(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2024-0099167
(43) 공개일자 2024년06월28일

(51) 국제특허분류(Int. Cl.)
C12N 15/113 (2010.01) *C12N 15/10* (2017.01)
C12N 15/86 (2006.01) *C12N 15/90* (2006.01)
C12N 9/22 (2006.01)
(52) CPC특허분류
C12N 15/113 (2013.01)
C12N 15/102 (2013.01)
(21) 출원번호 10-2024-7011702
(22) 출원일자(국제) 2022년09월07일
심사청구일자 없음
(85) 번역문제출일자 2024년04월08일
(86) 국제출원번호 PCT/US2022/076064
(87) 국제공개번호 WO 2023/039441
국제공개일자 2023년03월16일
(30) 우선권주장
63/242,003 2021년09월08일 미국(US)

(71) 출원인
플래그쉽 파이어니어링 이노베이션스 브이아이, 엘엘씨
미국 02142 매사추세츠주 캠브리지 55 캠브리지 파크웨이 8 플로어
(72) 발명자
보트머, 앤 헬렌
미국 02142 매사추세츠 캠브리지 캠브리지 파크웨이 55 스위트 800이
바우처, 제프리 이안
미국 02142 매사추세츠 캠브리지 캠브리지 파크웨이 55 스위트 800이
(뒷면에 계속)
(74) 대리인
양영준, 이상남

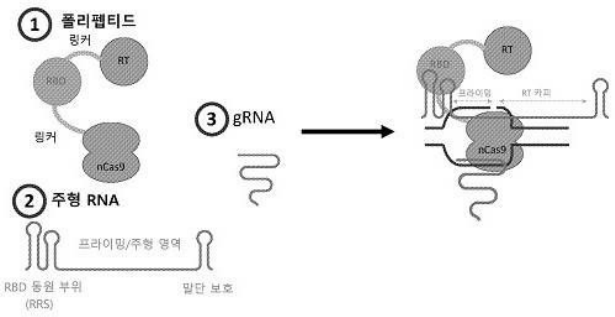
전체 청구항 수 : 총 91 항

(54) 발명의 명칭 **유전자 편집 시스템 구성요소의 트랜스로의 동원**

(57) 요 약

본 개시내용은, 예를 들어 세포, 조직, 또는 대상체에서 DNA 서열의 하나 이상의 위치에서 숙주 세포의 게놈을 표적화, 편집, 변형, 또는 조작하기 위한 조성물, 시스템, 및 방법을 제공한다.

대 표 도



(52) CPC특허분류
C12N 15/86 (2013.01)
C12N 15/907 (2013.01)
C12N 9/22 (2013.01)
C12N 2310/20 (2017.05)
C12N 2740/13022 (2013.01)
C12N 2740/16043 (2013.01)
(72) 발명자
코타-라무시노, 세실리아 지오반나 실비아
미국 02142 매사추세츠 캠브리지 캠브리지 파크웨이 55 스위트 800이
레이, 아나냐
미국 02142 매사추세츠 캠브리지 캠브리지 파크웨이 55 스위트 800이
산체스, 카를로스
미국 02142 매사추세츠 캠브리지 캠브리지 파크웨이 55 스위트 800이
스타인버그, 배럿 에단
미국 02142 매사추세츠 캠브리지 캠브리지 파크웨이 55 스위트 800이

명 세 서

청구범위

청구항 1

주형 RNA로서,

a) 표적 핵산 서열에 돌연변이를 도입하기 위한 돌연변이 영역을 포함하는 이종 목적 서열(여기서, 선택적으로 이종 목적 서열은 5'에서 3'으로 사후 편집 상동성 영역, 돌연변이 영역, 및 사전 편집 상동성 영역을 포함함), 및

b) 표적 핵산 서열의 제1 부분에 결합하되, 제1 부분은 표적 핵산 서열의 제1 가닥에 있고, PBS 서열은 이종 목적 서열의 3'인, 프라이머 결합 부위 서열(PBS 서열), 및

c) RBD 동원 부위(RRS)이되, PBS 서열의 3' 또는 이종 목적 서열의 5'인, RRS

를 포함하는, 주형 RNA.

청구항 2

제1항에 있어서, 말단 블록 서열, 예를 들어 표 41의 말단 블록 서열 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 추가로 포함하는, 주형 RNA.

청구항 3

제2항에 있어서, 상기 말단 블록 서열은 이종 목적 서열의 5'이고 RRS는 PBS 서열의 3'인, 주형 RNA.

청구항 4

제2항에 있어서, 상기 말단 블록 서열은 PBS 서열의 3'이고 RRS는 이종 목적 서열의 5'인, 주형 RNA.

청구항 5

제1항 내지 제4항 중 어느 한 항에 있어서, 상기 RRS는 표 40에 따른 서열 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 갖는, 주형 RNA.

청구항 6

제1항 내지 제5항 중 어느 한 항에 있어서, 복수의 RRS, 예를 들어 2, 3, 4, 5, 또는 10개 RRS의 탠덤 어레이를 포함하는, 주형 RNA.

청구항 7

제1항 내지 제6항 중 어느 한 항에 있어서, 상기 PBS 서열은 8 내지 17개 뉴클레오티드, 예를 들어 표적 핵산 서열에 대하여 100% 동일성을 갖는 8 내지 17개 뉴클레오티드를 포함하는, 주형 RNA.

청구항 8

제1항 내지 제7항 중 어느 한 항에 있어서, 상기 사전 편집 상동성 영역은 최대 20개 뉴클레오티드, 예를 들어, 표적 핵산 서열에 대하여 100% 동일성을 갖는 최대 20개 뉴클레오티드를 포함하는, 주형 RNA.

청구항 9

제1항 내지 제8항 중 어느 한 항에 있어서, 상기 사후 편집 상동성 영역은 5 내지 500개 뉴클레오티드, 예를 들어 표적 핵산 서열에 대하여 100% 동일성을 갖는 5 내지 500개 뉴클레오티드를 포함하는, 주형 RNA.

청구항 10

제1항 내지 제9항 중 어느 한 항에 있어서, 상기 돌연변이 영역은 표적 핵산에 삽입, 결실, 또는 치환을 생성하

도록 구성되는, 주형 RNA.

청구항 11

제1항 내지 제10항 중 어느 한 항에 있어서,

표적 핵산 서열의 상이한 부분(예를 들어, 제3 부분)에 상보적인 gRNA 스페이서(예를 들어, 여기서 상이한 부분(예를 들어, 제3 부분)은 표적 핵산 서열의 제1 가닥 상에 있음); 및

gRNA 스캐폴드

를 추가로 포함하는, 주형 RNA.

청구항 12

제11항에 있어서, 상기 gRNA 스페이서는 이종 목적 서열의 5'인, 주형 RNA.

청구항 13

제11항 또는 제12항에 있어서, 상기 gRNA 스캐폴드는 gRNA 스페이서와 이종 목적 서열 사이에 위치하는, 주형 RNA.

청구항 14

제11항 내지 제13항 중 어느 한 항에 있어서, 상기 gRNA 스페이서 및 PBS 서열은 표적 핵산 서열의 동일한 가닥에 결합하는, 주형 RNA.

청구항 15

제11항 내지 제14항 중 어느 한 항에 있어서, 상기 gRNA 스페이서, 이종 목적 서열, 및 PBS 서열은 표적 핵산 서열의 동일한 가닥에 결합하는, 주형 RNA.

청구항 16

제1항 내지 제4항 중 어느 한 항에 있어서, gRNA 스페이서 또는 gRNA 스캐폴드를 포함하지 않는, 주형 RNA.

청구항 17

제1항 내지 제16항 중 어느 한 항에 있어서, RRS와 PBS 서열 사이에 최대 20개의 뉴클레오티드의 링커를 포함하는, 주형 RNA.

청구항 18

유전자 변형 폴리펩티드로서,

역전사효소(RT) 도메인; 및

표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

DBD 및 RT 도메인에 이종성인 RNA-결합 도메인(RBD)

을 포함하며,

여기서, 도메인은 N-말단에서 C-말단 방향으로

m) DBD, RT 도메인, RBD;

n) RT 도메인, DBD, RBD;

o) RBD, DBD, RT 도메인;

p) RBD, RT 도메인, DBD;

q) DBD, RBD, RT 도메인; 또는

r) RT 도메인, RBD, DBD

로 배열되는, 유전자 변형 폴리펩티드.

청구항 19

유전자 변형 폴리펩티드로서

역전사효소(RT) 도메인; 및

표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

DBD 및 RT 도메인에 이종성인 복수(예를 들어, 2, 3, 4, 또는 5개)의 RNA-결합 도메인(RBD)

을 포함하는, 유전자 변형 폴리펩티드.

청구항 20

제6항에 있어서, 상기 RBD는 표 31에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 유전자 변형 폴리펩티드.

청구항 21

제1항 내지 제20항 중 어느 한 항에 있어서, 상기 복수의 RBD는 서로 동일한 아미노산 서열을 갖는, 유전자 변형 폴리펩티드.

청구항 22

제1항 내지 제21항 중 어느 한 항에 있어서, 상기 복수의 RBD는 서로 상이한 아미노산 서열을 갖는, 유전자 변형 폴리펩티드.

청구항 23

제1항 내지 제22항 중 어느 한 항에 있어서, 상기 DBD는 표 7 또는 8에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 유전자 변형 폴리펩티드.

청구항 24

제1항 내지 제23항 중 어느 한 항에 있어서, 상기 RT 도메인은 레트로바이러스로부터 유래되거나, 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 폴리펩티드 도메인인, 유전자 변형 폴리펩티드.

청구항 25

제1항 내지 제24항 중 어느 한 항에 있어서, 상기 RT 도메인은 표 6에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 유전자 변형 폴리펩티드.

청구항 26

제1항 내지 제25항 중 어느 한 항에 있어서, 상기 유전자 변형 폴리펩티드는 링커를 포함하는, 유전자 변형 폴리펩티드.

청구항 27

제1항 내지 제26항 중 어느 한 항에 있어서, 상기 링커는 표 10에 따른 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는, 유전자 변형 폴리펩티드.

청구항 28

제26항 또는 제27항에 있어서, 상기 링커는 DBD와 RT 도메인 사이, RT 도메인과 RBD 사이, 또는 RBD와 DBD 사이에 배치되는, 유전자 변형 폴리펩티드.

청구항 29

제1항 내지 제28항 중 어느 한 항에 있어서, 상기 유전자 변형 폴리펩티드는 N-말단에서 C-말단 방향으로

m) DBD, 제1 링커, RT 도메인, 제2 링커, RBD;

n) RT 도메인, 제1 링커, DBD, 제2 링커, RBD;

o) RBD, 제1 링커, DBD, 제2 링커, RT 도메인;

p) RBD, 제1 링커, RT 도메인, 제2 링커, DBD;

q) DBD, 제1 링커, RBD, 제2 링커, RT 도메인; 또는

r) RT 도메인, 제1 링커, RBD, 제2 링커, DBD

를 포함하는, 유전자 변형 폴리펩티드.

청구항 30

제1항 내지 제29항 중 어느 한 항에 있어서, 인테인-N 도메인을 포함하는 N-말단 부분 및 인테인-C 도메인을 포함하는 C-말단 부분의 인테인-매개 융합에 의해 생성된, 유전자 변형 폴리펩티드.

청구항 31

폴리펩티드 시스템(예를 들어, 폴리펩티드 복합체)으로서,

a) 역전사효소(RT) 도메인; 및

b) 표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas9 도메인, 예를 들어 Cas9 닉카제 도메인); 및

c) DBD 및 RT 도메인에 이종성인 RNA-결합 도메인(RBD)

을 포함하며,

여기서, (a), (b), 및 (c) 중 적어도 2개(예를 들어, 전부)는 별도의 폴리펩티드, 예를 들어 비공유적으로 복합체를 형성하는 별도의 폴리펩티드에 있는, 폴리펩티드 시스템.

청구항 32

제31항에 있어서, 복합체 형성은 제2의 양립 가능한 이량체화 도메인에 결합하는 제1 이량체화 도메인에 의해 매개되는, 폴리펩티드 시스템.

청구항 33

제32항에 있어서, 복합체 형성은 제4의 양립 가능한 이량체화 도메인에 결합하는 제3 이량체화 도메인에 의해 매개되는, 폴리펩티드 시스템.

청구항 34

제31항 내지 제33항 중 어느 한 항에 있어서,

RBD는 제1 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되고;

DBD는 제1 이량체화 도메인에 결합하는 제2 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되고;

DBD는 제3 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되고;

RT 도메인은 제3 이량체화 도메인에 결합하는 제4 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되는, 폴리펩티드 시스템.

청구항 35

제31항 내지 제34항 중 어느 한 항에 있어서, 상기 제1 및 제2 이량체화 도메인은 화학물질-유도 이량체화 도메인, 광-유도 이량체화 도메인, 항체-펩티드 이량체화 도메인, 또는 코일드 코일(coiled coil) 이량체화 도메인인, 폴리펩티드 시스템.

청구항 36

제31항 내지 제35항 중 어느 한 항에 있어서, 상기 제3 및 제4 이량체화 도메인은 화학물질-유도 이량체화 도메인, 광-유도 이량체화 도메인, 항체-펩티드 이량체화 도메인, 또는 코일드 코일 이량체화 도메인인, 폴리펩티드 시스템.

청구항 37

제31항 내지 제36항 중 어느 한 항에 있어서, 상기 제1 이량체화 도메인 및 제2 이량체화 도메인은 각각 복수의 카피, 예를 들어 2, 3, 4, 5, 10, 15, 20, 또는 30개 카피로 존재하는, 폴리펩티드 시스템.

청구항 38

제31항 내지 제37항 중 어느 한 항에 있어서, 상기 제3 이량체화 도메인 및 제4 이량체화 도메인은 각각 복수의 카피, 예를 들어 2, 3, 4, 5, 10, 15, 20, 또는 30개 카피로 존재하는, 폴리펩티드 시스템.

청구항 39

제31항 내지 제38항 중 어느 한 항에 있어서, 상기 제1 이량체화 도메인 및 제2 이량체화 도메인은 동일한 서열을 갖는(예를 들어, 제1 이량체화 도메인 및 제2 이량체화 도메인은 동종이량체를 형성함), 폴리펩티드 시스템.

청구항 40

제31항 내지 제39항 중 어느 한 항에 있어서, 상기 제3 이량체화 도메인 및 제4 이량체화 도메인은 동일한 서열을 갖는(예를 들어, 제3 이량체화 도메인 및 제4 이량체화 도메인은 동종이량체를 형성함), 폴리펩티드 시스템.

청구항 41

제31항 내지 제38항 중 어느 한 항에 있어서, 상기 제1 이량체화 도메인 및 제2 이량체화 도메인은 상이한 서열을 갖는(예를 들어, 제1 이량체화 도메인 및 제2 이량체화 도메인은 이종이량체를 형성함), 폴리펩티드 시스템.

청구항 42

제31항 내지 제41항 중 어느 한 항에 있어서, 상기 제3 이량체화 도메인 및 제4 이량체화 도메인은 상이한 서열을 갖는(예를 들어, 제3 이량체화 도메인 및 제4 이량체화 도메인은 이종이량체를 형성함), 폴리펩티드 시스템.

청구항 43

제31항 내지 제42항 중 어느 한 항에 있어서, 상기 DBD는 하나 이상의 추가적인 DBD에 작동 가능하게 연결되며, 여기서 선택적으로 추가적인 DBD는 DBD와 동일한 서열을 갖는, 폴리펩티드 시스템.

청구항 44

제31항 내지 제43항 중 어느 한 항에 있어서, 상기 RBD는 표 31에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 폴리펩티드 시스템.

청구항 45

제31항 내지 제44항 중 어느 한 항에 있어서, 상기 복수의 RBD는 서로 동일한 아미노산 서열을 갖는, 폴리펩티드 시스템.

청구항 46

제31항 내지 제45항 중 어느 한 항에 있어서, 상기 복수의 RBD는 서로 상이한 아미노산 서열을 갖는, 폴리펩티드 시스템.

청구항 47

제31항 내지 제46항 중 어느 한 항에 있어서, 상기 DBD는 표 31에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 폴리펩티드 시스템.

청구항 48

제31항 내지 제47항 중 어느 한 항에 있어서, 상기 RT 도메인은 레트로바이러스로부터 유래되거나, 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 폴리펩티드 도메인인, 폴리펩티드 시스템.

청구항 49

제31항 내지 제48항 중 어느 한 항에 있어서, 상기 RT 도메인은 표 6에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 폴리펩티드 시스템.

청구항 50

제31항 내지 제49항 중 어느 한 항에 있어서, 각각의 링커는 독립적으로 표 10에 따른 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는, 폴리펩티드 시스템.

청구항 51

제31항 내지 제50항 중 어느 한 항의 시스템의 폴리펩티드를 인코딩하는 핵산 또는 복수의 핵산.

청구항 52

시스템으로서,

제1항 내지 제17항 중 어느 한 항의 주형 RNA;

제18항 내지 제30항 중 어느 한 항의 유전자 변형 폴리펩티드 또는 제31항 내지 제50항 중 어느 한 항의 폴리펩티드 시스템; 및

표적 핵산 서열의 제2 부분에 결합하는 gRNA 스페이서(여기서, 제2 부분은 표적 핵산 서열의 제2 가닥 중 하나임); 및

유전자 변형 폴리펩티드 또는 폴리펩티드 시스템의 DBD에 결합하는 gRNA 스캐폴드

를 포함하는 제1 gRNA

를 포함하는, 시스템.

청구항 53

제52항에 있어서, 상기 주형 RNA는 gRNA 스페이서 또는 gRNA 스캐폴드를 포함하지 않는, 시스템.

청구항 54

제52항 또는 제53항에 있어서, 상기 gRNA 스페이서는 PBS 서열에 의해 결합된 표적 핵산 서열의 영역의 약 5, 10, 15, 20, 25, 30, 또는 40개 뉴클레오티드 내에 있는 표적 핵산 서열의 영역에 결합하는, 시스템.

청구항 55

제52항 내지 제54항 중 어느 한 항에 있어서,

제2 Cas 단백질(예를 들어, 사멸 Cas 단백질) 및

PBS 서열에 의해 결합된 위치의 3' 위치에서 표적 핵산의 제1 가닥에 결합하는 gRNA 스페이서, 및

제2 Cas 단백질에 결합하는 gRNA 스캐폴드

를 포함하는 제2 gRNA

를 추가로 포함하는, 시스템.

청구항 56

제55항에 있어서, 상기 제2 Cas 단백질은 사멸 Cas 단백질(예를 들어, 사멸 Cas9 단백질) 또는 Cas 닉카제 단백질(예를 들어, Cas9 닉카제 단백질)인, 시스템.

청구항 57

제55항에 있어서, 상기 제2 gRNA의 gRNA 스페이서는 적어도 18개 뉴클레오티드(예를 들어, 18 내지 28개 뉴클레오티드, 예를 들어, 18 내지 21개 뉴클레오티드)의 길이를 갖고, 제2 Cas 단백질은 사멸 Cas 단백질인, 시스템.

청구항 58

제55항에 있어서, 상기 제2 gRNA의 gRNA 스페이서는 17개 뉴클레오티드 이하(예를 들어, 14 내지 17개 뉴클레오티드)의 길이를 갖고, 선택적으로 제2 Cas 단백질은 Cas 닉카제 단백질인, 시스템.

청구항 59

제52항에 있어서, 상기 주형 RNA는

표적 핵산 서열의 제3 부분에 상보적인 gRNA 스페이서(여기서, 제3 부분은 표적 핵산 서열의 제1 가닥 상에 있음); 및

gRNA 스캐폴드

를 추가로 포함하는, 시스템.

청구항 60

제59항에 있어서, 상기 gRNA 스캐폴드는 유전자 변형 폴리펩티드 또는 폴리펩티드 시스템의 DBD에 결합하는, 시스템.

청구항 61

제59항 또는 제60항에 있어서, 상기 gRNA 스페이서는 17개 뉴클레오티드 이하의 길이를 갖는, 시스템.

청구항 62

제52항 내지 제61항 중 어느 한 항에 있어서, 상기 주형 RNA의 gRNA 스페이서는, 예를 들어 표적 핵산 서열의 제2 가닥에서, 주형 핵산의 닉 생성(nicking)을 유도하는, 시스템.

청구항 63

제52항 내지 제61항 중 어느 한 항에 있어서, 상기 주형 RNA의 gRNA 스페이서는 주형 핵산의 닉 생성을 유도하지 않는, 시스템.

청구항 64

시스템으로서,

i) 제1항 내지 제17항 중 어느 한 항의 주형 RNA(예를 들어, 제16항의 주형 RNA);

ii) DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

DBD에 이종성인 RNA-결합 도메인(RBD)(여기서, RBD는 주형 RNA의 RRS에 결합함)

을 포함하는 제1 폴리펩티드;

iii) 제1 폴리펩티드의 DBD를 표적 핵산 서열의 제2 부분으로 지시하는 gRNA 스페이서(여기서, 표적 핵산 서

열의 제2 부분은 핵산 서열의 제2 가닥 상에 있음); 및

제1 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

를 포함하는 제1 gRNA;

iv) RT 도메인, 및

RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인)(여기서, 제2 폴리펩티드의 DBD는 제1 폴리펩티드의 DBD와 상이한 서열을 가짐)

을 포함하는 제2 폴리펩티드; 및

v) 제2 폴리펩티드의 DBD를 표적 핵산 서열의 제3 부분으로 지시하는 gRNA 스페이서(여기서, 제3 부분은 표적 핵산의 제1 가닥 상에 있음), 및

제2 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

를 포함하는 제2 gRNA

를 포함하는, 시스템.

청구항 65

제64항에 있어서, 상기 제2 폴리펩티드의 DBD는 Cas 닉카제 도메인 또는 사멸 Cas 도메인을 포함하는, 시스템.

청구항 66

제64항에 있어서, 상기 제2 RNA의 gRNA 스페이서는, 예를 들어 표적 핵산 서열의 제2 가닥에서, 주형 핵산의 닉 생성을 유도하는, 시스템.

청구항 67

제64항에 있어서, 상기 제2 RNA의 gRNA 스페이서는 주형 핵산의 닉 생성을 유도하지 않는, 시스템.

청구항 68

제64항에 있어서, 상기 제1 gRNA는 제2 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

청구항 69

제64항에 있어서, 상기 제2 gRNA는 제1 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

청구항 70

시스템으로서,

i) 제1항 내지 제69항 중 어느 한 항의 주형 RNA

(여기서, 상기 주형 RNA는

표적 핵산 서열의 제3 부분에 상보적인 gRNA 스페이서(여기서, 제3 부분은 표적 핵산 서열의 제1 가닥 상에 있음); 및

gRNA 스캐폴드

를 포함함);

ii) DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

DBD에 이종성인 RNA-결합 도메인(RBD)(여기서, RBD는 주형 RNA의 RRS에 결합함)

을 포함하는 제1 폴리펩티드;

iii) 제1 폴리펩티드의 DBD를 표적 핵산 서열의 제2 부분으로 지시하는 gRNA 스페이서(여기서, 표적 핵산 서

열의 제2 부분은 핵산 서열의 제2 가닥 상에 있음); 및

제1 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

를 포함하는 제1 gRNA; 및

iv) RT 도메인, 및

RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인)(여기서, 제2 폴리펩티드의 DBD는 제1 폴리펩티드의 DBD와 상이한 서열을 갖고, 주형 RNA의 gRNA 스캐폴드는 제2 폴리펩티드의 DBD에 결합함)

을 포함하는 제2 폴리펩티드

를 포함하는, 시스템.

청구항 71

제70항에 있어서, 상기 제2 폴리펩티드의 DBD는 Cas 닉카제 도메인 또는 사멸 Cas 도메인을 포함하는, 시스템.

청구항 72

제70항에 있어서, 상기 주형 RNA의 gRNA 스페이서는, 예를 들어 표적 핵산 서열의 제2 가닥에서, 주형 핵산의 닉 생성을 유도하는, 시스템.

청구항 73

제70항에 있어서, 상기 주형 RNA의 gRNA 스페이서는 주형 핵산의 닉 생성을 유도하지 않는, 시스템.

청구항 74

제70항 내지 제73항 중 어느 한 항에 있어서, 상기 제1 gRNA는 제2 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

청구항 75

제70항 내지 제74항 중 어느 한 항에 있어서, 상기 주형 RNA의 gRNA는 제1 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

청구항 76

폴리펩티드 시스템으로서,

DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인);

DBD에 이종성인 RNA-결합 도메인(RBD); 및

선택적으로, DBD와 RBD 사이에 배치된 링커

를 포함하는 제1 폴리펩티드; 및

RT 도메인, 및

RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

선택적으로, RT 도메인과 DBD 사이에 배치된 링커

를 포함하는 제2 폴리펩티드

를 포함하는, 폴리펩티드 시스템.

청구항 77

제1항 내지 제76항 중 어느 한 항에 있어서, 상기 표적 핵산 서열은 표적 유전자, 인핸서, 또는 프로모터인, 주

형 RNA 또는 시스템.

청구항 78

제1항 내지 제77항 중 어느 한 항에 있어서, 상기 표적 핵산 서열은 인간 표적 유전자, 인간 인핸서, 또는 인간 프로모터인, 주형 RNA 또는 시스템.

청구항 79

제1항 내지 제78항 중 어느 한 항에 있어서, 상기 RBD는 표 31의 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 갖는, 시스템 또는 폴리펩티드 시스템.

청구항 80

세포(예를 들어, 인간 세포)에서 표적 핵산을 변형시키는 방법으로서, 세포를 제1항 내지 제79항 중 어느 한 항의 시스템, 또는 이를 인코딩하는 핵산과 접촉시키는 단계를 포함하며, 이에 의해 표적 핵산을 변형시키는, 방법.

청구항 81

제80항에 있어서, 상기 제2 폴리펩티드의 존재는 제2 폴리펩티드가 결여된 다른 유사한 시스템과 비교하여

표적 핵산의 풀림 증가;

변형된 표적 핵산의 수 증가;

표적 핵산으로의 삽입 길이 증가; 또는

표적 핵산에서 MMR 활성 감소

중 하나 이상을 초래하는, 방법.

청구항 82

제80항 또는 제81항에 있어서, 상기 세포는 생체 내 또는 생체 외에 있는, 방법.

청구항 83

주형 RNA로서,

a) 표적 핵산 서열에 돌연변이를 도입하기 위한 돌연변이 영역을 포함하는 이종 목적 서열(여기서, 선택적으로 이종 목적 서열은 5'에서 3'으로 사후 편집 상동성 영역, 돌연변이 영역, 및 사전 편집 상동성 영역을 포함함), 및

b) 표적 핵산 서열의 제1 부분에 결합하되, 제1 부분은 표적 핵산 서열의 제1 가닥에 있고, PBS 서열은 이종 목적 서열의 3'인, 프라이머 결합 부위 서열(PBS 서열), 및

c) RBD 동원 부위(RRS)이되, PBS 서열의 3' 또는 이종 목적 서열의 5'인, RRS

를 포함하는, 주형 RNA.

청구항 84

유전자 변형 폴리펩티드로서,

역전사효소(RT) 도메인; 및

표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

DBD 및 RT 도메인에 이종성인 RNA-결합 도메인(RBD)

을 포함하며,

여기서, 도메인은 N-말단에서 C-말단 방향으로

s) DBD, RT 도메인, RBD;

t) RT 도메인, DBD, RBD;

u) RBD, DBD, RT 도메인;

v) RBD, RT 도메인, DBD;

w) DBD, RBD, RT 도메인; 또는

x) RT 도메인, RBD, DBD

로 배열되는, 유전자 변형 폴리펩티드.

청구항 85

폴리펩티드 시스템(예를 들어, 폴리펩티드 복합체)으로서,

a) 역전사효소(RT) 도메인; 및

b) 표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas9 도메인, 예를 들어 Cas9 닉카제 도메인); 및

c) DBD 및 RT 도메인에 이종성인 RNA-결합 도메인(RBD)

을 포함하며,

여기서, (a), (b), 및 (c) 중 적어도 2개(예를 들어, 전부)는 별도의 폴리펩티드, 예를 들어 비공유적으로 복합체를 형성하는 별도의 폴리펩티드에 있는, 폴리펩티드 시스템.

청구항 86

제85항의 시스템의 폴리펩티드를 인코딩하는 핵산 또는 복수의 핵산.

청구항 87

시스템으로서,

제83항의 주형 RNA;

유전자 변형 폴리펩티드, 예를 들어 제84항의 유전자 변형 폴리펩티드, 또는 폴리펩티드 시스템, 예를 들어 제85항의 폴리펩티드 시스템; 및

표적 핵산 서열의 제2 부분에 결합하는 gRNA 스페이서(여기서, 제2 부분은 표적 핵산 서열의 제2 가닥 중 하나임); 및

유전자 변형 폴리펩티드 또는 폴리펩티드 시스템의 DBD에 결합하는 gRNA 스캐폴드

를 포함하는 제1 gRNA

를 포함하는, 시스템.

청구항 88

시스템으로서,

i) 제83항의 주형 RNA;

ii) DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

DBD에 이종성인 RNA-결합 도메인(RBD)(여기서, RBD는 주형 RNA의 RRS에 결합함)

을 포함하는 제1 폴리펩티드:

iii) 제1 폴리펩티드의 DBD를 표적 핵산 서열의 제2 부분으로 지시하는 gRNA 스페이서(여기서, 표적 핵산 서열의 제2 부분은 핵산 서열의 제2 가닥 상에 있음); 및

제1 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

를 포함하는 제1 gRNA;

iv) RT 도메인, 및

RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인)(여기서, 제2 폴리펩티드의 DBD는 제1 폴리펩티드의 DBD와 상이한 서열을 가짐)

을 포함하는 제2 폴리펩티드; 및

v) 제2 폴리펩티드의 DBD를 표적 핵산 서열의 제3 부분으로 지시하는 gRNA 스페이서(여기서, 제3 부분은 표적 핵산의 제1 가닥 상에 있음), 및

제2 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

를 포함하는 제2 gRNA

를 포함하는, 시스템.

청구항 89

시스템으로서,

i) 제1항 내지 제88항 중 어느 한 항의 주형 RNA

(여기서, 상기 주형 RNA는

표적 핵산 서열의 제3 부분에 상보적인 gRNA 스페이서(여기서, 제3 부분은 표적 핵산 서열의 제1 가닥 상에 있음); 및

gRNA 스캐폴드

를 포함함);

ii) DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

DBD에 이종성인 RNA-결합 도메인(RBD)(여기서, RBD는 주형 RNA의 RRS에 결합함)

을 포함하는 제1 폴리펩티드;

iii) 제1 폴리펩티드의 DBD를 표적 핵산 서열의 제2 부분으로 지시하는 gRNA 스페이서(여기서, 표적 핵산 서열의 제2 부분은 핵산 서열의 제2 가닥 상에 있음); 및

제1 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

를 포함하는 제1 gRNA; 및

iv) RT 도메인, 및

RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인)(여기서, 제2 폴리펩티드의 DBD는 제1 폴리펩티드의 DBD와 상이한 서열을 갖고, 주형 RNA의 gRNA 스캐폴드는 제2 폴리펩티드의 DBD에 결합함)

을 포함하는 제2 폴리펩티드

를 포함하는, 시스템.

청구항 90

폴리펩티드 시스템으로서,

DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인);

DBD에 이종성인 RNA-결합 도메인(RBD); 및

선택적으로, DBD와 RBD 사이에 배치된 링커

를 포함하는 제1 폴리펩티드; 및

RT 도메인, 및

RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

선택적으로, RT 도메인과 DBD 사이에 배치된 링커

를 포함하는 제2 폴리펩티드

를 포함하는, 폴리펩티드 시스템.

청구항 91

세포(예를 들어, 인간 세포)에서 표적 핵산을 변형시키는 방법으로서, 세포를 제1항 내지 제90항 중 어느 한 항의 시스템, 또는 이를 인코딩하는 핵산과 접촉시키는 단계를 포함하며, 이에 의해 표적 핵산을 변형시키는, 방법.

발명의 설명

기 술 분 야

[0001] **서열 목록**

[0002] 본 출원은 XML 서식으로 전자적으로 제출된 서열 목록을 포함하며, 이는 본 명세서에 전체적으로 참고로 포함된다. 2022년 9월 2일에 생성된 상기 XML 사본은 파일명이 V2065-7030WO_SL.xml이며, 크기가 15,727,041 바이트이다.

[0003] **관련 출원에 대한 상호 참조**

[0004] 본 출원은 2021년 9월 8일자로 출원된 미국 가출원 63/242,003의 이익을 주장한다. 상기 언급된 출원의 내용은 전체적으로 본 명세서에 참고로 포함된다.

배 경 기 술

[0005] 관심 핵산을 게놈으로 통합하는 것은, 삽입 사건을 촉진하는 특수화된 단백질의 부재 하에, 낮은 빈도 및 적은 부위 특이성으로 발생한다. CRISPR/Cas9와 같은 기존의 여러 접근법은 숙주 복구 경로에 의존하는 작은 편집체에 더 적합하고, 더 긴 서열의 통합에는 덜 효과적이다. Cre/loxP와 같은 기존의 다른 접근법은, loxP 부위를 게놈으로 삽입하는 제1 단계, 및 이어서 관심 서열을 loxP 부위로 삽입하는 제2 단계가 필요하다. 게놈에서 관심 서열을 삽입, 변경 또는 결실시키기 위한 개선된 조성물(예를 들어, 단백질 및 핵산) 및 방법이 당업계에 필요하다.

발명의 내용

[0006] 본 개시내용은, 생체 내 또는 시험관 내에서, 숙주 세포, 조직 또는 대상체의 하나 이상의 위치에서 게놈을 변경시키기 위한 새로운 조성물, 시스템, 및 방법에 관한 것이다. 구체적으로, 본 발명은 숙주 게놈에서 관심 서열을 삽입, 변경, 또는 결실시키기 위한 조성물, 시스템 및 방법을 특징으로 한다.

[0007] 본 개시내용에서 입증된 바와 같이, 출원인은 트랜스 주형 RNA 요소를 포함하는 별도의 RNA 주형 요소와 함께 유전자 변형 폴리펩티드, 또는 그러한 폴리펩티드를 인코딩하는 폴리뉴클레오티드를 전달함으로써 숙주 게놈에서 관심 서열을 편집할 수 있는 조성물 및 메커니즘을 알아내었다. 본 개시내용은 부분적으로 2개 이상의 상호작용에 의해 트랜스 주형 RNA를 유전자 변형 폴리펩티드:sgRNA:표적 게놈 DNA 복합체에 회합시키는 것에 관한 것이다. 이론에 구속되고자 함이 없이, 2개 이상의 상호작용 또는 앵커링 지점을 통한 그러한 회합은, 예를 들어 단일 또는 여러 뉴클레오티드 길이의 편집을 달성하기 위한 높은 재작성 활성을 달성할 수 있는 것으로 밝혀졌다. 본 명세서에 기재된 바와 같이, 2개 이상의 상호작용의 예는, 예를 들어 1) 통상적으로 유전자 변형 폴리펩티드와 트랜스 주형의 3' 말단 사이의 RRS:RBP 상호작용, 및 2) 5' 말단 블록 Cas9 스캐폴드 및 표적 DNA 상호작용에 대한 스페이서(추가적인 유전자 변형 폴리펩티드를 통해 매개됨)를 포함한다. 이러한 구성은 트랜스

주형 RNA를 유전자 변형 폴리펩티드:sgRNA:표적 게놈 DNA 복합체에 함께 고정시켜 재작성을 가능하게 하는 예시적인 상호작용을 예시한다. 5' 말단 블록 스페이서가 없는 경우 RRS:RBP 상호작용이 중요한 것으로 고려된다. RRS" RBD 상호작용과 5' 말단 블록 스페이서 둘 다의 존재는 높은 재작성 활성을 제공할 수 있고, 5' 말단 블록 스페이서의 존재는 더 약한 RRS:RBP 상호작용을 갖는 트랜스 주형으로 관찰된 재작성 활성을 구제하는 것으로 추가로 고려된다.

[0008] 본 조성물 또는 방법의 특징은 다음 열거된 구현예들 중 하나 이상을 포함할 수 있다.

[0009] 1. 주형 RNA로서,

[0010] a) 표적 핵산 서열에 돌연변이를 도입하기 위한 돌연변이 영역을 포함하는 이종 목적 서열(여기서, 선택적으로 이종 목적 서열은 5'에서 3'으로 사후 편집 상동성 영역, 돌연변이 영역, 및 사전 편집 상동성 영역을 포함함), 및

[0011] b) 표적 핵산 서열의 제1 부분에 결합하되, 제1 부분은 표적 핵산 서열의 제1 가닥에 있고, PBS 서열은 이종 목적 서열의 3'인, 프라이머 결합 부위 서열(PBS 서열), 및

[0012] c) RBD 동원 부위(RRS)이되, PBS 서열의 3' 또는 이종 목적 서열의 5'인, RRS

[0013] 를 포함하는, 주형 RNA.

[0014] 2. 주형 RNA로서,

[0015] a) 표적 핵산 서열에 돌연변이를 도입하기 위한 돌연변이 영역을 포함하는 이종 목적 서열(여기서, 선택적으로 이종 목적 서열은 5'에서 3'으로 사후 편집 상동성 영역, 돌연변이 영역, 및 사전 편집 상동성 영역을 포함함), 및

[0016] b) 표적 핵산 서열의 제1 부분에 결합하되, 제1 부분은 표적 핵산 서열의 제1 가닥에 있고, PBS 서열은 이종 목적 서열의 3'인, 프라이머 결합 부위 서열(PBS 서열), 및

[0017] c) RBD 동원 부위(RRS)이되, 선택적으로 RRS는 PBS 서열과 이종 목적 서열 사이에 위치하거나, 이종 목적 서열 내부에(예를 들어, 사전 편집 상동성 영역과 돌연변이 영역 사이에) 위치하는, RRS

[0018] 를 포함하는, 주형 RNA.

[0019] 3. 구현예 1 또는 2에 있어서, 말단 블록 서열, 예를 들어 표 41의 말단 블록 서열 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 추가로 포함하는, 주형 RNA.

[0020] 4. 선행하는 구현예 중 어느 하나에 있어서, 이종 목적 서열의 5'에 말단 블록을 포함하는, 주형 RNA.

[0021] 5. 선행하는 구현예 중 어느 하나에 있어서, PBS 서열의 3'에 말단 블록을 포함하며, 선택적으로 RRS는 말단 블록과 PBS 서열 사이에 위치하는, 주형 RNA.

[0022] 6. 선행하는 구현예 중 어느 하나에 있어서, PBS 서열의 3'에 제1 말단 블록 서열 및 이종 목적 서열의 5'에 제2 말단 블록 서열을 포함하는, 주형 RNA.

[0023] 7. 구현예 3 내지 6 중 어느 하나에 있어서, 말단 블록 서열은 이종 목적 서열의 5'이고 RRS는 PBS 서열의 3'인, 주형 RNA.

[0024] 8. 구현예 3 내지 6 중 어느 하나에 있어서, 말단 블록 서열은 PBS 서열의 3'이고 RRS는 이종 목적 서열의 5'인, 주형 RNA.

[0025] 9. 선행하는 구현예 중 어느 하나에 있어서, RRS는 표 40에 따른 서열 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열, 또는 이의 역상보체를 갖는, 주형 RNA.

[0026] 10. 선행하는 구현예 중 어느 하나에 있어서, 복수의 RRS, 예를 들어 2, 3, 4, 5, 또는 10개 RRS의 탠덤 어레이를 포함하는, 주형 RNA.

[0027] 11. 선행하는 구현예 중 어느 하나에 있어서, PBS 서열은 길이가 5 내지 1000 nt인, 주형 RNA.

[0028] 12. 선행하는 구현예 중 어느 하나에 있어서, PBS 서열은 8 내지 17개 뉴클레오티드, 예를 들어 표적 핵산 서열에 대하여 100% 동일성을 갖는 8 내지 17개 뉴클레오티드를 포함하는, 주형 RNA.

[0029] 13. 선행하는 구현예 중 어느 하나에 있어서, 사전 편집 상동성 영역은 최대 30개 뉴클레오티드, 예를 들어

최대 20개 뉴클레오티드, 예를 들어, 표적 핵산 서열에 대하여 100% 동일성을 갖는 최대 20개 뉴클레오티드를 포함하는, 주형 RNA.

[0030] 14. 구현예 1 내지 12 중 어느 하나에 있어서, 사후 편집 상동성 영역을 포함하지 않는, 주형 RNA.

[0031] 15. 선행하는 구현예 중 어느 하나에 있어서, 사후 편집 상동성 영역은 5 내지 1000, 5 내지 500개 뉴클레오티드, 예를 들어 표적 핵산 서열에 대하여 100% 동일성을 갖는 5 내지 500개 뉴클레오티드를 포함하는, 주형 RNA.

[0032] 16. 구현예 114에 있어서, 사후 편집 상동성 영역을 포함하지 않는, 주형 RNA.

[0033] 17. 선행하는 구현예 중 어느 하나에 있어서, 돌연변이 영역은 표적 핵산에 삽입, 결실, 또는 치환을 생성하도록 구성되는, 주형 RNA.

[0034] 18. 선행하는 구현예 중 어느 하나에 있어서,

[0035] 표적 핵산 서열의 상이한 부분(예를 들어, 제3 부분)에 상보적인 gRNA 스페이서(예를 들어, 여기서 상이한 부분(예를 들어, 제3 부분)은 표적 핵산 서열의 제1 가닥 상에 있음); 및

[0036] gRNA 스캐폴드

[0037] 를 추가로 포함하는, 주형 RNA.

[0038] 19. 구현예 18에 있어서, gRNA 스페이서는 이종 목적 서열의 5'인, 주형 RNA.

[0039] 20. 구현예 18 또는 19에 있어서, gRNA 스캐폴드는 gRNA 스페이서와 이종 목적 서열 사이에 위치하는, 주형 RNA.

[0040] 21. 구현예 18 내지 20 중 어느 하나에 있어서, gRNA 스페이서 및 PBS 서열은 표적 핵산 서열의 동일한 가닥에 결합하는, 주형 RNA.

[0041] 22. 구현예 18 내지 21 중 어느 하나에 있어서, gRNA 스페이서, 이종 목적 서열, 및 PBS 서열은 표적 핵산 서열의 동일한 가닥에 결합하는, 주형 RNA.

[0042] 23. 구현예 1 내지 8 중 어느 하나에 있어서, gRNA 스페이서 또는 gRNA 스캐폴드를 포함하지 않는, 주형 RNA.

[0043] 24. 선행하는 구현예 중 어느 하나에 있어서, RRS와 PBS 서열 사이에 최대 20개의 뉴클레오티드의 링커를 포함하는, 주형 RNA.

[0044] 25. 유전자 변형 폴리펩티드로서,

[0045] 역전사효소(RT) 도메인; 및

[0046] 표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0047] DBD 및 RT 도메인에 이종성인 RNA-결합 도메인(RBD)

[0048] 을 포함하는 유전자 변형 폴리펩티드.

[0049] 26. 유전자 변형 폴리펩티드로서,

[0050] 역전사효소(RT) 도메인; 및

[0051] 표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0052] DBD 및 RT 도메인에 이종성인 RNA-결합 도메인(RBD)

[0053] 을 포함하며, 여기서, 도메인은 N-말단에서 C-말단 방향으로

[0054] a) DBD, RT 도메인, RBD;

[0055] b) RT 도메인, DBD, RBD;

[0056] c) RBD, DBD, RT 도메인;

[0057] d) RBD, RT 도메인, DBD;

[0058] e) DBD, RBD, RT 도메인; 또는

[0059] f) RT 도메인, RBD, DBD

[0060] 로 배열되는, 유전자 변형 폴리펩티드.

[0061] 27. 유전자 변형 폴리펩티드로서,

[0062] 역전사효소(RT) 도메인; 및

[0063] 표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0064] DBD 및 RT 도메인에 이종성인 복수(예를 들어, 2, 3, 4, 또는 5개)의 RNA-결합 도메인(RBD)

[0065] 을 포함하는, 유전자 변형 폴리펩티드.

[0066] 28. 구현예 27에 있어서, RBD는 표 31에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 유전자 변형 폴리펩티드.

[0067] 29. 선행하는 구현예 중 어느 하나에 있어서, 복수의 RBD는 서로 동일한 아미노산 서열을 갖는, 유전자 변형 폴리펩티드.

[0068] 30. 선행하는 구현예 중 어느 하나에 있어서, 복수의 RBD는 서로 상이한 아미노산 서열을 갖는, 유전자 변형 폴리펩티드.

[0069] 31. 선행하는 구현예 중 어느 하나에 있어서, DBD는 표 7 또는 8에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 유전자 변형 폴리펩티드.

[0070] 32. 선행하는 구현예 중 어느 하나에 있어서, RT 도메인은 레트로바이러스로부터 유래되거나, 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 폴리펩티드 도메인인, 유전자 변형 폴리펩티드.

[0071] 33. 선행하는 구현예 중 어느 하나에 있어서, RT 도메인은 표 6에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 유전자 변형 폴리펩티드.

[0072] 34. 선행하는 구현예 중 어느 하나에 있어서, 유전자 변형 폴리펩티드는 링커를 포함하는, 유전자 변형 폴리펩티드.

[0073] 35. 선행하는 구현예 중 어느 하나에 있어서, 링커는 표 10에 따른 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는, 유전자 변형 폴리펩티드.

[0074] 36. 구현예 34 또는 35에 있어서, 링커는 DBD와 RT 도메인 사이, RT 도메인과 RBD 사이, 또는 RBD와 DBD 사이에 배치되는, 유전자 변형 폴리펩티드.

[0075] 37. 선행하는 구현예 중 어느 하나에 있어서, 유전자 변형 폴리펩티드는 N-말단에서 C-말단 방향으로

[0076] a) DBD, 제1 링커, RT 도메인, 제2 링커, RBD;

[0077] b) RT 도메인, 제1 링커, DBD, 제2 링커, RBD;

[0078] c) RBD, 제1 링커, DBD, 제2 링커, RT 도메인;

[0079] d) RBD, 제1 링커, RT 도메인, 제2 링커, DBD;

[0080] e) DBD, 제1 링커, RBD, 제2 링커, RT 도메인; 또는

[0081] f) RT 도메인, 제1 링커, RBD, 제2 링커, DBD

[0082] 를 포함하는, 유전자 변형 폴리펩티드.

[0083] 38. 선행하는 구현예 중 어느 하나에 있어서, 인테인-N 도메인을 포함하는 N-말단 부분 및 인테인-C 도메인

을 포함하는 C-말단 부분의 인테인-매개 융합에 의해 생성된, 유전자 변형 폴리펩티드.

[0084] 39. 폴리펩티드 시스템(예를 들어, 폴리펩티드 복합체)으로서,

[0085] a) 역전사효소(RT) 도메인; 및

[0086] b) 표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas9 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0087] c) DBD 및 RT 도메인에 이종성인 RNA-결합 도메인(RBD)

[0088] 을 포함하며,

[0089] 여기서, (a), (b), 및 (c) 중 적어도 2개(예를 들어, 전부)는 별도의 폴리펩티드, 예를 들어 비공유적으로 복합체를 형성하는 별도의 폴리펩티드에 있는, 폴리펩티드 시스템.

[0090] 40. 구현예 39에 있어서, 복합체 형성은 제2의 양립 가능한 이량체화 도메인에 결합하는 제1 이량체화 도메인에 의해 매개되는, 폴리펩티드 시스템.

[0091] 41. 구현예 40에 있어서, 복합체 형성은 제4의 양립 가능한 이량체화 도메인에 결합하는 제3 이량체화 도메인에 의해 매개되는, 폴리펩티드 시스템.

[0092] 42. 구현예 39 내지 41 중 어느 하나에 있어서,

[0093] RBD는 제1 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되고;

[0094] DBD는 제1 이량체화 도메인에 결합하는 제2 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되고;

[0095] DBD는 제3 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되고;

[0096] RT 도메인은 제3 이량체화 도메인에 결합하는 제4 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되는, 폴리펩티드 시스템.

[0097] 43. 구현예 39 내지 42 중 어느 하나에 있어서, 제1 및 제2 이량체화 도메인은 화학물질-유도 이량체화 도메인, 광-유도 이량체화 도메인, 항체-펩티드 이량체화 도메인, 또는 코일드 코일(coiled coil) 이량체화 도메인인, 폴리펩티드 시스템.

[0098] 44. 구현예 39 내지 43 중 어느 하나에 있어서, 제3 및 제4 이량체화 도메인은 화학물질-유도 이량체화 도메인, 광-유도 이량체화 도메인, 항체-펩티드 이량체화 도메인, 또는 코일드 코일 이량체화 도메인인, 폴리펩티드 시스템.

[0099] 45. 구현예 39 내지 44 중 어느 하나에 있어서, 제1 이량체화 도메인 및 제2 이량체화 도메인은 각각 복수의 카피, 예를 들어 2, 3, 4, 5, 10, 15, 20, 또는 30개 카피로 존재하는, 폴리펩티드 시스템.

[0100] 46. 구현예 39 내지 45 중 어느 하나에 있어서, 제3 이량체화 도메인 및 제4 이량체화 도메인은 각각 복수의 카피, 예를 들어 2, 3, 4, 5, 10, 15, 20, 또는 30개 카피로 존재하는, 폴리펩티드 시스템.

[0101] 47. 구현예 39 내지 46 중 어느 하나에 있어서, 제1 이량체화 도메인 및 제2 이량체화 도메인은 동일한 서열을 갖는(예를 들어, 제1 이량체화 도메인 및 제2 이량체화 도메인은 동종이량체를 형성함), 폴리펩티드 시스템.

[0102] 48. 구현예 39 내지 47 중 어느 하나에 있어서, 제3 이량체화 도메인 및 제4 이량체화 도메인은 동일한 서열을 갖는(예를 들어, 제3 이량체화 도메인 및 제4 이량체화 도메인은 동종이량체를 형성함), 폴리펩티드 시스템.

[0103] 49. 구현예 39 내지 48 중 어느 하나에 있어서, 제1 이량체화 도메인 및 제2 이량체화 도메인은 상이한 서열을 갖는(예를 들어, 제1 이량체화 도메인 및 제2 이량체화 도메인은 이종이량체를 형성함), 폴리펩티드 시스템.

[0104] 50. 구현예 39 내지 49 중 어느 하나에 있어서, 제3 이량체화 도메인 및 제4 이량체화 도메인은 상이한 서열을 갖는(예를 들어, 제3 이량체화 도메인 및 제4 이량체화 도메인은 이종이량체를 형성함), 폴리펩티드 시스템.

[0105] 51. 구현예 39 내지 50 중 어느 하나에 있어서, DBD는 하나 이상의 추가적인 DBD에 작동 가능하게 연결되며, 여기서 선택적으로 추가적인 DBD는 DBD와 동일한 서열을 갖는, 폴리펩티드 시스템.

[0106] 52. 구현예 39 내지 51 중 어느 하나에 있어서, RBD는 표 31에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 폴리펩티드 시스템.

[0107] 53. 구현예 39 내지 52 중 어느 하나에 있어서, 복수의 RBD는 서로 동일한 아미노산 서열을 갖는, 폴리펩티드 시스템.

[0108] 54. 구현예 39 내지 52 중 어느 하나에 있어서, 복수의 RBD는 서로 상이한 아미노산 서열을 갖는, 폴리펩티드 시스템.

[0109] 55. 구현예 39 내지 54 중 어느 하나에 있어서, DBD는 표 7 또는 8에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 폴리펩티드 시스템.

[0110] 56. 구현예 39 내지 55 중 어느 하나에 있어서, RT 도메인은 레트로바이러스로부터 유래되거나, 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 폴리펩티드 도메인인, 폴리펩티드 시스템.

[0111] 57. 구현예 39 내지 56 중 어느 하나에 있어서, RT 도메인은 표 6에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 폴리펩티드 시스템.

[0112] 58. 구현예 39 내지 57 중 어느 하나에 있어서, 각각의 링커는 독립적으로 표 10에 따른 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는, 폴리펩티드 시스템.

[0113] 59. 구현예 39 내지 57 중 어느 하나의 시스템의 폴리펩티드를 인코딩하는 핵산 또는 복수의 핵산.

[0114] 60. 시스템으로서,

[0115] 구현예 1 내지 24 중 어느 하나의 주형 RNA;

[0116] 구현예 25 내지 38 중 어느 하나의 유전자 변형 폴리펩티드 또는 구현예 39 내지 58 중 어느 하나의 폴리펩티드 시스템; 및

[0117] 표적 핵산 서열의 제2 부분에 결합하는 gRNA 스페이서(여기서, 제2 부분은 표적 핵산 서열의 제2 가닥 중 하나임); 및

[0118] 유전자 변형 폴리펩티드 또는 폴리펩티드 시스템의 DBD에 결합하는 gRNA 스캐폴드

[0119] 를 포함하는 제1 gRNA

[0120] 를 포함하는, 시스템.

[0121] 61. 구현예 60에 있어서, 주형 RNA는 gRNA 스페이서 또는 gRNA 스캐폴드를 포함하지 않는, 시스템.

[0122] 62. 구현예 60 또는 61에 있어서, gRNA 스페이서는 PBS 서열에 의해 결합된 표적 핵산 서열의 영역의 약 5, 10, 15, 20, 25, 30, 또는 40개 뉴클레오티드 내에 있는 표적 핵산 서열의 영역에 결합하는, 시스템.

[0123] 63. 구현예 60 내지 62 중 어느 하나에 있어서,

[0124] 제2 Cas 단백질(예를 들어, 사멸 Cas 단백질) 및

[0125] PBS 서열에 의해 결합된 위치의 3' 위치에서 표적 핵산의 제1 가닥에 결합하는 gRNA 스페이서, 및

[0126] 제2 Cas 단백질에 결합하는 gRNA 스캐폴드.

[0127] 를 포함하는 제2 gRNA

[0128] 를 추가로 포함하는, 시스템.

[0129] 64. 구현예 63에 있어서, 제2 Cas 단백질은 사멸 Cas 단백질(예를 들어, 사멸 Cas9 단백질) 또는 Cas 닉카제 단백질(예를 들어, Cas9 닉카제 단백질)인, 시스템.

[0130] 65. 구현예 63에 있어서, 제2 gRNA의 gRNA 스페이서는 적어도 18개 뉴클레오티드(예를 들어, 18 내지 28개

뉴클레오티드, 예를 들어, 18 내지 21개 뉴클레오티드)의 길이를 갖고, 제2 Cas 단백질은 사멸 Cas 단백질인, 시스템.

[0131] 66. 구현예 63에 있어서, 제2 gRNA의 gRNA 스페이서는 17개 뉴클레오티드 이하(예를 들어, 14 내지 17개 뉴클레오티드)의 길이를 갖고, 선택적으로 제2 Cas 단백질은 Cas 닉카제 단백질인, 시스템.

[0132] 67. 구현예 60에 있어서, 주형 RNA는

[0133] 표적 핵산 서열의 제3 부분에 상보적인 gRNA 스페이서(여기서, 제3 부분은 표적 핵산 서열의 제1 가닥 상에 있음); 및

[0134] gRNA 스캐폴드

[0135] 를 추가로 포함하는, 시스템.

[0136] 68. 구현예 67에 있어서, gRNA 스캐폴드는 유전자 변형 폴리펩티드 또는 폴리펩티드 시스템의 DBD에 결합하는, 시스템.

[0137] 69. 구현예 67 또는 68에 있어서, gRNA 스페이서는 17개 뉴클레오티드 이하의 길이를 갖는, 시스템.

[0138] 70. 구현예 60 내지 69 중 어느 하나에 있어서, 주형 RNA의 gRNA 스페이서는, 예를 들어 표적 핵산 서열의 제2 가닥에서, 주형 핵산의 닉 생성(nicking)을 유도하는, 시스템.

[0139] 71. 구현예 60 내지 69 중 어느 하나에 있어서, 주형 RNA의 gRNA 스페이서는 주형 핵산의 닉 생성을 유도하지 않는, 시스템.

[0140] 72. 시스템으로서,

[0141] i) 구현예 1 내지 24 중 어느 하나의 주형 RNA(예를 들어, 구현예 23의 주형 RNA);

[0142] ii) DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0143] DBD에 이종성인 RNA-결합 도메인(RBD)(여기서, RBD는 주형 RNA의 RRS에 결합함)

[0144] 을 포함하는 제1 폴리펩티드;

[0145] iii) 제1 폴리펩티드의 DBD를 표적 핵산 서열의 제2 부분으로 지시하는 gRNA 스페이서(여기서, 표적 핵산 서열의 제2 부분은 핵산 서열의 제2 가닥 상에 있음); 및

[0146] 제1 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

[0147] 를 포함하는 제1 gRNA;

[0148] iv) RT 도메인, 및

[0149] RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인)(여기서, 제2 폴리펩티드의 DBD는 제1 폴리펩티드의 DBD와 상이한 서열을 가짐)

[0150] 을 포함하는 제2 폴리펩티드; 및

[0151] v) 제2 폴리펩티드의 DBD를 표적 핵산 서열의 제3 부분으로 지시하는 gRNA 스페이서(여기서, 제3 부분은 표적 핵산의 제1 가닥 상에 있음), 및

[0152] 제2 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

[0153] 를 포함하는 제2 gRNA

[0154] 를 포함하는, 시스템.

[0155] 73. 구현예 72에 있어서, 제2 폴리펩티드의 DBD는 Cas 닉카제 도메인 또는 사멸 Cas 도메인을 포함하는, 시스템.

[0156] 74. 구현예 72에 있어서, 제2 RNA의 gRNA 스페이서는, 예를 들어 표적 핵산 서열의 제2 가닥에서, 주형 핵산의 닉 생성을 유도하는, 시스템.

[0157] 75. 구현예 72에 있어서, 제2 RNA의 gRNA 스페이서는 주형 핵산의 닉 생성을 유도하지 않는, 시스템.

[0158] 76. 구현예 72에 있어서, 제1 gRNA는 제2 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

[0159] 77. 구현예 72에 있어서, 제2 gRNA는 제1 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

[0160] 78. 시스템으로서,

[0161] i) 구현예 1 내지 24 중 어느 하나의 주형 RNA

[0162] (여기서, 주형 RNA는

[0163] 표적 핵산 서열의 제3 부분에 상보적인 gRNA 스페이서(여기서, 제3 부분은 표적 핵산 서열의 제1 가닥 상에 있음); 및

[0164] gRNA 스캐폴드

[0165] 를 포함함);

[0166] ii) DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0167] DBD에 이종성인 RNA-결합 도메인(RBD)(여기서, RBD는 주형 RNA의 RRS에 결합함)

[0168] 을 포함하는 제1 폴리펩티드;

[0169] iii) 제1 폴리펩티드의 DBD를 표적 핵산 서열의 제2 부분으로 지시하는 gRNA 스페이서(여기서, 표적 핵산 서열의 제2 부분은 핵산 서열의 제2 가닥 상에 있음); 및

[0170] 제1 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

[0171] 를 포함하는 제1 gRNA; 및

[0172] iv) RT 도메인, 및

[0173] RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인)(여기서, 제2 폴리펩티드의 DBD는 제1 폴리펩티드의 DBD와 상이한 서열을 갖고, 주형 RNA의 gRNA 스캐폴드는 제2 폴리펩티드의 DBD에 결합함)

[0174] 을 포함하는 제2 폴리펩티드

[0175] 를 포함하는, 시스템.

[0176] 79. 구현예 78에 있어서, 제2 폴리펩티드의 DBD는 Cas 닉카제 도메인 또는 사멸 Cas 도메인을 포함하는, 시스템.

[0177] 80. 구현예 78에 있어서, 주형 RNA의 gRNA 스페이서는, 예를 들어 표적 핵산 서열의 제2 가닥에서, 주형 핵산의 닉 생성을 유도하는, 시스템.

[0178] 81. 구현예 78에 있어서, 주형 RNA의 gRNA 스페이서는 주형 핵산의 닉 생성을 유도하지 않는, 시스템.

[0179] 82. 구현예 78 내지 중 어느 하나에 있어서, 제1 gRNA는 제2 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

[0180] 83. 구현예 78 내지 82 중 어느 하나에 있어서, 주형 RNA의 gRNA는 제1 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

[0181] 84. 폴리펩티드 시스템으로서,

[0182] DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인);

[0183] DBD에 이종성인 RNA-결합 도메인(RBD); 및

[0184] 선택적으로, DBD와 RBD 사이에 배치된 링커

[0185] 를 포함하는 제1 폴리펩티드; 및

[0186] RT 도메인, 및

[0187] RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0188] 선택적으로, RT 도메인과 DBD 사이에 배치된 링커

[0189] 를 포함하는 제2 폴리펩티드

[0190] 를 포함하는, 폴리펩티드 시스템.

[0191] 85. 구현예 1 내지 24 또는 60 내지 83 중 어느 하나에 있어서, 표적 핵산 서열은 표적 유전자, 인핸서, 또는 프로모터인, 주형 RNA 또는 시스템.

[0192] 86. 구현예 85에 있어서, 표적 핵산 서열은 인간 표적 유전자, 인간 인핸서, 또는 인간 프로모터인, 주형 RNA 또는 시스템.

[0193] 87. 선행하는 구현예 중 어느 하나에 있어서, RBD는 표 31의 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 갖는, 시스템 또는 폴리펩티드 시스템.

[0194] 88. 세포(예를 들어, 인간 세포)에서 표적 핵산을 변형시키는 방법으로서, 세포를 구현예 60 내지 83 중 어느 하나의 시스템, 또는 이를 인코딩하는 핵산과 접촉시키는 단계를 포함하며, 이에 의해 표적 핵산을 변형시키는, 방법.

[0195] 89. 구현예 88에 있어서, 제2 폴리펩티드의 존재는 제2 폴리펩티드가 결여된 다른 유사한 시스템과 비교하여

[0196] 표적 핵산의 풀림 증가;

[0197] 변형된 표적 핵산의 수 증가;

[0198] 표적 핵산으로의 삽입 길이 증가; 또는

[0199] 표적 핵산에서 MMR 활성 감소

[0200] 중 하나 이상을 초래하는, 방법.

[0201] 90. 구현예 88 또는 89에 있어서, 세포는 생체 내 또는 생체 외에 있는, 방법.

[0202] 91. 주형 RNA로서,

[0203] a) 표적 핵산 서열에 돌연변이를 도입하기 위한 돌연변이 영역을 포함하는 이종 목적 서열(여기서, 선택적으로 이종 목적 서열은 5'에서 3'으로 사후 편집 상동성 영역, 돌연변이 영역, 및 사전 편집 상동성 영역을 포함함), 및

[0204] b) 표적 핵산 서열의 제1 부분에 결합하되, 제1 부분은 표적 핵산 서열의 제1 가닥에 있고, PBS 서열은 이종 목적 서열의 3'인, 프라이머 결합 부위 서열(PBS 서열), 및

[0205] c) RBD 동원 부위(RRS)이되, PBS 서열의 3' 또는 이종 목적 서열의 5'인, RRS

[0206] 를 포함하는 주형 RNA.

[0207] 92. 구현예 91에 있어서, RRS는 표 S4에 열거된 바와 같은 주형 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열의 RRS를 포함하는, 주형 RNA.

[0208] 93. 구현예 91 또는 92에 있어서, 말단 블록 서열, 예를 들어 표 41의 말단 블록 서열을 추가로 포함하거나, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는, 주형 RNA.

[0209] 94. 구현예 93에 있어서, 말단 블록 서열은 이종 목적 서열의 5'이고(예를 들어, 주형 RNA의 5' 말단에 위치함), 선택적으로 RRS는 PBS 서열의 3'인, 주형 RNA.

[0210] 95. 구현예 94에 있어서, 말단 블록 서열은 gRNA 스캐폴드를 포함하는, 주형 RNA.

[0211] 96. 구현예 95에 있어서, gRNA 스캐폴드는 표 41로부터 선택되는, 주형 RNA.

[0212] 97. 구현예 95에 있어서, gRNA 스캐폴드는 Cas9 스캐폴드인, 주형 RNA.

[0213] 98. 구현예 93 내지 97 중 어느 하나에 있어서, 말단 블록 서열은, 예를 들어 말단 블록의 5' 말단에 위치하는(예를 들어, gRNA 스캐폴드의 5' 및/또는 주형 RNA의 5' 말단에 위치하는) gRNA 스페이서를 포함하는, 주형 RNA.

[0214] 99. 구현예 94 내지 98 중 어느 하나에 있어서, gRNA 스페이서는 프로-스페이서(예를 들어, 본 명세서에 기재된 바와 같음)인, 주형 RNA.

[0215] 100. 구현예 98에 있어서, 말단 블록은, 예를 들어 유전자 변형 폴리펩티드(예를 들어, 본 명세서에 기재된 바와 같음)의 DNA 결합 도메인에 결합하는, 주형 RNA.

[0216] 101. 구현예 100에 있어서, 말단 블록에 결합된 유전자 변형 폴리펩티드는 표적 핵산 서열의 제2 가닥에 닉(nick)을 형성하지 않는, 주형 RNA.

[0217] 102. 구현예 98 내지 101 중 어느 하나에 있어서, gRNA 스페이서는 표적 핵산 서열의 제1 부분에 대해 3'에 위치한 표적 핵산 서열의 제1 가닥의 제2 부분에 결합하는, 주형 RNA.

[0218] 103. 구현예 102에 있어서, gRNA 스페이서에 의해 결합된 제1 가닥 부분의 5' 말단은 제1 부분의 3' 말단으로부터 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 100, 100 내지 150, 또는 150 내지 200개 뉴클레오티드인, 주형 RNA.

[0219] 104. 구현예 98 내지 103 중 어느 하나에 있어서, 다음과 같은 주형 RNA:

[0220] (i) gRNA 스페이서는 17개 뉴클레오티드 이하, 예를 들어 약 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 또는 17개 뉴클레오티드의 길이를 갖고/갖거나;

[0221] (ii) gRNA 스페이서는 표적 핵산 서열의 제1 가닥 상의 제2 부분에 대해 100% 상보성을 갖고/갖거나;

[0222] (iii) gRNA 스페이서는 Cas 도메인에 의해 닉 생성 활성을 지시함.

[0223] 105. 구현예 104에 있어서, 다음과 같은 주형 RNA:

[0224] (i) gRNA 스페이서는 17개 뉴클레오티드 이하, 예를 들어 약 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 또는 17개 뉴클레오티드의 길이를 갖고;

[0225] (ii) gRNA 스페이서는 표적 핵산 서열의 제1 가닥 상의 제2 부분에 대해 100% 상보성을 갖음.

[0226] 106. 구현예 104에 있어서, 다음과 같은 주형 RNA:

[0227] (ii) gRNA 스페이서는 표적 핵산 서열의 제1 가닥 상의 제2 부분에 대해 100% 상보성을 갖고;

[0228] (iii) gRNA 스페이서는 Cas 도메인에 의해 닉 생성 활성을 지시함.

[0229] 107. 구현예 93 내지 106 중 어느 하나에 있어서, 말단 블록 서열은 PBS 서열 및/또는 RRS의 3'이고(예를 들어, 주형 RNA의 3' 말단에 위치함), 선택적으로 RRS는 이종 목적 서열의 5'인, 주형 RNA.

[0230] 108. 구현예 107에 있어서, 말단 블록 서열은 GGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC(서열번호 18,101), 표 41의 말단 블록 서열을 포함하거나, 이 중 임의의 것에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는, 주형 RNA.

[0231] 109. 구현예 93 내지 108 중 어느 하나에 있어서, 말단 블록 서열은 압타머를 포함하는, 주형 RNA.

[0232] 110. 구현예 93 내지 109 중 어느 하나에 있어서, 말단 블록 서열은 RNA 압타머-결합 단백질(예를 들어, DBD에서, 예를 들어 유전자 변형 폴리펩티드에 부착된 RNA 압타머-결합 단백질)에 결합할 수 있는, 주형 RNA.

[0233] 111. 구현예 93 내지 110 중 어느 하나에 있어서, 말단 블록은 하나 이상의 헤어핀(예를 들어, 1, 2, 3, 4, 또는 5개 헤어핀)을 포함하는, 주형 RNA.

[0234] 112. 구현예 93 내지 111 중 어느 하나에 있어서, 말단 블록은 ePEG 말단 블록을 포함하는, 주형 RNA.

[0235] 113. 구현예 91 내지 92 중 어느 하나에 있어서,

[0236] 5' 말단 블록 서열, 예를 들어 표 41의 말단 블록 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 포함함(여기서, 5' 말단 블록 서열은 이종 목적 서열의 5'이고(예를

들어, 주형 RNA의 5' 말단에 위치함), 선택적으로 RRS는 PBS 서열의 3'임); 및

[0237] 3' 말단 블록 서열, 예를 들어 표 41의 말단 블록 서열 또는 서열 GGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC(서열번호 18,101), 또는 이들 중 임의의 것에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 포함함(여기서, 3' 말단 블록 서열은 PBS 서열 및/또는 RRS의 3'이고(예를 들어, 주형 RNA의 3' 말단에 위치함), 선택적으로 RRS는 이종 목적 서열의 5'임)

[0238] 를 추가로 포함하는, 주형 RNA.

[0239] 114. 선행하는 구현예 중 어느 하나에 있어서, RRS는 MS2 서열을 포함하는, 주형 RNA.

[0240] 115. 선행하는 구현예 중 어느 하나에 있어서, RRS는 MCP 폴리펩티드에 결합하는, 주형 RNA.

[0241] 116. 선행하는 구현예 중 어느 하나에 있어서, RRS는 PP7 서열을 포함하는, 주형 RNA.

[0242] 117. 선행하는 구현예 중 어느 하나에 있어서, RRS 및 PBS는 약 5 내지 10, 10 내지 15, 또는 15 내지 20개 뉴클레오티드(예를 들어, 약 8개 뉴클레오티드 또는 약 16개 뉴클레오티드)의 길이를 갖는 영역에 의해 분리되는, 주형 RNA.

[0243] 118. 선행하는 구현예 중 어느 하나에 있어서, RRS는 표 40에 따른 서열 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 갖는, 주형 RNA.

[0244] 119. 선행하는 구현예 중 어느 하나에 있어서, 복수의 RRS(예를 들어, 동일하거나 상이한 RRS), 예를 들어 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개의 RRS, 예를 들어 2, 3, 4, 5, 또는 10개 RRS의 탠덤 어레이를 포함하는, 주형 RNA.

[0245] 120. 구현예 119에 있어서, 복수의 RRS는 각각 MS2 서열을 포함하는, 주형 RNA.

[0246] 121. 구현예 119 또는 120에 있어서, 복수의 RRS는 MS2 서열의 4개 반복부를 포함하는, 주형 RNA.

[0247] 122. 선행하는 구현예 중 어느 하나에 있어서, PBS 서열은 8 내지 17개 뉴클레오티드, 예를 들어 표적 핵산 서열에 대하여 100% 동일성을 갖는 8 내지 17개 뉴클레오티드를 포함하는, 주형 RNA.

[0248] 123. 구현예 122에 있어서, PBS 서열은 약 8, 13, 또는 17개 뉴클레오티드의 길이를 갖는, 주형 RNA.

[0249] 124. 구현예 122에 있어서, PBS 서열은 약 13개 뉴클레오티드의 길이를 갖는, 주형 RNA.

[0250] 125. 선행하는 구현예 중 어느 하나에 있어서, 사전 편집 상동성 영역은 최대 20개 뉴클레오티드, 예를 들어, 표적 핵산 서열에 대하여 100% 동일성을 갖는 최대 20개 뉴클레오티드를 포함하는, 주형 RNA.

[0251] 126. 선행하는 구현예 중 어느 하나에 있어서, 사후 편집 상동성 영역은 5 내지 500개 뉴클레오티드, 예를 들어 표적 핵산 서열에 대하여 100% 동일성을 갖는 5 내지 500개 뉴클레오티드를 포함하는, 주형 RNA.

[0252] 127. 선행하는 구현예 중 어느 하나에 있어서, 사후 편집 상동성 영역은 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 또는 60 내지 70개 뉴클레오티드, 예를 들어 약 12개 뉴클레오티드 또는 약 63개 뉴클레오티드를 포함하는, 주형 RNA.

[0253] 128. 구현예 127에 있어서, 사후 편집 상동성 영역은, 예를 들어 거의 일정한 간격(예를 들어, 약 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 또는 20개 뉴클레오티드 간격으로 떨어져 있음)으로, 하나 이상(예를 들어, 1, 2, 3, 4, 또는 5개)의 단일 뉴클레오티드 치환을 포함하는, 주형 RNA.

[0254] 129. 선행하는 구현예 중 어느 하나에 있어서, 돌연변이 영역은 표적 핵산에 삽입, 결실, 또는 치환을 생성하도록 구성되는, 주형 RNA.

[0255] 130. 선행하는 구현예 중 어느 하나에 있어서, gRNA 스페이서는 표적 핵산 서열의 상이한 부분(예를 들어, 제3 부분)에 상보적이며, 예를 들어, 여기서 상이한 부분(예를 들어, 제3 부분)은 표적 핵산 서열의 제1 가닥 상에 있는, 주형 RNA.

[0256] 131. 구현예 130에 있어서, gRNA는 이종 목적 서열의 5'인, 주형 RNA.

[0257] 132. 구현예 130 또는 131에 있어서, gRNA 스캐폴드는 gRNA 스페이서와 이종 목적 서열 사이에 위치하는, 주형 RNA.

[0258] 133. 구현예 130 내지 132 중 어느 하나에 있어서, gRNA 스페이서 및 PBS 서열은 표적 핵산 서열의 동일한 가닥에 결합하는, 주형 RNA.

[0259] 134. 구현예 130 내지 133 중 어느 하나에 있어서, gRNA 스페이서, 이종 목적 서열, 및 PBS 서열은 표적 핵산 서열의 동일한 가닥에 결합하는, 주형 RNA.

[0260] 135. 구현예 91 내지 129 중 어느 하나에 있어서, gRNA 스페이서 또는 gRNA 스캐폴드를 포함하지 않는, 주형 RNA.

[0261] 136. 선행하는 구현예 중 어느 하나에 있어서, RRS와 PBS 서열 사이에 최대 20개의 뉴클레오티드의 링커를 포함하는, 주형 RNA.

[0262] 137. 선행하는 구현예 중 어느 하나에 있어서, 주형 RNA는 선형인, 주형 RNA.

