



등록특허 10-2522263



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2023년04월18일

(11) 등록번호 10-2522263

(24) 등록일자 2023년04월12일

(51) 국제특허분류(Int. Cl.)  
C12N 9/12 (2006.01) C12P 19/34 (2006.01)  
(52) CPC특허분류  
C12N 9/1252 (2013.01)  
C12N 9/1264 (2013.01)  
(21) 출원번호 10-2018-7034921  
(22) 출원일자(국제) 2017년06월13일  
심사청구일자 2020년06월04일  
(85) 번역문제출일자 2018년11월30일  
(65) 공개번호 10-2019-0024883  
(43) 공개일자 2019년03월08일  
(86) 국제출원번호 PCT/FR2017/051519  
(87) 국제공개번호 WO 2017/216472  
국제공개일자 2017년12월21일  
(30) 우선권주장  
1655475 2016년06월14일 프랑스(FR)  
(56) 선행기술조사문헌  
US20160108382 A1\*  
WO2016064880 A1  
비특허문헌 1  
\*는 심사관에 의하여 인용된 문헌

(73) 특허권자  
디엔에이 스크립트  
프랑스, 94270 르 크렘린-비세르트르, 67 아브뉴 드  
퐁텐블로, 이머블르 오카베  
앵스티띠 파스퇴르  
프랑스공화국 파리 세텍스 15 에프-75724 뤼 드락  
터르 루 25-28  
상뜨로 나쇼날 드 라 러쉴르쉴 상띠피크  
프랑스 파리 에프-75016 뤼 미셀 앙즈 3  
(72) 발명자  
이버트 토마스  
프랑스, 75012 파리, 4 뤼 데 라 베가  
텔라뤼 마크  
프랑스, 78000 베르사유, 98 블러바드 데 라 레인  
느  
(74) 대리인  
제일특허법인(유)

전체 청구항 수 : 총 16 항

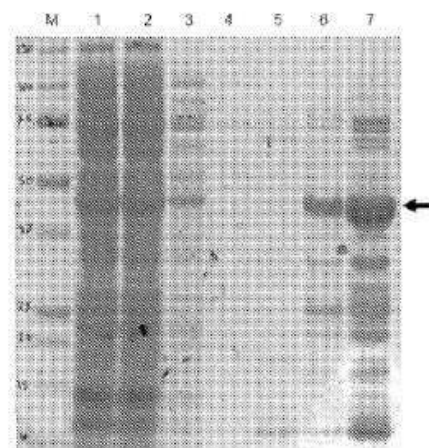
심사관 : 한지혜

(54) 발명의 명칭 poIX 패밀리의 DNA 폴리머라제의 변형체

### (57) 요약

본 발명은 적어도 하나의 특정 위치 내 잔기의 적어도 하나의 돌연변이를 포함하는, 주형 가닥 없이 핵산 분자를 합성할 수 있는 poIX 패밀리의 DNA 폴리머라제의 변형체, 또는 이러한 폴리머라제의 기능적 단편의 변형체, 및 특히 3'-OH 변형된 뉴클레오타이드를 포함하는 핵산 분자의 합성을 위한 상기 변형체의 용도에 관한 것이다.

대표도 - 도1



(52) CPC특허분류

*C12P 19/34* (2013.01)

*C12Y 207/07031* (2013.01)

---

## 명세서

### 청구범위

#### 청구항 1

주형 가닥(template strand) 없이 핵산 분자를 합성할 수 있는, 말단 데옥시리보뉴클레오티딜 트랜스퍼라제(TdT)의 변형체로서,

상기 변형체가 서열번호 1에 기재된 서열을 포함하되, R336N-E457N, R336N-R454A-E457N, R336N-R454A-E457G, R336N-E457G, R336G-R454A-E457N, 및 R336G-E457N으로부터 선택되는 치환의 조합을 포함하고, 상기 표시된 아미노산 위치가 서열번호 1을 기준으로 결정되는, TdT의 변형체.

#### 청구항 2

제1항에 있어서, 상기 변형체가 DNA 가닥, RNA 가닥, 또는 DNA 가닥 및 RNA 가닥 둘 모두를 합성할 수 있는 것인, TdT의 변형체.

#### 청구항 3

삭제

#### 청구항 4

제1항 또는 제2항에 있어서, 상기 변형체가 서열번호: 1 서열의 TdT의 변형체이고, C378 내지 L406 위치 사이의 잔기를 서열번호: 2 서열의 폴리머라제 Pol  $\mu$ 의 H363 내지 C390의 잔기로 치환하는 것을 추가로 포함하는 것인, TdT의 변형체.

#### 청구항 5

제1항 또는 제2항에 따른 TdT의 변형체를 코딩하는 핵산.

#### 청구항 6

제5항에 따른 핵산을 위한 발현 카세트.

#### 청구항 7

제5항에 따른 핵산 또는 상기 핵산을 위한 발현 카세트를 포함하는, 벡터.

#### 청구항 8

단리된 숙주 세포로서, 제5항에 따른 핵산; 또는 상기 핵산을 위한 발현 카세트; 또는 상기 핵산 또는 상기 발현 카세트를 포함하는 벡터를 포함하는, 단리된 숙주 세포.

#### 청구항 9

제1항 또는 제2항에 따른 TdT의 변형체를 생산하는 방법으로서, 상기 방법은 상기 TdT의 변형체를 코딩하는 핵산; 상기 핵산을 위한 발현 카세트; 상기 핵산 또는 상기 발현 카세트를 포함하는 벡터; 또는 상기 핵산, 상기 발현 카세트 또는 상기 벡터를 포함하는 단리된 세포에 의해 실시되는 방법.

#### 청구항 10

제1항 또는 제2항에 따른 TdT의 변형체를 생산하는 방법으로서, 단리된 숙주 세포가 상기 변형체를 코딩하는 핵산의 발현을 가능케 하는 배양 조건에서 배양되고, 선택적으로 그에 따라 발현된 변형체가 배양 배지 또는 상기 숙주 세포로부터 회수되고, 상기 숙주 세포가 상기 TdT의 변형체를 코딩하는 핵산; 또는 상기 핵산을 위한 발현 카세트; 또는 상기 핵산 또는 상기 발현 카세트를 포함하는 벡터를 포함하는 것인, 방법.

#### 청구항 11

주형 가닥 없이, 3'-OH-변형된 뉴클레오티드로부터 핵산 분자를 합성하는 방법으로서, 상기 방법이 제1항 또는 제2항의 TdT의 변형체에 의해 실시되는 방법.

#### 청구항 12

제11항에 있어서, DNA 가닥 또는 RNA 가닥을 합성하기 위한 것인, 방법.

#### 청구항 13

주형 가닥 없이 핵산 분자를 효소적으로 합성하는 방법으로서, 프라이머 가닥이 제1항 또는 제2항의 TdT의 변형체의 존재 하에서 적어도 하나의 뉴클레오티드와 접촉하게 되는, 방법.

#### 청구항 14

제13항에 있어서, 상기 뉴클레오티드가 3'-OH-변형된 뉴클레오티드인, 방법.

#### 청구항 15

주형 가닥 없이 핵산 분자를 효소적으로 합성하기 위한 키트로서, 제1항 또는 제2항에 따른 TdT의 적어도 하나의 변형체, 및 뉴클레오티드를 포함하는 것인, 키트.

#### 청구항 16

제15항에 있어서, 상기 뉴클레오티드가 3'-OH-변형된 뉴클레오티드인, 키트.

#### 청구항 17

제15항에 있어서, 상기 키트가 적어도 하나의 뉴클레오티드 프라이머를 추가로 포함하는 것인, 키트.

#### 청구항 18

삭제

#### 청구항 19

삭제

#### 청구항 20

삭제

#### 청구항 21

삭제

#### 청구항 22

삭제

#### 청구항 23

삭제

#### 청구항 24

삭제

#### 청구항 25

삭제

## 발명의 설명

### 기술 분야

- [0001] 본 발명은 효소 개선 분야에 관한 것이다. 본 발명은 polX 패밀리의 DNA 폴리머라제의 개선된 변형체, 이 변형체를 코딩하는 핵산, 숙주 세포에서 이 변형체의 생산, 주형 가닥 없이 핵산 분자의 합성을 위한 이의 용도, 및 주형 가닥 없이 핵산 분자의 합성을 위한 키트에 관한 것이다.

### 배경 기술

- [0002] 핵산 단편의 화학적 합성은 광범위하게 사용되는 실험실 기법이다 (Adams et al., 1983, J. Amer. Chem. Soc. 105: 661; Froehler et al., 1983; Tetrahedron Lett. 24: 3171). 이는 원하는 뉴클레오타이드 서열을 포함하는 핵산 분자를 신속하게 수득할 수 있게 한다. 5' 내지 3' 방향으로 합성을 수행하는 효소와는 달리, 화학적 합성은 3' 내지 5' 방향으로 수행된다. 그러나, 화학적 합성은 일정한 한계를 갖는다. 실제로, 이는 다수의 용매 및 시약의 사용을 필요로 한다. 또한, 이는 원하는 최종 핵산 가닥을 수득하기 위해 다른 하나와 조립되어야만 하는 짧은 핵산 단편을 수득할 수 있게 한다.
- [0003] 주형 가닥의 부재하에서 초기 핵산 단편 (프라이머)으로부터 뉴클레오타이드 사이의 커플링 반응을 수행하는 효소를 사용하는 대안적인 해법이 개발되었다. 몇몇 폴리머라제 효소는 이러한 유형의 합성 방법에 적합한 것처럼 보인다.
- [0004] 주형 가닥의 존재 또는 부재하에서 핵산 가닥의 합성을 촉매할 수 있는 매우 많은 수의 DNA 폴리머라제가 존재한다. 따라서, polX 패밀리의 DNA 폴리머라제는 다양한 범위의 생물학적 과정, 특히 DNA 복구 메커니즘 또는 DNA 서열에서 나타나는 오류의 정정을 위한 메커니즘에 관여한다. 이들 효소는 뉴클레오타이드를 삽입할 수 있고, 이는 핵산 가닥에서 서열 오류의 확인 후 제거되었다. polX 패밀리의 DNA 폴리머라제는 DNA 폴리머라제  $\beta$  (Pol  $\beta$ ),  $\lambda$  (Pol  $\lambda$ ),  $\mu$  (Pol  $\mu$ ), 효모 IV (Pol IV), 및 말단 데옥시리보뉴클레오타이드 트랜스퍼라제 (TdT)를 포함한다. 특히 TdT는 핵산 분자의 효소적 합성 방법에 매우 광범위하게 사용된다.
- [0005] 그러나 일반적으로 이들 DNA 폴리머라제는 천연 뉴클레오타이드의 혼입만을 허용한다. 모든 경우에, 천연 DNA 폴리머라제는 천연 뉴클레오타이드보다 더 큰 입체 장애를 나타내는 비-천연 뉴클레오타이드, 특히 3'-OH 변형된 뉴클레오타이드의 존재시에 그들의 촉매 활성을 상실한다.
- [0006] 그러나 변형된 뉴클레오타이드의 사용은 일부 특정 적용에 유용한 것으로 확인될 수 있다. 따라서, 이러한 뉴클레오타이드의 혼입에 의한 핵산 가닥의 합성을 촉매할 수 있는 효소가 개발되어야 한다. 이에 상당한 구조적 변형을 포함하는 뉴클레오타이드와 기능할 수 있는 DNA 폴리머라제 변형체가 개발되었다.
- [0007] 그러나 현재 이용가능한 변형체는, 특히 낮은 활성을 나타내고 실험실 규모의 효소적 합성에만 적합하기 때문에, 이들은 전혀 만족스럽지 않다. 따라서, 가능하다면 산업적 규모에서, 주형 가닥의 부재 시에 변형된 뉴클레오타이드를 사용하여 핵산을 합성할 수 있는 DNA 폴리머라제가 요구된다.

## 발명의 내용

### 해결하려는 과제

- [0008] 본 발명은 핵산의 효소적 합성을 위한 DNA 폴리머라제의 상업적 규모로의 사용을 저해하는 특정한 기술적 장벽을 극복한다.
- [0009] 따라서 본 발명은 주형 가닥(template strand)의 부재하에 변형된 뉴클레오타이드를 사용하는데 적합한 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제를 제안한다. 개발된 변형체는 그들이 유래하는 천연 DNA 폴리머라제에 비해 훨씬 더 큰 변형된 뉴클레오타이드의 혼입 능력을 나타낸다. 구체적으로, 본 발명의 주제인 DNA 폴리머라제 변형체는 당의 변형을 갖는 뉴클레오타이드의 혼입에 특히 효과적이다. 실제로, 본 발명자들은 그들이 유래하는 DNA 폴리머라제에 비해 증가된 촉매 포켓 용량 (catalytic pocket volume)을 갖고, 천연 뉴클레오타이드보다 더 큰 입체 장애를 나타내는 변형된 뉴클레오타이드의 혼입을 촉진하는 변형체를 개발하였다. 더욱 구체적으로, 본 발명의 주제인 polX 패밀리의 DNA 폴리머라제 변형체는 상기 효소의 촉매 공동 (catalytic cavity)의 수준에서 직접 개입하거나, 또는 뉴클레오타이드 수준에서 존재하는 변형에 기인한 입체 장애를 수용하도록 이 공동의 윤곽 기형 (deformation)을 허용하는, 아미노산에 대한 적어도 하나의 돌연변이를 포함한다.

예를 들어, 도입된 돌연변이는 변형된 뉴클레오티드의 3'-OH 말단이 수용되는 촉매 공동의 확대를 가능케 한다. 대안적으로 또는 추가적으로, 수행된 돌연변이는 촉매 활성 용량의 팽창 또는 증가, 3'-OH 변형된 뉴클레오티드에 의한 촉매 포켓으로의 접근의 증가를 가능케 하고/하거나, 이들은 3'-OH 변형된 뉴클레오티드의 큰 입체 장애를 초래하는 변형을 수용할 수 있도록 효소의 구조에 필요한 유연성을 부여한다. 이러한 돌연변이의 결과로서, 일단 폴리머라제가 신장되는 핵산 단편에 결합되면, 변형된 뉴클레오티드는 촉매 포켓의 코어 내로 침투하여 그의 접근이 확대되고 이는 핵산 가닥의 마지막 뉴클레오티드의 3'-OH 말단과 변형된 뉴클레오티드의 5'-트리포스페이트 사이를 형성하는 포스포디에스테르 결합인, 상기 촉매 포켓 내 최적의 공간적 형태를 취한다.

[0010] 따라서, 본 발명의 주제는 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체, 또는 이러한 폴리머라제의 기능적 단편의 변형체이고, 상기 변형체는 T331, G332, G333, F334, R336, K338, H342, D343, V344, D345, F346, A397, D399, D434, V436, A446, L447, L448, G449, W450, G452, R454, Q455, F456, E457, R458, R461, N474, E491, D501, Y502, I503, P505, R508, N509 및 A510으로 구성된 군으로부터 선택되는 적어도 하나의 위치 내 잔기, 또는 기능적으로 동등한 잔기의 적어도 하나의 돌연변이를 포함하고, 상기 표시된 위치는 서열번호 1과의 정렬에 의해 결정된다.

[0011] 특정 구현예에서, 변형체는 DNA 가닥 또는 RNA 가닥을 합성할 수 있다.

[0012] 구체적으로, 본 발명은 선택된 돌연변이(들)을 포함하는, polX 패밀리의 DNA 폴리머라제의 변형체, 특히 효모로부터의 Pol IV, Pol  $\mu$  또는 야생형 TdT에 관한 것이다. 특정 구현예에서, 본 발명에 따른 변형체는 서열번호: 1 서열의 TdT의 변형체 또는 서열번호: 1의 서열과 적어도 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일성을 갖는 상동 서열이고, 이는 선택된 돌연변이(들)을 보유한다.

[0013] 본 발명은 또한 본 발명에 따른 polX 패밀리의 DNA 폴리머라제의 변형체를 코딩하는 핵산, 본 발명에 따른 핵산을 포함하는 발현 카세트, 및 본 발명에 따른 핵산 또는 발현 카세트를 포함하는 벡터에 관한 것이다. 본 발명의 변형체를 코딩하는 핵산은 본 발명에 따른 DNA 폴리머라제의 성숙 형태 또는 전구체 형태의 핵산일 수 있다.

[0014] 본 발명은 또한 숙주 세포의 형질전환 또는 형질감염을 위한 본 발명에 따른 핵산, 발현 카세트 또는 벡터의 용도에 관한 것이다. 본 발명은 나아가 본 발명에 따른 polX 패밀리의 DNA 폴리머라제를 코딩하는 핵산, 발현 카세트 또는 벡터를 포함하는 숙주 세포에 관한 것이다. 본 발명은 본 발명에 따른 polX 패밀리의 DNA 폴리머라제의 변형체를 생산하기 위한 이러한 핵산, 이러한 발현 카세트, 이러한 벡터 또는 이러한 숙주 세포의 용도에 관한 것이다.

