1 ng 또는 초기 플라스미드를 실시예 1에 기재된 바와 같이 TthPrimPol 또는 랜덤 프라이머들로 프라이밍된 RCA 증폭에 적용하였고, 증폭 생성물을 정량화했다. 도 4의 A에 나타난 바와 같이, TthPrimPol은 TelN-생성 cDNA들의 증폭을 가능하게 하였지만, 랜덤 프라이머들로 증폭한 것보다 더 낮은 생산량을 가졌다.
- [0192] 이어서, 증폭 생성물 (DNA 콘카테머)에 TelN를 처리하여 다시 한번 cDNA들을 생성하였고, 이들을 실시예 2에 기재된 바와 같이 제한 효소 및 엑소뉴클레아제 처리에 의해 아가로스 겔에서 분석했다.
- [0193] 랜덤 프라이머 증폭이 더 높은 생산량의 DNA 생성물을 생성하였지만 (도 4의 A), TelN에 의해 증폭된 DNA로부터 생성된 cDNA의 분석은 TthPrimPol이 표적 DNA 서열을 함유하는 더 높은 수준의 cDNA들을 생성하는 것을 가능하게 함을 나타냈다 (도 4의 B).
- [0194] 이러한 결과는 TthPrimPol이 최소 어댑터 서열들을 함유하는 (즉, TthPrimPol의 인식 부위를 함유하지 않음)

cDNA를 프라이밍할 수 있을 뿐만 아니라, 보다 중요하게는, 랜덤 프라이머들보다 더 높은 품질의 증폭 생성물을 생성할 수 있어, 최종 cDNA들의 제조 효율을 증가시키는 것을 입증한다.

- [0195] **실시예 4: 본 발명의 방법에서 제조된 cDNA들의 기능적 검증**
- [0196] 상기 실시예 3에 개시된 바와 같이 제조되고, eGFP(enhanced green fluorescence protein)의 코딩 서열을 함유하는 cDNA들, eGFP의 코딩 서열을 함유하는 플라스미드 벡터, 빈 벡터(empty vector)를 Heinrich, M. et al. "Linear closed mini DNA generated by the prokaryotic cleaving-joining enzyme TelN is functional in mammalian cells", J Mol Med, 2002, vol. 80, pp. 648-654에 기재된 바와 같이 HEK293 세포에 일시적으로 형질감염시켰다.
- [0197] 표준 현미경 프로토콜에 따라 세포들의 형광 강도를 측정하기 위해 세포들을 현미경으로 24 h 및 48 h에 분석했다.
- [0198] 도 4 및 5에 나타난 바와 같이, 본 발명의 방법에 따라 합성된 cDNA들로 형질감염된 세포들은 eGFP의 강력한 발현을 나타냈다. 이러한 결과는 본 발명의 방법은 유전자 요법에 적합한 고도로 기능적인 cDNA들의 제조를 가능하게 함을 입증한다.
- [0199] **실시예 5: RCA를 통한 TelN-생성 cDNA 주형으로부터 맞춤형(customized) 단일 가닥 DNA 어댑터들을 함유하는 cDNA 제조**
- [0200] **천연 뉴클레오타이드 및 변형된 뉴클레오타이드를 함유하는 맞춤형 단일 가닥 DNA 어댑터들의 합성**
- [0201] 천연 뉴클레오타이드 및 변형된 뉴클레오타이드를 함유하는 맞춤형 단일 가닥 DNA 어댑터들을 표준 포스포아미다이트 화학에 따라 합성했으며 (Beaucage S. L. et al, 1981), 하기 변형된 뉴클레오타이드: 8-옥소-데옥시아데노신 (8-oxo-dA), 5-플루오로-데옥시우라실 (5FU), 이노신, 티오포스페이트 뉴클레오타이드, 또는 잠금 핵산(locked nucleic acid; LNA) 뉴클레오타이드 중 적어도 2개를 포함한다.
- [0202] 간략하게 말하자면, 포스포아미다이트 합성은 3'-측 뉴클레오타이드(3'-most nucleotide)로 시작하며 5'-측 뉴클레오타이드(5'-most nucleotide)가 부착될 때까지 반복되는 4개의 단계들로 구성된 일련의 사이클을 통해 진행된다. 이 단계들은 탈보호 (i), 커플링 (ii), 산화 (iii), 및 캡핑 (iv)이다.
- [0203] 이 사이클은 서열의 각 뉴클레오타이드마다 반복된다. 이 합성의 마지막에서 올리고뉴클레오타이드는 예를 들면, 3' 말단은 여전히 CPG에 부착되고 5' 말단은 트리틸기로 보호된 25-mer로 존재한다. 또한, 보호기들은 4개의 염기들 중 3개의 염기들에 남아 있어 염기들의 고리 구조의 완전성(integrity)을 유지시킨다. A 및 C에 대한 보호기는 벤조일이고 G에 대한 보호기는 N-2-이소부틸릴이다. 티미딘은 보호기가 필요 없다. 완료된 합성은 탈트리틸화된 다음에 조절 공극 유리(controlled pore glass)를 절단하여 3' 및 5' 말단 둘 모두에 하이드록실을 남긴다. 이 시점에서, 올리고 (염기 및 포스페이트)는 고온에서 수산화 암모늄을 사용한 염기 가수분해에 의해 탈보호된다. 최종 생성물은 기능성 단일 가닥 DNA 분자이다.
- [0204] 천연 올리고뉴클레오타이드들을 함유하는 상응하는 헤어핀 DNA 어댑터들을 또한 합성했다. 합성된 어댑터들의 목록은 표 2에 제공된다.
- [0205] 합성 완료시, 올리고뉴클레오타이드들은 지지체로부터 절단되고 보호기들은 제거됐다. 이어서, 표준 정제 단계 (예를 들면 PAGE, HPLC 및/또는 RNase 무함유 HPLC)를 사용하여 절단된 서열들로부터 전장 생성물을 분리했다.

표 2

[0206] 천연 뉴클레오타이드 및 변형된 뉴클레오타이드를 함유하는 단일 가닥 DNA 어댑터들

샘플 명칭	올리고 서열	SEQ ID
올리고 15	AGGGATCCACTCAGGAT	서열 번호 10
올리고 37	AGGGATCC*A*C*T*C*AGGAT	서열 번호 11
올리고 4	AGGGCTAACCCTCAGGTTAG	서열 번호 12
올리고 28	AGGGCTAACC/i8-옥소-dA/CTC/i8-옥소-dA/GGTTAG	서열 번호 13
올리고 29	AGGGCTAACCA/i5F-dU/T/i5FdU/AGGTTAG	서열 번호 14
올리고 17	AGGGATAACATGGCCACTCAGGCCATGTTAT	서열 번호 15
올리고 19	AGGGATAACA+T+G+G+C+CACTCAGGCCATGTTAT	서열 번호 16
올리고 22	AGGGATAACATGGCC/i8-옥소-dA/CTC/i8-옥소-dA/GGCCATGTTAT	서열 번호 17
올리고 21	AGGGATAACATGGCC/I/CTC/I/GGCCATGTTAT	서열 번호 18

올리고 41	AGGGCTTACG*C*G*C*GTAAG	서열 번호 19
--------	------------------------	----------

[0207] *, 우측에 포스포티오에이트화된 뉴클레오티드 (예를 들면, 올리고 37에서, 포스포티오에이트화된 뉴클레오티드는: ACTCA) +, 우측에 LNA 뉴클레오티드 (예를 들면, 올리고 19에서, LNA 뉴클레오티드는: TGGCC)

[0208] /I/ 이노신 뉴클레오티드

[0209] i8-옥소-dA, 8-옥소-데옥시아데노신 뉴클레오티드

[0210] i5F-dU, 5-플루오로-데옥시우라실 뉴클레오티드

[0211] **플라스미드 DNA로부터 맞춤형된 어댑터들을 갖는 cDNA 제조**

[0212] 플라스미드 DNA(pDNA)로부터 시작하여 표 2의 맞춤형된 어댑터들 중 일부를 사용하여 cDNA들을 제조했다. 먼저, pDNA, 예를 들면, 서열 번호 20을 갖는 eGFP 플라스미드 (BsaI 제한 자리들 및 프로텔로머라제 표적 서열에 의해 플랭킹된 Gfp를 인코딩하는 관심 서열을 포함함 (도 9 참고))에 프로텔로머라제를 처리하여 엔도뉴클레아제 제한 자리들에 의해 플랭킹된 관심 서열을 포함하는 cDNA를 생성했다. 이어서, 이 cDNA를 TthPrimPol 및 Phi29를 사용한 롤링 서클 증폭(RCA)을 통해 증폭시켰다. 생성된 콘카테머를 정제하고 상응하는 제한 효소 (예를 들면, BsaI)로 처리하고 맞춤형된 어댑터들; 예를 들면, 표 2의 올리고 21 및 41로 리게이팅했다. 예시적인 상세한 프로토콜은 하기에 제공된다.

[0213] **A. 플라스미드 DNA로부터 cDNA를 얻는 프로토콜**

표 4

실험 기구의 요약

기구	브랜드/제조업체	모델
저울	Mettler Toledo	ME4002E
pH 측정기	INSEA	PHSJ-5
원심분리기	ThermoFisher	Heraeus™ Pico™ 21
클린 벤치	AIRTECH	SW-CJ-2FD

표 5

물질 정보의 요약

No.	물질 명칭	브랜드 또는 제조업체	Cat. No.
1	엑소뉴클레아제 III	NEB	M0206
2	NEBuffer 1	NEB	M0206
3	TritonX-114	Solarbio	T8210
4	이소프로필 알코올	Sinopharm Chemical Reagent Co., Ltd.	67-63-0
5	KpnI	NEB	R3142L
6	HindIII	NEB	R3104S
7	CutSmart 완충액	NEB	B7204S
8	Te1N	GenScript	NA
9	Te1N 완충액	GenScript	NA

[0216] **1.1Te1N 분해**

[0217] 2h 동안 30°C에서 eGFP 플라스미드를 Te1N 효소로 분해하고 10 min 동안 75°C에서 불활성화시켰다. 동시에 여러 반응들을 수행할 때 그에 따라 스케일 업(scale up)을 한다.

표 6

[0218] TeIN 효소 분해 반응

구성성분	20 mL의 반응
10X 완충액	2mL
플라스미드	10 mg, 10 mL
TeIN	실제 첨가: 1.0×10^6 U(2 mL, 50 U/ μ L)
멸균수	20 mL까지 첨가

[0219] 1.2백본 제거

[0220] 1.2.1Kpn I 및 Hind III 분해

[0221] 마지막 단계의 생성물을 1h 동안 37°C에서 Kpn I 및 Hind III으로 분해하였다. 이어서, 샘플을 15분 동안 65°C에서 불활성화시켰다. 동시에 여러 반응들을 수행할 때 그에 따라 스케일 업을 한다.