[0263] 138. 선행하는 구현예 중 어느 하나에 있어서, 주형 RNA는 환형인, 주형 RNA.

[0264] 139. 유전자 변형 폴리펩티드로서,

[0265] 역전사효소(RT) 도메인; 및

[0266] 표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0267] DBD 및 RT 도메인에 이종성인 RNA-결합 도메인(RBD)

[0268] 을 포함하며,

[0269] 여기서 도메인은 N-말단에서 C-말단 방향으로

[0270] g) DBD, RT 도메인, RBD;

[0271] h) RT 도메인, DBD, RBD;

[0272] i) RBD, DBD, RT 도메인;

[0273] j) RBD, RT 도메인, DBD;

[0274] k) DBD, RBD, RT 도메인; 또는

[0275] l) RT 도메인, RBD, DBD

[0276] 로 배열되는, 유전자 변형 폴리펩티드.

[0277] 140. 구현예 139에 있어서, 하나 이상(예를 들어, 1, 2, 3, 또는 4개)의 추가의 RBD(예를 들어, RBD에 인접한, 예를 들어 RBD의 하나 이상의 추가 카피)를 추가로 포함하는, 유전자 변형 폴리펩티드.

[0278] 141. 유전자 변형 폴리펩티드로서,

[0279] 역전사효소(RT) 도메인; 및

[0280] 표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0281] DBD 및 RT 도메인에 이종성인 복수(예를 들어, 2, 3, 4, 또는 5개)의 RNA-결합 도메인(RBD)

[0282] 을 포함하는, 유전자 변형 폴리펩티드.

[0283] 142. 선행하는 구현예 중 어느 하나에 있어서, RBD는 표 31에 따른 아미노산 서열 또는 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 유전자 변형 폴리펩티드의 RBD의 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는, 유전자 변형 폴리펩티드.

[0284] 143. 선행하는 구현예 중 어느 하나에 있어서, 복수의 RBD는 서로 동일한 아미노산 서열을 갖는, 유전자 변형 폴리펩티드.

[0285] 144. 선행하는 구현예 중 어느 하나에 있어서, 복수의 RBD는 서로 상이한 아미노산 서열을 갖는, 유전자 변

형 폴리펩티드.

[0286] 145. 선행하는 구현예 중 어느 하나에 있어서, DBD는 표 7 또는 8에 따른 아미노산 서열 또는 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 유전자 변형 폴리펩티드의 DBD의 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는, 유전자 변형 폴리펩티드.

[0287] 146. 선행하는 구현예 중 어느 하나에 있어서, RT 도메인은 레트로바이러스로부터 유래되거나, 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 폴리펩티드 도메인인, 유전자 변형 폴리펩티드.

[0288] 147. 선행하는 구현예 중 어느 하나에 있어서, RT 도메인은 표 6에 따른 아미노산 서열 또는 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 유전자 변형 폴리펩티드의 RT 도메인의 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는, 유전자 변형 폴리펩티드.

[0289] 148. 선행하는 구현예 중 어느 하나에 있어서,

[0290] (a) RBD는 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 유전자 변형 폴리펩티드의 RBD의 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하고;

[0291] (b) DBD는 표 S1 내지 S3 중 임의의 것에 열거된 상기 유전자 변형 폴리펩티드의 DBD의 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하고;

[0292] (c) RT 도메인은 표 S1 내지 S3 중 임의의 것에 열거된 상기 유전자 변형 폴리펩티드의 RT 도메인의 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는,

[0293] 유전자 변형 폴리펩티드.

[0294] 149. 선행하는 구현예 중 어느 하나에 있어서, 유전자 변형 폴리펩티드는 링커를 포함하는, 유전자 변형 폴리펩티드.

[0295] 150. 구현예 149에 있어서, 링커는 길이가 2 내지 5개 아미노산(예를 들어, 길이가 4개 아미노산)인, 유전자 변형 폴리펩티드.

[0296] 151. 구현예 149에 있어서, 링커는 길이가 5 내지 10개 아미노산(예를 들어, 길이가 8개 아미노산)인, 유전자 변형 폴리펩티드.

[0297] 152. 구현예 149에 있어서, 링커는 길이가 10 내지 20개 아미노산(예를 들어, 길이가 16개 아미노산)인, 유전자 변형 폴리펩티드.

[0298] 153. 구현예 149 내지 152 중 어느 하나에 있어서, 링커는 표 10에 따른 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는, 유전자 변형 폴리펩티드.

[0299] 154. 구현예 149 내지 153 중 어느 하나에 있어서, 링커는 DBD와 RT 도메인 사이, RT 도메인과 RBD 사이, 또는 RBD와 DBD 사이에 배치되는, 유전자 변형 폴리펩티드.

[0300] 155. 구현예 149 내지 154 중 어느 하나에 있어서, 제1 링커 및 제2 링커를 포함하고,

[0301] (i) 제1 링커는 DBD와 RT 도메인 사이에 배치되고, 제2 링커는 RT 도메인과 RBD 사이에 배치되거나;

[0302] (ii) 제1 링커는 DBD와 RBD 사이에 배치되고, 제2 링커는 RBD와 RT 도메인 사이에 배치되거나;

[0303] (iii) 제1 링커는 RT 도메인과 DBD 사이에 배치되고, 제2 링커는 DBD와 RBD 사이에 배치되는,

[0304] 유전자 변형 폴리펩티드.

[0305] 156. 선행하는 구현예 중 어느 하나에 있어서, 유전자 변형 폴리펩티드는 N-말단에서 C-말단 방향으로

[0306] g) DBD, 제1 링커, RT 도메인, 제2 링커, RBD;

[0307] h) RT 도메인, 제1 링커, DBD, 제2 링커, RBD;

[0308] i) RBD, 제1 링커, DBD, 제2 링커, RT 도메인;

[0309] j) RBD, 제1 링커, RT 도메인, 제2 링커, DBD;

[0310] k) DBD, 제1 링커, RBD, 제2 링커, RT 도메인; 또는

[0311] l) RT 도메인, 제1 링커, RBD, 제2 링커, DBD

[0312] 를 포함하는, 유전자 변형 폴리펩티드.

[0313] 157. 선행하는 구현예 중 어느 하나에 있어서, 인테인-N 도메인을 포함하는 N-말단 부분 및 인테인-C 도메인을 포함하는 C-말단 부분의 인테인-매개 융합에 의해 생성된, 유전자 변형 폴리펩티드

[0314] 158. 선행하는 구현예 중 어느 하나에 있어서, DBD는 Cas 도메인, 예를 들어 Cas9 도메인, 예를 들어 Cas9 닉카제 도메인(예를 들어, 본 명세서에 기재된 바와 같음)을 포함하는, 유전자 변형 폴리펩티드.

[0315] 159. 구현예 158에 있어서, Cas 도메인은 dCas9인, 유전자 변형 폴리펩티드.

[0316] 160. 구현예 158에 있어서, Cas 도메인은 nCas9 도메인인, 유전자 변형 폴리펩티드.

[0317] 161. 선행하는 구현예 중 어느 하나에 있어서, RT 도메인은 AVIRE 도메인(예를 들어, 본 명세서에 기재된 바와 같음, 예를 들어 표 6에 열거된 바와 같은 AVIRE RT 도메인), 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함하는, 유전자 변형 폴리펩티드.

[0318] 162. 구현예 161에 있어서, PBS 서열은 8개 뉴클레오티드 초과, 예를 들어 약 9, 10, 11, 12, 13, 14, 15, 16, 또는 17개 뉴클레오티드의 길이를 갖는, 유전자 변형 폴리펩티드.

[0319] 163. 선행하는 구현예 중 어느 하나에 있어서, RT 도메인은 MLVMS 도메인(예를 들어, 본 명세서에 기재된 바와 같음, 예를 들어 표 6에 열거된 바와 같은 MLVMS RT 도메인), 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 서열 동일성을 갖는 아미노산 서열을 포함하는, 유전자 변형 폴리펩티드.

[0320] 164. 선행하는 구현예 중 어느 하나에 있어서, RT 도메인은 레트로트랜스포존 RT 도메인을 포함하는, 유전자 변형 폴리펩티드.

[0321] 165. 선행하는 구현예 중 어느 하나에 있어서, 도메인은 N-말단에서 C-말단 방향으로

[0322] a) DBD, RT 도메인, RBD;

[0323] b) RT 도메인, DBD, RBD;

[0324] c) RBD, DBD, RT 도메인;

[0325] d) RBD, RT 도메인, DBD;

[0326] e) DBD, RBD, RT 도메인; 또는

[0327] f) RT 도메인, RBD, DBD

[0328] 로 배열되는, 유전자 변형 폴리펩티드.

[0329] 166. 구현예 165에 있어서, 하나 이상(예를 들어, 1, 2, 3, 또는 4개)의 추가의 RBD(예를 들어, RBD에 인접한, 예를 들어 RBD의 하나 이상의 추가 카피)를 추가로 포함하는, 유전자 변형 폴리펩티드.

[0330] 167. 구현예 165 또는 166에 있어서, 하나 이상의 추가의 RT 도메인(예를 들어, RT 도메인에 인접한, 예를 들어 RT 도메인의 하나 이상의 추가 카피)를 추가로 포함하는, 유전자 변형 폴리펩티드.

[0331] 168. 구현예 167에 있어서, 추가 RT 도메인 중 하나 이상은 AVIRE 도메인(예를 들어, 본 명세서에 기재된 바와 같음)을 포함하는, 유전자 변형 폴리펩티드.

[0332] 169. 구현예 167 또는 168에 있어서, 추가 RT 도메인 중 하나 이상은 MLVMS 도메인(예를 들어, 본 명세서에 기재된 바와 같음)을 포함하는, 유전자 변형 폴리펩티드.

[0333] 170. 선행하는 구현예 중 어느 하나에 있어서, RNA 압타머-결합 도메인을 추가로 포함하는, 유전자 변형 폴리펩티드.

[0334] 171. 구현예 170에 있어서, DBD는, 예를 들어 링커를 통해, RNA 압타머-결합 도메인에 부착되는, 유전자 변형 폴리펩티드.

[0335] 172. 폴리펩티드 시스템(예를 들어, 폴리펩티드 복합체)으로서,

[0336] a) 역전사효소(RT) 도메인; 및

[0337] b) 표적 핵산 서열에 결합하고 RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas9 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0338] c) DBD 및 RT 도메인에 이종성인 RNA-결합 도메인(RBD)

[0339] 을 포함하며,

[0340] 여기서, (a), (b), 및 (c) 중 적어도 2개(예를 들어, 전부)는 별도의 폴리펩티드, 예를 들어 비공유적으로 복합체를 형성하는 별도의 폴리펩티드에 있는, 폴리펩티드 시스템.

[0341] 173. 구현예 172에 있어서, RT 도메인 및 DBD는 별도의 폴리펩티드에 있는, 폴리펩티드 시스템.

[0342] 174. 구현예 172에 있어서, RT 도메인 및 RBD는 별도의 폴리펩티드에 있는, 폴리펩티드 시스템.

[0343] 175. 구현예 172에 있어서, 복합체 형성은 제2의 양립 가능한 이량체화 도메인에 결합하는 제1 이량체화 도메인에 의해 매개되는, 폴리펩티드 시스템.

[0344] 176. 구현예 172에 있어서, 복합체 형성은 제4의 양립 가능한 이량체화 도메인에 결합하는 제3 이량체화 도메인에 의해 매개되는, 폴리펩티드 시스템.

[0345] 177. 구현예 172 내지 176 중 어느 하나에 있어서,

[0346] RBD는 제1 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되고;

[0347] DBD는 제1 이량체화 도메인에 결합하는 제2 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되고;

[0348] DBD는 제3 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되고;

[0349] RT 도메인은 제3 이량체화 도메인에 결합하는 제4 이량체화 도메인에 (예를 들어, 링커를 통해) 작동 가능하게 연결되는, 폴리펩티드 시스템.

[0350] 178. 구현예 172 내지 177 중 어느 하나에 있어서, 제1 및 제2 이량체화 도메인은 화학물질-유도 이량체화 도메인, 광-유도 이량체화 도메인, 항체-펩티드 이량체화 도메인, 또는 코일드 코일이량체화 도메인인, 폴리펩티드 시스템.

[0351] 179. 구현예 172 내지 178 중 어느 하나에 있어서, 제3 및 제4 이량체화 도메인은 화학물질-유도 이량체화 도메인, 광-유도 이량체화 도메인, 항체-펩티드 이량체화 도메인, 또는 코일드 코일 이량체화 도메인인, 폴리펩티드 시스템.

[0352] 180. 구현예 172 내지 179 중 어느 하나에 있어서, 제1 이량체화 도메인 및 제2 이량체화 도메인은 각각 복수의 카피, 예를 들어 2, 3, 4, 5, 10, 15, 20, 또는 30개 카피로 존재하는, 폴리펩티드 시스템.

[0353] 181. 구현예 172 내지 180 중 어느 하나에 있어서, 제3 이량체화 도메인 및 제4 이량체화 도메인은 각각 복수의 카피, 예를 들어 2, 3, 4, 5, 10, 15, 20, 또는 30개 카피로 존재하는, 폴리펩티드 시스템.

[0354] 182. 구현예 172 내지 181 중 어느 하나에 있어서, 제1 이량체화 도메인 및 제2 이량체화 도메인은 동일한 서열을 갖는(예를 들어, 제1 이량체화 도메인 및 제2 이량체화 도메인은 동종이량체를 형성함), 폴리펩티드 시스템.

[0355] 183. 구현예 172 내지 182 중 어느 하나에 있어서, 제3 이량체화 도메인 및 제4 이량체화 도메인은 동일한 서열을 갖는(예를 들어, 제3 이량체화 도메인 및 제4 이량체화 도메인은 동종이량체를 형성함), 폴리펩티드 시스템.

[0356] 184. 구현예 172 내지 181 중 어느 하나에 있어서, 제1 이량체화 도메인 및 제2 이량체화 도메인은 상이한 서열을 갖는(예를 들어, 제1 이량체화 도메인 및 제2 이량체화 도메인은 이종이량체를 형성함), 폴리펩티드 시스템.

[0357] 185. 구현예 172 내지 184 중 어느 하나에 있어서, 제3 이량체화 도메인 및 제4 이량체화 도메인은 상이한 서열을 갖는(예를 들어, 제3 이량체화 도메인 및 제4 이량체화 도메인은 이종이량체를 형성함), 폴리펩티드 시스템.

[0358] 186. 구현예 172 내지 185 중 어느 하나에 있어서, DBD는 하나 이상의 추가적인 DBD에 작동 가능하게 연결되며, 여기서 선택적으로 추가적인 DBD는 DBD와 동일한 서열을 갖는, 폴리펩티드 시스템.

[0359] 187. 구현예 172 내지 186 중 어느 하나에 있어서, RBD는 표 31에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 폴리펩티드 시스템.

[0360] 188. 구현예 172 내지 187 중 어느 하나에 있어서, 복수의 RBD는 서로 동일한 아미노산 서열을 갖는, 폴리펩티드 시스템.

[0361] 189. 구현예 172 내지 188 중 어느 하나에 있어서, 복수의 RBD는 서로 상이한 아미노산 서열을 갖는, 폴리펩티드 시스템.

[0362] 190. 구현예 172 내지 189 중 어느 하나에 있어서, DBD는 표 31에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 폴리펩티드 시스템.

[0363] 191. 구현예 172 내지 190 중 어느 하나에 있어서, RT 도메인은 레트로바이러스로부터 유래되거나, 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 폴리펩티드 도메인인, 폴리펩티드 시스템.

[0364] 192. 구현예 172 내지 191 중 어느 하나에 있어서, RT 도메인은 표 6에 따른 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는, 폴리펩티드 시스템.

[0365] 193. 구현예 172 내지 192 중 어느 하나에 있어서, 각각의 링커는 독립적으로 표 10에 따른 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함하는, 폴리펩티드 시스템.

[0366] 194. 구현예 172 내지 193 중 어느 하나의 시스템의 폴리펩티드를 인코딩하는 핵산 또는 복수의 핵산.

[0367] 195. 시스템으로서,

[0368] 구현예 91 내지 138 중 어느 하나의 주형 RNA;

[0369] 유전자 변형 폴리펩티드, 예를 들어 구현예 139 내지 171 중 어느 하나의 유전자 변형 폴리펩티드, 또는 폴리펩티드 시스템, 예를 들어 구현예 172 내지 193 중 어느 하나의 폴리펩티드 시스템; 및

[0370] 표적 핵산 서열의 제2 부분에 결합하는 gRNA 스페이서(여기서, 제2 부분은 표적 핵산 서열의 제2 가닥 중 하나임); 및

[0371] 유전자 변형 폴리펩티드 또는 폴리펩티드 시스템의 DBD에 결합하는 gRNA 스캐폴드

[0372] 를 포함하는 제1 gRNA

[0373] 를 포함하는 시스템.

[0374] 196. 구현예 195에 있어서, 제1 gRNA의 gRNA 스캐폴드는 주형 RNA의 gRNA 서열과 동일한 단백질 결합 특이성을 갖는, 시스템.

[0375] 197. 구현예 196에 있어서, 주형 RNA의 gRNA 서열은 (예를 들어, 유전자 변형 폴리펩티드의 DBD에서) 유전자 변형 폴리펩티드의 제1 카피에 결합하고, 제1 gRNA의 gRNA 스캐폴드는 (예를 들어, 유전자 변형 폴리펩티드의 DBD에서) 유전자 변형 폴리펩티드의 제2 카피에 결합하는, 시스템.

[0376] 198. 구현예 195에 있어서, 주형 RNA는 gRNA 스페이서 또는 gRNA 스캐폴드를 포함하지 않는, 시스템.

[0377] 199. 구현예 195 또는 198에 있어서, gRNA 스페이서는 PBS 서열에 의해 결합된 표적 핵산 서열의 영역의 약 5, 10, 15, 20, 25, 30, 또는 40개 뉴클레오티드 내에 있는 표적 핵산 서열의 영역에 결합하는, 시스템.

[0378] 200. 구현예 195 내지 199 중 어느 하나에 있어서,

[0379] 제2 Cas 단백질(예를 들어, 사멸 Cas 단백질) 및

[0380] PBS 서열에 의해 결합된 위치의 3' 위치에서 표적 핵산의 제1 가닥에 결합하는 gRNA 스페이서, 및

[0381] 제2 Cas 단백질에 결합하는 gRNA 스캐폴드

[0382] 를 포함하는 제2 gRNA

[0383] 를 추가로 포함하는, 시스템.

[0384] 201. 구현예 200에 있어서, 제2 Cas 단백질은 사멸 Cas 단백질(예를 들어, 사멸 Cas9 단백질) 또는 Cas 닉카제 단백질(예를 들어, Cas9 닉카제 단백질)인, 시스템.

[0385] 202. 구현예 200에 있어서, 제2 gRNA의 gRNA 스페이서는 적어도 18개 뉴클레오티드(예를 들어, 18 내지 28개 뉴클레오티드, 예를 들어, 18 내지 21개 뉴클레오티드)의 길이를 갖고, 제2 Cas 단백질은 사멸 Cas 단백질인, 시스템.

[0386] 203. 구현예 200에 있어서, 제2 gRNA의 gRNA 스페이서는 17개 뉴클레오티드 이하(예를 들어, 14 내지 17개 뉴클레오티드)의 길이를 갖고, 선택적으로 제2 Cas 단백질은 Cas 닉카제 단백질인, 시스템.

[0387] 204. 구현예 195에 있어서, 주형 RNA는

[0388] 표적 핵산 서열의 제3 부분에 상보적인 gRNA 스페이서(여기서, 제3 부분은 표적 핵산 서열의 제1 가닥 상에 있음); 및

[0389] gRNA 스캐폴드

[0390] 를 추가로 포함하는, 시스템.

[0391] 205. 구현예 204에 있어서, gRNA 스캐폴드는 유전자 변형 폴리펩티드 또는 폴리펩티드 시스템의 DBD에 결합하는, 시스템.

[0392] 206. 구현예 204 또는 205에 있어서, gRNA 스페이서는 17개 뉴클레오티드 이하의 길이를 갖는, 시스템.

[0393] 207. 구현예 195 내지 206 중 어느 하나에 있어서, 주형 RNA의 gRNA 스페이서는, 예를 들어 표적 핵산 서열의 제2 가닥에서, 주형 핵산의 닉 생성을 유도하는, 시스템.

[0394] 208. 구현예 195 내지 206 중 어느 하나에 있어서, 주형 RNA의 gRNA 스페이서는 주형 핵산의 닉 생성을 유도하지 않는, 시스템.

[0395] 209. 시스템으로서,

[0396] i) 구현예 91 내지 138 중 어느 하나의 주형 RNA(예를 들어, 구현예 16의 주형 RNA);

[0397] ii) DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0398] DBD에 이종성인 RNA-결합 도메인(RBD)(여기서, RBD는 주형 RNA의 RRS에 결합함)

[0399] 을 포함하는 제1 폴리펩티드;

[0400] iii) 제1 폴리펩티드의 DBD를 표적 핵산 서열의 제2 부분으로 지시하는 gRNA 스페이서(여기서, 표적 핵산 서열의 제2 부분은 핵산 서열의 제2 가닥 상에 있음); 및

[0401] 제1 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

[0402] 를 포함하는 제1 gRNA;

[0403] iv) RT 도메인, 및

[0404] RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인)(여기서, 제2 폴리펩티드의 DBD는 제1 폴리펩티드의 DBD와 상이한 서열을 가짐)

[0405] 을 포함하는 제2 폴리펩티드; 및

[0406] v) 제2 폴리펩티드의 DBD를 표적 핵산 서열의 제3 부분으로 지시하는 gRNA 스페이서(여기서, 제3 부분은 표적 핵산의 제1 가닥 상에 있음), 및

[0407] 제2 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

[0408] 를 포함하는 제2 gRNA

[0409] 를 포함하는, 시스템.

[0410] 210. 구현예 209에 있어서, 제2 폴리펩티드의 DBD는 Cas 닉카제 도메인 또는 사멸 Cas 도메인을 포함하는, 시스템.

[0411] 211. 구현예 209에 있어서, 제2 RNA의 gRNA 스페이서는, 예를 들어 표적 핵산 서열의 제2 가닥에서, 주형 핵산의 닉 생성을 유도하는, 시스템.

[0412] 212. 구현예 209에 있어서, 제2 RNA의 gRNA 스페이서는 주형 핵산의 닉 생성을 유도하지 않는, 시스템.

[0413] 213. 구현예 209에 있어서, 제1 gRNA는 제2 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

[0414] 214. 구현예 209에 있어서, 제2 gRNA는 제1 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

[0415] 215. 시스템으로서,

[0416] i) 선행하는 구현예 중 어느 하나의 주형 RNA로서,

[0417] 표적 핵산 서열의 제3 부분에 상보적인 gRNA 스페이서(여기서, 제3 부분은 표적 핵산 서열의 제1 가닥 상에 있음); 및

[0418] gRNA 스캐폴드

[0419] 를 포함하는, 주형 RNA;

[0420] ii) DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0421] DBD에 이종성인 RNA-결합 도메인(RBD)(여기서, RBD는 주형 RNA의 RRS에 결합함)

[0422] 을 포함하는 제1 폴리펩티드;

[0423] iii) 제1 폴리펩티드의 DBD를 표적 핵산 서열의 제2 부분으로 지시하는 gRNA 스페이서(여기서, 표적 핵산 서열의 제2 부분은 핵산 서열의 제2 가닥 상에 있음); 및

[0424] 제1 폴리펩티드의 DBD에 결합하는 gRNA 스캐폴드

[0425] 를 포함하는 제1 gRNA; 및

[0426] iv) RT 도메인, 및

[0427] RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인)(여기서, 제2 폴리펩티드의 DBD는 제1 폴리펩티드의 DBD와 상이한 서열을 갖고, 주형 RNA의 gRNA 스캐폴드는 제2 폴리펩티드의 DBD에 결합함)

[0428] 을 포함하는 제2 폴리펩티드

[0429] 를 포함하는, 시스템.

[0430] 216. 구현예 215에 있어서, 제2 폴리펩티드의 DBD는 Cas 닉카제 도메인 또는 사멸 Cas 도메인을 포함하는, 시스템.

[0431] 217. 구현예 215에 있어서, 주형 RNA의 gRNA 스페이서는, 예를 들어 표적 핵산 서열의 제2 가닥에서, 주형 핵산의 닉 생성을 유도하는, 시스템.

[0432] 218. 구현예 215에 있어서, 주형 RNA의 gRNA 스페이서는 주형 핵산의 닉 생성을 유도하지 않는, 시스템.

[0433] 219. 구현예 215 내지 218 중 어느 하나에 있어서, 제1 gRNA는 제2 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

[0434] 220. 구현예 215 내지 219 중 어느 하나에 있어서, 주형 RNA의 gRNA는 제1 폴리펩티드의 DBD에 검출 가능하게 결합하지 않는, 시스템.

[0435] 221. 폴리펩티드 시스템으로서,

[0436] DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인);

[0437] DBD에 이종성인 RNA-결합 도메인(RBD); 및

[0438] 선택적으로, DBD와 RBD 사이에 배치된 링커

[0439] 를 포함하는 제1 폴리펩티드; 및

[0440] RT 도메인, 및

[0441] RT 도메인에 이종성인 DNA 결합 도메인(DBD)(예를 들어, Cas 도메인, 예를 들어 Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 및

[0442] 선택적으로, RT 도메인과 DBD 사이에 배치된 링커

[0443] 를 포함하는 제2 폴리펩티드

[0444] 를 포함하는, 폴리펩티드 시스템.

[0445] 222. 선행하는 구현예 중 어느 하나에 있어서, 표적 핵산 서열은 표적 유전자, 인핸서, 또는 프로모터인, 주형 RNA 또는 시스템.

[0446] 223. 선행하는 구현예 중 어느 하나에 있어서, 표적 핵산 서열은 인간 표적 유전자, 인간 인핸서, 또는 인간 프로모터인, 주형 RNA 또는 시스템.

[0447] 224. 선행하는 구현예 중 어느 하나에 있어서, RBD는 표 31의 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 갖는, 시스템 또는 폴리펩티드 시스템.

[0448] 225. 세포(예를 들어, 인간 세포)에서 표적 핵산을 변형시키는 방법으로서, 세포를 선행하는 구현예 중 어느 하나의 시스템, 또는 이를 인코딩하는 핵산과 접촉시키는 단계를 포함하며, 이에 의해 표적 핵산을 변형시키는, 방법.

[0449] 226. 구현예 225에 있어서, 제2 폴리펩티드의 존재는 제2 폴리펩티드가 결여된 다른 유사한 시스템과 비교하여

[0450] 표적 핵산의 풀림 증가;

[0451] 변형된 표적 핵산의 수 증가;

[0452] 표적 핵산으로의 삽입 길이 증가; 또는

[0453] 표적 핵산에서 MMR 활성 감소

[0454] 중 하나 이상을 초래하는, 방법.

[0455] 227. 구현예 225 및 226 중 어느 하나에 있어서, 세포는 생체 내 또는 생체 외에 있는, 방법.

[0456] 일 양태에서, 본 개시내용은 (a) 표적 프라이밍된 역전사를 가능하게 하는 유전자 변형 폴리펩티드를 인코딩하는 핵산으로서, 폴리펩티드는 (i) 역전사효소 도메인 및 (ii) DNA에 결합하고 엔도뉴클레아제 활성을 갖는 Cas9 닉카제를 포함하는, 핵산, 및 (b) (i) 인간 유전자의 제1 부분에 상보적인 gRNA 스페이서, (ii) 폴리펩티드에 결합하는 gRNA 스캐폴드, (iii) 돌연변이 영역을 포함하는 이종 목적 서열, 및 (iv) 주형 RNA의 3' 말단에서 표적 DNA 가닥에 대하여 100% 상동성의 적어도 3, 4, 5, 6, 7 또는 8개의 염기를 포함하는 프라이머 결합 부위(PBS) 서열을 포함하는, 주형 RNA를 포함하는, DNA를 변형시키기 위한 시스템에 관한 것이다.

[0457] gRNA 스페이서는 주형 RNA의 5' 말단에서 표적 DNA에 대하여 100% 상동성의 적어도 15개 염기를 포함할 수 있다. 주형 RNA는 표적 DNA 가닥에 대하여 적어도 80% 상동성의 적어도 5개 염기를 포함하는 PBS 서열을 추가로 포함할 수 있다. 주형 RNA는 하나 이상의 화학적 변형을 포함할 수 있다.

[0458] 유전자 변형 폴리펩티드의 도메인은 펩티드 링커에 의해 연결될 수 있다. 폴리펩티드는 하나 이상의 펩티드 링

커를 포함할 수 있다. 유전자 변형 폴리펩티드는 핵 국재화 신호를 추가로 포함할 수 있다. 폴리펩티드는 하나 초과의 핵 국재화 신호, 예를 들어 다중 인접 핵 국재화 신호, 또는 폴리펩티드의 상이한 영역에서의 하나 이상의 핵 국재화 신호, 예를 들어 폴리펩티드의 N-말단에서의 하나 이상의 핵 국재화 신호 및 폴리펩티드의 C-말단에서의 하나 이상의 핵 국재화 신호를 포함할 수 있다. 유전자 변형 폴리펩티드를 인코딩하는 핵산은 하나 초과의 인테인 도메인을 인코딩할 수 있다.

[0459] 표적 세포에 시스템을 도입하면 적어도 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 500, 또는 1000개 염기쌍의 외인성 DNA의 삽입이 초래될 수 있다. 표적 세포에 시스템을 도입하면 결실이 초래될 수 있으며, 여기서 결실은 삽입의 상류 또는 하류에 있는 게놈 DNA의 2, 3, 4, 5, 10, 50, 또는 100개 염기쌍 미만이다. 표적 세포에 시스템을 도입하면 치환, 예를 들어 1, 2, 또는 3개 뉴클레오티드, 예를 들어 연속 뉴클레오티드의 치환이 초래될 수 있다.

[0460] 이종 목적 서열은 적어도 5, 10, 25, 50, 100, 150, 200, 250, 300, 400, 500, 600, 또는 700개 염기쌍일 수 있다.

[0461] 일 양태에서, 본 개시내용은 상기 기재된 시스템 및 약제학적으로 허용가능한 부형제 또는 담체를 포함하는 약제학적 조성물에 관한 것이며, 여기서 약제학적으로 허용가능한 부형제 또는 담체는 플라스미드 벡터, 바이러스 벡터, 소포, 및 지질 나노입자로 이루어진 군으로부터 선택된다. 일 양태에서, 본 개시내용은, 예를 들어 상기 기재된 시스템이 2가지 별개의 부형제 또는 담체, 예를 들어 2개의 지질 나노입자, 2개의 바이러스 벡터, 또는 1개의 지질 나노입자 및 1개의 바이러스 벡터에 의해 전달되는 경우, 상기 기재된 시스템 및 다수의 약제학적으로 허용가능한 부형제 또는 담체를 포함하는 약제학적 조성물에 관한 것이며, 여기서 약제학적으로 허용가능한 부형제 또는 담체는 플라스미드 벡터, 바이러스 벡터, 소포, 및 지질 나노입자로 이루어진 군으로부터 선택된다. 바이러스 벡터는 아데노-연관 바이러스(AAV)일 수 있다.

[0462] 일 양태에서, 본 개시내용은 상기 기재된 시스템을 포함하는 숙주 세포(예를 들어, 포유동물 세포, 예를 들어 인간 세포)에 관한 것이다.

[0463] 시스템은 생체 내, 시험관 내, 생체 외, 또는 제 위치(*in situ*)에서 도입될 수 있다. (a)의 핵산은 숙주 세포의 게놈 내로 통합될 수 있다. 일부 구현예에서, (a)의 핵산은 숙주 세포의 게놈 내로 통합되지 않는다. 일부 구현예에서, 이종 목적 서열은 숙주 세포 게놈의 하나의 표적 부위에만 삽입된다. 이종 목적 서열은 숙주 세포 게놈의 2개 이상의 표적 부위에서, 예를 들어 2개의 상동 염색체의 동일한 대응 부위에서 또는 동일하거나 상이한 염색체의 2개의 상이한 부위에서 삽입될 수 있다. 이종 목적 서열은 포유동물 폴리펩티드, 또는 이의 단편 또는 변이체를 인코딩할 수 있다. 시스템의 구성요소는 1, 2, 3, 4개, 또는 그 이상의 별개의 핵산 분자로 전달될 수 있다. 시스템은 전기천공에 의해, 또는 플라스미드 벡터, 바이러스 벡터, 소포, 및 지질 나노입자로부터 선택되는 적어도 하나의 비히클을 사용함으로써 숙주 세포 내로 도입될 수 있다.

도면의 간단한 설명

[0464] 특허 또는 출원 파일은 컬러로 작성된 적어도 하나의 도면을 포함한다. 컬러 도면(들)이 포함된 본 특허 또는 특허 출원 공개의 사본은 요청 및 필요한 수수료의 지불 시 사무국이 제공할 것이다.

도 1은 예시적인 트랜스 유전자 변형 시스템의 구성요소를 나타내는 일련의 다이어그램이다. 예시적인 시스템은 3가지 구성요소, 즉 (1) 유전자 변형 폴리펩티드, (2) 주형 RNA, 및 (3) gRNA를 포함한다. 유전자 변형 폴리펩티드는 닉카제 Cas9(nCas9), RNA 결합 도메인(RBD), 및 중합효소(이 예에서는 레트로바이러스 역전사효소(RT))를 포함한다. 주형은 RBD 동원 부위(RRS), 프라이머 결합 부위 서열(PBS 서열)(프라이밍) 및 이종 목적 서열(주형 영역)뿐만 아니라, (a) 구조를 엑소뉴클레아제로부터 보호하고/하거나 (b) 2차 구조로 인해 RT를 종료하는 말단 보호/말단 블록 서열을 함유한다. 제3 구성요소는 gRNA이다. 완전히 조립된 트랜스 유전자 변형 반응에서 gRNA는 유전자 변형 폴리펩티드의 nCas9와 결합하고 폴리펩티드를 DNA로 보낸다. 그 다음 nCas9는 DNA에 닉을 도입한다. 폴리펩티드의 RBD는 주형 RNA에서 RRS와의 상호작용을 통해 닉 부위로 주형을 동원한다. Cas9 유도 닉은 주형 RNA의 PBS 서열에 어닐링할 수 있는 3' 플랩(flap)을 생성한다. 그 다음 RT는 말단 보호 구조에 도달할 때까지 주형을 역전사할 수 있다. 고도로 구조화된 말단 보호는 역전사를 종료할 것이다. 세포 복구 과정은 편집된 가닥을 게놈 내로 통합할 것이다.

도 2a 및 2b는 본 명세서에 기재된 바와 같은 트랜스 유전자 변형 시스템에서 사용될 수 있는 예시적인 폴리펩티드를 나타내는 일련의 다이어그램이다. 다음과 같은 nCas9-RT-RBD를 함유하는 폴리펩티드를 조립할 수 있는 여러 가지 방법이 있다: (A) 방향 융합, (B) 각각 공유 또는 비공유적으로 전체 폴리펩티드를 조립하는 인테인

또는 이량체화(동종 또는 이종) 도메인의 사용. (A) 직접 융합 접근법에서, 링커는 nCas9를 RPD와 연결하고, 이는 결국 링커를 통해 RT에 연결된다(예를 들어, 나타낸 바와 같음). 예시적인 가능한 구성이 도 2a 아래 패널에 열거되어 있으며, RBD/링커는 별도의 표에 열거되어 있다. RBP는 한 번 또는 여러 번(예를 들어, n=1 내지 5) 존재할 수 있다. (B) 폴리펩티드는 또한 다양한 인테인 또는 이량체화 도메인을 사용하여 조립될 수 있다. 일부 경우에, nCas9는 이량체화 도메인(FD#1)에 연결되고, RPD는 파트너 이량체화 도메인에 연결된다. nCas9는 제2 이량체화 도메인(FD2)에 연결되는 한편, RT는 파트너에 연결된다. 이량체화 도메인은 공유 결합(예를 들어, 인테인을 사용하는 경우) 또는 폴리펩티드의 비공유 조립(예를 들어, 화학적 또는 광 유도 이량체화를 사용)을 초래할 수 있다. 2가지 이량체화 반응이 이용되며, 이때 폴리펩티드 복합체가 조립된다. 예시적인 가능한 변형이 본 명세서에 기재되어 있다(예를 들어, 인테인 이량체화 도메인, 화학적으로 유도된 이량체화 도메인, 광 유도된 이량체화 도메인, 항체-펩티드 이량체화 도메인, 코일드 코일 이량체화 도메인). 이량체화 도메인은 한 번 또는 여러 번(n=1 내지 30), 예를 들어 탠덤 반복부로 존재할 수 있다.

도 3a 내지 3c는 예시적인 주형 RNA 및 이의 하위영역을 나타내는 일련의 다이어그램이다. (A) 예시적인 주형 RNA의 개략도. 이 주형은 (3'에서 5'으로) 3' 말단에 하나 또는 여러 개(n=1 내지 10)의 RRS, 링커, 그 다음 PBS 서열(프라이밍)(8 내지 17 nt), 그 다음 이종 목적 서열(주형)을 포함한다. 일부 구현예에서 주형 영역은 사전 편집 상동성 영역(0 내지 20 nt), 게놈에 대한 원하는 변형(예를 들어, 삽입, 결실, 또는 점 돌연변이(들))을 갖는 돌연변이 영역, 및 사후 편집 상동성 영역(예를 들어, n=5 내지 500 nt)을 함유한다. 마지막으로, 말단 보호/말단 블록 서열이 주형 RNA의 5' 말단에 존재한다. 예시적인 가능한 구성이 도 3a 아래 패널에 열거되어 있다. (B) 다양한 주형 RNA 구성요소에 대한 예시적인 변형이 열거되어 있다. 그러한 구성요소에 대한 예시적인 서열이 본 명세서에 기재되어 있다. (C) RRS가 사전 편집 상동성 영역과 돌연변이 영역 사이에 위치하는 예시적인 주형 RNA의 개략도.

도 4a 및 4b는 특히, 예를 들어 유전자 변형 효율을 증가시키고/시키거나 긴 삽입부를 허용하기 위한, 표적 핵산의 제2 가닥의 결합 및 조절뿐만 아니라, 표적 핵산의 풀림 증가를 나타내는 일련의 다이어그램이다. 제2 가닥이 트랜스 유전자 변형의 맥락에서 관여할 수 있는 여러 가지 방법이 있다. (A) 한 가지 예시적인 구성에서, 제2 Cas9-gRNA 복합체는 트랜스로 도입될 수 있다. 이러한 제2 Cas9 복합체는, 예를 들어 제2 가닥에 닉을 지시하는 닉카제 Cas9(nCas9)일 수 있다. 이러한 닉은 RT 반응 후 제2 가닥 합성을 개시하고/하거나, 제1 (편집된) 가닥이 유지되고 카피되어야 한다는 신호를 세포 내인성 불일치 복구 시스템에 전달하는 데 사용될 수 있다. 대안적으로, Cas9는, 예를 들어 촉매적으로 비활성인(사멸) Cas9(dCas9)일 수 있다. 이론에 구속되고자 함이 없이, 일부 구현예에서 이는 DNA를 풀고 특히 더 긴 삽입부의 복구를 용이하게 할 수 있다. 이 시나리오에서 Cas9는 트랜스 재작성 폴리펩티드에 존재하는 Cas9와 동일한 종이거나 오르소고날(orthogonal) 종의 것일 수 있다. (B) 대체 구성에서, 제2 가닥 조절은 말단 구조로서 gRNA(전체 또는 부분)를 사용함으로써, 주형 RNA에 의해 동원된다. 이 gRNA는 스캐폴드와 20 nt 스페이서가 있는 전체 gRNA이거나, 스캐폴드와 17개 이하 뉴클레오티드의 스페이서가 있는 부분 gRNA일 수 있다. 전체 gRNA는 폴리펩티드 복합체와 결합할 것이며 폴리펩티드 복합체의 nCas9의 닉을 제2 가닥에 위치시킬 수 있다. 이러한 닉의 배치는 RT 반응 후 제2 가닥 합성을 개시하고/하거나, 제1 (편집된) 가닥이 유지되고 카피되어야 한다는 신호를 세포 내인성 불일치 복구 시스템에 전달하는 데 사용될 수 있다. 스페이서 영역(예를 들어, 17개 뉴클레오티드 이하, 예를 들어 약 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 또는 17개 뉴클레오티드의 길이를 가짐)은 폴리펩티드 복합체의 결합을 야기할 수 있지만, 닉을 생성하지 않을 것이다. 이는 DNA를 풀 것이며 삽입부(예를 들어, 더 긴 삽입부)의 복구를 용이하게 할 수 있다.

도 5a 및 5b는, 예를 들어 유전자 변형 효율을 증가시키고/시키거나 긴 삽입부를 허용하기 위한, 표적 핵산의 제2 가닥의 결합 및 조절을 위한 추가의 예시적인 구성을 나타내는 일련의 다이어그램이다. 이러한 대체 구성에서는, nCas9가 RBD에만 융합된다. nCas9-RBD 폴리펩티드와 연관된 gRNA는 이를 DNA에 동원하고, nCas9는 닉을 도입한다. RBD는 주형 RNA를 동원한다. 구성은 RT 도메인에 융합된 Cas9(예를 들어, 닉카제 또는 사멸 Cas9)로 이루어진 제2 폴리펩티드 복합체를 추가로 포함한다. 이러한 제2 복합체는 (A) 제2 gRNA를 사용하거나, (B) 주형 RNA의 5' 말단에 존재하는 gRNA를 사용하는 방법으로 DNA와 회합할 수 있다. 두 시나리오 모두에서, gRNA는 절단을 지시하기 위한 전체 20 nt 스페이서, 또는 17개 뉴클레오티드 이하(예를 들어, 약 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 또는 17개 뉴클레오티드)의 길이를 갖는 스페이서를 포함하여 닉을 도입하지 않고 DNA를 풀 수 있다.

도 6a는 예시적인 드라이버 구성을 나타내는 다이어그램이다.

도 6b는 예시적인 주형 핵산 구성을 나타내는 다이어그램이다.

도 7a는 세포에서의 재작성 활성을 분석하기 위한 예시적인 분석을 나타내는 다이어그램이다.

도 7b는 표시된 바와 같이 제1의 예시적인 RT 도메인 또는 제2의 RT 도메인을 포함하는 예시적인 유전자 변형 폴리펩티드에 대한 재작성 활성을 나타내는 그래프이다.

도 8은 예시적인 유전자 변형 시스템의 재작성 활성을 나타내는 다이어그램이다.

도 9는 예시적인 유전자 변형 시스템의 재작성 활성을 나타내는 다이어그램이다.

도 10는 예시적인 유전자 변형 시스템에 대한 재작성 활성을 나타내는 일련의 그래프이다.

도 11a 및 11b는 예시적인 유전자 변형 시스템에 대한 재작성 활성을 나타내는 일련의 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0465] *정의*

[0466] 본 명세서에 사용된 바와 같은 용어 "발현 카세트"는 본 발명의 핵산 분자의 발현에 충분한 핵산 요소를 포함하는 핵산 작제물을 지칭한다.

[0467] 본 명세서에 사용된 바와 같은 "gRNA 스페이서"는 표적 핵산에 대해 상보성을 갖고 gRNA 스캐폴드와 함께 Cas 단백질을 표적 핵산에 표적화할 수 있는 핵산의 일부를 지칭한다.

[0468] 본 명세서에 사용된 바와 같은 "gRNA 스캐폴드"는 Cas 단백질에 결합할 수 있고 gRNA 스페이서와 함께 Cas 단백질을 표적 핵산에 표적화할 수 있는 핵산의 일부를 지칭한다. 일부 구현예에서, gRNA 스캐폴드는 crRNA 서열, 테트라루프, 및 tracrRNA 서열을 포함한다.

[0469] 본 명세서에 사용된 바와 같은 "유전자 변형 폴리펩티드"는 (예를 들어, 포유동물 숙주 세포, 예컨대 숙주 세포의 게놈 DNA 분자에서) 핵산 서열(예를 들어, 주형 핵산에 제공된 서열)을 표적 DNA 분자에 통합할 수 있는, 레트로바이러스 역전사효소를 포함하는 폴리펩티드, 또는 레트로바이러스 역전사효소에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 아미노산 서열 동일성을 갖는 아미노산 서열을 포함하는 폴리펩티드를 지칭한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 숙주 기구에 실질적으로 의존하지 않으면서 서열을 통합할 수 있다. 일부 구현예에서, 유전자 변형 폴리펩티드는 서열을 게놈의 무작위 위치에 통합하고, 일부 구현예에서, 유전자 변형 폴리펩티드는 서열을 특정 표적 부위에 통합한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 종합적으로, 1) 주형 핵산의 결합, 2) 표적 DNA 분자의 결합, 및 3) 표적 DNA 내로 주형 핵산의 적어도 일부의 통합을 용이하게 하는 하나 이상의 도메인을 포함한다. 유전자 변형 폴리펩티드는 자연 발생 폴리펩티드뿐만 아니라, 예를 들어 자연 발생 서열에 대해 하나 이상의 아미노산 치환을 갖는, 상기한 조작된 변이체를 둘 다 포함한다. 유전자 변형 폴리펩티드는 또한, 예를 들어 이종성 하위도메인 또는 다른 치환된 도메인의 대체 또는 융합을 통해, 다른 야생형 도메인의 이종성 융합(또는 다른 접합체)뿐만 아니라 변형된 도메인의 융합을 통해서든 간에, 예를 들어 상기 언급된 도메인 중 하나 이상이 서로 이종성인 경우, 이종성 작제물을 포함한다. 본 명세서에 제공된 방법에서 사용될 수 있는 예시적인 유전자 변형 폴리펩티드, 및 이를 포함하는 시스템 및 이를 사용하는 방법은, 예를 들어 PCT/US2021/020948에 기재되어 있으며, 이는 레트로바이러스 역전사효소를 포함하는 유전자 변형 폴리펩티드에 대하여 본 명세서에 참고로 포함된다. 일부 구현예에서, 유전자 변형 폴리펩티드는 서열을 유전자에 통합시킨다. 일부 구현예에서, 유전자 변형 폴리펩티드는 서열을 유전자 외부의 서열에 통합시킨다. 본 명세서에 사용된 바와 같은 "유전자 변형 시스템"은 유전자 변형 폴리펩티드 및 주형 핵산을 포함하는 시스템을 지칭한다.

[0470] 본 명세서에 사용된 바와 같은 용어 "도메인"은 생체분자의 특정된 기능에 기여하는 생체분자의 구조를 지칭한다. 도메인은 생체분자의 인접(contiguous) 영역(예를 들어, 인접 서열) 또는 별개의 비인접 영역(예를 들어, 비인접 서열)을 포함할 수 있다. 단백질 도메인의 예는 엔도뉴클레아제 도메인, DNA 결합 도메인, 역전사 도메인을 포함하지만 이에 제한되지 않고; 핵산의 도메인의 예는 조절 도메인, 예컨대 전사 인자 결합 도메인이다. 일부 구현예에서, 도메인(예를 들어, Cas 도메인)은 2개 이상의 더 작은 도메인(예를 들어, DNA 결합 도메인 및 엔도뉴클레아제 도메인)을 포함할 수 있다.

[0471] 본 명세서에 사용된 바와 같은 용어 "말단 블록 서열"은 역전사를 손상시키고/시키거나 엑소뉴클레아제 활성을 손상시키는 2차 구조를 갖는 RNA 서열을 지칭한다. 일부 경우에, 말단 블록 서열은 스템-루프 서열을 포함한다.

[0472] 본 명세서에 사용된 바와 같은 용어 "외인성"은, 생체분자(예컨대, 핵산 서열 또는 폴리펩티드)와 관련하여 사

용되는 경우, 그 생체분자가 인간의 손에 의해 숙주 게놈, 세포 또는 유기체 내로 도입되었음을 의미한다. 예를 들어, 재조합 DNA 기법 또는 기타 방법을 사용하여 기존 게놈, 세포, 조직 또는 대상체로 첨가됨에 따른 핵산은 기존 핵산 서열, 세포 조직 또는 대상체에 대해 외인성이다.

[0473] 본 명세서에 사용된 바와 같은 "제1 가닥" 및 "제2 가닥"은 표적 DNA의 개별 DNA 가닥을 설명하기 위해 사용되며, 역전사효소 도메인의 중합을 개시하는 가닥, 예를 들어 표적 프라이밍된 합성이 개시된 위치를 기준으로 2개의 DNA 가닥을 구분한다. 제1 가닥은 역전사효소 도메인이 중합을 개시하는, 예를 들어 표적 프라이밍된 합성이 개시되는 표적 DNA의 가닥을 지칭한다. 제2 가닥은 표적 DNA의 다른 가닥을 지칭한다. 제1 및 제2 가닥 명칭은 다른 양태에서는 표적 부위 DNA 가닥을 설명하지 않으며; 예를 들어, 일부 구현예에서, 제1 및 제2 가닥은 본 명세서에 기재된 폴리펩티드에 의해 닉이 생기지만, '제1' 및 '제2' 가닥이라는 명칭은 그러한 닉이 발생하는 순서와는 관계가 없다.

[0474] "게놈 세이프 하버 부위"(GSH 부위)는, 예를 들어 삽입된 유전 요소가 숙주 세포 또는 유기체에 위험을 제기하는 숙주 게놈에서의 현저한 변경을 일으키지 않도록, 새로운 유전 물질의 통합을 수용할 수 있는 숙주 게놈 내 부위이다. GSH 부위는 다음 기준 중 일반적으로 1, 2, 3, 4, 5, 6, 7, 8 또는 9개를 충족한다: (i) 암-관련 유전자로부터 >300kb 떨어져 있음; (ii) miRNA/기타 기능성 소형 RNA로부터 >300kb 떨어져 있음; (iii) 5' 유전자 말단으로부터 >50kb 떨어져 있음; (iv) 복제 원점으로부터 >50kb 떨어져 있음; (v) 초보존된(ultraconservered) 요소로부터 >50kb 떨어져 있음; (vi) 낮은 전사 활성을 가짐(즉, mRNA +/- 25 kb 없음); (vii) 카피 수 가변 영역에 없음; (viii) 개방형 염색질에 존재함; 및/또는 (ix) 인간 게놈에 하나의 카피가 있는 유일한 것임. 이들 기준의 일부 또는 전부를 충족시키는 인간 게놈 내 GSH 부위의 예는, (i) 염색체 19 상의 AAV 바이러스의 통합의 자연 발생 부위인 아데노-연관 바이러스 부위 1(AAVS1); (ii) HIV-1 공동수용체로 알려진 케모카인 수용체 유전자인 케모카인(C-C 모티프) 수용체 5(CCR5) 유전자; (iii) 마우스 Rosa26 유전자좌의 인간 병렬상동체(ortholog); (iv) 리보솜 DNA("rDNA") 유전자자리를 포함한다. 추가의 GSH 부위가 알려져 있고, 예를 들어 문헌[Pellenz et al. epub August 20, 2018 (https://doi.org/10.1101/396390)]에 기재되어 있다.

[0475] 제2의 요소와 관련하여 제1 요소를 설명하기 위해 본 명세서에 사용된 바와 같이 용어 "이종"은, 제1 요소 및 제2 요소가 설명된 바와 같이 배치되어 자연에 존재하지 않음을 의미한다. 예를 들어, 이종 폴리펩티드, 핵산 분자, 작제물 또는 서열은 (a) 그가 발현되는 세포에서 천연이 아닌 폴리펩티드, 핵산 분자, 또는 폴리펩티드 또는 핵산 분자 서열의 일부, (b) 그의 천연 상태에 대해 변경 또는 돌연변이된 폴리펩티드, 핵산 분자, 또는 폴리펩티드 또는 핵산 분자의 일부, 또는 (c) 유사한 조건 하에서 천연 발현 수준과 비교하여 변경된 발현을 갖는 폴리펩티드 또는 핵산 분자를 지칭한다. 예를 들어, 이종 조절 서열(예를 들어, 프로모터, 인핸서)은 유전자 또는 핵산 분자의 발현을, 그 유전자 또는 핵산 분자가 자연에서 보통 발현되는 것과 상이한 방식으로 조절하는 데 사용될 수 있다. 또 다른 예에서, 폴리펩티드 또는 핵산 서열의 이종 도메인(예를 들어, 폴리펩티드의 DNA 결합 도메인 또는 폴리펩티드의 DNA 결합 도메인을 인코딩하는 핵산)은 다른 도메인에 대해 배치될 수 있거나, 폴리펩티드 또는 그의 인코딩 핵산의 다른 도메인 또는 일부에 대해 상이한 서열 또는 상이한 공급원으로부터의 것일 수 있다. 특정 구현예에서, 이종 핵산 분자는 천연 숙주 세포 게놈 내에 존재할 수 있지만, 변경된 발현 수준을 가질 수 있거나 또는 상이한 서열을 가질 수 있거나, 또는 둘 모두 가질 수 있다. 다른 구현예에서, 이종 핵산 분자는 숙주 세포 또는 숙주 게놈에 대해 내인성이 아닐 수 있지만, 대신 형질 전환(예를 들어, 형질주입(transfection), 전기천공)에 의해 숙주 세포 내로 도입되었을 수 있으며, 여기서 추가된 분자는 숙주 게놈 내로 통합될 수 있거나, 일시적으로(예를 들어, mRNA) 또는 한 세대 초과 동안 반안정적으로 염색체외 유전 물질(예를 들어, 에피좀 바이러스 벡터, 플라스미드 또는 기타 자기-복제 벡터)로 존재할 수 있다.

[0476] 본 명세서에 사용된 바와 같이, 표적 부위로의 서열의 "삽입"은, 예를 들어 편집되지 않은 표적 부위에 동족 위치가 없는 이종 목적 서열에 새로운 뉴클레오티드가 있는 경우, 표적 부위에 DNA 서열이 순 추가되는 것을 지칭한다. 일부 구현예에서, 표적 핵산 서열에 대한 PBS 서열 및 이종 목적 서열의 뉴클레오티드 정렬은 표적 핵산 서열에서 정렬 갭을 생성할 것이다.

[0477] 본 명세서에 사용된 바와 같은, 표적 부위에서 이종 목적 서열에 의해 생성된 "결실"은, 예를 들어 이종 목적 서열에 동족 위치가 없는 편집되지 않은 표적 부위에 뉴클레오티드가 있는 경우, 표적 부위에 DNA 서열이 순 결실되는 것을 지칭한다. 일부 구현예에서, 표적 핵산 서열에 대한 PBS 서열 및 이종 목적 서열의 뉴클레오티드 정렬은 PBS 서열 및 이종 목적 서열을 포함하는 분자에서 정렬 갭을 생성할 것이다.

[0478] 본 명세서에 사용된 바와 같은 용어 "역 말단 반복부" 또는 "ITR"은, 이들의 대칭성으로 인해 그렇게 명명된 AAV 바이러스 시스-요소를 지칭한다. 이들 요소는 AAV 게놈의 효율적인 증식을 촉진한다. ITR 기능을 위한 최소

요소는, Rep-결합 부위(RBS; AAV2의 경우 5'-GCGCGCTCGCTCGCTC-3') 및 말단 분해(resolution) 부위(TRS; AAV2의 경우 5'-AGTTGG-3')에 더하여, 헤어핀 형성을 가능하게 하는 가변 회문 서열인 것으로 가정된다. 본 발명에 따르면, ITR은 적어도 이들 3개의 요소(RBS, TRS 및 헤어핀 형성을 가능하게 하는 서열)을 포함한다. 또한, 본 발명에서 용어 "ITR"은 알려진 천연 AAV 혈청형의 ITR(예를 들어, 혈청형 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 또는 11 AAV의 ITR), 상이한 혈청형으로부터 유래된 ITR 요소들의 융합에 의해 형성된 키메라 ITR, 및 이의 기능성 변이체를 지칭한다. "기능성 변이체"는, 알려진 ITR과 적어도 80%, 85%, 90%, 바람직하게는 적어도 95%의 서열 동일성을 나타내어, Rep 단백질의 존재 하에 상기 ITR을 포함하는 서열의 증식을 가능하게 하는 서열을 지칭한다.

[0479] 본 명세서에 사용된 바와 같은 용어 "돌연변이 영역"은 표적 핵산의 상응하는 서열에 비해 하나 이상의 서열 차이를 갖는 주형 RNA의 영역을 지칭한다. 서열 차이는, 예를 들어 치환, 삽입, 프레임시프트, 또는 결실을 포함할 수 있다.

[0480] 용어 "돌연변이된"이 핵산 서열에 적용되는 경우, 이는 핵산 서열 내 뉴클레오티드가 참조(예를 들어, 천연) 핵산 서열과 비교하여 삽입, 결실 또는 변경됨을 의미한다. 유전자좌에서 단일 변경이 이루어질 수 있거나(점 돌연변이), 또는 단일 유전자좌에서 다수의 뉴클레오티드가 삽입, 결실, 또는 변경될 수 있다. 추가로, 핵산 서열 내의 임의의 수의 유전자자리에서 하나 이상의 변경이 이루어질 수 있다. 핵산 서열은 당업계에 알려진 임의의 방법에 의해 돌연변이될 수 있다.

[0481] "핵산 분자"는 제한 없이, 상보적 DNA("cDNA"), 게놈 DNA("gDNA") 및 메신저 RNA("mRNA")를 포함하는 RNA 및 DNA 분자를 모두 지칭하며, 본 명세서에 기재된 바와 같은, RNA 주형과 같이 화학적으로 합성되거나 재조합적으로 생성된 것들과 같은 합성 핵산 분자를 또한 포함한다. 핵산 분자는 이중 가닥 또는 단일 가닥, 환형, 또는 선형일 수 있다. 단일 가닥인 경우, 핵산 분자는 센스 가닥 또는 안티센스 가닥일 수 있다. 달리 나타내지 않는 한, 그리고 일반 형식 "서열번호"로 본 명세서에 기재된 모든 서열에 대한 예로서, "서열번호 1을 포함하는 핵산"은 적어도 일부가 (i) 서열번호 1의 서열, 또는 (ii) 서열번호 1에 상보적인 서열을 갖는 핵산을 지칭한다. 둘 사이의 선택은 서열번호 1이 사용되는 맥락에 따라 결정된다. 예를 들어, 핵산이 프로브로 사용되는 경우, 프로브가 원하는 표적에 상보적이어야 한다는 요건에 따라 둘 사이의 선택이 결정된다. 본 개시내용의 핵산 서열은 화학적 또는 생화학적으로 변형될 수 있거나, 또는 당업자에 의해 쉽게 이해될 것인 바와 같이 비-천연 또는 유도체화된 뉴클레오티드 염기를 함유할 수 있다. 그러한 변형은, 예를 들어 표지, 메틸화, 하나 이상의 자연에 존재하는 뉴클레오티드의 유사체로의 치환, 인터-뉴클레오티드 변형, 예컨대 비하전된 연결(예를 들어, 메틸 포스포네이트, 포스포트리에스테르, 포스포아미데이트, 카르바메이트 등), 하전된 연결(예를 들어, 포스포로티오에이트, 포스포로디티오에이트 등), 현수 모이어티(예를 들어, 폴리펩티드), 인터칼레이터(예를 들어, 아크리딘, 소랄렌 등), 킬레이터, 알킬화제, 및 변형된 연결(예를 들어, 알파 아노머 핵산 등)을 포함한다. 화학적으로 변형된 염기(예를 들어, 표 13 참조), 백본(예를 들어, 표 14 참조), 및 변형된 캡(예를 들어, 표 15 참조)도 포함된다. 수소 결합 및 기타 화학적 상호작용을 통해 지정된 서열에 결합하는 능력 면에서 폴리뉴클레오티드를 모방한 합성 분자도 포함된다. 그러한 분자는 당업계에 알려져 있으며, 예를 들어 펩티드 결합이 분자의 백본에서 포스페이트 연결을 치환한 것들, 예를 들어 펩티드 핵산(PNA)을 포함한다. 기타 변형은, 예를 들어 리보스 고리가 가교 모이어티 또는 "잠금" 핵산(LNA)에서 발견되는 변형과 같은 기타 구조를 함유한 유사체를 포함할 수 있다. 다양한 구현예에서, 핵산은 추가의 유전자 요소, 예컨대 조직-특이적 발현-제어 서열(들)(예를 들어, 조직-특이적 프로모터 및 조직-특이적 마이크로RNA 인식 서열), 및 추가의 요소, 예컨대 역전 반복부(예를 들어, 역 말단 반복부, 예컨대 바이러스, 예를 들어 AAV ITR로부터의 요소 또는 이로부터 유래된 요소) 및 탠덤 반복부(tandem repeat), 역전 반복부/직접 반복부, 상동성 영역(표적 DNA에 대해 다양한 정도의 상동성을 갖는 절편), 비번역 영역(UTR)(5', 3', 또는 5' 및 3' UTR), 및 상기 이들의 다양한 조합과 작동적으로 연관된다. 본 발명에 의해 제공된 시스템의 핵산 요소는 다양한 토폴로지에 제공될 수 있으며, 이는 단일 가닥, 이중 가닥, 환형, 선형, 개방형 말단을 갖는 선형, 폐쇄형 말단을 갖는 선형, 및 이들의 특정 형태, 예컨대 도기본(doggybone) DNA(dbDNA), 폐쇄형 말단 DNA(ceDNA)를 포함한다.

[0482] 본 명세서에 사용된 바와 같은 "유전자 발현 단위"는 적어도 하나의 이펙터 서열에 작동가능하게 연결된 적어도 하나의 조절 핵산 서열을 포함하는 핵산 서열이다. 제1 핵산 서열이 제2 핵산 서열과 기능적 관계에 위치할 때, 제1 핵산 서열은 제2 핵산 서열과 작동가능하게 연결된다. 예를 들어, 프로모터 또는 인핸서가 코딩 서열의 전사 또는 발현에 영향을 미치는 경우, 프로모터 또는 인핸서는 코딩 서열에 작동가능하게 연결된다. 작동가능하게 연결된 DNA 서열은 인접하거나 인접하지 않을 수 있다. 2개의 단백질-코딩 영역을 결합하는 데 필요한 경우, 작동가능하게 연결된 서열은 동일한 판독 프레임에 존재할 수 있다.

[0483] 본 명세서에 사용된 바와 같은 용어 "숙주 게놈" 또는 "숙주 세포"는, 단백질 및/또는 유전 물질이 도입된 세포 및/또는 그의 게놈을 지칭한다. 그러한 용어는 특정한 대상체 세포 및/또는 게놈뿐만 아니라, 그러한 세포의 자손 및/또는 그러한 세포의 자손의 게놈도 지칭하는 것으로 이해되어야 한다. 본 명세서에 사용된 바와 같은. 돌연변이 또는 환경적 영향으로 인해 후속 세대에서 특정 변형이 일어날 수 있기 때문에, 그러한 자손은 실제로 모세포와 동일하지 않을 수 있지만, 본 명세서에 사용된 바와 같은 용어 "숙주 세포"의 범주 내에 여전히 포함된다. 숙주 게놈 또는 숙주 세포는 배양물에서 성장된 단리된 세포 또는 세포주일 수 있거나, 또는 그러한 세포 또는 세포주로부터 단리된 게놈 물질, 또는 살아있는 조직 또는 유기체를 구성하는 숙주 세포 또는 숙주 게놈일 수 있다. 일부 경우에서, 숙주 세포는 예를 들어, 본 명세서에 기재된 바와 같은 동물 세포 또는 식물 세포일 수 있다. 특정 경우에, 숙주 세포는 포유동물 세포, 인간 세포, 조류 세포, 파충류 세포, 소 세포, 말 세포, 돼지 세포, 염소 세포, 양 세포, 닭 세포, 또는 칠면조 세포일 수 있다 특정 경우에, 숙주 세포는 옥수수 세포, 콩 세포, 밀 세포 또는 쌀 세포일 수 있다.

[0484] 본 명세서에 사용된 바와 같은 "작동적 연관"은 1) 프로모터 및 2) 이종 목적 서열과 같은 2개의 핵산 서열 사이의 기능적 관계를 설명하며, 이러한 예에서 프로모터와 이종 목적 서열(예를 들어, 관심 유전자)는, 적합한 조건 하에서 프로모터가 이종 목적 서열의 발현을 구동하도록 배향됨을 의미한다. 예를 들어, 프로모터 및 이종 목적 서열을 보유하는 주형 핵산은 단일 가닥, 예를 들어 (+) 또는 (-) 배향일 수 있다. 이 주형에서 프로모터와 이종 목적 서열 사이의 "작동적 연관"은 주형 핵산이 특정 상태에서 전사될 것인지 여부에 관계없이, 적합한 상태에 있을 때(예를 들어, 필요한 촉매 인자, 및 NPT 등의 존재 하에 (+) 배향에 있음) 정확하게 전사되는 것을 의미한다. 작동적 연관은 다른 조직-특이적 발현 제어 서열(예컨대 인핸서, 억제자(repressor) 및 마이크로RNA 인식 서열), IR/DR, ITR, UTR, 또는 상동성 영역 및 이종 목적 서열 또는 레트로바이러스 RT 도메인을 인코딩하는 서열을 포함하는 핵산의 다른 쌍에 유사하게 적용된다.

[0485] 본 명세서에 사용된 바와 같은 용어 "프라이머 결합 부위 서열" 또는 "PBS 서열"은 표적 핵산 서열에 포함된 영역에 결합할 수 있는 주형 RNA의 일부를 지칭한다. 일부 경우에, PBS 서열은 표적 핵산 서열에 포함된 영역에 대하여 100% 동일성을 갖는 적어도 3, 4, 5, 6, 7, 또는 8개 염기를 포함하는 핵산 서열이다. 일부 구현예에서 프라이머 영역은 표적 핵산 서열에 포함된 영역에 대하여 100% 동일성을 갖는 적어도 5, 6, 7, 8개 염기를 포함한다. 이론에 구속되고자 함이 없이, 일부 구현예에서 주형 RNA가 PBS 서열 및 이종 목적 서열을 포함하는 경우, PBS 서열은 표적 핵산 서열에 포함된 영역에 결합하여, 역전사효소 도메인이 해당 영역을 역전사를 위한 프라이머로서 사용할 수 있게 하고, 이종 목적 서열을 역전사를 위한 주형으로서 사용할 수 있게 한다.

[0486] 본 명세서에 사용된 바와 같은 "스템-루프 서열"은, 예를 들어 적어도 2개(예를 들어, 3, 4, 5, 6, 7, 8, 9, 또는 10개)의 염기 쌍을 포함하는 스템 및 적어도 3개(예를 들어, 4개)의 염기 쌍을 갖는 루프를 갖는 스템-루프를 형성하기에 충분한 자기-상보성을 갖는 핵산 서열(예를 들어, RNA 서열)을 지칭한다. 스템은 미스매치 또는 벌지를 포함할 수 있다.

[0487] 본 명세서에 사용된 바와 같은 "조직-특이적 발현-제어 서열"은 표적 조직에서 이종 목적 서열을 포함하는 전사체의 수준을, 조직-특이적 방식으로, 예를 들어 표적외 조직(들)에 비해 표적상 조직(들)에서 우선적으로, 증가 또는 감소시키는 핵산 요소를 의미한다. 일부 구현예에서, 조직-특이적 발현-제어 서열은 우선적으로 표적 조직에서 이종 목적 서열을 포함하는 전사체의 전사, 활성 또는 반감기를 조직-특이적 방식으로, 예를 들어 표적외 조직(들)에 비해 표적상 조직(들)에서 우선적으로, 구동 또는 억제한다. 예시적인 조직-특이적 발현-제어 서열은 조직-특이적 프로모터, 억제자, 인핸서, 또는 이의 조합, 및 조직-특이적 마이크로RNA 인식 서열을 포함한다. 조직 특이성은 표적상(주형 핵산의 발현 또는 활성이 바람직하거나, 허용가능한 조직(들)) 및 표적외(주형 핵산의 발현 또는 활성이 요구지 않거나, 허용가능하지 않은 조직(들))를 지칭한다. 예를 들어, 조직-특이적 프로모터는, 표적외 조직에 비해 표적상 조직에서 우선적으로 발현을 구동한다. 대조적으로, 조직-특이적 마이크로RNA 인식 서열에 결합하는 마이크로RNA는 표적상 조직에 비해, 표적외 조직에서 우선적으로 발현되어, 표적외 조직에서 주형 핵산의 발현을 감소시킨다. 따라서, 표적 조직과 같은 동일한 조직에 특이적인 프로모터 및 마이크로RNA 인식 서열은 그 조직 내 연관된 서열의 전사, 활성, 또는 반감기와 관련하여 대조적인 기능(각각 일치하는 발현 수준으로 촉진 및 억제함, 즉, 표적외 조직에서는 높은 마이크로RNA 수준이고 표적상 조직에서는 낮은 수준인 한편, 프로모터는 표적상 조직에서는 높은 발현을, 표적외 조직에서는 낮은 발현을 구동함)을 갖는다.

[0488] ***목차***

[0489] 1) 서론

[0490] 2) 유전자 변형 시스템

[0491] a) 유전자 변형 시스템의 폴리펩티드 구성요소

[0492] i) 도메인 작성

[0493] ii) 엔도뉴클레아제 도메인 및 DNA 결합 도메인

[0494] (1) Cas 도메인을 포함하는 유전자 변형 폴리펩티드

[0495] (2) TAL 이펙터 및 징크 핑거 뉴클레아제

[0496] iii) 링커

[0497] iv) 유전자 변형 시스템을 위한 국재화 서열

[0498] v) 유전자 변형 폴리펩티드 및 시스템의 진화된 변이체

[0499] vi) 인테인

[0500] vii) 추가 도메인

[0501] b) 주형 핵산

[0502] i) gRNA 스페이서 및 gRNA 스캐폴드

[0503] ii) 이종 목적 서열

[0504] iii) PBS 서열

[0505] iv) 예시적인 주형 서열

[0506] c) 유도성 활성을 갖는 gRNA

[0507] d) 유전자 변형 시스템의 환형 RNA 및 리보자임

[0508] e) 표적 핵산 부위

[0509] f) 제2 가닥 닉 생성

[0510] 3) 조성물 및 시스템의 생산

[0511] 4) 치료 적용

[0512] 5) 투여 및 전달

[0513] a) 조직 특이적 활성/투여

[0514] i) 프로모터

[0515] ii) 마이크로RNA

[0516] b) 바이러스 벡터 및 이의 구성요소

[0517] c) AAV 투여

[0518] d) 지질 나노입자

[0519] 6) 키트, 제조 물품, 및 약제학적 조성물

[0520] 7) 화학, 제조, 및 제어(CMC)

[0521] ***서론***

[0522] 본 개시내용은 세포, 조직 또는 대상체, 예를 들어 생체 내 또는 시험관 내에서, DNA 서열 내 하나 이상의 위치에서 DNA 서열을 표적화, 편집, 변형 또는 조작(예를 들어, 이종 목적 서열을 포유동물 게놈의 표적 부위 내로 삽입)하기 위한 방법 조성물에 관한 것이다. 이종 목적 DNA 서열은, 예를 들어 치환, 결실, 삽입, 예를 들어 코딩 서열, 조절 서열 또는 유전자 발현 단위를 포함할 수 있다.

[0523] 본 개시내용은 부분적으로 2개 이상의 상호작용에 의해 트랜스 주형 RNA를 유전자 변형 폴리펩티드:sgRNA:표적

게놈 DNA 복합체에 고정시키는 것에 관한 것이다. 이론에 구속되고자 함이 없이, 그러한 앵커링은, 예를 들어 단일 또는 여러 뉴클레오티드 길이의 편집을 달성하기 위한 높은 재작성 활성을 달성할 수 있는 것으로 고려된다. 예를 들어, 1) RRS:RBP 상호작용 및 2) (추가 유전자 변형 폴리펩티드를 통해 매개되는) 표적 DNA 상호작용에 대한 5' 말단 블록 Cas9 스캐폴드 및 스페이서는 트랜스 주형 RNA를 유전자 변형 폴리펩티드:sgRNA:표적 게놈 DNA 복합체에 함께 고정시켜 재작성을 가능하게 하는 예시적인 상호작용을 나타낸다. 5' 말단 블록 스페이서가 없는 경우 RRS:RBP 상호작용이 중요한 것으로 고려된다. 둘 다의 존재는 높은 재작성 활성을 제공할 수 있고 더 약한 RRS:RBP 상호작용과 조합된 5' 말단 블록 스페이서의 존재는 재작성 활성을 구제하는 것으로 추가로 고려된다.

[0524] 본 개시내용은 또한, 예를 들어 관심 서열에/로부터 하나 이상의 뉴클레오티드를 삽입하거나, 결실시키거나, 치환함으로써, 관심 게놈 DNA 서열을 변경시키기 위한 역전사효소-기반 시스템을 사용하여 질환을 치료하는 방법을 제공한다.

[0525] 본 개시내용은 부분적으로 유전자 변형 폴리펩티드 구성요소 및 주형 핵산(예를 들어, 주형 RNA) 구성요소를 포함하는 유전자 변형 시스템을 사용하여 질환을 치료하는 방법을 제공한다. 일부 구현예에서, 유전자 변형 시스템은 게놈 내 표적 부위로 변경을 도입하는 데 사용될 수 있다. 일부 구현예에서, 유전자 변형 폴리펩티드 구성요소는 라이팅 도메인(예를 들어, 역전사효소 도메인), DNA-결합 도메인, 및 엔도뉴클레아제 도메인(예를 들어, 닉카제 도메인)을 포함한다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 게놈 내 표적 부위에 결합하는(예를 들어, 표적 부위의 제2 가닥에 결합하는) 서열(예를 들어, gRNA 스페이서), 유전자 변형 폴리펩티드 구성요소에 결합하는 서열(예를 들어, gRNA 스캐폴드), 이종 목적 서열, 및 PBS 서열을 포함한다. 이론에 구속되고자 함이 없이, 주형 핵산(예를 들어, 주형 RNA)은 게놈 내 표적 부위의 제2 가닥에 결합하고, 유전자 변형 폴리펩티드 구성요소에 결합(예를 들어, 폴리펩티드 구성요소를 게놈 내 표적 부위에 국재화)하는 것으로 생각된다. 유전자 변형 폴리펩티드 구성요소의 엔도뉴클레아제(예를 들어, 닉카제)는, 표적 부위(예를 들어, 표적 부위의 제1 가닥)를 절단하여, 예를 들어, PBS 서열이 표적 부위의 제1 가닥 상에서 변경되는 부위에 인접한 서열에 결합하는 것을 가능하게 하는 것으로 생각된다. 폴리펩티드 구성요소의 라이팅 도메인(예를 들어, 역전사효소 도메인)은 주형 핵산의 PBS 서열을 프라이머로, 그리고 주형 핵산의 이종 목적 서열을 주형으로 포함하는 상보적 서열에 결합되는 표적 부위의 제1 가닥을 사용하여, 예를 들어 이종 목적 서열에 상보적인 서열을 중합하는 것으로 생각된다. 이론에 구속되고자 함이 없이, 적절한 이종 목적 서열의 선택은 표적 부위에서 하나 이상의 뉴클레오티드의 치환, 결실, 및/또는 삽입을 생성할 수 있는 것으로 생각된다.