[0015] 본 발명은 또한 본 발명에 따른 polX 패밀리의 DNA 폴리머라제의 변형체를 생산하는 방법에 관한 것으로, 본 발명에 따른 핵산, 발현 카세트 또는 벡터에 의한 숙주 세포의 형질전환 또는 형질감염, 상기 변형체를 코딩하는 핵산의 발현을 가능케 하는 배양 조건에서 형질전환/형질감염된 숙주 세포의 배양, 및 선택적으로 숙주 세포에 의해 생산된 polX 패밀리의 DNA 폴리머라제의 변형체의 수확을 포함한다.

[0016] 숙주 세포는 원핵생물 또는 진핵생물일 수 있다. 구체적으로, 숙주 세포는 미생물, 바람직하게는 세균, 호모 또는 버섯일 수 있다. 일 구현예에서, 숙주 세포는 세균, 바람직하게는 대장균 (*E. coli*)이다. 다른 구현예에서, 숙주 세포는 효모, 바람직하게는 *P. 파스토리스* (*P. pastoris*) 또는 *K. 락티스* (*K. lactis*)이다. 다른 구현예에서, 숙주 세포는 포유류 세포, 바람직하게는 COS7 또는 CHO 세포이다.

[0017] 본 발명은 또한 주형 가닥 없이 3'-OH 변형된 뉴클레오티드로부터 핵산 분자를 합성하기 위한 본 발명에 따른 polX 패밀리의 DNA 폴리머라제의 변형체의 용도에 관한 것이다. 자연적으로, 본 발명에 따른 polX 패밀리의 DNA 폴리머라제의 변형체는 또한 본 발명의 맥락에서 비-변형된 뉴클레오티드로부터 또는 변형 및 비-변형된 뉴클레오티드의 혼합물로부터, 주형 가닥 없이 핵산 분자를 합성하는데 사용될 수 있다.

[0018] 본 발명은 또한 주형 가닥 없이 핵산 분자의 효소적 합성 방법을 제안하고, 이에 따라 프라이머 가닥은 본 발명에 따른 polX 패밀리의 DNA 폴리머라제의 변형체의 존재하에 적어도 하나의 뉴클레오티드, 바람직하게는 3'-OH 변형된 뉴클레오티드와 접촉하게 된다. 상기 방법의 수행은 특히 정제된 변형체, 상기 변형체를 발현하도록 형질전환된 숙주 세포의 배양 배지, 및/또는 이러한 숙주 세포의 세포 추출물을 사용하여 일어날 수 있다.

[0019] 본 발명은 또한 본 발명에 따른 polX 패밀리의 DNA 폴리머라제의 적어도 하나의 변형체, 뉴클레오티드, 바람직하게는 3'-OH 변형된 뉴클레오티드, 및 선택적으로 적어도 하나의 프라이머 가닥, 또는 뉴클레오티드 프라이머, 및/또는 반응 완충제를 포함하는, 주형 가닥 없이 핵산 분자의 효소적 합성을 위한 키트에 관한 것이다.

## 도면의 간단한 설명



[0020]

도 1: 본 발명의 구현 실시예에 따른 TdT 변형체 분획의 SDS-PAGE 겔 (M: 분자량 마커; 1: 로우딩 전 원심분리; 2: 로우딩 후 원심분리; 3: 로우딩 후 세척 완충제; 4: 용출 분획 3 mL; 5: 용출 분획 30 mL; 6: 용출 피크 컴필레이션; 7: 농도);

도 2: 온라인 정렬 소프트웨어 (<http://multalin.toulouse.inra.fr/multalin/multalin.html>)에 의해 획득된 호모 사피엔스 DNA 폴리머라제 Pol  $\mu$  (UniProtKB Q9NP87), 팬 트로글로다이트 (Pan troglodytes) Pol  $\mu$  (UniProtKB H2QUI0), 무스 무스쿨러스 (Mus musculus) Pol  $\mu$  (UniProtKB Q924W4), 카니스 루페스 파밀리아리스 (Canis lupus familiaris) Pol  $\mu$  (UniProtKB F1P657), 무스 무스쿨러스 TdT (UniProtKB Q3UZ80), 갈루스 갈루스 (Gallus gallus) TdT (UniProtKB P36195) 및 호모 사피엔스 TdT (UniProtKB P04053)의 아미노산 서열의 정렬;

도 3: 5' 말단에서 미리 방사능으로 표지된 프라이머 및 3'-O-아미노-2',3'-디데옥시아데노신-5'-트리포스페이트 변형된 뉴클레오티드 (ONH2 겔) 또는 3'-biot-EDA-2',3'-디데옥시아데노신-5'-트리포스페이트 변형된 뉴클레오티드 (Biot-EDA 겔)의 존재하에서, 서열번호: 3 서열의 절두된 (truncated) 야생형 TdT 및 표 1에 제공된 상이한 치환을 포함하는 이 절두된 TdT의 몇몇 변형체의 활성 비교; SDS-PAGE 겔 상 (No: 존재하는 효소 부재; wt: 서열번호: 3 서열의 절두된 야생형 TdT; DSi: 표 1에 정의된 변형체 i);

도 4: SDS-PAGE 겔 상에서 5' 말단에 미리 방사능으로 표지된 프라이머 및 상이한 3'-O-아미노-2',3'-디데옥시아데노신-5'-트리포스페이트 변형된 뉴클레오티드의 존재하에서 본 발명에 따른 변형체 DS124 (표 1 참조)의 활성 연구;

도 5: SDS-PAGE 겔 상에서 5' 말단에 미리 방사능으로 표지된 프라이머 및 상이한 3'-O-아미노-2',3'-디데옥시아데노신-5'-트리포스페이트 변형된 뉴클레오티드의 존재하에서 변형체 DS22, DS24, DS124, DS125, DS126, DS127 및 DS128의 활성 연구;

도 6: 치환 R336N - R454A - E457G (DS125)의 조합을 갖는 본 발명에 따른 TdT의 변형체에 의한 서열 5'-AAAAAAAAAGGGG-3' (서열번호: 14)의 프라이머 후에 서열: 5'-GTACGCTAGT-3' (서열번호: 15)의 DNA 가닥의 합성.

### 발명을 실시하기 위한 구체적인 내용

[0021]

정의

[0022]

아미노산은 하기 명명법에 따라서 1문자 또는 3문자 코드에 의해 본 명세서에서 표시된다: A: Ala (알라닌); R: Arg (아르기닌); N: Asn (아스파라긴); D: Asp (아스파르트산); C: Cys (시스테인); Q: Gln (글루타민); E: Glu (글루탐산); G: Gly (글리신); H: His (히스티딘); I: Ile (이소류신); L: Leu (류신); K: Lys (리신); M: Met (메티오닌); F: Phe (페닐알라닌); P: Pro (프롤린); S: Ser (세린); T: Thr (트레오닌); W: Trp (트립토판); Y: Tyr (티로신); V: Val (발린).

[0023]

본 발명의 맥락에서 2개 핵산 또는 아미노산 서열 사이의 "동일성의 백분율"은, 최선의 정렬 후에 획득되는, 비교되는 두 서열간에서 동일한 뉴클레오티드 또는 아미노산 잔기의 백분율을 지정하는 것으로 이해되고, 이 백분율은 순전히 통계적이며 두 서열간의 차이는 무작위로 이들의 전체 길이에 걸쳐서 분포한다. 최선의 정렬 또는 최적의 정렬은, 하기에서 계산되는 바와 같이, 비교되는 두 서열 사이의 동일성의 백분율이 가장 높은 정렬이다. 2개 핵산 또는 아미노산 서열 사이의 서열 비교는 전통적으로 이들 서열을 최적의 방식으로 정렬시킨 후 이들을 비교함으로써 수행되고, 상기 비교는 서열 유사성의 국부적 영역을 동정하고 비교하기 위해 세그먼트에 의해 또는 비교 윈도우 (comparison window)에 의해 수행된다. 비교를 위한 서열의 최적 정렬은, 수동으로 이외에, Smith 및 Waterman의 국부적 상동성 알고리즘 (1981) (Ad. App. Math. 2: 482)에 의해, Needleman 및 Wunsch의 국부적 상동성 알고리즘 (1970) (J. Mol. Biol. 48: 443)에 의해, Pearson 및 Lipman의 유사성 조사 방법 (1988) (Proc. Natl. Acad. Sci. USA 85: 2444)에 의해, 이러한 알고리즘을 사용하는 컴퓨터 소프트웨어 (GAP, BESTFIT, FASTA 및 TFASTA, Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, WI)에 의해, 온라인 정렬 소프트웨어 Multalin (<http://multalin.toulouse.inra.fr/multalin/multalin.html>; 1988, Nucl. Acids Res., 16 (22), 10881-10890)에 의해 수행될 수 있다. 2개 핵산 또는 아미노산 서열 사이의 동일성의 백분율은, 비교되는 핵산 또는 아미노산 서열의 영역이 이들 두 서열간의 최적 정렬을 위해 기준 서열에 대해 부가 또는 결실을 포함할 수 있는 비교 윈도우에 의해 최적 방식으로 정렬되는 이들 두 서열을 비교함으로써 결정된다. 동일성의 백분율은 뉴클레오티드 또는 아미노산 잔기가 두 서열 사이에 동일한, 동일한 위치의 수를 결정하고, 이 동일한 위치의 수

를 비교 윈도우 내 총 위치의 수로 나누고, 이들 두 서열간의 동일성의 백분율을 얻기 위해 수득된 결과에 100을 곱함으로써 계산된다.

[0024] 본 발명의 주제인 변형체는 특정 잔기에 대한 이들의 돌연변이의 함수로 기술되며, 그 위치는 서열번호 1의 효소 서열과 정렬하거나 또는 이를 참조하여 결정된다. 본 발명의 맥락에서, 기능적으로 동등한 잔기에서 이들 동일한 돌연변이를 보유하는 임의의 변형체가 또한 포함된다. "기능적으로 동등한 잔기 (functionally equivalent residue)"는 서열번호: 1에 상동인 서열을 갖고 동일한 기능적 역할을 갖는 polX 패밀리의 DNA 폴리머라제의 서열 내 잔기를 의미하는 것으로 이해된다. 기능적으로 동등한 잔기는, 예를 들면 온라인 정렬 소프트웨어 Mutalin (<http://multalin.toulouse.inra.fr/multalin/multalin.html>; 1988, Nucl. Acids Res., 16 (22), 10881-10890)에 의해 수행되는 서열 정렬에 의해 동정된다. 정렬 후, 기능적으로 동등한 잔기는 고려되는 상이한 서열상의 상동인 위치에 존재한다. 서열의 정렬 및 기능적으로 동등한 잔기의 동정은 중간 변형체를 포함하는, polX 패밀리의 임의의 DNA 폴리머라제와 이들의 천연 변형체 사이에서 일어날 수 있다. 예를 들어, 인간 TdT (UniProtKB P04053)의 잔기 L40은 닭 TdT (UniProtKB P36195)의 잔기 M40 및 팬 트로글로다이크 Pol μ (UniProtKB H2QUI0)의 잔기 V40과 기능적으로 동등하고, 상기 잔기는 서열 정렬 후 고려된다 (도 2).

[0025] "기능적 단편"은 DNA 폴리머라제 활성을 나타내는 polX 패밀리의 DNA 폴리머라제의 단편을 의미하는 것으로 이해된다. 단편은 polX 패밀리의 DNA 폴리머라제의 100, 200, 300, 310, 320, 330, 340, 350, 360, 370, 380개 이상의 연속적인 아미노산을 포함한다. 바람직하게는, 단편은 상기 효소의 촉매 단편으로 구성된 polX 패밀리의 DNA 폴리머라제의 380개 연속적인 아미노산을 포함한다.

[0026] 용어 "변이체(mutant)" 및 "변형체(variant)"는 polX 패밀리의 DNA 폴리머라제로부터 유래한 폴리펩티드, 또는 이러한 DNA 폴리머라제의 기능적 단편의 유도체를 지칭하기 위해 호환적으로 사용되고, 이는 특히 서열번호: 1의 서열에 따른 뮤린 TdT와 같은 TdT로부터 유래하고, 하나 이상의 위치에서 변경, 즉 치환, 삽입 및/또는 결실을 포함하고 DNA 폴리머라제 활성을 갖는다. 변형체는 당해 분야에 익히 공지된 다양한 기법에 의해 수득될 수 있다. 구체적으로, 야생형 단백질을 코딩하는 DNA 서열을 변형하기 위한 기법의 예시에는 지정 돌연변이화, 무작위 돌연변이화, 및 합성 올리고뉴클레오티드의 제작을 포함하지만, 이들로 제한되는 것은 아니다.

[0027] 위치 또는 아미노산 잔기와 관련하여 본원에 사용된 바와 같이 용어 "변형" 또는 "돌연변이"는 고려되는 위치에 서의 아미노산이 기준 야생형 단백질의 아미노산에 대하여 변형되어 있음을 의미한다. 이러한 변형은 하나 이상의 위치, 특히 1, 2, 3, 4, 5개 이상의 위치에서 하나 이상의 아미노산, 특히 1 내지 5개, 1 내지 4개, 1 내지 3개, 1 내지 2개 아미노산의 치환, 결실 및/또는 삽입을 포함한다.

[0028] 위치 또는 아미노산 잔기와 관련하여 용어 "치환"은 특정 위치의 아미노산이 야생형 또는 모체 DNA 폴리머라제와는 다른 아미노산으로 대체되어 있는 것을 의미한다. 바람직하게는, 용어 "치환"은 하나의 아미노산 잔기가 20개 표준 천연 아미노산 잔기, 자연적 기원의 희귀 아미노산 잔기 (예를 들어, 히드록시프롤린, 히드록시리신, 알로히드록시리신, 6-N-메틸리신, N-에틸글리신, N-메틸글리신, N-에틸아스파라긴, 알로-이소류신, N-메틸이소류신, N-메틸발린, 피로글루타민, 아미노부티르산, 오르니틴), 및 종종 합성적으로 생산되는, 희귀한 비-천연 아미노산 잔기 (예를 들어, 노르류신, 노르발린 및 사이클로헥실알라닌)로부터 선택되는 다른 아미노산 잔기로 대체되는 것을 지칭한다. 바람직하게는, 용어 "치환"은 하나의 아미노산 잔기가 자연적 기원의 20개 표준 아미노산 잔기 (G, P, A, V, L, I, M, C, F, Y, W, H, K, R, Q, N, E, D, S 및 T)로부터 선택되는 다른 아미노산 잔기로 대체되는 것을 지칭한다. 치환은 보존적 또는 비-보존적 치환일 수 있다. 보존적 치환은 염기성 아미노산 (아르기닌, 리신 및 히스티딘), 산성 아미노산 (글루탐산 및 아스파르트산), 극성 아미노산 (글루타민 및 아스파라긴), 소수성 아미노산 (메티오닌, 류신, 이소류신 및 발린), 방향족 아미노산 (페닐알라닌, 트립토판 및 티로신), 및 소형 아미노산 (글리신, 알라닌, 세린 및 트레오닌) 중에서 동일한 그룹의 아미노산 내에서 일어난다. 본 명세서에서, 하기 용어가 치환을 지정하는데 사용된다: R454F는 서열번호: 1의 위치 454에서의 아미노산 잔기 (아르기닌, R)가 페닐알라닌 (F)으로 대체되는 것을 나타낸다. N474S/T/N/Q는 위치 474에서의 아미노산 (아스파라긴, N)이 세린 (S), 트레오닌 (T), 아스파라긴 (N) 또는 글루타민 (Q)으로 대체될 수 있음을 의미한다. "+"는 치환의 조합을 나타낸다.

[0029] 본 발명은 주형 가닥 없이 핵산 분자, 특히 DNA 또는 RNA 가닥을 합성할 수 있는 polX 패밀리의 (EC 2.7.7.7; Advances in Protein Chemistry, Vol. 71, 401-440)의 DNA 폴리머라제의 변형체에 관한 것이다. polX 패밀리의 DNA 폴리머라제는 구체적으로 DNA 폴리머라제 Pol β (인간의 경우 UniProt P06746; 마우스의 경우 Q8K409), Pol σ, Pol λ (인간의 경우 UniProt Q9UGP5; 마우스의 경우 Q9QUG2 및 Q9QXE2) 및 Pol μ (인간의 경우 UniProt Q9NP87; 마우스의 경우 Q9JIW4), Pol 4 (효모 반테르왈토지마 폴리스포라 (*Vanderwaltozyma polyspora*)의 경우



UniProt A7TER5; 효모 사카로마이세스 세레비시아에 (*Saccharomyces cerevisiae*)의 경우 P25615) 및 말단 데옥시리보뉴클레오티딜 트랜스퍼라제 또는 TdT (EC 2.7.7.31; 인간의 경우 UniProt P04053; 마우스의 경우 P09838)를 포함한다.