표 7

[0222] Kpn I 및 Hind III 분해 반응

구성성분	25 mL의 반응
10x Cutsmart	2.5 mL
마지막 단계의 플라스미드	20 ml
Kpn I	실제 첨가: 10000 U (500 μ L, 20 U/ μ L)
Hind III	실제 첨가: 10000 U (500 μ L, 20 U/ μ L)
멸균수	25 mL까지 첨가

[0223] 1.2.2Exo III 분해

[0224] 1h 동안 37°C에서 Exo III으로 분해하고 10min 동안 75°C에서 불활성화시켰다. 동시에 여러 반응들을 수행할 때 이에 따라 스케일 업을 한다.

표 8

[0225] Exo III 분해 반응

구성성분	28 mL의 반응
10x NEBuffer 1	2.8 mL
마지막 단계의 플라스미드	25 mL
Exo III	실제 첨가: 30000 U(300 μ L, 100U/ μ L)

[0226] 1.3겔 여과 크로마토그래피 및 이소프로판올로 cDNA 정제

[0227] 1.3.1겔 여과 크로마토그래피

[0228] 완충액 A: 10mM Tris-HCl, pH 7.5

[0229] 컬럼: 베스타로스 6 FF(Bestarose 6 FF) 153 mL

[0230] 샘플: 28 ml

[0231] 흐름: 60 cm/h

[0232] 수거 분획: 20mAU-20mAU, 40mL

[0233] CIP: 1 M NaOH + 순수(pure water)

[0234] 저장: 순수

[0235] 1.3.2엔도톡신 제거 및 이소프로판올 침전

[0236] 3M 아세트산 나트륨 및 15% Triton-114을 마지막 단계의 샘플에 첨가하고 표 6에 나타낸 바와 같이 볼텍싱에 의

해 혼합했다. 샘플을 4℃에서 5 min 동안 두었다. 이어서, 25℃에서 20 min 동안 12000g로 원심분리했다. 원심 분리 후에, 상층액을 수거하고 동일한 부피의 이소프로판올을 상층액에 첨가하고 완전히 혼합했다. 샘플을 실온에서 5 min 동안 두었다. 이후에, 20 min 동안 12000g로 원심분리하여 상층액을 제거했다. 마지막으로, 침전물을 10mM Tris-HCl (pH 7.5)로 현탁시켰다.

표 9

Triton-114 시스템

물질	첨가량 (A는 c1DNA의 부피)
c1DNA	A(40mL, ~100 µg/mL)
15% Triton 114	0.1A(4mL)

[0237]

[0238]

효소 분해, 겔 크로마토그래피, Triton 114 처리 및 이소프로판올 침전의 3단계 후에, eGFP_BSAI_c1DNA가 성공적으로 제조됐다. HPLC 크로마토그램에 따른 샘플의 DNA 균질도 (%)는 97%이었다. 이 샘플의 엔도톡신은 <10EU/mg이다.

[0239]

B. RCA를 통한 c1DNA로부터 맞춤형된 어댑터들을 함유하는 c1DNA를 얻는 프로토콜

[0240]

이 실험은 Trueprime-RCA 키트에 의해 상기 섹션에서 얻어진 eGFP_BSAI_c1DNA로부터 맞춤형된 어댑터들을 함유하는 c1DNA를 제조하기 위해 설계되었다.

표 10

실험 기구의 요약

기구	브랜드/제조업체	모델
저울	Mettler Toledo	ME4002E
pH 측정기	INSEA	PHSJ-5
원심분리기	ThermoFisher	Heraeus™ Pico™ 21
클린 벤치	AIRTECH	SW-CJ-2FD

[0241]

표 11

물질 정보의 요약

No.	물질 명칭	브랜드 또는 제조업체	Cat. No.
1	엑소뉴클레아제 III	NEB	M0206
2	NEBuffer 1	NEB	M0206
3	TritonX-114	Solarbio	T8210
4	이소프로필 알코올	Sinopharm Chemical Reagent Co., Ltd.	67-63-0
5	KpnI	NEB	R3142L
6	HindIII	NEB	R3104S
7	CutSmart 완충액	NEB	B7204S
8	4BBTM TruePrime® RCA 키트	4basebio	390100
9	완충액 D	4basebio	390100
10	완충액 N	4basebio	390100
11	반응 완충액	4basebio	390100
12	효소 1(TthPrimPol)	4basebio	390100
13	효소 2(Phi29 DNA 폴리머라제)	4basebio	390100
14	Te1N	GenScript	NA
15	Te1N 완충액	GenScript	NA
16	AxyPrep DNA 겔 추출 키트	Axygen	AP-GX-250
17	완충액 DE-B	Axygen	AP-GX-250
18	완충액 W1	Axygen	AP-GX-250
19	완충액 W2	Axygen	AP-GX-250

[0242]

20		T4 리가아제	NEB	M0202T
21		T4 리가아제 완충액	NEB	M0202T
22		BsaI	NEB	R3733L
23		T4 PNK	NEB	M0201L
24		T4 PNK 완충액	NEB	M0201L

[0243]

1. IRCA

[0244]

☆ 항상 피펫팅으로 혼합한다. 볼텍싱 하지 않는다

[0245]

☆ c1DNA 10 μl (≥1 ng/μl)를 깨끗한 튜브에 옮긴다

[0246]

☆ 완충액 D 10 μl를 첨가하고 3분 동안 실온에서 인큐베이팅한다

[0247]

☆ 완충액 N 10 μl를 각 튜브에 첨가하여 반응을 중화시킨다

[0248]

☆ 사용될 때까지 샘플을 실온에 둔다*

[0249]

☆ 하기 표에 나열된 순서로 구성성분들을 첨가하여 증폭 믹스(amplification mix)를 제조한다

[0250]

☆ 3시간 동안 30℃에서 인큐베이팅한다**. 10분 동안 65℃에서 반응을 불활성화시킨다

[0251]

☆ 4℃로 냉각시킨다. 증폭 DNA를 단기간 저장의 경우 4℃에서 저장하고 장기간 저장의 경우 -20℃에서 저장한다.

[0252]

(*) 증폭 반응은 샘플이 변성된 직후에 수행하는 것을 매우 권장한다.

[0253]

(**) 더 높은 증폭 생산량이 요구되는 경우에 인큐베이션 시간을 최대 6시간까지 늘릴 수 있다.

[0254]

동시에 여러 반응들을 수행할 때 그에 따라 스케일 업을 한다.

표 12

RCA-100ul 시스템

[0255]

물질	첨가량	코멘트
c1DNA	10 uL	(>1 μg/mL, 총 80 ng)
완충액 D	10 uL	RT에서 3 min
완충액 N	10 uL	중화
H ₂ O	37.2 μL	증폭 믹스
반응 완충액	10 μL	
dNTPs	10 μL	
효소 1 (TthPrimPol)	10 μL	
효소 2 (Phi29 DNA 폴리머라제)	2.8 μL	

[0256]

1. 2RCA 생성물 (콘카테머)를 이소프로판올로 정제한다 (상기 기재한 바와 같음)

[0257]

1. 3RCA 생성물 (콘카테머)를 Axygen 키트로 정제한다 (선택적임)

[0258]

샘플이 100 uL 이하인 경우에, c1DNA를 정제하는데 Axygen 키트를 사용할 수 있다. 프로토콜은 하기에 기술되며 병(bottle)들은 하기에 기재된 완충액 라벨을 함유한다:.

[0259]

1) 2x 샘플 부피의 완충액 DE-B를 첨가하고, 혼합한다.

[0260]

2) Miniprep 컬럼을 2 ml 마이크로퓨지 튜브(microfuge tube) 내에 배치한다. 마지막 단계의 샘플을 이 컬럼에 옮긴다. 1분 동안 12,000xg에서 원심분리한다.

[0261]

3) 2 ml 마이크로퓨지 튜브로터 여과액을 버린다. Miniprep 컬럼을 2 ml 마이크로퓨지 튜브에 되돌리고 완충액 W1 500 μl를 첨가한다. 30초 동안 12,000xg에서 원심분리한다.

[0262]

4) 2 ml 마이크로퓨지 튜브로부터 여과액을 버린다. Miniprep 컬럼을 2 ml 마이크로퓨지 튜브에 되돌리고 완충

액 W2 700 μ l를 첨가한다. 30초 동안 12,000xg에서 원심분리한다.

[0263] 5) 2 ml 마이크로퓨지 튜브로부터 여과액을 버린다. Miniprep 컬럼을 2 ml 마이크로퓨지 튜브에 다시 배치한다. 완충액 W2의 제2 700 μ l 분취량을 첨가하고 1분 동안 12,000xg에서 원심분리한다.

[0264] 6) Miniprep 컬럼을 깨끗한 1.5 ml 마이크로퓨지 튜브 (제공됨)에 옮긴다. DNA를 용출하기 위해, 10mM Tris-HCL (pH 7.5) 50 μ l를 멤브레인(membrane)의 중앙에 첨가한다. 1분 동안 실온에서 방치한다. 1분 동안 12,000 xg에서 원심분리한다.

[0265] 1.4올리고 변성 및 어닐링(*annealing*)

[0266] 올리고 (예를 들면, 표 2: 올리고 21 또는 올리고 41)를 10min 동안 95°C에서 변성시키고 30min 동안 실온에서 자연적으로 어닐링했다. 동시에 여러 반응들을 수행할 때 그에 따라 스케일 업을 한다.

표 13

올리고 변성 및 어닐링

물질	첨가량
인산화된 올리고	95 μ L
20X SSC	5 μ L

[0268] 1.5올리고 인산화 (선택적임, 올리고가 이미 인산화된 경우 이 단계를 건너 뛴)

[0269] 1h 동안 37°C에서 올리고 인산화를 수행했다.

표 14

올리고 인산화

물질	첨가량
인산화 없는 올리고	80 μ L
T4 PNK 완충액	10 μ L
T4 PNK	10 μ L

[0271] 1.6BsaI 분해

[0272] 2h 동안 37°C에서 BsaI로 분해하고 10min 동안 75°C에서 불활성화시켰다.

표 15

BsaI의 분해 시스템 100uL

물질	첨가량
정제된 RCA 생성물	85 μ L
CutSmart 완충액	10 μ L
BsaI	20U/ μ L, 1 μ L
멸균수	0.1mL까지 첨가

[0274] 1.7BsaI-분해 RCA 생성물을 이소프로판올로 정제한다 (상기 기재한 바와 같음)

[0275] 1.8BsaI-분해 RCA 생성물을 Axygen 키트로 정제한다. (선택적임, 샘플이 100 uL 이하인 경우; 상기 기재한 바와 같음)

[0276] 1.9T4 리게이션

[0277] 밤새 16°C에서 T4 리게이팅하고 10min 동안 75°C에서 불활성화시켰다.