[0526] ***유전자 변형 시스템***

[0527] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 다음을 포함한다: (A) 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산(여기서, 유전자 변형 폴리펩티드는 (i) 역전사효소 도메인, 및 (x) DNA 결합 기능을 함유하는 엔도뉴클레아제 도메인 또는 (y) 엔도뉴클레아제 도메인 및 별도의 DNA 결합 도메인을 포함함); 및 (B) 주형 RNA. 일부 구현예에서, 유전자 변형 폴리펩티드는 숙주 기구에 실질적으로 의존하지 않으면서, (예를 들어, 포유동물 숙주 세포, 예컨대 숙주 세포의 게놈 DNA 분자에서) 주형 핵산 서열을 표적 DNA 분자에 통합할 수 있는 실질적으로 자율적인 단백질 기계로서 작용한다. 예를 들어, 유전자 변형 단백질은 DNA-결합 도메인, 역전사효소 도메인, 및 엔도뉴클레아제 도메인을 포함할 수 있다. 일부 구현예에서, DNA-결합 기능은 단백질을 DNA 서열로 지시하는 RNA 구성요소, 예를 들어 gRNA 스페이서를 포함할 수 있다. 다른 구현예에서, 유전자 변형 폴리펩티드는 역전사효소 도메인 및 엔도뉴클레아제 도메인을 포함할 수 있다. 유전자 변형 시스템의 RNA 주형 요소는 통상적으로 유전자 변형 폴리펩티드 요소에 대해 이종성이며 숙주 게놈에 삽입될(역전사될) 목적 서열을 제공한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 표적 프라이밍된 역전사가 가능하다. 일부 구현예에서, 유전자 변형 폴리펩티드는 제2 가닥 합성이 가능하다.

[0528] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 아미노산 서열, 또는 이의 기능성 부분, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 DBD의 아미노산 서열, 또는 이에 대하여 적어

도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인, DBD, 및 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다.

[0529] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S1에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 아미노산 서열, 또는 이의 기능성 부분, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S1에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S1에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 DBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S1에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S1에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인, DBD, 및 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다.

[0530] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S2에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 아미노산 서열, 또는 이의 기능성 부분, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S2에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S2에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 DBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S2에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S2에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인, DBD, 및 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다.

[0531] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S3에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 아미노산 서열, 또는 이의 기능성 부분, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S3에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본

명세서에 기재된 유전자 변형 시스템은 표 S3에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 DBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S3에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S3에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인, DBD, 및 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 유전자 변형 폴리펩티드, 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자를 포함한다.

[0532] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S4에 열거된 바와 같은 핵산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함하는 주형 RNA를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S4에 열거된 바와 같은 주형 서열의 5' 말단 블록 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함하는 주형 RNA를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S4에 열거된 바와 같은 주형 서열의 PBS 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함하는 주형 RNA를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S4에 열거된 바와 같은 주형 서열의 링커 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함하는 주형 RNA를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S4에 열거된 바와 같은 주형 서열의 하나 이상(예를 들어, 1, 2, 3, 또는 4개)의 RRS 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함하는 주형 RNA를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S4에 열거된 바와 같은 주형 서열의 3' 말단 블록 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함하는 주형 RNA를 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표 S4에 열거된 바와 같은 주형 서열의 (예를 들어, 5'에서 3' 순서로) 5' 말단 블록 서열, 선택적으로 PBS 서열, 하나 이상(예를 들어, 1, 2, 3, 또는 4개)의 RRS 서열, 및 3' 말단 블록 서열 중 하나 이상(예를 들어, 1, 2, 3, 또는 4개), 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함하는 주형 RNA를 포함한다.

[0533] 일부 구현예에서 유전자 변형 시스템은 제2 폴리펩티드와 조합된다. 일부 구현예에서, 제2 폴리펩티드는 엔도뉴클레아제 도메인을 포함할 수 있다. 일부 구현예에서, 제2 폴리펩티드는 중합효소 도메인, 예를 들어 역전사효소 도메인을 포함할 수 있다. 일부 구현예에서, 제2 폴리펩티드는 DNA-의존성 DNA 중합효소 도메인을 포함할 수 있다. 일부 구현예에서, 제2 폴리펩티드는 예를 들어 제2 가닥 합성 또는 DNA 복구 해결에 기여함으로써 게놈 편집의 완료를 돕는다.

[0534] 기능성 유전자 변형 폴리펩티드는 관련없는 DNA 결합, 역전사, 및 엔도뉴클레아제 도메인으로 구성될 수 있다. 이러한 모듈 구조는 기능성 도메인들, 예를 들어 dCas9(DNA 결합), MMLV 역전사효소(역전사), FokI(엔도뉴클레아제)의 조합을 가능하게 한다. 일부 구현예에서, 다수의 기능성 도메인은 단일 단백질, 예를 들어 Cas9 또는 Cas9 닉카제(DNA 결합, 엔도뉴클레아제)로부터 발생할 수 있다.

[0535] 일부 구현예에서, 유전자 변형 폴리펩티드는 종합적으로, 1) 주형 핵산의 결합, 2) 표적 DNA 분자의 결합, 및 3) 표적 DNA 내로 주형 핵산의 적어도 일부의 통합을 용이하게 하는 하나 이상의 도메인을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 상응하는 자연 발생 서열에 대한 하나 이상의 아미노산 치환을 포함하는 조작된 폴리펩티드이다. 일부 구현예에서, 유전자 변형 폴리펩티드는, 예를 들어 이종성 하위도메인 또는 다른 치환된 도메인의 대체 또는 융합을 통해, 예를 들어 다른 야생형 도메인의 이종성 융합(또는 다른 접합체)뿐만 아니라 변형된 도메인의 융합을 통해, 서로에 대해 이종성인 2개 이상의 도메인을 포함한다. 예를 들어, 일부 구현예에서 다음 중 하나 이상이다: RT 도메인이 DBD에 대해 이종성이거나; DBD가 엔도뉴클레아제 도메인에 대해 이종성이거나; RT 도메인이 엔도뉴클레아제 도메인에 대해 이종성임.

[0536] 일부 구현예에서, 시스템에서 사용하기 위한 주형 RNA 분자는 5'에서 3'으로 (1) gRNA 스페이서; (2) gRNA 스캐폴드; (3) 이종 목적 서열 (4) 프라이머 결합 부위(PBS) 서열을 포함한다. 일부 구현예에서:

[0537] (1) gRNA 스페이서는 약 18 내지 22 nt, 예를 들어 20 nt의 gRNA 스페이서이다.

[0538] (2) gRNA 스캐폴드는 주형을 Cas 도메인, 예를 들어 닉카제 Cas9 도메인과 회합시키기 위한 하나 이상의 헤어핀 루프, 예를 들어 1, 2, 또는 3개의 루프를 포함하는 gRNA 스캐폴드이다. 일부 구현예에서, gRNA 스캐폴드는 5'에서 3'으로 서열 GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC(서열번호 8)를 포함한다.

[0539] (3) 일부 구현예에서, 이종 목적 서열은, 예를 들어 길이가 7 내지 74, 예를 들어 10 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 또는 70 내지 80 nt 또는, 80 내지 90 nt이다. 일부 구현예에서, 서열의 첫 번째(가장 5') 염기는 C가 아니다.

[0540] (4) 일부 구현예에서, 닉 생성이 발생한 후 표적 프라이밍 서열에 결합하는 PBS 서열은, 예를 들어 3 내지 20 nt, 예를 들어 7 내지 15 nt, 예를 들어 12 내지 14 nt이다. 일부 구현예에서, PBS 서열은 40 내지 60% GC 함량을 갖는다.

[0541] 일부 구현예에서, 시스템과 회합된 제2 gRNA는 완전한 통합의 구동을 도울 수 있다. 일부 구현예에서, 제2 gRNA는 제1 가닥 닉으로부터 0 내지 200 nt 떨어진 위치, 예를 들어 제1 가닥 닉으로부터 0 내지 50, 50 내지 100, 100 내지 200 nt 떨어진 위치를 표적화할 수 있다. 일부 구현예에서, 제2 gRNA는 편집이 이루어진 후 그의 표적 서열에만 결합할 수 있으며, 예를 들어 gRNA는 이종 목적 서열에는 존재하지만, 최초 표적 서열에는 존재하지 않는 서열에 결합한다.

[0542] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 HEK293, K562, U2OS, 또는 HeLa 세포에서 편집을 수행하는 데 사용된다. 일부 구현예에서, 유전자 변형 시스템은 1차 세포, 예를 들어 E18.5 마우스 유래의 1차 피질 뉴런에서 편집을 수행하는 데 사용된다.

[0543] 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드는 MoMLV RT 서열 또는 이의 변이체를 포함하는 (예를 들어, 본 명세서에 기재된 바와 같은) 역전사효소 또는 RT 도메인을 포함한다. 구현예에서, MoMLV RT 서열은 D200N, L603W, T330P, T306K, W313F, D524G, E562Q, D583N, P51L, S67R, E67K, T197A, H204R, E302K, F309N, L435G, N454K, H594Q, D653N, R110S, 및 K103L로부터 선택되는 하나 이상의 돌연변이를 포함한다. 구현예에서, MoMLV RT 서열은 선택적으로 T306K 및/또는 W313F를 추가로 포함하는 D200N, L603W, 및 T330P와 같은 돌연변이의 조합을 포함한다.

[0544] 일부 구현예에서, (예를 들어, 본 명세서에 기재된 바와 같은) 엔도뉴클레아제 도메인은, 예를 들어 H840A 돌연변이를 포함하는 nCAS9를 포함한다.

[0545] 일부 구현예에서, (예를 들어, 본 명세서에 기재된 바와 같은 시스템의) 이종 목적 서열은 길이가 약 1 내지 50, 50 내지 100, 100 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 900 내지 1000개, 또는 그 이상의 뉴클레오티드이다.

[0546] 일부 구현예에서, RT 및 엔도뉴클레아제 도메인은, 예를 들어 아미노산 서열 SGGSSGGSSGSETPGTSESATPESSGGSSGGSS(서열번호 6)를 포함하는 가요성 링커에 의해 연결된다.

[0547] 일부 구현예에서, 엔도뉴클레아제 도메인은 RT 도메인에 대해 N-말단이다. 일부 구현예에서, 엔도뉴클레아제 도메인은 RT 도메인에 대해 C-말단이다.

[0548] 일부 구현예에서, 시스템은, 예를 들어 본 명세서에 기재된 바와 같은, TPRT에 의해 이종 목적 서열을 표적 부위로 포함시킨다.

[0549] 일부 구현예에서, 유전자 변형 폴리펩티드는 DNA 결합 도메인을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 RNA 결합 도메인을 포함한다. 일부 구현예에서, RNA 결합 도메인은 B-박스 단백질, MS2 코트 단백질, dCas, 또는 본 명세서 표의 서열 요소의 RNA 결합 도메인을 포함한다. 일부 구현예에서, RNA 결합 도메인은 참조 RNA 결합 도메인보다 더 큰 친화도로 주형 RNA에 결합할 수 있다.

[0550] 일부 구현예에서, 유전자 변형 시스템은 적어도 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 또는 100개의 뉴클레오티드(및 선택적으로 500, 400, 300, 200, 또는 100개 이하의 뉴클레오티드)의, 표적 부위 내로의 삽입을 생성할 수 있다. 일부 구현예에서, 유전자 변형 시스템은 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 또는 100개의 뉴클레오티드(및 선택적으로 500, 400, 300, 200, 또는 100개 이하의 뉴클레오티드)의 표적 부위 내로의 삽입을 생성할 수 있다. 일부 구현

예에서, 유전자 변형 시스템은 적어도 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5 또는 10 킬로베이스(및 선택적으로 1, 5, 10, 또는 20 이하의 킬로베이스)의 표적 부위 내로의 삽입을 생성할 수 있다. 일부 구현예에서, 유전자 변형 시스템은 적어도 81, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 또는 200개의 뉴클레오티드(및 선택적으로 500, 400, 300, 또는 200개 이하의 뉴클레오티드)의 결실을 생성할 수 있다. 일부 구현예에서, 유전자 변형 시스템은 적어도 81, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 또는 200개의 뉴클레오티드(및 선택적으로 500, 400, 300, 또는 200개 이하의 뉴클레오티드)의 결실을 생성할 수 있다. 일부 구현예에서, 유전자 변형 시스템은 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 또는 200개의 뉴클레오티드(및 선택적으로 500, 400, 300, 또는 200개 이하의 뉴클레오티드)의 결실을 생성할 수 있다. 일부 구현예에서, 유전자 변형 시스템은 적어도 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5 또는 10 킬로베이스(및 선택적으로 1, 5, 10, 또는 20 이하의 킬로베이스)의 결실을 생성할 수 있다. 일부 구현예에서, 유전자 변형 시스템은 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 또는 100개 또는 그 이상의 뉴클레오티드의 표적 부위 내로의 치환을 생성할 수 있다. 일부 구현예에서, 유전자 변형 시스템은 1 내지 2, 2 내지 3, 3 내지 4, 4 내지 5, 5 내지 10, 10 내지 15, 15 내지 20, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 60, 60 내지 70, 70 내지 80, 80 내지 90, 또는 90 내지 100개 뉴클레오티드의 표적 부위에서의 치환을 생성할 수 있다.

[0551] 일부 구현예에서, 치환은 전위(transition) 돌연변이이다. 일부 구현예에서, 치환은 전환 돌연변이이다. 일부 구현예에서, 치환은 아데닌을 티민으로, 아데닌을 구아닌으로, 아데닌을 시토신으로, 구아닌을 티민으로, 구아닌을 시토신으로, 구아닌을 아데닌으로, 티민을 시토신으로, 티민을 아데닌으로, 티민을 구아닌으로, 시토신을 아데닌으로, 시토신을 구아닌으로, 또는 시토신을 티민으로 전환한다.

[0552] 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 유전자의 발현(예를 들어, 전사 또는 번역)을 증가시키거나 감소시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 프로모터 또는 인핸서, 예를 들어 전사 인자에 결합하는 서열에서 서열을 변경, 추가, 또는 결실시킴으로써 유전자의 발현(예를 들어, 전사 또는 번역)을 증가시키거나 감소시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 유전자의 번역을 변경시키고(예를 들어, 아미노산 서열을 변경함), 시작 또는 정지 코돈을 삽입 또는 결실시키고, 유전자의 번역 프레임을 변경 또는 고정시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은, 예를 들어 스플라이스 어셉터 수여 또는 공여 부위를 삽입, 결실, 또는 변경시킴으로써, 유전자의 스플라이싱을 변경시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 전사체 또는 단백질 반감기를 변경시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 (예를 들어, 세포질에서 미토콘드리아로, 세포질에서 세포외 공간으로(예를 들어, 분비 태그를 추가함)) 세포 내 단백질 국재화를 변경시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 (예를 들어, 잘못 접힌 단백질의 축적을 방지하기 위해) 단백질 접힘을 변경시킨다(예를 들어, 개선시킴). 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 유전자, 예를 들어 유전자에 의해 인코딩된 단백질의 활성을 변경, 증가, 감소시킨다.

[0553] 예시적인 유전자 변형 폴리펩티드, 및 이를 포함하는 시스템 및 이를 사용하는 방법은, 예를 들어 PCT/US2021/020948에 기재되어 있으며, 이는 그 안의 아미노산 및 핵산 서열을 포함하여 레트로바이러스 RT 도메인에 관하여 본 명세서에 참고로 포함된다.

[0554] 예시적인 유전자 변형 폴리펩티드 및 레트로바이러스 RT 도메인 서열은 또한, 예를 들어 2021년 3월 4일에 출원된 국제 출원 PCT/US21/20948, 예를 들어 그 안의 표 30, 표 31, 및 표 44에 기재되어 있으며; 전체 출원은, 예를 들어 상기 서열 및 표의 레트로바이러스 RT에 관하여 본 명세서에 참고로 포함된다. 따라서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 본 단락에 언급된 표 중 임의의 것에 따른 아미노산 서열, 또는 이의 도메인(예를 들어, 레트로바이러스 RT 도메인), 또는 상기한 것 중 임의의 것의 기능성 단편 또는 변이체, 또는 이에 대하여 적어도 70%, 80%, 85%, 90%, 95%, 또는 99% 동일성을 갖는 아미노산 서열을 포함할 수 있다.

[0555] 일부 구현예에서, 본 명세서에 기재된 시스템 중 임의의 것에서 사용하기 위한 폴리펩티드는 다중 상동성 단백질의 정렬된 폴리펩티드 서열에 기반한 분자 재구성 또는 조상(ancestral) 재구성일 수 있다. 일부 구현예에서, 본 명세서에 기재된 시스템 중 임의의 것에서 사용하기 위한 역전사효소 도메인은 동일하거나 상이한 공급원 유래의 역전사효소 도메인의 정렬에 기반한 분자 재구성 또는 조상 재구성일 수 있거나 특정 잔기에서 변형될 수 있다. 당업자는, 본 명세서에 제공된 수탁 번호를 기반으로, 예를 들어 BLAST(Basic Local Alignment Search Tool) 또는 보존된 도메인 분석을 위한 CD-검색과 같은 일상적인 서열 분석 도구를 사용함으로써, 폴리펩티드

또는 핵산 서열을 정렬할 수 있다. 분자 재구성은, 예를 들어 문헌[Ivics et al., Cell 1997, 501 - 510 ; Wagstaff et al., Molecular Biology and Evolution 2013, 88-99]에 기재된 접근법을 사용하여, 서열 공통성에 기반하여 창출될 수 있다.

[0556] *유전자 변형 시스템의 폴리펩티드 구성요소*

[0557] 일부 구현예에서, 유전자 변형 폴리펩티드는 DNA 표적 부위 결합, 주형 핵산(예를 들어, RNA) 결합, DNA 표적 부위 절단, 및 주형 핵산(예를 들어, RNA) 라이팅, 예를 들어, 역전사의 기능을 갖는다. 일부 구현예에서, 각각의 기능은 별개의 도메인 내에 포함된다. 일부 구현예에서, 기능은 둘 이상의 도메인(예를 들어 둘 이상의 도메인이 함께, 기능성을 나타냄)에 기인할 수 있다. 일부 구현예에서, 둘 이상의 도메인은 동일하거나 유사한 기능을 가질 수 있다(예를 들어, 둘 이상의 도메인은, 예를 들어 2개의 상이한 DNA 서열에 대해, 각각 독립적으로 DNA-결합 기능성을 가짐). 다른 구현예에서, 하나 이상의 도메인은 하나 이상의 기능을 가능하게 할 수 있으며, 예를 들어 Cas9 도메인은 DNA 결합 및 표적 부위 절단을 모두 가능하게 한다. 일부 구현예에서, 도메인은 모두 단일 폴리펩티드 내에 위치한다. 일부 구현예에서, 제1 도메인은 하나의 폴리펩티드에 있고, 제2 도메인은 제2 폴리펩티드에 있다. 예를 들어, 일부 구현예에서, 서열은 제1 폴리펩티드 및 제2 폴리펩티드 사이에서 분할될 수 있고, 예를 들어 여기서 제1 폴리펩티드는 역전사효소(RT) 도메인을 포함하고, 제2 폴리펩티드는 DNA-결합 도메인 및 엔도뉴클레아제 도메인, 예를 들어 닉카제 도메인을 포함한다. 추가의 예로서, 일부 구현예에서, 제1 폴리펩티드 및 제2 폴리펩티드는 각각 DNA 결합 도메인(예를 들어, 제1 DNA 결합 도메인 및 제2 DNA 결합 도메인)을 포함한다. 일부 구현예에서, 제1 및 제2 폴리펩티드는 스플릿(split)-인테인을 통해 번역 후 함께 합쳐져서 단일 유전자 변형 폴리펩티드를 형성할 수 있다.

[0558] 일부 양태에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 다음을 포함한다(예를 들어, 본 명세서에 기재된 시스템은 다음을 포함하는 유전자 변형 폴리펩티드를 포함함): 1) Cas 도메인(예를 들어, Cas 닉카제 도메인, 예를 들어 Cas9 닉카제 도메인); 2) 표 1의 역전사효소(RT) 도메인, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 또는 99% 동일성을 갖는 서열(여기서, RT 도메인은 Cas 도메인의 C-말단임); RT 도메인과 Cas 도메인 사이에 배치된 링커(여기서, 링커는 표 1의 RT 도메인과 동일한 행의 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 가짐).

[0559] 일부 구현예에서, RT 도메인은 표 1의 RT 도메인과 100% 동일성을 갖는 서열을 갖고 링커는 표 1의 RT 도메인과 동일한 행의 링커 서열에 대하여 100% 동일성을 갖는 서열을 갖는다. 일부 구현예에서, Cas 도메인은 표 8의 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 서열 목록의 서열번호 1 내지 3332 중 임의의 것에 따른 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함한다.

[0560] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 아미노산 서열, 또는 이의 기능성 부분, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 DBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S1 내지 S3 중 임의의 것에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인, DBD, 및 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다.

[0561] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S1에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 아미노산 서열, 또는 이의 기능성 부분, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S1에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인의 아미노산

서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S1에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 DBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S1에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S1에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인, DBD, 및 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다.

[0562] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S2에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 아미노산 서열, 또는 이의 기능성 부분, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S2에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S2에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 DBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S2에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S2에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인, DBD, 및 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다.

[0563] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S3에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 아미노산 서열, 또는 이의 기능성 부분, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S3에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S3에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 DBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S3에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 S3에 열거된 바와 같은 예시적인 유전자 변형 폴리펩티드의 RT 도메인, DBD, 및 RBD의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다.

[0564] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 DBD, RT 도메인, 및 하나 이상의 RBD(예를 들어, 본 명세서에 기재된 바와 같음)를 포함한다.

[0565] 특정 구현예에서, 유전자 변형 폴리펩티드는 N-말단에서 C-말단 순서로, DBD(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어 Cas 도메인, 예를 들어 Cas9 도메인), 하나 이상(예를 들어, 1, 2, 3, 또는 4개)의 RBD, 및 RT 도메인을 포함한다. 구현예에서, DBD 및 N-말단 RBD는 링커(예를 들어, 본 명세서에 기재된 바와 같음)에 의해 연결된다. 구현예에서, C-말단 RBD 및 RT 도메인은 링커(예를 들어, 본 명세서에 기재된 바와 같음)에 의해 연결된다.

[0566] 특정 구현예에서, 유전자 변형 폴리펩티드는 N-말단에서 C-말단 순서로, RT 도메인, 하나 이상(예를 들어, 1, 2, 3, 또는 4개)의 RBD, 및 DBD(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어 Cas 도메인, 예를 들어 Cas9 도메인)를 포함한다. 구현예에서, RT 도메인 및 N-말단 RBD는 링커(예를 들어, 본 명세서에 기재된 바와 같음)에 의해 연결된다. 구현예에서, C-말단 RBD 및 DBD는 링커(예를 들어, 본 명세서에 기재된 바와 같음)에 의해 연결된다.

[0567] 특정 구현예에서, 유전자 변형 폴리펩티드는 N-말단에서 C-말단 순서로, DBD(예를 들어, 본 명세서에 기재된 바

와 같은, 예를 들어 Cas 도메인, 예를 들어 Cas9 도메인), RT 도메인, 및 하나 이상(예를 들어, 1, 2, 3, 또는 4개)의 RBD를 포함한다. 구현예에서, DBD 및 RT 도메인은 링커(예를 들어, 본 명세서에 기재된 바와 같음)에 의해 연결된다. 구현예에서, RT 도메인 및 N-말단 RBD는 링커(예를 들어, 본 명세서에 기재된 바와 같음)에 의해 연결된다.

[0568] 일부 구현예에서, 유전자 변형 폴리펩티드는 N-말단 메티오닌 잔기를 포함한다.

[0569] 일부 구현예에서, 유전자 변형 폴리펩티드는, 예를 들어 본 명세서에 기재된 바와 같은, 하나 이상의 핵 국재화 서열(NLS)을 포함한다.

[0570] 일부 구현예에서, 유전자 변형 폴리펩티드는 Cas 도메인과 링커 사이에 GC 아미노산 서열, RT 도메인과 제2 NLS 사이에 AG 아미노산 서열, 및/또는 링커와 RT 도메인 사이에 GG 아미노산 서열을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 제1 NLS 및 Cas 도메인을 포함하는 서열번호 4000의 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 제2 NLS를 포함하는 서열번호 4001의 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 98%, 또는 99% 동일성을 갖는 서열을 포함한다.

[0571] **예시적인 N-말단 NLS-Cas9 도메인**

[0572] MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF
DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHP
IFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALS
LGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVN
TEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY
KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR
EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN
LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK
EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDR
EMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF
MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPE
NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQ
LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKL
IREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY
KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY
SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEI
IEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR
KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGG(**서열번호 4000**)

[0573] **NLS를 포함하는 예시적인 C-말단 서열**

[0574] AGKRTADGSEFEKRTADGSEFESPKKKAKVE(**서열번호 4001**)

[0575] *유전자 변형 도메인(RT 도메인)*

[0576] 본 발명의 특정 양태에서, 유전자 변형 시스템의 유전자 변형 도메인은 역전사효소 활성을 가지며, 역전사효소 도메인(RT 도메인)으로도 지칭된다. 일부 구현예에서, RT 도메인은 RT 촉매부 및 RNA-결합 영역(예를 들어, 주형 RNA에 결합하는 영역)을 포함한다.

[0577] 일부 구현예에서, 역전사효소를 인코딩하는 핵산은, 예를 들어 인간 세포에 대해 개선된, 변경된 코돈 사용을 갖기 위해 그의 천연 서열에서 변경된다. 일부 구현예에서 역전사효소 도메인은 레트로바이러스 유래의 이종성 역전사효소이다. 일부 구현예에서, 유전자 변형 폴리펩티드를 포함하는 RT 도메인은 원래의 아미노산 서열로부터 돌연변이되었으며, 예를 들어 적어도 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 또는 100개 치환을 갖는다. 일부 구현예에서, RT 도메인은 레트로바이러스의 RT, 예를 들어 HIV-1 RT, 몰로니 뮤린 백혈병 바이러스(MMLV) RT, 조류 골수아구증 바이러스(AMV) RT, 또는 라우스 육종 바이러스(RSV) RT로부터

유래된다.

[0578] 일부 구현예에서, 레트로바이러스 역전사효소(RT) 도메인은, 예를 들어 내인성 RT 도메인에 비해 표적-프라이밍된 역전사(TPRT) 개시의 향상된 엄격성을 나타낸다. 일부 구현예에서, RT 도메인은, 제1 가닥 닉의 바로 상류의 표적 부위에서 3 nt, 예를 들어 RNA 주형을 프라이밍하는 게놈 DNA가 RNA 주형에서 3 nt의 상동성에 대해 적어도 66% 또는 100% 상보성을 갖는 경우, TPRT를 개시한다. 일부 구현예에서, RT 도메인은, 주형 RNA 상동성과 표적 DNA 프라이밍 역전사 사이에 5 nt 미만의 미스매치(예를 들어, 1, 2, 3, 4, 또는 5 nt 미만의 미스매치)가 존재하는 경우 TPRT를 개시한다. 일부 구현예에서, RT 도메인은 TPRT 반응을 프라이밍하는 데 있어 미스매치에 대한 엄격성이 증가하도록 변형되며, 예를 들어 여기서 RT 도메인은 야생형(예를 들어, 비변형된) RT 도메인에 비해 프라이밍 영역에서 임의의 미스매치를 용인하지 않거나, 더 적은 미스매치를 용인한다. 일부 구현예에서, RT 도메인은 HIV-1 RT 도메인을 포함한다. 구현예에서, HIV-1 RT 도메인은 대안적인 RT 도메인에 비해 3개의 뉴클레오티드 미스매치가 있는 경우에도 더 낮은 수준의 합성을 개시한다(예를 들어, 문헌[Jamburuthugoda and Eickbush J Mol Biol 407(5):661-672 (2011)]에 기재된 바와 같으며; 그 전체가 본 명세서에 참고로 포함됨). 일부 구현예에서, RT 도메인은 이량체(예를 들어, 이종이량체 또는 동종이량체)를 형성한다. 일부 구현예에서, RT 도메인은 단량체성이다. 일부 구현예에서, RT 도메인은 단량체 또는 이량체(예를 들어, 이종이량체 또는 동종이량체)로서 자연적으로 기능한다. 일부 구현예에서, RT 도메인은 단량체로서 자연적으로 기능하며, 예를 들어 바이러스로부터 유래되고, 여기서 이는 단량체로서 기능한다. 구현예에서, RT 도메인은 뮤린 백혈병 바이러스(MLV; 때때로 MoMLV로 지칭됨)(예를 들어, P03355), 돼지 내인성 레트로바이러스(PERV)(예를 들어, UniProt Q4VFZ2), 마우스 유방 종양 바이러스(MMTV)(예를 들어, UniProt P03365), Mason-Pfizer 원숭이 바이러스(MPMV)(예를 들어, UniProt P07572), 소 백혈병 바이러스(BLV)(예를 들어, UniProt P03361), 인간 T-세포 백혈병 바이러스-1(HTLV-1)(예를 들어, UniProt P03362), 인간 포미(foamy) 바이러스(HFV)(예를 들어, UniProt P14350), 원숭이 포미 바이러스(SFV)(예를 들어, UniProt P23074), 또는 소 포미/세포융합 바이러스(BFV/BSV)(예를 들어, UniProt O41894), 또는 이의 기능성 단편 또는 변이체(예를 들어, 이와 적어도 70%, 80%, 90%, 95%, 또는 99% 동일성을 갖는 아미노산 서열)로부터의 RT 도메인으로부터 선택된다. 일부 구현예에서, RT 도메인은 그의 천연 기능에서 이량체성이다. 일부 구현예에서, RT 도메인은 바이러스로부터 유래되며, 여기서 이는 이량체로서 기능한다. 구현예에서, RT 도메인은 조류 육종/백혈병, 바이러스(ASLV)(예를 들어, UniProt A0A142BKH1), 라우스 육종 바이러스(RSV)(예를 들어, UniProt P03354), 조류 골수아구증 바이러스(AMV)(예를 들어, UniProt Q83133), 인간 면역결핍 바이러스 I형(HIV-1)(예를 들어, UniProt P03369), 인간 면역결핍 바이러스 II형(HIV-2)(예를 들어, UniProt P15833), 원숭이 면역결핍 바이러스(SIV)(예를 들어, UniProt P05896), 소 면역결핍 바이러스(BIV)(예를 들어, UniProt P19560), 말 전염성 빈혈 바이러스(EIAV)(예를 들어, UniProt P03371), 또는 고양이 면역결핍 바이러스(FIV)(예를 들어, UniProt P16088)(Herschhorn and Hizi Cell Mol Life Sci 67(16):2717-2747 (2010)), 또는 이의 기능성 단편 또는 변이체(예를 들어, 이와 적어도 70%, 80%, 90%, 95%, 또는 99% 동일성을 갖는 아미노산 서열)로부터의 RT 도메인으로부터 선택된다. 천연 이종이량체성 RT 도메인은 또한 일부 구현예에서 동종이량체로서 기능할 수 있다. 일부 구현예에서, 이량체성 RT 도메인은 융합 단백질, 예를 들어 동종이량체성 융합 단백질 또는 이종이량체성 융합 단백질로서 발현된다. 일부 구현예에서, 시스템의 RT 기능은 (예를 들어, 본 명세서에 기재된 바와 같은) 다수의 RT 도메인에 의해 수행된다. 추가의 구현예들에서, 다수의 RT 도메인은 융합 또는 분리되어, 예를 들어 동일한 폴리펩티드 상에 또는 상이한 폴리펩티드 상에 있을 수 있다.

[0579] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 인테그라제(integrase) 도메인을 포함하고, 예를 들어 여기서 인테그라제 도메인은 RT 도메인의 일부일 수 있다. 일부 구현예에서, (예를 들어, 본 명세서에 기재된 바와 같은) RT 도메인은 인테그라제 도메인을 포함한다. 일부 구현예에서, (예를 들어, 본 명세서에 기재된 바와 같은) RT 도메인은 인테그라제 도메인을 결여하거나, 또는 돌연변이에 의해 비활성화되거나 결실된 인테그라제 도메인을 포함한다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 RNase H 도메인을 포함하고, 예를 들어 여기서 RNase H 도메인은 RT 도메인의 일부일 수 있다. 일부 구현예에서, RNase H 도메인은 RT 도메인의 일부가 아니고 가요성 링커를 통해 공유결합적으로 연결된다. 일부 구현예에서, (예를 들어, 본 명세서에 기재된 바와 같은) RT 도메인은 RNase H 도메인, 예를 들어 내인성 RNase H 도메인 또는 이종 RNase H 도메인을 포함한다. 일부 구현예에서, (예를 들어, 본 명세서에 기재된 바와 같은) RT 도메인은 RNase H 도메인이 결여되어 있다. 일부 구현예에서, (예를 들어, 본 명세서에 기재된 바와 같은) RT 도메인은 이종 RNase H 도메인에 대해 추가, 결실, 돌연변이 또는 교환된 RNase H 도메인을 포함한다. 일부 구현예에서, 폴리펩티드는 비활성화된 내인성 RNase H 도메인을 포함한다. 일부 구현예에서, 폴리펩티드의 다른 도메인 중 하나로부터의 내인성 RNase H 도메인은 폴리펩티드에 포함되지 않도록 유전적으로 제거되며, 예를 들어 내인성 RNase H 도메인

은 포함하는 도메인으로부터 부분적으로 또는 완전히 절단된다. 일부 구현예에서, RNase H 도메인의 돌연변이는 일부 구현예에서, RNase H 도메인의 돌연변이는, 문헌[Kotewicz et al. Nucleic Acids Res 16(1):265-277 (1988)(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 방법에 의해 결정되는 바와 같은, 예를 들어 돌연변이가 없는 다른 유사한 도메인과 비교하여 적어도 10%, 20%, 30%, 40%, 50%, 60%, 70%, 80%, 또는 90%만큼 더 낮은 RNase 활성을 나타내는 폴리펩티드를 산출한다. 일부 구현예에서, RNase H 활성은 파괴된다.

[0580] 일부 구현예에서, RT 도메인은, 돌연변이가 없는 다른 유사한 도메인에 비해 정확도(fidelity)가 증가하도록 돌연변이된다. 예를 들어, 일부 구현예에서, RT 도메인(예를 들어, 역전사효소에서)에서 YADD 또는 YMDD 모티프는 YVDD로 대체된다. 구현예에서, YADD 또는 YMDD 또는 YVDD의 대체는 레트로바이러스 역전사효소 활성에서 더 높은 정확도를 생성한다(예를 들어, 문헌[Jamburuthugoda and Eickbush J Mol Biol 2011]에 기재된 바와 같으며; 이는 그 전체가 본 명세서에 참고로 포함됨).

[0581] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 표 6에 따른 아미노산 서열, 또는 이에 대하여 적어도 70%, 80%, 85%, 90%, 95%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 갖는 RT 도메인을 포함한다. 일부 구현예에서, 본 명세서에 기재된 핵산은 표 6에 따른 아미노산 서열, 또는 이에 대하여 적어도 70%, 80%, 85%, 90%, 95%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 갖는 RT 도메인을 인코딩한다.

[0582] **표 6: 레트로바이러스 유래의 예시적인 역전사효소 도메인**

RT 명칭	RT 아미노산 서열
AVIRE_P03360	TAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQY PITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPT VPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGF KNSPTLFDEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRV SGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGTIGYCRLWIPGFA ELAQPLYAATRGGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKG VLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLES LLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTST RPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTK ALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGLLTAGGKAIKNAPEILALLTAVWLPKRV AVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATIS
AVIRE_P03360_3mut	TAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQY PITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPT VPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGF KNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRV SGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGTIGYCRLWIPGFA ELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKG VLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLES LLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTST RPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTK ALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKR VAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATIS

[0583]

AVIRE_P03360_3mutA	TAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQY PITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPT VPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGF KNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRV SGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFA ELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKG VLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLES LLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTST RPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTK ALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKR VAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATIS
BAEVM_P10272	TVSLQDEHRLFDIPVTTSLPDVWLQDFPQAWAETGGLGRAKCQAPIIIDLKPTAVPVSIKQYP MSLEAHMGIRQHIIKFLELGVLRPCRSPWNTPLLPVKKPGTQDYRPVQDLREINKRTVDIHPT VPNPYNLLSTLKPDYSWYTVLDLKDAFFCLPLAPQSQELFAFEWKDPERGISGQLTWTRLPQ GFKNSPTLFDEALHRDLTDFRTQHPEVTLLQYVDDLLLAAPTKKACTQGTRHLLQELGEKGY RASAKKAQICQTKVTYLGYILSEGKRWLTPGRIETVARIPPPRNPREVREFLGTAGFCRLWIPG FAELAAPLYALTKESTPFTWQTEHQLAFEALKKALLSAPALGLPDTSKPFTLFLDERQGIAKG VLTQKLGPWKRPVAYLSKKLDPVAAGWPPCLRIMAATAMLVKDSAKLTLGQPLTVITPHTL EAIVRQPPDRWITNARLTHYQALLLDTDRVQFGPPVTLNPATLLPVPENQPSPHDCRQVLAET HGTREDLKDQELPDADHTWYTDGSSYLDSGTRRAGAAVVDGHNTIWAQSLPPGTSAQKAEL IALTKALELSKGKKANIYTDSRYAFATAHTHGSIYERRGLLTSEGKEIKNKAEIIALLKALFLP QEVAIIHCPGHQKGQDPVAVGNRQADRVARQAAMAEVLTLATEPDNTSHIT
BAEVM_P10272_3mut	TVSLQDEHRLFDIPVTTSLPDVWLQDFPQAWAETGGLGRAKCQAPIIIDLKPTAVPVSIKQYP MSLEAHMGIRQHIIKFLELGVLRPCRSPWNTPLLPVKKPGTQDYRPVQDLREINKRTVDIHPT VPNPYNLLSTLKPDYSWYTVLDLKDAFFCLPLAPQSQELFAFEWKDPERGISGQLTWTRLPQ GFKNSPTLFNEALHRDLTDFRTQHPEVTLLQYVDDLLLAAPTKKACTQGTRHLLQELGEKGY RASAKKAQICQTKVTYLGYILSEGKRWLTPGRIETVARIPPPRNPREVREFLGTAGFCRLWIPG FAELAAPLYALTKPSTPFTWQTEHQLAFEALKKALLSAPALGLPDTSKPFTLFLDERQGIAKG VLTQKLGPWKRPVAYLSKKLDPVAAGWPPCLRIMAATAMLVKDSAKLTLGQPLTVITPHTL EAIVRQPPDRWITNARLTHYQALLLDTDRVQFGPPVTLNPATLLPVPENQPSPHDCRQVLAET HGTREDLKDQELPDADHTWYTDGSSYLDSGTRRAGAAVVDGHNTIWAQSLPPGTSAQKAEL IALTKALELSKGKKANIYTDSRYAFATAHTHGSIYERRGWLTSEGKEIKNKAEIIALLKALFLP QEVAIIHCPGHQKGQDPVAVGNRQADRVARQAAMAEVLTLATEPDNTSHIT
BAEVM_P10272_3mutA	TVSLQDEHRLFDIPVTTSLPDVWLQDFPQAWAETGGLGRAKCQAPIIIDLKPTAVPVSIKQYP MSLEAHMGIRQHIIKFLELGVLRPCRSPWNTPLLPVKKPGTQDYRPVQDLREINKRTVDIHPT VPNPYNLLSTLKPDYSWYTVLDLKDAFFCLPLAPQSQELFAFEWKDPERGISGQLTWTRLPQ GFKNSPTLFNEALHRDLTDFRTQHPEVTLLQYVDDLLLAAPTKKACTQGTRHLLQELGEKGY RASAKKAQICQTKVTYLGYILSEGKRWLTPGRIETVARIPPPRNPREVREFLGKAGFCRLFIPG FAELAAPLYALTKPSTPFTWQTEHQLAFEALKKALLSAPALGLPDTSKPFTLFLDERQGIAKG VLTQKLGPWKRPVAYLSKKLDPVAAGWPPCLRIMAATAMLVKDSAKLTLGQPLTVITPHTL EAIVRQPPDRWITNARLTHYQALLLDTDRVQFGPPVTLNPATLLPVPENQPSPHDCRQVLAET HGTREDLKDQELPDADHTWYTDGSSYLDSGTRRAGAAVVDGHNTIWAQSLPPGTSAQKAEL IALTKALELSKGKKANIYTDSRYAFATAHTHGSIYERRGWLTSEGKEIKNKAEIIALLKALFLP QEVAIIHCPGHQKGQDPVAVGNRQADRVARQAAMAEVLTLATEPDNTSHIT
BLVAU_P25059	GVLDAPPSHIGLEHLPPPPEVPQFPLNLERLQALQDLVHRSLEAGYISPWDGPGNNPVFPVRKP NGAWRFVHDLRVTNALTKPIPALSPGPPDLTAIPTHLPHIICLDLKDAFFQIPVEDRFRSYFAFT LPTPGGLQPHRRFAWRVLPQGFINSPALFERALQEPLRQVSAAFSQSLLVSYMDDILYVSPTEE QRLQCYQTMAAHLRDLGFQVASEKTRQTPSPVPFLGQMVHERMVTYQSLPTLQISSPISLHQL QTVLGDLQWVSRGTPTTRRPLQLLYSSLKGIDDPRAIIHLSPEQQQGIAELRQALSHNARSRY NEQEPLLAYVHLTRAGSTLVLFQKGAQFPLAYFQTPLTDNQASPWGLLLLLGCQYLQAQALS SYAKTILKYYHNLPKTSLDNWIQSSEDPRVQELLQLWPQISSQGIQPPGPWKTLVTRAEVFLTP QFSPEPIPAALCLFSDGAARRGAYCLWKDHLLDFQAVPAPESAQKGELAGLLAGLAAAPPEP LNIWVDSKYLYSLLRTLVLGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSSASHPIASLN NYVDQL

[0584]

BLVAU_P25059_2mut	GVLDAPPSHIGLEHLPPPPEVPQFPLNLERLQALQDLVHRSLEAGYISPWDGPGNNPVFPVRKP NGAWRFVHDLRVTNALTKPIPALSPGPPDLTAIPTHLPHIICLDLKDAFFQIPVEDRFRSYFAFT LPTPGGLQPHRRFAWRVLPQGFINSPALFQRALQEPLRQVSAAFSQSLLVSYMDDILYVSPTE EQRLQCYQTMAAHLRDLGFQVASEKTRQTPSPVPFLGQMVHERMVTYQSLPTLQISSPISLHQ LQTVLGDLQWVSRGTPTTRRPLQLLYSSLKPIDDPRAIIHLSPEQQQGIAELRQALSHNARSRY NEQEPLLAYVHLTRAGSTLVLFQKGAQFPLAYFQTPLTDNQASPWGLLLLLGCQYLQAQALS SYAKTILKYYHNLPKTSLDNWIQSSEDPRVQELLQLWPQISSQGIQPPGPWKTLVTRAEVFLTP QFSPEPIPAALCLFSDGAARRGAYCLWKDHLLDFQAVPAPESAQKGELAGLLAGLAAAPPEP LNIWVDSKYLYSLLRTLVLGAWLQPDPVPSYALLYKSLLRHPAIFVGHVRSHSSASHPIASLN NYVDQL
BLVJ_P03361	GVLDTPPSHIGLEHLPPPPEVPQFPLNLERLQALQDLVHRSLEAGYISPWDGPGNNPVFPVRKP NGAWRFVHDLRATNALTKPIPALSPGPPDLTAIPTHPPHIICLDLKDAFFQIPVEDRFRFYLSFT LPSPGGLQPHRRFAWRVLPQGFINSPALFERALQEPLRQVSAAFSQSLLVSYMDDILYASPTEE QRSQCYQALAARLRDLGFQVASEKTSQTPSPVPFLGQMVHEQIVTYQSLPTLQISSPISLHQLQ AVLGDLQWVSRGTPTTRRPLQLLYSSLKRHHDPRAIIQLSPEQLQGIAELRQALSHNARSRYN EQEPLLAYVHLTRAGSTLVLFQKGAQFPLAYFQTPLTDNQASPWGLLLLLGCQYLQTQALSS YAKPILKYYHNLPKTSLDNWIQSSEDPRVQELLQLWPQISSQGIQPPGPWKTLITRAEVFLTPQ FSPDPIPAALCLFSDGATGRGAYCLWKDHLLDFQAVPAPESAQKGELAGLLAGLAAAPPEPV NIWVDSKYLYSLLRTLVLGAWLQPDPVPSYALLYKSLLRHPAIVVGHVRSHSSASHPIASLNN YVDQL
BLVJ_P03361_2mut	GVLDTPPSHIGLEHLPPPPEVPQFPLNLERLQALQDLVHRSLEAGYISPWDGPGNNPVFPVRKP NGAWRFVHDLRATNALTKPIPALSPGPPDLTAIPTHPPHIICLDLKDAFFQIPVEDRFRFYLSFT LPSPGGLQPHRRFAWRVLPQGFINSPALFNRALQEPLRQVSAAFSQSLLVSYMDDILYASPTEE QRSQCYQALAARLRDLGFQVASEKTSQTPSPVPFLGQMVHEQIVTYQSLPTLQISSPISLHQLQ AVLGDLQWVSRGTPTTRRPLQLLYSSLKRHHDPRAIIQLSPEQLQGIAELRQALSHNARSRYN EQEPLLAYVHLTRAGSTLVLFQKGAQFPLAYFQTPLTDNQASPWGLLLLLGCQYLQTQALSS YAKPILKYYHNLPKTSLDNWIQSSEDPRVQELLQLWPQISSQGIQPPGPWKTLITRAEVFLTPQ FSPDPIPAALCLFSDGATGRGAYCLWKDHLLDFQAVPAPESAQKGELAGLLAGLAAAPPEPV NIWVDSKYLYSLLRTWVLGAWLQPDPVPSYALLYKSLLRHPAIVVGHVRSHSSASHPIASLN NYVDQL
BLVJ_P03361_2mutB	GVLDTPPSHIGLEHLPPPPEVPQFPLNLERLQALQDLVHRSLEAGYISPWDGPGNNPVFPVRKP NGAWRFVHDLRATNALTKPIPALSPGPPDLTAPPTHPPHIICLDLKDAFFQIPVEDRFRFYLSFT LPSPGGLQPHRRFAWRVLPQGFINSPALFQRALQEPLRQVSAAFSQSLLVSYMDDILYASPTEE QRSQCYQALAARLRDLGFQVASEKTSQTPSPVPFLGQMVHEQIVTYQSLPTLQISSPISLHQLQ AVLGDLQWVSRGTPTTRRPLQLLYSSLKRHHDPRAIIQLSPEQLQGIAELRQALSHNARSRYN EQEPLLAYVHLTRAGSTLVLFQKGAQFPLAYFQTPLTDNQASPWGLLLLLGCQYLQTQALSS YAKPILKYYHNLPKTSLDNWIQSSEDPRVQELLQLWPQISSQGIQPPGPWKTLITRAEVFLTPQ FSPDPIPAALCLFSDGATGRGAYCLWKDHLLDFQAVPAPESAQKGELAGLLAGLAAAPPEPV NIWVDSKYLYSLLRTWVLGAWLQPDPVPSYALLYKSLLRHPAIVVGHVRSHSSASHPIASLN NYVDQL
FFV_O93209	MDLLKPLTVERKGVKIKGYWNSQADITCVPKDLLQGEEPVRQQNVTTIHGTQEGDVYYVNL KIDGRRINTEVIGTTLDYAIITPGDVPWILKKPLELTIKLDLEEQQGTLLNNSILSKKGKEELKQ LFEKYSALWQSWENQVGHRRIRPHKIATGTVKPTPQKQYHINPKAKPDIQIVINDLLKQGVLI QKESTMNTPVYPVPKPNGRWRMVLDYRAVNKVTPLIAVQNQHSYGILGSLFKGRYKTTIDLS NGFWAHPIVPEDYWITAFTWQGKQYCWTVLPQGFLNSPGLFTGDVVDLLQGIPNVEVYVDD VYISHDSEKEHLEYLDILFNRLKEAGYIISLKKSNIANSIVDFLGFQITNEGRGLTDTFKEKLENI TAPTTLKQLQSILGLLNFARNFIPDFTELIAPLYALIPKSTKNYVPWQIEHSTTLETLITKLNGAE YLQGRKGDKTLIMKVNASYTTGYIRYYNEGEKKPISYVSIVFSKTELKFTELEKLLTTVHKGL LKALDLSMGQNIHVYSPIVSMQNIQKTPQTAKKALASRWLSWLSYLEDPRIRFFYDPQMPAL KDLPAVDTGKDNKKHPSNFQHIFYTDGSAITSPTKEGHLNAGMGIVYFINKDGNLQKQQEWS ISLGNHTAQFAEIAAFEFALKKCLPLGGNILVVTDSNYVAKAYNEELDVWASNGFVNNRKKP LKHISKWKSVADLKRLRPDVVVTHEPGHQKLDSSPHAYGNNLADQLATQASFKVH
FFV_O93209_2mut	MDLLKPLTVERKGVKIKGYWNSQADITCVPKDLLQGEEPVRQQNVTTIHGTQEGDVYYVNL KIDGRRINTEVIGTTLDYAIITPGDVPWILKKPLELTIKLDLEEQQGTLLNNSILSKKGKEELKQ LFEKYSALWQSWENQVGHRRIRPHKIATGTVKPTPQKQYHINPKAKPDIQIVINDLLKQGVLI QKESTMNTPVYPVPKPNGRWRMVLDYRAVNKVTPLIAVQNQHSYGILGSLFKGRYKTTIDLS

[0585]

	NGFWAHPIVPEDYWITAFTWQGKQYCWTVLPQGFLNSPGLFNGDVVDLLQGIPNVEVYVDD VYISHDSEKEHLEYLDILFNRLKEAGYIISLKKSNIANSIVDFLGFQITNEGRGLTDTFKEKLENI TAPTTLKQLQSILGLLNFARNFIPDFTELIAPLYALIPKSPKNYVPWQIEHSTTLETLITKLNGAE YLQGRKGDKTLIMKVNASYTTGYIRYYNEGEKKPISYVSIVFSKTELKFTELEKLLTTVHKGL LKALDLSMGQNIHVYSPIVSMQNIQKTPQTAKKALASRWLSWLSYLEDPRIRFFYDPQMPAL KDLPAVDTGKDNKKHPSNFQHIFYTDGSAITSPTKEGHLNAGMGIVYFINKDGNLQKQQEWS ISLGNHTAQFAEIAAFEFALKKCLPLGGNILVVTDSNYVAKAYNEELDVWASNGFVNNRKKP LKHISKWKSVADLKRLRPDVVVTHEPGHQKLDSSPHAYGNNLADQLATQASFKVH
FFV_O93209_2mutA	MDLLKPLTVERKGVKIKGYWNSQADITCVPKDLLQGEEPVRQQNVTTIHGTQEGDVYYVNL KIDGRRINTEVIGTTLDYAIITPGDVPWILKKPLELTIKLDLEEQQGTLLNNSILSKKGKEELKQ LFEKYSALWQSWENQVGHRRIRPHKIATGTVKPTPQKQYHINPKAKPDIQIVINDLLKQGVLI QKESTMNTPVYPVPKPNGRWRMVLDYRAVNKVTPLIAVQNQHSYGILGSLFKGRYKTTIDLS NGFWAHPIVPEDYWITAFTWQGKQYCWTVLPQGFLNSPGLFNGDVVDLLQGIPNVEVYVDD VYISHDSEKEHLEYLDILFNRLKEAGYIISLKKSNIANSIVDFLGFQITNEGRGLTDTFKEKLENI TAPTTLKQLQSILGKLNFARNFIPDFTELIAPLYALIPKSPKNYVPWQIEHSTTLETLITKLNGAE YLQGRKGDKTLIMKVNASYTTGYIRYYNEGEKKPISYVSIVFSKTELKFTELEKLLTTVHKGL LKALDLSMGQNIHVYSPIVSMQNIQKTPQTAKKALASRWLSWLSYLEDPRIRFFYDPQMPAL KDLPAVDTGKDNKKHPSNFQHIFYTDGSAITSPTKEGHLNAGMGIVYFINKDGNLQKQQEWS ISLGNHTAQFAEIAAFEFALKKCLPLGGNILVVTDSNYVAKAYNEELDVWASNGFVNNRKKP LKHISKWKSVADLKRLRPDVVVTHEPGHQKLDSSPHAYGNNLADQLATQASFKVH
FFV_O93209-Pro	VPWILKKPLELTIKLDLEEQQGTLLNNSILSKKGKEELKQLFEKYSALWQSWENQVGHRRIRP HKIATGTVKPTPQKQYHINPKAKPDIQIVINDLLKQGVLIQKESTMNTPVYPVPKPNGRWRMV LDYRAVNKVTPLIAVQNQHSYGILGSLFKGRYKTTIDLSNGFWAHPIVPEDYWITAFTWQGK QYCWTVLPQGFLNSPGLFTGDVVDLLQGIPNVEVYVDDVYISHDSEKEHLEYLDILFNRLKE AGYIISLKKSNIANSIVDFLGFQITNEGRGLTDTFKEKLENITAPTTLKQLQSILGLLNFARNFIP DFTELIAPLYALIPKSTKNYVPWQIEHSTTLETLITKLNGAEYLQGRKGDKTLIMKVNASYTTG YIRYYNEGEKKPISYVSIVFSKTELKFTELEKLLTTVHKGLLKALDLSMGQNIHVYSPIVSMQN IQKTPQTAKKALASRWLSWLSYLEDPRIRFFYDPQMPALKDLPAVDTGKDNKKHPSNFQHIF YTDGSAITSPTKEGHLNAGMGIVYFINKDGNLQKQQEWSISLGNHTAQFAEIAAFEFALKKCL PLGGNILVVTDSNYVAKAYNEELDVWASNGFVNNRKKPLKHISKWKSVADLKRLRPDVVVT HEPGHQKLDSSPHAYGNNLADQLATQASFKVH
FFV_O93209-Pro_2mut	VPWILKKPLELTIKLDLEEQQGTLLNNSILSKKGKEELKQLFEKYSALWQSWENQVGHRRIRP HKIATGTVKPTPQKQYHINPKAKPDIQIVINDLLKQGVLIQKESTMNTPVYPVPKPNGRWRMV LDYRAVNKVTPLIAVQNQHSYGILGSLFKGRYKTTIDLSNGFWAHPIVPEDYWITAFTWQGK QYCWTVLPQGFLNSPGLFNGDVVDLLQGIPNVEVYVDDVYISHDSEKEHLEYLDILFNRLKE AGYIISLKKSNIANSIVDFLGFQITNEGRGLTDTFKEKLENITAPTTLKQLQSILGLLNFARNFIP DFTELIAPLYALIPKSPKNYVPWQIEHSTTLETLITKLNGAEYLQGRKGDKTLIMKVNASYTTG YIRYYNEGEKKPISYVSIVFSKTELKFTELEKLLTTVHKGLLKALDLSMGQNIHVYSPIVSMQN IQKTPQTAKKALASRWLSWLSYLEDPRIRFFYDPQMPALKDLPAVDTGKDNKKHPSNFQHIF YTDGSAITSPTKEGHLNAGMGIVYFINKDGNLQKQQEWSISLGNHTAQFAEIAAFEFALKKCL PLGGNILVVTDSNYVAKAYNEELDVWASNGFVNNRKKPLKHISKWKSVADLKRLRPDVVVT HEPGHQKLDSSPHAYGNNLADQLATQASFKVH
FFV_O93209-Pro_2mutA	VPWILKKPLELTIKLDLEEQQGTLLNNSILSKKGKEELKQLFEKYSALWQSWENQVGHRRIRP HKIATGTVKPTPQKQYHINPKAKPDIQIVINDLLKQGVLIQKESTMNTPVYPVPKPNGRWRMV LDYRAVNKVTPLIAVQNQHSYGILGSLFKGRYKTTIDLSNGFWAHPIVPEDYWITAFTWQGK QYCWTVLPQGFLNSPGLFNGDVVDLLQGIPNVEVYVDDVYISHDSEKEHLEYLDILFNRLKE AGYIISLKKSNIANSIVDFLGFQITNEGRGLTDTFKEKLENITAPTTLKQLQSILGKLNFARNFIP DFTELIAPLYALIPKSPKNYVPWQIEHSTTLETLITKLNGAEYLQGRKGDKTLIMKVNASYTTG YIRYYNEGEKKPISYVSIVFSKTELKFTELEKLLTTVHKGLLKALDLSMGQNIHVYSPIVSMQN IQKTPQTAKKALASRWLSWLSYLEDPRIRFFYDPQMPALKDLPAVDTGKDNKKHPSNFQHIF YTDGSAITSPTKEGHLNAGMGIVYFINKDGNLQKQQEWSISLGNHTAQFAEIAAFEFALKKCL PLGGNILVVTDSNYVAKAYNEELDVWASNGFVNNRKKPLKHISKWKSVADLKRLRPDVVVT HEPGHQKLDSSPHAYGNNLADQLATQASFKVH
FLV_P10273	TLQLEEEYRLFEPESTQKQEMDIWLKNFPQAWAETGGMGTAHCQAPVLIQLKATATPISIRQY PMPHEAYQGIKPHIRRMLDQGILKPCQSPWNTPLLPVKKPGTEDYRPVQDLREVNKRVEDIHP TVPNPYNLLSTLPPSHPWYTVLDLKDAFFCLRLHSESQLLFAFEWRDPEIGLSGQLTWTRLPQ

[0586]

	GFKNSPTLFDEALHSDLADFRVRYPALVLLQYVDDLLLAAATRTECLEGTKALLETLGNKGY RASAKKAQICLQEVTYLGYSLKDGQRWLTKARKEAILSIPVPKNSRQVREFLGTAGYCRLWI PGFAELAAPLYPLTRPGTLFQWGTEQQLAFEDIKKALLSSPALGLPDITKPFELFIDENSGFAK GVLVQKLGPWKRPVAYLSKKLDTVASGWPPCLRMVAAIAILVKDAGKLTLGQPLTILTSHPV EALVRQPPNKWLSNARMTHYQAMLLDAERVHFGPTVSLNPATLLPLPSGGNHHDCLQILAE THGTRPDLTDQPLPDADLTWYTDGSSFIRNGEREAGAAVTTESEVIWAAPLPPGTSAQRAELI ALTQALKMAEGKKLTVYTDSRYAFATTHVHGEIYRRRGLLTSEGKEIKNKNEILALLEALFLP KRLSIIHCPGHQKGDSPQAKGNRLADDTAKKAATETHSSLTVLP
FLV_P10273_3mut	TLQLEEEYRLFEPESTQKQEMDIWLKNFPQAWAETGGMGTAHCQAPVLIQLKATATPISIRQY PMPHEAYQGIKPHIRRMLDQGILKPCQSPWNTPLLPVKKPGTEDYRPVQDLREVNKRVEDIHP TVPNPYNLLSTLPPSHPWYTVLDLKDAFFCLRLHSESQLLFAFEWRDPEIGLSGQLTWTRLPQ GFKNSPTLFNEALHSDLADFRVRYPALVLLQYVDDLLLAAATRTECLEGTKALLETLGNKGY RASAKKAQICLQEVTYLGYSLKDGQRWLTKARKEAILSIPVPKNSRQVREFLGTAGYCRLWI PGFAELAAPLYPLTRPGTLFQWGTEQQLAFEDIKKALLSSPALGLPDITKPFELFIDENSGFAK GVLVQKLGPWKRPVAYLSKKLDTVASGWPPCLRMVAAIAILVKDAGKLTLGQPLTILTSHPV EALVRQPPNKWLSNARMTHYQAMLLDAERVHFGPTVSLNPATLLPLPSGGNHHDCLQILAE THGTRPDLTDQPLPDADLTWYTDGSSFIRNGEREAGAAVTTESEVIWAAPLPPGTSAQRAELI ALTQALKMAEGKKLTVYTDSRYAFATTHVHGEIYRRRGWLTSEGKEIKNKNEILALLEALFL PKRLSIIHCPGHQKGDSPQAKGNRLADDTAKKAATETHSSLTVLP
FLV_P10273_3mutA	TLQLEEEYRLFEPESTQKQEMDIWLKNFPQAWAETGGMGTAHCQAPVLIQLKATATPISIRQY PMPHEAYQGIKPHIRRMLDQGILKPCQSPWNTPLLPVKKPGTEDYRPVQDLREVNKRVEDIHP TVPNPYNLLSTLPPSHPWYTVLDLKDAFFCLRLHSESQLLFAFEWRDPEIGLSGQLTWTRLPQ GFKNSPTLFNEALHSDLADFRVRYPALVLLQYVDDLLLAAATRTECLEGTKALLETLGNKGY RASAKKAQICLQEVTYLGYSLKDGQRWLTKARKEAILSIPVPKNSRQVREFLGKAGYCRLFIP GFAELAAPLYPLTRPGTLFQWGTEQQLAFEDIKKALLSSPALGLPDITKPFELFIDENSGFAKG VLVQKLGPWKRPVAYLSKKLDTVASGWPPCLRMVAAIAILVKDAGKLTLGQPLTILTSHPVE ALVRQPPNKWLSNARMTHYQAMLLDAERVHFGPTVSLNPATLLPLPSGGNHHDCLQILAET HGTRPDLTDQPLPDADLTWYTDGSSFIRNGEREAGAAVTTESEVIWAAPLPPGTSAQRAELIA LTQALKMAEGKKLTVYTDSRYAFATTHVHGEIYRRRGWLTSEGKEIKNKNEILALLEALFLP KRLSIIHCPGHQKGDSPQAKGNRLADDTAKKAATETHSSLTVLP
FOAM_V_P14350	MNPLQLLQPLPAEIKGTKLLAHWNSGATITCIPESFLEDEQPIKKTLIKTIHGEKQQNVYYVTF KVKGRKVEAEVIASPYEYILLSPTDVPWLTQQPLQLTILVPLQEYQEKILSKTALPEDQKQQLK TLFVKYDNLWQHWENQVGHRKIRPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQGVL TPQNSTMNTPVYPVPKPDGRWRMVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTLDL ANGFWAHPITPESYWLTAFTWQGKQYCWTRLPQGFLNSPALFTADVVDLLKEIPNVQVYVD DIYLSHDDPKEHVQQLEKVFQILLQAGYVVSLKKSEIGQKTVEFLGFNITKEGRGLTDTFKTK LLNITPPKDLKQLQSILGLLNFARNFIPNFAELVQPLYNLIASAKGKYIEWSEENTKQLNMVIE ALNTASNLEERLPEQRLVIKVNTSPSAGYVRYYNETGKKPIMYLNYVFSKAELKFSMLEKLLT TMHKALIKAMDLAMGQEILVYSPIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFHYDK TLPELKHIPDVYTSSQSPVKHPSQYEGVFYTDGSAIKSPDPTKSNNAGMGIVHATYKPEYQVL NQWSIPLGNHTAQMAEIAAVEFACKKALKIPGPVLVITDSFYVAESANKELPYWKSNGFVNN KKKPLKHISKWKSIAECLSMKPDITIQHEKGISLQIPVFILKGNALADKLATQGSYVVN
FOAM_V_P14350_2mut	MNPLQLLQPLPAEIKGTKLLAHWNSGATITCIPESFLEDEQPIKKTLIKTIHGEKQQNVYYVTF KVKGRKVEAEVIASPYEYILLSPTDVPWLTQQPLQLTILVPLQEYQEKILSKTALPEDQKQQLK TLFVKYDNLWQHWENQVGHRKIRPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQGVL TPQNSTMNTPVYPVPKPDGRWRMVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTLDL ANGFWAHPITPESYWLTAFTWQGKQYCWTRLPQGFLNSPALFNADVVDLLKEIPNVQVYVD DIYLSHDDPKEHVQQLEKVFQILLQAGYVVSLKKSEIGQKTVEFLGFNITKEGRGLTDTFKTK LLNITPPKDLKQLQSILGLLNFARNFIPNFAELVQPLYNLIAPAKGKYIEWSEENTKQLNMVIE ALNTASNLEERLPEQRLVIKVNTSPSAGYVRYYNETGKKPIMYLNYVFSKAELKFSMLEKLLT TMHKALIKAMDLAMGQEILVYSPIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFHYDK TLPELKHIPDVYTSSQSPVKHPSQYEGVFYTDGSAIKSPDPTKSNNAGMGIVHATYKPEYQVL NQWSIPLGNHTAQMAEIAAVEFACKKALKIPGPVLVITDSFYVAESANKELPYWKSNGFVNN KKKPLKHISKWKSIAECLSMKPDITIQHEKGISLQIPVFILKGNALADKLATQGSYVVN
FOAM_V_P143	MNPLQLLQPLPAEIKGTKLLAHWNSGATITCIPESFLEDEQPIKKTLIKTIHGEKQQNVYYVTF KVKGRKVEAEVIASPYEYILLSPTDVPWLTQQPLQLTILVPLQEYQEKILSKTALPEDQKQQLK

[0587]

[0588]

50_2mu tA	TLFVKYDNLWQHWENQVGHRKIRPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQGVL TPQNSTMNTPVYPVPKPDGRWRMVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTLDL ANGFWAHPITPESYWLTAFTWQGKQYCWTRLPQGFLNSPALFNADVVDLLKEIPNVQVYVD DIYLSHDDPKEHVQQLEKVFQILLQAGYVVSLKKSEIGQKTVEFLGFNITKEGRGLTDTFKTK LLNITPPKDLKQLQSILGKLNFARNFIPNFAELVQPLYNLIAPAKGKYIEWSEENTKQLNMVIE ALNTASNLEERLPEQRLVIKVNTSPSAGYVRYYNETGKKPIMYLNYVFSKAELKFSMLEKLLT TMHKALIKAMDLAMGQEILVYSPIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFHYDK TLPELKHIPDVYTSSQSPVKHPSQYEGVFYTDGSAIKSPDPTKSNNAGMGIVHATYKPEYQVL NQWSIPLGNHTAQMAEIAAVEFACKKALKIPGPVLVITDSFYVAESANKELPYWKSNGFVNN KKKPLKHISKWKSIAECLSMKPDITIQHEKGISLQIPVFILKGNALADKLATQGSYVVN
FOAM V_P143 50-Pro	VPWLTQQPLQLTILVPLQEYQEKILSKTALPEDQKQQLKTLFVKYDNLWQHWENQVGHRKI RPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQGVLTPQNSTMNTPVYPVPKPDGRWR MVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTLDLANGFWAHPITPESYWLTAFTWQ GKQYCWTRLPQGFLNSPALFTADVVDLLKEIPNVQVYVDDIYLSHDDPKEHVQQLEKVFQIL LQAGYVVSLKKSEIGQKTVEFLGFNITKEGRGLTDTFKTKLLNITPPKDLKQLQSILGLLNFAR NFIPNFAELVQPLYNLIASAKGKYIEWSEENTKQLNMVIEALNTASNLEERLPEQRLVIKVNTS PSAGYVRYYNETGKKPIMYLNYVFSKAELKFSMLEKLLTTMHKALIKAMDLAMGQEILVYS PIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFHYDKTLPELKHIPDVYTSSQSPVKHPS QYEGVFYTDGSAIKSPDPTKSNNAGMGIVHATYKPEYQVLNQWSIPLGNHTAQMAEIAAVEF ACKKALKIPGPVLVITDSFYVAESANKELPYWKSNGFVNNKKKPLKHISKWKSIAECLSMKP DITIQHEKGISLQIPVFILKGNALADKLATQGSYVVN
FOAM V_P143 50-Pro_2mut	VPWLTQQPLQLTILVPLQEYQEKILSKTALPEDQKQQLKTLFVKYDNLWQHWENQVGHRKI RPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQGVLTPQNSTMNTPVYPVPKPDGRWR MVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTLDLANGFWAHPITPESYWLTAFTWQ GKQYCWTRLPQGFLNSPALFNADVVDLLKEIPNVQVYVDDIYLSHDDPKEHVQQLEKVFQIL LQAGYVVSLKKSEIGQKTVEFLGFNITKEGRGLTDTFKTKLLNITPPKDLKQLQSILGLLNFAR NFIPNFAELVQPLYNLIAPAKGKYIEWSEENTKQLNMVIEALNTASNLEERLPEQRLVIKVNTS PSAGYVRYYNETGKKPIMYLNYVFSKAELKFSMLEKLLTTMHKALIKAMDLAMGQEILVYS PIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFHYDKTLPELKHIPDVYTSSQSPVKHPS QYEGVFYTDGSAIKSPDPTKSNNAGMGIVHATYKPEYQVLNQWSIPLGNHTAQMAEIAAVEF ACKKALKIPGPVLVITDSFYVAESANKELPYWKSNGFVNNKKKPLKHISKWKSIAECLSMKP DITIQHEKGISLQIPVFILKGNALADKLATQGSYVVN
FOAM V_P143 50-Pro_2mutA	VPWLTQQPLQLTILVPLQEYQEKILSKTALPEDQKQQLKTLFVKYDNLWQHWENQVGHRKI RPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQGVLTPQNSTMNTPVYPVPKPDGRWR MVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTLDLANGFWAHPITPESYWLTAFTWQ GKQYCWTRLPQGFLNSPALFNADVVDLLKEIPNVQVYVDDIYLSHDDPKEHVQQLEKVFQIL LQAGYVVSLKKSEIGQKTVEFLGFNITKEGRGLTDTFKTKLLNITPPKDLKQLQSILGKLNFAR NFIPNFAELVQPLYNLIAPAKGKYIEWSEENTKQLNMVIEALNTASNLEERLPEQRLVIKVNTS PSAGYVRYYNETGKKPIMYLNYVFSKAELKFSMLEKLLTTMHKALIKAMDLAMGQEILVYS PIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFHYDKTLPELKHIPDVYTSSQSPVKHPS QYEGVFYTDGSAIKSPDPTKSNNAGMGIVHATYKPEYQVLNQWSIPLGNHTAQMAEIAAVEF ACKKALKIPGPVLVITDSFYVAESANKELPYWKSNGFVNNKKKPLKHISKWKSIAECLSMKP DITIQHEKGISLQIPVFILKGNALADKLATQGSYVVN
GALV_P21414	VLNLEEEYRLHEKPVPSSIDPSWLQLFPTVWAERAGMGLANQVPPVVVELRSGASPVAVRQY PMSKEAREGIRPHIQKFLDLGVLVPCRSPWNTPLLPVKKPGTNDYRPVQDLREINKRVQDIHP TVPNPYNLLSSLPPSYTWYSVLDLKDAFFCLRLHPNSQPLFAFEWKDPEKGNTGQLTWTRLP QGFKNSPTLFDEALHRDLAPFRALNPQVVLLQYVDDLLVAAPTYEDCKKGTQKLLQELSKLG YRVSAKKAQLCQREVTYLGYLLKEGKRWLTPARKATVMKIPVPTTPRQVREFLGTAGFCRL WIPGFASLAAPLYPLTKESIPFIWTEEHQQAFDHIKKALLSAPALALPDLTKPFTLYIDERAGV ARGVLTQTLGPWRRPVAYLSKKLDPVASGWPTCLKAVAAVALLLKDADKLTLGQNVTVIAS HSLESIVRQPPDRWMTNARMTHYQSLLLNERVSFAPPAVLNPATLLPVESEATPVHRCSEILA EETGTRRDLEDQPLPGVPTWYTDGSSFITEGKRRAGAPIVDGKRTVWASSLPEGTSAQKAEL VALTQALRLAEGKNINIYTDSRYAFATAHIHGAIYKQRGLLTSAGKDIKNKEEILALLEAIHLP RRVAIIHCPGHQRGSNPVATGNRRADEAAKQAALSTRVLAGTTKP