- [0030] 더욱 구체적으로, 본 발명은 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체, 또는 이러한 폴리머라제의 기능적 단편의 변형체에 관한 것으로, 상기 변형체는 T331, G332, G333, F334, R336, K338, H342, D343, V344, D345, F346, A397, D399, D434, V436, A446, L447, L448, G449, W450, G452, R454, Q455, F456, E457, R458, R461, N474, E491, D501, Y502, I503, P505, R508, N509 및 A510으로 구성된 군으로부터 선택되는 적어도 하나의 위치 내 잔기, 또는 기능적으로 동등한 잔기의 적어도 하나의 돌연변이를 포함하고, 상기 위치는 서열번호: 1 서열과 정렬하거나 이를 기준으로 결정된다.
- [0031] 일 구현예에서, 변형체는 DNA 가닥 및/또는 RNA 가닥을 합성할 수 있다.
- [0032] "적어도 하나의 돌연변이를 포함한다" 또는 "적어도 하나의 돌연변이를 포함하는"은 변형체가 서열번호: 1 서열의 폴리펩티드에 대해 지시되는 바와 같은 하나 이상의 돌연변이를 갖지만, 다른 변형, 특히 치환, 결실 또는 부가를 가질 수 있는 것을 의미하는 것으로 이해된다.
- [0033] 일반적으로, 상기 위치에서 하나 이상의 잔기의 돌연변이는 촉매 포켓을 확장시킬 수 있고 (예를 들어, 위치 W450, D434, D435, H342, D343, T331, R336, D399, R461, 및/또는 R508의 표적에 의함), 촉매 포켓에 대한 접근성을 증가시킬 수 있고/있거나 (예를 들어, 위치 R458, E455, R454, A397, K338, 및/또는 N509의 표적에 의함), 효소의 구조에 더 큰 유연성을 부여하여 큰 입체 장애를 나타내는 변형된 뉴클레오티드를 수용할 수 있게 한다 (예를 들어, 위치 V436, F346, V344, F334, M330, L448, E491, E457 및/또는 N474의 표적에 의함).
- [0034] 본 발명의 주제인 변형체는 Pol IV, Pol  $\mu$ , Pol  $\beta$ , Pol  $\lambda$  또는 TdT의 변형체, 바람직하게는 Pol IV, Pol  $\mu$ , 또는 TdT의 변형체일 수 있다. 대안적으로, 변형체는, 예를 들어 polX 패밀리의 적어도 2개 DNA 폴리머라제의 상이한 서열의 부분을 조합하는, 키메라 효소의 변형체일 수 있다.
- [0035] 특정 구현예에서, 변형체는 서열번호: 1에 따른 서열과 적어도 60% 동일성, 바람직하게는 서열번호: 1에 따른 서열과 적어도 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 및 100% 미만의 동일성을 갖는다.
- [0036] 본 발명에 따르면, 돌연변이는 하나 이상의 아미노산 잔기의 치환, 결실 또는 부가로 이루어질 수 있다. 결실의 경우에, 고려되는 잔기를 코딩하는 코돈이 정지 코돈으로 대체됨을 나타내는, 주석 X가 사용되고; 이후의 아미노산뿐만 아니라 문제의 잔기 모두가 결실된다. 따라서, 돌연변이 D501X는 효소가 위치 501의 아스파르트산 (D) 이전의 잔기, 즉 위치 500의 류신 (L)에서 끝나고, 그 외의 모든 잔기는 결실되는 것을 의미한다. 반면, 주석 Ø는 고려되는 잔기의 단일 점 결실을 지칭한다. 따라서, 돌연변이 D501Ø는 위치 501의 아스파르트산 (D)이 결실되었음을 의미한다.
- [0037] 바람직하게는, 본 발명에 따른 변형체는 T331, G332, G333, F334, R336, D343, L447, L448, G449, W450, G452, R454, Q455, E457 및 R508로 구성된 군으로부터 선택되는 적어도 하나의 위치의 잔기, 또는 기능적으로 동등한 잔기의 적어도 하나의 돌연변이, 바람직하게는 R336, R454, E457로 구성된 군으로부터 선택되는 적어도 하나의 위치의 잔기, 또는 기능적으로 동등한 잔기의 적어도 하나의 돌연변이를 포함하고, 상기 표시된 위치는 서열번호 1과의 정렬에 의해 결정된다.
- [0038] 특정 구현예에서, 상기 변형체는 R336, R454 및 E457로 구성된 군으로부터 선택되는 적어도 2개 위치의 잔기의 적어도 하나의 돌연변이, 바람직하게는 상기 3개 위치 R336, R454 및 E457의 잔기, 또는 기능적으로 동등한 잔기의 돌연변이를 포함하고, 상기 표시된 위치는 서열번호 1과의 정렬에 의해 결정된다.
- [0039] 특정 구현예에서, 변형체는 나아가 서열  $X_1X_2GGFR_1R_2GKX_3X_4$  (서열번호: 4)의 적어도 반-보존적 영역 내 잔기의 적어도 하나의 돌연변이를 포함하고, 상기에서
- [0040]  $X_1$ 은 M, I, V, L로부터 선택된 잔기를 나타내고,
- [0041]  $X_2$ 는 T, A, M, Q로부터 선택된 잔기를 나타내고,
- [0042]  $X_3$ 은 M, K, E, Q, L, S, P, R, D로부터 선택된 잔기를 나타내고,
- [0043]  $X_4$ 는 T, I, M, F, K, V, Y, E, Q, H, S, R, D로부터 선택된 잔기를 나타낸다.

- [0044] 바람직하게는, 상기 변형체는 서열번호: 4 서열의 반-보존적 영역의 적어도 하나의 위치  $R_1$ ,  $R_2$  및/또는 K 내 잔기의 적어도 하나의 치환을 갖는다.
- [0045] 또 다른 특정 구현예에서, 변형체는 또한 서열  $X_1X_2LGX_3X_4GSR_1X_5X_6ER_2$  (서열번호: 5)의 적어도 하나의 반-보존적 영역 내 잔기의 적어도 하나의 돌연변이를 포함하고, 상기에서
- [0046]  $X_1$ 은 A, C, G, S로부터 선택된 잔기를 나타내고,
- [0047]  $X_2$ 는 L, T, R로부터 선택된 잔기를 나타내고,
- [0048]  $X_3$ 은 W, Y로부터 선택된 잔기를 나타내고,
- [0049]  $X_4$ 는 T, S, I로부터 선택된 잔기를 나타내고,
- [0050]  $X_5$ 는 Q, L, H, F, Y, N, E, D 또는  $\emptyset$ 로부터 선택된 잔기를 나타내고,
- [0051]  $X_6$ 은 F, Y로부터 선택된 잔기를 나타낸다.
- [0052] 바람직하게는, 상기 변형체는 서열번호: 5 서열의 반-보존적 영역의 적어도 하나의 위치 S,  $R_1$  및/또는 E 내 잔기의 적어도 하나의 치환을 갖는다.
- [0053] 또 다른 특정 구현예에서, 변형체는 또한 서열  $LX_1YX_2X_3PX_4X_5RNA$  (서열번호: 6)의 적어도 하나의 반-보존적 영역 내 잔기의 적어도 하나의 돌연변이를 포함하고, 상기에서
- [0054]  $X_1$ 은 D, E, S, P, A, K로부터 선택된 잔기를 나타내고,
- [0055]  $X_2$ 는 I, L, M, V, A, T로부터 선택된 잔기를 나타내고,
- [0056]  $X_3$ 은 E, Q, P, Y, L, K, G, N으로부터 선택된 잔기를 나타내고,
- [0057]  $X_4$ 는 W, S, V, E, R, Q, T, C, K, H로부터 선택된 잔기를 나타내고,
- [0058]  $X_5$ 는 E, Q, D, H, L로부터 선택된 잔기를 나타낸다.
- [0059] 바람직하게는, 상기 변형체는 서열번호: 6 서열의 반-보존적 영역의 위치  $X_1$  내 잔기의 적어도 하나의 결실 및/또는 위치 R 및/또는 N 내 잔기의 적어도 하나의 치환을 갖는다.
- [0060] 특정 구현예에서, 변형체는 R336, K338, H342, A397, S453, R454, E457, N474, D501, Y502, I503, R508 및 N509로 구성된 군으로부터 선택되는 적어도 하나의 위치 내 잔기, 또는 기능적으로 동등한 잔기의 치환, 바람직하게는 R336, A397, R454, E457, N474, D501, Y502 및 I503으로 구성된 군으로부터 선택되는 적어도 하나의 위치 내 잔기, 또는 기능적으로 동등한 잔기의 치환, 더욱 바람직하게는 R336, R454 및 E457로 구성된 군으로부터 선택되는 적어도 하나의 위치 내 잔기, 또는 기능적으로 동등한 잔기의 적어도 하나의 치환을 포함하고, 상기 표시된 위치는 서열번호 1과의 정렬에 의해 결정된다.
- [0061] 본 발명은 바람직하게는 R336K/H/G/N/D, K338A/C/G/S/T/N, H342A/C/G/S/T/N, A397R/H/K/D/E, S453A/C/G/S/T, R454F/Y/W/A, E457G/N/S/T, N474S/T/N/Q, D501A/G/X, Y502A/G/X, I503A/G/X, R508A/C/G/S/T, N509A/C/G/S/T로 구성된 군으로부터의 적어도 하나의 치환을 포함하는 polX 패밀리의 DNA 폴리머라제의 변형체에 관한 것이다. 특정 구현예에서, 변형체는 R336, R454, E457로 구성된 군으로부터 선택되는 적어도 2개 위치의 잔기, 또는 기능적으로 동등한 잔기의 치환, 바람직하게는 상기 3개 위치의 잔기, 또는 기능적으로 동등한 잔기의 치환을 포함하고, 상기 표시된 위치는 서열번호 1과의 정렬에 의해 결정된다. 구체적으로, 치환은 R336K/H/G/N/D, R454F/Y/W/A 및 E457N/D/G/S/T로 구성된 군으로부터, 바람직하게는 R336N/G, R454A 및 E457G/N/S/T로 구성된 군으로부터 선택된다.
- [0062] 일 구현예에서, 변형체는 E457G/N/S/T에 따른 적어도 하나의 치환을 포함한다.
- [0063] 유리하게, 변형체는 상기에 언급된 그룹으로부터 선택되는 치환의 조합을 포함한다. 이 조합은 이 그룹으로부터 선택된 2, 3, 4, 5, 6, 7, 8, 9, 10 또는 11개 치환으로 이루어질 수 있다.

[0064] 본 발명은 더욱 구체적으로 주형 가닥 없이 DNA 또는 RNA 가닥과 같은 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체, 또는 이러한 폴리머라제의 기능적 단편에 관한 것이고, 상기 변형체는 표 1에 기술된 돌연변이의 적어도 하나의 조합을 포함하고, 상기 표시된 위치는 서열번호 1과의 정렬에 의해 결정된다.

[0065] 일 구현예에서, polX 패밀리의 DNA 폴리머라제의 변형체는 R336G - E457N; R336N - E457N; R336N - R454A - E457N; R336N - E454A - E457G; R336N - E457G; 및 R336G - R454A - E457N으로부터의 치환의 조합을 포함한다.

[0066] [표 1] polX 패밀리의 DNA 폴리머라제의 변형체의 돌연변이 조합의 예시

	돌연변이의 조합
DS1	R454F - E457N - A397D
DS2	R454F - E457N
DS3	R454Y - E457N - A397D
DS4	R454Y - E457N
DS5	R454W - E457N - A397D
DS6	R454W - E457N
DS7	R335A - E457N - A397D
DS8	R335A - E457N
DS9	R335G - E457N - A397D
DS10	R335G - E457N
DS11	R335N - E457N - A397D
DS12	R335N - E457N
DS13	R335D - E457N - A397D
DS14	R335D - E457N
DS15	R336K - E457N - A397D
DS16	R336K - E457N
DS17	R336H - E457N - A397D
DS18	R336H - E457N
DS21	R336G - E457N - A397D
DS22	R336G - E457N
DS23	R336N - E457N - A397D
DS24	R336N - E457N
DS25	R336D - E457N - A397D
DS26	R336D - E457N
DS27	R454A - E457N
DS28	R454A - E457A
DS29	R454A - E457G
DS30	R454A - E457D

[0067]

DS31	E457N
DS32	E457D
DS33	R454A - E457N - A397D
DS34	R454A - E457N - A397K
DS35	R454A - E457N - N474S
DS36	R454A - E457D - A397D
DS37	D501X
DS38	D501X - E457N
DS39	D501X - E457N - A397D
DS40	R454F - E457S - A397D
DS41	R454F - E457S
DS42	R454Y - E457S - A397D
DS43	R454Y - E457S
DS44	R454W - E457S - A397D
DS45	R454W - E457S
DS46	R335A - E457S - A397D
DS47	R335A - E457S
DS48	R335G - E457S - A397D
DS49	R335G - E457S
DS50	R335N - E457S - A397D
DS51	R335N - E457S
DS52	R335D - E457S - A397D
DS53	R335D - E457S
DS54	R336K - E457S - A397D
DS55	R336K - E457S
DS56	R336H - E457S - A397D
DS57	R336H - E457S
DS60	R336G - E457S - A397D
DS61	R336G - E457S
DS62	R336N - E457S - A397D
DS63	R336N - E457S
DS64	R336D - E457S - A397D

[0068]

DS65	R336D - E457S
DS66	R454A - E457S
DS70	E457S
DS72	R454A - E457S - A397D
DS73	R454A - E457S - A397K
DS74	R454A - E457S - N474S
DS75	D501X - E457S
DS76	D501X - E457S - A397D
DS77	R454F - E457T - A397D
DS78	R454F - E457T
DS79	R454Y - E457T - A397D
DS80	R454Y - E457T
DS81	R454W - E457T - A397D
DS82	R454W - E457T
DS83	R335A - E457T - A397D
DS84	R335A - E457T
DS85	R335G - E457T - A397D
DS86	R335G - E457T
DS87	R335N - E457T - A397D
DS88	R335N - E457T
DS89	R335D - E457T - A397D
DS90	R335D - E457T
DS91	R336K - E457T - A397D
DS92	R336K - E457T
DS93	R336H - E457T - A397D
DS94	R336H - E457T
DS97	R336G - E457T - A397D
DS98	R336G - E457T
DS99	R336N - E457T - A397D
DS100	R336N - E457T
DS101	R336D - E457T - A397D
DS102	R336D - E457T

[0069]

DS103	R454A - E457T
DS104	E457T
DS105	R454A - E457T - A397D
DS106	R454A - E457T - A397K
DS107	R454A - E457T - N474S
DS108	D501X - E457T
DS109	D501X - E457T - A397D
DS110	D502X
DS111	D502X - E457N
DS112	D502X - E457TN - A397D
DS113	D502X - E457S
DS114	D502X - E457S - A397D
DS115	D502X - E457T
DS116	D502X - E457T - A397D
DS117	D503X
DS118	D503X - E457N
DS119	D503X - E457TN - A397D
DS120	D503X - E457S
DS121	D503X - E457S - A397D
DS122	D503X - E457T
DS123	D503X - E457T - A397D
DS124	R336N - R454A - E457N
DS125	R336N - R454A - E457G
DS126	R336N - E457G
DS127	R336G - R454A - E457N

[0070]

[0071]

특정 구현예에서, 변형체는 polX 패밀리의 DNA 폴리머라제의 키메라 작제물이다. "키메라 작제물 (chimeric construct)"은 고려되는 DNA 폴리머라제 변형체 내 하나 이상의 상동성 서열의 대체로서 polX 패밀리의 일원인 효소의 하나 이상의 미리 결정된 서열의 부가, 특히 융합 또는 컨주게이션에 의해 형성된 키메라 효소를 의미하는 것으로 이해된다.

[0072]

따라서, 본 발명은 상기 위치의 하나 및/또는 다른 곳에서 하나 이상의 점 돌연변이 이외에, 위치 C378 내지 L406 사이의 잔기, 또는 기능적으로 동등한 잔기의 서열번호: 2 서열의 폴리머라제 Pol $\mu$ 의 잔기 H363 내지 C390, 또는 기능적으로 동등한 잔기로의 치환을 포함하는, 서열번호: 1 서열의 TdT의 변형체를 제한한다.