표 16

[0278]

T4 리게이션 시스템 100uL

물질	첨가량
올리고	5uL (~1 μg/ μL)
BsaI-분해 cDNA	85uL (~100ng/uL)
T4 리가아제 완충액	10 μL
T4 리가아제	20000U, 1 μL
멸균수	0.1mL까지 첨가

[0279]

1.10개선했된 골든 게이트 조립(Advanced Golden Gate Assembly) (선택적임)

[0280]

기존의 연결 방법은 일반적으로 관심 구조체를 생성하기 위해 여러 클로닝 단계들을 필요로 한다. 각 단계에서, 단일 DNA 단편은 공여자 플라스미드 또는 PCR 생성물로부터 수용자 벡터로 옮겨진다.

[0281]

골든 게이트 클로닝 동안에, 수용자 플라스미드에서 한 번에 최대 15개의 단편들을 조립하는 것이 가능하다. 모든 플라스미드 공여자, 수용자 벡터, IIS형 제한 효소 및 리가아제를 단일 튜브에 피펫팅하고, 열 순환기 (thermal cycler)에서 믹스를 인큐베이팅함으로써 클로닝을 수행한다. 또한, 본 발명자들은 또한 골든 게이트 조립으로 oDNA를 만드는 것을 제안한다. 이 시스템 및 조건은 각각 표 14 및 15에 기재되어 있다. 동시에 여러 반응들을 수행할 때 그에 따라 스케일 업을 한다.

표 17

[0282]

개선했된 골든 게이트 조립 시스템

물질	첨가량
10X T4 리가아제 완충액	10 μL
cDNA(RCA에 의해 증폭됨)	24 μg
올리고	72 μg
BsaI	(100U) 5uL
T4 DNA 리가아제	(20000U) 1uL
뉴클레아제 무함유 물	최대 200 μL

표 18

[0283]

개선했된 골든 게이트 조립 조건

온도	시간	
37°C	3min	25 사이클
22°C*	5min	
22°C	60min	
50°C	5min	
80°C	10min	
4°C	사용될 때까지 4° C에서 샘플을 저장	

[0284]

*NEB의 T4 리가아제의 최적 온도는 16°C이고 Thermofisher의 T4 리가아제의 최적 온도는 22°C이다.

[0285]

1.11예상치 못한 DNA의 분해

[0286]

1h 동안 37°C에서 Exo III으로 분해하고 10 min 동안 75°C에서 불활성화시켰다. 동시에 여러 반응들을 수행할 때 그에 따라 스케일 업을 한다.

표 19

[0287]

Exo III 분해 반응

구성성분	0.3 mL의 반응
10x NEBuffer 1	30 μL
마지막 단계의 플라스미드	0.2 mL

Exo III	200 U(2 μ L, 100U/ μ L)
뉴클레아제-무함유 물	최대 300 μ L

[0288] 1.12oDNA를 이소프로판올로 정제한다 (상기 기재한 바와 같음)

[0289] 1.13oDNA를 Axygen 키트로 정제한다. (선택적임, 샘플이 100 μ L 이하인 경우; 상기 기재된 바와 같음)

[0290] 올리고 21 및 41로 eGFP_BSaI_oDNA를 성공적으로 제조했다.

표 20

[0291]

샘플	농도 (ng/ μ L)	부피 (ml)	총 (μ g)	균질도 (%)
oDNA 21	119.6	0.05	6.0	95.6
oDNA 22	130.0	0.05	6.5	96.2

[0292] 동일한 절차를 사용하여 서열 번호 21을 갖는 Luc 플라스미드 (BsaI 제한 자리들, 및 프로텔로머라제 표적 서열들에 의해 플랭킹된 루시페라제를 인코딩하는 서열을 포함함) 및 표 2의 올리고 15, 37, 4, 28, 29, 17, 22, 37, 28, 29, 19 및 22로부터 출발하여 cDNA들을 제조했다. 또한, 동일한 절차를 사용하여 서열 번호 22 (이때, 관심 서열은 BsaI 제한 자리 및 프로텔로머라제 표적 서열들에 의해 플랭킹된 루시페라제를 인코딩하는 서열을 플랭킹하는 ITRs를 추가로 포함함, 도 10 참고)를 갖는 Luc-ITR 플라스미드 및 올리고 4로부터 출발하여 cDNAs를 제조하고; 그 결과, oDNA 4_{ITR} (총 6.2 μ g)을 생성했다 - 도 11 (아가로스 겔 전기영동).

[0293] 얻어진 cDNA의 품질을 표준 절차, 특히, 아가로스 겔 전기영동, 그레이스케일 분석, 음이온 교환 크로마토그래피-HPLC 및 생어 서열 분석에 의해 결정했다. 모든 cDNA들은 순도, 피크 분리도 및 서열 확인(sequence confirmation) 관점에서 우수한 품질 특성을 나타냈다. 예를 들면, oDNA41 및 oDNA21에 대한 결과는 각각 도 7 및 8에 나타나 있다.

[0294] **실시예 6. RCA에 이어서 TeIN 처리를 통한 TeIN-생성 cDNA 주형으로부터 cDNA 제조**

[0295] 대안적으로는, 본 발명의 cDNA는 실시예 5 섹션 A, 1.1부터 섹션 B, 1.3까지 기재된 절차에 이어서, 생성된 콘카테머의 TeIN 처리에 의해 제조될 수 있다. 이 마지막 단계에서, RCA로부터 얻어진 콘카테머 (섹션 B, 1.3)를 프로텔로머라제로 처리하는 것은 하기에 상세히 설명된다.

[0296] 1.4TeIN 분해

[0297] 정제된 RCA 생성물을 2h 동안 30°C에서 TeIN로 분해하고 10 min 동안 75°C에서 불활성화시켰다. 동시에 여러 반응들을 수행할 때 그에 따라 스케일 업을 한다.

표 21

TeIN 효소 분해 반응

[0298]

구성성분	0.5 mL의 반응
10X TeIN Buffer	0.05 mL
정제된 RCA 생성물 (콘카테머)	80 μ g, 0.3 mL
TeIN	1500 U(0.03 mL, 50 U/ μ L)
멸균수	0.5 mL까지 첨가

[0299] 1.5백본 제거

[0300] 1.4.1 Kpn I 및 Hind III 분해

[0301] 마지막 단계의 생성물을 1h 동안 37°C에서 Kpn I 및 Hind III으로 분해한다. 이어서, 15분 동안 65°C에서 샘플을 불활성화시켰다. 동시에 여러 반응들을 수행할 때 그에 따라 스케일 업을 한다.

표 22

[0302]

Kpn I 및 Hind III 분해

구성성분	0.6 mL의 반응
10x Cutsmart	60 μ L
마지막 단계의 cDNA	0.5 mL
Kpn I	실제 첨가: 400 U (20 μ L, 20 U/ μ L)
Hind III	실제 첨가: 400 U (20 μ L, 20 U/ μ L)

[0303]

1.4.2 Exo III 분해

[0304]

1h 동안 37°C에서 Exo III으로 분해하고 10 min 동안 75°C에서 불활성화시켰다. 동시에 여러 반응들을 수행할 때 그에 따라 스케일 업을 한다.

표 23

[0305]

Exo III 분해 반응

구성성분	0.67 mL의 반응
10x NEBuffer 1	62 μ L
마지막 단계의 cDNA	0.6 mL
Exo III	800 U (80 μ L, 100U/ μ L)

[0306]

1.6이소프로판올로 cDNA를 정제한다 (상기 기재된 바와 같음)

[0307]

1.7Axygen 키트로 cDNA 생성물을 정제한다 (선택적임, 샘플이 100 μ L 이하인 경우; 상기 정의된 바와 같음)

[0308]

GFP를 인코딩하는 관심 서열인 (도 9 참고), 합성된 cDNA는 절단/결합 후에 얻어진 일정한 28개의 염기쌍 프로 텔로머라제 서열을 보유한다. 합성 성능에 대한 세부사항들은, 단계별로 표 24에 기재되어 있고 (하기); 얻어진 cDNA는 아가로스 겔 전기영동(AGE)에 따르면 96.6%의 균질도를 나타냈다 - 도 12.

표 24

[0309]

주요 단계들의 회수율(recovery)은 하기와 같다:

	부피 (μ L)	농도 (μ g/mL)	총 (μ g)	회수율 (%)
주형 플라스미드	300	15	4.5	NA
정제된 RCA 생성물	1500	260	390	~86 배의 생산량
정제된 cDNA	1200	121.6	145.9	37.41%

[0310]

인용 문헌

[0311]

Heinrich, M. et al. "Linear closed mini DNA generated by the prokaryotic cleaving-joining enzyme TelN is functional in mammalian cells", J Mol Med, 2002, vol. 80, pp. 648-654

[0312]

Altschul et al., "Basic local alignment search tool", 1990, J. Mol. Biol, vol. 215, pp. 403-410

[0313]

Xiao X. et al., "A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle", 1997, J Virol., vol. 71(2), pp. 941-948.

[0314]

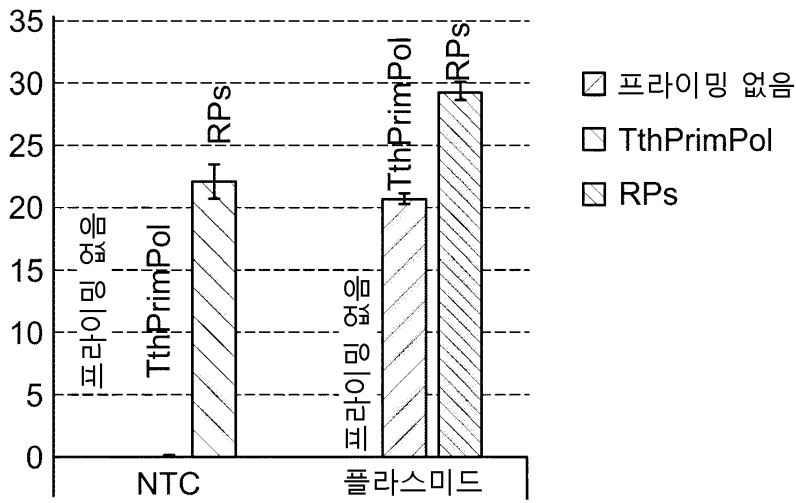
W02011000997

[0315]

Beaucage S. L. et al, Deoxynucleoside phosphoramidites—A new class of key intermediates for deoxypolynucleotide synthesis. Tetrahedron Letters, Volume 22, Issue 20, 1981, Pages 1859-1862