GALV_P21414_3mut	VLNLEEEYRLHEKPVPSSIDPSWLQLFPTVWAERAGMGLANQVPPVVVELRSGASPVAVRQY PMSKEAREGIRPHIQKFLDLGVLVPCRSPWNTPLLPVKKPGTNDYRPVQDLREINKRVQDIHP TVPNPYNLLSSLPPSYTWYSVLDLKDAFFCLRLHPNSQPLFAFEWKDPEKGNTGQLTWTRLP QGFKNSPTLFNEALHRDLAPFRALNPQVVLLQYVDDLLVAAPTYEDCKKGTQKLLQELSKLG YRVSAKKAQLCQREVTYLGYLLKEGKRWLTPARKATVMKIPVPTTPRQVREFLGTAGFCRL WIPGFASLAAPLYPLTKPSIPFIWTEEHQQAFDHIKKALLSAPALALPDLTKPFTLYIDERAGVA RGVLTQTLGPWRRPVAYLSKKLDPVASGWPTCLKAVAAVALLLKDADKLTLGQNVTVIASH SLESIVRQPPDRWMTNARMTHYQSLLLNERVSFAPPAVLNPATLLPVESEATPVHRCSEILAE ETGTRRDLEDQPLPGVPTWYTDGSSFITEGKRRAGAPIVDGKRTVWASSLPEGTSAQKAELV ALTQALRLAEGKNINIYTDSRYAFATAHIHGAIYKQRGWLTSAGKDIKNKEEILALLEAIHLPR RVAIIHCPGHQRGSNPVATGNRRADEAAKQAALSTRVLAGTTKP
GALV_P21414_3mutA	VLNLEEEYRLHEKPVPSSIDPSWLQLFPTVWAERAGMGLANQVPPVVVELRSGASPVAVRQY PMSKEAREGIRPHIQKFLDLGVLVPCRSPWNTPLLPVKKPGTNDYRPVQDLREINKRVQDIHP TVPNPYNLLSSLPPSYTWYSVLDLKDAFFCLRLHPNSQPLFAFEWKDPEKGNTGQLTWTRLP QGFKNSPTLFNEALHRDLAPFRALNPQVVLLQYVDDLLVAAPTYEDCKKGTQKLLQELSKLG YRVSAKKAQLCQREVTYLGYLLKEGKRWLTPARKATVMKIPVPTTPRQVREFLGKAGFCRL FIPGFASLAAPLYPLTKPSIPFIWTEEHQQAFDHIKKALLSAPALALPDLTKPFTLYIDERAGVA RGVLTQTLGPWRRPVAYLSKKLDPVASGWPTCLKAVAAVALLLKDADKLTLGQNVTVIASH SLESIVRQPPDRWMTNARMTHYQSLLLNERVSFAPPAVLNPATLLPVESEATPVHRCSEILAE ETGTRRDLEDQPLPGVPTWYTDGSSFITEGKRRAGAPIVDGKRTVWASSLPEGTSAQKAELV ALTQALRLAEGKNINIYTDSRYAFATAHIHGAIYKQRGWLTSAGKDIKNKEEILALLEAIHLPR RVAIIHCPGHQRGSNPVATGNRRADEAAKQAALSTRVLAGTTKP
HTL1A_P03362	AVLGLEHLPRPPQISQFPLNPERLQALQHLVRKALEAGHIEPYTGPGNNPVFPVKKANGTWRF IHDLRATNSLTIDLSSSSPGPPDLSSLPTTLAHLQTIDLRDAFFQIPLPKQFQPYFAFTVPQQCNY GPGTRYAWKVLPQGFKNSPTLFEMQLAHILQPIRQAFPQCTILQYMDDILLASPSHEDLLLLSE ATMASLISHGLPVSENKTQQTPGTIKFLGQIISPNHLTYDAVPTVPIRSRWALPELQALLGEIQ WVSKGTPTLRQPLHSLYCALQRHTDPRDQIYLNPSQVQSLVQLRQALSQNCRSRLVQTLPLL GAIMLTLTGTTTVVFQSKEQWPLVWLHAPLPHTSQCPWGQLLASAVLLLDKYTLQSYGLLC QTIHHNISTQTFNQFIQTSDHPSVPILLHHSHRFKNLGAQTGELWNTFLKTAAPLAPVKALMP VFTLSPVIINTAPCLFSDGSTSRAAYILWDKQILSQRSFPLPPPHKSAQRAELLGLLHGLSSARS WRCLNIFLDSKYLYHYLRTLALGTFQGRSSQAPFQALLPRLLSRKVVYLHHVRSHTNLPDPIS RLNALTDALLITPVLQL
HTL1A_P03362_2mut	AVLGLEHLPRPPQISQFPLNPERLQALQHLVRKALEAGHIEPYTGPGNNPVFPVKKANGTWRF IHDLRATNSLTIDLSSSSPGPPDLSSLPTTLAHLQTIDLRDAFFQIPLPKQFQPYFAFTVPQQCNY GPGTRYAWKVLPQGFKNSPTLFQMQLAHILQPIRQAFPQCTILQYMDDILLASPSHEDLLLLSE ATMASLISHGLPVSENKTQQTPGTIKFLGQIISPNHLTYDAVPTVPIRSRWALPELQALLGEIQ WVSKGTPTLRQPLHSLYCALQPHTDPRDQIYLNPSQVQSLVQLRQALSQNCRSRLVQTLPLL GAIMLTLTGTTTVVFQSKEQWPLVWLHAPLPHTSQCPWGQLLASAVLLLDKYTLQSYGLLC QTIHHNISTQTFNQFIQTSDHPSVPILLHHSHRFKNLGAQTGELWNTFLKTAAPLAPVKALMP VFTLSPVIINTAPCLFSDGSTSRAAYILWDKQILSQRSFPLPPPHKSAQRAELLGLLHGLSSARS WRCLNIFLDSKYLYHYLRTLALGTFQGRSSQAPFQALLPRLLSRKVVYLHHVRSHTNLPDPIS RLNALTDALLITPVLQL
HTL1A_P03362_2mutB	AVLGLEHLPRPPQISQFPLNPERLQALQHLVRKALEAGHIEPYTGPGNNPVFPVKKANGTWRF IHDLRATNSLTIDLSSSSPGPPDLSSPPTTLAHLQTIDLRDAFFQIPLPKQFQPYFAFTVPQQCNY GPGTRYAWKVLPQGFKNSPTLFQMQLAHILQPIRQAFPQCTILQYMDDILLASPSHEDLLLLSE ATMASLISHGLPVSENKTQQTPGTIKFLGQIISPNHLTYDAVPTVPIRSRWALPELQALLGEIQ WVSKGTPTLRQPLHSLYCALQPHTDPRDQIYLNPSQVQSLVQLRQALSQNCRSRLVQTLPLL GAIMLTLTGTTTVVFQSKEQWPLVWLHAPLPHTSQCPWGQLLASAVLLLDKYTLQSYGLLC QTIHHNISTQTFNQFIQTSDHPSVPILLHHSHRFKNLGAQTGELWNTFLKTAAPLAPVKALMP VFTLSPVIINTAPCLFSDGSTSRAAYILWDKQILSQRSFPLPPPHKSAQRAELLGLLHGLSSARS WRCLNIFLDSKYLYHYLRTLALGTFQGRSSQAPFQALLPRLLSRKVVYLHHVRSHTNLPDPIS RLNALTDALLITPVLQL
HTL1C_P14078	AVLGLEHLPRPPEISQFPLNPERLQALQHLVRKALEAGHIEPYTGPGNNPVFPVKKANGTWRF IHDLRATNSLTIDLSSSSPGPPDLSSLPTTLAHLQTIDLKDAFFQIPLPKQFQPYFAFTVPQQCNY GPGTRYAWRVLPQGFKNSPTLFEMQLAHILQPIRQAFPQCTILQYMDDILLASPSHADLQLLS EATMASLISHGLPVSENKTQQTPGTIKFLGQIISPNHLTYDAVPKVPIRSRWALPELQALLGEIQ

[0589]

	WVSKGTPTLRQPLHSLYCALQRHTDPRDQIYLNPSQVQSLVQLRQALSQNCRSRLVQTLPLL GAIMLTLTGTTTVVFQSKQQWPLVWLHAPLPHTSQCPWGQLLASAVLLLDKYTLQSYGLLC QTIHHNISTQTFNQFIQTSDHPSVPILLHHSHRFKNLGAQTGELWNTFLKTTAPLAPVKALMPV FTLSPVIINTAPCLFSDGSTSQAAYILWDKHILSQRSFPLPPPHKSAQRAELLGLLHGLSSARSW RCLNIFLDSKYLYHYLRTLALGTFQGRSSQAPFQALLPRLLSRKVVYLHHVRSHTNLPDPISRL NALTDALLITPVLQL
HTL1C_P14078_2mut	AVLGLEHLPRPPEISQFPLNPERLQALQHLVRKALEAGHIEPYTGPGNNPVFPVKKANGTWRF IHDLRATNSLTIDLSSSSPGPPDLSSLPTTLAHLQTIDLKDAFFQIPLPKQFQPYFAFTVPQQCNY GPGTRYAWRVLPQGFKNSPTLFQMQLAHILQPIRQAFPQCTILQYMDDILLASPSHADLQLLS EATMASLISHGLPVSENKTQQTPGTIKFLGQIISPNHLTYDAVPKVPIRSRWALPELQALLGEIQ WVSKGTPTLRQPLHSLYCALQPHTDPRDQIYLNPSQVQSLVQLRQALSQNCRSRLVQTLPLL GAIMLTLTGTTTVVFQSKQQWPLVWLHAPLPHTSQCPWGQLLASAVLLLDKYTLQSYGLLC QTIHHNISTQTFNQFIQTSDHPSVPILLHHSHRFKNLGAQTGELWNTFLKTTAPLAPVKALMPV FTLSPVIINTAPCLFSDGSTSQAAYILWDKHILSQRSFPLPPPHKSAQRAELLGLLHGLSSARSW RCLNIFLDSKYLYHYLRTLALGTFQGRSSQAPFQALLPRLLSRKVVYLHHVRSHTNLPDPISRL NALTDALLITPVLQL
HTL1L_P0C211	GLEHLPRPPEISQFPLNPERLQALQHLVRKALEAGHIEPYTGPGNNPVFPVKKANGTWRFIHD LRATNSLTVDLSSSSPGPPDLSSLPTTLAHLQTIDLKDAFFQIPLPKQFQPYFAFTVPQQCNYGP GTRYAWKVLPQGFKNSPTLFEMQLASILQPIRQAFPQCVILQYMDDILLASPSPEDLQQLSEAT MASLISHGLPVSQDKTQQTPGTIKFLGQIISPNHITYDAVPTVPIRSRWALPELQALLGEIQWVS KGTPTLRQPLHSLYCALQGHTDPRDQIYLNPSQVQSLMQLQQALSQNCRSRLAQTLPLLGAI MLTLTGTTTVVFQSKQQWPLVWLHAPLPHTSQCPWGQLLASAVLLLDKYTLQSYGLLCQTI HHNISIQTFNQFIQTSDHPSVPILLHHSHRFKNLGAQTGELWNTFLKTAAPLAPVKALTPVFTL SPIIINTAPCLFSDGSTSQAAYILWDKHILSQRSFPLPPPHKSAQQAELLGLLHGLSSARSWHCL NIFLDSKYLYHYLRTLALGTFQGKSSQAPFQALLPRLLAHKVIYLHHVRSHTNLPDPISKLNAL TDALLITPIL
HTL1L_P0C211_2mut	GLEHLPRPPEISQFPLNPERLQALQHLVRKALEAGHIEPYTGPGNNPVFPVKKANGTWRFIHD LRATNSLTVDLSSSSPGPPDLSSLPTTLAHLQTIDLKDAFFQIPLPKQFQPYFAFTVPQQCNYGP GTRYAWKVLPQGFKNSPTLFQMQLASILQPIRQAFPQCVILQYMDDILLASPSPEDLQQLSEA TMASLISHGLPVSQDKTQQTPGTIKFLGQIISPNHITYDAVPTVPIRSRWALPELQALLGEIQWV SKGTPTLRQPLHSLYCALQGHTDPRDQIYLNPSQVQSLMQLQQALSQNCRSRLAQTLPLLGAI MLTLTGTTTVVFQSKQQWPLVWLHAPLPHTSQCPWGQLLASAVLLLDKYTLQSYGLLCQTI HHNISIQTFNQFIQTSDHPSVPILLHHSHRFKNLGAQTGELWNTFLKTAAPLAPVKALTPVFTL SPIIINTAPCLFSDGSTSQAAYILWDKHILSQRSFPLPPPHKSAQQAELLGLLHGLSSARSWHCL NIFLDSKYLYHYLRTLAWGTFQGKSSQAPFQALLPRLLAHKVIYLHHVRSHTNLPDPISKLNA LTDALLITPIL
HTL1L_P0C211_2mutB	GLEHLPRPPEISQFPLNPERLQALQHLVRKALEAGHIEPYTGPGNNPVFPVKKANGTWRFIHD LRATNSLTVDLSSSSPGPPDLSSPPTTLAHLQTIDLKDAFFQIPLPKQFQPYFAFTVPQQCNYGP GTRYAWKVLPQGFKNSPTLFQMQLASILQPIRQAFPQCVILQYMDDILLASPSPEDLQQLSEA TMASLISHGLPVSQDKTQQTPGTIKFLGQIISPNHITYDAVPTVPIRSRWALPELQALLGEIQWV SKGTPTLRQPLHSLYCALQGHTDPRDQIYLNPSQVQSLMQLQQALSQNCRSRLAQTLPLLGAI MLTLTGTTTVVFQSKQQWPLVWLHAPLPHTSQCPWGQLLASAVLLLDKYTLQSYGLLCQTI HHNISIQTFNQFIQTSDHPSVPILLHHSHRFKNLGAQTGELWNTFLKTAAPLAPVKALTPVFTL SPIIINTAPCLFSDGSTSQAAYILWDKHILSQRSFPLPPPHKSAQQAELLGLLHGLSSARSWHCL NIFLDSKYLYHYLRTLAWGTFQGKSSQAPFQALLPRLLAHKVIYLHHVRSHTNLPDPISKLNA LTDALLITPIL
HTL32_Q0R5R2	GLEHLPPPPEVSQFPLNPERLQALTDLVSRALEAKHIEPYQGPGNNPIFPVKKPNGKWRFIHDL RATNSVTRDLASPSPGPPDLTSLPQGLPHLRTIDLTDAFFQIPLPTIFQPYFAFTLPQPNNYGPGT RYSWRVLPQGFKNSPTLFEQQLSHILTPVRKTFPNSLIIQYMDDILLASPAPGELAALTDKVTN ALTKEGLPLSPEKTQATPGPIHFLGQVISQDCITYETLPSINVKSTWSLAELQSMLGELQWVSK GTPVLRSSLHQLYLALRGHRDPRDTIKLTSIQVQALRTIQKALTLNCRSRLVNQLPILALIMLR PTGTTAVLFQTKQKWPLVWLHTPHPATSLRPWGQLLANAVIILDKYSLQHYGQVCKSFHHNI SNQALTYYLHTSDQSSVAILLQHSHRFHNLGAQPSGPWRSLLQMPQIFQNIDVLRPPFTISPVV INHAPCLFSDGSASKAAFIIWDRQVIHQQVLSLPSTCSAQAGELFGLLAGLQKSQPWVALNIFL DSKFLIGHLRRMALGAFPGPSTQCELHTQLLPLLQGKTVYVHHVRSHTLLQDPISRLNEATDA LMLAPLLPL

[0590]

HTL32_Q0R5R2_2mut	GLEHLPPPPEVSQFPLNPERLQALTDLVSRALEAKHIEPYQGPGNNPIFPVKKPNGKWRFIHDL RATNSVTRDLASPSPGPPDLTSLPQGLPHLRTIDLTDAFFQIPLPTIFQPYFAFTLPQPNNYGPGT RYSWRVLPQGFKNSPTLFQQQLSHILTPVRKTFPNSLIIQYMDDILLASPAPGELAALTDKVTN ALTKEGLPLSPEKTQATPGPIHFLGQVISQDCITYETLPSINVKSTWSLAELQSMLGELQWVSK GTPVLRSSLHQLYLALRGHRDPRDTIKLTSIQVQALRTIQKALTLNCRSRLVNQLPILALIMLR PTGTTAVLFQTKQKWPLVWLHTPHPATSLRPWGQLLANAVIILDKYSLQHYGQVCKSFHHNI SNQALTYYLHTSDQSSVAILLQHSHRFHNLGAQPSGPWRSLLQMPQIFQNIDVLRPPFTISPVV INHAPCLFSDGSASKAAFIIWDRQVIHQQVLSLPSTCSAQAGELFGLLAGLQKSQPWVALNIFL DSKFLIGHLRRMAWGAFPGPSTQCELHTQLLPLLQGKTVYVHHVRSHTLLQDPISRLNEATD ALMLAPLLPL
HTL32_Q0R5R2_2mutB	GLEHLPPPPEVSQFPLNPERLQALTDLVSRALEAKHIEPYQGPGNNPIFPVKKPNGKWRFIHDL RATNSVTRDLASPSPGPPDLTSPPQGLPHLRTIDLTDAFFQIPLPTIFQPYFAFTLPQPNNYGPGT RYSWRVLPQGFKNSPTLFQQQLSHILTPVRKTFPNSLIIQYMDDILLASPAPGELAALTDKVTN ALTKEGLPLSPEKTQATPGPIHFLGQVISQDCITYETLPSINVKSTWSLAELQSMLGELQWVSK GTPVLRSSLHQLYLALRGHRDPRDTIKLTSIQVQALRTIQKALTLNCRSRLVNQLPILALIMLR PTGTTAVLFQTKQKWPLVWLHTPHPATSLRPWGQLLANAVIILDKYSLQHYGQVCKSFHHNI SNQALTYYLHTSDQSSVAILLQHSHRFHNLGAQPSGPWRSLLQMPQIFQNIDVLRPPFTISPVV INHAPCLFSDGSASKAAFIIWDRQVIHQQVLSLPSTCSAQAGELFGLLAGLQKSQPWVALNIFL DSKFLIGHLRRMAWGAFPGPSTQCELHTQLLPLLQGKTVYVHHVRSHTLLQDPISRLNEATD ALMLAPLLPL
HTL3P_Q4U0X6	GLEHLPPPPEVSQFPLNPERLQALTDLVSRALEAKHIEPYQGPGNNPIFPVKKPNGKWRFIHDL RATNSLTRDLASPSPGPPDLTSLPQDLPHLRTIDLTDAFFQIPLPAVFQPYFAFTLPQPNNHGPG TRYSWRVLPQGFKNSPTLFEQQLSHILAPVRKAFPNSLIIQYMDDILLASPALRELTALTDKVT NALTKEGLPMSLEKTQATPGSIHFLGQVISPDCITYETLPSIHVKSIWSLAELQSMLGELQWVS KGTPVLRSSLHQLYLALRGHRDPRDTIELTSTQVQALKTIQKALALNCRSRLVSQLPILALIILR PTGTTAVLFQTKQKWPLVWLHTPHPATSLRPWGQLLANAIITLDKYSLQHYGQICKSFHHNIS NQALTYYLHTSDQSSVAILLQHSHRFHNLGAQPSGPWRSLLQVPQIFQNIDVLRPPFIISPVVID HAPCLFSDGATSKAAFILWDKQVIHQQVLPLPSTCSAQAGELFGLLAGLQKSKPWPALNIFLD SKFLIGHLRRMALGAFLGPSTQCDLHARLFPLLQGKTVYVHHVRSHTLLQDPISRLNEATDAL MLAPLLPL
HTL3P_Q4U0X6_2mut	GLEHLPPPPEVSQFPLNPERLQALTDLVSRALEAKHIEPYQGPGNNPIFPVKKPNGKWRFIHDL RATNSLTRDLASPSPGPPDLTSLPQDLPHLRTIDLTDAFFQIPLPAVFQPYFAFTLPQPNNHGPG TRYSWRVLPQGFKNSPTLFQQQLSHILAPVRKAFPNSLIIQYMDDILLASPALRELTALTDKVT NALTKEGLPMSLEKTQATPGSIHFLGQVISPDCITYETLPSIHVKSIWSLAELQSMLGELQWVS KGTPVLRSSLHQLYLALRGHRDPRDTIELTSTQVQALKTIQKALALNCRSRLVSQLPILALIILR PTGTTAVLFQTKQKWPLVWLHTPHPATSLRPWGQLLANAIITLDKYSLQHYGQICKSFHHNIS NQALTYYLHTSDQSSVAILLQHSHRFHNLGAQPSGPWRSLLQVPQIFQNIDVLRPPFIISPVVID HAPCLFSDGATSKAAFILWDKQVIHQQVLPLPSTCSAQAGELFGLLAGLQKSKPWPALNIFLD SKFLIGHLRRMAWGAFLGPSTQCDLHARLFPLLQGKTVYVHHVRSHTLLQDPISRLNEATDA LMLAPLLPL
HTL3P_Q4U0X6_2mutB	GLEHLPPPPEVSQFPLNPERLQALTDLVSRALEAKHIEPYQGPGNNPIFPVKKPNGKWRFIHDL RATNSLTRDLASPSPGPPDLTSPPQDLPHLRTIDLTDAFFQIPLPAVFQPYFAFTLPQPNNHGPG TRYSWRVLPQGFKNSPTLFQQQLSHILAPVRKAFPNSLIIQYMDDILLASPALRELTALTDKVT NALTKEGLPMSLEKTQATPGSIHFLGQVISPDCITYETLPSIHVKSIWSLAELQSMLGELQWVS KGTPVLRSSLHQLYLALRGHRDPRDTIELTSTQVQALKTIQKALALNCRSRLVSQLPILALIILR PTGTTAVLFQTKQKWPLVWLHTPHPATSLRPWGQLLANAIITLDKYSLQHYGQICKSFHHNIS NQALTYYLHTSDQSSVAILLQHSHRFHNLGAQPSGPWRSLLQVPQIFQNIDVLRPPFIISPVVID HAPCLFSDGATSKAAFILWDKQVIHQQVLPLPSTCSAQAGELFGLLAGLQKSKPWPALNIFLD SKFLIGHLRRMAWGAFLGPSTQCDLHARLFPLLQGKTVYVHHVRSHTLLQDPISRLNEATDA LMLAPLLPL
HTLV2_P03363_2mut	HLPPPPQVDQFPLNLPERLQALNDLVSKALEAGHIEPYSGPGNNPVFPVKKPNGKWRFIHDLR ATNAITTTLTSPSPGPPDLTSLPTALPHLQTIDLTDAFFQIPLPKQYQPYFAFTIPQPCNYGPGTR YAWTVLPQGFKNSPTLFQQQLAAVLNPMRKMFPTSTIVQYMDDILLASPTNEELQQLSQLTL QALTTHGLPISQEKTQQTPGQIRFLGQVISPNHITYESTPTIPIKSQWTLTELQVILGEIQWVSKG TPILRKHLQSLYSALHPYRDPRACTLTPQQLHALHAIQQALQHNCRGRLNPALPLLGLISLST SGTTSVIFQPKQNWPLAWLHTPHPPTSLCPWGHLLACTILTLDKYTLQHYGQLCQSFHHNMS

[0591]

	KQALCDFLRNSPHPSVGILIHHMGRFHNLGSQPSGPWKTLLHLPTLLQEPRLLRPIFTLSPVVL DTAPCLFSDGSPQKAAYVLWDQTILQQDITPLPSHETHSAQKGELLALICGLRAAKPWPSLNIF LDSKYLIKYLHSLAIGAFLGTSAHQTLQAALPPLLQGKTIYLHHVRSHTNLPDPISTFNEYTDS LILAPLVPL
JSRV_P31623	PLGTSDSPVTHADPIDWKSEEPVWVDQWPLTQEKLSAAQQLVQEQLRLGHIEPSTSAWNSPIF VIKKKSGKWRLLQDLRKVNETMMHMGALQPGLPTPSAIPDKSYIIVIDLKDCFYTIPLAPQDC KRFAFSLPSVNFKEPMQRYQWRVLPQGMTNSPTLCQKFVATAIAPVRQRFPQLYLVHYMDDI LLAHTDEHLLYQAFSILKQHLSLNGLVIADEKIQTHFPYNYLGFSLYPRVYNTQLVKLQTDHL KTLNDFQKLLGDINWIRPYLKLPTYTLQPLFDILKGDSDPASPRTLSLEGRTALQSIEEAIRQQQ ITYCDYQRSWGLYILPTPRAPTGVLYQDKPLRWIYLSATPTKHLLPYYELVAKIIAKGRHEAIQ YFGMEPPFICVPYALEQQDWLFQFSDNWSIAFANYPGQITHHYPSDKLLQFASSHAFIFPKIVR RQPIPEATLIFTDGSSNGTAALIINHQTYYAQTSFSSAQVVELFAVHQALLTVPTSFNLFTDSSY VVGALQMIETVPIIGTTSPEVLNLFTLIQQVLHCRQHPCFFGHIRAHSTLPGALVQGNHTADVL TKQVFFQS
JSRV_P31623_2mutB	PLGTSDSPVTHADPIDWKSEEPVWVDQWPLTQEKLSAAQQLVQEQLRLGHIEPSTSAWNSPIF VIKKKSGKWRLLQDLRKVNETMMHMGALQPGLPTPSPIPDKSYIIVIDLKDCFYTIPLAPQDC KRFAFSLPSVNFKEPMQRYQWRVLPQGMTNSPTLCQKFVATAIAPVRQRFPQLYLVHYMDDI LLAHTDEHLLYQAFSILKQHLSLNGLVIADEKIQTHFPYNYLGFSLYPRVYNTQLVKLQTDHL KTLNDFQKLLGDINWIRPYLKLPTYTLQPLFDILKGDSDPASPRTLSLEGRTALQSIEEAIRQQQ ITYCDYQRSWGLYILPTPRAPTGVLYQDKPLRWIYLSATPTKHLLPYYELVAKIIAKGRHEAIQ YFGMEPPFICVPYALEQQDWLFQFSDNWSIAFANYPGQITHHYPSDKLLQFASSHAFIFPKIVR RQPIPEATLIFTDGSSNGTAALIINHQTYYAQTSFSSAQVVELFAVHQALLTVPTSFNLFTDSSY VVGALQMIETVPIIGTTSPEVLNLFTLIQQVLHCRQHPCFFGHIRAHSTLPGALVQGNHTADVL TKQVFFQS
KORV_Q9TTC1	TLGDQGSRGSDPLPEPRVTLTVEGIPTEFLVNTGAEHSVLTKPMGKMGSKRTVVAGATGSKV YPWTTKRLLKIGQKQVTHSFLVIPECPAPLLGRDLLTKLKAQIQFSTEGPQVTWEDRPAMCLV LNLEEEYRLHEKPVPPSIDPSWLQLFPMVWAEKAGMGLANQVPPVVVELKSDASPVAVRQY PMSKEAREGIRPHIQRFLDLGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVQDIHP TVPNPYNLLSSLPPSHTWYSVLDLKDAFFCLKLHPNSQPLFAFEWRDPEKGNTGQLTWTRLP QGFKNSPTLFDEALHRDLASFRALNPQVVMLQYVDDLLVAAPTYRDCKEGTRRLLQELSKL GYRVSAKKAQLCREEVTYLGYLLKGGKRWLTPARKATVMKIPTPTTPRQVREFLGTAGFCR LWIPGFASLAAPLYPLTREKVPFTWTEAHQEAFGRIKEALLSAPALALPDLTKPFALYVDEKE GVARGVLTQTLGPWRRPVAYLSKKLDPVASGWPTCLKAIAAVALLLKDADKLTLGQNVLVI APHNLESIVRQPPDRWMTNARMTHYQSLLLNERVSFAPPAILNPATLLPVESDDTPIHICSEIL AEETGTRPDLRDQPLPGVPAWYTDGSSFIMDGRRQAGAAIVDNKRTVWASNLPEGTSAQKA ELIALTQALRLAEGKSINIYTDSRYAFATAHVHGAIYKQRGLLTSAGKDIKNKEEILALLEAIH LPKRVAIIHCPGHQRGTDPVATGNRKADEAAKQAAQSTRILTETTKN
KORV_Q9TTC1_3mut	TLGDQGSRGSDPLPEPRVTLTVEGIPTEFLVNTGAEHSVLTKPMGKMGSKRTVVAGATGSKV YPWTTKRLLKIGQKQVTHSFLVIPECPAPLLGRDLLTKLKAQIQFSTEGPQVTWEDRPAMCLV LNLEEEYRLHEKPVPPSIDPSWLQLFPMVWAEKAGMGLANQVPPVVVELKSDASPVAVRQY PMSKEAREGIRPHIQRFLDLGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVQDIHP TVPNPYNLLSSLPPSHTWYSVLDLKDAFFCLKLHPNSQPLFAFEWRDPEKGNTGQLTWTRLP QGFKNSPTLFNEALHRDLASFRALNPQVVMLQYVDDLLVAAPTYRDCKEGTRRLLQELSKL GYRVSAKKAQLCREEVTYLGYLLKGGKRWLTPARKATVMKIPTPTTPRQVREFLGTAGFCR LWIPGFASLAAPLYPLTRPKVPFTWTEAHQEAFGRIKEALLSAPALALPDLTKPFALYVDEKE GVARGVLTQTLGPWRRPVAYLSKKLDPVASGWPTCLKAIAAVALLLKDADKLTLGQNVLVI APHNLESIVRQPPDRWMTNARMTHYQSLLLNERVSFAPPAILNPATLLPVESDDTPIHICSEIL AEETGTRPDLRDQPLPGVPAWYTDGSSFIMDGRRQAGAAIVDNKRTVWASNLPEGTSAQKA ELIALTQALRLAEGKSINIYTDSRYAFATAHVHGAIYKQRGWLTSAGKDIKNKEEILALLEAIH LPKRVAIIHCPGHQRGTDPVATGNRKADEAAKQAAQSTRILTETTKN
KORV_Q9TTC1_3mutA	TLGDQGSRGSDPLPEPRVTLTVEGIPTEFLVNTGAEHSVLTKPMGKMGSKRTVVAGATGSKV YPWTTKRLLKIGQKQVTHSFLVIPECPAPLLGRDLLTKLKAQIQFSTEGPQVTWEDRPAMCLV LNLEEEYRLHEKPVPPSIDPSWLQLFPMVWAEKAGMGLANQVPPVVVELKSDASPVAVRQY PMSKEAREGIRPHIQRFLDLGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVQDIHP TVPNPYNLLSSLPPSHTWYSVLDLKDAFFCLKLHPNSQPLFAFEWRDPEKGNTGQLTWTRLP QGFKNSPTLFNEALHRDLASFRALNPQVVMLQYVDDLLVAAPTYRDCKEGTRRLLQELSKL

[0592]

	GYRVSAKKAQLCREEVTYLGYLLKGGKRWLTPARKATVMKIPTPTTPRQVREFLGKAGFCR LFIPGFASLAAPLYPLTRPKVPFTWTEAHQEAFGRIKEALLSAPALALPDLTKPFALYVDEKEG VARGVLTQTLGPWRRPVAYLSKKLDPVASGWPTCLKAIAAVALLLKDADKLTLGQNVLVIA PHNLESIVRQPPDRWMTNARMTHYQSLLLNERVSFAPPAILNPATLLPVESDDTPIHICSEILAE ETGTRPDLRDQPLPGVPAWYTDGSSFIMDGRRQAGAAIVDNKRTVWASNLPEGTSAQKAELI ALTQALRLAEGKSINIYTDSRYAFATAHVHGAIYKQRGWLTSAGKDIKNKEEILALLEAIHLP KRVAIIHCPGHQRGTDPVATGNRKADEAAKQAAQSTRILTETTKN
KORV_Q9TTC1-Pro	LLGRDLLTKLKAQIQFSTEGPQVTWEDRPAMCLVLNLEEEYRLHEKPVPPSIDPSWLQLFPMV WAEKAGMGLANQVPPVVVELKSDASPVAVRQYPMSKEAREGIRPHIQRFLDLGILVPCQSPW NTPLLPVKKPGTNDYRPVQDLREVNKRVQDIHPTVPNPYNLLSSLPPSHTWYSVLDLKDAFF CLKLHPNSQPLFAFEWRDPEKGNTGQLTWTRLPQGFKNSPTLFDEALHRDLASFRALNPQVV MLQYVDDLLVAAPTYRDCKEGTRRLLQELSKLGYRVSAKKAQLCREEVTYLGYLLKGGKR WLTPARKATVMKIPTPTTPRQVREFLGTAGFCRLWIPGFASLAAPLYPLTREKVPFTWTEAHQ EAFGRIKEALLSAPALALPDLTKPFALYVDEKEGVARGVLTQTLGPWRRPVAYLSKKLDPVA SGWPTCLKAIAAVALLLKDADKLTLGQNVLVIAPHNLESIVRQPPDRWMTNARMTHYQSLL LNERVSFAPPAILNPATLLPVESDDTPIHICSEILAEETGTRPDLRDQPLPGVPAWYTDGSSFIM DGRRQAGAAIVDNKRTVWASNLPEGTSAQKAELIALTQALRLAEGKSINIYTDSRYAFATAH VHGAIYKQRGLLTSAGKDIKNKEEILALLEAIHLPKRVAIIHCPGHQRGTDPVATGNRKADEA AKQAAQSTRILTETTKN
KORV_Q9TTC1-Pro_3mut	LLGRDLLTKLKAQIQFSTEGPQVTWEDRPAMCLVLNLEEEYRLHEKPVPPSIDPSWLQLFPMV WAEKAGMGLANQVPPVVVELKSDASPVAVRQYPMSKEAREGIRPHIQRFLDLGILVPCQSPW NTPLLPVKKPGTNDYRPVQDLREVNKRVQDIHPTVPNPYNLLSSLPPSHTWYSVLDLKDAFF CLKLHPNSQPLFAFEWRDPEKGNTGQLTWTRLPQGFKNSPTLFNEALHRDLASFRALNPQVV MLQYVDDLLVAAPTYRDCKEGTRRLLQELSKLGYRVSAKKAQLCREEVTYLGYLLKGGKR WLTPARKATVMKIPTPTTPRQVREFLGTAGFCRLWIPGFASLAAPLYPLTRPKVPFTWTEAHQ EAFGRIKEALLSAPALALPDLTKPFALYVDEKEGVARGVLTQTLGPWRRPVAYLSKKLDPVA SGWPTCLKAIAAVALLLKDADKLTLGQNVLVIAPHNLESIVRQPPDRWMTNARMTHYQSLL LNERVSFAPPAILNPATLLPVESDDTPIHICSEILAEETGTRPDLRDQPLPGVPAWYTDGSSFIM DGRRQAGAAIVDNKRTVWASNLPEGTSAQKAELIALTQALRLAEGKSINIYTDSRYAFATAH VHGAIYKQRGWLTSAGKDIKNKEEILALLEAIHLPKRVAIIHCPGHQRGTDPVATGNRKADEA AKQAAQSTRILTETTKN
KORV_Q9TTC1-Pro_3mutA	LLGRDLLTKLKAQIQFSTEGPQVTWEDRPAMCLVLNLEEEYRLHEKPVPPSIDPSWLQLFPMV WAEKAGMGLANQVPPVVVELKSDASPVAVRQYPMSKEAREGIRPHIQRFLDLGILVPCQSPW NTPLLPVKKPGTNDYRPVQDLREVNKRVQDIHPTVPNPYNLLSSLPPSHTWYSVLDLKDAFF CLKLHPNSQPLFAFEWRDPEKGNTGQLTWTRLPQGFKNSPTLFNEALHRDLASFRALNPQVV MLQYVDDLLVAAPTYRDCKEGTRRLLQELSKLGYRVSAKKAQLCREEVTYLGYLLKGGKR WLTPARKATVMKIPTPTTPRQVREFLGKAGFCRLFIPGFASLAAPLYPLTRPKVPFTWTEAHQ EAFGRIKEALLSAPALALPDLTKPFALYVDEKEGVARGVLTQTLGPWRRPVAYLSKKLDPVA SGWPTCLKAIAAVALLLKDADKLTLGQNVLVIAPHNLESIVRQPPDRWMTNARMTHYQSLL LNERVSFAPPAILNPATLLPVESDDTPIHICSEILAEETGTRPDLRDQPLPGVPAWYTDGSSFIM DGRRQAGAAIVDNKRTVWASNLPEGTSAQKAELIALTQALRLAEGKSINIYTDSRYAFATAH VHGAIYKQRGWLTSAGKDIKNKEEILALLEAIHLPKRVAIIHCPGHQRGTDPVATGNRKADEA AKQAAQSTRILTETTKN
MLVAV_P03356	TLNLEDEYRLYETSAEPEVSPGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEAKLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHRWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGMGISGQLTWTRLP QGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDILLAATSELDCQQGTRALLLTLGNLGY RASAKKAQLCQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRL WIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQ GYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLRKDAGKLTMGQPLVI LAPHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDC LEILAETHGTRPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWARALPAGTS AQRAELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGREIKNKSEILAL LKALFLPKRLSIIHCLGHQKGDSAEARGNRLADQAAREAAIKTPPDTSTLL
MLVAV_P03356	TLNLEDEYRLYETSAEPEVSPGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEAKLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH

[0593]

56_3mut	PTVPNPYNLLSGLPPSHRWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDILLAATSELDCQQGTRALLLTLGNLGY RASAKKAQLCQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRL WIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQ GYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLRKDAGKLTMGQPLVI LAPHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDC LEILAETHGTRPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWARALPAGTS AQRAELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGREIKNKSEILAL LKALFLPKRLSIIHCLGHQKGDSAEARGNRLADQAARE AAIKTPPDTSTLL
MLVAV_P033 56_3mut A	TLNLEDEYRLYETSAEPEVSPGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEAKLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHRWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDILLAATSELDCQQGTRALLLTLGNLGY RASAKKAQLCQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLF IPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLRKDAGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDCL EILAETHGTRPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWARALPAGTSA QRAELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGREIKNKSEILALL KALFLPKRLSIIHCLGHQKGDSAEARGNRLADQAAREAAIKTPPDTSTLL
MLVBM_Q7SVK7	TLGIEDEYRLHETSTEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIQQY PMSHEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGMGISGQLTWTRLP QGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDILLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLREGQRWLTEARKETVMGQPVPKTPRQLREFLGTAGFCRLW IPGFAEMAAPLYPLTKTGTLFSWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDCLE ILAETHGTRPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWAGALPAGTSAQ RAELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGREIKNKSEILALLK ALFLPKRLSIIHCLGHQKGDSAEARGNRLADQAAREAAIKTPPDTSTLL
MLVBM_Q7SVK7	TLGIEDEYRLHETSTEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIQQY PMSHEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGMGISGQLTWTRLP QGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDILLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLREGQRWLTEARKETVMGQPVPKTPRQLREFLGTAGFCRLW IPGFAEMAAPLYPLTKPGTLFSWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDCLE ILAETHGTRPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWAGALPAGTSAQ RAELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGREIKNKSEILALLK ALFLPKRLSIIHCLGHQKGDSAEARGNRLADQAAREAAIKTPPDTSTLL
MLVBM_Q7SVK7_3mut	TLGIEDEYRLHETSTEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIQQY PMSHEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDILLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLREGQRWLTEARKETVMGQPVPKTPRQLREFLGTAGFCRLW IPGFAEMAAPLYPLTKPGTLFSWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDCLE ILAETHGTRPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWAGALPAGTSAQ RAELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGREIKNKSEILALLK ALFLPKRLSIIHCLGHQKGDSAEARGNRLADQAAREAAIKTPPDTSTLL
MLVBM_Q7S	TLGIEDEYRLHETSTEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIQQY PMSHEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGMGISGQLTWTRLP

[0594]

VK7_3 mut	QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDILLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLREGQRWLTEARKETVMGQPVPKTPRQLREFLGTAGFCRLW IPGFAEMAAPLYPLTKPGTLFSWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDCLE ILAETHGTRPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWAGALPAGTSAQ RAELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGREIKNKSEILALLK ALFLPKRLSIIHCLGHQKGDSAEARGNRLADQAARE AAIKTPPDTSTLL
MLVB M_Q7S VK7_3 mutA_ WS	LGIEDEYRLHETSTEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIQQYP MSHEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPT VPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGMGISGQLTWTRLPQ GFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDILLAATSELDCQQGTRALLQTLGDLGYR ASAKKAQICQKQVKYLGYLLREGQRWLTEARKETVMGQPVPKTPRQLREFLGKAGFCRLFIP GFAEMAAPLYPLTKPGTLFSWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYA KGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAP HAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDCLEI LAETHGTRPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWAGALPAGTSAQR AELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGREIKNKSEILALLKA LFLPKRLSIIHCLGHQKGDSAEARGNRLADQAAREAAIKTPPDTSTLLI
MLVB M_Q7S VK7_3 mutA_ WS	LGIEDEYRLHETSTEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIQQYP MSHEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPT VPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGMGISGQLTWTRLPQ GFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDILLAATSELDCQQGTRALLQTLGDLGYR ASAKKAQICQKQVKYLGYLLREGQRWLTEARKETVMGQPVPKTPRQLREFLGKAGFCRLFIP GFAEMAAPLYPLTKPGTLFSWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYA KGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAP HAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDCLEI LAETHGTRPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWAGALPAGTSAQR AELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGREIKNKSEILALLKA LFLPKRLSIIHCLGHQKGDSAEARGNRLADQAAREAAIKTPPDTSTLLI
MLVC B_P083 61	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFDEALHRDLAGFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPIPKTPRQLREFLGTAGFCRLWI PGFAEMAAPLYPLTKTGTLFNWGPDQQKAFQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHDCLD ILAEAHGTRSDLMDQPLPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWARALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGNSAEARGNRMADQAAREVATRETPETSTLL
MLVC B_P083 61_3mu t	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLAGFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPIPKTPRQLREFLGTAGFCRLWI PGFAEMAAPLYPLTKPGTLFNWGPDQQKAFQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHDCLD ILAEAHGTRSDLMDQPLPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWARALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGNSAEARGNRMADQAAREVATRETPETSTLL
MLVC B_P083 61_3mu tA	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLAGFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGY

[0595]

	RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPIPKTPRQLREFLGKAGFCRLFI PGFAEMAAPLYPLTKPGTLFNWGPDQQKAFQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHDCLD ILAEAHGTRSDLMDQPLPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWARALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGNSAEARGNRMADQAARFVATRETPETSTLL
MLVF5_P26810	TLNIEDEYRLHETSKGPDVPLGSTWLSDFPQAWAETGGMGLAFRQAPLIISLKATSTPVSIKQY PMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQSLFAFEWKDPEMGISGQLTWTRLP QGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGLCRLW IPGFAEMAAPLYPLTKTGTLFKWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDVGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPIVALNPATLLPLPEEGLQHDCL DILAEAHGTRPDLTDQPLPDADHTWYTDGSSFLQEGQRRAGAAVTTETEVIWAKALPAGTSA QRAELIALTQALKMAAGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALL KALFLPKRLSIIHCPGHQKGNHAEARGNRMADQAARFVATRETPETSTLL
MLVF5_P26810_3mut	TLNIEDEYRLHETSKGPDVPLGSTWLSDFPQAWAETGGMGLAFRQAPLIISLKATSTPVSIKQY PMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQSLFAFEWKDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGLCRLW IPGFAEMAAPLYPLTKPGTLFKWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDVGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPIVALNPATLLPLPEEGLQHDCL DILAEAHGTRPDLTDQPLPDADHTWYTDGSSFLQEGQRRAGAAVTTETEVIWAKALPAGTSA QRAELIALTQALKMAAGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGNHAEARGNRMADQAARFVATRETPETSTLL
MLVF5_P26810_3mutA	TLNIEDEYRLHETSKGPDVPLGSTWLSDFPQAWAETGGMGLAFRQAPLIISLKATSTPVSIKQY PMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQSLFAFEWKDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGLCRLF IPGFAEMAAPLYPLTKPGTLFKWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDVGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPIVALNPATLLPLPEEGLQHDCL DILAEAHGTRPDLTDQPLPDADHTWYTDGSSFLQEGQRRAGAAVTTETEVIWAKALPAGTSA QRAELIALTQALKMAAGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGNHAEARGNRMADQAARFVATRETPETSTLL
MLVFF_P26809_3mut	TLNIEDEYRLHETSKGPDVPLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQSLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLW IPGFAEMAAPLYPLTKPGTLFEWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPIVALNPATLLPLPEEGLQHDCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVVWAKALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGNRAEARGNRMADQAARFVATRETPETSTLL
MLVFF_P26809_3mutA	TLNIEDEYRLHETSKGPDVPLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQSLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGDLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFI

[0596]

	PGFAEMAAPLYPLTKPGTLFEWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPIVALNPATLLPLPEEGLQHDCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVVWAKALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGNRAEARGNRMADQAAREVATRETPETSTLL
MLVM S_P033 55	TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLW IPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCL DILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL
MLVM S_P033 55	TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLW IPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCL DILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL
MLVM S_P033 55_3mu t	TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLW IPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCL DILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILA LLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL
MLVM S_P033 55_3mu t	TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLW IPGFAEMAAPLYPLTKPGTLFEWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCL DILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILA LLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL
MLVM S_P033 55_3mu tA_WS	TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFI PGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY

[0597]

	AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMfTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLD ILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL
MLVMS_P03355_3mutA_WS	TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFI PGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMfTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLD ILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL
MLVMS_P03355_PLV919	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFI PGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMfTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLD ILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRT ADGSEFE
MLVMS_P03355_PLV919	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGY RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFI PGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA PHAVEALVKQPPDRWLSNARMfTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLD ILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRT ADGSEFE
MLVRD_P11227	TLNIEDEYRLHEISTEPDVSPGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQY PMSQEAKLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQGLREVNKRVEDIHP TVPNPYNLLSGLPTSHRWYTVLDLKDAFFCLRLHPTSQPLFASEWRDPGMGISGQLTWTRLP QGFKNSPTLFDEALHRGLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLKTLGNLGY RASAKKAQICQKQVKYLGYLLREGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLW IPRFAEMAAPLYPLTKTGTLFNWGPDQQKAYHEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDCL EILAETHGTEPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWARALPAGTSA QRAELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYKRRGLLTSEGREIKNKSEILALL KALFLPKRLSIIHCLGHQKGDSAEARGNRLADQAAREAAIKTPPDTSTLL
MLVRD_P11227_3mut	TLNIEDEYRLHEISTEPDVSPGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQY PMSQEAKLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQGLREVNKRVEDIHP TVPNPYNLLSGLPTSHRWYTVLDLKDAFFCLRLHPTSQPLFASEWRDPGMGISGQLTWTRLP QGFKNSPTLFNEALHRGLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLKTLGNLGY RASAKKAQICQKQVKYLGYLLREGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLW

[0598]

	IPRFAEMAAPLYPLTKPGTLFNWGPDQQKAYHEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEEGAPHDCL EILAETHGTEPDLTDQPIPDADHTWYTDGSSFLQEGQRKAGAAVTTETEVIWARALPAGTSA QRAELIALTQALKMAEGKRLNVYTDSRYAFATAHIHGEIYKRRGWLTSEGREIKNKSEILALL KALFLPKRLSIIHCLGHQKGDSAEARGNRLADQAAREAAIKTPPDTSTLL
MMTV_P03365	WVQEISDSRPMLHIYLNGRRFLGLLNTGADKTCIAGRDWPANWPIHQTESSLQGLGMACGV ARSSQPLRWQHEDKSGIIHPFVIPTLPFTLWGRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLF ADQISWKSDQPVWLNQWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKKSGK WRLLQDLRAVNATMHDMGALQPGLPSPVAVPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVP SPNFKRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYIVHYMDDILLAHPSRS IVDEILTSMIQALNKHGLVVSTEKIQKYDNLKYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQK LLGNINWIRPFLKLTTGELKPLFEILNGDSNPISTRKLTPEACKALQLMNERLSTARVKRLDLS QPWSLCILKTEYTPTACLWQDGVVEWIHLPHISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDY IVVPYTKVQFDLLLQEKEDWPISLLGFLGEVHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVI FTDGSANGRSVTYIQGREPIIKENTQNTAQQAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIE TATLSPRTKIYTELKHLQRLIHKRQEKFYIGHIRGHTGLPGPLAQGNAYADSLTRILT
MMTV_P03365_2mut	WVQEISDSRPMLHIYLNGRRFLGLLNTGADKTCIAGRDWPANWPIHQTESSLQGLGMACGV ARSSQPLRWQHEDKSGIIHPFVIPTLPFTLWGRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLF ADQISWKSDQPVWLNQWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKKSGK WRLLQDLRAVNATMHDMGALQPGLPSPVAVPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVP SPNFKRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYIVHYMDDILLAHPSRS IVDEILTSMIQALNKHGLVVSTEKIQKYDNLKYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQK LLGNINWIRPFLKLTTGELKPLFEILNPDSNPISTRKLTPEACKALQLMNERLSTARVKRLDLS QPWSLCILKTEYTPTACLWQDGVVEWIHLPHISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDY IVVPYTKVQFDLLLQEKEDWPISLLGFLGEVHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVI FTDGSANGRSVTYIQGREPIIKENTQNTAQQAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIE TATLSPRTKIYTELKHLQRLIHKRQEKFYIGHIRGHTGLPGPLAQGNAYADSLTRILT
MMTV_P03365_2mut_WS	VQEISDSRPMLHIYLNGRRFLGLLDTGADKTCIAGRDWPANWPIHQTESSLQGLGMACGVAR SSQPLRWQHEDKSGIIHPFVIPTLPFTLWGRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFAD QISWKSDQPVWLNQWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRL LQDLRAVNATMHDMGALQPGLPSPVAVPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNF KRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDE ILTSMIQALNKHGLVVSTEKIQKYDNLKYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGN INWIRPFLKLTTGELKPLFEILNPDSNPISTRKLTPEACKALQLMNERLSTARVKRLDLSQPWS LCILKTEYTPTACLWQDGVVEWIHLPHISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVP YTKVQFDLLLQEKEDWPISLLGFLGEVHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTD GSANGRSVTYIQGREPIIKENTQNTAQQAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETAT LSPRTKIYTELKHLQRLIHKRQEKFYIGHIRGHTGLPGPLAQGNAYADSLTRILTA
MMTV_P03365_2mut_WS	VQEISDSRPMLHIYLNGRRFLGLLDTGADKTCIAGRDWPANWPIHQTESSLQGLGMACGVAR SSQPLRWQHEDKSGIIHPFVIPTLPFTLWGRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFAD QISWKSDQPVWLNQWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRL LQDLRAVNATMHDMGALQPGLPSPVAVPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNF KRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDE ILTSMIQALNKHGLVVSTEKIQKYDNLKYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGN

[0599]

	INWIRPFLKLTTGELKPLFEILNPDSNPISTRKLTPEACKALQLMNERLSTARVKRLDLSQPWS LCILKTEYTPTACLWQDGVVEWIHLPHISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVP YTKVQFDLLLQEKEDWPISLLGFLGEVHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTD GSANGRSVTYIQGREPIIKENTQNTAQQAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETAT LSPRTKIYTELKHLQRLIHKRQEKFYIGHIRGHTGLPGPLAQGNAYADSLTRILTA
MMTV B_P033 65_2mu tB	WVQEISDSRPMLHIYLNGRRFLGLLNTGADKTCIAGRDWPANWPIHQTESSLQGLGMACGV ARSSQPLRWQHEDKSGIIHPFVIPTLPFTLWGRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLF ADQISWKSDQPVWLNQWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKKSGK WRLLQDLRAVNATMHDMGALQPGLPSPVAPPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVP SPNFKRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYIVHYMDDILLAHPSRS IVDEILTSMIQALNKHGLVVSTEKIQKYDNLKYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQK LLGNINWIRPFLKLTTGELKPLFEILNPDSNPISTRKLTPEACKALQLMNERLSTARVKRLDLS QPWSLCILKTEYTPTACLWQDGVVEWIHLPHISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDY IVVPYTKVQFDLLLQEKEDWPISLLGFLGEVHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVI FTDGSANGRSVTYIQGREPIIKENTQNTAQQAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIE TATLSPRTKIYTELKHLQRLIHKRQEKFYIGHIRGHTGLPGPLAQGNAYADSLTRILT
MMTV B_P033 65_2mu tB	WVQEISDSRPMLHIYLNGRRFLGLLNTGADKTCIAGRDWPANWPIHQTESSLQGLGMACGV ARSSQPLRWQHEDKSGIIHPFVIPTLPFTLWGRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLF ADQISWKSDQPVWLNQWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKKSGK WRLLQDLRAVNATMHDMGALQPGLPSPVAPPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVP SPNFKRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYIVHYMDDILLAHPSRS IVDEILTSMIQALNKHGLVVSTEKIQKYDNLKYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQK LLGNINWIRPFLKLTTGELKPLFEILNPDSNPISTRKLTPEACKALQLMNERLSTARVKRLDLS QPWSLCILKTEYTPTACLWQDGVVEWIHLPHISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDY IVVPYTKVQFDLLLQEKEDWPISLLGFLGEVHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVI FTDGSANGRSVTYIQGREPIIKENTQNTAQQAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIE TATLSPRTKIYTELKHLQRLIHKRQEKFYIGHIRGHTGLPGPLAQGNAYADSLTRILT
MMTV B_P033 65_2mu tB_WS	VQEISDSRPMLHIYLNGRRFLGLLDTGADKTCIAGRDWPANWPIHQTESSLQGLGMACGVAR SSQPLRWQHEDKSGIIHPFVIPTLPFTLWGRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFAD QISWKSDQPVWLNQWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRL LQDLRAVNATMHDMGALQPGLPSPPAVPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNF KRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDE ILTSMIQALNKHGLVVSTEKIQKYDNLKYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGN INWIRPFLKLTTGELKPLFEILNPDSNPISTRKLTPEACKALQLMNERLSTARVKRLDLSQPWS LCILKTEYTPTACLWQDGVVEWIHLPHISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVP YTKVQFDLLLQEKEDWPISLLGFLGEVHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTD GSANGRSVTYIQGREPIIKENTQNTAQQAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETAT LSPRTKIYTELKHLQRLIHKRQEKFYIGHIRGHTGLPGPLAQGNAYADSLTRILTA
MMTV B_P033 65_2mu tB_WS	VQEISDSRPMLHIYLNGRRFLGLLDTGADKTCIAGRDWPANWPIHQTESSLQGLGMACGVAR SSQPLRWQHEDKSGIIHPFVIPTLPFTLWGRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFAD QISWKSDQPVWLNQWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRL LQDLRAVNATMHDMGALQPGLPSPVAVPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNF KRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDE ILTSMIQALNKHGLVVSTEKIQKYDNLKYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGN INWIRPFLKLTTGELKPLFEILNGDSNPISTRKLTPEACKALQLMNERLSTARVKRLDLSQPWS

[0600]

	LCILKTEYTPTACLWQDGVVEWIHLPHISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVP YTKVQFDLLLQEKEDWPISLLGFLGEVHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTD GSANGRSVTYIQGREPIIKENTQNTAQQAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETAT LSPRTKIYTELKHLQRLIHKRQEKFYIGHIRGHTGLPGPLAQGNAYADSLTRILTA
MMTV B_P03365_WS	VQEISDSRPMLHIYLNGRRFLGLLDTGADKTCIAGRDWPANWPIHQTESSLQGLGMACGVAR SSQPLRWQHEDKSGIIHPFVIPTLPFTLWGRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFAD QISWKSDQPVWLNQWPLKQEKLQALQQLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRL LQDLRAVNATMHDMGALQPGLPSPVAVPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNF KRPYQRFQWKVLPQGMKNSPTLCQKFVDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDE ILTSMIQALNKHGLVVSTEKIQKYDNLKYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGN INWIRPFLKLTTGELKPLFEILNGDSNPISTRKLTPEACKALQLMNERLSTARVKRLDLSQPWS LCILKTEYTPTACLWQDGVVEWIHLPHISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVP YTKVQFDLLLQEKEDWPISLLGFLGEVHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTD GSANGRSVTYIQGREPIIKENTQNTAQQAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETAT LSPRTKIYTELKHLQRLIHKRQEKFYIGHIRGHTGLPGPLAQGNAYADSLTRILTA
MMTV B_P03365-Pro	GRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFADQISWKSDQPVWLNQWPLKQEKLQALQ QLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRLLQDLRAVNATMHDMGALQPGLPSPVA VPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNFKRPYQRFQWKVLPQGMKNSPTLCQKF VDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDEILTSMIQALNKHGLVVSTEKIQKYDNL KYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGNINWIRPFLKLTTGELKPLFEILNGDSNP ISTRKLTPEACKALQLMNERLSTARVKRLDLSQPWSLCILKTEYTPTACLWQDGVVEWIHLP HISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVPYTKVQFDLLLQEKEDWPISLLGFLGE VHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTDGSANGRSVTYIQGREPIIKENTQNTAQ QAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETATLSPRTKIYTELKHLQRLIHKRQEKFYI GHIRGHTGLPGPLAQGNAYADSLTRILT
MMTV B_P03365-Pro	GRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFADQISWKSDQPVWLNQWPLKQEKLQALQ QLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRLLQDLRAVNATMHDMGALQPGLPSPVA VPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNFKRPYQRFQWKVLPQGMKNSPTLCQKF VDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDEILTSMIQALNKHGLVVSTEKIQKYDNL KYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGNINWIRPFLKLTTGELKPLFEILNGDSNP ISTRKLTPEACKALQLMNERLSTARVKRLDLSQPWSLCILKTEYTPTACLWQDGVVEWIHLP HISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVPYTKVQFDLLLQEKEDWPISLLGFLGE VHFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTDGSANGRSVTYIQGREPIIKENTQNTAQ QAEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETATLSPRTKIYTELKHLQRLIHKRQEKFYI GHIRGHTGLPGPLAQGNAYADSLTRILT
MMTV B_P03365-Pro_2mut	GRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFADQISWKSDQPVWLNQWPLKQEKLQALQ QLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRLLQDLRAVNATMHDMGALQPGLPSPVA VPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNFKRPYQRFQWKVLPQGMKNSPTLCQKF VDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDEILTSMIQALNKHGLVVSTEKIQKYDNL KYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGNINWIRPFLKLTTGELKPLFEILNPDSNPI STRKLTPEACKALQLMNERLSTARVKRLDLSQPWSLCILKTEYTPTACLWQDGVVEWIHLPH ISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVPYTKVQFDLLLQEKEDWPISLLGFLGEV HFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTDGSANGRSVTYIQGREPIIKENTQNTAQQ AEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETATLSPRTKIYTELKHLQRLIHKRQEKFYIG HIRGHTGLPGPLAQGNAYADSLTRILT
MMTV B_P03365-Pro_2mut	GRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFADQISWKSDQPVWLNQWPLKQEKLQALQ QLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRLLQDLRAVNATMHDMGALQPGLPSPVA VPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNFKRPYQRFQWKVLPQGMKNSPTLCQKF VDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDEILTSMIQALNKHGLVVSTEKIQKYDNL KYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGNINWIRPFLKLTTGELKPLFEILNPDSNPI STRKLTPEACKALQLMNERLSTARVKRLDLSQPWSLCILKTEYTPTACLWQDGVVEWIHLPH ISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVPYTKVQFDLLLQEKEDWPISLLGFLGEV HFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTDGSANGRSVTYIQGREPIIKENTQNTAQQ AEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETATLSPRTKIYTELKHLQRLIHKRQEKFYIG HIRGHTGLPGPLAQGNAYADSLTRILT

[0601]

MMTV_B_P03365-Pro_2mutB	GRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFADQISWKSDQPVWLNQWPLKQEKLQALQ QLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRLLQDLRAVNATMHDMGALQPGLPSPVA PPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNFKRPYQRFQWKVLPQGMKNSPTLCQKF VDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDEILTSMIQALNKHGLVVSTEKIQKYDNL KYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGNINWIRPFLKLTTGELKPLFEILNPDSNPI STRKLTPEACKALQLMNERLSTARVKRLDLSQPWSLCILKTEYTPTACLWQDGVVEWIHLPH ISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVPYTKVQFDLLLQEKEDWPISLLGFLGEV HFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTDGSANGRSVTYIQGREPIIKENTQNTAQQ AEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETATLSPRTKIYTELKHLQRLIHKRQEKFYIG HIRGHTGLPGPLAQGNAYADSLTRILT
MMTV_B_P03365-Pro_2mutB	GRDIMKDIKVRLMTDSPDDSQDLMIGAIESNLFADQISWKSDQPVWLNQWPLKQEKLQALQ QLVTEQLQLGHLEESNSPWNTPVFVIKKKSGKWRLLQDLRAVNATMHDMGALQPGLPSPVA PPKGWEIIIIDLQDCFFNIKLHPEDCKRFAFSVPSPNFKRPYQRFQWKVLPQGMKNSPTLCQKF VDKAILTVRDKYQDSYIVHYMDDILLAHPSRSIVDEILTSMIQALNKHGLVVSTEKIQKYDNL KYLGTHIQGDSVSYQKLQIRTDKLRTLNDFQKLLGNINWIRPFLKLTTGELKPLFEILNPDSNPI STRKLTPEACKALQLMNERLSTARVKRLDLSQPWSLCILKTEYTPTACLWQDGVVEWIHLPH ISPKVITPYDIFCTQLIIKGRHRSKELFSKDPDYIVVPYTKVQFDLLLQEKEDWPISLLGFLGEV HFHLPKDPLLTFTLQTAIIFPHMTSTTPLEKGIVIFTDGSANGRSVTYIQGREPIIKENTQNTAQQ AEIVAVITAFEEVSQPFNLYTDSKYVTGLFPEIETATLSPRTKIYTELKHLQRLIHKRQEKFYIG HIRGHTGLPGPLAQGNAYADSLTRILT
MPMV_P07572	LTAAIDILAPQQCAEPITWKSDEPVWVDQWPLTNDKLAAAQQLVQEQLEAGHITESSSPWNT PIFVIKKKSGKWRLLQDLRAVNATMVLMGALQPGLPSPVAIPQGYLKIIIDLKDCFFSIPLHPS DQKRFAFSLPSTNFKEPMQRFQWKVLPQGMANSPTLCQKYVATAIHKVRHAWKQMYIIHY MDDILIAGKDGQQVLQCFDQLKQELTAAGLHIAPEKVQLQDPYTYLGFELNGPKITNQKAVI RKDKLQTLNDFQKLLGDINWLRPYLKLTTGDLKPLFDTLKGDSDPNSHRSLSKEALASLEKV ETAIAEQFVTHINYSLPLIFLIFNTALTPTGLFWQDNPIMWIHLPASPKKVLLPYYDAIADLIILG RDHSKKYFGIEPSTIIQPYSKSQIDWLMQNTEMWPIACASFVGILDNHYPPNKLIQFCKLHTFV FPQIISKTPLNNALLVFTDGSSTGMAAYTLTDTTIKFQTNLNSAQLVELQALIAVLSAFPNQPL NIYTDSAYLAHSIPLLETVAQIKHISETAKLFLQCQQLIYNRSIPFYIGHVRAHSGLPGPIAQGN QRADLATKIVASNINT
MPMV_P07572_2mutB	LTAAIDILAPQQCAEPITWKSDEPVWVDQWPLTNDKLAAAQQLVQEQLEAGHITESSSPWNT PIFVIKKKSGKWRLLQDLRAVNATMVLMGALQPGLPSPVAPPQGYLKIIIDLKDCFFSIPLHPS DQKRFAFSLPSTNFKEPMQRFQWKVLPQGMANSPTLCQKYVATAIHKVRHAWKQMYIIHY MDDILIAGKDGQQVLQCFDQLKQELTAAGLHIAPEKVQLQDPYTYLGFELNGPKITNQKAVI RKDKLQTLNDFQKLLGDINWLRPYLKLTTGDLKPLFDTLKPDSDPNSHRSLSKEALASLEKVE TAIAEQFVTHINYSLPLIFLIFNTALTPTGLFWQDNPIMWIHLPASPKKVLLPYYDAIADLIILGR DHSKKYFGIEPSTIIQPYSKSQIDWLMQNTEMWPIACASFVGILDNHYPPNKLIQFCKLHTFVF PQIISKTPLNNALLVFTDGSSTGMAAYTLTDTTIKFQTNLNSAQLVELQALIAVLSAFPNQPLNI YTDSAYLAHSIPLLETVAQIKHISETAKLFLQCQQLIYNRSIPFYIGHVRAHSGLPGPIAQGNQR ADLATKIVASNINT
PERV_Q4VFZ2	TLQLDDEYRLYSPLVKPDQNIQFWLEQFPQAWAETAGMGLAKQVPPQVIQLKASATPVSVR QYPLSKEAQEGIRPHVQRLIQQGILVPVQSPWNTPLLPVRKPGTNDYRPVQDLREVNKRVQDI HPTVPNPYNLLCALPPQRSWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGTGRTGQLTWTR LPQGFKNSPTIFDEALHRDLANFRIQHPQVTLLQYVDDLLLAGATKQDCLEGTKALLLELSDL GYRASAKKAQICRREVTYLGYSLRDGQRWLTEARKKTVVQIPAPTTAKQVREFLGTAGFCRL WIPGFATLAAPLYPLTKEKGEFSWAPEHQKAFDAIKKALLSAPALALPDVTKPFTLYVDERK GVARGVLTQTLGPWRRPVAYLSKKLDPVASGWPVCLKAIAAVAILVKDADKLTLGQNITVIA PHALENIVRQPPDRWMTNARMTHYQSLLLTERVTFAPPAALNPATLLPEETDEPVTHDCHQL LIEETGVRKDLTDIPLTGEVLTWFTDGSSYVVEGKRMAGAAVVDGTRTIWASSLPEGTSAQK AELMALTQALRLAEGKSINIYTDSRYAFATAHVHGAIYKQRGLLTSAGREIKNKEEILSLLEA LHLPKRLAIIHCPGHQKAKDPISRGNQMADRVAKQAAQGVNLL
PERV_Q4VFZ2	TLQLDDEYRLYSPLVKPDQNIQFWLEQFPQAWAETAGMGLAKQVPPQVIQLKASATPVSVR QYPLSKEAQEGIRPHVQRLIQQGILVPVQSPWNTPLLPVRKPGTNDYRPVQDLREVNKRVQDI HPTVPNPYNLLCALPPQRSWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGTGRTGQLTWTR LPQGFKNSPTIFDEALHRDLANFRIQHPQVTLLQYVDDLLLAGATKQDCLEGTKALLLELSDL GYRASAKKAQICRREVTYLGYSLRDGQRWLTEARKKTVVQIPAPTTAKQVREFLGTAGFCRL

[0602]

[0603]

	WIPGFATLAAPLYPLTKEKGEFSWAPEHQKAFDAIKKALLSAPALALPDVTKPFTLYVDERK GVARGVLTQTLGPWRRPVAYLSKKLDPVASGWPVCLKAIAAVAILVKDADKLTLGQNITVIA PHALENIVRQPPDRWMTNARMTHYQSLLLTERVTFAPPAALNPATLLPEETDEPVTHDCHQL LIEETGVRKDLTDIPLTGEVLTWFTDGSSYVVEGKRMAGAAVVDGTRTIWASSLPEGTSAQK AELMALTQALRLAEGKSINIYTDSRYAFATAHVHGAIYKQRGLLTSAGREIKNKEEILSLLEA LHLPKRLAIIHCPGHQKAKDPISRGNQMADRVAKQAAQGVNLL
PERV_Q4VFZ2_3mut	TLQLDDEYRLYSPLVKPDQNIQFWLEQFPQAWAETAGMGLAKQVPPQVIQLKASATPVSVR QYPLSKEAQEGIRPHVQRLIQQGILVPVQSPWNTPLLPVRKPGTNDYRPVQDLREVNKRVQDI HPTVPNPYNLLCALPPQRSWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGTGRTGQLTWTR LPQGFKNSPTIFNEALHRDLANFRIQHPQVTLLQYVDDLLLAGATKQDCLEGTKALLLELSDL GYRASAKKAQICRREVTYLGYSLRDGQRWLTEARKKTVVQIPAPTTAKQVREFLGTAGFCRL WIPGFATLAAPLYPLTKPKGEFSWAPEHQKAFDAIKKALLSAPALALPDVTKPFTLYVDERKG VARGVLTQTLGPWRRPVAYLSKKLDPVASGWPVCLKAIAAVAILVKDADKLTLGQNITVIAP HALENIVRQPPDRWMTNARMTHYQSLLLTERVTFAPPAALNPATLLPEETDEPVTHDCHQLL IEETGVRKDLTDIPLTGEVLTWFTDGSSYVVEGKRMAGAAVVDGTRTIWASSLPEGTSAQKA ELMALTQALRLAEGKSINIYTDSRYAFATAHVHGAIYKQRGWLTSAGREIKNKEEILSLLEAL HLPKRLAIIHCPGHQKAKDPISRGNQMADRVAKQAAQGVNLL
PERV_Q4VFZ2_3mut	TLQLDDEYRLYSPLVKPDQNIQFWLEQFPQAWAETAGMGLAKQVPPQVIQLKASATPVSVR QYPLSKEAQEGIRPHVQRLIQQGILVPVQSPWNTPLLPVRKPGTNDYRPVQDLREVNKRVQDI HPTVPNPYNLLCALPPQRSWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGTGRTGQLTWTR LPQGFKNSPTIFNEALHRDLANFRIQHPQVTLLQYVDDLLLAGATKQDCLEGTKALLLELSDL GYRASAKKAQICRREVTYLGYSLRDGQRWLTEARKKTVVQIPAPTTAKQVREFLGTAGFCRL WIPGFATLAAPLYPLTKPKGEFSWAPEHQKAFDAIKKALLSAPALALPDVTKPFTLYVDERKG VARGVLTQTLGPWRRPVAYLSKKLDPVASGWPVCLKAIAAVAILVKDADKLTLGQNITVIAP HALENIVRQPPDRWMTNARMTHYQSLLLTERVTFAPPAALNPATLLPEETDEPVTHDCHQLL IEETGVRKDLTDIPLTGEVLTWFTDGSSYVVEGKRMAGAAVVDGTRTIWASSLPEGTSAQKA ELMALTQALRLAEGKSINIYTDSRYAFATAHVHGAIYKQRGWLTSAGREIKNKEEILSLLEAL HLPKRLAIIHCPGHQKAKDPISRGNQMADRVAKQAAQGVNLL
PERV_Q4VFZ2_3mutA_WS	LDDEYRLYSPLVKPDQNIQFWLEQFPQAWAETAGMGLAKQVPPQVIQLKASATPVSVRQYP LSKEAQEGIRPHVQRLIQQGILVPVQSPWNTPLLPVRKPGTNDYRPVQDLREVNKRVQDIHPT VPNPYNLLCALPPQRSWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGTGRTGQLTWTRLPQ GFKNSPTIFNEALHRDLANFRIQHPQVTLLQYVDDLLLAGATKQDCLEGTKALLLELSDLGYR ASAKKAQICRREVTYLGYSLRDGQRWLTEARKKTVVQIPAPTTAKQVREFLGKAGFCRLFIP GFATLAAPLYPLTKPKGEFSWAPEHQKAFDAIKKALLSAPALALPDVTKPFTLYVDERKGVA RGVLTQTLGPWRRPVAYLSKKLDPVASGWPVCLKAIAAVAILVKDADKLTLGQNITVIAPHA LENIVRQPPDRWMTNARMTHYQSLLLTERVTFAPPAALNPATLLPEETDEPVTHDCHQLLIEE TGVRKDLTDIPLTGEVLTWFTDGSSYVVEGKRMAGAAVVDGTRTIWASSLPEGTSAQKAEL MALTQALRLAEGKSINIYTDSRYAFATAHVHGAIYKQRGWLTSAGREIKNKEEILSLLEALHL PKRLAIIHCPGHQKAKDPISRGNQMADRVAKQAAQGVNLLP
PERV_Q4VFZ2_3mutA_WS	LDDEYRLYSPLVKPDQNIQFWLEQFPQAWAETAGMGLAKQVPPQVIQLKASATPVSVRQYP LSKEAQEGIRPHVQRLIQQGILVPVQSPWNTPLLPVRKPGTNDYRPVQDLREVNKRVQDIHPT VPNPYNLLCALPPQRSWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPGTGRTGQLTWTRLPQ GFKNSPTIFNEALHRDLANFRIQHPQVTLLQYVDDLLLAGATKQDCLEGTKALLLELSDLGYR ASAKKAQICRREVTYLGYSLRDGQRWLTEARKKTVVQIPAPTTAKQVREFLGKAGFCRLFIP GFATLAAPLYPLTKPKGEFSWAPEHQKAFDAIKKALLSAPALALPDVTKPFTLYVDERKGVA RGVLTQTLGPWRRPVAYLSKKLDPVASGWPVCLKAIAAVAILVKDADKLTLGQNITVIAPHA LENIVRQPPDRWMTNARMTHYQSLLLTERVTFAPPAALNPATLLPEETDEPVTHDCHQLLIEE TGVRKDLTDIPLTGEVLTWFTDGSSYVVEGKRMAGAAVVDGTRTIWASSLPEGTSAQKAEL MALTQALRLAEGKSINIYTDSRYAFATAHVHGAIYKQRGWLTSAGREIKNKEEILSLLEALHL PKRLAIIHCPGHQKAKDPISRGNQMADRVAKQAAQGVNLLP
SFV1_P23074	MDPLQLLQPLEAEIKGTKLKAHWNSGATITCVPEAFLEDERPIQTMLIKTIHGEKQQDVYYLT FKVQGRKVEAEVLASPYDYILLNPSDVPWLMKKPLQLTVLVPLHEYQERLLQQTALPKEQKE LLQKLFLKYDALWQHWENQVGHRRIKPHNIATGTLAPRPQKQYPINPKAKPSIQIVIDDLLKQ GVLIQQNSTMNTPVYPVPKPDGKWRMVLDYREVNKTIPLIAAQNQHSAGILSSIYRGKYKTT LDLTNGFWAHPITPESYWLTAFTWQGKQYCWTRLPQGFLNSPALFTADVVDLLKEIPNVQA YVDDIYISHDDPQEHLEQLEKIFSILLNAGYVVSLKKSEIAQREVEFLGFNITKEGRGLTDTFKQ