[0073]

대안적으로 또는 부가적으로, 본 발명의 주제인 변형체는 N-말단에서 하나 이상의 연속적인 아미노산 잔기의 결실을 가질 수 있다. 이들 결실은 특히 다른 단백질과의 결합에 관여하고/하거나 세포 국소화에 관여하는 하나 이상의 효소 도메인을 표적할 수 있다. 예를 들어, TdT의 폴리펩티드 서열은 N-말단에 Ku70/80과 같은 다른 단백질과의 상호작용의 BRCT 도메인 및 핵 국소화 도메인 (NLS)을 포함한다.

[0074]

본 발명의 특정 구현예에서, 변형체는 상기에 기술된 하나 이상의 돌연변이 이외에, 야생형 TdT의 N-말단에 상응하는 잔기 1-129의 결실을 갖는, 서열번호: 1 서열의 TdT의 변형체이다.

[0075]

일부 특정 경우에, 돌연변이화 전략은 천연 변형체의 서열, 결합된 단백질과의 서열 비교, 물리적 특성, 이러한 실체를 포함하는 3차원 구조 또는 컴퓨터 시뮬레이션의 연구와 같은 공지된 정보에 의해 유도될 수 있다.

[0076]

본 발명은 본 발명에 따라 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체를 코딩하는 핵산에 관한 것이다. 본 발명은 또한 본 발명에 따른 핵산의 발현 카세트에 관한 것이다. 본 발명은 나아가 본 발명에 따른 핵산 또는 발현 카세트를 포함하는 벡터에 관한 것이다. 상기 벡터는 플라스미드 또는 바이러스 벡터로부터 선택될 수 있다.

[0077]

DNA 폴리머라제 변형체를 코딩하는 핵산은 DNA (cDNA 또는 gDNA), RNA, 이 둘의 혼합물일 수 있다. 이는 단일-가닥 형태 또는 이중가닥 형태 또는 이 두 형태의 혼합물일 수 있다. 이는, 예를 들어, 변형된 결합, 변형된



퓨린 또는 피리미딘 염기, 또는 변형된 당을 포함하는 변형된 뉴클레오타이드를 포함할 수 있다. 이는 화학적 합성, 재조합, 돌연변이화 등을 비롯해 당업자에게 공지된 임의의 방법에 의해 제조될 수 있다.

[0078] 발현 카세트는 본 발명에 따라 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체의 발현에 필요한 모든 요소들, 특히 숙주 세포에서의 전사 및 번역에 필요한 요소를 포함한다. 숙주 세포는 원핵생물 또는 진핵생물일 수 있다. 구체적으로, 발현 카세트는 프로모터 및 종결자, 선택적으로 증폭자 (amplifier)를 포함한다. 프로모터는 원핵생물 또는 진핵생물일 수 있다. 다음은 바람직한 원핵생물 프로모터의 예시이다: LacI, LacZ, pLacT, ptac, pARA, pBAD, 박테리오파지 T3 또는 T7 RNA 폴리머라제 프로모터, 폴리히드린 프로모터, 람다 파지 PR 또는 PL 프로모터. 다음은 바람직한 진핵생물 프로모터의 예시이다: 초기 CMV 프로모터, HSV 티미딘 키나아제 프로모터, 초기 또는 후기 SV40 프로모터, 뮤린 뮤린 메탈로티오네인 (murine murine metallothionein)-L 프로모터, 및 특정 레트로바이러스의 LTR 영역. 일반적으로, 적합한 프로모터의 선택을 위해, 당업자는 Sambrook 등 (1989)에 의한 연구 또는 Fuller 등에 의해 기술된 기법 (1996; Immunology in Current Protocols in Molecular Biology)을 유리하게 인용할 수 있다.

[0079] 본 발명은 본 발명에 따라 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체를 코딩하는 핵산 또는 발현 카세트를 보유하는 벡터에 관한 것이다. 벡터는 발현 벡터가 바람직하고, 다시 말해 이는 숙주 세포에서 변형체의 발현에 필요한 요소를 포함한다. 숙주 세포는 원핵생물, 예를 들어, 대장균 (*E. coli*) 또는 진핵생물일 수 있다. 진핵생물은 효모 (예를 들어, *P. 파스토리스* 또는 *K. 락티스*) 또는 진균 (예를 들어, *아스페르질러스* 속)와 같은 하등 진핵생물 또는 곤충 세포 (예를 들어, Sf9 또는 Sf21), 포유류 세포 또는 식물 세포와 같은 고등 진핵생물일 수 있다. 세포는 포유류 세포, 예를 들어, COS (녹색 원숭이 세포주) (예를 들어, COS 1 (ATCC CRL-1650), COS 7 (ATCC CRL-1651), CHO (US 4,889,803; US 5,047,335, CHO-K1 (ATCC CCL-61)), 뮤린 세포 및 인간 세포일 수 있다. 특정 구현예에서, 세포는 비-인간 및 비-배아성이다. 벡터는 플라스미드, 파지, 파지미드, 코스미드, 바이러스, YAC, BAC, 아그로박테리움 (*Agrobacterium*) pTi 플라스미드 등일 수 있다. 벡터는 바람직하게는 복제 기원, 다클로닝 부위 및 선별 유전자로부터 선택되는 하나 이상의 요소를 포함할 수 있다. 바람직한 구현예에서, 벡터는 플라스미드이다. 다음은 원핵생물 벡터의 비-포괄적인 예시이다: pQE70, pQE60, pQE-9 (Qiagen), pbs, pD10, 파지스크립트 (phagescript), psiX174, p블루스크립 (pbluescript) SK, pbsks, pNH8A, pNH16A, pNH18A, pNH46A (Stratagene); ptrc99a, pKK223-3, pKK233-3, pDR540, pBR322, 및 pRIT5 (Pharmacia), pET (Novagen). 다음은 진핵생물 벡터의 비-포괄적인 예시이다: pWLNCO, pSV2CAT, pPICZ, pcDNA3.1 (+) Hyg (Invitrogen), pOG44, pXT1, pSG (Stratagene); pSVK3, pBPV, pCI-neo (Stratagene), pMSG, pSVL (Pharmacia); 및 pQE-30 (QLAexpress). 바이러스 벡터는 비-포괄적 방식으로 아데노바이러스, AAV, HSV, 렌티바이러스 등일 수 있다. 바람직하게는, 발현 벡터는 플라스미드 또는 바이러스 벡터이다.

[0080] 본 발명에 따른 변형체를 코딩하는 서열은 시그널 펩티드를 포함하거나 포함하지 않을 수 있다. 시그널 펩티드를 포함하지 않는 경우, 메티오닌이 선택적으로 N-말단에 부가될 수 있다. 다른 대안으로, 이중 시그널 펩티드가 도입될 수 있다. 이러한 이중 시그널 펩티드는 대장균과 같은 원핵생물로부터 또는 진핵생물, 특히 포유류 세포, 곤충 세포 또는 효모로부터 유래할 수 있다.

[0081] 본 발명은 세포를 형질전환 또는 형질감염시키기 위한 본 발명에 따른 폴리뉴클레오타이드, 발현 카세트 또는 벡터의 용도에 관한 것이다. 본 발명은 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 폴리머라제 DNA의 변형체를 코딩하는 핵산, 발현 카세트 또는 벡터를 포함하는 숙주 세포, 및 본 발명에 따라 재조합 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체를 생산하는 이의 용도에 관한 것이다. 용어 "숙주 세포"는 이 세포의 배양 또는 성장으로부터 생성되는 딸세포를 포함한다. 특정 구현예에서, 세포는 비-인간 및 비-배아성이다. 본 발명은 또한 본 발명에 따라 재조합 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체를 생산하는 방법에 관한 것으로, 이는 본 발명에 따른 폴리뉴클레오타이드, 발현 카세트 또는 벡터에 의한 세포의 형질전환 또는 형질감염; 형질감염/형질전환된 세포의 배양; 및 세포에 의해 생산된 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체의 수확을 포함한다. 대안적 구현예에서, 본 발명에 따라 재조합 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체를 생산하는 방법은 본 발명에 따른 폴리뉴클레오타이드, 발현 카세트 또는 벡터를 포함하는 세포의 제공; 형질감염/형질전환된 세포의 배양; 및 세포에 의해 생산된 주형 가닥 없이 핵산 분자를 합성할 수 있는 polX 패밀리의 DNA 폴리머라제의 변형체의 수확을 포함한다. 구체적으로, 세포는 변형체를 코딩하는 핵산에 의해 일시적 또는 안정한 방식으로 형질전환/형질감염될 수 있다. 이러한 핵산은 에피솜의 형태 또는 염색체 형태로 세포 내에 함유될 수 있다. 재조합 단백질을 생산하는 방법은 당업자에게

익히 공지되어 있다. 예를 들어, 대장균에서의 생산을 위해 미국특허 제5,004,689호, EP 특허 제446 582호, 문헌 [Wang et al., Sci. Sin. B 24:1076-1084, 1994 and Nature 295, page 503], 및 포유류 세포에서의 생산을 위해 문헌 [JAMES et al., Protein Science (1996), 5: 331-340]에 기술된 구체적인 방법을 인용할 수 있다.

[0082] 본 발명에 따른 DNA 폴리머라제 변형체는 주형 가닥 없이 핵산을 합성하는데 특히 유리하다. 더욱 구체적으로, 본 발명에 따른 변형체는 천연 뉴클레오티드보다 더 큰 입체 장애를 나타내는 변형된 뉴클레오티드에 의한 핵산의 합성에 특히 적합한 확대된 촉매 포켓을 갖는다. 본 발명에 따른 변형체는 특히 출원 W02016/034807에 기재된 것과 같은 변형된 뉴클레오티드를 핵산 가닥 내로 혼입하는 것을 가능케 할 수 있다.

[0083] 상기에 기술된 돌연변이 또는 특정 돌연변이의 조합을 제시하는, 본 발명에 따른 DNA 폴리머라제 변형체, 특히 TdT의 변형체의 혼입 동역학은 야생형 DNA 폴리머라제의 혼입 동역학에 비해 현저히 개선된다. 이러한 변형체는 고-성능 효소적 DNA 합성 방법의 맥락에서 유리하게 사용될 수 있다.

[0084] 따라서, 본 발명은 또한 3'-OH 변형된 뉴클레오티드, 특히 출원 W02016034807에 기술된 것들로부터 주형 가닥 없이 핵산 분자를 합성하기 위한 본 발명에 따른 polX 패밀리의 DNA 폴리머라제의 변형체의 용도에 관한 것이다.

[0085] 본 발명은 또한 프라이머 가닥이 본 발명에 따른 polX 패밀리의 DNA 폴리머라제의 변형체의 존재하에서 적어도 하나의 뉴클레오티드, 바람직하게는 3'-OH 변형된 뉴클레오티드와 접촉하게 되는, 주형 가닥 없이 핵산 분자의 효소적 합성을 위한 방법에 관한 것이다.

[0086] 유리하게, 본 발명에 따른 변형체는 출원 W02015/159023에 기술된 합성 방법을 수행하는데 사용될 수 있다.

[0087] 본 발명은 또한 본 발명에 따른 polX 패밀리의 DNA 폴리머라제의 적어도 하나의 변형체, 뉴클레오티드, 바람직하게는 3'-OH 변형된 뉴클레오티드, 및 선택적으로 적어도 하나의 뉴클레오티드 프라이머를 포함하는, 주형 가닥 없이 핵산 분자의 효소적 합성을 위한 키트에 관한 것이다.

[0088] 본 명세서에 언급된 모든 참고문헌은 본 출원에 참조로서 포함된다. 본 발명의 다른 특징 및 이점은 당연히 예시적이고 비-제한적인 하기 실시예의 이해 시 더욱 명확해질 것이다.

## [0089] 실시예

### [0090] 실시예 1 - 본 발명에 따른 polX 패밀리의 DNA 폴리머라제 변형체의 생성, 생산 및 정제

#### [0091] 생산자 균주의 생성

[0092] 뮤린 TdT의 절두된 유전자를, 그의 구성이 문헌 [Boule et al., 1998, *Mol. Biotechnol.*, **10**, 199-208]에 기술되어 있는, 플라스미드 pET28b로부터 생성하였다. 상응하는 서열번호: 3 서열 (처음 120개 아미노산이 절두된 서열번호: 1에 상응함)을 하기 프라이머를 사용하여 통상의 PCR 증폭 및 분자생물학 기법에 따라 증폭하였다:

[0093] - T7-pro: TAATACGACTCACTATAGGG (서열번호: 7)

[0094] - T7-ter: GCTAGTTATTGCTCAGCGG (서열번호: 8)

[0095] 이를 플라스미드 pET32에 클로닝하여 벡터 pET32-SEQ ID NO. 3을 생성하였다.

[0096] 플라스미드 pET32-SEQ ID NO. 3의 서열을 먼저 분석한 후, 상업적인 대장균 (*E. coli*) 균주 BL21 (DE3) (Novagen)에 형질전환시켰다. 카나마이신/클로람페니콜 페트리 디쉬에서 성장할 수 있는 콜로니를 분리하고 Ec-SEQ ID NO. 3으로 표지하였다.

#### [0097] 변형체의 생성

[0098] 벡터 pET32-SEQ ID NO. 3을 출발 벡터로 사용하였다. 점 돌연변이 (또는 일부 경우에는 이들이 충분히 가깝다면 점 돌연변이들)를 포함하는 프라이머를 Agilent의 온라인 도구: (<http://www.genomics.agilent.com/primerDesignProgram.jsp>)로부터 생성하였다.

[0099] QuickChange II (Agilent) 키트를 사용하여 목적하는 돌연변이(들)을 포함하는 변형체의 플라스미드를 생성하였다. 혈장 pET32-DSi를 얻기 위해 제조사에 의해 제공된 돌연변이화 프로토콜을 철저히 따랐다 (i는 표 1에 제공된 당해 변형체의 수입). 프로토콜의 마지막에, 플라스미드 pET32-DSx의 서열을 먼저 분석한 후, 상업적 대장균 균주 BL21 (DE3) (Novagen)에 형질전환시켰다. 카나마이신/클로람페니콜 페트리 디쉬에서 성장할 수 있는 콜로니를 분리하고 Ec-DSx로 표지하였다.

[0100] 생산

[0101] 세포 Ec-SEQ ID NO. 3 및 Ec-DSx를 적당량의 카나마이신 및 클로람페니콜이 첨가된 50 mL의 LB 배지를 함유하는 250 mL 엘른메이어 (Erlenmeyer) 플라스크에서 전-배양하였다. 배양액을 밤새 교반 하에 37℃에서 인큐베이션 하였다. 그 후에 전-배양액을 사용하여 적당량의 카나마이신 및 클로람페니콜이 첨가된 2 L의 LB 배지를 함유하는 5 L 엘른메이어 플라스크에서 인큐베이션하였다. 출발 광학 밀도 (OD)는 0.01이었다. 배양액을 교반 하에 37℃에서 인큐베이션하였다. OD 값이 0.6 및 0.9 사이에 도달할 때까지 OD를 주기적으로 측정하였다. 이 값에 도달하면, 1 mL의 이소프로필 β-D-1-티오갈락토피라노시드 1 M을 배양 배지에 첨가하였다. 배양액을 다음 날까지 37℃에서 다시 인큐베이션하였다. 그 후에 세포를 5,000 rpm을 초과하지 않으면서 원심분리에 의해 수확하였다. 수득된 상이한 펠렛을 용해 완충제 (20 mM Tris-HCl, pH 8.3, 0.5 M NaCl)로 세척하는 동안 단일 펠렛을 형성하도록 수집하였다. 세포 펠렛을 -20℃에서 냉동시켰다. 이러한 방식으로 수 개월 동안 이를 저장할 수 있다.

[0102] 추출

[0103] 이전 단계 도중에 냉동된 세포 펠렛을 25 내지 37℃로 가열된 수조에서 해동시켰다. 해동이 완료되면, 세포 펠렛을 대략 100 mL의 용해 완충제에 재현탁하였다. 매우 균질한 용액, 특히 응집체의 완전한 부재를 유도해야만 하는 재현탁에 특별한 주의를 기울였다. 그렇게 재현탁된, 세포를 14,000 psi의 압력에서 프랜치 프레스를 사용하여 용해시켰다. 수집된 용해물을 1시간 내지 1시간 30분간 고속, 10,000 g에서 원심분리하였다. 원심분리물을 0.2 μM 필터를 통해 여과하고 충분한 용량의 튜브에 수집하였다.