도면

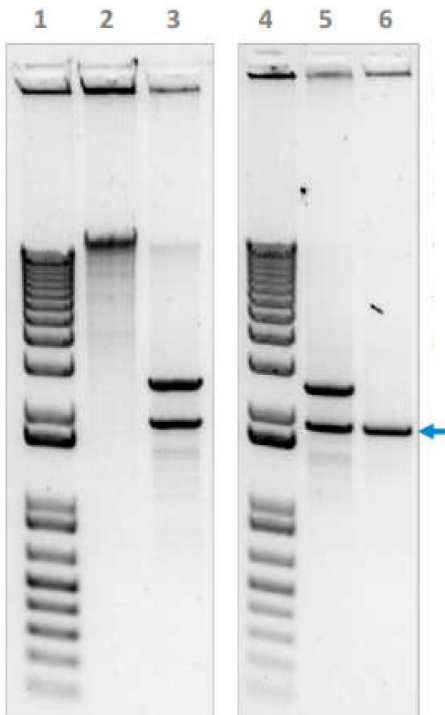
도면1



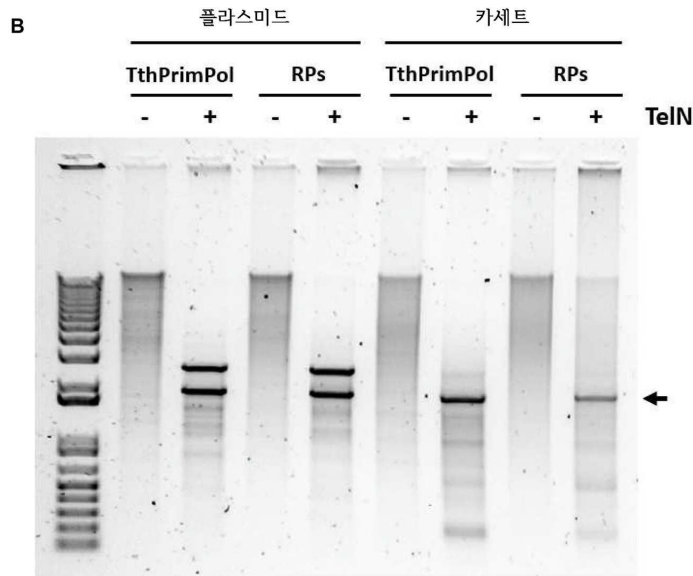
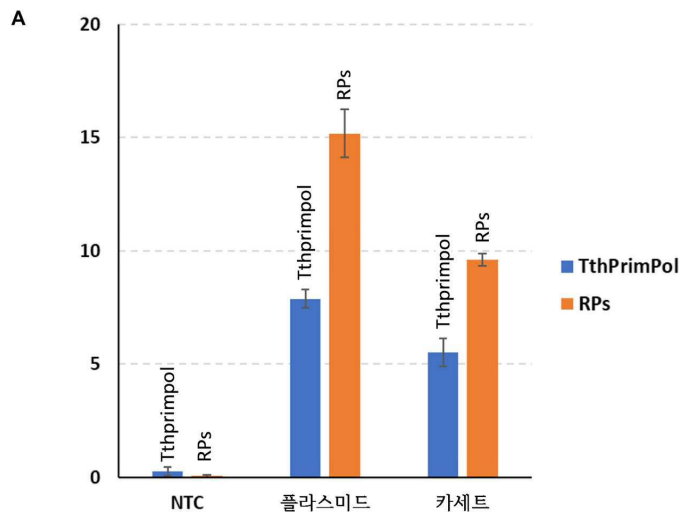
도면2

샘플	생성된 판독	매핑된 판독	백분율
PrimPol	10,628,762	8,419,939 100%	79.22%
RPs	12,467,258	8,198,854 97.4%	65.76%