	KLLNITPPKDLKQLQSILGLLNFARNFIPNYSELVKPLYTIVANANGKFISWTEDNSNQLQHIIS VLNQADNLEERNPETRLIIKVNSSPSAGYIRYYNEGSKRPIMYVNYIFSKAEAKFTQTEKLLTT MHKGLIKAMDLAMGQEILVYSPIVSMTKIQRTPLPERKALPVRWITWMTYLEDPRIQFHYDK SLPELQQIPNVTEDVIAKTKHPSEFAMVFYTDGSAIKHPDVNKSHSAGMGIAQVQFIPEYKIVH QWSIPLGDHTAQLAEIAAVEFACKKALKISGPVLIVTDSFYVAESANKELPYWKSNGFLNNK KKPLRHVSKWKSIAECLQLKPDIIIMHEKGHQQPMTTLHTEGNNLADKLATQGSYVVH
SFV1_P23074_2mut	MDPLQLLQPLEAEIKGTKLKAHWNSGATITCVPEAFLEDERPIQTMLIKTIHGEKQQDVYYLT FKVQGRKVEAEVLASPYDYILLNPSDVPWLMKKPLQLTVLVPLHEYQERLLQQTALPKEQKE LLQKLFLKYDALWQHWENQVGHRRIKPHNIATGTLAPRPQKQYPINPKAKPSIQIVIDDLLKQ GVLIQQNSTMNTPVYPVPKPDGKWRMVLDYREVNKTIPLIAAQNQHSAGILSSIYRGKYKTT LDLTNGFWAHPITPESYWLTAFTWQGKQYCWTRLPQGFLNSPALFNADVVDLLKEIPNVQA YVDDIYISHDDPQEHLEQLEKIFSILLNAGYVVSLKKSEIAQREVEFLGFNITKEGRGLTDTFKQ KLLNITPPKDLKQLQSILGLLNFARNFIPNYSELVKPLYTIVAPANGKFISWTEDNSNQLQHIIS VLNQADNLEERNPETRLIIKVNSSPSAGYIRYYNEGSKRPIMYVNYIFSKAEAKFTQTEKLLTT MHKGLIKAMDLAMGQEILVYSPIVSMTKIQRTPLPERKALPVRWITWMTYLEDPRIQFHYDK SLPELQQIPNVTEDVIAKTKHPSEFAMVFYTDGSAIKHPDVNKSHSAGMGIAQVQFIPEYKIVH QWSIPLGDHTAQLAEIAAVEFACKKALKISGPVLIVTDSFYVAESANKELPYWKSNGFLNNK KKPLRHVSKWKSIAECLQLKPDIIIMHEKGHQQPMTTLHTEGNNLADKLATQGSYVVH
SFV1_P23074_2mutA	MDPLQLLQPLEAEIKGTKLKAHWNSGATITCVPEAFLEDERPIQTMLIKTIHGEKQQDVYYLT FKVQGRKVEAEVLASPYDYILLNPSDVPWLMKKPLQLTVLVPLHEYQERLLQQTALPKEQKE LLQKLFLKYDALWQHWENQVGHRRIKPHNIATGTLAPRPQKQYPINPKAKPSIQIVIDDLLKQ GVLIQQNSTMNTPVYPVPKPDGKWRMVLDYREVNKTIPLIAAQNQHSAGILSSIYRGKYKTT LDLTNGFWAHPITPESYWLTAFTWQGKQYCWTRLPQGFLNSPALFNADVVDLLKEIPNVQA YVDDIYISHDDPQEHLEQLEKIFSILLNAGYVVSLKKSEIAQREVEFLGFNITKEGRGLTDTFKQ KLLNITPPKDLKQLQSILGKLNFARNFIPNYSELVKPLYTIVAPANGKFISWTEDNSNQLQHIIS VLNQADNLEERNPETRLIIKVNSSPSAGYIRYYNEGSKRPIMYVNYIFSKAEAKFTQTEKLLTT MHKGLIKAMDLAMGQEILVYSPIVSMTKIQRTPLPERKALPVRWITWMTYLEDPRIQFHYDK SLPELQQIPNVTEDVIAKTKHPSEFAMVFYTDGSAIKHPDVNKSHSAGMGIAQVQFIPEYKIVH QWSIPLGDHTAQLAEIAAVEFACKKALKISGPVLIVTDSFYVAESANKELPYWKSNGFLNNK KKPLRHVSKWKSIAECLQLKPDIIIMHEKGHQQPMTTLHTEGNNLADKLATQGSYVVH
SFV1_P23074-Pro	VPWLMKKPLQLTVLVPLHEYQERLLQQTALPKEQKELLQKLFLKYDALWQHWENQVGHRR IKPHNIATGTLAPRPQKQYPINPKAKPSIQIVIDDLLKQGVLIQQNSTMNTPVYPVPKPDGKWR MVLDYREVNKTIPLIAAQNQHSAGILSSIYRGKYKTTLDLTNGFWAHPITPESYWLTAFTWQG KQYCWTRLPQGFLNSPALFTADVVDLLKEIPNVQAYVDDIYISHDDPQEHLEQLEKIFSILLNA GYVVSLKKSEIAQREVEFLGFNITKEGRGLTDTFKQKLLNITPPKDLKQLQSILGLLNFARNFIP NYSELVKPLYTIVANANGKFISWTEDNSNQLQHIISVLNQADNLEERNPETRLIIKVNSSPSAG YIRYYNEGSKRPIMYVNYIFSKAEAKFTQTEKLLTTMHKGLIKAMDLAMGQEILVYSPIVSMT KIQRTPLPERKALPVRWITWMTYLEDPRIQFHYDKSLPELQQIPNVTEDVIAKTKHPSEFAMV FYTDGSAIKHPDVNKSHSAGMGIAQVQFIPEYKIVHQWSIPLGDHTAQLAEIAAVEFACKKAL KISGPVLIVTDSFYVAESANKELPYWKSNGFLNNKKKPLRHVSKWKSIAECLQLKPDIIIMHE KGHQQPMTTLHTEGNNLADKLATQGSYVVH
SFV1_P23074-Pro_2mut	VPWLMKKPLQLTVLVPLHEYQERLLQQTALPKEQKELLQKLFLKYDALWQHWENQVGHRR IKPHNIATGTLAPRPQKQYPINPKAKPSIQIVIDDLLKQGVLIQQNSTMNTPVYPVPKPDGKWR MVLDYREVNKTIPLIAAQNQHSAGILSSIYRGKYKTTLDLTNGFWAHPITPESYWLTAFTWQG KQYCWTRLPQGFLNSPALFNADVVDLLKEIPNVQAYVDDIYISHDDPQEHLEQLEKIFSILLN AGYVVSLKKSEIAQREVEFLGFNITKEGRGLTDTFKQKLLNITPPKDLKQLQSILGLLNFARNF IPNYSELVKPLYTIVAPANGKFISWTEDNSNQLQHIISVLNQADNLEERNPETRLIIKVNSSPSA GYIRYYNEGSKRPIMYVNYIFSKAEAKFTQTEKLLTTMHKGLIKAMDLAMGQEILVYSPIVSM TKIQRTPLPERKALPVRWITWMTYLEDPRIQFHYDKSLPELQQIPNVTEDVIAKTKHPSEFAM VFYTDGSAIKHPDVNKSHSAGMGIAQVQFIPEYKIVHQWSIPLGDHTAQLAEIAAVEFACKKA LKISGPVLIVTDSFYVAESANKELPYWKSNGFLNNKKKPLRHVSKWKSIAECLQLKPDIIIMHE KGHQQPMTTLHTEGNNLADKLATQGSYVVH
SFV1_P23074-Pro_2mutA	VPWLMKKPLQLTVLVPLHEYQERLLQQTALPKEQKELLQKLFLKYDALWQHWENQVGHRR IKPHNIATGTLAPRPQKQYPINPKAKPSIQIVIDDLLKQGVLIQQNSTMNTPVYPVPKPDGKWR MVLDYREVNKTIPLIAAQNQHSAGILSSIYRGKYKTTLDLTNGFWAHPITPESYWLTAFTWQG KQYCWTRLPQGFLNSPALFNADVVDLLKEIPNVQAYVDDIYISHDDPQEHLEQLEKIFSILLN

[0604]

	AGYVVSLKKSEIAQREVEFLGFNITKEGRGLTDTFKQKLLNITPPKDLKQLQSILGKLNFARNF IPNYSELVKPLYTIVAPANGKFISWTEDNSNQLQHIISVLNQADNLEERNPETRLIIKVNSSPSA GYIRYYNEGSKRPIMYVNYIFSKAEAKFTQTEKLLTTMHKGLIKAMDLAMGQEILVYSPIVSM TKIQRTPLPERKALPVRWITWMTYLEDPRIQFHYDKSLPELQQIPNVTEDVIAKTKHPSEFAM VFYTDGSAIKHPDVNKSHSAGMGIAQVQFIPEYKIVHQWSIPLGDHTAQLAEIAAVEFACKKA LKISGPVLIVTDSFYVAESANKELPYWKSNGFLNNKKKPLRHVSKWKSIAECLQLKPDIIIMHE KGHQQPMTTLHTEGNNLADKLATQGSYVVH
SFV3L_ P27401	MDPLQLLQPLEAEIKGTKLKAHWNSGATITCVPQAFLEEEVPIKNIWIKTIHGEKEQPVYYLTF KIQGRKVEAEVISSPYDYILVSPSDIPWLMKKPLQLTTLVPLQEYEERLLKQTMLTGSYKEKL QSLFLKYDALWQHWENQVGHRRIKPHHIATGTVNPRPQKQYPINPKAKASIQTVINDLLKQG VLIQQNSIMNTPVYPVPKPDGKWRMVLDYREVNKTIPLIAAQNQHSAGILSSIFRGKYKTTLD LSNGFWAHSITPESYWLTAFTWLGQQYCWTRLPQGFLNSPALFTADVVDLLKEVPNVQVYV DDIYISHDDPREHLEQLEKVFSLLLNAGYVVSLKKSEIAQHEVEFLGFNITKEGRGLTETFKQK LLNITPPRDLKQLQSILGLLNFARNFIPNFSELVKPLYNIIATANGKYITWTTDNSQQLQNIISML NSAENLEERNPEVRLIMKVNTSPSAGYIRFYNEFAKRPIMYLNYVYTKAEVKFTNTEKLLTTI HKGLIKALDLGMGQEILVYSPIVSMTKIQKTPLPERKALPIRWITWMSYLEDPRIQFHYDKTLP ELQQVPTVTDDIIAKIKHPSEFSMVFYTDGSAIKHPNVNKSHNAGMGIAQVQFKPEFTVINTW SIPLGDHTAQLAEVAAVEFACKKALKIDGPVLIVTDSFYVAESVNKELPYWQSNGFFNNKKK PLKHVSKWKSIADCIQLKPDIIIIHEKGHQPTASTFHTEGNNLADKLATQGSYVVN
SFV3L_ P27401 _2mut	MDPLQLLQPLEAEIKGTKLKAHWNSGATITCVPQAFLEEEVPIKNIWIKTIHGEKEQPVYYLTF KIQGRKVEAEVISSPYDYILVSPSDIPWLMKKPLQLTTLVPLQEYEERLLKQTMLTGSYKEKL QSLFLKYDALWQHWENQVGHRRIKPHHIATGTVNPRPQKQYPINPKAKASIQTVINDLLKQG VLIQQNSIMNTPVYPVPKPDGKWRMVLDYREVNKTIPLIAAQNQHSAGILSSIFRGKYKTTLD LSNGFWAHSITPESYWLTAFTWLGQQYCWTRLPQGFLNSPALFNADVVDLLKEVPNVQVYV DDIYISHDDPREHLEQLEKVFSLLLNAGYVVSLKKSEIAQHEVEFLGFNITKEGRGLTETFKQK LLNITPPRDLKQLQSILGLLNFARNFIPNFSELVKPLYNIIATAPGKYITWTTDNSQQLQNIISML NSAENLEERNPEVRLIMKVNTSPSAGYIRFYNEFAKRPIMYLNYVYTKAEVKFTNTEKLLTTI HKGLIKALDLGMGQEILVYSPIVSMTKIQKTPLPERKALPIRWITWMSYLEDPRIQFHYDKTLP ELQQVPTVTDDIIAKIKHPSEFSMVFYTDGSAIKHPNVNKSHNAGMGIAQVQFKPEFTVINTW SIPLGDHTAQLAEVAAVEFACKKALKIDGPVLIVTDSFYVAESVNKELPYWQSNGFFNNKKK PLKHVSKWKSIADCIQLKPDIIIIHEKGHQPTASTFHTEGNNLADKLATQGSYVVN
SFV3L_ P27401 _2mutA	MDPLQLLQPLEAEIKGTKLKAHWNSGATITCVPQAFLEEEVPIKNIWIKTIHGEKEQPVYYLTF KIQGRKVEAEVISSPYDYILVSPSDIPWLMKKPLQLTTLVPLQEYEERLLKQTMLTGSYKEKL QSLFLKYDALWQHWENQVGHRRIKPHHIATGTVNPRPQKQYPINPKAKASIQTVINDLLKQG VLIQQNSIMNTPVYPVPKPDGKWRMVLDYREVNKTIPLIAAQNQHSAGILSSIFRGKYKTTLD LSNGFWAHSITPESYWLTAFTWLGQQYCWTRLPQGFLNSPALFNADVVDLLKEVPNVQVYV DDIYISHDDPREHLEQLEKVFSLLLNAGYVVSLKKSEIAQHEVEFLGFNITKEGRGLTETFKQK LLNITPPRDLKQLQSILGKLNFARNFIPNFSELVKPLYNIIATAPGKYITWTTDNSQQLQNIISML NSAENLEERNPEVRLIMKVNTSPSAGYIRFYNEFAKRPIMYLNYVYTKAEVKFTNTEKLLTTI HKGLIKALDLGMGQEILVYSPIVSMTKIQKTPLPERKALPIRWITWMSYLEDPRIQFHYDKTLP ELQQVPTVTDDIIAKIKHPSEFSMVFYTDGSAIKHPNVNKSHNAGMGIAQVQFKPEFTVINTW SIPLGDHTAQLAEVAAVEFACKKALKIDGPVLIVTDSFYVAESVNKELPYWQSNGFFNNKKK PLKHVSKWKSIADCIQLKPDIIIIHEKGHQPTASTFHTEGNNLADKLATQGSYVVN
SFV3L_ P27401- Pro	IPWLMKKPLQLTTLVPLQEYEERLLKQTMLTGSYKEKLQSLFLKYDALWQHWENQVGHRRI KPHHIATGTVNPRPQKQYPINPKAKASIQTVINDLLKQGVLIQQNSIMNTPVYPVPKPDGKWR MVLDYREVNKTIPLIAAQNQHSAGILSSIFRGKYKTTLDLSNGFWAHSITPESYWLTAFTWLG QQYCWTRLPQGFLNSPALFTADVVDLLKEVPNVQVYVDDIYISHDDPREHLEQLEKVFSLLL NAGYVVSLKKSEIAQHEVEFLGFNITKEGRGLTETFKQKLLNITPPRDLKQLQSILGLLNFARN FIPNFSELVKPLYNIIATANGKYITWTTDNSQQLQNIISMLNSAENLEERNPEVRLIMKVNTSPS AGYIRFYNEFAKRPIMYLNYVYTKAEVKFTNTEKLLTTIHKGLIKALDLGMGQEILVYSPIVS MTKIQKTPLPERKALPIRWITWMSYLEDPRIQFHYDKTLPELQQVPTVTDDIIAKIKHPSEFSM VFYTDGSAIKHPNVNKSHNAGMGIAQVQFKPEFTVINTWSIPLGDHTAQLAEVAAVEFACKK ALKIDGPVLIVTDSFYVAESVNKELPYWQSNGFFNNKKKPLKHVSKWKSIADCIQLKPDIIIIH EKGHQPTASTFHTEGNNLADKLATQGSYVVN
SFV3L_ P27401-	IPWLMKKPLQLTTLVPLQEYEERLLKQTMLTGSYKEKLQSLFLKYDALWQHWENQVGHRRI KPHHIATGTVNPRPQKQYPINPKAKASIQTVINDLLKQGVLIQQNSIMNTPVYPVPKPDGKWR

[0605]

Pro_2mut	MVLDYREVNKTIPLIAAQNQHSAGILSSIFRGKYKTTLDLSNGFWAHSITPESYWLTAFTWLG QQYCWTRLPQGFLNSPALFNADVVDLLKEVPNVQVYVDDIYISHDDPREHLEQLEKVFSLLL NAGYVVSLKKSEIAQHEVEFLGFNITKEGRGLTETFKQKLLNITPPRDLKQLQSILGLLNFARN FIPNFSELVKPLYNIIATAPGKYITWTTDNSQQLQNIISMLNSAENLEERNPEVRLIMKVNTSPS AGYIRFYNEFAKRPIMYLNYVYTKAEVKFTNTEKLLTTIHKGLIKALDLGMGQEILVYSPIVS MTKIQKTPLPERKALPIRWITWMSYLEDPRIQFHYDKTLPELQQVPTVTDDIIAKIKHPSEFSM VFYTDGSAIKHPNVNKSHNAGMGIAQVQFKPEFTVINTWSIPLGDHTAQLAEVAAVEFACKK ALKIDGPVLIVTDSFYVAESVNKELPYWQSNGFFNNKKKPLKHVSKWKSIADCIQLKPDIIIIH EKGHQPTASTFHTEGNNLADKLATQGSYVVN
SFV3L_P27401-Pro_2mutA	IPWLMKKPLQLTTLVPLQEYEERLLKQTMLTGSYKEKLQSLFLKYDALWQHWENQVGHRRI KPHHIATGTVNPRPQKQYPINPKAKASIQTVINDLLKQGVLIQQNSIMNTPVYPVPKPDGKWR MVLDYREVNKTIPLIAAQNQHSAGILSSIFRGKYKTTLDLSNGFWAHSITPESYWLTAFTWLG QQYCWTRLPQGFLNSPALFNADVVDLLKEVPNVQVYVDDIYISHDDPREHLEQLEKVFSLLL NAGYVVSLKKSEIAQHEVEFLGFNITKEGRGLTETFKQKLLNITPPRDLKQLQSILGKLNFARN FIPNFSELVKPLYNIIATAPGKYITWTTDNSQQLQNIISMLNSAENLEERNPEVRLIMKVNTSPS AGYIRFYNEFAKRPIMYLNYVYTKAEVKFTNTEKLLTTIHKGLIKALDLGMGQEILVYSPIVS MTKIQKTPLPERKALPIRWITWMSYLEDPRIQFHYDKTLPELQQVPTVTDDIIAKIKHPSEFSM VFYTDGSAIKHPNVNKSHNAGMGIAQVQFKPEFTVINTWSIPLGDHTAQLAEVAAVEFACKK ALKIDGPVLIVTDSFYVAESVNKELPYWQSNGFFNNKKKPLKHVSKWKSIADCIQLKPDIIIIH EKGHQPTASTFHTEGNNLADKLATQGSYVVN
SFVCP_Q87040	MNPLQLLQPLPAEVKGTKLLAHWNSGATITCIPESFLEDEQPIKQTLIKTIHGEKQQNVYYLTF KVKGRKVEAEVIASPYEYILLSPTDVPWLTQQPLQLTILVPLQEYQDRILNKTALPEEQKQQL KALFTKYDNLWQHWENQVGHRKIRPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQG VLTPQNSTMNTPVYPVPKPDGRWRMVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTL DLANGFWAHPITPDSYWLTAFTWQGKQYCWTRLPQGFLNSPALFTADAVDLLKEVPNVQV YVDDIYLSHDNPHEHIQQLEKVFQILLQAGYVVSLKKSEIGQRTVEFLGFNITKEGRGLTDTFK TKLLNVTPPKDLKQLQSILGLLNFARNFIPNFAELVQTLYNLIASSKGKYIEWTEDNTKQLNK VIEALNTASNLEERLPDQRLVIKVNTSPSAGYVRYYNESGKKPIMYLNYVFSKAELKFSMLEK LLTTMHKALIKAMDLAMGQEILVYSPIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFH YDKTLPELKHIPDVYTSSIPPLKHPSQYEGVFCTDGSAIKSPDPTKSNNAGMGIVHAIYNPEYKI LNQWSIPLGHHTAQMAEIAAVEFACKKALKVPGPVLVITDSFYVAESANKELPYWKSNGFVN NKKEPLKHISKWKSIAECLSIKPDITIQHEKGHQPINTSIHTEGNALADKLATQGSYVVN
SFVCP_Q87040_2mut	MNPLQLLQPLPAEVKGTKLLAHWNSGATITCIPESFLEDEQPIKQTLIKTIHGEKQQNVYYLTF KVKGRKVEAEVIASPYEYILLSPTDVPWLTQQPLQLTILVPLQEYQDRILNKTALPEEQKQQL KALFTKYDNLWQHWENQVGHRKIRPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQG VLTPQNSTMNTPVYPVPKPDGRWRMVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTL DLANGFWAHPITPDSYWLTAFTWQGKQYCWTRLPQGFLNSPALFNADAVDLLKEVPNVQV YVDDIYLSHDNPHEHIQQLEKVFQILLQAGYVVSLKKSEIGQRTVEFLGFNITKEGRGLTDTFK TKLLNVTPPKDLKQLQSILGLLNFARNFIPNFAELVQTLYNLIASSPGKYIEWTEDNTKQLNKV IEALNTASNLEERLPDQRLVIKVNTSPSAGYVRYYNESGKKPIMYLNYVFSKAELKFSMLEKL LTTMHKALIKAMDLAMGQEILVYSPIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFHY DKTLPELKHIPDVYTSSIPPLKHPSQYEGVFCTDGSAIKSPDPTKSNNAGMGIVHAIYNPEYKIL NQWSIPLGHHTAQMAEIAAVEFACKKALKVPGPVLVITDSFYVAESANKELPYWKSNGFVN NKKEPLKHISKWKSIAECLSIKPDITIQHEKGHQPINTSIHTEGNALADKLATQGSYVVN
SFVCP_Q87040_2mutA	MNPLQLLQPLPAEVKGTKLLAHWNSGATITCIPESFLEDEQPIKQTLIKTIHGEKQQNVYYLTF KVKGRKVEAEVIASPYEYILLSPTDVPWLTQQPLQLTILVPLQEYQDRILNKTALPEEQKQQL KALFTKYDNLWQHWENQVGHRKIRPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQG VLTPQNSTMNTPVYPVPKPDGRWRMVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTL DLANGFWAHPITPDSYWLTAFTWQGKQYCWTRLPQGFLNSPALFNADAVDLLKEVPNVQV YVDDIYLSHDNPHEHIQQLEKVFQILLQAGYVVSLKKSEIGQRTVEFLGFNITKEGRGLTDTFK TKLLNVTPPKDLKQLQSILGKLNFARNFIPNFAELVQTLYNLIASSPGKYIEWTEDNTKQLNK VIEALNTASNLEERLPDQRLVIKVNTSPSAGYVRYYNESGKKPIMYLNYVFSKAELKFSMLEK LLTTMHKALIKAMDLAMGQEILVYSPIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFH YDKTLPELKHIPDVYTSSIPPLKHPSQYEGVFCTDGSAIKSPDPTKSNNAGMGIVHAIYNPEYKI LNQWSIPLGHHTAQMAEIAAVEFACKKALKVPGPVLVITDSFYVAESANKELPYWKSNGFVN NKKEPLKHISKWKSIAECLSIKPDITIQHEKGHQPINTSIHTEGNALADKLATQGSYVVN

[0606]

SFVCP_Q87040-Pro	VPWLTQQPLQLTILVPLQEYQDRILNKTALPEEQKQQLKALFTKYDNLWQHWENQVGHRKI RPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQGVLTPQNSTMNTPVYPVPKPDGRWR MVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTLDLANGFWAHPITPDSYWLTAFTWQ GKQYCWTRLPQGFLNSPALFTADAVDLLKEVPNVQVYVDDIYLSHDNPHEHIQQLEKVFQIL LQAGYVVSLKKSEIGQRTVEFLGFNITKEGRGLTDTFKTKLLNVTPPKDLKQLQSILGLLNFA RNFIPNFAELVQTLYNLIASSKGKYIEWTEDNTKQLNKVIEALNTASNLEERLPDQRLVIKVNT SPSAGYVRYYNESGKKPIMYLNYVFSKAELKFSMLEKLLTTMHKALIKAMDLAMGQEILVY SPIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFHYDKTLPELKHIPDVYTSSIPPLKHPS QYEGVFCTDGSAIKSPDPTKSNNAGMGIVHAIYNPEYKILNQWSIPLGHHTAQMAEIAAVEFA CKKALKVPGPVLVITDSFYVAESANKELPYWKSNGFVNNKKEPLKHISKWKSIAECLSIKPDI TIQHEKGHQPINTSIHTEGNALADKLATQGSYVVN
SFVCP_Q87040-Pro_2mut	VPWLTQQPLQLTILVPLQEYQDRILNKTALPEEQKQQLKALFTKYDNLWQHWENQVGHRKI RPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQGVLTPQNSTMNTPVYPVPKPDGRWR MVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTLDLANGFWAHPITPDSYWLTAFTWQ GKQYCWTRLPQGFLNSPALFNADAVDLLKEVPNVQVYVDDIYLSHDNPHEHIQQLEKVFQIL LQAGYVVSLKKSEIGQRTVEFLGFNITKEGRGLTDTFKTKLLNVTPPKDLKQLQSILGLLNFA RNFIPNFAELVQTLYNLIASSPGKYIEWTEDNTKQLNKVIEALNTASNLEERLPDQRLVIKVNT SPSAGYVRYYNESGKKPIMYLNYVFSKAELKFSMLEKLLTTMHKALIKAMDLAMGQEILVY SPIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFHYDKTLPELKHIPDVYTSSIPPLKHPS QYEGVFCTDGSAIKSPDPTKSNNAGMGIVHAIYNPEYKILNQWSIPLGHHTAQMAEIAAVEFA CKKALKVPGPVLVITDSFYVAESANKELPYWKSNGFVNNKKEPLKHISKWKSIAECLSIKPDI TIQHEKGHQPINTSIHTEGNALADKLATQGSYVVN
SFVCP_Q87040-Pro_2mutA	VPWLTQQPLQLTILVPLQEYQDRILNKTALPEEQKQQLKALFTKYDNLWQHWENQVGHRKI RPHNIATGDYPPRPQKQYPINPKAKPSIQIVIDDLLKQGVLTPQNSTMNTPVYPVPKPDGRWR MVLDYREVNKTIPLTAAQNQHSAGILATIVRQKYKTTLDLANGFWAHPITPDSYWLTAFTWQ GKQYCWTRLPQGFLNSPALFNADAVDLLKEVPNVQVYVDDIYLSHDNPHEHIQQLEKVFQIL LQAGYVVSLKKSEIGQRTVEFLGFNITKEGRGLTDTFKTKLLNVTPPKDLKQLQSILGKLNFA RNFIPNFAELVQTLYNLIASSPGKYIEWTEDNTKQLNKVIEALNTASNLEERLPDQRLVIKVNT SPSAGYVRYYNESGKKPIMYLNYVFSKAELKFSMLEKLLTTMHKALIKAMDLAMGQEILVY SPIVSMTKIQKTPLPERKALPIRWITWMTYLEDPRIQFHYDKTLPELKHIPDVYTSSIPPLKHPS QYEGVFCTDGSAIKSPDPTKSNNAGMGIVHAIYNPEYKILNQWSIPLGHHTAQMAEIAAVEFA CKKALKVPGPVLVITDSFYVAESANKELPYWKSNGFVNNKKEPLKHISKWKSIAECLSIKPDI TIQHEKGHQPINTSIHTEGNALADKLATQGSYVVN
SMRVH_P03364	PRSRAIDIPVPHADKISWKITDPVWVDQWPLTYEKTLAAIALVQEQLAAGHIEPTNSPWNTPIF IIKKKSGSWRLLQDLRAVNKVMVPMGALQPGLPSPVAIPLNYHKIVIDLKDCFFTIPLHPEDRP YFAFSVPQINFQSPMPRYQWKVLPQGMANSPTLCQKFVAAAIAPVRSQWPEAYILHYMDDIL LACDSAEAAKACYAHIISCLTSYGLKIAPDKVQVSEPFSYLGFELHHQQVFTPRVCLKTDHLK TLNDFQKLLGDIQWLRPYLKLPTSALVPLNNILKGDPNPLSVRALTPEAKQSLALINKAIQNQS VQQISYNLPLVLLLLPTPHTPTAVFWQPNGTDPTKNGSPLLWLHLPASPSKVLLTYPSLLAMLI IKGRYTGRQLFGRDPHSIIIPYTQDQLTWLLQTSDEWAIALSSFTGDIDNHYPSDPVIQFAKLH QFIFPKITKCAPIPQATLVFTDGSSNGIAAYVIDNQPISIKSPYLSAQLVELYAILQVFTVLAHQP FNLYTDSAYIAQSVPLLETVPFIKSSTNATPLFSKLQQLILNRQHPFFIGHLRAHLNLPGPLAEG NALADAATQIFPIISD
SMRVH_P03364_2mut	PRSRAIDIPVPHADKISWKITDPVWVDQWPLTYEKTLAAIALVQEQLAAGHIEPTNSPWNTPIF IIKKKSGSWRLLQDLRAVNKVMVPMGALQPGLPSPVAIPLNYHKIVIDLKDCFFTIPLHPEDRP YFAFSVPQINFQSPMPRYQWKVLPQGMANSPTLCQKFVAAAIAPVRSQWPEAYILHYMDDIL LACDSAEAAKACYAHIISCLTSYGLKIAPDKVQVSEPFSYLGFELHHQQVFTPRVCLKTDHLK TLNDFQKLLGDIQWLRPYLKLPTSALVPLNNILKPDPNPLSVRALTPEAKQSLALINKAIQNQS VQQISYNLPLVLLLLPTPHTPTAVFWQPNGTDPTKNGSPLLWLHLPASPSKVLLTYPSLLAMLI IKGRYTGRQLFGRDPHSIIIPYTQDQLTWLLQTSDEWAIALSSFTGDIDNHYPSDPVIQFAKLH QFIFPKITKCAPIPQATLVFTDGSSNGIAAYVIDNQPISIKSPYLSAQLVELYAILQVFTVLAHQP FNLYTDSAYIAQSVPLLETVPFIKSSTNATPLFSKLQQLILNRQHPFFIGHLRAHLNLPGPLAEG NALADAATQIFPIISD
SMRVH_P033	PRSRAIDIPVPHADKISWKITDPVWVDQWPLTYEKTLAAIALVQEQLAAGHIEPTNSPWNTPIF IIKKKSGSWRLLQDLRAVNKVMVPMGALQPGLPSPVAPPLNYHKIVIDLKDCFFTIPLHPEDR PYFAFSVPQINFQSPMPRYQWKVLPQGMANSPTLCQKFVAAAIAPVRSQWPEAYILHYMDDI

[0607]

64_2mu tB	LLACDSAEAAKACYAHIISCLTSYGLKIAPDKVQVSEPFSYLGFELHHQQVFTPRVCLKTDHL KTLNDFQKLLGDIQWLRPYLKLPTSALVPLNNILKPDPNPLSVRALTPEAKQSLALINKAIQNQ SVQQISYNLPLVLLLLPTPHTPTAVFWQPNGTDPTKNGSPLLWLHLPASPSKVLLTYPSLLAM LIIKGRYTGRQLFGRDPHSIIIPYTQDQLTWLLQTSDEWAIALSSFTGDIDNHYPSDPVIQFAKL HQFIFPKITKCAPIPQATLVFTDGSSNGIAAYVIDNQPISIKSPYLSAQLVELYAILQVFTVLAHQ PFNLYTDSAYIAQSVPLLETVPFIKSSTNATPLFSKLQQLILNRQHPFFIGHLRAHLNLPGPLAE GNALADAATQIFPIISD
SRV2_ P51517	LATAVDILAPQRYADPITWKSDEPVWVDQWPLTQEKLAAAQQLVQEQLQAGHIIESNSPWN TPIFVIKKKSGKWRLLQDLRAVNATMVLMGALQPGLPSPVAIPQGYFKIVIDLKDCFFTIPLQP VDQKRFAFSLPSTNFKQPMKRYQWKVLPQGMANSPTLCQKYVAAAIEPVRKSWAQMYIIHY MDDILIAGKLGEQVLQCFAQLKQALTTTGLQIAPEKVQLQDPYTYLGFQINGPKITNQKAVIR RDKLQTLNDFQKLLGDINWLRPYLHLTTGDLKPLFDILKGDSNPNSPRSLSEAALASLQKVET AIAEQFVTQIDYTQPLTFLIFNTTLTPTGLFWQNNPVMWVHLPASPKKVLLPYYDAIADLIILG RDNSKKYFGLEPSTIIQPYSKSQIHWLMQNTETWPIACASYAGNIDNHYPPNKLIQFCKLHAV VFPRIISKTPLDNALLVFTDGSSTGIAAYTFEKTTVRFKTSHTSAQLVELQALIAVLSAFPHRAL NVYTDSAYLAHSIPLLETVSHIKHISDTAKFFLQCQQLIYNRSIPFYLGHIRAHSGLPGPLSQGN HITDLATKVVATTLTT
SRV2_ P51517 _2mutB	LATAVDILAPQRYADPITWKSDEPVWVDQWPLTQEKLAAAQQLVQEQLQAGHIIESNSPWN TPIFVIKKKSGKWRLLQDLRAVNATMVLMGALQPGLPSPVAPPQGYFKIVIDLKDCFFTIPLQP VDQKRFAFSLPSTNFKQPMKRYQWKVLPQGMANSPTLCQKYVAAAIEPVRKSWAQMYIIHY MDDILIAGKLGEQVLQCFAQLKQALTTTGLQIAPEKVQLQDPYTYLGFQINGPKITNQKAVIR RDKLQTLNDFQKLLGDINWLRPYLHLTTGDLKPLFDILKGDSNPNSPRSLSEAALASLQKVET AIAEQFVTQIDYTQPLTFLIFNTTLTPTGLFWQNNPVMWVHLPASPKKVLLPYYDAIADLIILG RDNSKKYFGLEPSTIIQPYSKSQIHWLMQNTETWPIACASYAGNIDNHYPPNKLIQFCKLHAV VFPRIISKTPLDNALLVFTDGSSTGIAAYTFEKTTVRFKTSHTSAQLVELQALIAVLSAFPHRAL NVYTDSAYLAHSIPLLETVSHIKHISDTAKFFLQCQQLIYNRSIPFYLGHIRAHSGLPGPLSQGN HITDLATKVVATTLTT
WDSV_ O92815	SCQTKNTLNIDEYLLQFPDQLWASLPTDIGRMLVPPITIKIKDNASLPSIRQYPLPKDKTEGLRP LISSLENQGILIKCHSPCNTPIFPIKKAGRDEYRMIHDLRAINNIVAPLTAVVASPTTVLSNLAPS LHWFTVIDLSNAFFSVPIHKDSQYLFAFTFEGHQYTWTVLPQGFIHSPTLFSQALYQSLHKIKF KISSEICIYMDDVLIASKDRDTNLKDTAVMLQHLASEGHKVSKKKLQLCQQEVVYLGQLLTP EGRKILPDRKVTVSQFQQPTTIRQIRAFLGLVGYCRHWIPEFSIHSKFLEKQLKKDTAEPFQLD DQQVEAFNKLKHAITTAPVLVVPDPAKPFQLYTSHSEHASIAVLTQKHAGRTRPIAFLSSKFD AIESGLPPCLKACASIHRSLTQADSFILGAPLIIYTTHAICTLLQRDRSQLVTASRFSKWEADLL RPELTFVACSAVSPAHLYMQSCENNIPPHDCVLLTHTISRPRPDLSDLPIPDPDMTLFSDGSYTT GRGGAAVVMHRPVTDDFIIIHQQPGGASAQTAELLALAAACHLATDKTVNIYTDSRYAYGV VHDFGHLWMHRGFVTSAGTPIKNHKEIEYLLKQIMKPKQVSVIKIEAHTKGVSMEVRGNAA ADEAAKNAVFLVQR
WDSV_ O92815 _2mut	SCQTKNTLNIDEYLLQFPDQLWASLPTDIGRMLVPPITIKIKDNASLPSIRQYPLPKDKTEGLRP LISSLENQGILIKCHSPCNTPIFPIKKAGRDEYRMIHDLRAINNIVAPLTAVVASPTTVLSNLAPS LHWFTVIDLSNAFFSVPIHKDSQYLFAFTFEGHQYTWTVLPQGFIHSPTLFNQALYQSLHKIKF KISSEICIYMDDVLIASKDRDTNLKDTAVMLQHLASEGHKVSKKKLQLCQQEVVYLGQLLTP EGRKILPDRKVTVSQFQQPTTIRQIRAFLGLVGYCRHWIPEFSIHSKFLEKQLKPDTAEPFQLD DQQVEAFNKLKHAITTAPVLVVPDPAKPFQLYTSHSEHASIAVLTQKHAGRTRPIAFLSSKFD AIESGLPPCLKACASIHRSLTQADSFILGAPLIIYTTHAICTLLQRDRSQLVTASRFSKWEADLL RPELTFVACSAVSPAHLYMQSCENNIPPHDCVLLTHTISRPRPDLSDLPIPDPDMTLFSDGSYTT GRGGAAVVMHRPVTDDFIIIHQQPGGASAQTAELLALAAACHLATDKTVNIYTDSRYAYGV VHDFGHLWMHRGFVTSAGTPIKNHKEIEYLLKQIMKPKQVSVIKIEAHTKGVSMEVRGNAA ADEAAKNAVFLVQR
WDSV_ O92815 _2mutA	SCQTKNTLNIDEYLLQFPDQLWASLPTDIGRMLVPPITIKIKDNASLPSIRQYPLPKDKTEGLRP LISSLENQGILIKCHSPCNTPIFPIKKAGRDEYRMIHDLRAINNIVAPLTAVVASPTTVLSNLAPS LHWFTVIDLSNAFFSVPIHKDSQYLFAFTFEGHQYTWTVLPQGFIHSPTLFNQALYQSLHKIKF KISSEICIYMDDVLIASKDRDTNLKDTAVMLQHLASEGHKVSKKKLQLCQQEVVYLGQLLTP EGRKILPDRKVTVSQFQQPTTIRQIRAFLGKVGYCRHFIPEFSIHSKFLEKQLKPDTAEPFQLDD QQVEAFNKLKHAITTAPVLVVPDPAKPFQLYTSHSEHASIAVLTQKHAGRTRPIAFLSSKFDAI ESGLPPCLKACASIHRSLTQADSFILGAPLIIYTTHAICTLLQRDRSQLVTASRFSKWEADLLRP

[0608]

	ELTFVACSAVSPAHLYMQSCENNIPPHDCVLLTHTISRPRPDLSDLPIPDPDMTLFSDGSYTTG RGGAAVVMHRPVTDDFIIIHQQPGGASAQTAELLALAAACHLATDKTVNIYTDSRYAYGVV HDFGHLWMHRGFVTSAGTPIKNHKEIEYLLKQIMKPKQVSVIKIEAHTKGVSMEVRGNAAA DEAAKNAVFLVQR
WMSV_P03359	VLNLEEEYRLHEKPVPSSIDPSWLQLFPTVWAERAGMGLANQVPPVVVELRSGASPVAVRQY PMSKEAREGIRPHIQRFLDLGVLVPCQSPWNTPLLPVKKPGTNDYRPVQDLREINKRVQDIHP TVPNPYNLLSSLPPSHTWYSVLDLKDAFFCLKLHPNSQPLFAFEWRDPEKGNTGQLTWTRLP QGFKNSPTLFDEALHRDLAPFRALNPQVVLLQYVDDLLVAAPTYRDCKEGTQKLLQELSKLG YRVSAKKAQLCQKEVTYLGYLLKEGKRWLTPARKATVMKIPPPTTPRQVREFLGTAGFCRL WIPGFASLAAPLYPLTKESIPFIWTEEHQKAFDRIKEALLSAPALALPDLTKPFTLYVDERAGV ARGVLTQTLGPWRRPVAYLSKKLDPVASGWPTCLKAVAAVALLLKDADKLTLGQNVTVIAS HSLESIVRQPPDRWMTNARMTHYQSLLLNERVSFAPPAVLNPATLLPVESEATPVHRCSEILA EETGTRRDLKDQPLPGVPAWYTDGSSFIAEGKRRAGAAIVDGKRTVWASSLPEGTSAQKAEL VALTQALRLAEGKDINIYTDSRYAFATAHIHGAIYKQRGLLTSAGKDIKNKEEILALLEAIHLP KRVAIIHCPGHQKGNDPVATGNRRADEAAKQAALSTRVLAETTKP
WMSV_P03359_3mut	VLNLEEEYRLHEKPVPSSIDPSWLQLFPTVWAERAGMGLANQVPPVVVELRSGASPVAVRQY PMSKEAREGIRPHIQRFLDLGVLVPCQSPWNTPLLPVKKPGTNDYRPVQDLREINKRVQDIHP TVPNPYNLLSSLPPSHTWYSVLDLKDAFFCLKLHPNSQPLFAFEWRDPEKGNTGQLTWTRLP QGFKNSPTLFNEALHRDLAPFRALNPQVVLLQYVDDLLVAAPTYRDCKEGTQKLLQELSKLG YRVSAKKAQLCQKEVTYLGYLLKEGKRWLTPARKATVMKIPPPTTPRQVREFLGTAGFCRL WIPGFASLAAPLYPLTKPSIPFIWTEEHQKAFDRIKEALLSAPALALPDLTKPFTLYVDERAGV ARGVLTQTLGPWRRPVAYLSKKLDPVASGWPTCLKAVAAVALLLKDADKLTLGQNVTVIAS HSLESIVRQPPDRWMTNARMTHYQSLLLNERVSFAPPAVLNPATLLPVESEATPVHRCSEILA EETGTRRDLKDQPLPGVPAWYTDGSSFIAEGKRRAGAAIVDGKRTVWASSLPEGTSAQKAEL VALTQALRLAEGKDINIYTDSRYAFATAHIHGAIYKQRGWLTSAGKDIKNKEEILALLEAIHLP KRVAIIHCPGHQKGNDPVATGNRRADEAAKQAALSTRVLAETTKP
WMSV_P03359_3mutA	VLNLEEEYRLHEKPVPSSIDPSWLQLFPTVWAERAGMGLANQVPPVVVELRSGASPVAVRQY PMSKEAREGIRPHIQRFLDLGVLVPCQSPWNTPLLPVKKPGTNDYRPVQDLREINKRVQDIHP TVPNPYNLLSSLPPSHTWYSVLDLKDAFFCLKLHPNSQPLFAFEWRDPEKGNTGQLTWTRLP QGFKNSPTLFNEALHRDLAPFRALNPQVVLLQYVDDLLVAAPTYRDCKEGTQKLLQELSKLG YRVSAKKAQLCQKEVTYLGYLLKEGKRWLTPARKATVMKIPPPTTPRQVREFLGKAGFCRL FIPGFASLAAPLYPLTKPSIPFIWTEEHQKAFDRIKEALLSAPALALPDLTKPFTLYVDERAGVA RGVLTQTLGPWRRPVAYLSKKLDPVASGWPTCLKAVAAVALLLKDADKLTLGQNVTVIASH SLESIVRQPPDRWMTNARMTHYQSLLLNERVSFAPPAVLNPATLLPVESEATPVHRCSEILAE ETGTRRDLKDQPLPGVPAWYTDGSSFIAEGKRRAGAAIVDGKRTVWASSLPEGTSAQKAELV ALTQALRLAEGKDINIYTDSRYAFATAHIHGAIYKQRGWLTSAGKDIKNKEEILALLEAIHLPK RVAIIHCPGHQKGNDPVATGNRRADEAAKQAALSTRVLAETTKP
XMRV6_A1Z651	TLNIEDEYRLHETSKEPDVPLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSEQDCQRGTRALLQTLGNLG YRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRL WIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQ GYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI LAPHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEKEAPHDC LEILAETHGTRPDLTDQPIPDADYTWYTDGSSFLQEGQRRAGAAVTTETEVIWARALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHVHGEIYRRRGLLTSEGREIKNKNEILAL LKALFLPKRLSIIHCPGHQKGNSAEARGNRMADQAAREAAMKAVLETSTLL
XMRV6_A1Z651_3mut	TLNIEDEYRLHETSKEPDVPLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSEQDCQRGTRALLQTLGNLG YRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRL WIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQ GYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVI LAPHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEKEAPHDC

[0609]

	LEILAETHGTRPDLTDQPIPDADYTWYTDGSSFLQEGQRRAGAAVTTETEVIWARALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHVHGEIYRRRGWLTSEGREIKNKNEILAL LKALFLPKRLSIIHCPGHQKGNSAEARGNRMADQAAREAAMKAVLETSTLL
XMRV6_A1Z651_3mutA	TLNIEDEYRLHETSKEPDVPLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQ YPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIH PTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP QGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSEQDCQRGTRALLQTLGNLG YRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRL FIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVIL APHAVEALVKQPPDRWLSNARMTHYQAMLLDTDRVQFGPVVALNPATLLPLPEKEAPHDCL EILAETHGTRPDLTDQPIPDADYTWYTDGSSFLQEGQRRAGAAVTTETEVIWARALPAGTSA QRAELIALTQALKMAEGKKLNVYTDSRYAFATAHVHGEIYRRRGWLTSEGREIKNKNEILAL LKALFLPKRLSIIHCPGHQKGNSAEARGNRMADQAAREAAMKAVLETSTLL

[0610]

[0611] 일부 구현예에서, 역전사효소 도메인은, 예를 들어 부위-특이적 돌연변이에 의해 변형된다. 일부 구현예에서,

역전사효소 도메인은 개선된 특성을 갖도록 조작되며, 예를 들어 SuperScript IV(SSIV) 역전사효소는 MMLV RT로부터 유래되었다. 일부 구현예에서, 역전사효소 도메인은 예를 들어 본 명세서에 참고로 포함된 WO2001068895에 기재된 바와 같이 더 낮은 오류율을 갖도록 조작될 수 있다. 일부 구현예에서, 역전사효소 도메인은 더 열안정성이 있도록 조작될 수 있다. 일부 구현예에서, 역전사효소 도메인은 더 진행성(processive)이 되도록 조작될 수 있다. 일부 구현예에서, 역전사효소 도메인은 억제제에 대한 내성을 갖도록 조작될 수 있다. 일부 구현예에서, 역전사효소 도메인은 더 신속하도록 조작될 수 있다. 일부 구현예에서, 역전사효소 도메인은 RNA 주형에서 변형된 뉴클레오티드를 더 양호하게 용인하도록 조작될 수 있다. 일부 구현예에서, 역전사효소 도메인은 변형된 DNA 뉴클레오티드를 삽입하도록 조작될 수 있다. 일부 구현예에서, 역전사효소 도메인은 주형 RNA에 결합하도록 조작된다. 일부 구현예에서, 하나 이상의 돌연변이는 뮤린 백혈병 바이러스 역전사효소의 RT 도메인 내 D200N, L603W, T330P, D524G, E562Q, D583N, P51L, S67R, E67K, T197A, H204R, E302K, F309N, W313F, L435G, N454K, H594Q, L671P, E69K, 또는 D653N로부터 선택되거나, 또는 또 다른 RT 도메인의 상응하는 위치에서의 상응하는 돌연변이이다.

[0612] 일부 구현예에서, RT 도메인(예를 들어, 표 6에 열거된 바와 같음)은 하기 표 2에 열거된 바와 같은 하나 이상의 돌연변이를 포함한다. 일부 구현예에서, 표 6에 열거된 바와 같은 RT 도메인은 하기 표 2의 상응하는 행에 열거된 돌연변이 중 1, 2, 4, 5, 6, 또는 6개를 포함한다.

[0613] 표 2. 예시적인 RT 도메인 돌연변이(표 6의 상응하는 행에 열거된 바와 같은 상응하는 야생형 서열과 비교)

RT 도메인 명칭	돌연변이(들)					
AVIRE_P03360						
AVIRE_P03360_3mut	D200N	G330P	L605W			
AVIRE_P03360_3mutA	D200N	G330P	L605W	T306K	W313F	
BAEVM_P10272						
BAEVM_P10272_3mut	D198N	E328P	L602W			
BAEVM_P10272_3mutA	D198N	E328P	L602W	T304K	W311F	
BLVAU_P25059						
BLVAU_P25059_2mut	E159Q	G286P				
BLVJ_P03361						
BLVJ_P03361_2mut	E159Q	L524W				
BLVJ_P03361_2mutB	E159Q	L524W	I97P			
FFV_O93209	D21N					
FFV_O93209_2mut	D21N	T293N	T419P			
FFV_O93209_2mutA	D21N	T293N	T419P	L393K		
FFV_O93209-Pro						
FFV_O93209-Pro_2mut	T207N	T333P				
FFV_O93209-Pro_2mutA	T207N	T333P	L307K			
FLV_P10273						
FLV_P10273_3mut	D199N	L602W				
FLV_P10273_3mutA	D199N	L602W	T305K	W312F		
FOAMV_P14350	D24N					
FOAMV_P14350_2mut	D24N	T296N	S420P			
FOAMV_P14350_2mutA	D24N	T296N	S420P	L396K		
FOAMV_P14350-Pro						
FOAMV_P14350-Pro_2mut	T207N	S331P				
FOAMV_P14350-Pro_2mutA	T207N	S331P	L307K			
GALV_P21414						
GALV_P21414_3mut	D198N	E328P	L600W			
GALV_P21414_3mutA	D198N	E328P	L600W	T304K	W311F	
HTL1A_P03362						
HTL1A_P03362_2mut	E152Q	R279P				
HTL1A_P03362_2mutB	E152Q	R279P	L90P			
HTL1C_P14078						
HTL1C_P14078_2mut	E152Q	R279P				
HTL1L_P0C211						
HTL1L_P0C211_2mut	E149Q	L527W				
HTL1L_P0C211_2mutB	E149Q	L527W	L87P			
HTL32_Q0R5R2						
HTL32_Q0R5R2_2mut	E149Q	L526W				

[0614]

HTL32_Q0R5R2_2mutB	E149Q	L526W	L87P			
HTL3P_Q4U0X6						
HTL3P_Q4U0X6_2mut	E149Q	L526W				
HTL3P_Q4U0X6_2mutB	E149Q	L526W	L87P			
HTLV2_P03363_2mut	E147Q	G274P				
JSRV_P31623						
JSRV_P31623_2mutB	A100P					
KORV_Q9TTC1	D32N					
KORV_Q9TTC1_3mut	D32N	D322N	E452P	L724W		
KORV_Q9TTC1_3mutA	D32N	D322N	E452P	L724W	T428K	W435F
KORV_Q9TTC1-Pro						
KORV_Q9TTC1-Pro_3mut	D231N	E361P	L633W			
KORV_Q9TTC1-Pro_3mutA	D231N	E361P	L633W	T337K	W344F	
MLVAV_P03356						
MLVAV_P03356_3mut	D200N	T330P	L603W			
MLVAV_P03356_3mutA	D200N	T330P	L603W	T306K	W313F	
MLVBM_Q7SVK7						
MLVBM_Q7SVK7						
MLVBM_Q7SVK7_3mut	D200N	T330P	L603W			
MLVBM_Q7SVK7_3mut	D200N	T330P	L603W			
MLVBM_Q7SVK7_3mutA_WS	D199N	T329P	L602W	T305K	W312F	
MLVBM_Q7SVK7_3mutA_WS	D199N	T329P	L602W	T305K	W312F	
MLVCB_P08361						
MLVCB_P08361_3mut	D200N	T330P	L603W			
MLVCB_P08361_3mutA	D200N	T330P	L603W	T306K	W313F	
MLVF5_P26810						
MLVF5_P26810_3mut	D200N	T330P	L603W			
MLVF5_P26810_3mutA	D200N	T330P	L603W	T306K	W313F	
MLVFF_P26809_3mut	D200N	T330P	L603W			
MLVFF_P26809_3mutA	D200N	T330P	L603W	T306K	W313F	
MLVMS_P03355						
MLVMS_P03355						
MLVMS_P03355_3mut	D200N	T330P	L603W			
MLVMS_P03355_3mut	D200N	T330P	L603W			
MLVMS_P03355_3mutA_WS	D200N	T330P	L603W	T306K	W313F	
MLVMS_P03355_3mutA_WS	D200N	T330P	L603W	T306K	W313F	
MLVMS_P03355_PLV919	D200N	T330P	L603W	T306K	W313F	H8Y
MLVMS_P03355_PLV919	D200N	T330P	L603W	T306K	W313F	H8Y
MLVRD_P11227						

[0615]

MLVRD_P11227_3mut	D200N	T330P	L603W			
MMTVB_P03365	D26N					
MMTVB_P03365	D26N					
MMTVB_P03365_2mut	D26N	G401P				
MMTVB_P03365_2mut_WS	G400P					
MMTVB_P03365_2mut_WS	G400P					
MMTVB_P03365_2mutB	D26N	G401P	V215P			
MMTVB_P03365_2mutB	D26N	G401P	V215P			
MMTVB_P03365_2mutB_WS	G400P	V212P				
MMTVB_P03365_2mutB_WS	G400P	V212P				
MMTVB_P03365_WS						
MMTVB_P03365_WS						
MMTVB_P03365-Pro						
MMTVB_P03365-Pro						
MMTVB_P03365-Pro_2mut	G309P					
MMTVB_P03365-Pro_2mut	G309P					
MMTVB_P03365-Pro_2mutB	G309P	V123P				
MMTVB_P03365-Pro_2mutB	G309P	V123P				
MPMV_P07572						
MPMV_P07572_2mutB	G289P	I103P				
PERV_Q4VFZ2						
PERV_Q4VFZ2						
PERV_Q4VFZ2_3mut	D199N	E329P	L602W			
PERV_Q4VFZ2_3mut	D199N	E329P	L602W			
PERV_Q4VFZ2_3mutA_WS	D196N	E326P	L599W	T302K	W309F	
PERV_Q4VFZ2_3mutA_WS	D196N	E326P	L599W	T302K	W309F	
SFV1_P23074	D24N					
SFV1_P23074_2mut	D24N	T296N	N420P			
SFV1_P23074_2mutA	D24N	T296N	N420P	L396K		
SFV1_P23074-Pro						
SFV1_P23074-Pro_2mut	T207N	N331P				
SFV1_P23074-Pro_2mutA	T207N	N331P	L307K			
SFV3L_P27401	D24N					
SFV3L_P27401_2mut	D24N	T296N	N422P			
SFV3L_P27401_2mutA	D24N	T296N	N422P	L396K		
SFV3L_P27401-Pro						
SFV3L_P27401-Pro_2mut	T307N	N333P				
SFV3L_P27401-Pro_2mutA	T307N	N333P	L307K			
SFVCP_Q87040	D24N					

[0616]

SFVCP_Q87040_2mut	D24N	T296N	K422P			
SFVCP_Q87040_2mutA	D24N	T296N	K422P	L396K		
SFVCP_Q87040-Pro						
SFVCP_Q87040-Pro_2mut	T207N	K333P				
SFVCP_Q87040-Pro_2mutA	T207N	K333P	L307K			
SMRVH_P03364						
SMRVH_P03364_2mut	G288P					
SMRVH_P03364_2mutB	G288P	I102P				
SRV2_P51517						
SRV2_P51517_2mutB	I103P					
WDSV_O92815						
WDSV_O92815_2mut	S183N	K312P				
WDSV_O92815_2mutA	S183N	K312P	L288K	W295F		
WMSV_P03359						
WMSV_P03359_3mut	D198N	E328P	L600W			
WMSV_P03359_3mutA	D198N	E328P	L600W	T304K	W311F	
XMRV6_A1Z651						
XMRV6_A1Z651_3mut	D200N	T330P	L603W			
XMRV6_A1Z651_3mutA	D200N	T330P	L603W	T306K	W313F	

[0617]

[0618] 일부 구현예에서, 유전자 변형 폴리펩티드는, 예를 들어 다음 서열을 포함하는, 레트로바이러스 역전사효소, 예를 들어 야생형 M-MLV RT로부터 유래된 RT 도메인을 포함한다:

[0619] *M-MLV (WT):*

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYP
MSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVP
NPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN
SPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKA
QICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAA
PLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLG
PWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPD
RWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQP
LPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGK
KLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGH
[0620] SAEARGNRMADQAARKAAITETPDTSTLLI (서열번호 2)

[0621] 일부 구현예에서, 유전자 변형 폴리펩티드는, 예를 들어 다음 서열을 포함하는, 레트로바이러스 역전사효소, 예를 들어 M-MLV RT로부터 유래된 RT 도메인을 포함한다:

TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYP
MSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVP
NPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN
SPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKA
QICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAA
PLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLG
PWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPD
RWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQP
LPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGK
KLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGH
[0622] SAEARGNRMADQAARKAAITETPDTSTLL (서열번호 3)

[0623] 일부 구현예에서, 유전자 변형 폴리펩티드는 NP_057933의 아미노산 659 내지 1329의 서열을 포함하는 레트로바이러스 역전사효소로부터 유래된 RT 도메인을 포함한다. 구현예에서, 유전자 변형 폴리펩티드는, 예를 들어 하기 나타낸 바와 같이, NP_057933의 아미노산 659 내지 1329의 서열의 N-말단에 하나의 추가 아미노산을 추가로 포함한다:

TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYP
MSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPT
VPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLP
QGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL
GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRL
WIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY
AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPH
AVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAE
AHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELI
ALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKR
LSIIHCPGHQKGHSAEARGNRMADQAARKAA (서열번호 4)

[0624]

[0625] **코어** RT(굵은 글씨), 위에 주석이 달림

[0626] RNAseH(밑줄), 위에 주석이 달림

[0627] 구현예에서, 유전자 변형 폴리펩티드는 NP_057933의 아미노산 659 내지 1329의 서열의 C-말단에 하나의 추가 아미노산을 추가로 포함한다: 구현예에서, 유전자 변형 폴리펩티드는 RNaseH1 도메인(예를 들어, NP_057933의 아미노산 1178 내지 1318)을 포함한다.

[0628] 일부 구현예에서, 레트로바이러스 역전사효소 도메인, 예를 들어, M-MLV RT는 RT의 특징, 예를 들어 열안정성, 가공성, 및/또는 주형 결합을 개선시킬 수 있는 야생형 서열로부터 하나 이상의 돌연변이를 포함할 수 있다. 일부 구현예에서, M-MLV RT 도메인은 상기 M-MLV(WT) 서열에 대하여, 예를 들어 D200N, L603W, T330P, T306K, W313F, D524G, E562Q, D583N, P51L, S67R, E67K, T197A, H204R, E302K, F309N, L435G, N454K, H594Q, D653N, R110S, K103L, 예를 들어 D200N, L603W, 및 T330P와 같은 돌연변이의 조합(선택적으로 T306K 및 W313F를 추가로 포함함)으로부터 선택되는 하나 이상의 돌연변이를 포함한다. 일부 구현예에서, 본 명세서에 사용된 M-MLV RT는 돌연변이 D200N, L603W, T330P, T306K 및 W313F를 포함한다. 구현예에서, 돌연변이체 M-MLV RT는 다음 아미노산 서열을 포함한다:

[0629] *M-MLV(PE2):*

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYP
MSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVP
NPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN
SPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKA
QICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAP
LYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGP
WRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDR
WLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPL
PDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKK
LNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHS
AEARGNRMADQAARKAAITETPDTSTLLI (서열번호 5)

[0630]

[0631] 일부 구현예에서, 라이팅 도메인(예를 들어, RT 도메인)은, 예를 들어 RNA 서열에 특이적으로 결합하는 RNA-결합 도메인을 포함한다. 일부 구현예에서, 주형 RNA는 라이팅 도메인의 RNA-결합 도메인에 의해 특이적으로 결합된 RNA 서열을 포함한다.

[0632] 일부 구현예에서, 역전사 도메인은 특정 주형, 예를 들어 시스템의 주형 RNA만을 인식하고 역전사한다. 일부 구현예에서, 주형은 역전사 도메인에 의한 인식 및 역전사를 가능하게 하는 서열 또는 구조를 포함한다. 일부 구현예에서, 주형은 본 명세서에 기재된 게놈 조작 시스템의 폴리펩티드 구성요소의 RNA-결합 도메인과의 회합을 가능하게 하는 서열 또는 구조를 포함한다. 일부 실시형태에서, 게놈 조작 시스템은 바람직하게는 회합 서열이

결여된 주형에 비해 회합 서열을 포함하는 주형을 역전사한다.

[0633] 라이팅 도메인은 또한 DNA-의존성 DNA 중합효소 활성을 포함할 수 있으며, 예를 들어 주형 DNA 서열로부터 게놈 내로 DNA를 작성할 수 있는 효소 활성을 포함할 수 있다. 일부 구현예에서, DNA-의존성 DNA 중합은 표적 부위 편집의 제2 가닥 합성을 완료하는 데 이용된다. 일부 구현예에서, DNA-의존성 DNA 중합효소 활성은 폴리펩티드 내 DNA 중합효소 도메인에 의해 제공된다. 일부 구현예에서, DNA-의존성 DNA 중합효소 활성은 또한 DNA-의존성 DNA 중합, 예를 들어 제2 가닥 합성이 가능한 역전사효소 도메인에 의해 제공된다. 일부 구현예에서, DNA-의존성 DNA 중합효소 활성은 시스템의 제2 폴리펩티드에 의해 제공된다. 일부 구현예에서, DNA-의존성 DNA 중합효소 활성은 선택적으로 게놈 조작 시스템의 구성요소에 의해 표적 부위로 동원되는 내인성 숙주 세포 중합효소에 의해 제공된다.

[0634] 일부 구현예에서, 역전사효소 도메인은 참조 역전사효소 도메인에 비해 더 낮은 확률의 시험관 내 조기 종료율(P_{off})을 갖는다. 일부 구현예에서, 참조 역전사효소 도메인은 바이러스 역전사효소 도메인, 예를 들어 M-MLV 유래의 RT 도메인이다.

[0635] 일부 구현예에서, 역전사효소 도메인은, 예를 들어 1094 nt RNA 상에서 측정 시, 약 5 x 10^{-3}/nt, 5 x 10^{-4}/nt, 또는 5 x 10^{-6}/nt 미만의, 낮은 확률의 시험관 내 조기 종료율(P_{off})을 갖는다. 구현예에서, 시험관 내 조기 종료율은 문헌[Bibillo and Eickbush (2002) J Biol Chem 277(38):34836-34845(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이 결정된다.

[0636] 일부 구현예에서, 역전사효소 도메인은 적어도 약 30% 또는 50%의 세포 내 통합을 완료할 수 있다. 완전한 통합 백분율은 실질적으로 전장 통합 사건의 수(예를 들어, 예측된 통합된 서열의 적어도 98%를 포함하는 게놈 부위)를 세포 집단에서 총(실질적으로 전장 및 부분적 통합을 포함함) 통합 사건의 수로 나눔으로써 측정될 수 있다. 구현예에서, 세포 내 통합은, 예를 들어 문헌[Karst et al. (2020) *bioRxiv* doi.org/10.1101/645903(이는 그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, (예를 들어, 통합 부위에 걸쳐) 긴-판독 암플리콘 시퀀싱을 이용하여 결정된다.

[0637] 구현예에서, 세포 내 통합의 정량화는 주형 RNA에 상응하는 DNA 서열을 적어도 약 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 함유하는 통합의 분획을 계수하는 것을 포함한다(예를 들어, 주형 RNA는 적어도 0.05, 0.1, 0.5, 0.6, 0.7, 0.8, 0.9, 1, 1.5, 2, 3, 4, 또는 5 kb, 예를 들어 길이가 0.5 내지 0.6, 0.6 내지 0.7, 0.7 내지 0.8, 0.8 내지 0.9, 1.0 내지 1.2, 1.2 내지 1.4, 1.4 내지 1.6, 1.6 내지 1.8, 1.8 내지 2.0, 2 내지 3, 3 내지 4, 또는 4 내지 5 kb의 길이를 가짐).

[0638] 일부 구현예에서, 역전사효소 도메인은 시험관 내에서 dNTP를 중합할 수 있다. 구현예에서, 역전사효소 도메인은 시험관 내에서 0.1 내지 50 nt/sec(예를 들어, 0.1 내지 1, 1 내지 10, 또는 10 내지 50 nt/sec)의 비로 dNTP를 중합할 수 있다. 구현예에서, 역전사효소 도메인에 의한 dNTP의 중합은 단일 분자 분석에 의해, 예를 들어 문헌[Schwartz and Quake (2009) *PNAS* 106(48):20294-20299(그 전체가 참고로 포함됨)]에 기재된 바와 같이 측정된다.

[0639] 일부 구현예에서, 역전사효소 도메인은, 예를 들어 문헌[Yasukawa et al. (2017) *Biochem Biophys Res Commun* 492(2):147-153(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, 1 x 10^{-3} 내지 1 x 10^{-4} 또는 1 x 10^{-4} 내지 1 x 10^{-5} 치환/nt의 시험관 내 오류율(예를 들어, 뉴클레오티드의 오혼입(misincorporation))을 갖는다. 일부 구현예에서, 역전사효소 도메인은, 예를 들어 문헌[Karst et al. (2020) *bioRxiv* doi.org/10.1101/645903(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, 예를 들어 긴-판독 암플리콘 시퀀싱에 의해, 1 x 10^{-3} 내지 1 x 10^{-4} 또는 1 x 10^{-4} 내지 1 x 10^{-5} 치환/nt의 세포 내 오류율을 갖는다.

[0640] 일부 구현예에서, 역전사효소 도메인은 시험관 내 표적 RNA의 역전사를 수행할 수 있다. 일부 구현예에서, 역전사효소는 주형의 역전사를 개시하는 데 적어도 3개 뉴클레오티드의 프라이머를 필요로 한다. 일부 구현예에서, 표적 RNA의 역전사는, 예를 들어 문헌[Bibillo and Eickbush (2002) *J Biol Chem* 277(38):34836-34845(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이 표적 RNA로부터의 cDNA의 검출에 의해 결정된다(예를 들어, ssDNA 프라이머가 제공되는 경우, 예를 들어, 3' 말단에서 적어도 3, 4, 5, 6, 7, 8, 9, 또는 10 nt를 이용하여 표적에 어닐링함).

[0641] 일부 구현예에서, 예를 들어 그의 RNA 주형을 cDNA로 전환하는 경우, 역전사효소 도메인은, 예를 들어 단백질

결합 모티프(예를 들어, 3' UTR)가 결여된 RNA 주형과 비교하여, (예를 들어, cDNA 생산에 의해) 적어도 5 또는 10배 이상 효율적으로 역전사를 수행한다. 구현예에서, 역전사의 효율은 문헌[Yasukawa et al. (2017) *Biochem Biophys Res Commun* 492(2):147-153(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이 측정된다.

[0642] 일부 구현예에서, 예를 들어 세포(예를 들어, HEK293T 세포) 내에서 발현된 경우, 역전사효소 도메인은 임의의 내인성 세포 RNA보다 더 높은 빈도(예를 들어, 약 5 또는 10배 더 높은 빈도)로 특이적 RNA 주형에 특이적으로 결합한다. 구현예에서, 역전사효소 도메인과 주형 RNA간의 특이적 결합의 빈도는, 예를 들어 문헌[Lin and Miles (2019) *Nucleic Acids Res* 47(11):5490-5501(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이 CLIP-seq에 의해 측정된다.

[0643] *주형 핵산 결합 도메인*

[0644] 유전자 변형 폴리펩티드는 통상적으로 주형 핵산(예를 들어, 주형 RNA)과 회합할 수 있는 영역을 함유한다. 일부 구현예에서, 주형 핵산 결합 도메인은 RNA 결합 도메인이다. 일부 구현예에서, RNA 결합 도메인은 특이적 시그니처, 예를 들어 구조적 모티프를 함유하는 RNA 분자와 회합할 수 있는 모듈 도메인이다. 다른 구현예에서, 주형 핵산 결합 도메인(예를 들어, RNA 결합 도메인)은 역전사 도메인 내에 함유되고, 예를 들어 역전사효소-유래된 구성요소는 RNA 선호도에 대한 알려진 시그니처를 갖는다.

[0645] 다른 구현예에서, 주형 핵산 결합 도메인(예를 들어, RNA 결합 도메인)은 표적 DNA 결합 도메인 내에 포함된다. 예를 들어, 일부 구현예에서, DNA 결합 도메인은 gRNA를 포함하는 주형 핵산(예를 들어, 주형 RNA)의 구조를 인식하는 CRISPR-연관 단백질이다. 일부 구현예에서, 유전자 변형 폴리펩티드는 DNA-결합 도메인이 표적 게놈 DNA 서열에 결합하도록 허용하는 gRNA 스캐폴드와 회합하는 CRISPR-연관 단백질을 포함하는 DNA-결합 도메인을 포함한다. 일부 구현예에서, gRNA 스캐폴드 및 gRNA 스페이서는 주형 핵산(예를 들어, 주형 RNA) 내에 포함되며, 따라서 DNA-결합 도메인은 또한 주형 핵산 결합 도메인이다. 일부 구현예에서, 폴리펩티드는 다수의 도메인에서 RNA 결합 기능을 가지며, 예를 들어 CRISPR-연관 DNA 결합 도메인에서 gRNA 구조 및 역전사효소 도메인 내 추가의 서열 또는 구조에 결합할 수 있다.

[0646] 일부 구현예에서, RNA 결합 도메인은 참조 RNA 결합 도메인보다 더 큰 친화도로 주형 RNA에 결합할 수 있다. 일부 구현예에서, 참조 RNA 결합 도메인은 S. 피오게네스(S. pyogenes)의 Cas9로부터의 RNA 결합 도메인이다. 일부 구현예에서, RNA 결합 도메인은 100 pM 내지 10 nM(예를 들어, 100 pM 내지 1 nM 또는 1 nM 내지 10 nM)의 친화도로 주형 RNA에 결합할 수 있다. 일부 구현예에서, 그의 주형 RNA에 대한 RNA 결합 도메인의 친화도는 시험관 내에서, 예를 들어 문헌[Asmari et al. Methods 146:107-119 (2018)(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, 예를 들어 열영동에 의해 측정된다. 일부 구현예에서, 그의 주형 RNA에 대한 RNA 결합 도메인의 친화도는 세포 내에서 (예를 들어, FRET 또는 CLIP-Seq에 의해) 측정된다.

[0647] 일부 구현예에서, RNA 결합 도메인은 스크램블드(scrambled) RNA를 이용한 경우보다 적어도 약 5배 또는 10배 더 높은 빈도로 시험관 내에서 주형 RNA와 회합된다. 일부 구현예에서, RNA 결합 도메인과 주형 RNA 또는 스크램블드 RNA 간의 회합의 빈도는, 예를 들어 문헌[Lin and Miles (2019) *Nucleic Acids Res* 47(11):5490-5501(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이 CLIP-seq에 의해 측정된다. 일부 구현예에서, RNA 결합 도메인은 스크램블드 RNA를 이용한 경우보다 적어도 약 5배 또는 10배 더 높은 빈도로 세포(예를 들어, HEK293T 세포) 내에서 주형 RNA와 회합된다. 일부 구현예에서, RNA 결합 도메인과 주형 RNA 또는 스크램블드 RNA 간의 회합의 빈도는, 예를 들어 문헌[Lin and Miles (2019) 상동]에 기재된 바와 같이 CLIP-seq에 의해 측정된다.

[0648] *RNA 결합 도메인(RBD)*

[0649] 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드는 RNA 결합 도메인(RBD)을 포함한다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드는 표 31에 열거된 바와 같은 RBD의 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 RBD를 포함한다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드의 RBD는, 예를 들어 표 31에 열거된 바와 같은 RNA 결합 파트너에 결합한다. 구현예에서, RBD는 표 31의 어느 한 행에 열거된 바와 같은 RBD의 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하고, 표 31의 동일한 행에 열거된 RNA 결합 파트너에 결합한다.

[0650] **표 31. 예시적인 RNA 결합 도메인 서열**

명칭	RNA 결합 파트너	아미노산 서열
MCP_v1	MS2	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGRA
MCP_v2	MS2	MASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQKRKYTIKVEVPKVATQTVGGVELPVAAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGLY
PCP	PP7	MGSMSKTIVLSVGEATRTLTEIQSTADRQIFEEKVGPLVGRLRLTASLRQNGAKTAYRVNLKLDQADVVDSGLPKVRYTQVWSHDVTIVANSTEASRKSLYDLTKSLVATSQVEDLVVNLVPLGRGSGRA
Com	com	MGSMKSIRCKNCNKLLFKADSFDHIEIRCPRCKRHIIMLNACEHPTEKHCGKREKITHSDETVRYGSGRA
LS4	LS4-1, CS1	YVRFEVPEDMQNEALSLLEKVRESGKVKKGTNRTTHAVYRGLAKLVYIAEDVDPPEIVAHLPLLCEEKNVPYIYVKSKNDLGRAVGGWPIGASAAIINEGELRKELGSLVEKIKGLQKRSHMHLE
LS12	LS12-1, CS2	YVRFEVPEDMQNEALSLLEKVRESGKVKKGTNSTTLAVSRGLAKLVYIAEDVDPPEIVAHLPLLCEEKNVPYIYVKSKNDLGRAVGRVYPGASAAIINEGELRKELGSLVEKIKGLQKRSHMHLE
람다 N(1-22)	BoxB	MDAQTRRRERRAEKQAQWKAAN
L7Ae	Kt	MYVRFEVPEDMQNEALSLLEKVRESGKVKKGTNETTKAVERGLAKLVYIAEDVDPPEIVAHLPLLCEEKNVPYIYVKSKNDLGRAVGIEVPCASAAIINEGELRKELGSLVEKIKGLQK
L7Ae	Kt	YVRFEVPEDMQNEALSLLEKVRESGKVKKGTNETTKAVERGLAKLVYIAEDVDPPEIVAHLPLLCEEKNVPYIYVKSKNDLGRAVGIEVPCASAAIINEGELRKELGSLVEKIKGLQKRSHMHLE

[0651]

[0652] *엔도뉴클레아제 도메인 및 DNA 결합 도메인*

[0653] 일부 구현예에서, 유전자 변형 폴리펩티드는 엔도뉴클레아제 도메인을 통한 DNA 표적 부위 절단 기능을 갖는다. 일부 구현예에서, 유전자 변형 폴리펩티드는, 예를 들어 표적 핵산에 결합하기 위한, DNA 결합 도메인을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드의 도메인(예를 들어, Cas 도메인)은 2개 이상의 더 작은 도메인, 예를 들어 DNA 결합 도메인 및 엔도뉴클레아제 도메인을 포함한다. DNA 결합 도메인(예를 들어, Cas 도메인)이 표적 핵산 서열에 결합한다고 할 때, 일부 구현예에서 결합은 gRNA에 의해 매개되는 것으로 이해된다.

[0654] 일부 구현예에서, 도메인은 2가지 기능을 갖는다. 예를 들어, 일부 구현예에서, 엔도뉴클레아제 도메인은 또한 DNA-결합 도메인이다. 일부 구현예에서, 엔도뉴클레아제 도메인은 또한 주형 핵산(예를 들어, 주형 RNA) 결합 도메인이다. 예를 들어, 일부 구현예에서, 폴리펩티드는 gRNA를 포함하는 주형 RNA에 결합하고, (예를 들어, gRNA의 일부에 대한 상보성을 갖는) 표적 DNA 서열에 결합하고, 표적 DNA 서열을 절단하는 CRISPR-연관 엔도뉴클레아제 도메인을 포함한다. 일부 구현예에서, 이종성 공급원 유래의 엔도뉴클레아제 도메인 또는 엔도뉴클레아제/DNA-결합 도메인이 본 명세서에 기재된 유전자 변형 시스템에서 사용될 수 있거나 (예를 들어, 하나 이상의 잔기의 삽입, 결실, 또는 치환에 의해) 변형될 수 있다.