[0104] 정제

[0105] TdT를 친화도 칼럼 상에서 정제하였다. 5 mL His-Trap Crude (GE Life Sciences) 칼럼을 연동 펌프 (Peristaltic Pump - MINIIPULS® Evolution, Gilson)와 함께 사용하였다. 첫 번째 단계에서, 칼럼을 2 내지 3 CV (칼럼 용량)의 용해 완충제를 사용해 평형시켰다. 그 후에 이전 단계의 원심분리물을 대략 0.5 내지 5 mL/분의 속도로 상기 칼럼에 로우딩하였다. 모든 원심분리물이 로우딩되면, 칼럼을 3 CV의 용해 완충제로 세척한 후, 3 CV의 세척 완충제 (20 mM Tris-HCl, pH 8.3, 0.5 M NaCl, 60 mM 이미다졸)로 세척하였다. 이 단계의 마지막에, 용출 완충제 (20 mM Tris-HCl, pH 8.3, 0.5 M NaCl, 1 M 이미다졸)를 3 CV의 총 부피에 대해 대략 0.5 내지 1 mL/분으로 칼럼에 주입하였다. 전체 용출 기간 동안에, 칼럼의 유출물을 1 mL 분획으로 수집하였다. 어느 분획이 용출 피크를 함유하는지 여부를 결정하기 위해 이들 분획을 SDS-PAGE로 분석하였다. 분획이 결정되면, 이들을 폴링하여 단일 분획을 형성하고 투석 완충제 (20 mM Tris-HCl, pH 6.8, 200 mM NaCl, 50 mM MgOAc, 100 mM [NH<sub>4</sub>]<sub>2</sub>SO<sub>4</sub>)에 대해 투석하였다. 그 후에 TdT를 5 내지 15 mg/mL의 최종 농도로 농축하였다 (Amicon Ultra-30 원심분리 필터, Merk Millipore). 농축된 TdT를 50% 글리세롤의 첨가 후 장기간 보관을 위해 -20℃에서 냉동시켰다. 전체 정제 기간에 걸쳐서, 상이한 샘플의 분취액을 SDS-PAGE 겔 분석을 위해 수집 (대략 5 μL)하였고, 그 결과를 도 1에 나타내었다.

[0106] 실시예 2 - 본 발명에 따른 변형체의 생성에 사용될 수 있는 polX 패밀리의 상이한 폴리머라제들간의 서열 정렬

[0107] polX 패밀리의 상이한 DNA 폴리머라제를 온라인 정렬 소프트웨어 Mutalin (<http://multalin.toulouse.inra.fr/multalin/multalin.html>, 2016년 4월 4일에 접속함)을 사용하여 정렬시켰다.

**표 2**

[0108] 정렬된 서열

식별자	DNA 폴리머라제	종	길이
Q9NP87	Pol $\mu$ (서열번호: 2)	호모 사피엔스	494
H2QUI0	Pol $\mu$ (서열번호: 9)	팬 트로글로다이크	494
Q924W4	Pol $\mu$ (서열번호: 10)	무스 무스쿨러스	496
F1P657	TdT (서열번호: 11)	카니스 루퍼스 파밀리아리스	509
Q3UZ80	TdT (서열번호: 1)	무스 무스쿨러스	510
P36195	TdT (서열번호: 12)	갈루스 갈루스	506
P04053	TdT (서열번호: 13)	호모 사피엔스	509

[0109] 수득된 정렬을 도 2에 나타내었다.

[0110] **실시예 3 - 비-천연 기질의 존재하에서 변형체의 활성 연구**

[0111] 본 발명에 따른 상이한 변형체의 활성을 하기 시험에 의해 결정하였다. 결과를 각각의 변형체가 유래된 천연 효소로 수득된 결과와 비교하였다.

[0112] **활성 시험**

**표 3**

반응 혼합물

시약	농도	부피
H <sub>2</sub> O	-	15 $\mu$ L
프라이머	500 nM	2.5 $\mu$ L
완충제	10x	2.5 $\mu$ L
변형된 뉴클레오티드	250 $\mu$ M	2.5 $\mu$ L
효소	20 $\mu$ M	2.5 $\mu$ L

[0114] 서열 5'-AAAAAAAAAGGGG-3' (서열번호: 14)의 사용된 프라이머를 효소 PNK (NEB) 및 방사능 ATP (PerkinElmer)의 사용을 포함하는 표준 라벨링 프로토콜에 의해 미리 5'에 방사성으로 표지하였다.

[0115] 250 mM Tris-HCl pH 7.2, 80 mM MgCl<sub>2</sub>, 3.3 mM ZnSO<sub>4</sub>로 구성된 완충제 10x를 사용하였다.

[0116] 사용된 변형된 뉴클레오티드는, 예를 들어 3'-O-아미노-2',3'-디데옥시뉴클레오티드-5'-트리포스페이트 (ONH<sub>2</sub>, Firebird Biosciences) 또는 3'-biot-EDA-2',3'-디데옥시뉴클레오티드-5'-트리포스페이트 (Biot-EDA, Jena Biosciences), 예컨대 3'-O-아미노-2',3'-디데옥시아데노신-5'-트리포스페이트 또는 3'-biot-EDA-2',3'-디데옥시아데노신-5'-트리포스페이트이다. 3'-O-아미노 그룹은 3'-OH 말단에 결합된 더 큰 부피의 그룹이다. 3'-biot-EDA 그룹은 3'-OH 말단에 결합된 매우 큰 부피의 비가요성 (inflexible) 그룹이다.

[0117] 표 1에 열거된 변형체에 의해 생산된 변형체에 의해 제공되는 변형된 뉴클레오티드의 혼입 성능을 효소만 달리 하는 동시 활성 시험을 수행함으로써 천연 TdT (서열번호: 3)와 비교하여 평가하였다.

[0118] 시약을 상기 표 3에 제공된 순서대로 첨가한 다음 37°C에서 90분간 인큐베이션하였다. 그 후에 포름아미드 블루 (포름아미드 100%, 1 내지 5 mg의 브로모페놀 블루; Sigma)의 첨가에 의해 반응을 중단시켰다.

[0119] 겔 및 방사선투과시험 (radiography)

[0120] 이전 활성 시험의 분석을 위해 16% 폴리아크릴아미드 변성 겔 (Biorad)을 사용하였다. 겔을 먼저 붓고 중합을 유도하였다. 그 후에 이를 적당한 크기를 갖고 TBE 완충제 (Sigma)로 채워진 전기영동 탱크에 장착하였다. 상이한 샘플을 전처리 없이 겔 상에 직접 로우딩하였다.

[0121] 이어서 겔을 3 내지 6시간 동안 500 내지 2000 V의 전위차에 적용시켰다. 만족할 만큼 이동하였을 때, 겔을 분리하고 인큐베이션 카세트에 이동시켰다. 형광체 스크린 (Amersham)을 적절한 검출 방식으로 미리 파라미터화되어 있는 Typhoon 기기 (GE Life Sciences)를 사용하여 10 내지 60분간 이미징을 위해 사용하였다.

[0122] 결과

[0123] 사용된 두 효소의 비교 결과를 도 3에 나타내었다.

[0124] 보다 구체적으로, 첫 번째 겔 (ONH<sub>2</sub> 혼입) 상에서, 천연 TdT (wt 칼럼)는 음성 대조군 (칼럼 부재)과의 비교에 의해 나타난 바와 같이 3'-O-아미노-2',3'-디데옥시아데노신-5'-트리포스페이트 변형된 뉴클레오티드를 혼입할 수 없다.

[0125] 다양한 변형체들 중에서, 3개의 상이한 그룹이 관찰될 수 있다:

[0126] 변형체의 제1 그룹 (칼럼 DS7 내지 DS34)은 대략 50% 혼입할 수 있다.

[0127] 변형체의 제2 그룹 (칼럼 DS46 내지 DS73)은 95% 이상, 때로는 98% 이상 혼입할 수 있다.

- [0128] 변형체의 제3 그룹 (칼럼 DS83 내지 DS106)은 60 내지 80% 혼입할 수 있다.
- [0129] 두 번째 겔 (Biot-EDA 혼입) 상에서, 천연 TdT (wt 칼럼)는 또한 대조군 (칼럼 부재)과의 비교에 의해 나타난 바와 같이 3'-biot-EDA-2', 3'-디데옥시아데노신-5'-트리포스페이트 변형된 뉴클레오티드를 혼입할 수 없다.
- [0130] 다양한 변형체들 중에서, 3개의 상이한 그룹이 관찰될 수 있다:
- [0131] 변형체의 제1 그룹 (칼럼 DS7 내지 DS34)은 대략 5 내지 10% 혼입할 수 있다.
- [0132] 변형체의 제2 그룹 (칼럼 DS46 내지 DS73)은 30% 이상, 때로는 40% 이상 혼입할 수 있다.
- [0133] 변형체의 제3 그룹 (칼럼 DS83 내지 DS106)은 10 내지 25% 혼입할 수 있다.
- [0134] 이러한 결과는 야생형 효소와는 달리, 본 발명에 따른 TdT의 변형체가 모두 변형된 뉴클레오티드, 특히 3'-OH 변형된 뉴클레오티드를 기질로 사용할 수 있음을 확인한다. 특히 유리하게, 특정 변형체는 매우 높은 혼입율을 가지며 이는 상기 뉴클레오티드의 입체 장애를 매우 크게 증가시키는 결과를 초래하는 경향이 있는 변형을 보유하는 뉴클레오티드가 존재하는 경우에도 그러하다.
- [0135] **실시예 4 - 본 발명에 따른 변형체의 동역학 연구**
- [0136] 치환 R336N - R454A - E457N (DS124)의 조합을 갖는 변이체를 전술한 실시예 1에 따라 생성하고 생산하였다.
- [0137] 활성 시험
- [0138] 활성 시험에서, 효소를 ONH2 변형된 뉴클레오티드의 존재하에 놓이게 하고 상이한 시간 동안 37°C에서 인큐베이션하였다. DS124의 혼입 동역학을 관찰하고 이를 천연 WT 효소 (서열번호: 3)의 동역학과 비교하기 위해 반응을 중단시켰다.

**표 4**

- [0139] 반응 혼합물

시약	농도	용량
H <sub>2</sub> O	-	15 $\mu$ L
완충제	10x	2.5 $\mu$ L
뉴클레오티드	2.5 $\mu$ M	2.5 $\mu$ L
효소	80 $\mu$ M	2.5 $\mu$ L
프라이머	1 $\mu$ M	2.5 $\mu$ L

- [0140] 사용된 프라이머 및 완충제는 실시예 3에 따른다.
- [0141] 사용된 변형된 뉴클레오티드는 3'-O-아미노-2', 3'-디데옥시뉴클레오티드-5'-트리포스페이트 (ONH2, Firebird Biosciences): 3'-O-아미노-2', 3'-디데옥시구아노신-5'-트리포스페이트, 3'-O-아미노-2', 3'-디데옥시시티딘-5'-트리포스페이트 및 3'-O-아미노-2', 3'-디데옥시티미딘-5'-트리포스페이트이다. 3'-O-아미노 그룹은 3'-OH 말단에 결합된 더 큰 부피의 그룹이다.
- [0142] 효소 DS124에 의한 뉴클레오티드 혼합물의 혼입 성능을 프라이머를 제외한 모든 시약을 포함하는 프리믹스 (표 4의 순서대로 첨가됨)를 제조한 활성 시험을 수행하여 평가하였다. 이들을 상이한 반응 웰에 분주한다. 초기 시점 t = 0에서, 프라이머를 모든 웰에 동시에 첨가한다. 상이한 시점 t = 2분, t = 5분, t = 10분, t = 15분, t = 30분 및 t = 90분에서, 포름아미드 블루 (포름아미드 100%, 1 내지 5 mg의 브로모페놀 블루; Sigma)의 첨가에 의해 반응을 중단시킨다.
- [0143] 겔 및 방사선투과시험
- [0144] 활성 시험의 분석을 실시예 3에 기술된 프로토콜에 따라 폴리아크릴아미드 겔 내 상이한 샘플의 이동에 의해 수행하였다.
- [0145] 결과
- [0146] 두 효소 (DS124 및 WT)의 비교 결과를 도 4에 나타내었다.
- [0147] 보다 구체적으로, 이 겔 상에서, 음성 대조군 (칼럼 부재)은 이것이 신장되지 않았을 때, 즉 뉴클레오티드의 혼



입이 없었을 때 사용된 프라이머의 예상 크기를 제공한다. 천연 TdT (WT 칼럼)는 변형된 뉴클레오타이드 (여기서 ONH<sub>2</sub>-dGTP)를 혼입할 수 없다: 밴드는 칼럼 부재와 동일한 수준에서 관찰될 수 있다.

[0148] 시험된 모든 뉴클레오타이드, 및 90분 (여기서 양성 대조군으로 사용됨)에서, 인큐베이션 시간의 45분의 1 감소에 상응하는, 2분까지 모든 시간의 경우에, 변형체 DS124는 100%의 걸보기 유효성으로 변형된 뉴클레오타이드를 혼입할 수 있다.

[0149] 이러한 결과는 본 발명에 따른 TdT의 변형체가 혼입 효율 및 혼입의 신속성의 측면에서 천연 TdT보다 훨씬 더 높은 혼입 성능을 나타낼 수 있음을 확인한다. 본 발명에 따른 TdT 변형체의 동역학은 본 발명에 의해 기술된 돌연변이 또는 특정 돌연변이의 조합에 의해 현저히 개선된다.

[0150] **실시예 5 - 본 발명에 따른 변형체의 특이성 연구**

[0151] 하기 표 5에 따른 치환 조합을 갖는 변이체를 실시예 1에 따라 생성하고 생산하였다.

**표 5**

사용된 효소적 변형체의 목록

[0152]

#	돌연변이의 조합
DS124	R336N - R454A - E457N
DS24	R336N - E457N
DS125	R336N - R454A - E457G
DS126	R336N - E457G
DS127	R336G - R454A - E457N
DS22	R336G - E457N
DS128	R336A - R454A - E457G
WT	서열번호: 3

[0153] 활성 시험

[0154] 활성 시험에서, 상이한 변형체를 천연 뉴클레오타이드의 혼합물 및 고도로 농축된 변형된 뉴클레오타이드의 존재하에 놓아두었다. 인큐베이션 시간을 단축시키고 정량적 첨가를 달성하기 위해 효소의 농도를 또한 증가시켰다 (비교예 4).

[0155] 생성된 상이한 변형체의 활성을 하기 시험에 의해 결정하였다:

[0156] 각 변형체를 2가지 조건에 따라 시험한다: (1) 뉴클레오타이드의 부재 (H<sub>2</sub>O로 대체됨) 또는 (2) 뉴클레오타이드 혼합물의 존재. 상이한 변형체의 결과를 서로 비교한다. 대조군 샘플을 첨가하였고; 이는 뉴클레오타이드도 효소도 포함하지 않았다 (H<sub>2</sub>O로 대체되었음).

**표 6**

반응 혼합물

[0157]

시약	농도	용량
H <sub>2</sub> O	-	15 µL
프라이머	1 µM	2.5 µL
완충제	10x	2.5 µL
혼합 뉴클레오타이드 (10:90)	2.5 µM	2.5 µL
효소	80 µM	2.5 µL

[0158] 사용된 프라이머 및 완충제는 실시예 3과 동일하다.

[0159] 존재하는 경우, 뉴클레오타이드의 혼합물은, 예를 들어 2'-데옥시구아노신 5'-트리포스페이트 (dGTP)와 같은 천연 2'-데옥시뉴클레오타이드 5'-트리포스페이트 뉴클레오타이드 (Nuc, Sigma-Aldrich) 및 3'-O-아미노-2',3'-디데옥시구아노신-5'-트리포스페이트와 같은 3'-O-아미노-2',3'-디데옥시뉴클레오타이드-5'-트리포스페이트 변형된 뉴클레오타이드 (ONH<sub>2</sub>, Firebird Biosciences)로 구성된다. 더 큰 부피의 3'-O 아미노 그룹이 3'-OH 말단에 결합하였



다. 혼합물은 90%의 ONH2-dGTP 변형된 뉴클레오티드 및 10%의 천연 dGTP 뉴클레오티드로 구성된다.

[0160] 서로 비교하여 표 5에 열거된 변형체에 의한 뉴클레오티드 혼합물의 혼입 성능을 효소만 달리하는 활성 시험을 동시에 수행함으로써 평가하였다.

[0161] 시약을 상기 표 6에 제공된 순서대로 첨가한 후, 37℃에서 15분간 인큐베이션하였다. 그 후에 포름아미드 블루 (포름아미드 100%, 1 내지 5 mg의 브로모페놀 블루; Simga)의 첨가에 의해 반응을 중단시켰다.

[0162] 겔 및 방사선투과시험

[0163] 활성 시험의 분석을 실시예 3에 기술된 프로토콜에 따라 폴리아크릴아미드 겔 내 상이한 샘플의 이동에 의해 수행하였다.

[0164] 결과

[0165] 사용된 효소의 비교 결과를 도 5에 나타내었다.