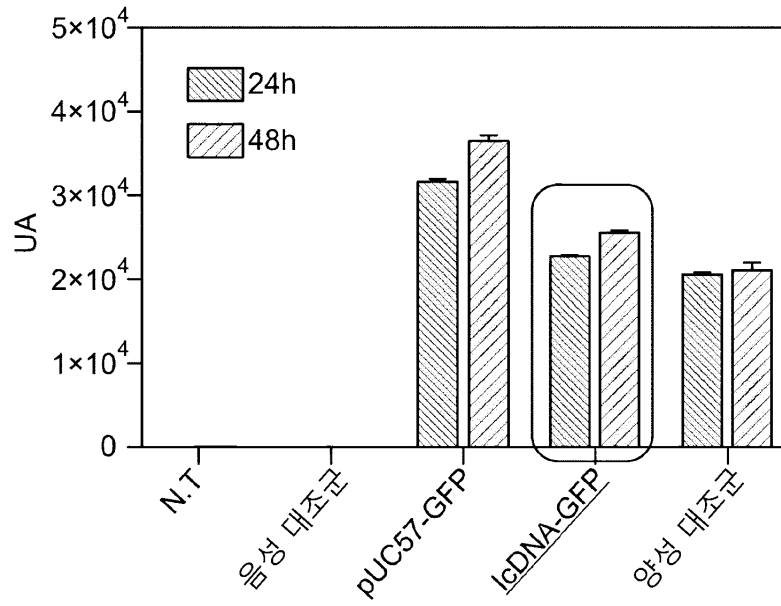
도면3



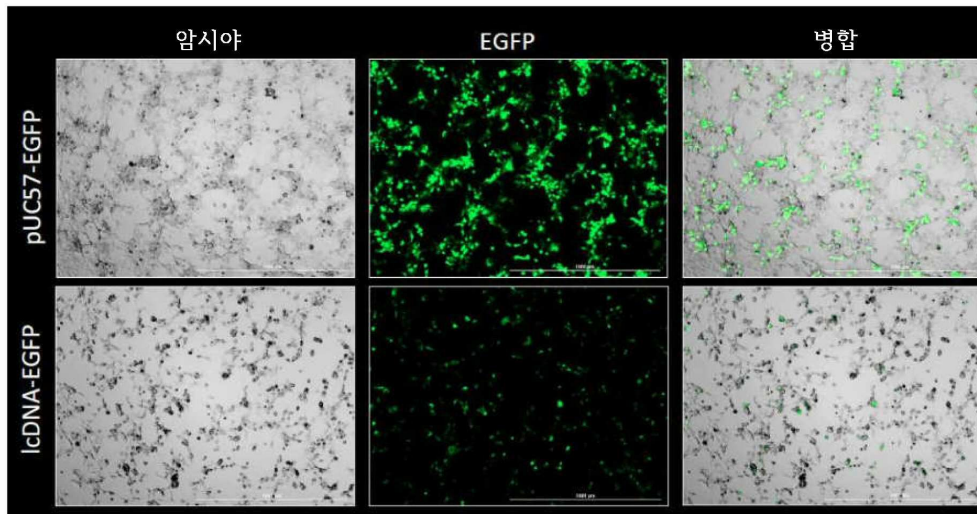
도면4



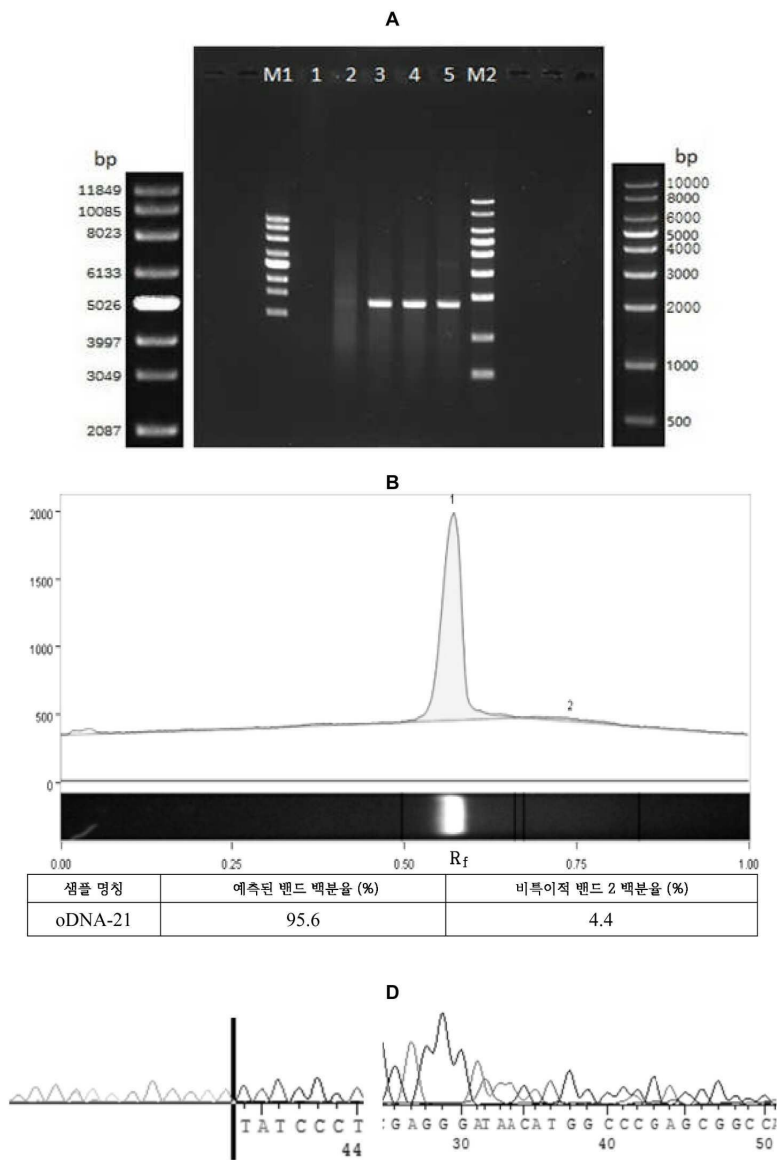
도면5



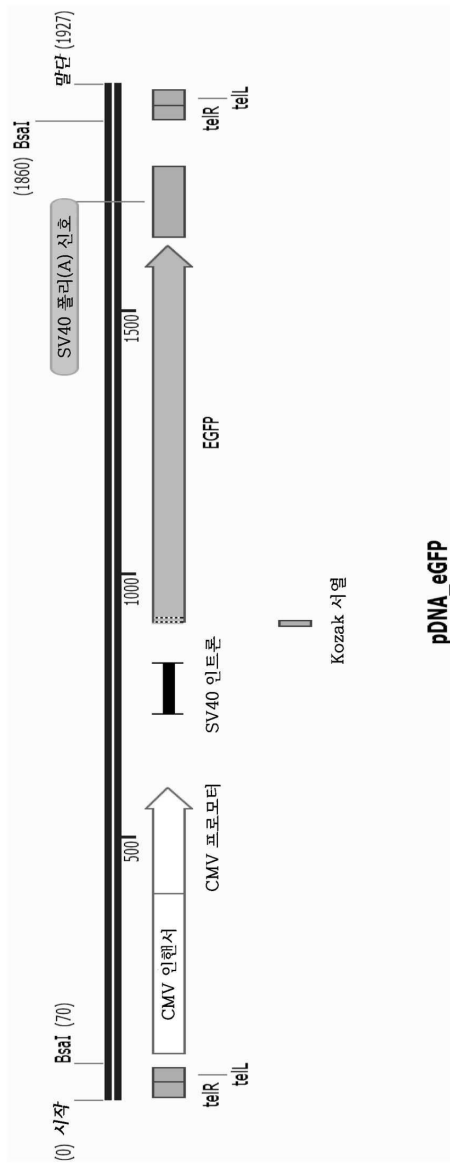
도면6



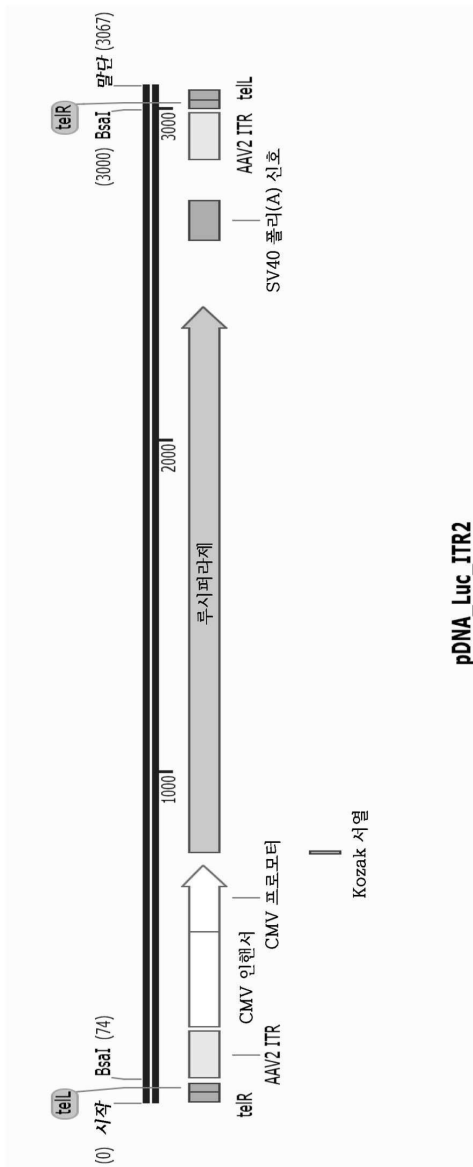
도면8



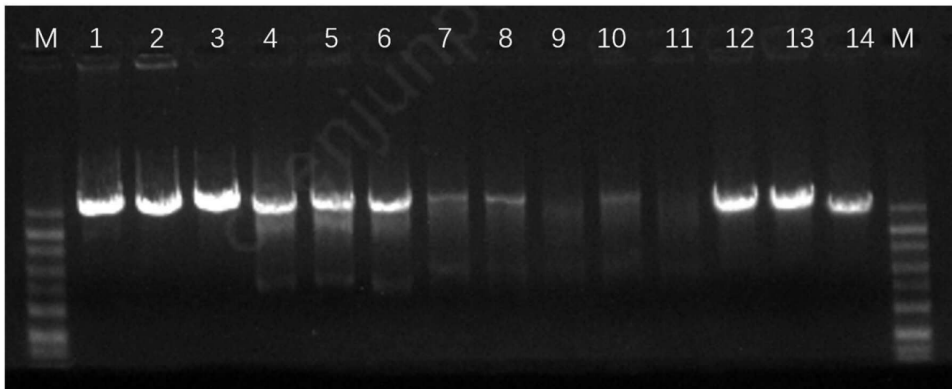
도면9



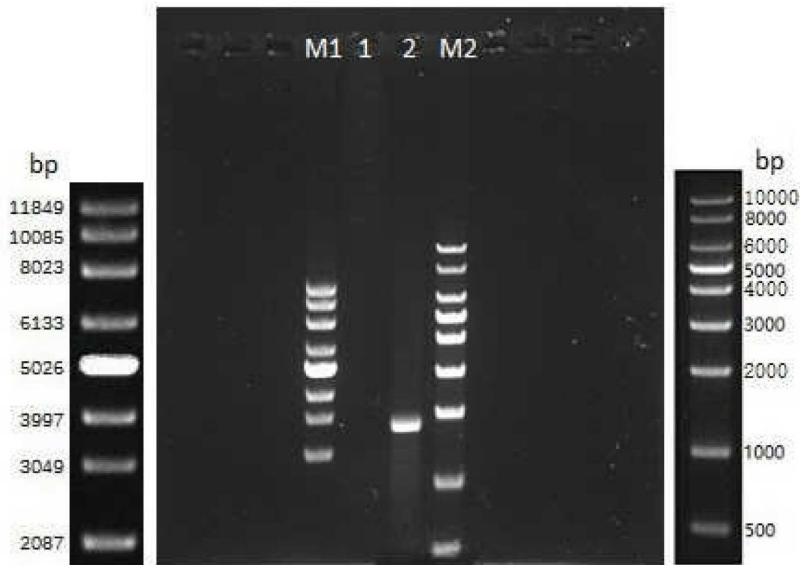
도면10



도면11



도면12



서열 목록

SEQUENCE LISTING

<110> COLUMBUS EXPLORER, S.L.

<120> Process for the production of closed linear DNA

<130> P5393PC00

<150> EP20382064

<151> 2020-01-31

<160> 23

<170> PatentIn version 3.5

<210> 1

<211> 293

<212> PRT

<213> Thermus thermophilus

<400> 1

Met Arg Pro Ile Glu His Ala Leu Ser Tyr Ala Ala Gln Gly Tyr Gly

1 5 10 15

Val Leu Pro Leu Arg Pro Gly Gly Lys Glu Pro Leu Gly Lys Leu Val

20 25 30

Pro His Gly Leu Lys Asn Ala Ser Arg Asp Pro Ala Thr Leu Glu Ala

35 40 45

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

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

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

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

 Met Arg Arg Trp Val

290

<210> 2

<211> 631

<212> PRT

<213> Artificial

<220><223> TeIN

<400> 2

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

1 5 10 15

Val Glu Ala Ile Asp Ala Ser Asp Arg Pro Gln Gly Asp Lys Thr Lys

 20 25 30

Arg Ile Lys Ala Ala Ala Ala Arg Tyr Lys Asn Ala Leu Phe Asn Asp

 35 40 45

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

50 55 60

Thr Phe Asn Ala Tyr Met Ser Arg Ala Arg Lys Arg Phe Asp Asp Lys

65 70 75 80

Leu His His Ser Phe Asp Lys Asn Ile Asn Lys Leu Ser Glu Lys Tyr

 85 90 95

Pro Leu Tyr Ser Glu Glu Leu Ser Ser Trp Leu Ser Met Pro Thr Ala

100 105 110

Asn Ile Arg Gln His Met Ser Ser Leu Gln Ser Lys Leu Lys Glu Ile

115 120 125

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

130 135 140

Ser Asp Ala Lys Ile Ala Arg Leu Ile Lys Lys Tyr Pro Asp Trp Ser

145 150 155 160

Phe Ala Leu Ser Asp Leu Asn Ser Asp Asp Trp Lys Glu Arg Arg Asp

165 170 175

Tyr Leu Tyr Lys Leu Phe Gln Gln Gly Ser Ala Leu Leu Glu Glu Leu

180 185 190

His Gln Leu Lys Val Asn His Glu Val Leu Tyr His Leu Gln Leu Ser

Val Ala Leu Gln Lys Leu Asp Asp Glu Met Pro Gly Phe Ala Arg Gly
 450 455 460
 Asp Ala Gly Val Arg Leu His Glu Thr Val Lys Gln Leu Val Glu Gln
 465 470 475 480
 Asp Pro Ser Ala Lys Ile Thr Asn Ser Thr Leu Arg Ala Phe Lys Phe
 485 490 495

Ser Pro Thr Met Ile Ser Arg Tyr Leu Glu Phe Ala Ala Asp Ala Leu
 500 505 510
 Gly Gln Phe Val Gly Glu Asn Gly Gln Trp Gln Leu Lys Ile Glu Thr
 515 520 525
 Pro Ala Ile Val Leu Pro Asp Glu Glu Ser Val Glu Thr Ile Asp Glu
 530 535 540
 Pro Asp Asp Glu Ser Gln Asp Asp Glu Leu Asp Glu Asp Glu Ile Glu
 545 550 555 560

Leu Asp Glu Gly Gly Gly Asp Glu Pro Thr Glu Glu Glu Gly Pro Glu
 565 570 575
 Glu His Gln Pro Thr Ala Leu Lys Pro Val Phe Lys Pro Ala Lys Asn
 580 585 590
 Asn Gly Asp Gly Thr Tyr Lys Ile Glu Phe Glu Tyr Asp Gly Lys His
 595 600 605
 Tyr Ala Trp Ser Gly Pro Ala Asp Ser Pro Met Ala Ala Met Arg Ser
 610 615 620

Ala Trp Glu Thr Tyr Tyr Ser
 625 630

<210> 3

<211> 56

<212> DNA

<213> Artificial

<220><223> TelN target sequence

<400> 3

tatcagcaca caattgccca ttatcgcgc gtataatgga ctattgtgtg ctgata

56

<210> 4

<211> 41
 <212> DNA
 <213> Artificial
 <220><223> AflIII adaptor
 <400> 4
 ttaagtaaca ttgttgcc actcaggcca acaaatgtta c 41
 <210> 5
 <211> 41
 <212> DNA
 <213> Artificial
 <220><223> NheI-HF adaptor
 <400> 5
 ctagctaaca ttgttgcc actcaggcca acaaatgtta g 41

 <210> 6
 <211> 40
 <212> DNA
 <213> Artificial
 <220><223> EcoRV adaptor
 <400> 6
 tctaacattt gttggccact caggccaaca aatgtagat 40
 <210> 7
 <211> 40
 <212> DNA
 <213> Artificial
 <220><223> ScaI adaptor
 <400> 7
 cttaacattt gttggccact caggccaaca aatgtaagt 40
 <210> 8
 <211> 141
 <212> DNA
 <213> Artificial
 <220><223> 5' ITR sequence
 <400> 8

cctgcaggca gctgcgcgct cgctcgctca ctgaggccgc ccgggcaaag cccgggcgctc 60

gggcgacctt tggtcgcccc gcctcagtga gcgagcgagc ggcagagag ggagtggcca 120

actccatcac taggggttcc t 141

<210> 9

<211> 141

<212> DNA

<213> Artificial

<220><223> 3' ITR sequence

<400> 9

aggaaccctt agtgatggag ttggccactc cctctctgcg cgctcgctcg ctcaactgagg 60

ccggcgacc aaaggtcgcc gcacgcccg gctttgcccc ggccgctca gtgagcgagc 120

gagcgcgag ctgcctgcag g 141

<210> 10

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> oligo 15

<400> 10

agggatccac tcaggat 17

<210> 11

<211> 17

<212> DNA

<213> Artificial Sequence

<220><223> oligo 37

<220><221> modified_base

<222> (9)..(13)