[0655] 일부 구현예에서, 엔도뉴클레아제 도메인 또는 엔도뉴클레아제/DNA 결합 도메인을 인코딩하는 핵산은, 예를 들어 인간 세포에 대해 개선된, 변경된 코돈 사용을 갖기 위해 그의 천연 서열에서 변경된다. 일부 구현예에서, 엔도뉴클레아제 요소는 이종성 엔도뉴클레아제 요소, 예컨대 Cas 엔도뉴클레아제(예를 들어, Cas9), II형 제한 엔도뉴클레아제(예를 들어, Fok1), 메가뉴클레아제(예를 들어, I-SceI), 또는 다른 엔도뉴클레아제 도메인이다.

[0656] 특정 양태에서, 본 명세서에 기재된 유전자 변형 폴리펩티드의 DNA-결합 도메인은 원하는 숙주 DNA 표적 서열에 대한 결합을 위해 선택, 설계 또는 구축된다. 특정 구현예에서, 폴리펩티드의 DNA-결합 도메인은 이종성 DNA-결합 요소이다. 일부 구현예에서, 이종성 DNA 결합 요소는 징크-핑거 요소 또는 TAL 이펙터 요소, 예를 징크-핑거 또는 TAL 폴리펩티드 또는 이의 기능성 단편이다. 일부 구현예에서, 이종성 DNA 결합 요소는 서열-가이드된 DNA 결합 요소, 예컨대 Cas9, Cpf1, 또는 엔도뉴클레아제 활성을 갖지 않도록 변경된 기타 CRISPR-관련 단백질이다. 일부 구현예에서, 이종성 DNA 결합 요소는 엔도뉴클레아제 활성을 보유한다. 일부 구현예에서, 이종성 DNA 결합 요소는 ssDNA를 절단하는 부분적인 엔도뉴클레아제 활성을 보유하며, 예를 들어 닉카제 활성을 갖는다. 특정 구현예에서, 이종성 DNA 결합 도메인은 Cas9, TAL 도메인, ZF 도메인, Myb 도메인, 이의 조합, 또는 이의 다수 중

임의의 하나 이상일 수 있다.

[0657] 일부 구현예에서, DNA-결합 도메인은, 예를 들어 부위-특이적 돌연변이, DNA-결합 요소(예를 들어, 징크 핑거의 수 및/또는 특이성)의 증가 또는 감소 등에 의해 변형되어, DNA-결합 특이성 및 친화도를 변경시킨다. 일부 구현예에서, DNA 결합 도메인을 인코딩하는 핵산 서열은, 예를 들어 인간 세포에 대해 개선된, 변경된 코돈 사용을 갖기 위해 그의 천연 서열에서 변경된다. 구현예에서, DNA 결합 도메인은 야생형 DNA 결합 도메인에 대한 하나 이상의 변형, 예를 들어 지향된 진화, 예를 들어 파지-보조된 지속적 진화(phage-assisted continuous evolution; PACE)를 통한 변형을 포함한다.

[0658] 일부 구현예에서, DNA 결합 도메인은 메가뉴클레아제 도메인(예를 들어, 본 명세서에, 예를 들어 엔도뉴클레아제 도메인 섹션에 기재된 바와 같음), 또는 이의 기능성 단편을 포함한다. 일부 구현예에서, 메가뉴클레아제 도메인은 엔도뉴클레아제 활성, 예를 들어 이중 가닥 절단 및/또는 닉카제 활성을 갖는다. 다른 구현예에서, 메가뉴클레아제 도메인은 감소된 활성을 갖고, 예를 들어 엔도뉴클레아제 활성이 결여되고, 예를 들어 메가뉴클레아제는 촉매적으로 비활성이다. 일부 구현예에서, 촉매적으로 비활성인 메가뉴클레아제는, 예를 들어 그 전체가 본 명세서에 참고로 포함된 문헌[Fonfara et al. Nucleic Acids Res 40(2):847-860 (2012)]에 기재된 바와 같이, DNA 결합 도메인으로 사용된다.

[0659] 일부 구현예에서, 유전자 변형 폴리펩티드는, 예를 들어 야생형 폴리펩티드에 비해 DNA-결합 도메인에 대한 변형을 포함한다. 일부 구현예에서, DNA-결합 도메인은 원래 DNA-결합 도메인의 아미노산 서열에 대한 추가, 결실, 대체, 또는 변형을 포함한다. 일부 구현예에서, DNA-결합 도메인은 관심 표적 핵산(예를 들어, DNA) 서열에 특이적으로 결합하는 이종 기능성 도메인을 포함하도록 변형된다. 일부 구현예에서, 기능성 도메인은 폴리펩티드의 이전 DNA-결합 도메인의 적어도 일부(예를 들어, 전체)를 대체한다. 일부 구현예에서, 기능성 도메인은 징크 핑거(예를 들어, 관심 표적 핵산(예를 들어, DNA)에 특이적으로 결합하는 징크 핑거)를 포함한다. 일부 구현예에서, 기능성 도메인은 Cas 도메인(예를 들어, 관심 표적 핵산(예를 들어, DNA)에 특이적으로 결합하는 Cas 도메인)을 포함한다. 일부 구현예에서, Cas 도메인은 Cas9 또는 이의 돌연변이체 또는 변이체(예를 들어, 본 명세서에 기재된 바와 같음)를 포함한다. 구현예에서, Cas 도메인은, 예를 들어 본 명세서에 기재된 바와 같이 가이드 RNA(gRNA)와 회합된다. 구현예에서, Cas 도메인은 gRNA에 의해 관심 표적 핵산(예를 들어, DNA) 서열로 지시된다. 구현예에서, Cas 도메인은 gRNA와 동일한 핵산(예를 들어, RNA) 분자에서 인코딩된다. 구현예에서, Cas 도메인은 gRNA와 상이한 핵산(예를 들어, RNA) 분자에서 인코딩된다.

[0660] 일부 구현예에서, DNA 결합 도메인은 참조 DNA 결합 도메인보다 더 큰 친화도로 표적 서열(예를 들어, dsDNA 표적 서열)에 결합할 수 있다. 일부 구현예에서, 참조 DNA 결합 도메인은 S. 피오게네스의 Cas9로부터의 DNA 결합 도메인이다. 일부 구현예에서, DNA 결합 도메인은 100 pM 내지 10 nM(예를 들어, 100 pM 내지 1 nM 또는 1 nM 내지 10 nM)의 친화도로 표적 서열(예를 들어, dsDNA 표적 서열)에 결합할 수 있다.

[0661] 일부 구현예에서, 그의 표적 서열(예를 들어, dsDNA 표적 서열)에 대한 DNA 결합 도메인의 친화도는 시험관 내에서, 예를 들어 문헌[Asmari et al. Methods 146:107-119 (2018)(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, 예를 들어 열영동에 의해 측정된다.

[0662] 구현예에서, DNA 결합 도메인은 몰 과량, 예를 들어 약 100배 몰 과량의 스크램블드 서열 경쟁자 dsDNA 존재 하에서, 100 pM 내지 10 nM(예를 들어, 100 pM 내지 1 nM 또는 1 nM 내지 10 nM) 사이의 친화도로 표적 서열(예를 들어, dsDNA 표적 서열)에 결합할 수 있다.

[0663] 일부 구현예에서, DNA 결합 도메인은, 예를 들어 ChIP-seq(예를 들어, HEK293T 세포에서)에 의해, 예를 들어 문헌[He and Pu (2010) *Curr. Protoc Mol Biol* Chapter 21(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이 측정 시, 표적 세포, 예를 들어 인간 표적 세포의 게놈 내의 임의의 다른 서열보다 더 빈번하게 그의 표적 서열(예를 들어, dsDNA 표적 서열)과 회합하는 것으로 발견된다. 일부 구현예에서, DNA 결합 도메인은, 예를 들어 ChIP-seq(예를 들어, HEK293T 세포에서)에 의해, 예를 들어 문헌[He and Pu (2010), 상동]에 기재된 바와 같이 측정 시, 표적 세포의 게놈 내의 임의의 다른 서열보다 적어도 약 5배 또는 10배 더 빈번하게 그의 표적 서열(예를 들어, dsDNA 표적 서열)과 회합하는 것으로 발견된다.

[0664] 일부 구현예에서, 엔도뉴클레아제 도메인은 닉카제 활성을 갖고, 표적 DNA의 하나의 가닥을 절단한다. 일부 구현예에서, 닉카제 활성은 표적 부위에서 이중 가닥 파손의 형성을 감소시킨다. 일부 구현예에서, 엔도뉴클레아제 도메인은 표적 DNA의 제1 가닥과 제2 가닥에 엇갈린(staggered) 닉 구조를 형성한다. 일부 구현예에서, 엇갈린 닉 구조는 표적 부위에 자유 3' 오버행(overhang)을 생성한다. 일부 구현예에서, 표적 부위의 자유 3' 오버

행은, 예를 들어 주형 핵산의 3' 상동성 영역의 접근 및 어닐링을 향상시킴으로써 편집 효율성을 개선시킨다. 일부 구현예에서, 엇갈린 닉 구조는 표적 부위에서 이중 가닥 파손의 형성을 감소시킨다.

[0665] 일부 구현예에서, 엔도뉴클레아제 도메인은 표적 DNA의 가닥 둘 다를 절단하며, 예를 들어 절단 부위의 양쪽에 ssDNA 오버행이 없는 표적의 평활 말단(blunt-end) 절단을 초래한다. 본 명세서에 기재된 유전자 변형 시스템의 엔도뉴클레아제 도메인의 아미노산 서열은 본 명세서에 기재된 엔도뉴클레아제 도메인, 예를 들어 본 명세서에 기재된 바와 같은 엔도뉴클레아제 도메인의 아미노산 서열에 대하여 적어도 약 50%, 적어도 약 60%, 적어도 약 70%, 적어도 약 80%, 적어도 약 85%, 적어도 약 90%, 적어도 약 95%, 적어도 약 96%, 적어도 약 97%, 적어도 약 98%, 적어도 약 99% 동일할 수 있다.

[0666] 특정 구현예에서, 이종성 엔도뉴클레아제는 Fok1 또는 이의 기능성 단편이다. 특정 구현예에서, 이종성 엔도뉴클레아제는 홀리데이 접합 리졸바제(resolvase) 또는 이의 동족체, 예컨대 솔폴로부스 솔파타리쿠스-솔헤(Sulfolobus solfataricus--Ssol Hje)로부터의 홀리데이 접합 리졸바제이다(Govindaraju et al., *Nucleic Acids Research* 44:7, 2016). 특정 구현예에서, 이종성 엔도뉴클레아제는 Prp8과 같은 스플라이스좀 단백질의 큰 단편의 엔도뉴클레아제이다(Mahbub et al., *Mobile DNA* 8:16, 2017). 특정 구현예에서, 이종성 엔도뉴클레아제는 CRISPR-연관 단백질, 예를 들어 Cas9로부터 유래된다. 특정 구현예에서, 이종성 엔도뉴클레아제는 단지 ssDNA 절단 활성만, 예를 들어 닉카제 활성만 갖도록 조작되며, 예를 들어 Cas9 닉카제, 예를 들어 D10A, H840A, 또는 N863A 돌연변이를 갖는 SpCas9가 된다. 표 8은 예시적인 Cas 단백질 및 닉카제 활성과 연관된 돌연변이를 제공한다. 또 다른 구현예에서, 동종성 엔도뉴클레아제 도메인은, 예를 들어 부위-특이적 돌연변이에 의해 변형되어, DNA 엔도뉴클레아제 활성을 변경시킨다. 또 다른 구현예에서, 엔도뉴클레아제 도메인은, 예를 들어 DNA-서열 특이성을 부여하는 도메인을 제거하기 위한 절단 또는 DNA-서열 특이성을 부여하는 영역을 비활성화시키기 위한 돌연변이에 의해, DNA-서열 특이성을 감소시키도록 변형된다.

[0667] 일부 구현예에서, 엔도뉴클레아제 도메인은 닉카제 활성을 갖고, 이중 가닥 파손을 형성하지 않는다. 일부 구현예에서, 엔도뉴클레아제 도메인은 이중 가닥 파손보다 더 높은 빈도로 단일 가닥 파손을 형성하며, 예를 들어 적어도 90%, 95%, 96%, 97%, 98%, 또는 99%의 절단은 단일 가닥 파손이거나, 또는 10%, 5%, 4%, 3%, 2%, 또는 1% 미만의 절단은 이중 가닥 파손이다. 일부 구현예에서, 엔도뉴클레아제는 실질적으로 이중 가닥 파손을 형성하지 않는다. 일부 구현예에서, 엔도뉴클레아제는 검출가능한 수준의 이중 가닥 파손을 형성하지 않는다.

[0668] 일부 구현예에서, 엔도뉴클레아제 도메인은 제1 가닥의 표적 부위 DNA에 닉을 생성하는 닉카제 활성을 갖고; 예를 들어, 일부 구현예에서 엔도뉴클레아제 도메인은 라이팅 도메인에 의해 연장될 가닥 상에서 변경 부위 근처의 표적 부위의 게놈 DNA를 절단한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 제1 가닥의 표적 부위 DNA에 닉을 생성하고 제2 가닥의 표적 부위 DNA에 닉을 생성하지 않는 닉카제 활성을 갖는다. 예를 들어, 폴리펩티드가 닉카제 활성을 갖는 CRISPR-연관 엔도뉴클레아제 도메인을 포함하는 경우, 일부 구현예에서 상기 CRISPR-연관 엔도뉴클레아제 도메인은 PAM 부위를 함유하는 표적 부위 DNA 가닥에 닉을 생성한다(예를 들어, 그리고 PAM 부위를 함유하지 않는 표적 부위 DNA 가닥에 닉을 생성하지 않음). 추가의 예로서, 폴리펩티드가 닉카제 활성을 갖는 CRISPR-연관 엔도뉴클레아제 도메인을 포함하는 경우, 일부 구현예에서 상기 CRISPR-연관 엔도뉴클레아제 도메인은 PAM 부위를 함유하지 않는 표적 부위 DNA 가닥에 닉을 생성한다(예를 들어, 그리고 PAM 부위를 함유하는 표적 부위 DNA 가닥에 닉을 생성하지 않음).

[0669] 일부 다른 구현예에서, 엔도뉴클레아제 도메인은 제1 가닥 및 제2 가닥의 표적 부위 DNA에 닉을 생성하는 닉카제 활성을 갖는다. 이론에 구속되고자 함이 없이, 본 명세서에 기재된 폴리펩티드의 라이팅 도메인(예를 들어, RT 도메인)이 주형 핵산(예를 들어, 주형 RNA)의 이종 목적 서열로부터 중합(예를 들어, 역전사)된 후, 세포 DNA 복구 기구는 제1 DNA 가닥 상의 닉을 복구해야 한다. 표적 부위 DNA는 이제 제1 DNA 가닥에 대해 2개의 상이한 서열을 함유하며, 하나는 원래 게놈 DNA(예를 들어, 자유 5' 말단을 가짐)에 상응하고 두 번째는 이종 목적 서열(예를 들어, 자유 3' 말단을 가짐)로부터 중합된 것에 상응한다. 2개의 상이한 서열은 서로 평형을 이루는 것으로 여겨지며, 첫 번째 서열이 두 번째 가닥을 혼성화한 다음 다른 하나를 혼성화하고, 세포 DNA 복구 장치의 서열이 복구된 표적 부위로 통합하는 것은 확률적 공정일 수 있다. 이론에 구속되고자 함이 없이, 제2 가닥에 추가의 닉을 도입하는 것은 세포 DNA 복구 기구가 원래 게놈 서열보다 이종 목적 서열 기반 서열을 더 빈번하게 채용하도록 편향시킬 수 있는 것으로 생각된다(Anzalone et al. Nature 576:149-157 (2019)). 일부 구현예에서, 추가의 닉은 표적 부위 변형(예를 들어, 삽입, 결실, 또는 치환)의 적어도 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 또는 150개의 뉴클레오티드 5' 또는 3' 또는 제1 가닥의 닉에 위치한다.

[0670] 대안적으로 또는 추가적으로, 이론에 구속되고자 함이 없이, 제2 가닥에 대한 추가의 닉은 제2 가닥 합성을 촉진할 수 있는 것으로 여겨진다. 일부 구현예에서, 유전자 변형 시스템이 제1 가닥의 일부를 삽입하거나 치환하는 경우, 제2 가닥에서의 삽입/치환에 상응하는 새로운 서열의 합성이 필요하다.

[0671] 일부 구현예에서, 폴리펩티드는 엔도뉴클레아제 활성을 갖는 단일 도메인(예를 들어, 단일 엔도뉴클레아제 도메인)을 포함하고, 상기 도메인은 제1 가닥 및 제2 가닥 둘 모두에 닉을 생성한다. 예를 들어, 이러한 구현예에서 엔도뉴클레아제 도메인은 CRISPR-연관 엔도뉴클레아제 도메인일 수 있고, 주형 핵산(예를 들어, 주형 RNA)은 제1 가닥의 닉 생성을 지시하는 gRNA 스페이서 및 제2 가닥의 닉 생성을 지시하는 추가의 gRNA 스페이서를 포함한다. 일부 구현예에서, 폴리펩티드는 엔도뉴클레아제 활성을 갖는 복수의 도메인을 포함하고, 제1 엔도뉴클레아제 도메인은 제1 가닥에 닉을 생성하고 제2 엔도뉴클레아제 도메인은 제2 가닥에 닉을 생성한다(선택적으로, 제1 엔도뉴클레아제 도메인은 제2 가닥에 닉을 생성하지 않고(예를 들어, 닉을 생성할 수 없음), 제2 엔도뉴클레아제 도메인은 제1 가닥에 닉을 생성하지 않는다(예를 들어, 닉을 생성할 수 없음).

[0672] 일부 구현예에서, 엔도뉴클레아제 도메인은 제1 가닥 및 제2 가닥에 닉을 생성할 수 있다. 일부 구현예에서, 제1 및 제2 가닥 닉은 표적 부위 내 동일 위치에서 발생하지만 반대 가닥에서 발생한다. 일부 구현예에서, 제2 가닥 닉은 엇갈린(staggered) 위치, 예를 들어 제1 닉으로부터 상류 또는 하류에서 발생한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 제2 가닥 닉이 제1 가닥 닉의 상류에 있는 경우 표적 부위 결실을 생성한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 제2 가닥 닉이 제1 가닥 닉의 하류에 있는 경우 표적 부위 복제를 생성한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 제1 및 제2 가닥 닉이 표적 부위의 동일 위치에서 발생하는 경우 복제 및/또는 결실을 생성하지 않는다. 일부 구현예에서, 엔도뉴클레아제 도메인은, 예를 들어 제1 또는 제2 가닥의 닉 생성을 촉진하는 단백질 적합(conformation) 또는 RNA-결합 상태에 따라 변경된 활성을 갖는다(예를 들어, 그 전체가 본 명세서에 참고로 포함된 문헌[Christensen et al. PNAS 2006]에 기재된 바와 같음).

[0673] 일부 구현예에서, 엔도뉴클레아제 도메인은 메가뉴클레아제, 또는 이의 기능성 단편을 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 호밍(homing) 엔도뉴클레아제, 또는 이의 기능성 단편을 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 LAGLIDADG, GIY-YIG, HNH, His-Cys Box, 또는 PD-(D/E) XK 패밀리로부터의 예를 들어 패밀리 명칭에 표시된 바와 같이 보존된 아미노산 모티프를 갖는 메가뉴클레아제, 또는 이의 기능성 단편 또는 변이체를 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 예를 들어 I-SmaMI(Uniprot F7WD42), I-SceI(Uniprot P03882), I-AniI(Uniprot P03880), I-DmoI(Uniprot P21505), I-CreI(Uniprot P05725), I-TevI(Uniprot P13299), I-OnuI(Uniprot Q4VWW5), 또는 I-BmoI(Uniprot Q9ANR6)로부터 선택되는 메가뉴클레아제, 또는 이의 단편을 포함한다. 일부 구현예에서, 메가뉴클레아제는 기능성 형태에서 자연적으로 단량체, 예를 들어 I-SceI, I-TevI이며, 또는 이량체, 예를 들어 I-CreI이다. 예를 들어, LAGLIDADG 모티프의 단일 카피를 갖는 LAGLIDADG 메가뉴클레아제는 일반적으로 동종이량체를 형성하는 한편, LAGLIDADG 모티프의 2개의 카피를 갖는 구성원은 일반적으로 단량체로 발견된다. 일부 구현예에서, 보통 이량체로 형성되는 메가뉴클레아제는 융합물로서 발현되며, 예를 들어 2개의 서브유닛은 단일 ORF로 발현되고, 선택적으로, 링커, 예를 들어 I-CreI 이량체 융합물에 의해 연결된다(Rodriguez-Fornes et al. Gene Therapy 2020; 그 전체가 본 명세서에 참고로 포함됨). 일부 구현예에서, 메가뉴클레아제 또는 이의 기능성 단편은 이중 가닥 DNA 분자, 예를 들어, I-SceI(K122I 및/또는 K223I)(Niu et al. J Mol Biol 2008), I-AniI(K227M)(McConnell Smith et al. PNAS 2009), I-DmoI(Q42A 및/또는 K120M)(Molina et al. J Biol Chem 2015) 중 하나의 가닥에 대한 닉카제 활성을 선호하도록 변경된다. 일부 구현예에서, 메가뉴클레아제 또는 단일가닥 절단에 대한 이러한 선호도를 갖는 이의 기능성 단편은, 예를 들어 닉카제 활성을 갖는 엔도뉴클레아제 도메인으로 사용된다. 일부 구현예에서, 엔도뉴클레아제 도메인은 세이프 하버 부위, 예를 들어 I-CreI 표적화 SH6 부위를 자연적으로 표적화하거나 표적화하도록 조작된 메가뉴클레아제 또는 이의 기능성 단편을 포함한다(Rodriguez-Fornes et al., 상동). 일부 구현예에서, 엔도뉴클레아제 도메인은 서열 내성 촉매 도메인, 예를 들어 최소 모티프 CNNNG를 인식하는 I-TevI를 갖는 메가뉴클레아제 또는 이의 기능성 단편을 포함한다(Kleinstiver et al. PNAS 2012). 일부 구현예에서, 표적 서열 내성 촉매 도메인은, 예를 들어 활성을 지시하기 위해, 예를 들어 I-TevI를 (i) Tev-ZFE를 생성하기 위한 징크 핑거(Kleinstiver et al. PNAS 2012), (ii) MegaTevs를 생성하기 위한 기타 메가뉴클레아제(Wolfs et al. Nucleic Acids Res 2014), 및/또는 (iii) TevCas9를 생성하기 위한 Cas9(Wolfs et al. PNAS 2016)에 융합시킴으로써, DNA 결합 도메인에 융합된다.

[0674] 일부 구현예에서, 엔도뉴클레아제 도메인은 제한 효소, 예를 들어 IIS형 또는 IIP형 제한 효소를 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 IIS형 제한 효소, 예를 들어 FokI, 또는 이의 단편 또는 변이체를 포

함한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 IIP형 제한 효소, 예를 들어 PvuII, 또는 이의 단편 또는 변이체를 포함한다. 일부 구현예에서, 이량체성 제한 효소는, 그가 단일 사슬로서 기능하도록 하는 융합물, 예를 들어 FokI 이량체 융합물로서 발현된다(Minczuk et al. Nucleic Acids Res 36(12):3926-3938 (2008)).

[0675] 추가적인 엔도뉴클레아제 도메인의 사용은, 예를 들어 문헌[Guha and Edgell Int J Mol Sci 18(22):2565 (2017)]에 설명되어 있으며, 이는 그 전체가 본 명세서에 참고로 포함된다.

[0676] 일부 구현예에서, 유전자 변형 폴리펩티드는, 예를 들어 야생형 Cas 단백질에 비해 엔도뉴클레아제 도메인에 대한 변형을 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 야생형 Cas 단백질의 아미노산 서열에 대한 추가, 결실, 대체, 또는 변형을 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인은 관심 표적 핵산(예를 들어, DNA) 서열에 특이적으로 결합하고/하거나 이의 엔도뉴클레아제 절단을 유도하는 이종 기능성 도메인을 포함하도록 변형된다. 일부 구현예에서, 엔도뉴클레아제 도메인은 징크 핑거를 포함한다. 구현예에서, Cas 도메인을 포함하는 엔도뉴클레아제 도메인은, 예를 들어 본 명세서에 기재된 바와 같이 가이드 RNA(gRNA)와 회합된다. 일부 구현예에서, 엔도뉴클레아제 도메인은 특정 표적 핵산(예를 들어, DNA) 서열을 표적화하지 않는 기능성 도메인을 포함하도록 변형된다. 구현예에서, 엔도뉴클레아제 도메인은 Fok1 도메인을 포함한다.

[0677] 일부 구현예에서, 엔도뉴클레아제 도메인은 스크램블드 dsDNA를 이용한 경우보다 적어도 약 5배 또는 10배 더 높은 빈도로 시험관 내에서 표적 dsDNA와 회합된다. 일부 구현예에서, 엔도뉴클레아제 도메인은, 예를 들어 세포(예를 들어, HEK293T 세포)에서, 스크램블드 dsDNA를 이용한 경우보다 적어도 약 5배 또는 10배 더 높은 빈도로 시험관 내에서 표적 dsDNA와 회합된다. 일부 구현예에서, 엔도뉴클레아제 도메인과 표적 DNA 또는 스크램블드 DNA 간의 회합의 빈도는, 예를 들어 문헌[He and Pu (2010) *Curr. Protoc Mol Biol* Chapter 21(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이 ChIP-seq에 의해 측정된다.

[0678] 일부 구현예에서, 엔도뉴클레아제 도메인은, 예를 들어 비-표적 서열과 비교하여(예를 들어, 표적 세포의 게놈에서 임의의 다른 게놈 서열과 비교하여) 적어도 약 5배 또는 약 10배의 증가로 표적 서열에서 닉의 형성을 촉매할 수 있다. 일부 구현예에서, 닉 형성의 수준은, 예를 들어 문헌[Elacqua et al. (2019) *bioRxiv* doi.org/10.1101/867937(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, NickSeq을 사용하여 결정된다.

[0679] 일부 구현예에서, 엔도뉴클레아제 도메인은 시험관 내에서 DNA에 닉을 생성할 수 있다. 구현예에서, 닉은 노출된 염기를 생성한다. 구현예에서, 노출된 염기는, 예를 들어 문헌[Chaudhry and Weinfeld (1995) *Nucleic Acids Res* 23(19):3805-3809(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, 뉴클레아제 민감도 분석을 사용하여 검출될 수 있다. 구현예에서, (예를 들어, 뉴클레아제 민감도 분석에 의해 검출된) 노출된 염기의 수준은 참조 엔도뉴클레아제 도메인과 비교하여 적어도 10%, 50% 또는 그 이상 증가된다. 일부 구현예에서, 참조 엔도뉴클레아제 도메인은 S. 피오게네스의 Cas9로부터의 엔도뉴클레아제 도메인이다.

[0680] 일부 구현예에서, 엔도뉴클레아제 도메인은 세포 내에서 DNA에 닉을 생성할 수 있다. 구현예에서, 엔도뉴클레아제 도메인은 HEK293T 세포 내에서 DNA에 닉을 생성할 수 있다. 구현예에서, Rad51 부재 하에서 복제를 진행하는 복구되지 않는 닉은 닉 부위에서 증가된 NHEJ 비를 초래하며, 이는, 예를 들어 문헌[Bothmer et al. (2017) *Nat Commun* 8:13905(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, 예를 들어 Rad51 억제 분석을 사용하여 검출될 수 있다. 구현예에서, NHEJ 비는 0 내지 5% 넘게 증가된다. 구현예에서, NHEJ 비는, 예를 들어 Rad51 억제시, 20 내지 70%(예를 들어, 30% 내지 60% 또는 40 내지 50%)로 증가된다.

[0681] 일부 구현예에서, 엔도뉴클레아제 도메인은 절단 후 표적을 방출한다. 일부 구현예에서, 표적의 방출은, 예를 들어 문헌[Yourik at al. RNA 25(1):35-44 (2019)(그 전체가 본 명세서에 참고로 포함됨)]에 기재되고, 도 2에 나타낸 바와 같이, 효소에 의한 다수의 전환수(turnover)를 평가함으로써 간접적으로 표시된다. 일부 구현예에서, 엔도뉴클레아제 도메인의 k_{exp}는 그러한 방법에 의해 측정 시, 1 x 10^{-3} 내지 1 x 10^{-}5 분-1이다.

[0682] 일부 구현예에서, 엔도뉴클레아제 도메인은 시험관 내에서 약 1 x 10^{8} s^{-1} M^{-1} 초과의 촉매 효율(k_{cat}/K_m)을 갖는다. 구현예에서, 엔도뉴클레아제 도메인은 시험관 내에서 약 1 x 10^{5}, 1 x 10^{6}, 1 x 10^{7}, 또는 1 x 10^{8}, s^{-1} M^{-1} 초과의 촉매 효율을 갖는다. 구현예에서, 촉매 효율은 문헌[Chen et al. (2018) *Science* 360(6387):436-439(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이 결정된다. 일부 구현예에서, 엔도뉴클레아제 도메인은 세포 내에서 약 1 x 10^{8} s^{-1} M^{-1} 초과의 촉매 효율(k_{cat}/K_m)을 갖는다. 구현예에서, 엔도뉴클레아제 도메인은 세포

내에서 약 1 x 10^5, 1 x 10^6, 1 x 10^7, 또는 1 x 10^8, s^{-1} M^{-1} 초과의 촉매 효율을 갖는다.

[0683] *Cas 도메인을 포함하는 유전자 변형 폴리펩티드*

[0684] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드는 Cas 도메인을 포함한다. 일부 구현예에서, Cas 도메인은 gRNA 스페이서에 의해 지정된 표적 부위로 유전자 변형 폴리펩티드를 지시하여, 이에 의해 "시스"로 표적 핵산 서열을 변형시킬 수 있다. 일부 구현예에서, 유전자 변형 폴리펩티드는 Cas 도메인에 융합된다. 유전자 변형 폴리펩티드는 CRISPR/Cas 도메인(CRISPR-연관 단백질로도 지칭됨)을 포함한다. 일부 구현예에서, CRISPR/Cas 도메인은 클러스터링된 조절 간격의 짧은 회문 반복(CRISPR) 시스템에 관여하는 단백질, 예를 들어, Cas 단백질을 포함하고, 선택적으로 가이드 RNA, 예를 들어 단일 가이드 RNA(sgRNA)에 결합한다.

[0685] CRISPR 시스템은 원래 박테리아 및 고세균에서 발견된 적응 방어 시스템이다. CRISPR 시스템은 CRISPR-연관 또는 "Cas" 엔도뉴클레아제(예를 들어, Cas9 또는 Cpf1)로 지칭되는 RNA-유도 뉴클레아제를 사용하여 외래 DNA를 절단한다. 예를 들어, 통상적인 CRISPR-Cas 시스템에서, 엔도뉴클레아제는 단일- 또는 이중-가닥 DNA 서열을 표적화하는 서열-특이적, 비-코딩 "가이드 RNA"에 의해 표적 뉴클레오티드 서열(예를 들어 서열-편집할 게놈 내의 부위)로 지시된다. CRISPR 시스템의 3가지 클래스(I 내지 III)가 확인되었다. 클래스 II CRISPR 시스템은 (다수의 Cas 단백질이 아닌) 단일 Cas 엔도뉴클레아제를 사용한다. 하나의 클래스 II CRISPR 시스템은 Cas9와 같은 II형 Cas 엔도뉴클레아제, CRISPR RNA("crRNA"), 및 트랜스-활성화 crRNA("tracrRNA")를 포함한다. crRNA는 통상적으로 표적 DNA 서열에 상응하는 약 20-뉴클레오티드 RNA 서열인 "스페이서" 서열("프로토스페이서")을 함유한다. 야생형 시스템에서, 그리고 일부 조작된 시스템에서, crRNA는 또한 tracrRNA에 결합하여 RNase III에 의해 절단되는 부분적으로 이중 가닥인 구조를 형성하여 crRNA/tracrRNA 하이브리드 분자를 생성하는 영역을 함유한다. 그 다음 crRNA/tracrRNA 하이브리드는 Cas 엔도뉴클레아제가 표적 DNA 서열을 인식하고 절단하도록 지시한다. 표적 DNA 서열은 일반적으로 주어진 Cas 엔도뉴클레아제에 특이적이고 crRNA의 스페이서와 일치하는 표적 부위에서 절단 활성에 필요한 "프로토스페이서 인접 모티프"(protospacer adjacent motif; "PAM")에 인접한다. 다양한 원핵생물 종으로부터 확인된 CRISPR 엔도뉴클레아제는, 예를 들어 표 7의 예시적인 Cas 효소에 대해 열거된 바와 같이 고유한 PAM 서열 요건을 가지며; PAM 서열의 예로는 5'-NGG(스트렙토코커스 피오게네스(Streptococcus pyogenes)), 5'-NNAGAA(스트렙토코커스 써모필루스(Streptococcus thermophilus) CRISPR1), 5'-NGGNG(스트렙토코커스 써모필루스 CRISPR3), 및 5'-NNNGATT(네이세리아 메닝지디티스(Neisseria meningiditis))가 있다. 일부 엔도뉴클레아제, 예를 들어 Cas9 엔도뉴클레아제는 G-풍부 PAM 부위, 예를 들어 5'-NGG와 회합되며, PAM 부위로부터(5'로부터)의 위치 3인 뉴클레오티드 상류에서 표적 DNA의 평활 말단 절단을 수행한다. 또 다른 클래스 II CRISPR 시스템은, Cas9보다 작은 V 형 엔도뉴클레아제 Cpf1를 포함하며, 예에는 AsCpf1((애시드아미노코커스(Acidaminococcus) 종 유래) 및 LbCpf1((라크노스피라세(Lachnospiraceae) 종 유래)가 포함된다. Cpf1-연관 CRISPR 어레이는 tracrRNA의 필요 없이 성숙한 crRNA로 처리되며; 즉, Cpf1 시스템은 일부 구현예에서 표적 DNA 서열을 절단하기 위해 Cpf1 뉴클레아제 및 crRNA만을 포함한다. Cpf1 엔도뉴클레아제는 통상적으로 T-풍부 PAM 부위, 예를 들어 5'-TTN과 회합된다. Cpf1은 5'-CTA PAM 모티프도 또한 인식할 수 있다. Cpf1은 통상적으로 4- 또는 5-뉴클레오티드 5' 오버행을 갖는 오프셋 또는 엇갈린 이중 가닥 파손을 도입함으로써, 예를 들어 코딩 가닥 상에서 PAM 부위로부터(3'로부터)의 18개 뉴클레오티드 하류, 및 상보적 가닥 상에서 PAM 부위로부터 23개 뉴클레오티드 하류에 위치한 5-뉴클레오티드 오프셋 또는 엇갈린 절단물을 갖는 표적 DNA를 절단하며; 그러한 오프셋 절단으로 인한 5-뉴클레오티드 오버행은 평활 말단 절단된 DNA에서의 삽입에 의한 것보다는, 상동성 재조합에 의한 DNA 삽입에 의해 더 정확한 게놈 편집을 가능하게 한다. 예를 들어 문헌[Zetsche et al. (2015) Cell, 163:759 - 771]을 참조한다.

[0686] 다양한 CRISPR 연관 (Cas) 유전자 또는 단백질이 본 개시내용에 의해 제공되는 기술에 사용될 수 있으며, Cas 단백질의 선택은 본 방법의 특정 조건에 따라 달라질 것이다. Cas 단백질의 특정 예는 Cas1, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9, Cas10, Cpf1, C2C1, 또는 C2C3을 포함하는 클래스 II 시스템을 포함한다. 일부 구현예에서, Cas 단백질, 예를 들어 Cas9 단백질은 다양한 원핵생물 종들 중 임의의 것으로부터 유래될 수 있다. 일부 구현예에서, 특정 Cas 단백질, 예를 들어 특정 Cas9 단백질은 특정 프로토스페이서-인접 모티프(PAM) 서열을 인식하도록 선택된다. 일부 구현예에서, DNA-결합 도메인 또는 엔도뉴클레아제 도메인은 서열 표적화 폴리펩티드, 예컨대 Cas 단백질, 예를 들어 Cas9를 포함한다. 특정 구현예에서, Cas 단백질, 예를 들어 Cas9 단백질은 박테리아 또는 고세균으로부터 수득되거나, 또는 알려진 방법을 이용하여 합성될 수 있다. 특정 구현예에서, Cas 단백질은 그람 양성 박테리아 또는 그람 음성 박테리아에서 유래할 수 있다. 특정 구현예에서, Cas 단백질은 스트렙토코커스(예를 들어, S. 피오게네스, 또는 S. 써모필루스), 프란시셀라(Francisella)(예를 들어, F. 노비시다(novicida)), 스태필로코커스(예를 들어, S. 아우레우스(aureus)), 애시드아미노코커스(예를

들어, 애시드아미노코커스 BV3L6), 나이세리아(예를 들어, N. 메닝기티디스), 크립토코커스, 코리네박테리움, 헤모필루스(Haemophilus), 유박테리움(Eubacterium), 파스퇴렐라(Pasteurella), 프레보텔라(Prevotella), 베일로넬라(Veillonella), 또는 마리노박터(Marinobacter)에서 유래할 수 있다.

[0687] 일부 구현예에서, 유전자 변형 폴리펩티드는 하기 서열번호 4000의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일성을 갖는 서열을 포함할 수 있다. 구현예에서, 하기 서열번호 4000의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일성을 갖는 서열은 유전자 변형 폴리펩티드의 N-말단 끝에 위치한다. 구현예에서, 하기 서열번호 4000의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일성을 갖는 서열은 유전자 변형 폴리펩티드의 N-말단 끝의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 또는 30개 아미노산 내에 위치한다.

[0688] **예시적인 N-말단 NLS-Cas9 도메인**

[0689] MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF
DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHP
IFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALS
LGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVN
TEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY
KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR
EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN
LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK
EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDR
EMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF
MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPE
NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQ
LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKL
IREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY
KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY
SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEI
IEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR
KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGG (서열번호 4000)

[0690] 일부 구현예에서, 유전자 변형 폴리펩티드는 하기 서열번호 4001의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일성을 갖는 서열을 포함할 수 있다. 구현예에서, 하기 서열번호 4001의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일성을 갖는 서열은 유전자 변형 폴리펩티드의 C-말단 끝에 위치한다. 구현예에서, 하기 서열번호 4001의 아미노산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일성을 갖는 서열은 유전자 변형 폴리펩티드의 C-말단 끝의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 또는 30개 아미노산 내에 위치한다.

[0691] **NLS를 포함하는 예시적인 C-말단 서열**

[0692] AGKRTADGSEFEKRTADGSEFESPKKKAKVE (서열번호 4001)

[0693] **예시적인 벤치마킹 서열**

MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF
DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHP
IFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALS
LGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVN
TEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY
KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR
EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN
LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK
EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDR
EMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF
MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPE
NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQ
LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKL
IREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY
KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY
SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEI
IEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR
KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGSGGSSGGSSGSETPGTSESATPESSGG
SSGGSSGGTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATS
TPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK
RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLT
WTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLG
NLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRL
FIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGY
AKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAV
EALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHG
TRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQ
ALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSI
IHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSEFEAGKRT
ADGSEFEKRTADGSEFESPKKKAKVE(서열번호 4002)

[0694]

[0695] 일부 구현예에서, 유전자 변형 폴리펩티드는 표 7 또는 8에 열거된 바와 같은 Cas 도메인, 또는 이의 기능성 단편, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일성을 갖는 서열을 포함할 수 있다.

[0696] 표 7. CRISPR/Cas 단백질, 종, 및 돌연변이

명칭	효소	종	AA의 수	PAM	PAM 인식을 변경시키기 위한 돌연변이	촉매적으로 사멸시키기 위한 돌연변이
FnCas9	Cas9	프란시셀라 노비시다	1629	5'-NGG-3'	Wt	D11A/H969A/N995A
FnCas9 RHA	Cas9	프란시셀라 노비시다	1629	5'-YG-3'	E1369R/E1449H/R1556A	D11A/H969A/N995A
SaCas9	Cas9	스태필로코커스 아우레우스	1053	5'-NNGRRT-3'	Wt	D10A/H557A
SaCas9 KKH	Cas9	스태필로코커스 아우레우스	1053	5'-NNNRRT-3'	E782K/N968K/R1015H	D10A/H557A
SpCas9	Cas9	스트렙토코커스 피오게네스	1368	5'-NGG-3'	Wt	D10A/D839A/H840A/N863A
SpCas9 VQR	Cas9	스트렙토코커스 피오게네스	1368	5'-NGA-3'	D1135V/R1335Q/T1337R	D10A/D839A/H840A/N863A
AsCpf1 RR	Cpf1	애시드아미노코커스 종 BV3L6	1307	5'-TYCV-3'	S542R/K607R	E993A
AsCpf1 RVR	Cpf1	애시드아미노코커스 종 BV3L6	1307	5'-TATV-3'	S542R/K548V/N552R	E993A
FnCpf1	Cpf1	프란시셀라 노비시다	1300	5'-NTTN-3'	Wt	D917A/E1006A/D1255A
NmCas9	Cas9	네이세리아 메닝지티디스	1082	5'-NNNGATT-3'	Wt	D16A/D587A/H588A/N611A

[0697]

[0698] *표 8 CRISPR/Cas 단백질의 아미노산 서열, 종, 및 돌연변이*

변이체	모 숙주(들)	단백질 서열	닉카제 (HNH)	닉카제 (HNH)	닉카제 (RuvC)
Nme2Cas9	네이세리아 메닝지티디스	MAAFKPNPINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPK TGDSLAMARRLARSVRRLTRRRAHRLLRARRLLKREGVLQAADFDENGLIKS LPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELG ALLKGVANNAHALQTGDFRTPAELALNKFEKESGHIRNQRGDYSHTFSRKD LQAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCT FEPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRK SKLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEG LKDKKSPLNLSSELQDEIGTAFSLFKTDEDITGRLKDRVQPEILEALLKHISFDKF VQISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRN PVVLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENR KDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLVRLNE KGYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSR	N611A	H588A	D16A

[0699]

		EWQEFKARVETSRFPRSKKQRILLQKFDEDGFKECNLNDTRYVNRFLCQFVA DHILLTGKGKRRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVACS TVAMQQKITRFVRYKEMNAFDGKTIDKETGKVLHQKTHFPQPWEFFAQEV MIRVFGKPDGKPEFEEADTPEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPNR KMSGAHKDTLRSAKRFVKHNEKISVKRVWLTEIKLADLENMVNYKNGREIEL YEALKARLEAYGGNAKQAFDPKDNPFYKKGGQLVKAVRVEKTQESGVLLNK KNAYTIADNGDMVRVDVFCKVDKKGKNQYFIVPIYAWQVAENILPDIDCKG YRIDDSYTFCFSLHKYDLIAFQKDEKSKVEFAYYINCDSSNGRFYLAWHDKGS KEQQFRISTQNLVLIQKYQVNELGKEIRPCRLKKRPPVR			
PpnCas9	파스퇴렐라 뉴모트로피카	MQNNPLNYILGLDLGIASIGWAVVEIDEESSPIRLIDVGVRTFERAEVAKTGE SLALSRRLARSSRRLIKRRAERLKKAKRLLKAEKILHSIDEKLPINVWQLRVKGL KEKLERQEWAAVLLHLSKHRGYLSQRKNEGKSDNKELGALLSGIASNHQML QSSEYRTPAEIAVKKFQVEEGHIRNQRGSYTHTFSRLDLLAEMELLFQRQAEL GNSYTSTTLLENLTALLMWQKPALAGDAILKMLGKCTFEPSEYKAAKNSYSA ERFVWLTKLNNLRILENGTERALNDNERFALLEQPYEKSKLTYAQVRAMLAL SDNAIFKGVRYLGEDKKTVESKTTLIEMKFYHQIRKTLGSAELKKEWNELKGN SDLLDEIGTAFSLYKTDDDICRYLEGKLPERVLNALLENLNFDKFIQLSLKALHQ ILPLMLQGQRYDEAVSAIYGDHYGKKSTETTRLLPTIPADEIRNPVVLRTLTQA RKVINAVVRLYGSPARIHIETAREVGKSYQDRKKLEKQQEDNRKQRESAVKK FKEMFPHFVGEPKGKDILKMRLYELQQAKCLYSGKSLELHRLLEKGYVEVDH ALPFSRTWDDSFNNKVLVLANENQNKGNLTPYEWLDGKNNSERWQHFVV RVQTSGFSYAKKQRILNHKLDEKGFIERNLNDTRYVARFLCNFIADNMLLVG KGKRNVFASNGQITALLRHRWGLQKVREQNDRHHALDAVVVACSTVAMQ QKITRFVRYNEGNVFSGERIDRETGEIIPLHFPSPWAFFKEENVEIRIFSENPKLE LENRLPDYPQYNHEWVQPLFVSRMPTRKMTGQGHMETVKSAKRLNEGLS VLKVPLTQLKLSDLERMVNRDREIALYESLKARLEQFGNDPAKAFAEPFYKKG GALVKAVRLEQTQKSGVLVRDGNGVADNASMVRVDVFTKGGKYFLVPIYT WQVAKGILPNRAATQGKDENDWDIMDEMATFQFSLCQNDLIKLVTKKKTI FGYFNGLNRATSNIIIKEHDLDKSKGKLGIYLEVGVKLAISLEKYQVDELGKNI RPCRPTKRQHVR			
SauCas9	스태필로코커스 아우레우스	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGA RRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSA ALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKK DGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYE GPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLN NLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTST GKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSE LTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSK DAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYS LEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQ YLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNL VDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKG YKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQ EYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVN NLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPL YKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKL SLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQA EFIASFYNNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPP RIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG	N580A	H557A	D10A
SauCas9-KKH	스태필로코커스 아우레우스	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGA RRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSA ALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKK DGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYE GPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLN NLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTST GKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSE LTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSK	N580A	H557A	D10A

[0700]

		DAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYS LEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQ YLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNL VDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKG YKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQ EYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIV NNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNP LYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVK LSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQ AEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRP PHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG			
Saur\Cas9	스타필로코커스 아우리쿨라리스	MQENQQKQNYILGLDIGITSVGYGLIDSKTREVIDAGVRLFPEADSENNSNR RSKRGARRLKRRRIHRLNRVKDLLADYQMIDLNNVPKSTDPYTIRVKGLREPL TKEEFAIALLHIAKRRGLHNISVSMGDEEQDNELSTKQQLQKNAQQLQDKY VCELQLERLTNINKVRGEKNRFKTEDIFVKEVKQLCETQRQYHNIDDQFIQQY IDLVSTRREYFEGPGNGSPYGWDGDLLKWYEKLMGRCTYFPEELRSVKYAYS ADLFNALNDLNNLVVTRDDNPKLEYYEKYHIIENVFKQKKNPTLKQIAKEIGV QDYDINGYRITKSGKPQFTSFKLYHDLKNIFEQAKYLEDVEMLDEIAKILTIYQ DESIKKALDQLPELLTESEKSQIAQLTGYTGTHRLSLKCIHIVIDELWESPENQ MEIFTRLNLKPKKVEMSEIDSIPTTLVDEFILSPVVKRAFIQSIKVINAVINRFGL PEDIIIELAREKNSKDRRKFINKLQKQNEATRKKIEQLLAKYGNTNAKYMIEKI KLHDMQEGKCLYSLEAIPLEDLLSNPTHYEVDHIIPRSVSFDNSLNNKVLVKQ SENSKKGNRTPYQYLSSNESKISYNQFKQHILNLSKAKDRISKKKRDMLLEER DINKFEVQKEFINRNLVDTRYATRELSNLLKTYFSTHDYAVKVKTINGGFTNH LRKVWDFKKHRNHGYKHHAEDALVIANADFLFKTHKALRRTDKILEQPGLE VNDTTVKVDTEEKYQELFETPKQVKNIKQFRDFKYSHRVDKKPNRQLINDTL YSTREIDGETYVVQTLKDLYAKDNEKVKKLFTERPQKILMYQHDPKTFEKLM TILNQYAEAKNPLAAYYEDKGEYVTKYAKKGNGPAIHKIKYIDKKLGSYLDVS NKYPETQNKLVKLSLKSFRFDIYKCEQGYKMVSIGYLDVLKKDNYYYIPKDKYE AEKQKKKIKESDLFVGSFYYNDLIMYEDELFRVIGVNSDINNLVELNMVDITY KDFCEVNNVTGEKRIKKTIGKRVVLIEKYTTDILGNLYKTPLPKKPQLIFKRGEL	NS88A	HS65A	D15A
Saur\Cas9-KKH	스타필로코커스 아우리쿨라리스	MQENQQKQNYILGLDIGITSVGYGLIDSKTREVIDAGVRLFPEADSENNSNR RSKRGARRLKRRRIHRLNRVKDLLADYQMIDLNNVPKSTDPYTIRVKGLREPL TKEEFAIALLHIAKRRGLHNISVSMGDEEQDNELSTKQQLQKNAQQLQDKY VCELQLERLTNINKVRGEKNRFKTEDFVKEVKQLCETQRQYHNIDDQFIQQY IDLVSTRREYFEGPGNGSPYGWDGDLLKWYEKLMGRCTYFPEELRSVKYAYS ADLFNALNDLNNLVVTRDDNPKLEYYEKYHIIENVFKQKKNPTLKQIAKEIGV QDYDIRGYRITKSGKPQFTSFKLYHDLKNIFEQAKYLEDVEMLDEIAKILTIYQ DESIKKALDQLPELLTESEKSQIAQLTGYTGTHRLSLKCIHIVIDELWESPENQ MEIFTRLNLKPKKVEMSEIDSIPTTLVDEFILSPVVKRAFIQSIKVINAVINRFGL PEDIIIELAREKNSKDRRKFINKLQKQNEATRKKIEQLLAKYGNTNAKYMIEKI KLHDMQEGKCLYSLEAIPLEDLLSNPTHYEVDHIIPRSVSFDNSLNNKVLVKQ SENSKKGNRTPYQYLSSNESKISYNQFKQHILNLSKAKDRISKKKRDMLLEER DINKFEVQKEFINRNLVDTRYATRELSNLLKTYFSTHDYAVKVKTINGGFTNH LRKVWDFKKHRNHGYKHHAEDALVIANADFLFKTHKALRRTDKILEQPGLE VNDTTVKVDTEEKYQELFETPKQVKNIKQFRDFKYSHRVDKKPNRKLINDTL YSTREIDGETYVVQTLKDLYAKDNEKVKKLFTERPQKILMYQHDPKTFEKLM TILNQYAEAKNPLAAYYEDKGEYVTKYAKKGNGPAIHKIKYIDKKLGSYLDVS NKYPETQNKLVKLSLKSFRFDIYKCEQGYKMVSIGYLDVLKKDNYYYIPKDKYE AEKQKKKIKESDLFVGSFYKNDLIMYEDELFRVIGVNSDINNLVELNMVDITY KDFCEVNNVTGEKHIKKTIGKRVVLIEKYTTDILGNLYKTPLPKKPQLIFKRGEL	NS88A	HS65A	D15A
ScaCas9-Sc++	스트랩토코커스 카니스	MEKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTNRKSIKKNLMGALL FDSGETAEATRLKRTARRRYTRRKNRIRYLQEIFANEMAKLDDSFFQRLEESF LVEEDKKNERHPIFGNLADEVAYHRNYPTIYHLRKKLADSPEKADLRLIYLALA HIIKFRGHFLIEGKLNAENSDVAKLFYQLIQTYNQLFEESPLDEIEVDAKGILSA RLSKSKRLEKLIAVFPNEKKNGLFGNIIALALGLTPNFKSNFDLTEDAKLQLSKD TYDDDLDELLGQIGDQYADLFSAAKNLSDAILLSDILRSNSEVTKAPLSASMV KRYDEHHQDLALLKTLVRQQLPEKYAEIFKDDTKNGYAGYVGADKKLRKRS GKLATEEFYKFIKPILEKMDGAEELLAKLNREDLLRKQRTFDNGSIPHQIHLK	N872A	H849A	D10A

[0701]

		ELHAILRRQEEFYPFLKENREKIEKILTFRIPYYVGPLARGNSRFAWLTRKSEEA ITPWNFEEVVDKGASAQSFIERMTNFDEQLPNEKVLPKHSLLYEYFTVYNEL TKVKYVTERMRKPEFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEIIGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRHYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFSNRNFMQLIHDDSLTFKEEIEKAQVSGQGDSLHEQIANLAGSPAIKKGIL QTVKIVDELVKVMGHKPENIVIEMARENQTTTKGLQQSRERIKKIEEGIKELE SQILKENPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFIKDDSIDNKVLTRSVENRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI DKPIREVKVITLKSKLVSDFRKDFQLYKVREINNYHHAHDAYLNAVVGTALIK KYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKRFFYSNIMNFFKTEVKL ANGEIRKRPLIETNGETGEVVWNKEKDFATVRKVLAMPQVNIVKKTEVQTG GFSKESILSKRESAKLIPRKKGWDTRKYGGFGSPTVAYSILVVAKVEKGKAKKL KSVKVLVGITIMEKGSYEKDPIGFLEAKGYKDIKKELIIKLPKYSLFELENGRRR MLASAKELQKANELVLPQHLVRLLYYTQNISATTGSNNLGYIEQHREEFKEIF EKIIDFSEKYILKNKVNSNLKSSFDEQFAVSDSILLSNSFVSLLKYTSFGASGGFT FLDLDVKQGRLRYQTVTEVLDATLIYQSITGLYETRTDLSQLGGD			
SpyCas9	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKE LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASA GELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	N863A	H840A	D10A
SpyCas9-NG	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ	N863A	H840A	D10A

[0702]

		ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES IRPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKE LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASA RFLQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPRAF KYFDTTIDRKVYRSTKEVLDATLIHQSITGLYETRIDLSQLGGD			
SpyCas9-SpRY	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAERTRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES IRPKRNSDKLIARKKDWDPKKYGGFLWPTVAYSVLVVAKVEKGKSKKLKSVK ELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AKQLQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTRLGAPRAF KYFDTTIDPKQYRSTKEVLDATLIHQSITGLYETRIDLSQLGGD	N863A	H840A	D10A
St1Cas9	스트렙토코커스 써모필루스	MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQG RRLARRKKHRRVRLNRLFEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFI ALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSKQLETKTPGQIQLER YQTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEF INRYLEILTGKRKYYHGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEF RAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEKAMGPAK LFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETL DKLAYVLTLNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGW HNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIY NPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKAN KDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYT GKTSIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKGQRTPYQ ALDSMDDAWSFRELKAFVRESKTLSNKKKEYLLTEEDISKFDVRKKFIERNLV DTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYH HHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFK APYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKADE TYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPN KQINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDIT PKDSNNKVVLQSVSPWRADVYFNKTTGKYEILGLKYADLQFEKGTGTYKISQ EKYNDIKKKEGVDSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTMPKQKH YVELKPYDKQKFEGGEALIKVLGNVANSGQCKKGLGKSNISIYKVRTDVLGN QHIIKNEGDKPKLDF	N622A	H599A	D9A

[0703]

BatCas9	브레비바실러스 라테로스포루스	MAYTMGIDVGIASCGWAIVDLERQRIIDIGVRTFEKAENPKNGEALAVPRRE ARSSRRRLRRKKHRIERLKHMFVRNGLAVDIQHLEQTLRSQNEIDVWQLRV DGLDRMLTQKEWLRVLIHLAQRRGFQSNRKTDGSSEDGQVLVNVTENDRL MEEKDYRTVAEMMVKDEKFSDHKRNKNGNYHGVVSRSSLLVEEHTLFETQ RQHHNSLASKDFELEYVNIWSAQRPVATKDQIEKMIGTCTFLPKEKRAPKAS WHFQYFMLLQTINHIRITNVQGTRSLNKEEIEQVVNMALTKSKVSYHDTRKI LDLSEEYQFVGLDYGKEDEKKKVESKETTIKLDDYHKLNKIFNEVELAKGETWE ADDYDTVAYALTFFKDDEDIRDYLQNKYKDSKNRLVKNLANKEYTNELIGKV STLSFRKVGHLSLKALRKIIPFLEQGMTYDKACQAAGFDFQGISKKKRSVVLP VIDQISNPVVNRALTQTRKVINALIKKYGSPETIHIETARELSKTFDERKNITKD YKENRDKNEHAKKHLSELGIINPTGLDIVKYKLWCEQQGRCMYSNQPISFER LKESGYTEVDHIIPYSRSMNDSYNNRVLVMTRENREKGNQTPFEYMGNDT QRWYEFEQRVTTNPQIKKEKRQNLLLKGFTNRRELEMLERNLNDTRYITKYL SHFISTNLEFSPSDKKKKVVNTSGRITSHLRSRWGLEKNRGQNDLHHAMDAI VIAVTSDSFIQQVTNYYKRKERRELNGDDKFFLPWKFFREEVIARLSPNPKEQ IEALPNHFYSEDELADLQPIFVSRMPKRSITGEAHQAQFRRVVGKTKEGKNIT AKKTALVDISYDKNGDFNMYGRETDPATYEAIKERYLEFGGNVKKAFSTDLH KPKKDGTKGPLIKSVRIMENKTLVHPVNKGKGVVYNSSIVRTDVFQRKEKYY LLPVYVTDVTKGKLPNKVIVAKKGYHDWIEVDDSFTFLFSLYPNDLIFIRQNPK KKISLKKRIESHSISDSKEVQEIHAYYKGVDSSTAAIEFIIHDGSYYAKGVGVQN LDCFEKYQVDILGNYFKVKGEKRLELETSDSNHKGKDVNSIKSTSR	N607A	H584A	D8A
cCas9-v16	스태필로코커스 아우레우스	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGA RRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSA ALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKK DGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYE GPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLN NLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTST GKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSE LTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSK DAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYS LEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQ YLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNL VDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKG YKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQ EYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIV NNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNP LYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVK LSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQ AEFIASFYKNDLIKINGELYRVIGVNSDKNNLIEVNMIDITYREYLENMNDKRP PHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG	N580A	H557A	D10A
cCas9-v17	스태필로코커스 아우레우스	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGA RRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSA ALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKK DGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYE GPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLN NLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTST GKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSE LTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSK DAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYS LEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQ YLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNL VDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKG YKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQ EYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIV NNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNP LYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVK LSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQ	N580A	H557A	D10A

[0704]

		AEFIASFYKNDLIKINGELYRVIGVNNSTRNIVELNMIDITYREYLENMNDKRP PHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG			
cCas9-v21	스테필로코커스 아우레우스	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGA RRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSA ALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKK DGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYE GPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLN NLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTST GKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSE LTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSK DAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYS LEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQ YLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNL VDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKG YKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQ EYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIV NNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNP LYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVK LSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQ AEFIASFYKNDLIKINGELYRVIGVNSDDRNIIELNMIDITYREYLENMNDKRP PHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG	N580A	H557A	D10A
cCas9-v42	스테필로코커스 아우레우스	MKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGA RRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSA ALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKK DGEVRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYE GPGEGSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLN NLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTST GKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSE LTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKV DLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSK DAQKMINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYS LEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQ YLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNL VDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKG YKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEIETEQ EYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKDDKGNTLIV NNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNP LYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVK LSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQ AEFIASFYKNDLIKINGELYRVIGVNNNRLNKIELNMIDITYREYLENMNDKRP PHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG	N580A	H557A	D10A
CdCas9	코리네박테리움 디프테리아	MKYHVGIDVGTFSVGLAAIEVDDAGMPIKTLSLVSHIHDSGLDPDEIKSAVT RLASSGIARRTRRLYRRKRRRLQQLDKFIQRQGWPVIELEDYSDPLYPWKVR AELAASYIADEKERGEKLSVALRHIARHRGWRNPYAKVSSLYLPDGPSDAFK AIREEIKRASGQPVPETATVGQMVTLCELGTLKLRGEGGVLSARLQQSDYAR EIQEICRMQEIGQELYRKIIDVVFAAESPKGSASSRVGKDPLQPGKNRAIKAS DAFQRYRIAALIGNLRVRVDGEKRILSVEEKNLVFDHLVNLTPKKEPEWVTIA EILGIDRGQLIGTATMTDDGERAGARPPTHDTNRSIVNSRIAPLVDAWKTA SALEQHAMVKALSNAEVDDFDSPEGAKVQAFFADLDDDVHAKLDSLHLPV GRAAYSEDTLVRLTRRMLSDGVDLYTARLQEFGIEPSWTPPTPRIGEPVGNP AVDRVLKTVSRWLESATKTWGAPERVIIEHVREGFVTEKRAREMDGDMRR RAARNAKLFQEMQEKLNVQGKPSRADLWRYQSVQRQNCQCAYCGSPITF SNSEMDHIVPRAGQGSTNTRENLVAVCHRCNQSKGNTPFAIWAKNTSIEG VSVKEAVERTRHWVTDTGMRSTDFKKFTKAVVERFQRATMDEEIDARSME SVAWMANELRSRVAQHFASHGTTVRVYRGSLTAEARRASGISGKLKFFDGV GKSRLDRRHHAIDAAVIAFTSDYVAETLAVRSNLKQSQAHRQEAPQWREFT GKDAEHRAAWRVWCQKMEKLSALLTEDLRDDRVVVMSNVRLRLGNGSA HKETIGKLSKVKLSSQLSVSDIDKASSEALWCALTREPGFDPKEGLPANPERHI	N597A	H573A	D8A

[0705]

		RVNGTHVYAGDNGLFPVSAGSIALRGGYAELGSSFHHARVYKITSGKKPAF AMLRVYTIDLLPYRNQDLFSVELKPQTMSMRQAEKKLRDALATGNAEYLG WLVVDDELVVDTSKIATDQVKAVEAELGTIRRWRVDGFFSPSKLRLRPLQM SKEGIKKESAPELSKIIDRPGWLPAVNKLFSDGNVTVVRRDSLGRVRLESTAH LPVTWKVQ			
CjeCas9	캄필로박터 제주니	MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLARSA RKRLARRKARLNHLKHLIANEFKLNYEDYQSFDESLAKAYKGSLISPYELRFRA LNELLSKQDFARVILHIAKRRGYDDIKNSDDKEKGAILKAIKQNEEKLANYQS VGEYLYKEYFQKFKENSKEFTNVRNKKESYERCIAQSFLKDELKLIFKKQREFG FSFSKKFEEEVLSVAFYKRALKDFSHLVGNCSFFTDEKRAPKNSPLAFMFVAL TRIINLLNNLKNTEGILYTKDDLNALLNEVLKNGTLTYKQTKKLLGLSDDYEFK GEKGTYFIEFKKYKEFIKALGEHNLSQDDLNEIAKDITLIKDEIKLKKALAKYDLN QNQSDSLSKLEFKDHLNISFKALKLVTPLMLEGKKYDEACNELNLKVAINEDK KDFLPAFNETYYKDEVTNPVVLRAIKEYRKVLNALLKKYGKVHKINIELAREVG KNHSQRAKIEKEQNENYKAKKDAELECEKLGLKINSKNILKLRLFKEQKEFCAY SGEKIKISDLQDEKMLEIDHIYPYSRSFDDSYMNKVLVFTKQNQEKLNQTPFE AFGNDSAKWQKIEVLAKNLPTKKQKRILDKNYKDKEQKNFKDRNLNDTRYI ARLVLNYTKDYLDFLPLSDDENTKLNDTQKGSKVHVEAKSGMLTSALRHTW GFSAKDRNNHLHHAIDAVIIAYANNSIVKAFSDFKKEQESNSAELYAKKISELD YKNKRKFFEPFSGFRQKVLDKIDEIFVSKPERKKPSGALHEETFRKEEEFYQSY GGKEGVLKALELGKIRKVNGKIVKNGDMFRVDIFKHKKTNKFYAVPIYTMDF ALKVLPNKAVARSKKGEIKDWILMDENYEFCFSLYKDSLILIQTKDMQEPEFV YYNAFTSSTVSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKSIGIQNLKVFEK YIVSALGEVTKAEFRQREDFKK			
GeoCas9	제오바실러스 스테아로써모필루스	MRYKIGLDIGITSVGWAVMNLDIPRIEDLGVRIFDRAENPQTGESLALPRRLA RSARRRLRRRKHRLERIRRLVIREGILTKEELDKLFEEKHEIDVWQLRVEALDR KLNNDELARVLLHLAKRRGFKSNRKSERSNKENSTMLKHIEENRAILSSYRTV GEMIVKDPKFALHKRNKGENYTNTIARDDLEREIRLIFSKQREFGNMSCTEEF ENEYITFWASQRPVASKDDIEKKVGFCTFEPKEKRAPKATYTFQSFIAWEHIN KLRLISPSGARGLTDEERRLLYEQAFQKNKITYHDIRTLLHLPDDTYPKGIVYDR GESRKQNENIRFLELDAYHQIRKAVDKVYGKGKSSSFLPIDFDTFGYALTLFKD DADIHSYLRNEYEQNGKRMPNLANKVYDNELIEELLNLSFTKFGHLSLKALRS ILPYMEQGEVYSSACERAGYTFTGPKKKQKTMLLPNIPPIANPVVMRALTQA RKVVNAIIKKYGSPVSIHIELARDLSQTFDERRKTKKEQDENRKKNETAIRQL MEYGLTLNPTGHDIVKFKLWSEQNGRCAYSLQPIEIERLLEPGYVEVDHVIPY SRSLDDSYTNKVLVLTRENREKGNRIPAEYLGVGTERWQQFETFVLTNKQFS KKKRDRLLRLHYDENEETEFKNRNLNDTRYISRFFANFIREHLKFAESDDKQK VYTVNGRVTAHLRSRWEFNKNREESDLHHAVDAVIVACTTPSDIAKVTAFY QRREQNKELAKKTEPHFPQPWPHFADELRARLSKHPKESIKALNLGNYDDQ KLESLQPVFVSRMPKRSVTGAAHQETLRRYVGIDERSGKIQTVVKTKLSEIKL DASGHFPMYGKESDPRTYEAIRQRLLEHNNDPKKAFQEPLYKPKKNGEPGP VIRTVKIIDTKNQVIPLNDGKTVAYNSNIVRVDVFEKDGKYYCVPVYTMDIM KGILPNKAIEPNKPYSEWKEMTEDYTFRFSLYPNDLIRIELPREKTVKTAAGEE INVKDVFVYYKTIDSANGGLELISHDHRFSLRGVGSRTLKRFEKYQVDVLGNI YKVRGEKRVGLASSAHSKPGKTIRPLQSTRD			
iSpyMacCas9	스트렙토코커스 종	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRKLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV			

[0706]

		VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEIQTVGQNGG LFDDNPKSPLEVTPSKLVPLKKELNPKKYGGYQKPTTAYPVLLIDTKQLIPISV MNKKQFEQNPVKFLRDRGYQQVGKNDFIKLPKYTLVDIGDGIKRLWASSKEI HKGNQLVVSKKSQILLYHAHHLDSDLSNDYLQNHNQQFDVLFNEIISFSKKC KLGKEHIQKIENVYSNKKNSASIEELAESFIKLLGFTQLGATSPFNFLGVKLNQ KQYKGKKDYILPCTEGTLIRQSITGLYETRVDLSKIGEDSGGSGGSKRTADGSE FES			
NmeCas9	네이세리아 메닝지티디스	MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPK TGDSLAMARRLARSVRRLTRRRAHRLLRTRRLLKREGVLQAANFDENGLIKS LPNTPWQLRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELG ALLKGVAGNAHALQTGDFRTPAELALNKFEKESGHIRNQRSDYSHTFSRKDL QAELILLFEKQKEFGNPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTF EPAEPKAAKNTYTAERFIWLTKLNNLRILEQGSERPLTDTERATLMDEPYRKS KLTYAQARKLLGLEDTAFFKGLRYGKDNAEASTLMEMKAYHAISRALEKEGL KDKKSPLNLSPELQDEIGTAFSLFKTDEDITGRLKDRIQPEILEALLKHISFDKFV QISLKALRRIVPLMEQGKRYDEACAEIYGDHYGKKNTEEKIYLPPIPADEIRNP VVLRALSQARKVINGVVRRYGSPARIHIETAREVGKSFKDRKEIEKRQEENRK DREKAAAKFREYFPNFVGEPKSKDILKLRLYEQQHGKCLYSGKEINLGRLNEK GYVEIDHALPFSRTWDDSFNNKVLVLGSENQNKGNQTPYEYFNGKDNSRE WQEFKARVETSRFPRSKKQRILLQKFDEDGFKERNLNDTRYVNRFLCQFVA DRMRLTGKGKKRVFASNGQITNLLRGFWGLRKVRAENDRHHALDAVVVA CSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKTHFPQPWEFFAQ EVMIRVFGKPDGKPEFEEADTLEKLRTLIAEKLSSRPEAVHEYVTPLFVSRAP NRKMSGQGHMETVKSAKRLDEGVSVLRVPLTQLKLKDLEKMVNREREPKL YEALKARLEAHKDDPAKAFAEPFYKYDKAGNRTQQVKAVRVEQVQKTGVW VRNHNGIADNATMVRVDVFEKGDKYYLVPIYSWQVAKGILPDRAVVQGKD EEDWQLIDDSFNFKFSLHPNDLVEVITKKARMFGYFASCHRGTGNINIRIHD LDHKIGKNGILEGIGVKTALSFQKYQIDELGKEIRPCRLKKRPPVR	N611A	H588A	D16A
ScaCas9	스트렙토코커스 카니스	MEKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTNRKSIKKNLMGALL FDSGETAEATRLKRTARRRYTRRKNRIRYLQEIFANEMAKLDDSFFQRLEESF LVEEDKKNERHPIFGNLADEVAYHRNYPTIYHLRKKLADSPEKADLRLIYLALA HIIKFRGHFLIEGKLNAENSDVAKLFYQLIQTYNQLFEESPLDEIEVDAKGILSA RLSKSKRLEKLIAVFPNEKKNGLFGNIIALALGLTPNFKSNFDLTEDAKLQLSKD TYDDDLDELLGQIGDQYADLFSAAKNLSDAILLSDILRSNSEVTKAPLSASMV KRYDEHHQDLALLKTLVRQQFPEKYAEIFKDDTKNGYAGYVDGKIKHRKRTT KLATQEEFYKFIKPILEKMDGAEELLAKLNREDLLRKQRTFDNGSIPHQIHLKE LHAILRRQEEFYPFLKENREKIEKILTFRIPYYVGPLARGNSRFAWLTRKSEEAI TPWNFEEVVDKGASAQSFIERMTNFDEQLPNEKVLPKHSLLYEYFTVYNELT KVKYVTERMRKPEFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE RLKTYAHLFDDKVMKQLKRRHYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFSNRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEQIANLAGSPAIKKGIL QTVKIVDELVKVMGHKPENIVIEMARENQTTTKGLQQSRERKKRIEEGIKELG SQILKENPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFIKDDSIDNKVLTRSVENRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSEADKAGFIKRQLVETRQITKHVARILDSRMNTKRDKN DKPIREVKVITLKSKLVSDFRKDFQLYKVRDINNYHHAHDAYLNAVVGTALIK KYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKRFFYSNIMNFFKTEVKL ANGEIRKRPLIETNGETGEVVWNKEKDFATVRKVLAMPQVNIVKKTEVQTG GFSKESILSKRESAKLIPRKKGWDTRKYGGFGSPTVAYSILVVAKVEKGKAKKL KSVKVLVGITIMEKGSYEKDPIGFLEAKGYKDIKKELIFKLPKYSLFELENGRRR MLASATELQKANELVLPQHLVRLLYYTQNISATTGSNNLGYIEQHREEFKEIF	N872A	H849A	D10A

[0707]

		EKIIDFSEKYILKNKVNSNLKSSFDEQFAVSDSILLSNSFVSLLKYTSFGASGGFT FLDLDVKQGRLRYQTVTEVLDATLIYQSITGLYETRTDLSQLGGD			
ScaCas9-HiFi-Sc++	스트렙토코커스 카니스	MEKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTNRKSIKKNLMGALL FDSGETAEATRLKRTARRRYTRRKNRIRYLQEIFANEMAKLDDSFFQRLEESF LVEEDKKNERHPIFGNIVDEVAYHRNYPTIYHLRKKLADSPEKADLRLIYLALA HIIKFRGHFLIEGKLNAENSDVAKLFYQLIQTYNQLFEESPLDEIEVDAKGILSA RLSKSKRLEKLIAVFPNEKKNGLFGNIIALALGLTPNFKSNFDLTEDAKLQLSKD TYDDDLDELLGQIGDQYADLFSAAKNLSDAILLSDILRSNSEVTKAPLSASMV KRYDEHHQDLALLKTLVRQQFPEKYAEIFKDDTKNGYAGYVDGADKKLRKRS GKLATEEEFYKFIKPILEKMDGAEELLAKLNREDLLRKQRTFDNGSIPHQIHLK ELHAILRRQEEFYPFLKENREKIEKILTFRIPYYVGPLARGNSRFAWLTRKSEEA ITPWNFEEVVDKGASAQSFIERMTNFDENLPNKKVLPKHSLLYEYFTVYNEL TKVKYVTERMRKPEFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEIIGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIE ERLKTYAHLFDDKVMKQLKRRHYTGWGRLSRKMINGIRDKQSGKTILDFLKS DGFSNANFMQLIHDDSLTFKEEIEKAQVSGQGDSLHEQIANLAGSPAIKKGIL QTVKIVDELVKVMGHKPENIVIEMARENQTTTKGLQQSRERKKRIEEGIKELG SQILKENPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFIKDDSIDNKVLTRSVENRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKGNSDKLIARKKDWDPKKYGGFNSPTAAYSVLVVAKVEKGKSKKLKSVK ELLGITIMERSSFEKNPIGFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGVLHKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGVPAA FKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	NB72A	H849A	D10A
SpyCas9-3var-NRRH	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MVKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEE FYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGIIPHQIHLGELHAILRRQ GDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRLRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVV DELVKVMGGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQI LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKGNSDKLIARKKDWDPKKYGGFNSPTAAYSVLVVAKVEKGKSKKLKSVK ELLGITIMERSSFEKNPIGFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGVLHKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGVPAA FKYFDTTIDKKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	NB63A	H840A	D10A
SpyCas9-3var-NRTH	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS	NB63A	H840A	D10A

[0708]

		KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MVKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEE FYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGIIPHQIHLGELHAILRRQ GDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRLRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVV DELVKVMGGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQI LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKGNSDKLIARKKDWDPKKYGGFNSPTVAYSVLVVAKVEKGKSKKLKSVK ELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS ASVLHKGNELALPSKYVNFLYLASHYEKLKGSSEDNKQKQLFVEQHKHYLDEI IEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGASAAF KYFDTTIGRKLYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD			
SpyCas9-3var-NRCH	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MVKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEE FYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGIIPHQIHLGELHAILRRQ GDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRLRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVV DELVKVMGGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQI LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKGNSDKLIARKKDWDPKKYGGFNSPTVAYSVLVVAKVEKGKSKKLKSVK ELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGVLQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAA FKYFDTTINRKQYNTTKEVLDATLIRQSITGLYETRIDLSQLGGD	N863A	H840A	D10A
SpyCas9-HF1	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA	N863A	H840A	D10A

[0709]

		HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKE LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASA GELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD			
SpyCas9-QQR1	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKE LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASA RELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADAQLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTFKQKQYRSTKEVLDATLIHQSITGLYETRIDLSQLGGD	N863A	H840A	D10A
SpyCas9-SpG	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK	N863A	H840A	D10A

[0710]

		LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFLWPTVAYSVLVVAKVEKGKSKKLKSVK ELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AKQLQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDE IIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAA FKYFDTTIDRKQYRSTKEVLDATLIHQSITGLYETRIDLSQLGGD			
SpyCas9-VQR	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKE LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASA GELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKQYRSTKEVLDATLIHQSITGLYETRIDLSQLGGD	N863A	H840A	D10A
SpyCas9-VRER	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEF YKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED RFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR NFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQ ILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPK LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEI RKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKES ILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKE LLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASA RELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEII EQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKEYRSTKEVLDATLIHQSITGLYETRIDLSQLGGD	N863A	H840A	D10A

[0711]

SpyCas9-xCas	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDTKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKLYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGIIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEKVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGDQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFIQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGVLQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	N863A	H840A	D10A
SpyCas9-xCas-NG	스트렙토코커스 피오게네스	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDTKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKLYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGIIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEKVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGDQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFIQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASARFLQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPRAFKYFDTTIDRKVYRSTKEVLDATLIHQSITGLYETRIDLSQLGGD	N863A	H840A	D10A
St1Cas9-CNRZ1066	스트렙토코커스 써모필루스	MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQGRRLARRKKHRRVRLNRLFEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFIALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSKQLETKTPGQIQLERYQTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEFINRYLEILTGKRKYYHGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEFRAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEKAMGPAKLFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETL	N622A	H599A	D9A

[0712]

		DKLAYVLTLNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGW HNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIY NPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKAN KDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYT GKTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKGQRTPYQ ALDSMDDAWSFRELKAFVRESKTLSNKKKEYLLTEEDISKFDVRKKFIERNLV DTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYH HHAVDALIIAASSQLNLWKKQKNTLVSYSEEQLLDIETGELISDDEYKESVFKA PYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKKDET YVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNK QMNEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLLGNPIDI TPENSKNKVVLQSLKPWRTDVYFNKATGKYEILGLKYADLQFEKGTGTYKIS QEKYNDHKKEGVDSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTLPKQK HYVELKPYDKQKFEGGEALIKVLGNVANGGQCIKGLAKSNISIYKVRTDVLG NQHIIKNEGDKPKLDF			
St1Cas9-LMG1831	스트렙토코커스 씨모필루스	MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQG RRLARRKKHRRVRLNRLFEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFI ALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSKQLETKTPGQIQLER YQTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEF INRYLEILTGKRKYYHGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEF RAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEKAAMGPAK LFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETL DKLAYVLTLNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGW HNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIY NPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKAN KDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYT GKTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKGQRTPYQ ALDSMDDAWSFRELKAFVRESKTLSNKKKEYLLTEEDISKFDVRKKFIERNLV DTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYH HHAVDALIIAASSQLNLWKKQKNTLVSYSEEQLLDIETGELISDDEYKESVFKA PYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKKDET YVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPNK QMNEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLLGNPIDI TPENSKNKVVLQSLKPWRTDVYFNKNTGKYEILGLKYADLQFEKKTGTYKISQ EKYNGIMKEEGVDSDSEFKFTLYKNDLLLVKDTETKEQQLFRFLSRTMPNVK YYVELKPYSKDKFEKNESLIEILGSADKSGRCIKGLGKSNISIYKVRTDVLGNQH IIKNEGDKPKLDF	N622A	H599A	D9A
St1Cas9-MTH17CL396	스트렙토코커스 씨모필루스	MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQG RRLARRKKHRRVRLNRLFEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFI ALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSKQLETKTPGQIQLER YQTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEF INRYLEILTGKRKYYHGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEF RAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEKAAMGPAK LFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETL DKLAYVLTLNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGW HNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIY NPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKAN KDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYT GKTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKGQRTPYQ ALDSMDDAWSFRELKAFVRESKTLSNKKKEYLLTEEDISKFDVRKKFIERNLV DTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYH HHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFK APYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKADE TYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPN KQINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDIT PKDSNNKVVLQSLKPWRTDVYFNKNTGKYEILGLKYSDMQFEKGTGKYSISK EQYENIKVREGVDENSEFKFTLYKNDLLLLKDSENGEQILLRFTSRNDTSKHYV ELKPYNRQKFEGSEYLIKSLGTVAKGGQCIKGLGKSNISIYKVRTDVLGNQHII KNEGDKPKLDF	N622A	H599A	D9A

[0713]

[0714]

St1Cas9-TH1477	스트렙토코커스 써모필루스	MSDLVLGLDIGIGSVGVGILNKVTGEIIHKNSRIFPAAQAENNLVRRTNRQG RRLARRKKHRRVRLNRLFEESGLITDFTKISINLNPYQLRVKGLTDELSNEELFI ALKNMVKHRGISYLDDASDDGNSSVGDYAQIVKENSKQLETKTPGQIQLER YQTYGQLRGDFTVEKDGKKHRLINVFPTSAYRSEALRILQTQQEFNPQITDEF INRYLEILTGKRKYYHGPGNEKSRTDYGRYRTSGETLDNIFGILIGKCTFYPDEF RAAKASYTAQEFNLLNDLNNLTVPTETKKLSKEQKNQIINYVKNEKAMGPAK LFKYIAKLLSCDVADIKGYRIDKSGKAEIHTFEAYRKMKTLETLDIEQMDRETL DKLAYVLTLNTEREGIQEALEHEFADGSFSQKQVDELVQFRKANSSIFGKGW HNFSVKLMMELIPELYETSEEQMTILTRLGKQKTTSSSNKTKYIDEKLLTEEIY NPVVAKSVRQAIKIVNAAIKEYGDFDNIVIEMARETNEDDEKKAIQKIQKAN KDEKDAAMLKAANQYNGKAELPHSVFHGHKQLATKIRLWHQQGERCLYT GKTISIHDLINNSNQFEVDHILPLSITFDDSLANKVLVYATANQEKGQRTPYQ ALDSMDDAWSFRELKAFVRESKTLSNKKKEYLLTEEDISKFDVRKKFIERNLV DTRYASRVVLNALQEHFRAHKIDTKVSVVRGQFTSQLRRHWGIEKTRDTYH HHAVDALIIAASSQLNLWKKQKNTLVSYSEDQLLDIETGELISDDEYKESVFK APYQHFVDTLKSKEFEDSILFSYQVDSKFNRKISDATIYATRQAKVGKDKADE TYVLGKIKDIYTQDGYDAFMKIYKKDKSKFLMYRHDPQTFEKVIEPILENYPN KQINEKGKEVPCNPFLKYKEEHGYIRKYSKKGNGPEIKSLKYYDSKLGNHIDIT PKDSNNKVVLQSLKPWRTDVYFNKNTGKYEILGLKYSDMQFEKGTGKYSISK EQYENIKVKEGVDENSEFKFTLYKNDLLLLKDSENGEQILLRFTSRNDTSKHYV ELKPYNRQKFEGSEYLIKSLGTVVKGGRCKGLGKSNISIYKVRTDVLGNQHIIK NEGDKPKLDF	N622A	H599A	D9A
sRGN3.1	스태필로코커스 종	MNQKFILGLDIGITSVGYGLIDYETKNIIDAGVRLFPEANVENNEGRRSKRGS RRLKRRRIHRLERVKLLLTEYDLINKEQIPTSNNPYQIRVKGLSEILSKDELAIAL LHLAKRRGIHNVDVAADKEETASDSLSTKDQINKNAKFLESRYVCELQKERLE NEGHVRGVENRFLTKDIVREAKKIIQTQMQYYPEIDETFKEKYISLVETRREYF EGPGQGSPFGWNGDLKKWYEMLMGHCTYFPQELRSVKYAYSADLFNALN DLNNLIIQRDNSEKLEYHEKYHIIENVFKQKKKPTLKQIAKEIGVNPEDIKGYRI TKSGTPEFTSFKLFHDLKKVVKDHAILDDIDLLNQIAEILTIYQDKDSIVAELGQ LEYLMSEADKQSISELTGYTGTHSLSLKCMNMIIDELWHSSMNQMEVFTYL NMRPKKYELKGYQRIPTDMIDDAILSPVVKRTFIQSINVINKVIEKYGIPEDIIIE LARENNSDDRKKFINNLQKKNEATRKRINEIIGQTGNQNAKRIVEKIRLHDQ QEGKCLYSLESIPLEDLLNNPNHYEVDHIIPRSVSFDNSYHNKVLVKQSENSK KSNLTPYQYFNSGKSKLSYNQFKQHILNLSKSQDRISKKKKEYLLEERDINKFE VQKEFINRNLVDTRYATRELTNYLKAYFSANNMNVKVKTINGSFTDYLRKV WKFKKERNHGYKHHAEDALIIANADFLFKENKKLKAVNSVLEKPEIETKQLDI QVDSEDNYSEMFIIPKQVQDIKDFRNFKYSHRVDKKPNRQLINDTLYSTRKK DNSTYIVQTIKDIYAKDNTTLKKQFDKSPEKFLMYQHDPRTFEKLEVIMKQYA NEKNPLAKYHEETGEYLTKYSKKNNGPIVKKIKLIGNKVGNHLDVTNKYEN KKLVKLSIKNYRFDVYLTEKGYKFVTIAYLNVFKKDNYYYIPKDKYQELKEKKKI KDTDQFIASFYKNDLIKLNGDLYKIIGVNSDDRNIIELDYYDIKYKDYCEINNIK GEPRIKKTIGKKTESIEKFTTDVLGNLYLHSTEKAPQLIFKRGL	N585A	H562A	D10A
sRGN3.3	스태필로코커스 종	MNQKFILGLDIGITSVGYGLIDYETKNIIDAGVRLFPEANVENNEGRRSKRGS RRLKRRRIHRLERVKLLLTEYDLINKEQIPTSNNPYQIRVKGLSEILSKDELAIAL LHLAKRRGIHNVDVAADKEETASDSLSTKDQINKNAKFLESRYVCELQKERLE NEGHVRGVENRFLTKDIVREAKKIIQTQMQYYPEIDETFKEKYISLVETRREYF EGPGQGSPFGWNGDLKKWYEMLMGHCTYFPQELRSVKYAYSADLFNALN DLNNLIIQRDNSEKLEYHEKYHIIENVFKQKKKPTLKQIAKEIGVNPEDIKGYRI TKSGTPEFTSFKLFHDLKKVVKDHAILDDIDLLNQIAEILTIYQDKDSIVAELGQ LEYLMSEADKQSISELTGYTGTHSLSLKCMNMIIDELWHSSMNQMEVFTYL NMRPKKYELKGYQRIPTDMIDDAILSPVVKRTFIQSINVINKVIEKYGIPEDIIIE LARENNSDDRKKFINNLQKKNEATRKRINEIIGQTGNQNAKRIVEKIRLHDQ QEGKCLYSLESIPLEDLLNNPNHYEVDHIIPRSVSFDNSYHNKVLVKQSENSK KSNLTPYQYFNSGKSKLSYNQFKQHILNLSKSQDRISKKKKEYLLEERDINKFE VQKEFINRNLVDTRYATRELTSYLKAYFSANNMDVKVKTINGSFTNHLRKV WRFDKYRNHGYKHHAEDALIIANADFLFKENKKLQNTNKILEKPTIENNTKK VTVEKEEDYNNVFETPKLVEDIKQYRDYKFSHRVDKKPNRQLINDTLYSTRM KDEHDYIVQTITDIYGKDNTNLKKQFNKNPEKFLMYQNDPKTFEKLSIIMKQ YSDEKNPLAKYYEETGEYLTKYSKKNNGPIVKKIKLLGNKVGNHLDVTNKYEN	N585A	H562A	D10A

[0715]

		STKKLVKLSIKNYRFDVYLTEKGYKFVTIAYLNVFKKDNYYYIPKDKYQELKEKK KIKDTDQFIASFYKNDLIKLNGDLYKIIGVNSDDRNIIELDYYDIKYKDYCEINNI KGEPRIKKTIGKKTESIEKFTTDVLGNLYLHSTEKAPQLIFKRGL			

[0716] 일부 구현예에서, Cas 단백질은 Cas 단백질이 결합 및/또는 기능하기 위한 표적 DNA 서열에 존재하거나 이에 인접하기 위해 프로토스페이서 인접 모티프(PAM)를 필요로 한다. 일부 구현예에서, PAM은 5'에서 3'로 NGG, YG, NNGRRT, NNNRRT, NGA, TYCV, TATV, NTTN, 또는 NNNGATT이거나, 이를 포함하며, 여기서 N은 임의의 뉴클레오티드를 나타내고, Y는 C 또는 T를 나타내고, R은 A 또는 G를 나타내고, V는 A 또는 C 또는 G를 나타낸다. 일부 구현예에서, Cas 단백질은 표 7 또는 8에 열거된 단백질이다. 일부 구현예에서, Cas 단백질은 그의 PAM을 변경하는 하나 이상의 돌연변이를 포함한다. 일부 구현예에서, Cas 단백질은 E1369R, E1449H, 및 R1556A 돌연변이 또는 상기 위치에 상응하는 아미노산에 대한 유사한 치환을 포함한다. 일부 구현예에서, Cas 단백질은 E782K, N968K, 및 R1015H 돌연변이 또는 상기 위치에 상응하는 아미노산에 대한 유사한 치환을 포함한다. 일부 구현예

에서, Cas 단백질은 D1135V, R1335Q, 및 T1337R 돌연변이 또는 상기 위치에 상응하는 아미노산에 대한 유사한 치환을 포함한다. 일부 구현예에서, Cas 단백질은 S542R 및 K607R 돌연변이 또는 상기 위치에 상응하는 아미노산에 대한 유사한 치환을 포함한다. 일부 구현예에서, Cas 단백질은 S542R, K548V, 및 N552R 돌연변이 또는 상기 위치에 상응하는 아미노산에 대한 유사한 치환을 포함한다. 변경된 PAM 서열을 인식하기 위한 Cas 효소의 조작에서의 예시적인 발전은, 그 전체가 본 명세서에 참고로 포함된 문헌[Collias et al Nature Communications 12:555 (2021)]에 검토되어 있다.

[0717] 일부 구현예에서, Cas 단백질은 촉매적으로 활성이며, 표적 DNA 부위의 하나 또는 2개의 가닥을 절단한다. 일부 구현예에서, 표적 DNA 부위의 절단 후 변경, 예를 들어 세포 복구 기구에 의한, 예를 들어 삽입 또는 결실이 형성된다.

[0718] 일부 구현예에서, Cas 단백질은 뉴클레아제, 예를 들어, 뉴클레아제-결핍 Cas9를 탈활성화 또는 부분적으로 탈활성화하도록 변형된다. 야생형 Cas9는 gRNA에 의해 표적화된 특이적인 DNA 서열에서 이중 가닥 파손(DSB)을 생성하는 한편, 변형된 기능기들을 갖는 다수의 CRISPR 엔도뉴클레아제가 이용가능하며, 예를 들어: 부분적으로 탈활성화된 Cas9의 "닉카제" 형태는 단일 가닥 파손만을 생성하고; 촉매적으로 비활성인 Cas9("dCas9")는 표적 DNA를 절단하지 않는다. 일부 구현예에서, DNA 서열에 결합하는 dCas9는 입체 장애에 의해 그 부위에서 전사를 간섭할 수 있다. 일부 구현예에서, 앵커 서열에 결합하는 dCas9는 기하학적 복합체(예를 들어 ASMC) 형성 및/또는 유지를 간섭(예를 들어, 감소 또는 방지)할 수 있다. 일부 구현예에서, DNA-결합 도메인은 촉매적으로 비활성인 Cas9, 예를 들어 dCas9를 포함한다. 많은 촉매적으로 비활성인 Cas9 단백질이 당업계에 알려져 있다. 일부 구현예에서, dCas9는 Cas 단백질의 각각의 엔도뉴클레아제 도메인에서 돌연변이, 예를 들어 D10A 및 H840A 또는 N863A 돌연변이를 포함한다. 일부 구현예에서, 촉매적으로 비활성이거나 부분적으로 비활성인 CRISPR/Cas 도메인은 하나 이상의 돌연변이, 예를 들어 표 7에 열거된 하나 이상의 돌연변이를 포함하는 Cas 단백질을 포함한다. 일부 구현예에서, 표 7의 제공된 열에 기재된 Cas 단백질은 표 7의 동일한 열에 열거된 1, 2, 3, 또는 모든 돌연변이를 포함한다. 일부 구현예에서, 예를 들어 표 7에 기재되지 않은, Cas 단백질은 표 7의 열에 열거된 1, 2, 3, 또는 모든 돌연변이, 또는 Cas 단백질에서 상응하는 부위에서 상응하는 돌연변이를 포함한다.

[0719] 일부 구현예에서, 향상된 활성을 갖는 Cas9 유도체가 유전자 변형 폴리펩티드에 사용될 수 있다. 일부 구현예에서, Cas9 유도체는 HNH 엔도뉴클레아제 도메인의 활성을 개선시키는 돌연변이, 예를 들어 SpyCas9 R221K, N394K, 또는 R-루프 형성을 개선시키는 돌연변이, 예를 들어 SpyCas9 L1245V를 포함할 수 있거나, 그러한 돌연변이의 조합, 예를 들어 SpyCas9 R221K/N394K, SpyCas9 N394K/L1245V, SpyCas9 R221K/L1245V, 또는 SpyCas9 R221K/N394K/L1245V를 포함할 수 있다(예를 들어, 문헌[Spencer and Zhang Sci Rep 7:16836 (2017)] 참조, Cas 유도체 및 돌연변이를 포함함, 이는 본 명세서에 참고로 포함됨). 일부 구현예에서, Cas9 유도체는 본 명세서에 기재된 돌연변이 중 하나 이상의 유형, 예를 들어 PAM-변형 돌연변이, 단백질 안정화 돌연변이, 활성 향상 돌연변이, 및/또는 모 효소에 비해 하나 또는 2개 엔도뉴클레아제 도메인을 부분적으로 또는 완전히 비활성화시키는 돌연변이(예를 들어, 표적 DNA의 하나 또는 두 가닥 모두에 대한 엔도뉴클레아제 활성을 없애는 하나 이상의 돌연변이, 예를 들어 닉카제 또는 촉매적 사멸 효소)을 포함할 수 있다. 일부 구현예에서, 본 명세서에 기재된 시스템에 사용된 Cas9 효소는 촉매 효율을 개선시키는 돌연변이(예를 들어, SpyCas9 R221K, N394K, 및/또는 L1245V)에 추가적으로, 효소에 대한 닉카제 활성을 부여하는 돌연변이(예를 들어, SpyCas9 N863A 또는 H840A)를 포함할 수 있다. 일부 구현예에서, 본 명세서에 기재된 시스템에 사용된 Cas9 효소는 촉매 효율을 개선시키는 R221K 및 N394K 돌연변이에 추가적으로, 닉카제 활성을 부여하는 N863A 돌연변이를 추가로 포함하는 SpyCas9 효소 또는 유도체이다.

[0720] 일부 구현예에서, 촉매적으로 비활성인, 예를 들어 dCas9 또는 부분적으로 탈활성화된 Cas9 단백질은 D11 돌연변이(예를 들어, D11A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사한 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9 또는 부분적으로 탈활성화된 Cas9 단백질은 H969 돌연변이(예를 들어, H969A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9 또는 부분적으로 탈활성화된 Cas9 단백질은 N995 돌연변이(예를 들어, N995A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9는 위치 D11, H969, 및 N995 중 1, 2 또는 3개에서의 돌연변이(예를 들어, D11A, H969A, 및 N995A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다.

[0721] 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9 또는 부분적으로 탈활성화된 Cas9 단백질은 D10 돌연변이(예를 들어, D10A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을

포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9 또는 부분적으로 탈활성화된 Cas9 단백질은 H557 돌연변이(예를 들어, H557A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9는 D10 돌연변이(예를 들어, D10A 돌연변이) 및 H557 돌연변이(예를 들어, H557A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다.

[0722] 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 D839 돌연변이(예를 들어, D839A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 H840 돌연변이(예를 들어, H840A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 N863 돌연변이(예를 들어, N863A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9는 D10 돌연변이(예를 들어, D10A), D839 돌연변이(예를 들어, D839A), H840 돌연변이(예를 들어, H840A), 및 N863 돌연변이(예를 들어, N863A) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다.

[0723] 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 E993 돌연변이(예를 들어, E993A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다.

[0724] 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 D917 돌연변이(예를 들어, D917A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 E1006 돌연변이(예를 들어, E1006A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 D1255 돌연변이(예를 들어, D1255A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9는 D917 돌연변이(예를 들어, D917A), E1006 돌연변이(예를 들어, E1006A), 및 D1255 돌연변이(예를 들어, D1255A) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다.

[0725] 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 D16 돌연변이(예를 들어, D16A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 D587 돌연변이(예를 들어, D587A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 부분적으로 탈활성화된 Cas 도메인은 닉카제 활성을 갖는다. 일부 구현예에서, 부분적으로 탈활성화된 Cas9 도메인은 Cas9 닉카제 도메인이다. 일부 구현예에서, 촉매적으로 비활성인 Cas 도메인 또는 사멸 Cas 도메인은 검출가능한 이중 가닥 파손 형성을 생성하지 않는다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 H588 돌연변이(예를 들어, H588A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9, 또는 부분적으로 탈활성화된 Cas9 단백질은 N611 돌연변이(예를 들어, N611A 돌연변이) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다. 일부 구현예에서, 촉매적으로 비활성인 Cas9 단백질, 예를 들어 dCas9는 D16 돌연변이(예를 들어, D16A), D587 돌연변이(예를 들어, D587A), H588 돌연변이(예를 들어, H588A), 및 N611 돌연변이(예를 들어, N611A) 또는 상기 위치에 상응하는 아미노산에 대한 유사 치환을 포함한다.

[0726] 일부 구현예에서, DNA-결합 도메인 또는 엔도뉴클레아제 도메인은 gRNA(예를 들어, 주형 핵산, 예를 들어 gRNA를 포함하는 주형 RNA)를 포함하거나 그에 연결된(예를 들어, 공유결합적으로) Cas 분자를 포함할 수 있다.

[0727] 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 스트렙토코커스 피오게네스 Cas9(SpCas9) 또는 이의 기능성 단편 또는 변이체를 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 변형된 SpCas9를 포함한다. 구현예에서, 변형된 SpCas9는 프로토스페이서-인접 모티프(PAM) 특이성을 변경시키는 변형을 포함한다. 구현예에서, PAM은 핵산 서열 5'-NGT-3'에 대한 특이성을 갖는다. 구현예에서, 변형된 SpCas9는 하나 이상의 아미노산 치환을, 예를 들어 위치 L1111, D1135, G1218, E1219, A1322, 또는 R1335 중 하나 이상에서, 예를 들어 L1111R, D1135V, G1218R, E1219F, A1322R, R1335V로부터 선택되는 위치에서 포함한

다. 구현예에서, 변형된 SpCas9는 아미노산 치환 T1337R 및 예를 들어 L1111, D1135L, S1136R, G1218S, E1219V, D1332A, D1332S, D1332T, D1332V, D1332L, D1332K, D1332R, R1335Q, T1337, T1337L, T1337Q, T1337I, T1337V, T1337F, T1337S, T1337N, T1337K, T1337H, T1337Q, 및 T1337M으로부터 선택되는 하나 이상의 추가의 아미노산 치환, 또는 이에 상응하는 아미노산 치환을 포함한다. 구현예에서, 변형된 SpCas9는 (i) D1135L, S1136R, G1218S, E1219V, A1322R, R1335Q, 및 T1337로부터 선택되는 하나 이상의 아미노산 치환; 및 (ii) L1111R, G1218R, E1219F, D1332A, D1332S, D1332T, D1332V, D1332L, D1332K, D1332R, T1337L, T1337I, T1337V, T1337F, T1337S, T1337N, T1337K, T1337R, T1337H, T1337Q, 및 T1337M으로부터 선택되는 하나 이상의 치환 또는 이에 상응하는 아미노산 치환을 포함한다.

[0728] 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 Cas 도메인, 예를 들어 Cas9 도메인을 포함한다. 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 뉴클레아제 활성 Cas 도메인, Cas 닉카제(nCas) 도메인, 또는 뉴클레아제-비활성 Cas(dCas) 도메인을 포함한다. 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 뉴클레아제-활성 Cas9 도메인, Cas9 닉카제(nCas9) 도메인, 또는 뉴클레아제-비활성 Cas9(dCas9) 도메인을 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 Cas9(예를 들어, dCas9 및 nCas9), Cas12a/Cpf1, Cas12b/C2c1, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, 또는 Cas12i의 Cas9 도메인을 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 Cas9(예를 들어, dCas9 및 nCas9), Cas12a/Cpf1, Cas12b/C2c1, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, 또는 Cas12i를 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 S. 피오게네스 또는 S. 써모필루스 Cas9, 또는 이의 기능성 단편을 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은, 예를 들어 본 명세서에 참고로 포함된 문헌[Chylinski, Rhun, and Charpentier (2013) RNA Biology 10:5, 726-737]에 기재된 바와 같은, Cas9 서열을 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 HNH 뉴클레아제 서브도메인 및/또는 Cas, 예를 들어 본 명세서에 기재된 바와 같은 Cas9의 RuvC1 서브도메인, 또는 이의 변이체를 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 Cas12a/Cpf1, Cas12b/C2c1, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, 또는 Cas12i를 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 Cas 폴리펩티드(예를 들어, 효소), 또는 이의 기능성 단편을 포함한다. 구현예에서, Cas 폴리펩티드(예를 들어, 효소)는 Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5d, Cas5t, Cas5h, Cas5a, Cas6, Cas7, Cas8, Cas8a, Cas8b, Cas8c, Cas9(예를 들어, Csn1 또는 Csx12), Cas10, Cas10d, Cas12a/Cpf1, Cas12b/C2c1, Cas12c/C2c3, Cas12d/CasY, Cas12e/CasX, Cas12g, Cas12h, Cas12i, Csy1 , Csy2, Csy3, Csy4, Cse1, Cse2, Cse3, Cse4, Cse5e, Csc1, Csc2, Csa5, Csn1, Csn2, Csm1, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx1S, Csx11, Csf1, Csf2, CsO, Csf4, Csd1, Csd2, Cst1, Cst2, Csh1, Csh2, Csa1, Csa2, Csa3, Csa4, Csa5, II형 Cas 이펙터 단백질, V형 Cas 이펙터 단백질, VI형 Cas 이펙터 단백질, CARF, DinG, Cpf1, Cas12b/C2c1, Cas12c/C2c3, Cas12b/C2c1, Cas12c/C2c3, SpCas9(K855A), eSpCas9(1.1), SpCas9-HF1, 초정확한 Cas9 변이체(HypaCas9), 이의 동족체, 변형된 또는 이의 조작된 형태들, 및/또는 이의 기능성 단편들로부터 선택된다. 구현예에서, Cas9는 예를 들어 H840A, D10A, P475A, W476A, N477A, D1125A, W1126A, 및 D1127A로부터 선택되는 하나 이상의 치환을 포함한다. 구현예에서, Cas9는 D10, G12, G17, E762, H840, N854, N863, H982, H983, A984, D986, 및/또는 A987로부터 선택되는 위치에서 하나 이상의 돌연변이, 예를 들어 D10A, G12A, G17A, E762A, H840A, N854A, N863A, H982A, H983A, A984A, 및/또는 D986A로부터 선택되는 하나 이상의 치환을 포함한다. 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 코리네박테리움 울세란스(Corynebacterium ulcerans), 코리네박테리움 디프테리아(Corynebacterium diphtheria), 스피로플라스마 시르피디콜라(Spiroplasma syrphidicola), 프레보텔라 인터미디아(Prevotella intermedia), 스피로플라스마 타이와넨스(Spiroplasma taiwanense), 스트렙토코커스 이니에(Streptococcus iniae), 벨리엘라 발티카(Belliella baltica), 사이크로플렉서스 토르퀴스(Psychroflexus torquis), 스트렙토코커스 써모필루스, 리수테리아 이노쿠아(Listeria innocua), 캄필로박터 제주니(Campylobacter jejuni), 나이세리아 메닝기티디스, 스트렙토코커스 피오게네스, 또는 스태필로코커스 아우레우스(Staphylococcus aureus), 또는 이들의 단편 또는 변이체로부터의 Cas(예를 들어, Cas9) 서열을 포함한다.

[0729] 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은, 예를 들어 D917A, E1006A, D1255A, D917A/E1006A, D917A/D1255A, E1006A/D1255A, 및 D917A/E1006A/D1255A로부터 선택된, 예를 들어 D917, E1006A, D1255 위치 또는 이들의 임의의 조합에서, 예를 들어 하나 이상의 치환을 포함하는 Cpf1 도메인을 포함한다.

[0730] 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA 결합 도메인은 spCas9, spCas9-VRQR(서열번호 19), spCas9-VRER(서열번호 20), xCas9(sp), saCas9, saCas9-KKH, spCas9-MQKSER(서열번호 21), spCas9-LRKIQK(서열번호

22), 또는 spCas9- LRVSQL(서열번호 23)를 포함한다.

[0731] 일부 구현예에서, 유전자 변형 폴리펩티드는 Cas9 닉카제, 예를 들어 Cas9 H840A를 포함하는 엔도뉴클레아제 도메인을 갖는다. 구현예에서, Cas9 H840A는 다음 아미노산 서열을 갖는다:

[0732] *Cas9 닉카제(H840A):*

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR
TARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEK
YPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLF
EENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDA
KLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK
LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN
SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNE
LTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRF
NASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG
DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM
KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSF
LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYK
VREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT
GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI
TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY
VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH
RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL
GGD

[0733]

[0734] 일부 구현예에서, 유전자 변형 폴리펩티드는 D10A 및/또는 H840A 돌연변이를 포함하는 dCas9 서열, 예를 들어 다음 서열을 포함한다:

SMDKKYSIGLAIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL
KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYH
EKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQ
LFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED
AKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYD
EHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV
KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARG
NSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYN
ELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRF
NASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG
DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM
KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSF
LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYK
VREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF
FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQT
GGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI
TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY
VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH
RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQL
GGD (서열번호 7)

[0735]

[0736] *TAL 이펙터 및 징크 핑거 뉴클레아제*

[0737] 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA-결합 도메인은 TAL 이펙터 분자를 포함한다. TAL 이펙터 분자, 예를 들어 DNA 서열에 특이적으로 결합하는 TAL 이펙터 분자는 복수의 TAL 이펙터 도메인들 또는 이의 단편들을 통상적으로 포함하며, 선택적으로 자연에 존재하는 TAL 이펙터들의 하나 이상의 추가의 부분들을 포함한다(예를 들어, TAL 이펙터 도메인들의 N- 및/또는 C-말단). 많은 TAL 이펙터가 당업자에게 알려져 있으며, 예를 들어 Thermo Fisher Scientific으로부터, 상업적으로 구매가능하다.

[0738] 자연에 존재하는 TALE는 숙주 식물에서의 유전자 발현을 조절하고, 박테리아 콜로니화 및 생존을 용이하게 하는 식물 병원체 잔토모나스(*Xanthomonas*)를 포함하는 박테리아 병원체의 다양한 종에 의해 분비된 천연 이펙터 단백질이다. TAL 이펙터의 특이적 결합은 통상적으로 33 또는 234개의 아미노산의 나란히 정렬된 거의 동일한 반복부들의 중심 반복 도메인을 기준으로 한다(반복부-가변 이(di)-잔기, RVD 도메인).

[0739] TAL 이펙터 과의 원들은 그들의 반복부들의 수와 순서에서 주로 상이하다. 반복부들의 수는 통상적으로 1.5 내지 33.5개의 반복부 범위이며, C-말단 반복부는 일반적으로 길이가 더 짧고(예를 들어, 약 20개 아미노산), 일반적으로 "절반-반복부"로서 지칭된다. TAL 이펙터의 각각의 반복부는 일반적으로, 상이한 염기쌍 특이성을 나타내는 상이한 반복부 유형과 하나의 반복-대-하나의 염기쌍 상관관계를 특징으로 한다(하나의 반복부는 표적 유전자 서열 상에서 일 염기쌍을 인식함). 일반적으로, 반복부의 수가 더 적을수록, 단백질-DNA 상호작용은 더 약해진다. 다수의 6.5개 반복부는 리포터 유전자의 전사를 활성화하기에 충분한 것으로 보인다(Scholze et al., 2010).

[0740] 반복부 대 반복부 변이는 아미노산 위치 12 및 13에서 주로 발생하며, 따라서 이는 "초가변"으로 지칭되고, 표적 DNA 프로모터 서열과의 상호작용의 특이성에 책임이 있고, 이는 예시적인 반복부 가변 2잔기(RVD) 및 이들의 핵산 염기 표적에 대한 상동성 열거한 표 9에 나타낸 바와 같다.

[0741] 표 9 - RVD 및 핵산 염기 특이성

[0742]

표적	가능한 RVD 아미노산 조합												
A	NI	NN	CI	HI	KI								
G	NN	GN	SN	VN	LN	DN	QN	EN	HN	RH	NK	AN	FN
C	HD	RD	KD	ND	AD								
T	NG	HG	VG	IG	EG	MG	YG	AA	EP	VA	QG	KG	RG

[0743] 이에 따라, TAL 이펙터의 반복부를 표적 특이적 DAN 서열로 변형하는 것이 가능하다. 추가의 연구는, RVD NK가 G를 표적화할 수 있음을 보여준다. TAL 이펙터의 표적 부위는 제1 반복부에 의해 표적화되는 5' 염기에 측접하는 T를 포함하는 경향도 있지만, 이러한 인식의 정확한 기작은 알려지지 않았다. 113개가 넘는 TAL 이펙터 서열이 오늘날까지 알려졌다. 잔토모나스로부터의 TAL 이펙터의 비제한적 예는 Hax2, Hax3, Hax4, AvrXa7, AvrXa10 및 AvrBs3를 포함한다.

[0744] 이에 따라, 본 명세서에 기재된 TAL 이펙터 분자의 TAL 이펙터 도메인은 임의의 박테리아 종(예를 들어, 잔토모나스 종, 예컨대 아프리카 균주인, 잔토모나스 오리제(oryzae) pv. 오리제(Yu et al. 2011), 잔토모나스 캄페스트리스 pv. 라파니(campestris pv. raphani) 균주 756C 및 잔토모나스오리제 pv. 오리지콜라(oryzicola) 균주 BLS256 (Bogdanove et al. 2011)으로부터의 TAL 이펙터로부터 유도될 수 있다. 일부 구현예에서, TAL 이펙터 도메인은 RVD 도메인뿐만 아니라 자연에 존재하는 TAL 이펙터로부터의 측접 서열(들)(RVD 도메인의 N-말단 및/또는 C-말단 측부 상에서의 서열)도 포함한다. 이는 자연에 존재하는 TAL 이펙터의 RVD보다 더 많거나 더 적은 반복부들을 포함할 수 있다. TAL 이펙터 분자는 상기 코드를 기준으로 한 소정의 DNA 서열 및 본 기술 분야에 알려진 다른 것들을 표적화하도록 설계될 수 있다. TAL 이펙터 도메인의 수(예를 들어, 반복부(단량체 또는 모듈)) 및 그의 특이적 서열은 원하는 DNA 표적 서열을 기준으로 선택될 수 있다. 예를 들어, TAL 이펙터 도메인, 예를 들어 반복부들은 특이적 표적 서열에 맞도록 제거 또는 부가될 수 있다. 구현예에서, 본 발명의 TAL 이펙터 분자는 6.5 내지 33.5 TAL 이펙터 도메인, 예를 들어 반복부를 포함한다. 구현예에서, 본 발명의 TAL 이펙터 분자는 8 내지 33.5 TAL 이펙터 도메인, 예를 들어 반복부, 예를 들어 10 내지 25 TAL 이펙터 도메인, 예를 들어 반복부, 예를 들어 10 내지 14 TAL 이펙터 도메인, 예를 들어 반복부를 포함한다.

[0745] 일부 구현예에서, TAL 이펙터 분자는 DNA 표적 서열에 대한 완전한 매치에 상응하는 TAL 이펙터 도메인을 포함한다. 일부 구현예에서, DNA 표적 서열 상에서 반복부와 표적 염기쌍 사이의 미스매치는 TAL 이펙터 분자를 포

함하는 폴리펩티드의 기능을 가능하게 하는 한 허용된다. 일반적으로, TALE 결합은 미스매치의 수와 역의 상관관계를 갖는다. 일부 구현예에서, 본 발명의 폴리펩티드의 TAL 이펙터 분자는 표적 DNA 서열과, 7개의 미스매치, 6개의 미스매치, 5개의 미스매치, 4개의 미스매치, 3개의 미스매치, 2개의 미스매치, 또는 1개의 미스매치 이하를 포함하고, 선택적으로 미스매치가 없다. 이론에 구속되고자 함이 없이, 일반적으로 TAL 이펙터 분자 내 TAL 이펙터 도메인의 수가 더 적을수록, 더 적은 수의 미스매치가 용인될 것이고, TAL 이펙터 분자를 포함하는 폴리펩티드의 기능을 여전히 허용할 것이다. 결합 친화도는 매칭되는 반복부-DNA 조합의 합에 따라 달라지는 것으로 생각된다. 예를 들어, 25 이상의 TAL 이펙터 도메인을 갖는 TAL 이펙터 분자는 7 이하의 미스매치를 용인할 수 있다.

[0746] TAL 이펙터 도메인에 더하여, 본 발명의 TAL 이펙터 분자는 자연에 존재하는 TAL 이펙터로부터 유도된 추가의 서열들을 포함할 수 있다. TAL 이펙터 분자의 TAL 이펙터 도메인의 각각의 측면에 포함된 C-말단 및/또는 N-말단 서열(들)의 길이는 다양할 수 있으며, 예를 들어 문헌[Zhang et al. (2011)]의 연구를 기준으로 당업자에 의해 선택될 수 있다. 상기 문헌[Zhang et al.]은 Hax3 유도된 TAL-이펙터를 기초로 한 단백질에서 많은 C-말단 및 N-말단 절단 돌연변이체를 특징화하였으며, 표적 서열에 대한 최적 결합 및 이에 따라 전사 활성화에 기여하는, 주요 요소를 확인하였다. 일반적으로, 전사 활성은 N-말단의 길이와 역의 상관관계에 있는 것으로 밝혀졌다. C-말단에 대하여, Hax 3 서열의 최초 68개 아미노산 내 DNA 결합 잔기에 중요한 요소가 확인되었다. 이에 따라 일부 구현예에서, 자연에 존재하는 TAL 이펙터의 TAL 이펙터 도메인의 C-말단부 상에서 최초 68개 아미노산은 TAL 이펙터 분자 내에 포함된다. 따라서, 구현예에서 TAL 이펙터 분자는 1) 자연에 존재하는 TAL 이펙터로부터 유래된 하나 이상의 TAL 이펙터 도메인; 2) TAL 이펙터의 N-말단부에 대한 자연에 존재하는 TAL 이펙터로부터 적어도 70, 80, 90, 100, 110, 120, 130, 140, 150, 170, 180, 190, 200, 220, 230, 240, 250, 260, 270, 280개 이상의 아미노산; 및/또는 3) TAL 이펙터 도메인의 C-말단부 상에서 자연에 존재하는 TAL 이펙터로부터의 적어도 68, 80, 90, 100, 110, 120, 130, 140, 150, 170, 180, 190, 200, 220, 230, 240, 250, 260 이상의 아미노산을 포함한다.

[0747] 일부 구현예에서, 엔도뉴클레아제 도메인 또는 DNA-결합 도메인은 Zn 핑거 분자이거나, 또는 이를 포함한다. Zn 핑거 분자는 Zn 핑거 단백질, 예를 들어, 자연에 존재하는 Zn 핑거 단백질 또는 조작된 Zn 핑거 단백질, 또는 이의 단편을 포함한다. 많은 Zn 핑거 단백질이 당업자에게 알려져 있으며, 예를 들어 Sigma-Aldrich로부터, 상업적으로 구매가능하다.

[0748] 일부 구현예에서, Zn 핑거 분자는, 선택된 표적 DNA 서열에 결합하도록 조작된, 자연에 존재하지 않는 Zn 핑거 단백질을 포함한다. 예를 들어, 그 전체가 본 명세서에 참고로 포함된, 문헌[Beerli, et al. (2002) Nature Biotechnol. 20:135-141; Pabo, et al. (2001) Ann. Rev. Biochem. 70:313-340; Isalan, et al. (2001) Nature Biotechnol. 19:656-660; Segal, et al. (2001) Curr. Opin. Biotechnol. 12:632-637; Choo, et al. (2000) Curr. Opin. Struct. Biol. 10:411-416]; 미국 특허 6,453,242; 6,534,261; 6,599,692; 6,503,717; 6,689,558; 7,030,215; 6,794,136; 7,067,317; 7,262,054; 7,070,934; 7,361,635; 7,253,273; 및 미국 특허 공개 2005/0064474; 2007/0218528; 2005/0267061을 참조한다.

[0749] 조작된 Zn 핑거 단백질은, 자연에 존재하는 Zn 핑거 단백질에 비해, 새로운 결합 특이성을 가질 수 있다. 조작 방법은 합리적 설계(rational design) 및 다양한 유형의 선택을 포함하지만 이에 제한되지 않는다. 합리적 설계는, 예를 들어 삼중자(또는 사중자) 뉴클레오티드 서열 및 개별적인 Zn 핑거 아미노산 서열을 포함하는 데이터베이스의 이용을 포함하며, 여기서 각각의 삼중자 또는 사중자 뉴클레오티드 서열은, 특정 삼중자 또는 사중자 서열에 결합하는 징크 핑거의 아미노산 서열 하나 이상과 회합된다. 예를 들어, 그 전체가 본 명세서에 참고로 포함된, 미국 특허 6,453,242 및 6,534,261호를 참조한다.

[0750] 예시적인 선택 방법은, 파지 디스플레이 및 2-하이브리드 시스템을 포함하고, 미국 특허 5,789,538; 5,925,523; 6,007,988; 6,013,453; 6,410,248; 6,140,466; 6,200,759; 및 6,242,568; 국제 특허 공개 WO 98/37186; WO 98/53057; WO 00/27878; 및 WO 01/88197 및 GB 2,338,237에 개시되어 있다. 추가적으로, 징크 핑거 단백질에 대한 결합 특이성의 향상은, 예를 들어 국제 특허 공개 WO 02/077227에 기재되어 있다.

[0751] 추가적으로, 이들 및 기타 참고에 개시된 것과 같이, 징크 핑거 도메인 및/또는 다중-핑거의 징크 핑거 단백질은 임의의 적합한 링커 서열을 함께 이용하여 링크될 수 있으며, 이는 예를 들어 길이 5개 이상의 아미노산의 링커를 포함한다. 또한, 6개 이상의 아미노산 길이의 예시적인 링커 서열에 대해, 미국 특허 6,479,626; 6,903,185; 및 7,153,949를 참조한다. 본 명세서에 기재된 단백질은 단백질의 개별적인 징크 핑거들 간에 적합한 링커의 임의의 조합을 포함할 수 있다. 추가적으로, 징크 핑거 결합 도메인에 대한 결합 특이성의 향상은,

예를 들어 공동 소유된 국제 특허 공개 WO 02/077227에 기재되어 있다.

[0752] Zn 핑거 단백질 및 융합 단백질의 설계 및 구축 방법(및 이를 인코딩하는 폴리뉴클레오티드)는 당업자에게 알려져 있으며, 미국 특허 6,140,0815; 789,538; 6,453,242; 6,534,261; 5,925,523; 6,007,988; 6,013,453; 및 6,200,759; 국제 특허 공개 WO 95/19431; WO 96/06166; WO 98/53057; WO 98/54311; WO 00/27878; WO 01/60970; WO 01/88197; WO 02/099084; WO 98/53058; WO 98/53059; WO 98/53060; WO 02/016536; 및 WO 03/016496에 상세하게 기재되어 있다.

[0753] 추가적으로, 이들 및 기타 참고에 개시된 바와 같이, Zn 핑거 단백질 및/또는 다중-핑거 Zn 핑거 단백질은, 예를 들어 길이 5개 이상의 아미노산의 링커를 포함하는, 임의의 적합한 링커 서열을 이용하여, 예를 들어 융합 단백질로서, 함께 연결될 수 있다. 또한, 길이 6개 이상의 예시적인 아미노산 링커 서열에 대해 미국 특허 6,479,626; 6,903,185; 및 7,153,949를 참조한다. 본 명세서에 기재된 Zn 핑거 분자는, 개별적인 징크 핑거 단백질 및/또는 Zn 핑거 분자의 다중-핑거 Zn 핑거 단백질 간에 적합한 링커의 임의의 조합을 포함할 수 있다.

[0754] 특정 구현예에서, DNA-결합 도메인 또는 엔도뉴클레아제 도메인은 표적 DNA 서열에 (서열-특이적 방식으로) 결합하는 조작된 징크 핑거 단백질을 포함하는 Zn 핑거 분자를 포함한다. 일부 구현예에서, Zn 핑거 분자는 하나의 Zn 핑거 단백질 또는 이의 단편을 포함한다. 다른 구현예에서, Zn 핑거 분자는 복수의 Zn 핑거 단백질들(또는 이의 단편), 예를 들어, 2, 3, 4, 5, 6개 이상의 Zn 핑거 단백질(및 선택적으로 12, 11, 10, 9, 8, 7, 6, 5, 4, 3, 또는 2 이하의 Zn 핑거 단백질)을 포함한다. 일부 구현예에서, Zn 핑거 분자는 적어도 3개의 Zn 핑거 단백질을 포함한다. 일부 구현예에서, Zn 핑거 분자는 4, 5, 또는 6개의 핑거를 포함한다. 일부 구현예에서, Zn 핑거 분자는 8, 9, 10, 11 또는 12개의 핑거를 포함한다. 일부 구현예에서, 3개의 Zn 핑거 단백질을 포함하는 Zn 핑거 분자는 9 또는 10개의 뉴클레오티드를 포함하는 표적 DNA 서열을 인식한다. 일부 구현예에서, 4개의 Zn 핑거 단백질을 포함하는 Zn 핑거 분자는 12 내지 14개의 뉴클레오티드를 포함하는 표적 DNA 서열을 인식한다. 일부 구현예에서, 6개의 Zn 핑거 단백질을 포함하는 Zn 핑거 분자는 18 내지 21개의 뉴클레오티드를 포함하는 표적 DNA 서열을 인식한다.

[0755] 일부 구현예에서, Zn 핑거 분자는 양손(two-handed) Zn 핑거 단백질을 포함한다. 양손 징크 핑거 단백질은, 여기서 징크 핑거 단백질의 2개의 덩어리는 개재하는 아미노산들에 의해 분리되어, 2개의 징크 핑거 도메인이 2개의 불연속적인 표적 DNA 서열에 결합되도록 하는 단백질들이다. 징크 핑거 결합 단백질의 양손 유형의 예는 SIP1이며, 여기서 4개의 징크 핑거 단백질의 덩어리는 단백질의 아미노 말단에 위치되고, 3개의 징크 핑거 단백질의 덩어리는 카르복실 말단에 위치한다(문헌[Remade, et al. (1999) EMBO Journal 18(18):5073-5084] 참조). 이들 단백질의 징크 핑거의 각각의 덩어리는 이들 단백질에서 징크 핑거의 각각의 클러스터가 독특한 표적 서열에 결합할 수 있고, 2개의 표적 서열 사이의 간격은 많은 뉴클레오티드를 포함할 수 있다.

[0756] *링커*

[0757] 일부 구현예에서, 유전자 변형 폴리펩티드는 링커, 예를 들어 펩티드 링커, 예를 들어 표 1 또는 표 10에 기재된 바와 같은 링커를 포함할 수 있다. 일부 구현예에서, 유전자 변형 폴리펩티드는 N-말단에서 C-말단 방향으로, Cas 도메인(예를 들어, 표 8의 Cas 도메인), 표 10의 링커(또는 이에 대하여 적어도 70%, 80%, 85%, 90%, 95%, 또는 99% 동일성을 갖는 서열), 및 RT 도메인(예를 들어, 표 6의 RT 도메인)을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 엔도뉴클레아제와 RT 도메인 사이에 가요성 링커, 예를 들어 아미노산 서열 SGGSSGGSSGSETPGTSESATPESSGGSSGGSS를 포함하는 링커를 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드의 RT 도메인은 엔도뉴클레아제 도메인의 C-말단에 위치할 수 있다. 일부 구현예에서, 유전자 변형 폴리펩티드의 RT 도메인은 엔도뉴클레아제 도메인의 N-말단에 위치할 수 있다.

[0758] *표 10. 예시적인 링커 서열*

아미노산 서열	서열번호
GGS	101
GGSGGS	102
GGSGGSGGS	103
GGSGGSGGSGGS	104
GGSGGSGGSGGSGGS	105
GGSGGSGGSGGSGGSGGS	106
GGGGS	107
GGGGSGGGGS	108
GGGGSGGGGSGGGGS	109
GGGGSGGGGSGGGGSGGGGS	110

[0759]

GGGGSGGGGSGGGGSGGGGSGGGGS	111
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	112
GGG	113
GGGG	114
GGGGG	115
GGGGGG	116
GGGGGGG	117
GGGGGGGG	118
GSS	119
GSSGSS	120
GSSGSSGSS	121
GSSGSSGSSGSS	122
GSSGSSGSSGSSGSS	123
GSSGSSGSSGSSGSSGSS	124
EAAAK	125
EAAAKEAAAK	126
EAAAKEAAAKEAAAK	127
EAAAKEAAAKEAAAKEAAAK	128
EAAAKEAAAKEAAAKEAAAKEAAAK	129
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	130
PAP	131
PAPAP	132
PAPAPAP	133
PAPAPAPAP	134
PAPAPAPAPAP	135
PAPAPAPAPAPAP	136
GGSGGG	137
GGGGGS	138
GGSGSS	139

[0760]

GSSGGS	140
GGSEAAAK	141
EAAAKGGS	142
GGSPAP	143
PAPGGS	144
GGGGSS	145
GSSGGG	146
GGGEAAAK	147
EAAAKGGG	148
GGGPAP	149
PAPGGG	150
GSSEAAAK	151
EAAAKGSS	152
GSSPAP	153
PAPGSS	154
EAAAKPAP	155
PAPEAAAK	156
GGSGGGGSS	157
GGSGSSGGG	158
GGGGGSGSS	159
GGGGSSGGS	160
GSSGGSGGG	161
GSSGGGGGS	162
GGSGGGEAAAK	163
GGSEAAAKGGG	164
GGGGGSEAAAK	165
GGGEAAAKGGS	166
EAAAKGGSGGG	167
EAAAKGGGGGS	168

[0761]

[0762]

GGSGGGPAP	169
GGSPAPGGG	170
GGGGGSPAP	171
GGGPAPGGS	172
PAPGGSGGG	173
PAPGGGGGS	174
GGSGSSEAAAK	175
GGSEAAAKGSS	176
GSSGGSEAAAK	177
GSSEAAAKGGS	178
EAAAKGGSGSS	179
EAAAKGSSGGS	180
GGSGSSPAP	181
GGSPAPGSS	182
GSSGGSPAP	183
GSSPAPGGS	184
PAPGGSGSS	185
PAPGSSGGS	186
GGSEAAAKPAP	187
GGSPAPEAAAK	188
EAAAKGGSPAP	189
EAAAKPAPGGS	190
PAPGGSEAAAK	191
PAPEAAAKGGS	192
GGGGSSEAAAK	193
GGGEAAAKGSS	194
GSSGGGEAAAK	195
GSSEAAAKGGG	196
EAAAKGGGGSS	197

EAAAKGSSGGG	198
GGGGSSPAP	199
GGGPAPGSS	200
GSSGGGPAP	201
GSSPAPGGG	202
PAPGGGGSS	203
PAPGSSGGG	204
GGGEAAAKPAP	205
GGGPAPEAAAK	206
EAAAKGGGPAP	207
EAAAKPAPGGG	208
PAPGGGEAAAK	209
PAPEAAAKGGG	210
GSSEAAAKPAP	211
GSSPAPEAAAK	212
EAAAKGSSPAP	213
EAAAKPAPGSS	214
PAPGSSEAAAK	215
PAPEAAAKGSS	216
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	217
GGGGSEAAAKGGGGS	218
EAAAKGGGGSEAAAK	219
SGSETPGTSESATPES	220
GSAGSAAGSGEF	221
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	222

[0763]

[0764] 일부 구현예에서, 유전자 변형 폴리펩티드의 링커는 $(SGGS)_n$(서열번호 25), $(GGGS)_n$(서열번호 26), $(GGGGS)_n$(서열번호 27), $(G)_n$, $(EAAAK)_n$(서열번호 28), $(GGS)_n$, 또는 $(XP)_n$으로부터 선택되는 모티프를 포함한다.

[0765] *풀링된 스크리닝에 의한 유전자 변형 폴리펩티드 선택*

[0766] 후보 유전자 변형 폴리펩티드를 스크리닝하여 후보 유전자 편집 능력을 평가할 수 있다. 예를 들어, 인간 게놈에서 코딩 서열의 표적화된 편집을 위해 설계된 RNA 유전자 변형 시스템이 사용될 수 있다. 특정 구현예에서, 그러한 유전자 변형 시스템은 풀링된 스크리닝 접근법과 함께 사용될 수 있다.

[0767] 예를 들어, 유전자 변형 폴리펩티드 후보 및 주형 가이드 RNA(tgRNA)의 라이브러리를 포유동물 세포 내로 도입하여 풀링된 스크리닝 접근법에 의해 후보의 유전자 편집 능력을 시험할 수 있다. 특정 구현예에서, 유전자 변형 폴리펩티드 후보의 라이브러리를 포유동물 세포 내로 도입한 후 tgRNA를 세포 내로 도입한다.

[0768] 스크리닝에 사용될 수 있는 포유동물 세포의 대표적인 비제한적 예는 HEK293T 세포, U2OS 세포, HeLa 세포, HepG2 세포, Huh7 세포, K562 세포, 또는 iPS 세포를 포함한다.

[0769] 유전자 변형 폴리펩티드 후보는 1) Cas-뉴클레아제, 예를 들어 야생형 Cas 뉴클레아제, 예를 들어 야생형 Cas9 뉴클레아제, 돌연변이 Cas 뉴클레아제, 예를 들어 Cas 닉카제, 예를 들어 Cas9 닉카제, 예컨대 Cas9 N863A 닉카제, 또는 **표 7 또는 8**로부터 선택되는 Cas 뉴클레아제, 2) 다양한 정도의 길이, 가요성, 소수성, 및/또는 2차

구조를 나타낼 수 있는 펩티드 링커, 예를 들어 **표 1 또는 10**의 서열; 및 3) 역전사효소(RT), 예를 들어 **표 1 또는 6**의 RT 도메인을 포함할 수 있다. 유전자 변형 폴리펩티드 후보 라이브러리는 Cas 뉴클레아제, 펩티드 링커 또는 RT 도메인 구성요소 중 1개, 2개 또는 3개 전부와 관련하여 서로 상이한 복수의 상이한 유전자 변형 폴리펩티드 후보, 또는 그러한 유전자 변형 폴리펩티드 후보를 인코딩하는 복수의 핵산 발현 벡터를 포함한다.

[0770] 유전자 변형 폴리펩티드 후보의 스크리닝을 위해, 유전자 변형 폴리펩티드 구성요소 및 tgRNA 구성요소를 포함하는 2-구성요소 시스템이 사용될 수 있다. 유전자 변형 구성요소는, 예를 들어 유전자 변형 폴리펩티드 후보를 인코딩하고, 예를 들어 유전자 변형 폴리펩티드 후보를 인코딩하는 인간 코돈-최적화 핵산, 예를 들어 상기 기재된 바와 같은 Cas-링커-RT 융합체를 포함하는 발현 벡터, 예를 들어 발현 플라스미드 또는 렌티바이러스 벡터를 포함할 수 있다. 특정 구현예에서, (i) 포유동물 세포에서의 발현을 위한 프로모터, 예를 들어 CMV 프로모터; (ii) 유전자 변형 라이브러리 후보, 예를 들어 표 CC의 Cas 뉴클레아제, 표 AA의 펩티드 링커 및 표 BB의 RT를 포함하는 Cas-링커-RT 융합체, 예를 들어 **표 1**의 Cas-링커-RT 융합체; (iii) 자가-절단 폴리펩티드, 예를 들어 T2A 펩티드; (iv) 포유동물 세포에서 선택을 가능하게 하는 마커, 예를 들어 퓨로마이신 내성 유전자; 및 (v) 종결 신호, 예를 들어 폴리 A 테일을 포함하는 렌티바이러스 카세트가 이용된다.

[0771] tgRNA 구성요소는 tgRNA를 생산하고, tgRNA의 발현을 구동하는 U6 프로모터를 이용하며(여기서, tgRNA는 Cas에 의해 인식되는 비-코딩 RNA 서열임), tgRNA를 관심 게놈 유전자좌에 국재화시키며, 또한 RT 도메인에 의해 원하는 편집물을 게놈으로 역전사하는 주형을 제공하는, tgRNA 또는 발현 벡터, 예를 들어 발현 플라스미드를 포함할 수 있다.

[0772] 유전자 변형 폴리펩티드 라이브러리 후보를 발현하는 세포의 풀을 준비하기 위해, 포유동물 세포, 예를 들어 HEK293T 또는 U2OS 세포는 유전자 변형 후보 폴리펩티드 라이브러리의 풀링된 유전자 변형 폴리펩티드 후보 발현 벡터 제제, 예를 들어 렌티바이러스 제제를 이용하여 형질도입될 수 있다. 특정 구현예에서, 렌티바이러스 플라스미드가 이용되고, HEK293 Lenti-X 세포는 렌티바이러스 플라스미드 형질감염 전에 15 cm 플레이트(약 $12x10^6$개 세포)에 시딩된다. 그러한 구현예에서, 렌티바이러스 플라스미드 형질감염은 Lentiviral Packaging Mix(Biosettia)를 사용하여 수행될 수 있으며 유전자 변형 후보 라이브러리에 대한 플라스미드 DNA의 형질감염은 제조업체의 프로토콜에 따라 Lipofectamine 2000 및 Opti-MEM 배지를 사용하여 다음 날 수행된다. 그러한 구현예에서, 세포외 DNA는 다음 날 전체 배지 교체에 의해 제거될 수 있고 바이러스-함유 배지는 48시간 후에 수확될 수 있다. 렌티바이러스 배지는 Lenti-X Concentrator(TaKaRa Biosciences)를 사용하여 농축될 수 있으며 5 mL의 렌티바이러스 분취액이 제조되고 -80° C에서 보관될 수 있다. 렌티바이러스 적정은 선택 후, 예를 들어 퓨로마이신 선택 후 콜로니 형성 단위를 계수함으로써 수행된다.

[0773] 표적 DNA의 유전자 편집을 모니터링하기 위해, 표적 DNA를 운반하는 포유동물 세포, 예를 들어 HEK293T 또는 U2OS 세포가 이용될 수 있다. 표적 DNA의 유전자 편집을 모니터링하기 위한 다른 구현예에서, 표적 DNA 게놈 랜딩 패드(landing pad)를 운반하는 포유동물 세포, 예를 들어 HEK293T 또는 U2OS 세포가 이용될 수 있다. 특정 구현예에서, 표적 DNA 게놈 랜딩 패드는 관심 질환 또는 장애의 치료를 위해 편집될 유전자를 포함할 수 있다. 다른 특정 구현예에서, 표적 DNA는 유전자 편집이 발생했는지 여부를 결정하기 위해 모니터링될 수 있는 검출가능한 특성을 나타내는 단백질을 발현하는 유전자 서열이다. 예를 들어, 특정 구현예에서, 청색 형광 단백질(BFP)- 또는 녹색 형광 단백질(GFP)-발현 게놈 랜딩 패드가 이용된다. 특정 구현예에서, 표적 DNA, 예를 들어 표적 DNA 게놈 랜딩 패드를 포함하는 포유동물 세포, 예를 들어 HEK293T 또는 U2OS 세포는 유전자 변형 라이브러리 후보당 500x 내지 3000x 세포로 배양 플레이트에 시딩되고 0.2 내지 0.3의 감염다중도(MOI)로 형질도입되어 세포당 다중 감염을 최소화한다. 감염된 세포의 선택을 가능하게 하기 위해 감염 후 48시간 후에 퓨로마이신(2.5 ug/mL)이 첨가될 수 있다. 그러한 구현예에서, 세포는 적어도 7일 동안 퓨로마이신 선택 하에 유지된 다음, tgRNA 도입, 예를 들어 tgRNA 전기천공을 위해 규모가 확대될 수 있다.

[0774] 유전자 편집이 발생하는지 여부를 확인하기 위해, 편집될 표적 DNA를 함유하는 포유동물 세포를 유전자 변형 폴리펩티드 라이브러리 후보로 감염시킨 다음 표적 DNA 편집에 사용하도록 설계된 tgRNA로 형질감염시킬 수 있다. 이어서, 세포를 분석하여 설계된 결과에 따라 표적 유전자좌의 편집이 발생했는지, 또는 편집이 발생하지 않았거나 불완전한 편집이 발생했는지 여부를, 예를 들어 세포 분류 및 서열 분석을 사용하여 결정할 수 있다.

[0775] 특정 구현예에서, 게놈 편집이 발생하는지 여부를 확인하기 위해, BFP- 또는 GFP-발현 포유동물 세포, 예를 들어 HEK293T 또는 U2OS 세포를 유전자 변형 라이브러리 후보로 감염시킨 다음, 예를 들어 라이브러리 후보당 250x 내지 1000x 초과의 커버리지를 보장하는 세포 수에서 BFP에서 GFP로 또는 GFP에서 BFP로 전환시키도록 설계된 20 ng의 tgRNA 플라스미드를 이용하여 웰당 250,000개 세포를 전기천공시킴으로써, tgRNA 플라스미드 또는

RNA로 형질감염시키거나 전기천공시킬 수 있다. 그러한 구현예에서, 이 분석에서 다양한 작제물의 게놈 편집 능력은 전기천공 후 4 내지10일 후에 색-전환 형광 단백질(FP)의 발현을 위해 형광 활성화 세포 분류(FACS)에 의해 세포를 분류함으로써 평가될 수 있다. 세포는 편집되지 않은 세포(원래의 형광 단백질 신호를 나타냄), 편집된 세포(전환된 형광 단백질 신호를 나타냄) 및 불완전한 편집(형광 단백질 신호를 나타내지 않음) 세포의 별개의 집단으로 분류 및 수확된다. 분석 동안에 후보 농축을 결정하기 위해 분류되지 않은 세포 샘플은 또한 입력 집단으로 수확될 수 있다.

[0776] 어떤 유전자 변형 라이브러리 후보가 분석에서 게놈 편집 능력을 나타내는지를 결정하기 위해, 분류된 세포 집단에서 게놈 DNA(gDNA)를 수확하고, 각각의 집단에서 유전자 변형 라이브러리 후보를 시퀀싱함으로써 분석한다. 간략하게, 유전자 변형 후보는 유전자 변형 폴리펩티드 발현 벡터, 예를 들어 렌티바이러스 카세트에 특이적인 프라이머를 사용하여 게놈으로부터 증폭될 수 있고, 게놈 DNA를 희석하기 위해 PCR의 두 번째 라운드에서 증폭된 다음 시퀀싱되며, 예를 들어 차세대 시퀀싱 플랫폼에 의해 시퀀싱될 수 있다. 시퀀싱 판독물의 품질 관리 후, 적어도 약 1500개 뉴클레오티드 및 일반적으로 약 3200개 이하의 뉴클레오티드의 판독물이 유전자 변형 폴리펩티드 라이브러리 서열에 맵핑되고 라이브러리 서열과 최소 약 80% 일치하는 판독물은 이러한 풀링된 스크린의 목적을 위해 주어진 후보에 대해 성공적으로 정렬된 것으로 간주된다. 분석에서 유전자 편집, 예를 들어 BFP에서 GFP로 또는 GFP에서 BFP로의 편집을 수행할 수 있는 후보를 확인하기 위해, 편집된 집단에 있는 각각의 라이브러리 후보의 판독 수를 초기의 분류되지 않은 집단에서 판독 수와 비교한다.

[0777] 풀링된 스크리닝의 목적을 위해, 게놈 편집 능력이 있는 유전자 변형 후보는 분류되지 않은 (입력) 세포에 비해 편집된(전환된 FP) 집단의 농축을 기준으로 확인된다. 일부 구현예에서, 입력에 대한 적어도 1.0, 1.5, 2.0, 2.5, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90배, 또는 적어도 100배의 농축, 예를 들어 적어도 2배의 농축은 잠재적으로 유용한 유전자 편집 활성을 나타낸다. 일부 구현예에서, 농축은 농축 비율의 로그 베이스 2를 위함으로써 로그 값으로 전환된다. 일부 구현예에서, 적어도 0, 1, 2, 3, 4, 5, 5.5, 6.0, 6.2, 6.3, 6.4, 6.5, 또는 적어도 6.6의 log2 농축 점수, 예를 들어 적어도 1.0의 log2 농축 점수는 잠재적으로 유용한 유전자 편집 활성을 나타낸다. 특정 구현예에서, 유전자 변형 후보에 대해 관찰된 농축 값은 참조, 예를 들어 요소 ID 번호 17380을 이용하여 유사한 조건 하에 관찰된 농축 값과 비교될 수 있다.

[0778] 일부 구현예에서, 다중 tgRNA가 유전자 변형 후보 라이브러리를 스크리닝하는 데 사용될 수 있다. 특정 구현예에서, 복수의 tgRNA가, 예를 들어 특정 표적 유전자, 예를 들어 질환 치료를 위한 유전자 표적의 유전자 편집을 위해, 주형/Cas-링커-RT 융합 쌍을 최적화하는 데 이용될 수 있다. 특정 구현예에서, 유전자 변형 후보를 스크리닝하기 위한 풀링된 접근법은 배열된 형식으로 다수의 상이한 tgRNA를 사용하여 수행될 수 있다.

[0779] 일부 구현예에서, 다양한 유형의 편집, 예를 들어 다양한 길이의 삽입, 치환, 및/또는 결실이 유전자 변형 후보 라이브러리를 스크리닝하는 데 사용될 수 있다.

[0780] 일부 구현예에서, 다중 표적 서열, 예를 들어 다양한 형광 단백질이 유전자 변형 후보 라이브러리를 스크리닝하는 데 사용될 수 있다. 일부 구현예에서, 다중 표적 서열, 예를 들어 다양한 형광 단백질이 유전자 변형 후보 라이브러리를 스크리닝하는 데 사용될 수 있다. 일부 구현예에서, 다중 세포 유형, 예를 들어 HEK293T 또는 U2OS가 유전자 변형 후보 라이브러리를 스크리닝하는 데 사용될 수 있다. 당업자는 주어진 후보가 변경된 편집 능력을 나타낼 수 있거나 심지어 tgRNA 서열(예를 들어, 뉴클레오티드 변형, PBS 길이, RT 주형 길이), 표적 서열, 표적 위치, 편집 유형, 유전자 변형 폴리펩티드의 제1 가닥 닉에 대한 돌연변이의 위치, 또는 세포 유형을 포함하여, 다양한 조건에 걸쳐 임의의 관찰가능하거나 유용한 활성의 이득 또는 손실을 나타낼 수 있음을 이해할 것이다. 따라서, 일부 구현예에서, 유전자 변형 라이브러리 후보는, 예를 들어 적어도 2가지의 세포 유형에서 적어도 2개의 별개의 tgRNA를 이용하여 다수의 파라미터에 걸쳐 스크리닝되고, 유전자 편집 활성은 임의의 단일 조건에서의 강화에 의해 확인된다. 다른 구현예에서, 다양한 tgRNA 및 세포 유형에 걸쳐 더 강력한 활성을 갖는 후보는 적어도 2가지 조건, 예를 들어 스크리닝된 모든 조건에서의 강화에 의해 확인된다. 명확성을 위해, 임의의 주어진 조건 하에서 강화가 거의 나타나지 않거나 전혀 나타나지 않은 것으로 확인된 후보는 모든 조건에 걸쳐 비활성인 것으로 가정되지 않으며 다양한 파라미터로 스크리닝되거나, 예를 들어 도메인(예를 들어, RT 도메인), 링커, 또는 다른 신호(예를 들어, NLS)를 교환, 셔플링, 또는 변화시킴으로써, 폴리펩티드 수준에서 재구성될 수 있다.

[0781] *예시적인 Cas9-링커-RT 융합체의 서열*

[0782] 일부 구현예에서, 유전자 변형 폴리펩티드는 링커 서열 및 RT 서열을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 표 1에 열거된 바와 같은 링커 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%,

97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 표 1에 열거된 바와 같은 RT 도메인의 아미노산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 표 1에 열거된 바와 같은 링커 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열; 및 표 1에 열거된 바와 같은 RT 도메인의 아미노산, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 (i) 표 1의 행에 열거된 바와 같은 링커 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열; 및 (ii) 표 1의 동일한 행에 열거된 바와 같은 RT 도메인의 아미노산, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함한다. 표 1에 명명된 각각의 RT 도메인에 대해, 해당하는 아미노산 서열은 본 명세서의 표 6에서 확인할 수 있다.

[0783] *이량체화 도메인*

[0784] 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 시스템은, 예를 들어 Cas 도메인(예를 들어, Cas9 도메인, 예를 들어 nCas9 또는 dCas9 도메인)을 포함하는, DNA 결합 도메인(DBD); RNA 결합 도메인(RBD); 및 레트로바이러스 역전사효소(RT) 도메인을 포함한다. 일부 구현예에서, DBD는 2개의 이량체화 도메인 사이의 결합을 통해 RBD에 부착된다. 일부 구현예에서, DBD는 2개의 이량체화 도메인 사이의 결합을 통해 RT 도메인에 부착된다. 일부 구현예에서, RT 도메인은 2개의 이량체화 도메인 사이의 결합을 통해 RBD에 부착된다.

[0785] 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드 또는 복합체에 포함된 한 쌍의 이량체화 도메인은 화합물(예를 들어, 소분자)에 의해 이량체화하도록 유도될 수 있다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드 또는 복합체에 포함된 한 쌍의 이량체화 도메인은 빛(예를 들어, 특정 색상 및/또는 파장의 빛)에 대한 노출에 의해 이량체화하도록 유도될 수 있다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드 또는 복합체에 포함된 한 쌍의 이량체화 도메인은 표 34의 단일 행에 열거된 바와 같은 사슬 A 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열) 및 사슬 B 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열)을 포함한다. 구현예에서, 한 쌍의 이량체화 도메인은 표 34의 동일한 행에 열거된 유도인자에 의해 유도될 수 있다.