[0166] 보다 구체적으로, 이 겔 상에서, 음성 대조군 (칼럼 부재)은 이것이 신장되지 않았을 때, 즉 뉴클레오티드의 혼입이 없었을 때 사용된 프라이머의 예상 크기를 제공한다. 하기 샘플을 쌍으로 사용하는데, 각 쌍은 2가지 조건하에서 시험된 동일한 효소적 변형체에 상응한다: 뉴클레오티드의 부재 및 존재 (이들이 존재하는 경우 혼합물의 형태).

[0167] 시험된 상이한 변형체들 중에서, 3개의 상이한 그룹이 관찰될 수 있다:

[0168] 제1 그룹은 변형체 DS128로, 이는 음성 대조군을 구성한다. 이 변형체는 매우 낮은 비율의 뉴클레오티드 혼입을 갖는데: 뉴클레오티드 혼합물이 존재하는 경우 5% 내지 10% 혼입이 관찰되고; 이는 혼합물에 존재하는 천연 뉴클레오티드의 비율에 상응한다.

[0169] 제2 그룹은 변형체 DS127 및 DS22로 구성된다. 이들 변형체는 높은 비율의 뉴클레오티드 혼입을 갖는데: 뉴클레오티드 혼합물이 존재하는 경우 50% 내지 60%의 혼입이 관찰된다. 이 경우에, 2개 뉴클레오티드의 연속적 혼입에 상응하는 추가적인 부가의 밴드가 이들 2개 변형체에 대해서 항상 관찰된다. 이 밴드의 강도는 뉴클레오티드 혼합물에 존재하는 천연 뉴클레오티드의 비율에 상응한다.

[0170] 마지막 그룹은 변형체 DS124, DS24, DS125 및 DS126으로 구성된다. 이들 변형체는 매우 높은 비율의 뉴클레오티드 혼입을 갖는데: 뉴클레오티드 혼합물이 존재하는 경우 DS124 및 DS125에 대해 80% 내지 100%이다. 이 경우에, 추가적인 부가의 밴드는 존재하지 않는다. 변형체 DS24 및 DS126의 경우에는, 비-혼입의 비율이 혼합물 내 존재하는 천연 뉴클레오티드의 비율과 유사하다.

[0171] 이러한 결과는 본 발명에 따른 TdT 변형체가 변형된 뉴클레오티드 및 천연 뉴클레오티드의 혼합물 중에서 변형된 뉴클레오티드를 우선적으로 사용할 수 있음을 확인한다. 특히 유리한 방식으로, 이들 변형체는 변형된 뉴클레오티드의 매우 높은 비율의 혼입을 가지며, 이들을 혼입하지 않는 방식으로 천연 뉴클레오티드를 구별할 수 있고, 따라서 추가적인 부가를 회피함으로써 합성되는 DNA의 품질을 매우 향상시킨다.

[0172] **실시예 6 - 주형 가닥 없이 DNA 가닥의 합성 예시**

[0173] 치환 R336N - R454A - E457G (DS125)의 조합을 갖는 TdT의 변형체를 실시예 1에 따라 생성하고 생산하였다.

[0174] 변형체 DS125를 사용하여 서열 5'-AAAAAAAAAAGGGG-3' (서열번호: 14)의 프라이머 후에 서열 5'-GTACGCTAGT-3' (서열번호: 15)을 합성한다. 상기 프라이머를 효소 PNK (NEB) 및 방사능 ATP (PerkinElmer)의 사용을 포함하는 표준 라벨링 프로토콜에 의해 미리 5'에 방사성으로 표지하였다.

[0175] 프라이머는 상보적 서열: 5'-CCTTTTTTTTTT-3' (서열번호: 16)의 포획 단편과의 상호작용에 의해 고체 지지체에 결합된다. 포획 단편은 표면에 결합된 반응기와 공유적으로 반응하게 할 수 있는 그룹을 그의 3' 말단에 보유한다. 예를 들어, 이 그룹은 NH<sub>2</sub>이고, 반응기는 N-히드록시석시니미드이고, 자기 비드 (Dynabeads, Thermofisher)의 표면일 수 있다. 프라이머와 포획 단편의 상호작용은 표준 DNA 단편 혼성화 조건하에서 수행된다.

[0176] 사용된 변형된 뉴클레오티드는 3'-O-아미노-2',3'-디데옥시뉴클레오티드-5'-트리포스페이트 (ONH2, Firebird Biosciences), 예컨대 3'-O-아미노-2',3'-디데옥시구아노신-5'-트리포스페이트, 3'-O-아미노-2',3'-디데옥시시티딘-5'-트리포스페이트, 3'-O-아미노-2',3'-디데옥시티미딘-5'-트리포스페이트 또는 3'-O-아미노-2',3'-디데옥

시아테노신-5'-트리포스페이트이다. 3'-O-아미노 그룹은 3'-OH 말단에 결합된 더 큰 부피의 그룹이다.

합성

### 표 7

반응 혼합물

시약	농도	용량
H <sub>2</sub> O	-	210 $\mu$ L
완충제	10x	70 $\mu$ L
뉴클레오티드	2.5 $\mu$ M	35 $\mu$ L
효소	80 $\mu$ M	35 $\mu$ L
고체 지지체 상의 프라이머	1 $\mu$ M	-

250 mM Tris-HCl pH 7.2, 80 mM MgCl<sub>2</sub>, 3.3 mM ZnSO<sub>4</sub>로 구성된 완충제 10x를 사용하였다.

사용된 세척 완충제 L은 Tris-HCl 25 mM, pH 7.2로 구성된다.

사용된 탈보호 완충제 D는 10 mM MgCl<sub>2</sub>의 존재하에서 아세트산나트륨 50 mM, pH 5.5로 구성된다.

합성을 개시하기 전에, 그 위에 프라이머가 35 pmol 프라이머의 총 등가량에 대해 혼성화되어 있는 고체 지지체를 구성하는 비드를 완충제 L로 수 차례 세척하였다. 이들 세척 후, 비드를 자석에 고정시키고, 상층액을 완전히 제거하였다.

표 7의 순서대로 첨가된 상이한 시약으로 구성된 몇몇 프리믹스를 제조하였다. 이들 프리믹스 각각은 하기 표 8에 따른 상이한 뉴클레오티드를 함유한다.

### 표 8

프리믹스의 조성

프리믹스 번호	프리믹스의 뉴클레오티드
1	G
2	T
3	A
4	C
5	G
6	C
7	T
8	A
9	G
10	T

미리 세척되고 이들로부터 상층액이 제거된 비드에 프리믹스 1이 첨가되면 합성이 개시된다. 하기 표 9에 따른 합성 단계가 새로운 서열 5'-GTACGCTAGT-3'을 생성하기 위해 서로 뒤따른다.

### 표 9

주형 가닥 없이 DNA 가닥의 합성 방법의 단계

단계	행위	부피	기간
신장 1	첨가 프리믹스 1	350 $\mu$ L	15분
샘플링 1	샘플링	5 $\mu$ L	< 1분
1차 세척 1	첨가 완충제 L	350 $\mu$ L	5분
1차 탈보호 1	추가 완충제 D	350 $\mu$ L	15분
2차 탈보호 1	추가 완충제 D	350 $\mu$ L	15분
2차 세척 1	첨가 완충제 L	350 $\mu$ L	5분
신장 2	첨가 프리믹스 2	350 $\mu$ L	15분

샘플링 2	샘플링	5 $\mu$ L	< 1분
1차 세척 2	첨가 완충제 L	350 $\mu$ L	5분
1차 탈보호 2	첨가 완충제 D	350 $\mu$ L	15분
2차 탈보호 2	첨가 완충제 D	350 $\mu$ L	15분
2차 세척 2	첨가 완충제 L	350 $\mu$ L	5분
신장 3	첨가 프리믹스 3	350 $\mu$ L	15분
샘플링 3	샘플링	5 $\mu$ L	< 1분
1차 세척 3	첨가 완충제 L	350 $\mu$ L	5분
1차 탈보호 3	첨가 완충제 D	350 $\mu$ L	15분
2차 탈보호 3	첨가 완충제 D	350 $\mu$ L	15분
2차 세척 3	첨가 완충제 L	350 $\mu$ L	5분
신장 4	첨가 프리믹스 4	350 $\mu$ L	15분
샘플링 4	샘플링	5 $\mu$ L	< 1분
1차 세척 4	첨가 완충제 L	350 $\mu$ L	5분
1차 탈보호 4	첨가 완충제 D	350 $\mu$ L	15분
2차 탈보호 4	첨가 완충제 D	350 $\mu$ L	15분
2차 세척 4	첨가 완충제 L	350 $\mu$ L	5분
신장 5	첨가 프리믹스 5	350 $\mu$ L	15분
샘플링 5	샘플링	5 $\mu$ L	< 1분
1차 세척 5	첨가 완충제 L	350 $\mu$ L	5분
1차 탈보호 5	첨가 완충제 D	350 $\mu$ L	15분
2차 탈보호 5	첨가 완충제 D	350 $\mu$ L	15분
2차 세척 5	첨가 완충제 L	350 $\mu$ L	5분
신장 6	첨가 프리믹스 6	350 $\mu$ L	15분
샘플링 6	샘플링	5 $\mu$ L	< 1분
1차 세척 6	첨가 완충제 L	350 $\mu$ L	5분
1차 탈보호 6	첨가 완충제 D	350 $\mu$ L	15분
2차 탈보호 6	첨가 완충제 D	350 $\mu$ L	15분
2차 세척 6	첨가 완충제 L	350 $\mu$ L	5분
신장 7	첨가 프리믹스 7	350 $\mu$ L	15분
샘플링 7	샘플링	5 $\mu$ L	< 1분
1차 세척 7	첨가 완충제 L	350 $\mu$ L	5분
1차 탈보호 7	첨가 완충제 D	350 $\mu$ L	15분
2차 탈보호 7	첨가 완충제 D	350 $\mu$ L	15분
2차 세척 7	첨가 완충제 L	350 $\mu$ L	5분
신장 8	첨가 프리믹스 8	350 $\mu$ L	15분
샘플링 8	샘플링	5 $\mu$ L	< 1분
1차 세척 8	첨가 완충제 L	350 $\mu$ L	5분
1차 탈보호 8	첨가 완충제 D	350 $\mu$ L	15분
2차 탈보호 8	첨가 완충제 D	350 $\mu$ L	15분
2차 세척 8	첨가 완충제 L	350 $\mu$ L	5분
신장 9	첨가 프리믹스 9	350 $\mu$ L	15분
샘플링 9	샘플링	5 $\mu$ L	< 1분
1차 세척 9	첨가 완충제 L	350 $\mu$ L	5분
1차 탈보호 9	첨가 완충제 D	350 $\mu$ L	15분
2차 탈보호 9	첨가 완충제 D	350 $\mu$ L	15분
2차 세척 9	첨가 완충제 L	350 $\mu$ L	5분
신장 10	첨가 프리믹스 10	350 $\mu$ L	15분
샘플링 10	샘플링	5 $\mu$ L	< 1분

[0187] 샘플링 단계를 제외하고, 각 단계 사이에서, 비드를 자석을 이용해 수집하고, 상층액을 완전히 제거한다.

[0188] 반응을 중단시키고 분석을 준비하기 위해 각 샘플에 15  $\mu$ L의 포름아미드 블루 용액 (포름아미드 100%, 1 내지 5 mg의 브로모페놀 블루; Sigma)을 첨가한다.

[0189] 겔 및 방사선투과시험

[0190] 활성 시험의 분석을 실시예 3에 기술된 프로토콜에 따라 폴리아크릴아미드 겔 내 상이한 샘플의 이동에 의해 수

행하였다.

[0191] 결과

[0192] 이 합성의 결과를 도 6에 나타내었다.

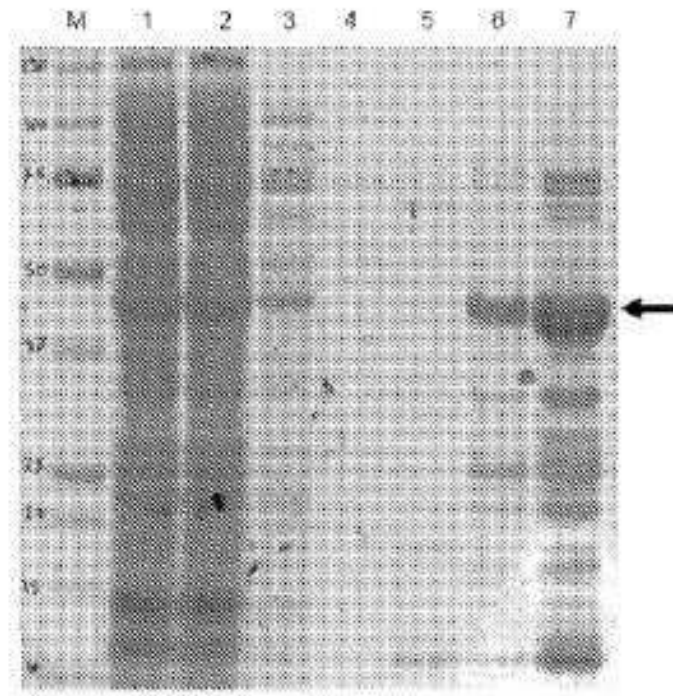
[0193] 칼럼 0 (No, 뉴클레오티드 부재)은 이것이 신장되지 않았을 때, 즉 뉴클레오티드의 혼입이 없었을 때 사용된 프라이머의 예상 크기를 제공한다.

[0194] 칼럼 1 내지 10은 합성 도중의 샘플 1 내지 10에 상응한다. 뉴클레오티드의 각각의 혼입을 최대 성능을 갖는 효소에 의해 수행하였다. 추가적인 정제 단계는 수행하지 않는다.

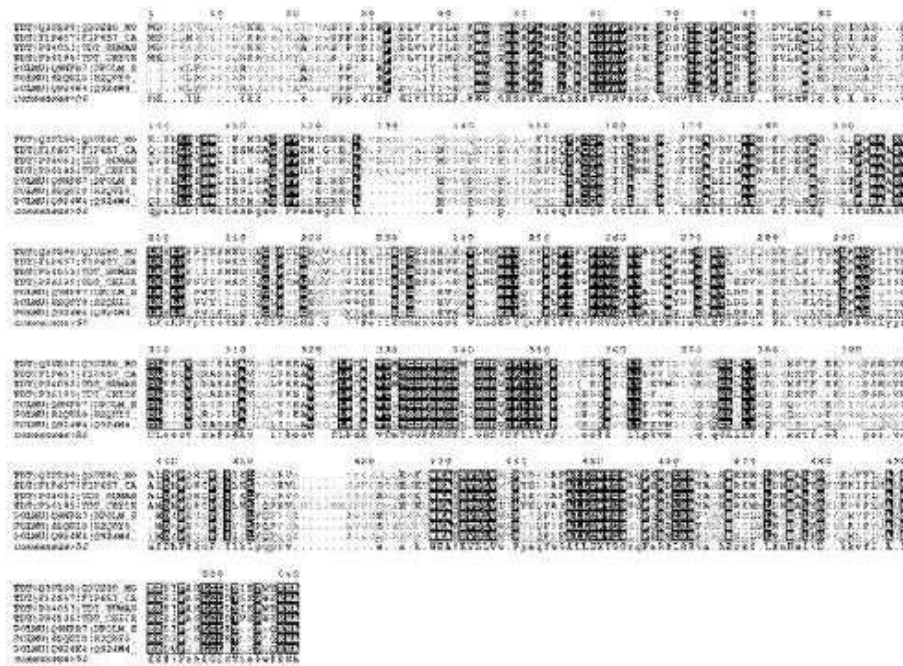
[0195] 유사한 합성 실험을 천연 TdT를 사용하여 수행하였다. 후자는 변형된 뉴클레오티드를 혼입할 수 없기 때문에, 원하는 서열을 합성할 수 없었다.