<223> phosphothioated nucleotides

<400> 11

agggatccac tcaggat 17

<210> 12

<211> 21

<212> DNA

<213> Artificial Sequence

<220><223> oligo 4
 <400> 12
 agggctaacc actcaggtta g 21

<210> 13
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> oligo 28
 <220><221> modified_base
 <222> (11)..(11)
 <223> 8-oxo-deoxyadenosine
 <220><221> modified_base
 <222> (15)..(15)
 <223> 8-oxo-deoxyadenosine
 <400> 13
 agggctaacc actcaggtta g 21

<210> 14
 <211> 21
 <212> DNA
 <213> Artificial Sequence
 <220><223> oligo 29
 <220><221> modified_base
 <222> (12)..(12)
 <223> 5-Fluoro-deoxyuracil
 <220><221> misc_feature

 <222> (12)..(12)
 <223> n is a, c, g, or t
 <220><221> modified_base
 <222> (14)..(14)
 <223> 5-Fluoro-deoxyuracil
 <220><221> misc_feature
 <222> (14)..(14)
 <223> n is a, c, g, or t

<400> 14	
agggctaacc antnaggtta g	21
<210> 15	
<211> 31	
<212> DNA	
<213> Artificial Sequence	
<220><223> oligo 17	
<400> 15	
agggataaca tggccactca ggccatgta t	31
<210> 16	
<211> 25	
<212> DNA	
<213> Artificial Sequence	
<220><223> oligo 19	
<220><221> modified_base	
<222> (11)..(15)	
<223> LNA nucleotides	
<400> 16	
agggataaca tggccactca ggcca	25
<210> 17	
<211> 31	
<212> DNA	
<213> Artificial Sequence	
<220><223> oligo 22	
<220><221> modified_base	
<222> (16)..(16)	
<223> 8-oxo-deoxyadenosine	
<220><221> modified_base	
<222> (20)..(20)	
<223> 8-oxo-deoxyadenosine	
<400> 17	
agggataaca tggccactca ggccatgta t	31
<210> 18	

<211> 31
 <212> DNA
 <213> Artificial Sequence
 <220><223> oligo 21
 <220><221> modified_base
 <222> (16)..(16)
 <223> I
 <220><221> misc_feature
 <222> (16)..(16)
 <223> n is a, c, g, or t
 <220><221> modified_base
 <222> (20)..(20)
 <223> I
 <220><221> misc_feature
 <222> (20)..(20)
 <223> n is a, c, g, or t
 <400> 18
 agggataaca tggccnctcn ggccatgta t
 <210> 19
 <211> 18
 <212> DNA
 <213> Artificial Sequence
 <
 <220><223> oligo 41
 <220><221> modified_base
 <222> (11)..(14)
 <223> phosphothioated nucleotide
 <400> 19
 agggcttacg cgcgtaag
 <210> 20
 <211> 3762
 <212> DNA
 <213> Artificial Sequence
 <220><223> eGFP plasmid
 <400> 20

31

18

ccaatgatgg tacctatca gcacacaatt gccattata cgcgcgtata atggactatt 60
 gtgtgctgat aggtctcgcc ctgatttaa tcgtcgaccg ttacataact tacggtaaat 120
 ggccccctg gctgaccgcc caacgacccc cgccattga cgtcaataat gacgtatgtt 180

cccatagtaa cgccaatagg gactttccat tgacgtcaat gggaggagta ttacggtaa 240
 actgcccact tggcagtaca tcaagtgtat catatgcaa gtacgcccc tattgacgtc 300
 aatgacggta aatggcccc ctggcattat gcccagtaca tgacctatg ggactttcct 360
 acttggcagt acatctacgt attagtcac gctattacca tggatgacg gttttggcag 420
 tacatcaatg ggcgtggata gcggtttgac tcacggggat ttccaagtct ccacccatt 480
 gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg actttccaaa atgtcgtaac 540
 aactccgcc cattgacgca aatggcggtt aggcgtgtac ggtgggaggt ctatataagc 600

agagctcgtt tagtgaaccg tcagatcgcc tggagacgcc atccacgtg ttttgacctc 660
 catagaagac accgggaccg atccagcctc cggactctag aggatccggt actcgaggaa 720
 ctgaaaaacc agaaagttaa ctggttaagt tagtcttttt gctttttatt tcaggtcccg 780
 gatccggtgg tgggtcaaat caaagaactg ctctcagtg gatgttcct ttacttctag 840
 gcctgtacgg aagtgttact tctgtctaa aagctgcgga attgtaccg cggccccgga 900
 tccaccggtc gccaccatgg tgagcaaggc cgaggagctg ttaccgggg tggtgccat 960
 cctggtcgag ctggacggcg acgtaaacgg ccacaagttc agcgtgtccg gcgagggcga 1020

gggcgatgcc acctacggca agctgacct gaagttcacc tgcaccaccg gcaagctgcc 1080
 cgtgccctgg cccacctcg tgaccacct gacctacggc gtgcagtgtc tcagccgcta 1140
 cccgaccac atgaagcagc acgacttctt caagtccgcc atgcccgaag gctacgtcca 1200
 ggagcgcacc atcttttca aggacgacgg caactacaag acccgcccg aggtgaagtt 1260
 cgagggcgac accctggtga accgcatcga gctgaagggc atcgacttca aggaggacgg 1320
 caacatcctg gggcacaagc tggagtaca ctacaacagc cacaacgtct atatcatggc 1380
 cgacaagcag aagaacggca tcaagtgaa ctcaagatc cgccacaaca tcgaggacgg 1440

cagcgtgcag ctgcccacc actaccagca gaacaccccc atcggcgacg gccccgtgct 1500
 gctgcccgac aaccactacc tgagcacca gtccgccctg agcaaagacc ccaacgagaa 1560
 gcgcatcac atggtctcgt tggagtctg gaccgcccgg gggatcactc tcggcatgga 1620
 cgagctgtac aagtaaagcg gcccggggga tccagacatg ataagataca ttgatgagtt 1680
 tggacaaacc acaactagaa tgacgtgaaa aaaatgcttt atttgtgaaa tttgtgatgc 1740
 tattgcttta tttgtaacca ttataagctg caataaaca gttacaaca acaattgcat 1800

tcattttatg tttcaggttc agggggagggt gtgggagggt ttttagtcga ccatttaaat 1860

cggtctcgag ggtatcagca cacaattgcc cattatacgc gcgtataatg gactattgtg 1920

tgctgataga agcttatcgg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 1980

cgtaaaaagg ccgcgttgct ggcggttttc cataggctcc gccccctga cgagcatcac 2040

aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg 2100

tttccccctg gaagctccct cgtgcgctct cctgttccga cctgcccgt taccggatac 2160

ctgtccgctt ttctccctc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat 2220

ctcagttcgg ttaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 2280

cccgaccgct gcgccttacc cggtaactat cgtcttgagt ccaaccggt aagacacgac 2340

ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggt ttaggcggt 2400

gctacagagt tctgaagtg ttggcctaac tacggctaca ctagaagaac agtatttgg 2460

atctgcgctc tgctgaagcc agttacctc ggaaaaagag ttgtagctc ttgatccggc 2520

aaacaaacca ccgctggtag cgggtggttt tttgtttgca agcagcagat tacgcgcaga 2580

aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac 2640

gaaaactcac gtaaggat tttggtcatg agattatcaa aaaggatctt cacctagatc 2700

cttttaaat aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggctc 2760

gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca 2820

tccatagtgt cctgactccc cgctgtgtag ataactacga tacgggaggg cttaccatct 2880

ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggetccaga tttatcagca 2940

ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgctcc 3000

atccagtcta ttaattgtg ccgggaagct agagtaagta gttcgccagt taatagttg 3060

cgcaacgttg ttgccattgc tacaggcatc gtgggtgtcac gctcgtcgtt tggtaggct 3120

tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat gttgtgcaaa 3180

aaagcggtta gtccttcgg tcctccgacg gttgtcagaa gtaagttggc cgcagtgtta 3240

tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 3300

ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 3360

agttgtctct gcccgcgctc aatacgggat aataccgctc cacatagcag aactttaaaa 3420

gtgctcatca ttgaaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 3480

agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttacttcc 3540

accagcgttt ctgggtgagc aaaaacagga aggcaaatg cgcacaaaaa gggaataagg 3600
 gcgacacgga aatgttgaat actcactc ttctttttc aatattattg aagcatttat 3660
 cagggttatt gtctcatgag cggatacata ttggaatgta tttagaaaaa taaacaata 3720
 ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tc 3762

<210> 21
 <211> 4695
 <212> DNA
 <213> Artificial Sequence
 <220><223> Luc plasmid
 <400> 21

ccaatgatgg taccgtatca gcacacaatt gccattata cgcgcgtata atggactatt 60
 gtgtgctgat aggtctcgcc ctgatttaa tcgtcgaccg ttacataact tacggtaaat 120

ggccccctg gctgaccgcc caacacccc cgccattga cgtcaataat gacgtatgtt 180
 cccatagtaa cgccaatagg gactttccat tgacgtcaat ggggtggagta ttacggtaa 240
 actgcccact tggcagtaca tcaagtgtat catatgcaa gtacgcccc tattgacgtc 300
 aatgacggta aatggccgc ctggcattat gccagtaca tgacctatg ggactttcct 360
 acttggcagt acatctacgt attagtcatc gctattacca tggatgatgc gttttggcag 420
 tacatcaatg ggcgtggata gcggtttgac tcacggggat ttccaagtct ccacccatt 480
 gacgtcaatg ggagtttgtt ttggcaccaa aatcaacggg actttccaaa atgtcgtaac 540

aactccgcc cattgacgca aatggcggt aggcgtgtac ggtgggaggt ctatataagc 600
 agagctcgt tagtgaaccg tcagatcgcc tggagacgcc atccacgtg ttttgacctc 660
 catagaagac accgggaccg atccagcctc cggactctag aggatccgt actcgaggaa 720
 ctgaaaaacc agaaagttaa ctggttaagt tagtcttttt gtcttttatt tcagggtccc 780
 gatccggtg tggtgcaaat caaagaactg ctctcagtg gatgttcct ttacttctag 840
 gcctgtacgg aagtgttact tctgctctaa aagctgcgga attgtaccg cggccccgga 900
 tccaccggtc gccaccatgg aagacgcaa aaacataaag aaaggccccg gccatttcta 960

tcctctagag gatggaaccg ctggagagca actgcataag gctatgaaga gatacgcct 1020
 ggttctgga acaattgctt ttacagatgc acatatcgag gtgaacatca cgtacgcgga 1080
 atactcgaa atgtccgttc ggttggcaga agctatgaaa cgatatgggc tgaatacaaa 1140
 tcacagaatc gtcgtatgca gtgaaaactc tcttcaattc tttatgccgg tgttgggcgc 1200
 gtattttatc ggagttgcag ttgcggccgc gaacgacatt tataatgaac gtgaattgct 1260
 caacagtatg aacatttcgc agcctaccgt agtgtttgtt tccaaaaagg ggttgcaaaa 1320

aattttgaac gtgcaaaaaa aattaccaat aatccagaaa attattatca tggattctaa 1380

aacggattac cagggatttc agtcgatgta cacgttcgtc acatctcatc tacctcccgg 1440
ttttaatgaa tacgattttg taccagagtc ctttgatcgt gacaaaacaa ttgcaactgat 1500
aatgaattcc tciggatcta ctgggttacc taagggtgtg gcccttccgc atagaactgc 1560
ctgcgtcaga ttctcgcagc ccagagatcc ttttttggc aatcaaatca ttccggatac 1620
tgcgatttta agtgttgttc cattccatca cggttttgga atgtttacta cactcggata 1680
tttgatatgt ggatttcgag tcgtcttaat gtatagattt gaagaagagc tgtttttacg 1740
atcccttcag gattacaaaa ttcaaagtgc gttgctagta ccaaccctat tttcattctt 1800

cgccaaaagc acictgattg acaaatacga tttatctaat ttacacgaaa ttgcttctgg 1860
gggcgcacct ctttcgaaag aagtcgggga agcggttgca aaacgcttcc atcttcagg 1920
gatacgacaa ggatatgggc tcactgagac tacatcagct attctgatta cacccgaggg 1980
ggatgataaa ccgggcgcgg tcggtaaagt tgttccattt ttgaaagca aggttgtgga 2040
tctggatacc gggaaaacgc tgggcgttaa tcagagaggc gaattatgtg tcagaggacc 2100
tatgattatg tccggttatg taaacaatcc ggaagcgacc aacgccttga ttgacaagga 2160
tggatggcta cattctggag acatagctta ctgggacgaa gacgaacact tcttcatagt 2220

tgaccgcttg aagtctttaa ttaaatacaa aggatafcag gtggcccccg ctgaattgga 2280
atcgatattg ttacaacacc ccaacatctt cgacgcgggc gtggcaggtc ttcccacga 2340
tgacccggt gaacttcccg ccgccgttgt tgttttggag cacgaaaga cgatgacgga 2400
aaaagagatc gtggattacg tcgccagtca agtaacaacc gcgaaaaagt tgcgcccagg 2460
agttgtgttt gtggacgaag taccgaaagg tcttaccgga aaactcgacg caagaaaaat 2520
cagagagatc ctataaagg ccaagaaggg cggaaagtcc aaattgtaa gcggcccgg 2580
ggatccagac atgataagat acattgatga gtttggacaa accacaacta gaatgcagtg 2640

aaaaaatgc tttatttggg aaatttgtga tgctattgct ttatttghaa ccattataag 2700
ctgcaataaa caagttaaca acaacaattg cattcatctt atgtttcagg ttcaggggga 2760
ggtgtgggag gttttttagt cgaccattta aatcggcttc gaggtatca gcacacaatt 2820
gccattata cgcgcgtata atggactatt gtgtgctgat agaagcttat cggaaagaac 2880
atgtgagcaa aaggccagca aaaggccagg aaccgtaaaa aggcccgctt gctggcgttt 2940
ttccataggc tccgcccc tgacgagcat cacaaaaatc gacgctcaag tcagaggtgg 3000
cgaaacccga caggactata aagataccag gcgtttcccc ctggaagctc cctcgtgcgc 3060

tctcctgttc cgaccctgcc gcttaccgga tacctgtccg cttttctccc ttcgggaagc 3120
 gtggcgcttt ctcatagctc acgtgtagg tatctcagtt cgggttaggt cgttcgctcc 3180
 aagctgggct gtgtgcacga accccccgtt cagccccgacc gctgcgctt atccggtaac 3240
 tatcgtcttg agtccaacct ggtaagacac gacttatcgc cactggcagc agccactggt 3300
 aacaggatta gcagagcgag gtatgtaggc ggtgctacag agttctttaa gtggtgcct 3360
 aactacggtt aactagaag aacagtattt ggtatctgcg ctctgctgaa gccagttacc 3420
 ttcggaaaaa gagtggtag ctcttgatcc ggcaaaaaa ccaccgctgg tagcggtggt 3480

ttttttgttt gcaagcagca gattacgcgc agaaaaaaag gatctcaaga agatcctttg 3540
 atcttttcta cggggtctga cgctcagtgg aacgaaaact cacgttaagg gatttttgct 3600
 atgagattat caaaaaggat cttcacctag atccttttaa attaaaaatg aagttttaa 3660
 tcaatctaaa gtatatatga gtaaacttgg tctgacagtt accaatgctt aatcagtgag 3720
 gcacctatct cagcgatctg tctatttctg tcatccatag ttgctgact ccccgctgtg 3780
 tagataacta cgatacggga gggcttacca tctggcccca gtgctgcaat gataccgca 3840
 gaccacgct caccgctcc agatttatca gcaataaacc agccagccgg aaggcccgag 3900

cgcagaagtg gtctgcaac tttatccgc tccatccagt ctattaattg ttgccgggaa 3960
 gctagagtaa gtagttcgc agttaatagt ttgcgcaac ttgttgccat tgctacaggc 4020
 atcgtgggtg cagctcctg gtttggatg gcttcattca gctccggtt ccaacgatca 4080
 aggcgagtta catgatcccc catgttgtgc aaaaaagcgg ttagctcctt cggctcctccg 4140
 atcgttgtea gaagtaagtt ggccgcagtg ttatcactca tggttatggc agcactgcat 4200
 aattctctta ctgcatgcc atccgtaaga tgcttttctg tgactggtga gtactcaacc 4260
 aagtcattct gagaatagt gtagcggcga ccgagttgct cttgcccggc gtcaatcagg 4320

gataataccg cgccacatag cagaacttta aaagtgetca tcattggaaa acgttcttctg 4380
 gggcgaaaac tctcaaggat cttaccgctg ttgagatcca gttcgatgta acccactcgt 4440
 gcaccaact gatcttcagc atcttttact ttaccagcg tttctgggtg agcaaaaaa 4500
 ggaaggcaaa atgccgcaaa aaaggaata agggcgacac ggaaatgtg aataactcata 4560
 ctcttctttt ttcaatatta ttgaagcatt taccagggtt attgtctcat gagcggatac 4620
 atatttgaat gtatttagaa aaataaaca ataggggttc cgcgcacatt tccccgaaa 4680
 gtgccacctg acgtc 4695

<210> 22

<211> 4902

<212> DNA

<213> Artificial Sequence

<220><223> Luc-ITR plasmid

<400> 22

ccaatgatgg taccgtatca gcacacaatt gccattata cgcgcgtata atggactatt 60
 gtgtgctgat agataggctc cgccctcctg caggcagctg cgcgctcgct cgctcactga 120
 ggccgcccgg gcaaagcccc ggcgctgggc gacctttggt cggccggcct cagtgagcga 180
 gcgagcgcgc agagaggaggg tggccaactc catcactagg ggttcctgcg gccgcacgcg 240
 tcgttacata acttacgga aatggccccg ctggctgacc gccaacgac ccccgccat 300
 tgacgtcaat agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac 360

 ggtaaactgc ccaacttgca gtacatcaag tgtatcatat gccaagtacg ccccctattg 420
 acgtcaatga cggtaaatgg cccgctggc attgtgcca gtacatgacc ttatgggact 480
 ttctacttg gcagtacatc tacgtattag tcatcgctat taccatgggtg atgcggtttt 540
 ggcagtacat caatgggctg ggatagcggg ttgactcacg gggatttcca agtctccacc 600
 ccattgacgt caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgac 660
 gtaacaactc cgccccattg acgcaaatgg gcggtaggcg tgtacggtgg gaggtctata 720
 taagcagagc tgctagcttg gcattccggt actgttggtg aagccaccat ggaagacgcc 780

 aaaaacataa agaaaggccc ggcgccattc tatccgctgg aagatggaac cgctggagag 840
 caactgcata aggctatgaa gagatagccc ctggttcctg gaacaattgc ttttacagat 900
 gcacatatcg aggtggacat cacttacgct gagtacttcg aaatgtccgt tcggttggca 960
 gaagctatga aacgatatgg gctgaataca aatcacagaa tcgtcgatg cagtgaaaac 1020
 tctcttcaat tctttatgcc ggtgttgggc gcgttattta tcggagttgc agttgcgcc 1080
 gcgaacgaca ttataatga acgtgaattg ctcaacagta tgggcatttc gcagcctacc 1140
 gtggtgttcg ttccaaaaa ggggttgcaa aaaattttga acgtgcaaaa aaagctccca 1200

 atcatccaaa aaattattat catggattct aaaacggatt accagggatt tcagtcgatg 1260
 tacacgttcg teacatctca tctacctecc ggttttaatg aatacgattt tgtgccagag 1320
 tccttcgata gggacaagac aattgcactg atcatgaact cctctggatc tactggtctg 1380
 cctaaaggtg tcgctctgcc tcatagaact gcctgcgtga gattctcgca tgccagagat 1440
 cctatTTTTG gcaatcaaat cattccggat actgcgattt taagtgttgt tccattccat 1500
 cacggttttg gaatgtttac tacactcgga tatttgatg gtggatttcg agtcgtctta 1560
 atgtatagat ttgaagaaga gctgtttctg aggagccttc aggattacaa gattcaaagt 1620

 gcgctgctgg tgccaacctt atttcccttc ttgcctcaaaa gcactctgat tgacaatac 1680

gatttatcta atttacacga aattgcttct ggtggcgctc ccctctctaa ggaagtcggg 1740
gaagcggttg ccaagaggtt ccactcgcca ggtatcaggc aaggatatgg gctcactgag 1800
actacatcag ctattctgat tacacccgag ggggatgata aaccgggagc ggtcggtaaa 1860
gttgttccat ttttgaagc gaaggttgtg gatctggata ccgggaaaac gctgggcgtt 1920
aatcaagag gcgaactgtg tgtgagaggt cctatgatta tgtccggtta tgtaaacat 1980
ccggaagcga ccaacgcctt gattgacaag gatggatggc tacattctgg agacatagct 2040

tactgggacg aagacgaaca cttcttcac c gttgaccgcc tgaagtctct gattaagtac 2100
aaaggctatc aggtggctcc cgctgaattg gaatccatct tgctccaaca cccaacatc 2160
ttcgacgcag gtgtcgcagg tcttcccgac gatgacgccg gtgaacttcc cgcccgctt 2220
gttgttttgg agcacggaaa gacgatgacg gaaaaagaga tcgtggatta cgtcgccagt 2280
caagtaacaa ccgcgaaaaa gttgcgcgga ggagttgtgt ttgtggacga agtaccgaaa 2340
ggtcttaccg gaaaactcga cgcaagaaaa atcagagaga tcctcataaa ggccaagaag 2400
ggcggaaaaga tcgccgtgtc aagccatggc ttcccgcggc cggaggcggc gcaggatgat 2460