## 도면

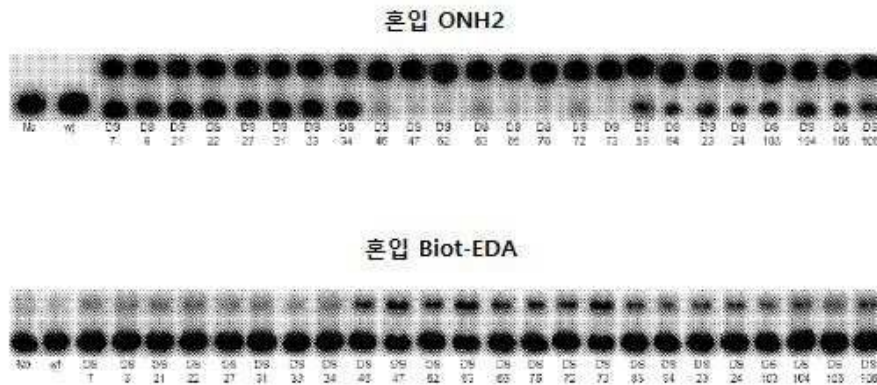
### 도면1



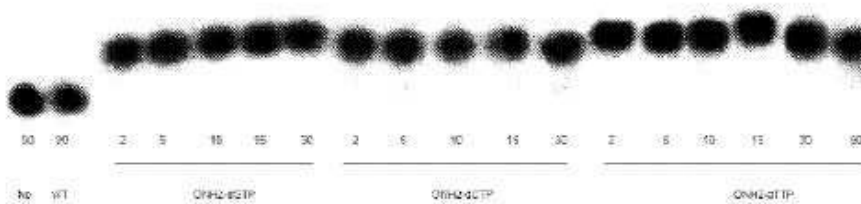
도면2



도면3

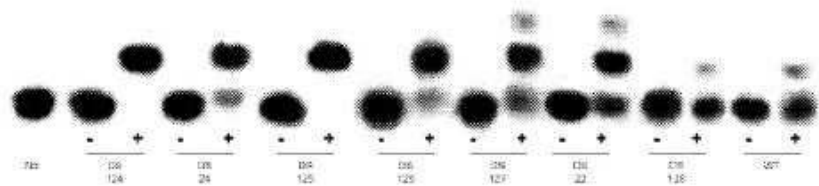


도면4

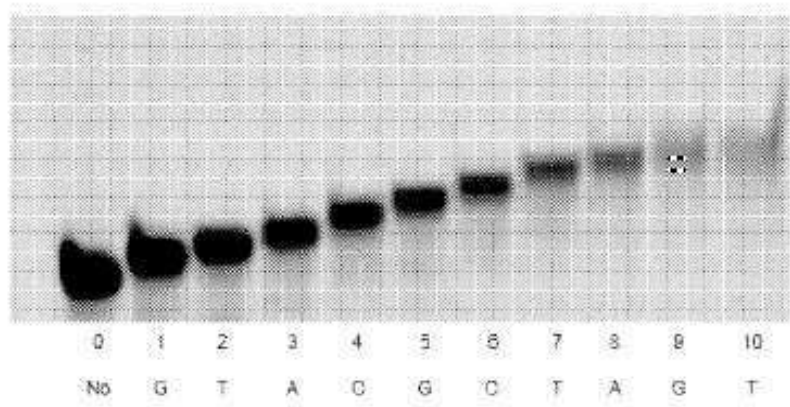




도면5



도면6



서열 목록

SEQUENCE LISTING

<110> DNA SCRIPT

INSTITUT PASTEUR

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE

<120> VARIANTS OF A DNA POLYMERASE OF THE POLX FAMILY

<130> IPA181416-FR

<150> FR 1655475

<151> 2016-06-14

<160> 16

<170> PatentIn version 3.5

<210> 1

<211> 510

<212> PRT

<213> artificial sequence

<220><223> TdT de souris

<400> 1



Met Asp Pro Leu Gln Ala Val His Leu Gly Pro Arg Lys Lys Arg Pro  
1 5 10 15

Arg Gln Leu Gly Thr Pro Val Ala Ser Thr Pro Tyr Asp Ile Arg Phe  
20 25 30

Arg Asp Leu Val Leu Phe Ile Leu Glu Lys Lys Met Gly Thr Thr Arg  
35 40 45

Arg Ala Phe Leu Met Glu Leu Ala Arg Arg Lys Gly Phe Arg Val Glu  
50 55 60

Asn Glu Leu Ser Asp Ser Val Thr His Ile Val Ala Glu Asn Asn Ser  
65 70 75 80

Gly Ser Asp Val Leu Glu Trp Leu Gln Leu Gln Asn Ile Lys Ala Ser  
85 90 95

Ser Glu Leu Glu Leu Leu Asp Ile Ser Trp Leu Ile Glu Cys Met Gly  
100 105 110

Ala Gly Lys Pro Val Glu Met Met Gly Arg His Gln Leu Val Val Asn  
115 120 125

Arg Asn Ser Ser Pro Ser Pro Val Pro Gly Ser Gln Asn Val Pro Ala  
130 135 140

Pro Ala Val Lys Lys Ile Ser Gln Tyr Ala Cys Gln Arg Arg Thr Thr  
145 150 155 160

Leu Asn Asn Tyr Asn Gln Leu Phe Thr Asp Ala Leu Asp Ile Leu Ala  
165 170 175

Glu Asn Asp Glu Leu Arg Glu Asn Glu Gly Ser Cys Leu Ala Phe Met  
180 185 190

Arg Ala Ser Ser Val Leu Lys Ser Leu Pro Phe Pro Ile Thr Ser Met  
195 200 205

Lys Asp Thr Glu Gly Ile Pro Cys Leu Gly Asp Lys Val Lys Ser Ile  
210 215 220

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

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

245 250 255  
 Gly Val Gly Leu Lys Thr Ala Glu Lys Trp Phe Arg Met Gly Phe Arg  
 260 265 270

Thr Leu Ser Lys Ile Gln Ser Asp Lys Ser Leu Arg Phe Thr Gln Met  
 275 280 285

Gln Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Asn  
 290 295 300

Arg Pro Glu Ala Glu Ala Val Ser Met Leu Val Lys Glu Ala Val Val  
 305 310 315 320

Thr Phe Leu Pro Asp Ala Leu Val Thr Met Thr Gly Gly Phe Arg Arg  
 325 330 335

Gly Lys Met Thr Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Glu  
 340 345 350

Ala Thr Glu Asp Glu Glu Gln Gln Leu Leu His Lys Val Thr Asp Phe  
 355 360 365

Trp Lys Gln Gln Gly Leu Leu Leu Tyr Cys Asp Ile Leu Glu Ser Thr  
 370 375 380

Phe Glu Lys Phe Lys Gln Pro Ser Arg Lys Val Asp Ala Leu Asp His  
 385 390 395 400

Phe Gln Lys Cys Phe Leu Ile Leu Lys Leu Asp His Gly Arg Val His  
 405 410 415

Ser Glu Lys Ser Gly Gln Gln Glu Gly Lys Gly Trp Lys Ala Ile Arg  
 420 425 430

Val Asp Leu Val Met Cys Pro Tyr Asp Arg Arg Ala Phe Ala Leu Leu  
 435 440 445

Gly Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr Ala  
 450 455 460

Thr His Glu Arg Lys Met Met Leu Asp Asn His Ala Leu Tyr Asp Arg  
 465 470 475 480

Thr Lys Arg Val Phe Leu Glu Ala Glu Ser Glu Glu Glu Ile Phe Ala  
 485 490 495

His Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn Ala

500 505 510

<210> 2

<211> 494

<212> PRT

<213> artificial sequence

<220><223> Pol ?humaine

<400> 2

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

1 5 10 15

Ala Ala Ser Ser Thr Pro Pro Ser Thr Arg Phe Pro Gly Val Ala Ile

20 25 30

Tyr Leu Val Glu Pro Arg Met Gly Arg Ser Arg Arg Ala Phe Leu Thr

35 40 45

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

50 55 60

Glu Ala Thr His Val Val Met Glu Glu Thr Ser Ala Glu Glu Ala Val

65 70 75 80

Ser Trp Gln Glu Arg Arg Met Ala Ala Ala Pro Pro Gly Cys Thr Pro

85 90 95

Pro Ala Leu Leu Asp Ile Ser Trp Leu Thr Glu Ser Leu Gly Ala Gly

100 105 110

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

115 120 125

Arg Lys Gly Pro Leu Ser Pro Ala Trp Met Pro Ala Tyr Ala Cys Gln

130 135 140

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

145 150 155 160

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

165 170 175

Leu Thr Phe Cys Arg Ala Ala Ser Val Leu Lys Ala Leu Pro Ser Pro

180 185 190

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

195 200 205

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

210 215 220

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

225 230 235 240

Thr Gln Ile Phe Gly Val Gly Val Lys Thr Ala Asp Arg Trp Tyr Arg

245 250 255

Glu Gly Leu Arg Thr Leu Asp Asp Leu Arg Glu Gln Pro Gln Lys Leu

260 265 270

Thr Gln Gln Gln Lys Ala Gly Leu Gln His His Gln Asp Leu Ser Thr

275 280 285

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

290 295 300

Ala Val Gly Gln Ala Leu Pro Gly Ala Thr Val Thr Leu Thr Gly Gly

305 310 315 320

Phe Arg Arg Gly Lys Leu Gln Gly His Asp Val Asp Phe Leu Ile Thr

325 330 335

His Pro Lys Glu Gly Gln Glu Ala Gly Leu Leu Pro Arg Val Met Cys

340 345 350

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

355 360 365

Cys Cys Glu Ser Pro Thr Arg Leu Ala Gln Gln Ser His Met Asp Ala

370 375 380

Phe Glu Arg Ser Phe Cys Ile Phe Arg Leu Pro Gln Pro Pro Gly Ala

385 390 395 400

Ala Val Gly Gly Ser Thr Arg Pro Cys Pro Ser Trp Lys Ala Val Arg

405 410 415

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

420 425 430

Gly Trp Thr Gly Ser Lys Leu Phe Gln Arg Glu Leu Arg Arg Phe Ser

435 440 445  
Arg Lys Glu Lys Gly Leu Trp Leu Asn Ser His Gly Leu Phe Asp Pro

450 455 460  
Glu Gln Lys Thr Phe Phe Gln Ala Ala Ser Glu Glu Asp Ile Phe Arg  
465 470 475 480  
His Leu Gly Leu Glu Tyr Leu Pro Pro Glu Gln Arg Asn Ala

485 490

<210> 3

<211> 401

<212> PRT

<213> artificial sequence

<220><223> TdT de souris tronqu?

<400> 3

Thr Met Gly Ser Ser His His His His His His Ser Ser Gly Leu Val

1 5 10 15

Pro Arg Gly Ser His Met Ser Pro Ser Pro Val Pro Gly Ser Gln Asn

20 25 30

Val Pro Ala Pro Ala Val Lys Lys Ile Ser Gln Tyr Ala Cys Gln Arg

35 40 45

Arg Thr Thr Leu Asn Asn Tyr Asn Gln Leu Phe Thr Asp Ala Leu Asp

50 55 60

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

65 70 75 80

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

85 90 95

Thr Ser Met Lys Asp Thr Glu Gly Ile Pro Cys Leu Gly Asp Lys Val

100 105 110

Lys Ser Ile Ile Glu Gly Ile Ile Glu Asp Gly Glu Ser Ser Glu Ala

115 120 125

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

130 135 140

Ser Val Phe Gly Val Gly Leu Lys Thr Ala Glu Lys Trp Phe Arg Met  
 145 150 155 160  
 Gly Phe Arg Thr Leu Ser Lys Ile Gln Ser Asp Lys Ser Leu Arg Phe  
 165 170 175  
 Thr Gln Met Gln Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser  
 180 185 190  
 Cys Val Asn Arg Pro Glu Ala Glu Ala Val Ser Met Leu Val Lys Glu  
 195 200 205  
  
 Ala Val Val Thr Phe Leu Pro Asp Ala Leu Val Thr Met Thr Gly Gly  
 210 215 220  
 Phe Arg Arg Gly Lys Met Thr Gly His Asp Val Asp Phe Leu Ile Thr  
 225 230 235 240  
 Ser Pro Glu Ala Thr Glu Asp Glu Glu Gln Gln Leu Leu His Lys Val  
 245 250 255  
 Thr Asp Phe Trp Lys Gln Gln Gly Leu Leu Leu Tyr Cys Asp Ile Leu  
 260 265 270  
  
 Glu Ser Thr Phe Glu Lys Phe Lys Gln Pro Ser Arg Lys Val Asp Ala  
 275 280 285  
 Leu Asp His Phe Gln Lys Cys Phe Leu Ile Leu Lys Leu Asp His Gly  
 290 295 300  
 Arg Val His Ser Glu Lys Ser Gly Gln Gln Glu Gly Lys Gly Trp Lys  
 305 310 315 320  
 Ala Ile Arg Val Asp Leu Val Met Cys Pro Tyr Asp Arg Arg Ala Phe  
 325 330 335  
  
 Ala Leu Leu Gly Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg  
 340 345 350  
 Arg Tyr Ala Thr His Glu Arg Lys Met Met Leu Asp Asn His Ala Leu  
 355 360 365  
 Tyr Asp Arg Thr Lys Arg Val Phe Leu Glu Ala Glu Ser Glu Glu Glu  
 370 375 380  
 Ile Phe Ala His Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn



385                      390                      395                      400

Ala

<210> 4

<211> 11

<212> PRT

<213> artificial sequence

<220><223> s?uence semi-conserv?

<220><221> MISC\_FEATURE

<222> (1)..(1)

<223> X=M, I, V, L

<220><221> MISC\_FEATURE

<222> (2)..(2)

<223> X=T, A, M, Q

<220><221> MISC\_FEATURE

<222> (10)..(10)

<223> X=M, K, E, Q, L, S, P, R, D

<220><221> MISC\_FEATURE

<222> (11)..(11)

<223> X=T, I, M, F, K, V, Y, E, Q, H, S, R, D

<400> 4

Xaa Xaa Gly Gly Phe Arg Arg Gly Lys Xaa Xaa

1                      5                      10

<210> 5

<211> 13

<212> PRT

<213> artificial sequence

<220><223> r?ion semi-conserv?

<220><221> MISC\_FEATURE

<222> (1)..(1)

<223> X=A, C, G, S

<220><221> MISC\_FEATURE

<222> (2)..(2)

<223> X=L, T, R

<220><221> MISC\_FEATURE

<222> (5)..(5)

<223> X=W, Y

<220><221> MISC\_FEATURE

<222> (6)..(6)

<223> X=T, S, I

<220><221> MISC\_FEATURE

<222> (10)..(10)

<223> X=Q, L, H, F, Y, N, E, D,

<220><221> MISC\_FEATURE

<222> (11)..(11)

<223> X=F, Y

<400> 5

Xaa Xaa Leu Gly Xaa Xaa Gly Ser Arg Xaa Xaa Glu Arg

1 5 10

<210> 6

<211> 11

<212> PRT

<213> artificial sequence

<220><223> r?ion semi-conserv?

<220><221> MISC\_FEATURE

<222> (2)..(2)

<223> X=D, E, S, P, A, K

<220><221> MISC\_FEATURE

<222> (4)..(4)

<223> X=I, L, M, V, A, T

<220><221> MISC\_FEATURE

<222> (5)..(5)

<223> X=E, Q, P, Y, L, K, G, N

<220><221> MISC\_FEATURE

<222> (7)..(7)

<223> X=W, S, V, E, R, Q, T, C, K, H

<220><221> MISC\_FEATURE

<222> (8)..(8)

<223

> X=E, Q, D, H, L

<400> 6

Leu Xaa Tyr Xaa Xaa Pro Xaa Xaa Arg Asn Ala

1 5 10

<210> 7

<211> 20

<212> DNA

<213> artificial sequence

<220><223> amorce

<400> 7

taatacgact cactataggg 20

<210> 8

<211> 19

<212> DNA

<213> artificial sequence

<220><223> amorce

<400> 8

gctagttatt gctcagcgg 19

<210> 9

<211> 494

<212> PRT

<213

> Pan troglodytes

<400> 9

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

1 5 10 15

Ala Ala Ser Ser Thr Pro Pro Ser Thr Arg Phe Pro Gly Val Ala Ile

20 25 30

Tyr Leu Val Glu Pro Arg Met Gly Arg Ser Arg Arg Ala Phe Leu Thr

35 40 45

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

50 55 60

Glu	Ala	Thr	His	Val	Val	Met	Glu	Glu	Thr	Ser	Ala	Glu	Glu	Ala	Val
65				70				75				80			
Ser	Trp	Gln	Glu	Arg	Arg	Met	Ala	Ala	Ala	Pro	Pro	Gly	Cys	Thr	Pro
85				90				95							
Pro	Ala	Leu	Leu	Asp	Ile	Ser	Trp	Leu	Thr	Glu	Ser	Leu	Gly	Ala	Gly
100				105				110							
Gln	Pro	Val	Pro	Val	Glu	Cys	Arg	His	Arg	Leu	Glu	Val	Ala	Gly	Pro
115				120				125							
Arg	Lys	Gly	Pro	Leu	Ser	Pro	Ala	Trp	Met	Pro	Ala	Tyr	Val	Cys	Gln
130				135				140							
Arg	Pro	Thr	Pro	Leu	Thr	His	His	Asn	Thr	Gly	Leu	Ser	Glu	Ala	Leu
145				150				155				160			
Glu	Thr	Leu	Ala	Glu	Ala	Ala	Gly	Phe	Glu	Gly	Ser	Glu	Gly	Arg	Leu
165				170				175							
Leu	Thr	Phe	Cys	Arg	Ala	Ala	Ser	Val	Leu	Lys	Ala	Leu	Pro	Ser	Pro
180				185				190							
Val	Thr	Thr	Leu	Ser	Gln	Leu	Gln	Gly	Leu	Pro	His	Phe	Gly	Glu	His
195				200				205							
Ser	Ser	Arg	Val	Val	Gln	Glu	Leu	Leu	Glu	His	Gly	Val	Cys	Glu	Glu
210				215				220							
Val	Glu	Arg	Val	Gln	Arg	Ser	Glu	Arg	Tyr	Gln	Thr	Met	Lys	Leu	Phe
225				230				235				240			
Thr	Gln	Ile	Phe	Gly	Val	Gly	Val	Lys	Thr	Ala	Asp	Arg	Trp	Tyr	Arg
245				250				255							
Glu	Gly	Leu	Arg	Thr	Leu	Asp	Asp	Leu	Arg	Glu	Gln	Pro	Gln	Lys	Leu
260				265				270							
Thr	Gln	Gln	Gln	Lys	Ala	Gly	Leu	Gln	His	His	Gln	Asp	Leu	Ser	Thr
275				280				285							
Pro	Val	Leu	Arg	Ser	Asp	Val	Asp	Ala	Leu	Gln	Gln	Val	Val	Glu	Glu
290				295				300							
Ala	Val	Gly	Gln	Ala	Leu	Pro	Gly	Ala	Thr	Val	Thr	Leu	Thr	Gly	Gly