ggcacgctgc ccattgcttg tgcccaggag agcgggatgg accgtcaccc tgcagcctgt 2520
gcttctgcta ggatcaatgt gtaggcggcc aacaatgcga tccgatggcc gcgactctag 2580
agtcggggcg gccggccgct tcgagcagac atgataagat acattgatga gtttgacaa 2640
accacaacta gaatgcagtg aaaaaaatgc tttatttgtg aaatttgtga tgctattgct 2700
ttatttghaa ccattataag ctgcaataaa caagttaaca acaacaattg cattcatttt 2760
atgtttcagg ttcaggggga ggtgtgggag gttttttaa gcaagtaaaa cctctacaaa 2820
tgtggtaaaa tcgataagcc cgtgcggacc gagcggccgc aggaaccctt agtgatggag 2880

ttggccactc cctctctcgg cgctcgtctg ctactgagg ccgggcgacc aaaggtcgcc 2940
cgacgcccg gctttgccc ggccgcctca gtgagcgagc gagcgcgag ctgcctgcag 3000
gggtctcgag ggtatcagca cacaattgcc cattatacgc gcgtataatg gactattgtg 3060
tgctgataga agcttatcgg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 3120
cgtaaaaagg ccgctgtgct ggcgtttttc cataggetcc gccccctga cgageatcac 3180
aaaaatcgac gctcaagtca gaggtggcga aaccgacag gactataaag ataccaggcg 3240
tttccccctg gaagctccct cgtgcgctct cctgttccga cctgcccgt taccggatac 3300

ctgtccgctt ttctccctc gggaagcgtg gcgctttctc atagctcacg ctgtaggtat 3360
ctcagttcgg ttaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 3420
cccgaccgct gcgccttacc cggttaactat cgtcttgagt ccaaccggt aagacacgac 3480
ttatcggcac tggcagcagc cactggtaac aggattagca gagcgagta ttaggcggt 3540

gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaagaac agtatttgg 3600
atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc 3660
aaacaaacca ccgctggtag cggtaggtttt tttgtttgca agcagcagat tacgcgcaga 3720

aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac 3780
gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc 3840
cttttaatt aaaaatgaag ttttaaatca atctaaagta tataatgagta aacttggctc 3900
gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca 3960
tccatagtgg cctgactccc cgctgtgtag ataactacga tacgggaggg cttaccatct 4020
ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca 4080
ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc 4140

atccagtcta ttaattgttg cggggaagct agagtaagta gttcgccagt taatagtgtg 4200
cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggatggct 4260
tcattcagct ccggttccca acgatcaagg cgagttacat gatccccat gttgtgcaaa 4320
aaagcggtta gtccttccgg tcctccgatc gttgtcagaa gtaagtggc cgcagtgtta 4380
tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 4440
ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gggcgaccg 4500
agttgctctt gcccgcgctc aatacgggat aataccgcgc cacatagcag aactttaaaa 4560

gigctcatca ttgaaaaacg ttcttcgggg cgaaaactct caaggatctt accgctgttg 4620
agatccagtt cgatgtaacc cactcgtgca cccaactgat cttcagcatc ttttactttc 4680
accagcgttt ctgggtgagc aaaaacagga aggcaaatg ccgcaaaaaa gggaataagg 4740
gcgacacgga aatgttgaat actcatactc ttctttttc aatattattg aagcatttat 4800
cagggttatt gtctcatgag cggatacata tttgaatgta tttagaaaaa taaacaata 4860
ggggttccgc gcacatttcc ccgaaaagtg ccacctgacg tc 4902

<210> 23
<211> 4403
<212> DNA
<213> Artificial Sequence

<220><223> pUC57-Kan_TEL0-CMV-EGFP
<400> 23
tcgcgcgttt cggatgatgac ggtgaaaacc tetgacacat gcagctcccg gagacggtea 60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcaggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180
 accatatgcg gtgtgaaata ccgcacagat gcgtaaggag aaaataccgc atcaggcgcc 240
 attcgcatt caggctgcbc aactgttggg aagggcgatc ggtgcbggcc tcttcgctat 300
 tacgccagct ggcgaaaggg ggatgtgctg caagcgatt aagtgggta acgccagggt 360
 tttcccagtc acgacgttgt aaaacgacgg ccagagaatt cgtatcagca cacaattgcc 420

 cattatacgc gcgtataatg gactattgtg tgctgataga tttaaatcgt cgaccgttac 480
 ataacttacg gtaaatggcc cgcctggctg accgcccaac gacccccgcc cattgacgtc 540
 aataatgacg tatgttccca tagtaacgcc aatagggact ttccattgac gtcaatgggt 600
 ggagtattta cggtaaacctg cccacttggc agtacaatca gtgtatcata tgccaagtac 660
 gccccctatt gacgtcaatg acggtaaatg gcccgctgg cattatgccc agtacaatgac 720
 cttatgggac tttcctactt ggacgtacat ctacgtatta gtcatcgcta ttaccatggt 780
 gatgcbggtt tggcagtaca tcaatggcg tgatagcgg tttgactcac ggggatttcc 840

 aagtctccac ccattgacg tcaatgggag tttgttttg caccaaaatc aacgggactt 900
 tccaaaatgt cgtaaacact ccgccccatt gacgcaaatg ggcggtagcc gtgtacggtg 960
 ggaggtctat ataagcagag ctcgtttagt gaaccgtcag atcgctgga gacgcatcc 1020
 acgtgtttt gacctccata gaagacaccg ggaccgatcc agcctccgga ctctagagga 1080
 tccggtactc gaggaactga aaaaccagaa agttaactgg taagtttagt cttttgtct 1140
 tttatttcag gtcccgatc cgggtggtgt gcaaatcaaa gaactgctcc tcagtggatg 1200
 ttgcctttac ttctaggcct gtacggaagt gttacttctg ctctaaaagc tgcggaattg 1260

 taccgcggc ccgggatcca ccggtcgcca ccatggtgag caagggcgag gagctgttca 1320
 ccgggtggt gccatcctg gtcgagctgg acggcgact aaacggccac aagttcagcg 1380
 tgtccggca ggcgagggc gatgccact acggcaagct gacctgaag ttcactgca 1440
 ccaccggcaa gtcgccgtg ccctggcca cctcgtgac caccctgacc tacggcgtgc 1500
 agtgtttcag ccgtacccc gaccacatga agcagcacga cttcttaag tccgcatgc 1560
 ccgaaggcta cgtccaggag cgcaccatct tcttaagga cgacggcaac tacaagacc 1620
 gcgccgaggt gaagttcgag ggcgacacc tggtaaccg catcgagctg aaggcatcg 1680

 acttcaagga ggacggcaac atcctggggc acaagctgga gtacaactac aacagccaca 1740
 acgtctatat catggccgac aagcagaaga acggcatcaa ggtgaacttc aagatccgcc 1800
 acaacatcga ggacggcagc gtgcagctcg ccgaccacta ccagcagaac accccatcg 1860
 gcgacggccc cgtgctgctg cccgacaacc actacctgag caccagtcg gccctgagca 1920
 aagaccccaa cgagaagcgc gatcacatgg tctgctgga gttcgtgacc gccgcbgga 1980

tcaactctcgg catggacgag ctgtacaagt aaagcggccg cggggatcca gacatgataa 2040
 gatacattga tgagtttggg caaaccacaa ctagaatgca gtgaaaaaaaa tgctttat 2100

 gtgaaatttg tgatgctatt gctttat 2160
 taaccattat aagctgcaat aaacaagtta
 acaacaacaa ttgacattcat tttatgtttc aggttcaggg ggaggtgtgg gaggtttttt 2220
 agtcgacat ttaaatctat cagcacacaa ttgccatta tacgcgctg taatggacta 2280
 ttgtgtgctg atagaagctt ggtgtaatca tggcatagc tgtttcctgt gtgaaattgt 2340
 tatccgctca caattccaca caacatacga gccggaagca taaagtgtaa agcctgggg 2400
 gcctaagtag tgagctaact cacattaatt gcgttgctg cactgcccgc tttccagtcg 2460
 ggaaacctgt cgtgccagct gcattaatga atcggccaac gcgcggggag aggcggtttg 2520

 cgtattgggc gctcttcgc ttcctcgtc actgactgc tgcgctcggg cgttcggctg 2580
 cggcgagcgg tatcagetca ctcaaaggcg gtaatacggg tatccacaga atcaggggat 2640
 aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 2700
 gcgttgctgg cgtttttcca taggctcgc cccctgacg agcatcaca aaatcgacgc 2760
 tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tcccctgga 2820
 agtccctcgc tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgcttt 2880
 ctcccttcgg gaagcgtggc gctttctcat agctcacgct gtaggtatct cagttcggtg 2940

 taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc 3000
 gccttatccg gtaactatcg tcttgagtcc aaccggtaa gacacgactt atcgccactg 3060
 gcagcagcca ctgtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc 3120
 ttgaagtggg ggctaacta cggctacact agaagaacag tatttggat ctgcgctctg 3180
 ctgaagccag ttacctcgg aaaaagagtt gtagctctt gatccggcaa acaaaccacc 3240
 gctggtagcg gtggttttt tgtttgcaag cagcagatta cgccagaaa aaaaggatct 3300
 caagaagatc cttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt 3360

 taagggattt tggcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa 3420
 aaatgaagtt taaatcaag cccaatctga ataatgttac aaccaattaa ccaattctga 3480
 ttagaaaaac tcatcgagca tcaaatgaaa ctgcaattta tcatatcag gattatcaat 3540
 accatatttt tgaaaaagcc gtttctgtaa tgaaggagaa aactcaccga ggcagttcca 3600
 taggatggca agatcctggg atcggctctg gattccgact cgtccaacat caatacaacc 3660
 tattaatttc cctcgtcaa aaataaggtt atcaagtgag aaatcaccat gagtgcagac 3720

tgaatccggt gagaatggca aaagtttatg cttttctttc cagacttggt caacaggcca	3780
gccattacgc tgcicatcaa aatcactcgc atcaacccaaa ccgttattca ttcgtgattg	3840
cgcttgagcg agacgaaata cgcgatcgtt gttaaaagga caattacaaa caggaatcga	3900
atgcaaccgg cgcaggaaca ctgccagcgc atcaacaata ttttcacctg aatcaggata	3960
tttttctaata acctggaatg ctgtttttcc ggggatcgca gtggtgagta accatgcac	4020
atcaggagta cggataaaat gcttgatggt cggagagagc ataaattccg tcagccagtt	4080
tagtctgacc atctcatctg taacatcatt ggcaacgcta cctttgccat gtttcagaaa	4140
caactctggc gcatcgggct tccatacaa gcgatagatt gtcgcacctg attgcccgc	4200
attatcgga gccatttat acccatataa atcagcatcc atgttggaaat ttaatcgagg	4260
cctcgacgtt tcccgttgaa tatggetcat aacaccctt gtattactgt ttatgtaagc	4320
agacagtttt attgttcatg atgatataat tttatcttgt gcaatgtaac atcagagatt	4380
ttgagacacg ggccagagct gca	4403