305                      310                      315                      320

Phe Arg Arg Gly Lys Leu Gln Gly His Asp Val Asp Phe Leu Ile Thr

325                      330                      335

His Pro Lys Glu Gly Gln Glu Ala Gly Leu Leu Pro Arg Val Met Cys

340                      345                      350

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

355 360 365

Cys Trp Glu Ser Pro Thr Arg Leu Ala Gln Gln Ser His Met Asp Ala

370                      375                      380

Phe Glu Arg Ser Phe Cys Ile Phe Arg Leu Pro Gln Pro Pro Gly Ala

385                      390                      395                      400

Ala Val Gly Gly Ser Thr Arg Pro Cys Pro Ser Trp Lys Ala Val Arg

405                      410                      415

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

420 425 430

Gly Trp Thr Gly Ser Lys Leu Phe Gln Arg Glu Leu Arg Arg Phe Ser

435                      440                      445

Arg Lys Glu Lys Gly Leu Trp Leu Asn Ser His Gly Leu Phe Asp Pro

450                      455                      460

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

465                      470                      475                      480

His Leu Gly Leu Glu Tyr Leu Pro Pro Glu Gln Arg Asn Ala

485 490

<210> 10

<211> 496

<212> PRT

<213> Mus musculus

 $\langle 400 \rangle$  10

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

1                      5                      10                      15

Val Ala Ser Ser Thr Pro Pro Ser Val Val Arg Phe Pro Asp Val Ala

20 25 30  
 Ile Tyr Leu Ala Glu Pro Arg Met Gly Arg Ser Arg Arg Ala Phe Leu  
 35 40 45  
 Thr Arg Leu Ala Arg Ser Lys Gly Phe Arg Val Leu Asp Ala Tyr Ser  
 50 55 60  
 Ser Lys Val Thr His Val Val Met Glu Gly Thr Ser Ala Lys Glu Ala  
  
 65 70 75 80  
 Ile Cys Trp Gln Lys Asn Met Asp Ala Leu Pro Thr Gly Cys Pro Gln  
 85 90 95  
 Pro Ala Leu Leu Asp Ile Ser Trp Phe Thr Glu Ser Met Ala Ala Gly  
 100 105 110  
 Gln Pro Val Arg Glu Glu Gly Arg His His Leu Glu Val Ala Glu Pro  
 115 120 125  
 Arg Lys Glu Pro Pro Val Ser Ala Ser Met Pro Ala Tyr Ala Cys Gln  
  
 130 135 140  
 Arg Pro Ser Pro Leu Thr His His Asn Thr Leu Leu Ser Glu Ala Leu  
 145 150 155 160  
 Glu Thr Leu Ala Glu Ala Ala Gly Phe Glu Ala Asn Glu Gly Arg Leu  
 165 170 175  
 Leu Ser Phe Ser Arg Ala Asp Ser Val Leu Lys Ser Leu Pro Cys Pro  
 180 185 190  
 Val Ala Ser Leu Ser Gln Leu His Gly Leu Pro Tyr Phe Gly Glu His  
  
 195 200 205  
 Ser Thr Arg Val Ile Gln Glu Leu Leu Glu His Gly Thr Cys Glu Glu  
 210 215 220  
 Val Lys Gln Val Arg Cys Ser Glu Arg Tyr Gln Thr Met Lys Leu Phe  
 225 230 235 240  
 Thr Gln Val Phe Gly Val Gly Val Lys Thr Ala Asn Arg Trp Tyr Gln  
 245 250 255  
 Glu Gly Leu Arg Thr Leu Asp Glu Leu Arg Glu Gln Pro Gln Arg Leu  
  
 260 265 270

Thr Gln Gln Gln Lys Ala Gly Leu Gln Tyr Tyr Gln Asp Leu Ser Thr  
 275 280 285  
 Pro Val Arg Arg Ala Asp Ala Glu Ala Leu Gln Gln Leu Ile Glu Ala  
 290 295 300  
 Ala Val Arg Gln Thr Leu Pro Gly Ala Thr Val Thr Leu Thr Gly Gly  
 305 310 315 320  
 Phe Arg Arg Gly Lys Leu Gln Gly His Asp Val Asp Phe Leu Ile Thr  
 325 330 335  
 His Pro Glu Glu Gly Gln Glu Val Gly Leu Leu Pro Lys Val Met Ser  
 340 345 350  
 Cys Leu Gln Ser Gln Gly Leu Val Leu Tyr His Gln Tyr His Arg Ser  
 355 360 365  
 His Leu Ala Asp Ser Ala His Asn Leu Arg Gln Arg Ser Ser Thr Met  
 370 375 380  
 Asp Ala Phe Glu Arg Ser Phe Cys Ile Leu Gly Leu Pro Gln Pro Gln  
 385 390 395 400  
 Gln Ala Ala Leu Ala Gly Ala Leu Pro Pro Cys Pro Thr Trp Lys Ala  
 405 410 415  
 Val Arg Val Asp Leu Val Val Thr Pro Ser Ser Gln Phe Pro Phe Ala  
 420 425 430  
 Leu Leu Gly Trp Thr Gly Ser Gln Phe Phe Glu Arg Glu Leu Arg Arg  
 435 440 445  
 Phe Ser Arg Gln Glu Lys Gly Leu Trp Leu Asn Ser His Gly Leu Phe  
 450 455 460  
 Asp Pro Glu Gln Lys Arg Val Phe His Ala Thr Ser Glu Glu Asp Val  
 465 470 475 480  
 Phe Arg Leu Leu Gly Leu Lys Tyr Leu Pro Pro Glu Gln Arg Asn Ala  
 485 490 495

<210> 11

<211> 509

<212> PRT

<213> Canis lupus



<400> 11

Met Asp Pro Leu Gln Met Ala His Ser Gly Pro Arg Lys Lys Arg Pro  
1 5 10 15

Arg Gln Met Gly Ala Pro Met Val Ser Pro Pro His Asn Ile Lys Phe  
20 25 30

Gln Asp Leu Val Leu Tyr Ile Leu Glu Lys Lys Met Gly Thr Thr Arg  
35 40 45

Arg Ala Phe Leu Met Glu Leu Ala Arg Arg Lys Gly Phe Arg Val Asp  
50 55 60

Asn Glu Phe Ser Asp Ser Ile Thr His Ile Val Ala Glu Asn Asn Ser  
65 70 75 80

Gly Ser Asp Val Leu Glu Trp Leu Gln Val Gln Asn Ile Lys Ala Ser  
85 90 95

Ser Gln Leu Glu Leu Leu Asp Ile Ser Trp Leu Ile Glu Ser Met Gly  
100 105 110

Ala Gly Lys Pro Val Glu Met Thr Gly Lys His Gln Leu Met Arg Arg  
115 120 125

Asp Tyr Thr Ala Ser Pro Asn Pro Glu Leu Gln Lys Thr Leu Pro Val  
130 135 140

Ala Val Lys Lys Ile Ser Gln Tyr Ala Cys Gln Arg Arg Thr Thr Leu  
145 150 155 160

Asn Asn Tyr Asn Asn Val Phe Thr Asp Ala Phe Glu Val Leu Ala Glu  
165 170 175

Asn Tyr Glu Phe Arg Glu Asn Glu Val Phe Ser Leu Thr Phe Met Arg  
180 185 190

Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Thr Ile Ile Ser Met Lys  
195 200 205

Asp Thr Glu Gly Ile Pro Cys Leu Gly Asp Gln Val Lys Cys Ile Ile  
210 215 220

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

Asn Asp Glu Arg Tyr Gln Ser Phe Lys Leu Phe Thr Ser Val Phe Gly  
 245 250 255  
 Val Gly Leu Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Phe Arg Thr  
 260 265 270  
  
 Leu Ser Lys Ile Lys Ser Asp Lys Ser Leu Lys Phe Thr Pro Met Gln  
 275 280 285  
 Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Thr Arg  
 290 295 300  
 Ala Glu Ala Glu Ala Val Gly Val Leu Val Lys Glu Ala Val Gly Ala  
 305 310 315 320  
 Phe Leu Pro Asp Ala Phe Val Thr Met Thr Gly Gly Phe Arg Arg Gly  
 325 330 335  
  
 Lys Lys Met Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Gly Ser  
 340 345 350  
 Thr Asp Glu Asp Glu Glu Gln Leu Leu Pro Lys Val Ile Asn Leu Trp  
 355 360 365  
 Glu Arg Lys Gly Leu Leu Leu Tyr Cys Asp Leu Val Glu Ser Thr Phe  
 370 375 380  
 Glu Lys Leu Lys Leu Pro Ser Arg Lys Val Asp Ala Leu Asp His Phe  
 385 390 395 400  
  
 Gln Lys Cys Phe Leu Ile Leu Lys Leu His His Gln Arg Val Asp Gly  
 405 410 415  
 Gly Lys Cys Ser Gln Gln Glu Gly Lys Thr Trp Lys Ala Ile Arg Val  
 420 425 430  
 Asp Leu Val Met Cys Pro Tyr Glu Arg Arg Ala Phe Ala Leu Leu Gly  
 435 440 445  
 Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr Ala Ser  
 450 455 460  
  
 His Glu Arg Lys Met Ile Leu Asp Asn His Ala Leu Tyr Asp Lys Thr  
 465 470 475 480  
 Lys Lys Ile Phe Leu Lys Ala Glu Ser Glu Glu Glu Ile Phe Ala His

485                      490                      495  
 Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn Ala  
 500                      505  
 <210> 12  
 <211> 506  
 <212> PRT  
 <213> Gallus gallus  
 <400> 12  
 Met Glu Arg Ile Arg Pro Pro Thr Val Val Ser Gln Arg Lys Arg Gln  
  
 1                      5                      10                      15  
 Lys Gly Met Tyr Ser Pro Lys Leu Ser Cys Gly Tyr Glu Ile Lys Phe  
 20                      25                      30  
 Asn Lys Leu Val Ile Phe Ile Met Gln Arg Lys Met Gly Met Thr Arg  
 35                      40                      45  
 Arg Thr Phe Leu Met Glu Leu Ala Arg Ser Lys Gly Phe Arg Val Glu  
 50                      55                      60  
 Ser Glu Leu Ser Asp Ser Val Thr His Ile Val Ala Glu Asn Asn Ser  
  
 65                      70                      75                      80  
 Tyr Pro Glu Val Leu Asp Trp Leu Lys Gly Gln Ala Val Gly Asp Ser  
 85                      90                      95  
 Ser Arg Phe Glu Ile Leu Asp Ile Ser Trp Leu Thr Ala Cys Met Glu  
 100                      105                      110  
 Met Gly Arg Pro Val Asp Leu Glu Lys Lys Tyr His Leu Val Glu Gln  
 115                      120                      125  
 Ala Gly Gln Tyr Pro Thr Leu Lys Thr Pro Glu Ser Glu Val Ser Ser  
  
 130                      135                      140  
 Phe Thr Ala Ser Lys Val Ser Gln Tyr Ser Cys Gln Arg Lys Thr Thr  
 145                      150                      155                      160  
 Leu Asn Asn Cys Asn Lys Lys Phe Thr Asp Ala Phe Glu Ile Met Ala  
 165                      170                      175  
 Glu Asn Tyr Glu Phe Lys Glu Asn Glu Ile Phe Cys Leu Glu Phe Leu  
 180                      185                      190

Arg Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Pro Val Thr Arg Met

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

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

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

385 390 395 400  
Tyr Gln Pro Arg Val Asp Asn Ser Ser Tyr Asn Met Ser Lys Lys Cys  
405 410 415  
Asp Met Ala Glu Val Lys Asp Trp Lys Ala Ile Arg Val Asp Leu Val  
420 425 430  
Ile Thr Pro Phe Glu Gln Tyr Ala Tyr Ala Leu Leu Gly Trp Thr Gly

435 440 445  
Ser Arg Gln Phe Gly Arg Asp Leu Arg Arg Tyr Ala Thr His Glu Arg

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

500 505

<210> 13

<211> 509

<212> PRT

<213> Homo sapiens

<400> 13

Met Asp Pro Pro Arg Ala Ser His Leu Ser Pro Arg Lys Lys Arg Pro

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

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

130	135	140	
Ile Ala Val Gln Lys Ile Ser Gln Tyr Ala Cys Gln Arg Arg Thr Thr			
145	150	155	160
Leu Asn Asn Cys Asn Gln Ile Phe Thr Asp Ala Phe Asp Ile Leu Ala			
165	170	175	
Glu Asn Cys Glu Phe Arg Glu Asn Glu Asp Ser Cys Val Thr Phe Met			
180	185	190	
Arg Ala Ala Ser Val Leu Lys Ser Leu Pro Phe Thr Ile Ile Ser Met			
195	200	205	
Lys Asp Thr Glu Gly Ile Pro Cys Leu Gly Ser Lys Val Lys Gly Ile			
210	215	220	
Ile Glu Glu Ile Ile Glu Asp Gly Glu Ser Ser Glu Val Lys Ala Val			
225	230	235	240
Leu Asn Asp Glu Arg Tyr Gln Ser Phe Lys Leu Phe Thr Ser Val Phe			
245	250	255	
Gly Val Gly Leu Lys Thr Ser Glu Lys Trp Phe Arg Met Gly Phe Arg			
260	265	270	
Thr Leu Ser Lys Val Arg Ser Asp Lys Ser Leu Lys Phe Thr Arg Met			
275	280	285	
Gln Lys Ala Gly Phe Leu Tyr Tyr Glu Asp Leu Val Ser Cys Val Thr			
290	295	300	
Arg Ala Glu Ala Glu Ala Val Ser Val Leu Val Lys Glu Ala Val Trp			
305	310	315	320
Ala Phe Leu Pro Asp Ala Phe Val Thr Met Thr Gly Gly Phe Arg Arg			
325	330	335	
Gly Lys Lys Met Gly His Asp Val Asp Phe Leu Ile Thr Ser Pro Gly			
340	345	350	
Ser Thr Glu Asp Glu Glu Gln Leu Leu Gln Lys Val Met Asn Leu Trp			
355	360	365	
Glu Lys Lys Gly Leu Leu Leu Tyr Tyr Asp Leu Val Glu Ser Thr Phe			
370	375	380	

Glu Lys Leu Arg Leu Pro Ser Arg Lys Val Asp Ala Leu Asp His Phe

385 390 395 400

Gln Lys Cys Phe Leu Ile Phe Lys Leu Pro Arg Gln Arg Val Asp Ser

405 410 415

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

420 425 430

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

435 440 445

Trp Thr Gly Ser Arg Gln Phe Glu Arg Asp Leu Arg Arg Tyr Ala Thr

450 455 460

His Glu Arg Lys Met Ile Leu Asp Asn His Ala Leu Tyr Asp Lys Thr

465 470 475 480

Lys Arg Ile Phe Leu Lys Ala Glu Ser Glu Glu Glu Ile Phe Ala His

485 490 495

Leu Gly Leu Asp Tyr Ile Glu Pro Trp Glu Arg Asn Ala

500 505

<210> 14

<211> 14

<212> DNA

<213> artificial sequence

<220><

223> amorce

<400> 14

aaaaaaaaa gggg

14

<210> 15

<211> 10

<212> DNA

<213> artificial sequence

<220><223> s?uence synth?is?

<400> 15

gtacgctagt

10

<210> 16

<211> 12



<212> DNA

<213> artificial sequence

<220><223> fragment de capture

<400> 16

cctttttttt tt

12