[0786] 표 34. 예시적인 화학물질- 또는 광-유도 이량체화 도메인

유도인자(들)	사슬 A 명칭	사슬 A 서열	예시적인 사슬 A 공급원	사슬 B 명칭	사슬 B 서열	예시적인 사슬 B 공급원
라파마이신/라파로그	FKBP	GVQVETISPGDGRTFPKRGQTCVVHY TGMLEDGKKFDSSRDRNKPFKFMLGK QEVIRGWEEGVAQMSVGQRAKLTISP DYAYGATGHPGIIPPHATLVFDVELL KLE	snapgene	FRB	ILWHEMWHEGLEEASRLYFGERNV KGMFEVLEPLHAMMERGPQTLKET SFNQAYGRDLMEAQEWCRKYMKSG NVKDLTQAWDLYYHVFRRISKQL	snapgene
라파마이신/라파로그	FKBP	GVQVETISPGDGRTFPKRGQTCVVHY TGMLEDGKKFDSSRDRNKPFKFMLGK QEVIRGWEEGVAQMSVGQRAKLTISP DYAYGATGHPGIIPPHATLVFDVELL KLE	snapgene	FRB	EMWHEGLEEASRLYFGERNVKGMF EVLEPLHAMMERGPQTLKETSFNQ AYGRDLMEAQEWCRKYMKSGNVKD LTQAWDLYYHVFRRISKQL	snapgene
라파마이신/라파로그	FKBP	GVQVETISPGDGRTFPKRGQTCVVHY TGMLEDGKKFDSSRDRNKPFKFMLGK QEVIRGWEEGVAQMSVGQRAKLTISP DYAYGATGHPGIIPPHATLVFDVELL KLE	snapgene	FRB*	ILWHEMWHEGLEEASRLYFGERNV KGMFEVLEPLHAMMERGPQTLKET SFNQAYGRDLMEAQEWCRKYMKSG NVKDLLQAWDLYYHVFRRIS	addgene 108836 (pBW1308)
라파마이신/라파로그	FKBP	GVQVETISPGDGRTFPKRGQTCVVHY TGMLEDGKKFDSSRDRNKPFKFMLGK QEVIRGWEEGVAQMSVGQRAKLTISP DYAYGATGHPGIIPPHATLVFDVELL KLE	snapgene	FRB*	ILWHEMWHEGLEEASRLYFGERNV KGMFEVLEPLHAMMERGPQTLKET SFNQAYGRDLMEAQEWCRKYMKSG NVKDLLQAWDLYYHVFRRISK	snapgene
라파마이신/라파로그	FKBP	SRGVQVETISPGDGRTFPKRGQTCVV HYTGMLEDGKKFDSSRDRNKPFKFML GKQEVIRGWEEGVAQMSVGQRAKLTI SPDYAYGATGHPGIIPPHATLVFDVE LLKLE	addgene 108837 (pBHW1309)	FRB	ILWHEMWHEGLEEASRLYFGERNV KGMFEVLEPLHAMMERGPQTLKET SFNQAYGRDLMEAQEWCRKYMKSG NVKDLTQAWDLYYHVFRRISK	snapgene
라파마이신/라파로그	FKBP	SRGVQVETISPGDGRTFPKRGQTCVV HYTGMLEDGKKFDSSRDRNKPFKFML GKQEVIRGWEEGVAQMSVGQRAKLTI	addgene 108837	FRB	EMWHEGLEEASRLYFGERNVKGMF EVLEPLHAMMERGPQTLKETSFNQ	snapgene

[0787]

			(pBHW1309)			
라파마이신/라파로그	FKBP	SRGVQVETISPGDGRTFPKRGQTCVV HYTGMLEDGKKFDSSRDRNKPFKFML GKQEVIRGWEEGVAQMSVGQRAKLTI SPDYAYGATGHPGIIPPHATLVFDVE LLKLE	addgene 108837 (pBHW1309)	FRB*	AYGRDLMEAQEWCRKYMKSGNVKD LTQAWDLYYHVFRRISKQL	addgene 108836 (pBW1308)
라파마이신/라파로그	FKBP	SRGVQVETISPGDGRTFPKRGQTCVV HYTGMLEDGKKFDSSRDRNKPFKFML GKQEVIRGWEEGVAQMSVGQRAKLTI SPDYAYGATGHPGIIPPHATLVFDVE LLKLE	addgene 108837 (pBHW1309)	FRB*	ILWHEMWHEGLEEASRLYFGERNV KGMFEVLEPLHAMMERGPQTLKET SFNQAYGRDLMEAQEWCRKYMKSG NVKDLLQAWDLYYHVFRRISK	snapgene
맵시스산	ABI	PLYGFTSICGRRPEMEAAVSTIPRFL QSSSGSMLDGRTDPQSAAHFFGVYDG HGGSQVANYCREMHLALAEEIAKEK PMLCDGDTWLEKWKKALFNSFLRVDS EIESVAPETVGSTSVVAVVFPSHIFV ANCGDSRAVLCRGKTALPLSVDHKPD REDEAARIEAAGGKVIQWNGARVFGV LAMSRSIGDRYLKPSIIPDPEVTAVK RVKEDDCLILASDGVWDVMTDEEACE MARKRILLWHKKNAVAGDASLLADER RKEGKDPAAMSAAEYLSKLAIQRGSK DNISVVVVDLK	addgene 135985 (TL)	PYL	APTQDEFTQLSQSIAEFHTYQLGN GRCSSLLAQRIHAPPETVWSVVRR FDRPQIYKHFIKSCNVSEDFEMRV GCTRDVNVISGLPANTSRERLDLL DDRRVTGFSITGGEHRLRNYKSV TTVHRFEKEEEEERIWTVVLESYV VDVPEGNSEEDTRLFADTVIRLNL QKLASITEAMN	addgene 135988 (TL)
맵시스산	ABI	PLYGFTSICGRRPEMEAAVSTIPRFL QSSSGSMLDGRFDPQSAAHFFGVYDG HGGSQVANYCREMHLALAEEIAKEK PMLCDGDTWLEKWKKALFNSFLRVDS EIESVAPETVGSTSVVAVVFPSHIFV ANCGDSRAVLCRGKTALPLSVDHKPD REDEAARIEAAGGKVIQWNGARVFGV LAMSRSIGDRYLKPSIIPDPEVTAVK RVKEDDCLILASDGVWDVMTDEEACE MARKRILLWHKKNAVAGDASLLADER	addgene 135985 (TL)	PYL	APTQDEFTQLSQSIAEFHTYQLGN GRCSSLLAQRIHAPPETVWSVVRR FDRPQIYKHFIKSCNVSEDFEMRV GCTRDVNVISGLPANTSRERLDLL DDRRVTGFSITGGEHRLRNYKSV TTVHRFEKEEEEERIWTVVLESYV VDVPEGNSEEDTRLFADTVIRLNL QKLASITEAMNYPYDVPDYA	addgene 108841 (pBW1313

[0788]

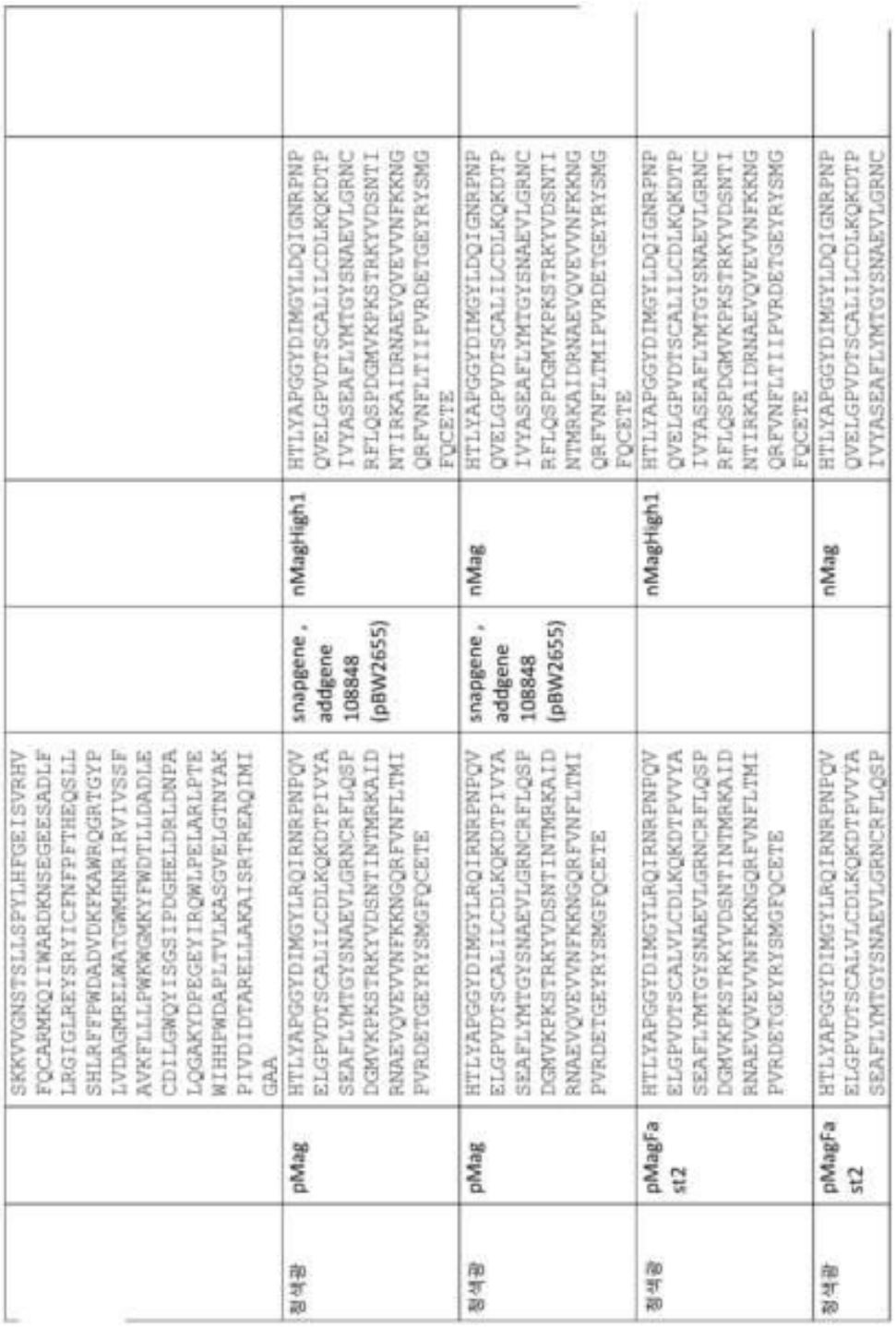
[0789]

		RKEGKDPAAMSAAEYLSKLAIQRGSK DNISVVVVDLK					
엡시스산	ABI	PLYGFTSICGRRPEMEDAVSTIPRFL QSSSGSMLDGRFDPQSAAHFFGVYDG HGGSQVANYCREHMHLALAEEIAKEK PMLCDGDTWLEKWKKALFNSFLRVDS EIGSVAPETVGSTSVVAVVFPSHIFV ANCGDSRAVLCRGKTALPLSVDHKPD REDEAARIEAAGGKVIQWNGARVFGV LAMSRSIGDRYLKPSIIPDPEVTAVK RVKEDDCLILASDGVWDVMTDEEACE MARKRILLWHKKNAVAGDASLLADER RKEGKDPAAMSAAEYLSKLAIQRGSK DNISVVVVDLK	addgene 108839 (pBW1311)	PYL	APTQDEFTQLSQSIAEFHTYQLGN GRCSSLLAQRIHAPPETVWSVVRR FDRPQIYKHFIKSCNVSEDFEMRV GCTRDVNVISGLPANTSRERLDLL DDDRRVTGFSITGGEHRLRNYKSV TTVHRFEKEEEEERIWTVVLESYV VDVPEGNSEEDTRLFADTVIRLNL QKLASITEAMNYPYDVPDYA	addgene 135988 (TL)
엡시스산	ABI	PLYGFTSICGRRPEMEDAVSTIPRFL QSSSGSMLDGRFDPQSAAHFFGVYDG HGGSQVANYCREHMHLALAEEIAKEK PMLCDGDTWLEKWKKALFNSFLRVDS EIGSVAPETVGSTSVVAVVFPSHIFV ANCGDSRAVLCRGKTALPLSVDHKPD REDEAARIEAAGGKVIQWNGARVFGV LAMSRSIGDRYLKPSIIPDPEVTAVK RVKEDDCLILASDGVWDVMTDEEACE MARKRILLWHKKNAVAGDASLLADER RKEGKDPAAMSAAEYLSKLAIQRGSK DNISVVVVDLK	addgene 108839 (pBW1311)	PYL	APTQDEFTQLSQSIAEFHTYQLGN GRCSSLLAQRIHAPPETVWSVVRR FDRPQIYKHFIKSCNVSEDFEMRV GCTRDVNVISGLPANTSRERLDLL DDDRRVTGFSITGGEHRLRNYKSV TTVHRFEKEEEEERIWTVVLESYV VDVPEGNSEEDTRLFADTVIRLNL QKLASITEAMNYPYDVPDYA	addgene 108841 (pBW1313)
지베렐린 /GA3-AM (지베렐린 에스테르)	GAI	KRDHHHHHHQDKKTMMMNEEDDGNGM DELLAVLGYKVRSSEMADVAQKLEQL EVMMSNVQEDDLSQLATETVHYNPAE LYTWLDSMLTDLN	addgene 108845 (pBW2067)	GID1	AASDEVNLIESRTVVPLNTWVLIS NFKVAYNILRRPDGTFNRHLAEYL DRKVTANANPVDGVFSFDVLIDRR INLLSRVYRPAYADQEQPPSILDL EKPVDGDIVPVILFFHGGSFAHSS ANSAIYDTLCRRLVGLCKCVVVSV NYRRAPENPYPCAYDDGWIALNWV NSRSWLKSKKDSKVHIFLAGDSSG GNIAHNVALRAGESGIDVLGNILL	addgene 108843 (pBW2065

[0790]

청색광	CRY2	KMDKKTIVWFRRDLRIEDNPALAAAA HEGSVFPVFIWCPEEEGQFYPGRASR WWMKQSLAHLSQSLKALGSDLTLIKT HNTISAILDCIRVTGATKVVFNHLYD PVSLVRDHTVKEKLVERGISVQSYNG DLLYEPWEIYCEKGKPFTSFNSYWKK CLDMSIESVMLPFPWRLMPITAAAEA IWACSIEELGLENEAEKPSNALLTRA WSPGWSNADKLLNEFIEKQLIDYAKN SKKVVGNSTSLLSPYLHFGEISVRHV FQCARMKQIIWARDKNSEGEESADLF LRGIGLREYSRYICFNFPFTHEQSLL SHLRFFPWDADVDKFKAWRQGRTGYP LVDAGMRELWATGWMHNRIRVIVSSF AVKFLLLPWKWGMKYFWDTLLDADLE CDILGWQYISGSIPDGHELDRLDNPA LQGAKYDPEGEYIRQWLPELARLPTE WIHHPWDAPLTVLKASGVELGTNYAK PIVDIDTARELLAKAISRTREAQIMI GAA	addgene 135989 (TL)	CIBN		NPMPGGNERTESEKSLDGKYFVTV RDRDWYWKAFLPEGEDREHPACNP FSPRGKSLEGVSFPKSLVVVAGLD LIRDWQLAYAEGLKKAGQEVKLMH LEKATVGFYLLPNNNHFHNVMDEI SAFVNAEC	addgene 135986 (TL)
청색광	CRY2	KMDKKTIVWFRRDLRIEDNPALAAAA HEGSVFPVFIWCPEEEGQFYPGRASR WWMKQSLAHLSQSLKALGSDLTLIKT HNTISAILDCIRVTGATKVVFNHLYD PVSLVRDHTVKEKLVERGISVQSYNG DLLYEPWEIYCEKGKPFTSFNSYWKK CLDMSIESVMLPFPWRLMPITAAAEA IWACSIEELGLENEAEKPSNALLTRA WSPGWSNADKLLNEFIEKQLIDYAKN	addgene 135989 (TL)	CIBN		NGAIGGDLLLNFPDMSVLERQRAH LKYLNPTFDSPLAGFFADSSMITG GEMDSYLSTAGLNLPMMYGETTVE GDSRLSISPETTLGTGNFKKRKFD TETKDCNEKKKMTMNRDDLVEEG EEEKSKITEQNNGSTKSIKKMKHK AKKEENNFSNDSSKVTKELEKTDY I	snapgene

[0791]



		SKKVVGNSTSLLSPYLHFGEISVRHV FQCARMKQIIWARDKNSEGEESADLF LRGIGLREYSRYICFNFPFTHEQSLL SHLRFFPWDADVDKFKAWRQGRTGYP LVDAGMREIWATGWMHNRIRVIVSSF AVKFLLLPWKWGMKYFWDTLLDADLE CDILGWQYISGSIPDGHELDRLDNPA LQGAKYDFEGEYIRQWLPELARLPTE WIHHPWDAPLTVLKASGVELGTNYAK PIVDIDTAREILARAISRTREAQIMI GAA			
청색광	pMag	HTLYAPGGYDIMGYLRQIRNRPNPQV ELGPVDTSCALILCDLKQKDTPIVYA SEAFLYMTGYSNAEVLGRNCRFLQSP DGMVKPKSTRKYVDSNTINTMRKAID RNAEVQVEVVNFKKNGQRFVNFLTMI PVRDETGEYRYSMGFQCETE	snapgene , addgene 108848 (pBW2655)	nMagHigh1		HTLYAPGGYDIMGYLDQIGNRPNP QVELGPVDTSCALILCDLKQKDTP IVYASEAFLYMTGYSNAEVLGRNC RFLQSPDGMVKPKSTRKYVDSNTI NTIRKAIDRNAEVQVEVVNFKKNG QRFVNFLTIIPVRDETGEYRYSMG FQCETE
청색광	pMag	HTLYAPGGYDIMGYLRQIRNRPNPQV ELGPVDTSCALILCDLKQKDTPIVYA SEAFLYMTGYSNAEVLGRNCRFLQSP DGMVKPKSTRKYVDSNTINTMRKAID RNAEVQVEVVNFKKNGQRFVNFLTMI PVRDETGEYRYSMGFQCETE	snapgene , addgene 108848 (pBW2655)	nMag		HTLYAPGGYDIMGYLDQIGNRPNP QVELGPVDTSCALILCDLKQKDTP IVYASEAFLYMTGYSNAEVLGRNC RFLQSPDGMVKPKSTRKYVDSNTI NTMRKAIDRNAEVQVEVVNFKKNG QRFVNFLTMIPVRDETGEYRYSMG FQCETE
청색광	pMagFa st2	HTLYAPGGYDIMGYLRQIRNRPNPQV ELGPVDTSCALVLCDLKQKDTPVVYA SEAFLYMTGYSNAEVLGRNCRFLQSP DGMVKPKSTRKYVDSNTINTMRKAID RNAEVQVEVVNFKKNGQRFVNFLTMI PVRDETGEYRYSMGFQCETE		nMagHigh1		HTLYAPGGYDIMGYLDQIGNRPNP QVELGPVDTSCALILCDLKQKDTP IVYASEAFLYMTGYSNAEVLGRNC RFLQSPDGMVKPKSTRKYVDSNTI NTIRKAIDRNAEVQVEVVNFKKNG QRFVNFLTIIPVRDETGEYRYSMG FQCETE
청색광	pMagFa st2	HTLYAPGGYDIMGYLRQIRNRPNPQV ELGPVDTSCALVLCDLKQKDTPVVYA SEAFLYMTGYSNAEVLGRNCRFLQSP		nMag		HTLYAPGGYDIMGYLDQIGNRPNP QVELGPVDTSCALILCDLKQKDTP IVYASEAFLYMTGYSNAEVLGRNC

[0792]

		DGMVKPKSTRKYVDSNTINTMRKAID RNAEVQVEVVNFKKNGQRFVNFLTMI PVRDETGEYRYSMGFQCETE				RFLQSPDGMVKPKSTRKYVDSNTI NTMRKAIDRNAEVQVEVVNFKKNG QRFVNFLTMIPVRDETGEYRYSMG FQCETE
적색광	PhyB	VSGVGGSGGGRGGGRGGEEEPSSSHT PNNRRGGEQAQSSGTKSLRPRSNTES MSKAIQQYTVDARLHAVFEQSGESGK SFDYSQSLKTTTYGSSVPEQQITAYL SRIQRGGYIQPFGCMIAVDESSFRII GYSENAREMLGIMPQSVPTLEKPEIL AMGTDVRSLFTSSSSILLERAFVARE ITLLNPVWIHSKNTGKPFYAILHRID VGVVIDLEPARTEDPALSIAGAVQSQ KLAVRAISQLQALPGGDIKLLCDTVV ESVRDLTGYDRVMVYKFHEDEHGEVV AESKRDDLEPYIGLHYPATDIPQASR FLFKQNRVRMIVDCNATPVLVVQDDR LTQSMCLVGSTLRAPHGCHSQYMANM GSIASLAMAVIINGNEDDGSNVASGR SSMRLMGLVVCHHTSSRCIPFPLRYA CEFLMQAFGLQLNMELQLALQMSEKR VLRTQTLLCDMLLRDSPAGIVTQSPS IMDLVKCDGAAFLYHGKYYPLGVAPS EVQIKDVVEWLLANHADSTGLSTDSL GDAGYPGAAALGDAVCGMAVAYITKR DFLFWFRSHTAKEIKWGGAKHHPEDK DDGQRMHPRSSFQAFLEVVKSRSQPW ETAEMDAIHSLQLILRDSFKESEAAM NSKVVDGVVQPCRDMAGEQGIDELGA VAREMVRLIETATVPIFAVDAGGCIN GWNAKIAELTGLSVEEAMGKSLVSDL IYKENEATVNKLLSRALRGDEEKNVE VKLKTFSPELQGKAVFVVVNACSSKD YLNNIVGVCFVGQDVTSQKIVMDKFI NIQGDYKAIVHSPNPLIPPIFAADEN	pBW2682	PIF6		pBW2684

[0793]

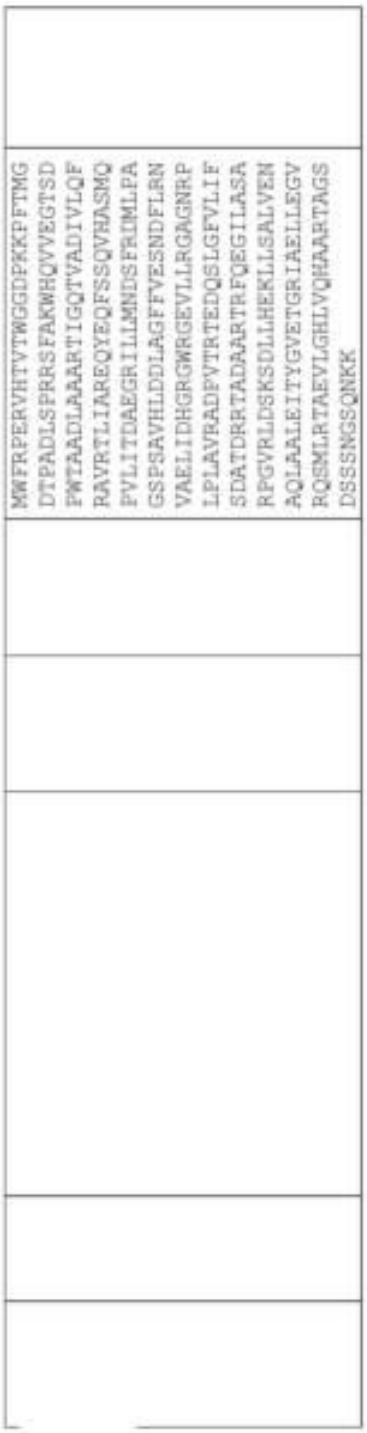
적색광	PhyB		pBW2682	PIF6		pBW2684
		TCCLEWNMAMEKLTGWSRSEVIGKMI				
		VGEVFGSCCMLKGPDALTKFMIVLHN				
		AIGGQDTDKFPFPFFDRNGKFVQALL				
		TANKVSLEGKVIGAFCFLQIPS				
	PhyB	VSGVGGSGGGRGGGRGGGEEEPSSSHT	pBW2682	PIF6		pBW2684
		PNNRRGGEQAQSSGTKSLRPRSNTES				
		MSKAIQQYTVDARLHAVFEQSGESGK				
		SFDYSQSLKTTTYGSSVPEQQITAYL				
		SRIQRGGYIQPFGCMIAVDESSFRII				
		GYSENAREMLGIMPQSVPTLEKPEIL				
		AMGTDVRSLFTSSSSILLERAFVARE				
		ITLLNPVWIHSKNTGKPFYAILHRID				
		VGVVIDLEPARTEDPALSIAGAVQSQ				
		KLAVRAISQLQALPGGDIKLLCDTVV				
		ESVKDLTGYDRVMVYKFHEDEHGEVV				
		AESKRDDLEPYIGLHYPATDIPQASR				
		FLFKQNAVNMIVDCNATFVLVVQDDR				
		LTQSMCLVGSTLRAPHGCHSQYMANM				
		GSIASLAMAVIINGNEDDGSNVASGR				
		SSMALWGLVVCHHTSSRCIPFPLRYA				
		CEFLMQAFGLQLNMELQLALQMSEKR				
		VLRTQTLLCDMLLRDSPAGIVTQSPS				
		IMDLVKCDGAAFLYHGKYYPLGVAPS				
		EVQIKDVVEWLLANHADSTGLSTDSL				
		GDAGYPGAAALGDAVCGMAVAYITKR				
		DFLFWFRSHTAKEIKWGGAKHHPEDK				
		DEGQRMHPRSSFQAFLEVVKSRSQPW				
		ETAEMDAIHSLQLILRDSFKESEAAM				
		NSKVVDGVVQPCRDMAGEQGIDELGA				
		VAREMVRLIETATVPIFAVDAGGCIN				
		GWNAKIAELTGLSVEEAMGKSLVSDL				
		IYKENEATVNKLLSRALRGDEEKNVE				
		VKLKTFSPELQGKAVFVVVNACSSKD				
		YLNNIVGVCFVGQDVTSQKIVMDKFI				
		NIQGDYKAIVHSPNPLIPPIFAADEN				

[0794]

		TCCLEWNMAMEKLTGWSRSEVIGRMI VGEVFGSCCMLKGPDALTKFMIVLHN AIGGQDTDKFPFFFFDRNGKFVQALL TANKRVSLEGKVIGAFCFLQIPS				
적색광	PhyBNT	VSGVGGSGGGRGGGRGGGEEEPSSSHT PNNRRGGEQAQSSGTKSLRPRSNTES MSKAIQQYTVDARLHAVFEQSGESGK SFDYSQSLKTTTYGSSVPEQQITAYL SRIQRGGYIQPFGCMIAVDESSFRII GYSENAREMLGIMPQSVPTLEKPEIL AMGTDVRSLFTSSSSILLERAFVARE ITLLNPVWIHSKNTGKPFYAILHRID VGVVIDLEPARTEDPALSIAGAVQSQ KLAVRAISQLQALPGGDIKLLCDTVV ESVRDLTGYDRVMVYKFHEDEHGEVV AESKRDDLEPYIGLHYPATDIPQASR FLFKQNRVRMIVDCNATPVLVVQDDR LTQSMCLVGSTLRAPHGCHSQYMANM GSIASLAMAVIINGNEDDGSNVASGR SSMRLWGLVVCHHTSSRCIPFPLRYA CEFLMQAFGLQLNMELQLALQMSEKR VLRTQTLLCDMLLRDSPAGIVTQSPS IMDLVKCDGAAFLYHGKYYPLGVAPS EVQIKDVVEWLLANHADSTGLSTDSL GDAGYPGAAALGDAVCGMAVAYITKR DFLFWFRSHTAKEIKWGGAKHHPEDK DDGQRMHPRSSFQAFLEVVKSRSQPW ETAEMDAIHSLQLILRDSFKES	pBW2682	PIF6	MELPTDYCCRLSDQEYMELVFENG QILAKGQRSNVSLHNQRTKSIMDL YEAEYNEDFMKSIIHGGGGAITNL GDTQVVPQSHVAAAHETNMLESNK HVD	pBW2684
적색광	PhyBNT	VSGVGGSGGGRGGGRGGGEEEPSSSHT PNNRRGGEQAQSSGTKSLRPRSNTES MSKAIQQYTVDARLHAVFEQSGESGK SFDYSQSLKTTTYGSSVPEQQITAYL SRIQRGGYIQPFGCMIAVDESSFRII GYSENAREMLGIMPQSVPTLEKPEIL AMGTDVRSLFTSSSSILLERAFVARE	pBW2682	PIF6	MELPTDYCCRLSDQEYMELVFENG QILAKGQRSNVSLHNQRTKSIMDL YEAEYNEDFMKSIIHGGGGAITNL GDTQVVPQSHVAAAHETNMLESNK HVD	pBW2684

[0795]

근적외광
PpsR2
ITLLNPVWIHSKNTGKPFYAILHRID
VGVVIDLEPARTEDPALSIAGAVQSQ
KLAVRAISQLQALPGGDIKLLCDTVV
ESVRDLTGYDRVMVYKFHEDEHGEVV
AESKRDDLEPYIGLHYPATDIPQASR
FLFKQNRVRMIVDCNATPVLVVQDDR
LTQSMCLVGSTLRAPHGCHSQYMANM
GSIASLAMAVIINGNEDDGSNVASGR
SSMRLWGLVVCHHTSSRCIPFPLRYA
CEFLMQAFGLQLNMELQLALQMSEKR
VLRTQTLLCDMLLRDSPAGIVTQSPS
IMDLVKCDGAAFLYHGKYYPLGVAPS
EVQIKDVVEWLLANHADSTGLSTDSL
GDAGYPGAAALGDAVCGMAVAYITKR
DFLFWFRSHTAKEIKWGGAKHHPEDK
DDGQRMHPRSSFQAFLEVVKSRSQFW
ETAEMDAIHSLQLLIRDSFKES
pBW2780
ASKSVHADITLLLIMEGVIREATLSP
TMAAESVDGWLGRRWSDIAGAEGGDK
VRRMVEDARRSGISAFRQINQPFPSG
VEIPIEFTTMLLGDRTGMIAVGKNMQ
AVTELHSRLIAAQQAMERDYWRLREL
ETRYRLVFDAAADAVMIVSAGDMRIV
EANRAAVNAISRVERGNDDLAGRDFL
AEVAAADRDAVRIMLAQVRQRGTALS
VLVHLGRYDRAWMLRGSLMSSERRQV
FLLHFTPVTTTPAIDDVDDDAVLRGL
IDRIPDGFVALDSEGVVRHANQAFLD
LVQIGSKPAAVGRSLGVWMGRPGADL
SSLLTLLRRYKTVRLFQTTIRGELGT
ETEVEVSAVDGEIDQYIGVLMRNTVAR
RLDAADDHDALRQALGPISKQLGRSS
LRKLVKNAVSIVEQHYVKEALLRSRG
NRTATAELLGLSHQSLYAKLNSYGFD
DKGVVASAADGAEGASDDAED
BphP1
pBW2779
VAGHASGSPAFGTADLSNCEREEI
HLAGSIQPHGALLVVSEPDHRIIQ
ASANAAEFLNLGSVLGVPLAEIDG
DLLIKILPHLDPTAEGMPVAVRCR
IGNPSTEYDGLMHRPPEGGLIIEL
ERAGPPIDLSGTLAPALERIRTAG
SLRALCDDTALLFQQCTGYDRVMV
YRFDEQGHGEVFSERHVPGLESYF
GNRYPSSDIPQMARRLYERQRVRV
LVDVSYQPVPLEPRLSPLTGRDLD
MSGCFLRSMSPIHLQYLKNMGVRA
TLVVSLVVGGKLWGLVACHHYLPR
FIHFELRAICELLAEAIATRITAL
ESFAQSSELFVQRLEQRMIEAIT
REGDWRAAIFDTSQSILQPLHADG
CALYYEDQIRTIGDVPSTQDVREI
AGWLDRQPRAAVTSTASLGLDVPE
LAHLTRMASGVVAAPISDHRGEFL



[0796]

[0797] 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드 또는 복합체에 포함된 한 쌍의 이량체화 도메인은 항체, 또는 이의 기능성 단편, 및 항체 또는 이의 단편에 의해 인식되는 펩티드를 포함한다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드 또는 복합체에 포함된 한 쌍의 이량체화 도메인은 표 35의 단일 행에 열거된 바와 같은 사슬 A 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열) 및 사슬 B 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열)을 포함한다.

[0798] **표 35. 예시적인 항체-펩티드 이량체화 도메인**

시스템	사슬 A 명칭	사슬 A 서열	예시적인 사슬 A 공급원	사슬 B 명칭	사슬 B 서열	예시적인 사슬 B 공급원
SunTag	GCN4_v4	EELLSKN YHLENEV ARLKK	snapgene	GCN4_scFv	GPDIVMTQSPSSLSASVGDRVTITCRSSTGAVTTSN YASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSGSLI GDKATLTISSLQPEDFATYFCALWYSNHWVFGQGTK VELKRGGGGSGGGGSGGGGSGGGSEVKLLESGGGL VQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEW IGVIWGDGITDYNSALKDRFIISKDNGKNTVYLQMS KVRSDDTALYYCVTGLFDYWGQGTLVTVSS	addgene 60904
SunTag	GCN4_v4	EELLSKN YHLENEV ARLKK	snapgene	GCN4_scFv	GPDIVMTQSPSSLSASVGDRVTITCRSSTGAVTTSN YASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSGSLI GDKATLTISSLQPEDFATYFCALWYSNHWVFGQGTK VELKRGGGGSGGGGSGGGGSGGGSEVKLLESGGGL VQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEW IGVIWGDGITDYNSALKDRFIISKDNGKNTVYLQMS KVRSDDTALYYCVTGLFDYWGQGTLVTVSS	addgene 60904
SunTag	GCN4_v1	LLPKNYH LENEVAR LKKLVGE R	snapgene	GCN4_scFv	GPDIVMTQSPSSLSASVGDRVTITCRSSTGAVTTSN YASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSGSLI GDKATLTISSLQPEDFATYFCALWYSNHWVFGQGTK VELKRGGGGSGGGGSGGGGSGGGSEVKLLESGGGL VQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEW IGVIWGDGITDYNSALKDRFIISKDNGKNTVYLQMS KVRSDDTALYYCVTGLFDYWGQGTLVTVSS	addgene 60904
SunTag	GCN4_v1	LLPKNYH LENEVAR LKKLVGE R	snapgene	GCN4_scFv	GPDIVMTQSPSSLSASVGDRVTITCRSSTGAVTTSN YASWVQEKPGKLFKGLIGGTNNRAPGVPSRFSGSLI GDKATLTISSLQPEDFATYFCALWYSNHWVFGQGTK VELKRGGGGSGGGGSGGGGSGGGSEVKLLESGGGL VQPGGSLKLSCAVSGFSLTDYGVNWVRQAPGRGLEW IGVIWGDGITDYNSALKDRFIISKDNGKNTVYLQMS KVRSDDTALYYCVTGLFDYWGQGTLVTVSS	addgene 60904
MoonTag	gp41_peptide	KNEQELL ELDKWAS L	addgene 128605	moontag_nanobody	EVQLVESGGGLVQPGGSLRLSCAASGSISSVDVMSW YRQAPGKQRELVAFITDRGRTNYKVSVKGRFTISRD NSKNMVYLQMNSLKPEDTADYLCRAESRTSWSSPSP LDVWGRGTQVTVSS	addgene 128602

[0799]

[0800] 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드 또는 복합체에 포함된 이량체화 도메인은 코일드 코일 이량체화 도메인을 포함한다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드 또는 복합체에 포함된 이량체화 도메인은 표 36의 단일 행에 열거된 바와 같은 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드 또는 복합체에 포함된 한 쌍의 이량체화 도메인은 동일한 코일드 코일 이량체화 도메인(또는 서로에 대해 적어도 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 코일드 코일 이량체화 도메인)의 카피를 포함한다.

[0801] 표 36. 예시적인 코일드 코일 이량체화 도메인

명칭	서열
P1	SPEDEIQALEEENAQLEQENAALEEEIAQLEYG
P2	SPEDKIAQLKEKNAALKEKNQQLKEKIQALKYG
P3	SPEDEIQQLEEEIAQLEQKNAALKEKNQALKYG
P4	SPEDKIAQLKQKIQALKQENQQLEEENAALEYG
P3S	SPEDEIQQLEEEISQLEQKNSQLKEKNQQLKYG
P4S	SPEDKISQLKQKIQQLKQENQQLEEENSQLEYG
P5	SPEDENAALEEKIAQLKQKNAALKEEIQALEYG
P6	SPEDKNAALKEEIQALEEENQALEEKIAQLKYG
P7	SPEDEIQALEEKNAQLKQEIAALEEKNQALKYG
P8	SPEDKIAQLKEENQQLEQKIQALKEENAALEYG
P9	SPEDENQALEQKNAQLKQEIAALEQEIAQLEYG
P10	SPEDKNAQLKEENAALEEKIQQLKEKIQALKYG
P11	SPEDENQALEQEIAQLEQEIAALEQKNAQLKYG
P12	SPEDKNAQLKEKIAALKEKIQQLKEENQALEYG
DHD9	GSPKEEARELIRKQKELIKEQKKLIKEAKQKSDSRDAERIWKRSREINRESKKINKRIKELIKS
DHD9	PKKEAEELAEESEELHDRSEKLHERAEQSSNSEEARKILEDIERISERIEEISDRIERLLRS
DHD13_XAAA	GTKEDILERQRKIIERAQEIHRRQQEILEELERIIRKPGSSEEAMKRMLKLLEESLRLLKELLELSEESAQLLYEQR
DHD13_XAAA	GTEKRLLEEAERAHREQKEIIKKAQELHRRLEEIVRQSGSSEEAKKEAKKILEEIRELSKRSLELLREILYLSQEQK GSLVPR
DHD13_XAXA	TKEDILERQRKIIERAQEIIRRQQEILEELERIIRKPGSSEEAMKRMLKLLEESLRLLKELLELLEESAQLLYEQR
DHD13_XAXA	GSTEKRLLEEAERAHREAKEIIKKAQELHRRLEEIVRQSGSSEEAKKEAKKILEEIRELSKRLLELLREILYLSQEQ K
DHD13_XAAX	TKEDILERARKIIERAQEIHRRQQEILEELERIIRKPGSSEEAMKRMLKLLEESLRLLKELLELSEELAQLLYEQR
DHD13_XAAX	GSTEKRLLEEAERAIREQKEIIKKAQELHRRLEEIVRQSGSSEEAKKEAKKILEEIRELSKRSLELLREILYLLQEQ K
DHD13_2:341	TKEDILERQRKIIERAQEIHRRQQEILEELEYIIR

[0802]

[0803]

DHD13_2:341	MSEEAMKRMLKLLEESLRLLKELLELSEESAQLLYEQRKANNGSETEKRLLEEAERAHREQKEIIKKAQELHRRLEE IVRQSGSSEEAKKEAKKILEEIRELSKRSLELLREILYLSQEQK
DHD13_AAAA	MTKEDILERQRKIIERAQEIHRRQQEILKEQEKIIRKPGSSEEAMKRSLKLIEESLRLLKELLELSEESAQLLYEQR
DHD13_AAAA	GTEKRLLEEAERAHREQKEIIKKAQELHKELTKIHQQSGSSEEAKKRALKISQEIRELSKRSLELLREILYLSQEQK
DHD13_BAAA	TKEDILERQRKIIERAQEIHRRQQEILKRSEEIIRKPGSSEEALETLRELQEESLRLLKELLELSEESAQLLYEQR
DHD13_BAAA	GSTEKRLLEEAERAHREQKEIIKKAQELHRRTEEIIRQSGSSEEAKDELRRIQEEIRELSKRSLELLREILYLSQEQ K
DHD13_4:123	TTKRYLEEAERAHREQKEIIKKAQELHRRLEEIVRQ
DHD13_4:123	GSSEEAKKEAKKILEEIRELSKRSLELLREILYLSQQVNDVDEKALERQRKIIERAQEIHRRQQEILEELERIIRKP GSSEEAMKRMLKLLEESLRLLKELLELSEESAQLLYEAR
DHD13_1:234	EAMKRMLKLIEESIRLILKELLELSEESAQLLYEAR
DHD13_1:234	TTKRYLEEAERAHREQKEIIKKAQELHRRLEEIVRQSGSSEEAKKEAKKILEEIRELSKRSLELLREILYLSQQVND VDEKALERQRKIIERAQEIHRRQQEILEELERIIRKPGS
DHD15	TREELLRENIELAKEHIEIMREILELLQKMEELLEKARGADEDVAKTIKELLRRLKEIIERNQRIAKEHEYIARERS
DHD15	GTERKLIERSRRLQEESKRLLDEMAEIMRRIKKLLKKARGADEKVLDEIRKIIERIRELLDRSRKIHERSEEIAYKE E
DHD20	GDRQELIRRNIELLKEHIKILEEISQLIEELSELLDKSSSEEVVKRYKKILERYKQLLRKSQEIHKESSEIAKKES
DHD20	GDEQKLIERSQRMQKESLELLKEIIKILDTIEKLLDKPDSEELLDTIKKLHDTLKKIHDRNKKLLKEHEEILRQRSG SLVPR
DHD21	DKEEEYKRLLDEIKEILKESKEVLKDSKRVLEDIKRKVPDDDLVKLLEKHVRLLEEHVKLLEQLIREAEKSSK
DHD21	QGSSAEELLKKIKESEKKIRDSLRKIKEIIKKSRKEGVDDKQLDLIRKVVESHRDLLRLHRDLLRLLREETS
DHD25	DIDESIKEVEKLLEEVEQSLQKLDDSLKKLLEKVNQDPDVDDSVRKIVKRHVEILKRHEEVLKRLIEVVKEHTKTVK
DHD25	GSDREEVHKEIVKLIREIIKIHKKILKIHEKIKNGEIDPSEILKLSEEIKKLTDTIIKIEDLEQLTRDLRR
DHD27	DRKEIVKRHQKVVELLKESSKLLRESSKLLQRLLDKTGDENLQKAVDDQDKAIKRQETAIRKSQEASKKLD
DHD27	DNSEEIKKVAKTSREVAEYSERVAKENDKVVKTLEEGKIDESELLRLLEESIKIFDTALKLHEEAYKLHQDLVRKVS
DHD30	DESEAASVAIESVQILVESVKLLEESVRILLDAVKKNGVEDLLRVAQRWEKLVDEWLKVVKRWLDNVRDIQR
DHD30	GSDKAEEVEKSVRKIEESIKKIRKSIKKAEDAVQLLKEGKIDAKDFLRIVREDLEVVKEDVEIVKEDVENVREFSS
DHD33	SDKEVSDKLLKASKKLLKVSEELLEVVRRLLKALKDDELIKKIADLLRKIIDKDKKFIRTSEEIVKESR
DHD33	GSDLKEVLKTVEEAVKEIIKSSEELLQISRKILEISRVGVDEHEYISAIREYLKALEKHIQILKKFIEILKELIRAV S

[0804]

DHD34_XAAX A	SKEEIDKIVKKHKKKKIEEHKKKVDELKKLVEEHDKRVSQDKDDKVKKLSEEVKKIIKRLEEVSKKLLKVI SDKR
DHD34_XAAX A	GSNDEELKKIILETLDRILKKLDKILTRLIEVLKKSEDPNLDDKDYTELVKQFIELIKKYEEVVKEYEEVVRQLIRLF S
DHD34_XAXX A	SKEEIDKIVKKHKKKKIEELKKLVDELKKLVEEHDKRVSQDKDDKVKKLSEEVKKIIKRVEEVAKRLEEVSKKLLKVI SDKR
DHD34_XAXX A	GSNDEELKKIILETLDRILKKLEKILTRLIEVLKKSEDPNLDDKDYTELVKQFIELIKKFEEVIKEYEEVVRQLIRLF S
DHD34_XAAA A	SKEEIDKIVKKHKKKKIEEHKKKVDEHKKLVEEHDKRVSQDKDDKVKKLSEELKKISKRLEEVSKKLLKVI SDKR
DHD34_XAAA A	GSNDEELKKIILETLDRILKKLDKILTRLDEVLKKSEDPNLDDKDYTELVKQYIELVKKYEEVVKEYEEVVRQLIRLF S
DHD36	DHSRKLKEILDRLRKHVKRLKEHLDELRDLVRQVPEDKLLEHVVKLSDKILQISERAVREFTKSVDKDS
DHD36	GSDKKDELERILDEIRRLIERLDEILSRLNKLLELLKHGVPNAKEVVKDYIRLLKEYLELVKEFLKLVKRHADLVS
DHD37_ABXB	DSDEHLKKLKTFLENLRRHLDRLDKHIKQLRDILSENPEDERVKDVIDLSERSVRIVKTVIKIFEDSVRKKE
DHD37_ABXB	GSDDKELDKLLDTLEKILQTATKIIDDANKLLEKLRRSERKDPKVVETYVELLKRHEKAVKELLEIAKTHAKKVE
DHD37_BBBB	MDEEDHLKKLKTHLEKLERHLKLLEDHAKKLEDILKERPEDSAVKESIDELRRSIEILVRESIEIFRQSVEEEE
DHD37_BBBB	GDVKELTKILDTLTKILETATKVIKDATKLLEEHRKSDKPDPRLIETHKKLVEEHETLVRQHKEIAEEHLKRTR
DHD37_XBXB	DSDEHLKKLKTFLENLRRHLDRLDKLLKELRDILSENPEDERVKDVIDELERVIRIVKTVIKIFEDSVRKKE
DHD37_XBXB	GSDDKELDKLLDTLEKILQTATKIIDDLNKVLEKLRRSERKDPKVIETVVELLKRHEKAVKELLEIAKTHAKKVE
DHD37_AXXB	DSDEHLKKLKTFLENLRKLEDLLDKHIKQLRDILSENPEDERVKDVIDLSERVVRTVKTVIKIFEDSVRKKE
DHD37_AXXB	GSDDKELDKLLDTLEKILQTATKVVDDANKLIEKLRRSERKDPKVVETYVELLKRLEKLIKELLEIAKTHAKKVE
DHD37_3:124	DSDEHLKKLKTFLENLRRHLDRLDKHIKQLRDILSEN
DHD37_3:124	EDERVKDVIDLSERSVRIVKTVIKIFEDSVRKLEKTKPDSKTAKELDKLLDTLEKILQTATKIIDDANKLIEKLRRS ERKDPKVVETYVELLKRHEKAVKELLEIAKTHAKKVE
DHD37_1:234	DSDEHLYKLKTFLENLRHLDRLDKHIKQLRDILSENPEDERVKDAIDLSERSVRIVKTVIKIFEDSVRKKEKRPID KRDDKELDKLLDTLEKILQTATKIIDDANKLLEYLRR
DHD37_1:234	GDPKVVETYVELLKRHEKAVKELLEIAKTHAKKVE
DHD37_AXBB	DSDEHLDRLDKHLKKLKTFLENLRRHIKQLRDILSENPEDERVKDVIDLSKTVIKIFEDSVRKKERSVRIVE
DHD37_AXBB	GSDDKEATKIIDDLDKLLDTLEKILQTANKLLEKLRRSERKDPKVVETYVKAVKELLEIAKTHAELLKRHEKKVE

[0805]

DHD37_X88A	DSDEHIKQLRDHLDRLDKHLKKLKTFLENLRRILSENPEDERVKTVIKIFEDSVRKKERSVRIVKDVIDLSE
DHD37_X88A	GSDDKEANKLLEKATKIIDDLDKLLDTLEKILQTLRRSERKDPKAVKELLEIAKTHAELLKRHEKVVETYVKKVE
DHD39	DHSRKLEEIILDRLRKHVKRLLEHLRELLSLVKENPEDKDLVEVLELSLAILRRSLEAVEAFLKSVTKKDPDDEDLRR KADEIRKEVEEIKKSLAEVEKEIYKLK
DHD39	GSSADDVLEDILKIIRELIEILDQILSLLNQLLKLLRHGVPNAKKVVEKYKEILELYLQLVSLFLKIVKTHADAVSG KIDKKAEEEIKKEEEKIKEKLRQAKDILKKLQEEIDKTR
DHD40	DRDAHLYKLLTFLEQLVRHLDRLVKHITQLRDIVKKDPEDERAVDVIRQSVRSLEIVITVLKIFVDSVSDAARSKEA EKIVRKIRKEIDEIRQKLREIDKEVKKTTS
DHD40	GSNDKVLDKILDILDRILRLATRVIDLANKLLQVKKKSTHKDPRIVETYKELLKIHETAVRLLLELADLHRRLKSKD EEANKRVETELDRIRKKVKDIEDKVRKLEDKVRKTAS
DHD43	NDLSKEVLKKLEKSVEELLRRVQKSVKEAQKRGLLSDELVDRHLKILNQLVKRHLELLQEVIKRSDKK
DHD43	GSDEAVKRVVEKSLKILDEVIKKSLDILRELIELQIRHAKDDESVIRASKSALKDAIEALKKSLDEIKKALKRSADE G
DHD65	SSEEVVKVHEKVVKLHKEILELLKKIIKIHETAARDPDDKDSIKKLSDEIKKIVKRIEDISDQAKRESSDAQRKQS
DHD65	DKEEESKELLKKLKEILKRSEELLEESKELLKLAKNGEIDESELADADRKLNKKHEKLVQDIQDLLREHERQDR
DHD70	DEKKKIDKIVKETEDLLQKSEKLLQQSKEAVKRIRSQVKENEIVDRLLRISEELLKISRRLVEISRRIASTLS
DHD70	GSSKEEVIRLLKENVRLIKENLELLTRNLKLITDLVRGSNGSEEKIKTLKELLKEYRELLKRYRKLVEDYKRLVDKH D
DHD88	EIQELIKSSRRIIEESKELIKESEEVLRRIKEILDRIRNGVDNQEDLLREILKLLTKNLKIIQRNLKLLQDNAEILK RLVS
DHD88	GSYIEDVIKKILDVSRELIKLSRTIIKISEEINKQLQQGRDTKDLVKKYDEIIKKYTRIVQHYTELIKELQKLLS
DHD89	SPTEEAIQLSQRVIELSKRVIELSKEILKLLKRVLDLLPDLDKNEEKRLDDYDKELKEYDKELKKYEKRLKDLAS
DHD89	GSEEEEILKIQKELLRIQSEILDKQKKILDTLRSNGAVTEEVRSILEKVERLSEEAKELSKEAKELTKEVSKLIS
DHD90	SPLKELNNQLLRLLRELVKVSKKIVDLSKTIIEVLKHTDLDPRLLDSLEKSQQELDKSQKELDKVVKELTKVNKKLQ
DHD90	GSPLEDLVRKYDELVKTYEKLVEEFKKAVDKYDKAVKKAPVSKEATDSLDLIRKVLELLDRNLKLIKENAKLIKELL K
DHD91	SPTRENEKVIKENEKVISDNERVLEEVVKVVETATDRKEIQDAVDEVRKSVDKLRDSVRKLEESVRTLD
DHD91	GSPIKDISKRLLEISKRLVEISDRIVELLQRIADSKDPNKDLQKEVKDVLEEYKRLVREYREVVKEYEKVVS
DHD92	DEDEHVKQLIKNADLLRKHAELLKELVKLFQEIASQIPDDRVAKKVTDVVDRIDKILKQTEKLVRRTKQILDYSR
DHD92	GSNLEELVKLLKEVLEMHERLLRIHEDLVEAHKSNASDKESERKLKKSDKDIKESLKKIKSIIDQVRYIQS
DHD93	PVEDIIEESLRLLEESLKLLNRILKLLEDSLRKLPRSEEWRQRLDEFRKKLEDWKEELERWIEDVRYKKT

[0806]

DHD93	GSDEDYESREIIDEIRKLLDRSKKIVHRSQRLVERVKSTPLSEDQEDLIRRHEETINRHRELVKELEKVLEDHERHIR
DHD94	PEEDSRRVLERFVRVSREVLKVLEEFLRVSEELLREADRDRDRRLEEYERQVDELREEIRRYKEEVDKFDKEVKYYKK
DHD94	GSPEKDENRKLLDKVRKLVEKSRRLVEELRKLVDQSTKNGLIDEKALRKQQEVLRKVEEVLEKQERVLREIEEISYRVI
DHD94_3:214	GSPERDENRKLLDKVRKLVEKSRRLVEELRKLVDQSTKN
DHD94_3:214	GSDEKALRKQQEVLRKVEEVLEKQERVLRELEEISYRVITRGEDHKAEEDSRRVLERFVRVSREVLKVLEEFLRVSEELLREADRDRDRRLEEYERQVDELREEIRRYKEEVDKFDKEVKYYKK
DHD94_2:143	GSDRRLEEYERQVDELREEIRRYKEEVDKFDKEVKYYKK
DHD94_2:143	GSPERDENRKLLDKVRKLVEKSRRLVEELRKLVDQSTKNGLIDEKALRKQQEVLRKVEEVLEKQERVLRELEEISYRVITRGEDHKAEEDSRRVLERFVRVSREVLKVLEEFLRVSEELLREADR
DHD95	DLSEESKKFVEKVKKLEKESRELEKQVKKIEEDSRSVENDVQKEFLELLKRLLDIQKKVVEVLREVVKVQQYVDS
DHD95	GSDSEYESRQVLRELDTVLKDSHTVLEALRQVIRDSQDVVSKSDEESRRVIDDLEKVIQDSKKVLDDIKRLIDKSKSIKS
DHD96	NEDELLKLLTENLKLLDENLKLLRENLSLLRQANNITDKNRIREIVKQSKEIVKQSREILKQSKEIVERIKYIVS
DHD96	GSSLYELTQRYEKLVQQYEELVKDYRRLVKKLEKLKRDNKPDKRLLKEIVDVIKKSVEIIDRSLKLLEESIKILEETD
DHD97	SQERSLEILKRILDVLKESLEILKESLSILRQLASRIKNPNRKIEEILKESDKIIKESDKVLKEIEEVIRYSS
DHD97	GSDIEYESKEILELIKELLKLSRELLKESRRALELVRKSRDDSIVEEVIQVHKKVLDIHKEVLKIVRKVVEVHRRVKS
DHD98	SKKDESTKLERLAEKIDEITKRIEELVKDVKRKSSEGVDKDQQQKIDEVFQKLLDLQREILEILDRILKVQQYILD
DHD98	GSDLEYLNRRLLQLIKTLIDLNRHLLKLIDKLKKLNSREGDEEKIKEESKQIQEQFKEIVERSKEIIKQIKEIIKRSQ
DHD99	DFERSSRRLEKVVEDLRRSSDRLREVIDELRKSADEKDEDEDLRRARKEHRDLIEELKRALEKQEEIIKHLQELVYRQL
DHD99	GSEESEEVRKVVERIKKISRELEEVVKELDRVSKEFDRHGETDEIVREHERIVEKLEEIVKKHTKIVEELAEIVYKQQ
DHD100	SDDDSVRVLDEIVKILDESVKLLKESLKLLDDFLRTKPDDHLKEVVKESKKVVEQSKKVLDRIKKIIYESK
DHD100	GSDLLYLSKELLKLVRELLKLSRELVELSRRLVNSTHKSPELVKKYDKLVKKYQDLLKKLADVADEYLRQRS
DHD101	DEKDYHRRLIEHLEDLVRRHEELIKRQKKVVEELERRGLDERLRRVVDRFRRSSERWEEVIERFRQVVDKLRKSVE

[0807]

DHD101	GSDAYDLDRIVKEHRRLVEEQREIVEELEKLVRRQEDHRVDKKESHEILERLERIIRRSTRILTELEKLTDEFERRT R
DHD102	DERYRAREHIRRVEEHTKRLRHILKRLREHEEKLRRELKPGDEITESVDRFKKIVDQFEESIKKFETVSEELRKSDS
DHD102	GSDRQRILDRLDKILEKLDDILKKLKDILETLSKDDVSDRRHKDLVEKFRELVDTHHKLVERYRELVYQNR
DHD102_1:243	GSDEITESVDRFKKIVDQFEESIKKFETVSEELRKSIS
DHD102_1:243	GSDPQRAADRLDKILEKLDDILKKLKDILETLSKDDVKDRRAKDLVEKFRELVDTHHKLVERYRELVYTATAGSDLA RELIRRVEEHTKRLRHILKRLREHEEKLRR
DHD103	NADDQLATSIKKLEDSIDQLIKIVRKFEESVKKLQKHGVDQHHVEILRKIVEIFRQHIEKLKKHLEKLRYTSS
DHD103	GSDKEYLVTEHEKLVREHEKIVSEIEKLVKKHEAGVDESELEEILKKVEKLLRKLDEILEQLTQLLRKTE
DHD103_1:423	GSDQHVVEILRKIVEIFRQHIEKLKKHLEKLRYTSS
DHD103_1:423	GSDAEYLVTEHEKLVREHEKIVSEIEKLVKKHEKGVDESELEEILKKVEKLLRKLDEILEQLTQLLRKAEKHIDKHS KAADQLATSIKKLEDSIDQLIKIVRKFEESVKKLQKH
DHD104	DEDDDIRRVLDESRRVLEHSRRVLKRSEEVLEKASRKKEKDTEEIEKHLKRLREHAKKLEKHRRELDDFLYKEI
DHD104	GSRDKYLLERLNDILKKLDEIVDKLSDILKRLKDVRHDDRLQELVERYKEIVKEYKRIVEEYEKLVREFEEQQR
DHD105	DRDYEDKEFKKIIKELEDVQEELKKLQEKIKRFSSELEEPNELLKEQLKVNEEQLEVNKKILKILRDQLKQNE
DHD105	GSDAEYKVRESVKRSKESVKHSEDVVDKLNKSVKLSESGHSDAEKASRELVKLVRSEVVELSREVIKLSEKVLRVIS
DHD106	DLQYKQEKLIRHFDRVVREWDKLVRKFSKVLEKQKHESKDKELEEASRRVDELIKRLREQLKRSKEILRRLKELSRK SS
DHD106	GSDWEELLRRLEKVLQEYEEIVKELIDLIERLIKVSEDKSKDASEYKKLVTELEKLISKLEELVKEYEYK TE
DHD107	DAKDELEKSLQEIEESLKELKKLLEELDKSLRELTSQGRNKKLEEHIKKVQKFIELVKKYIKAVQDYLKEVRYDNS
DHD107	GSDKERAARATEEMVKLTKKLLKAVEDLVRDVRRLLKEGLISEKHARIAETILEVFKKHAKIIKKHVDIVKYDES
DHD108	GSPLKERLLEIQRDLDRVLEEVVERLLRIQERLDSVVERKPPDVHEEYKYIVDEIREIVERVVREYEEIVKRIDEEV R
DHD108	GSEEDERIRYDLDRIRKDVRRKLEEIRQRVREIEKKLRDAGHRRDEKELLRELIETSKDILRLVEELLKKIIDKSED LLRKTE
DHD109	GSDEEDYINENVEKDVRDIEDDVRRINERIRELLEKIRTEEVLQRVLEEHHELVERVLRKLVEILRKHEEENR
DHD109	GSDEEEYYKEKLHKLLREIEELLKHYRELVRRLEELVKRGELDKDTAAHILERLSELLERIIRRVAHTLRRLSEERR

[0808]

DHD110	GSDEDEISYDSKRRVEEIVRQARE KSEKSRKDIEDVAEVLRKGDVSEKEVVDELVKVLEEQVKVLREAVERLREVLK KQVDDVR
DHD110	GSDIVELVDHLLKRSLKLLEELAELVRRLLEKSTELLKRRTEEHKEEVVEESEYMVRELEERLRRVVDESEKLVRDA DKHIR
DHD111	GSKEKDIVKTLVDLLRENLETLERLIEEVVRLLKENVDVRDEGRDDKDSERILRDIKRRIDEAAKESREIIERIEKE VEYRSR
DHD111	GSPEVDVLRRIVREILKASEELLRLLLRKLIDEALKLSERKRDSQEYREVVDRVKKELERLLDEYRKLVEELKEKLRY DTR
DHD112	GSDKRYESEKLKRRLDEAVEKVREVVERVERESDRVLEEVRRRRESKEVVDKVIEDNDKALEDVLRVVDEVAKVVRD VVRENTR
DHD112	GSPREYHSKDILRKVDEILERIRRHADRVKKKSERLKRENVDVNEHSKDVKRVIRELLELVKELLRLAKKHSDDQQE
DHD113	GSDEDEILYHSERLLQKLKKELDDLKEKSRELLEELKKEDPDDRLIERIIRLHDEVLKDLDEVLKNILEVHREVLER LR
DHD113	DKLDRLLKIHEEALRRAEELIKRLLDIHRRALDLARRGELDDYLLKESERELREIIRRAREELKESRDRLEEISR
DHD114	GSPKEELIRRVLEEVKRLNEKILEIIRRAAELVKRANDELPETEKLREIDRELEKKLKEIEDELRRIDKELDDALYE IED
DHD114	GSPKLDKLRELLERNLEKLREILEEVLKILRTNLERVREDIRDEDVLQEYERLIRKAEEDLRRVLKEYDDLLKKLVY ELR
DHD115	GSKEDESVKRAEEIVRTLLKLLEDSLREAERSLRDIKNGEDEHNLRRISEKLEELSKRITETETERLLREIQYTSR
DHD115	GSPNQELLDRVRKILEDLLRLNEELVRLNKELLKRALEMRRKNRDSEEVLERLAEEYRKRLEEYRRELEKLLEELEE TIYRYKR
DHD116	GSDESEEAQHEVEKVLDDIRRLSEHLQKRLEEVLEEVYELRREGSDRTEVVELLKEVIREIVRVNREALERLLRVVE EAVKRNE
DHD116	GSDEELVETVKRIQKEILDRLTELAKLLVEIQREIKKLKDEGEDDKELKRLSDELEEKVRQVVEEIKRLSDELEET VEYVSR
DHD117	GSDEEEEVVRRAEELVKEHEELIERVIRTHEELVYKLEDQGADKKLVDVLKRVVEESERVAREIVKVSRELIRLIEE ASR
DHD117	GSSKEEILKELEDLQRRLIEELKKLQERVVELLEELIKRLRDRGRDDKHLKRLVKEVRRLSEEVLRSIKEVSDRVRY QLR
DHD118	GSDKEEESEYLLRDLVRLLEKVKEKIEEVNREVEKLLKKVKDGRLDRREVLREILRLNRELAEIIKEVVDRIRHVVE RSER
DHD118	GSDLHEVVYETKELLKRIEEVVEELRKKSEDIIRKAERGEISEDELKRLQEEIAREAKKLLDEIKRVLERHLEQTL

[0809]

DHD119	GSPVEEIIKEVVKRVIEVQEKVLRIISHAVKRVVEVQKKYDPGSEESNRVVEEVKKTIEDAIRESDEVVDEVVKRIQ YTVR
DHD119	GSPEQEIADRILTEIRESQKELERLARKILKLLDESQEKAKRGRLSEEESDELIERIKKELDELLERSKELLKKIEY ELR
DHD120	GSDEDKEANRVLDEVLKTVRDLLETANEVLKEVLYRLKRTDDQEKVVRTLTEVLKEHLKLVEEIVRILDKVLKEHLE TEK
DHD120	GSPEDDVLRRLEEVSEKILRVAEDVARQLREVSEKITQGKVDRKEWEEDIKRLKRELEELLREWKEEIERLTYELR
DHD121	GSRREEVVKRIRELLKRNKELIDRIRELLEENEYLDKDARDKDVLRRSVELLEELVRILEESVELAKEIIKLLREVV E
DHD121	GSDEKEDNRRLQHKIERILEKNEDLQRKLEEILELLERGEADEEKIDRLRKAVEDYRRVVEEIKEDVKRHKYTVR
DHD122	GSDEKEEAKKASEESVRTVERILEELLKASEESVELLRRGEDAKDVVERSKEAIKRVKELLDEVVKRSDEILKYIHN
DHD122	GSDEKKLINEVVETQKRLIKEAAKRLSEVVRHQTELIRELREKNVDDKDVEKLIKESLDLAEEIVRRIKELLDESKK LVEYVSN
DHD123	GSPDMDEVKRVLDELIEIQEEILREIKRVLEKLIKIQEDNGSEYESREVVREIVEIARKLVERSRRVVKKITETLQ
DHD123	GSDERYATREIVERIERIAREILKRTEEIVREVREVLSRDVDQEEVVRRLADLLRESVELVQHLVRRVEELLQESVE RKK
DHD124	GSPEREALREVLEDLKRVTDRLREIVERVLEELKKVTDHVDSERILRESRRVLKELKDIIEEILRESEKVLEKLKYT ED
DHD124	GSPAREILEEVVKKHLEVVEDAARILEEIIREHEKAVREDRDKKELEEISRDLLRKAREALKKVKDISDDLSREIEY VAS
DHD125	GSPVEEAIKKVIDDLRDVQRKIRELVEELIRLLEEVQRDNDKRESEYVVERVEEILRRITETSREVVRKAVEDLS
DHD125	GSDSDEKAEYLLKEMERVVRESDEVVKKILRDLEEVLERLRRGEISEDDVTEILKELAERHIRAIEELVRRLRELLE RHKR
DHD126	GSPVEEVLKELSEVNERVRDIAREIIERLSEVNEEVKETDDEDELKKISKKVVDEVEDLLRKILEVSEEVVRRVEYH DR
DHD126	GSPKEDILREVLRRHKEIVREIVRLVREAVETHLELVKRNSDDRDAQDVIRKLEEDLERLVRHAQEVIEEIFYRLH
DHD127	GSPRSYLLKELADLSQHLVRLLERLVRESERVVEVLERGEVDEEELKRLEDLHRELEKAVREVRETHREIRERSR
DHD127	GSDREYIIKDILDSQEHLLRLIEELLETQKELLEILKRRPDSVERVRELVRRSKEIADEIRRQSDRNVRLLEEVSK
DHD128	GSDEKDEIRHVIESVERLIEDIKRLLKTLRELAHDDSDKKTVKEVLDRVKEMIERHRRELEEHRKELERAEYEVR
DHD128	GSESEDRIKELLKRHIELVERHEELLHEIKKLIDLEEKDDKDREEAVKRIDDAIKESSEEMLEESKEILEEIEYLNR
DHD129	GSSLEDSVRLNDEVVKVVERVVRLNQEVVRLIKHATDVEDEETVKYVLERVREVLDESREVLKRVHELLEESERRLE

[0810]

DHD129	GSHEKDIVYKVEDLVRKSDRIAERAREIVKRSRDIMREIRKDKDNKKLSDDLLKVTRDLQRVVDELEELSRELLRVA EESRK
DHD130	GSPELDEVKKLIDELKKSVERLEESIREVKESIKKLRKGDIDAEENIKLLKENIKIVRENIKIIKEIIDVVQYVLR
DHD130	GSDEEEIEELLRELEKLLKKSEEALEESKKLIDESEELLRRDRLDKEKHVRASEEHVKLSEEHLRISREIVKILEKA VYSTR
DHD131	GSDESDRIRKIVEESDEIVKESRKLAERARELIKESEDKRVSEERNERLLEELIRILDENAELLKRNLELLKEVLYR TR
DHD131	GSDEDDELERLLREYHRVLREYEKLLEELRRLYEEYKRGEVSEEESDRILREIKEILDKSERLWDLSEEVWRTLLYQ AE
DHD132	GSDKKDASRRAIRVLHEFVRVSEEVLEVLRKSVESLKRLDVDEKIKRTHDRIEEELRRWKRELEELIERLREWEYHQ D
DHD132	GSDDEEEDKRLLEEVKRSLDTDERILEKLRHSLERQLEDVDKDEDSRRVLRELDEITKRSREVVKRLRKLAYESK
DHD133	GSDKEYKLDRILRRLDELIKQLSRILEEIERLVDELEREPLDDKEVQDVIERIVELIDEHLELLKEYIKLLEEYIKT TK
DHD133	GSPSKEYQEKSAERQKELLHEYEKLVRHLREIVEKLQRREIDKEEVLRRLVEIIERLKDLHKKIEDAHRKNEEAHKE NK
DHD134	GSRDRKISEELIKALEDHIRMLEELIRAIEEHIKLAERGVDEKELRESLEELKKIVDELEKSLEELRKLAERYKYET R
DHD134	GSPKEESVEELKRVIDKHEEILRELKRVLEEHERVSHDEDENELRRSLERLKHILDRLHESLKELHELLKKNEYTER
DHD135	GSDHEYWVKIVERILRVMEKHAEIVKKHLEIVERVVREGPSEDLRRKLKESLRETEESLRELKELLDELDELSEKTR
DHD135	GSDEEYVTRSQRRLKRLLEEYIKVVEEHARLVERNERDDKELKRSIDELDKLTKELLELVKRYKELVDKTET
DHD136	GSDKEEIVKLQDEVIKTLERHLDILRKHIDLLEKLKDHLSEELKERVDRSIKKLEESIKRLERIIEELQELAEYSL
DHD136	GSREEELKESAEELERSVRELKKEADKYKEEVDRLHYRGKVDKDWVRVVEKLIKLVEEHLELIREHLELLKEERR
DHD137	GSDMEYELKKSAEELRKSLEELKRILDELHKSLRELRRHGDDEEYVQTVEELRKEIEEHAKKLEEHLKELERVAT
DHD137	PEYELKKSVDDLKRDVDRLVEEVEEVFELSKERLREDRKHLELVEEMVRLIEKHLELIKEHLKLADDHVR
DHD138	GSREKDESKELNDEYKKLLEEYERLLRRSEELVKRAKGPRDEKELKRILEENEDILRRTKEILERTKEISEEQKYRR R
DHD138	GSDKDERQERLNEESDKSNEESERSNRESEELNRRARGPNDEKELQETLDRHLELLERNQRLLDENKEILRESQYLN D
DHD139	GSENKYILKEILKLLRENLKLLHDILRLLDENLEELEKHGAKDLDDYRRKIEEIRKKVEDYREKIEEIEKKVERDR
DHD139	GSESEYTQEEILELLKESIKLLRELRLLEESEELWRRENTKSERSEEIKERAKEAIKRSEEILERVKRLSDHSR

[0811]

DHD140	GSDEEEANYVSDKAVKIAEDVQELLKELLELSEVVRRGEVDEDEYDRVLRKLQEVMKEYEEVLKEYEEVSRKHE
DHD140	GSPEKYLIKTQEELLRRHAEILEDLIRKVERQVDLRRKVDERDEDLKRELERSLRELERLVRESSRLVEEIRELSKE IKR
DHD141	GSDEEYELERISRESKELLERYKRLLREYQELLKELRHVKDLDRAVKIIHELMRVSKELVEISHRLLELHERLVRRR K
DHD141	GSEKEYIEKLSRKIEEDIRRSEERAKDSERLVRRLEELAKRKRLDLDDVLRVAEENLEILEDNLRILEEILKEQDKS NR
DHD142	GSPHEEVVELHERVMEISERAVELIQRIIDIIRRIREDDKDIEKLVKTIRDLVREYEELHRELEEIDEEIYKKSE
DHD142	GSDHEDVVRLHEDLVRKQEDARRVLEEIVRLAEEIVEVIKKDEKDKDRVTRLVEEIEKLVEEYKKKVDEMRKISDEI KYRSR
DHD143	GSRAREVVKRAKRIIEEWQKILEEWRRILEDERVDDRDNERIIRENERVIRENEKIIRDVIRLLEELLYE RR
DHD143	GSREDEELEEEIDRIRQMVEEYEELVKEYEELTEKYKQGKVDKEESKKIIEKSERLLDLSQDAVRKVKEIIRRILYT NR
DHD144	GSPKEEIVKLHDESAELHRRSVEVADEILKMHERSKDVDDERESRELSKEIERLIREVEEVSKRIKRLSEEVEYLVR
DHD144	GSPLEEILKIQRRINKIQDDINKILHEILRMQEKLNRSSDKDEVEESLRRIRELIKRIKDLSKEIEDLSREVKYRTT
DHD145	GSPEDEHVYVVREIYEVLREHAEVLEENREVIERLLEAKKRGDKSEELVKELKKSIDKLKEISRKLEEIVKELEKVS EKLK
DHD145	GSDEDETSYRILELLREIVRASRELIRLSEELLEVARRDDKDETVLETLIREYKELLDRYRRLIEELTRLVEEYEER SR
DHD146	GSTQEEINRIQHEVLRIQEEIDEILRDIVEKLKAISRGELDHEVVKDVEDKVREALEKSEELLDKSRKVEYKSE
DHD146	GSDEEELNRELLEKSKRLVDINRDIIRTAQELIEMLKDSKDGRVDEDTKRELRDKLRKLEEKLERVREELRKYEELL RYVQR
DHD147	GSDEKDRVYEILKEVQRLVKEYRDISKEIEDLVKHYEHITDDEAQEVSKELIDKSLRASEIVRELIRLIKELLDELE
DHD147	GSDEEDVLYHLRELLEELKRVSDDYERLVREIKETSERKDRDTKENKDMLDELVKAHREQEKLLERLVRLLEELFER KR
DHD1	PREQAIRISEEIIRISKKIEILERTRSSTAREAMKWAKDSIRLAEESKYLLDK
DHD1	IEDDVKKIQDSTKKAQKETIEALERSTSSTARKQMEEQKEQIRLQKEAMYLLKK
DHD2	SREEIAKLQEEVIKLQRRVIELQKEVIELQRRAKELTSSYTKEILEIQRRIEEIQREIEEIQKRIEEIQEEIQRRT
DHD2	SDEEIKRLSEEVIQLSRRVIKMSREAIKLSREVQKLTPSYQKRIKEIADRSIELARESIEIAKRSEKIAEESQRRT
DHD3	PAKDEALKMANESLELAKKSARLIQESSSKEILERIEKIQRRIAELQDRIAYLIKK

[0812]

DHD3	PAKDEALRMIDESRELIKKSNELIQRSSSKEILERILEIQRKIAELQKRIQYLLKS
DHD4	TDEARYRSERIVKEAKRLLDEARRRSEKIVREAKQRSNSEDAKRIMEENLRESEEAARRLREIIRRNLEESRETG
DHD4	TREALEYQRKMAEEIEDLLREALRRQEEMVREAKQRSLSEEFKRIMERILEEQERVMRLAKEALERILEEQKRTG
DHD5	SERTKREAKRSQEEILREAKEAMRRAKESQDHRQNRDGSNSEDLERLSQEQKRELEEVERRLKELAREQKYKLEDS
DHD5	SEDLKRILKEITERELKLMQDLMEILKKITEDENNLDSNNSEDLKRSIEKARRILDEALRKLEESARRAKYIQEDN
DHD6	TEDEIRESLKWLDEVLQELREIARESNEVLERNRQKSRSDKLREDIERYKKRMEEARKKLDDQLNKYKKRMDENRS
DHD6	TEEELKESKKFAEDLARSARRALKESKRVLEEISQASRSKKLEEIVRRYKEQVKRWQDEWDERAREYRKRMKENRS
DHD7	TKTEEIERLAREIKKLSEKVERLAQEIEELSRRVKEENSTDRELKEANREIERAIREIEKANKRMEEALRRMKYNG
DHD7	TKTEEHERLAREISKLADEHRKLAKIIEELARRIKEENLTDDELREAIRKIEDALRKNKEALKIMKEAAERNRYNT
DHD8	TKKEESRELARESEELARESEKLARKSLELARRAESSGSEEEKRRIIDENRKIIERNREIIERNKEIIEYNKELIS
DHD8	TKDEESLELNRESEELNRKSKELNDRAESSNSEEEEKEILREHKEILREHLEILRRHKEILRRHKYLTS
DHD16	TREELLRENIELAKEHIEIMREILELLQKMEELLERQSSEDILEELRKIIERIRELLDRSRKTHERSEEIAYKEE
DHD16	SEDIAREIKELLRRLKEIIERNQRIAKEHEYIARERKKLDPSNEKERKLLERSRRLQEESKRLLDEMAEIMRRIKKL LD
DHD18	DRQKLIEENIKLLDKHIKILEEILRLLKKDIDLLKKSSSEEVLEELKKIHRRIDKLLDESKKIHKRSSEIVKKRS
DHD18	DEQKLIETSQRLQEKSERLLEKFEQILREASDLYRKPDSEELLRRVEKLLRELEKLIRENQDLARKHEKILRDQS
DHD19	DRQELIRENIELLKKHIKIVKEIQKLIETFIELLKKSSSEEILRRLKKILKRIEKLYRESQEIHKRSEEIAKKRQ
DHD19	DEERLIDKSRELQKESEELLKELLKIFKRIEELLEKPDSEELIREIKKLLETLSEIHKRNEKLARTHEEILRQQS
DHD22	STRDVQREIAKAFKKMADVQKKLAEEIKRHVKNVEKKNKDNDEYRKIATELLKKATESQKKLKELLDRIRKSDS
DHD22	DKDDRSTSLLKRVEKLIDESDRIIDKFTTLIELSRNGKIDDDQYKKELKEILELLKKYDKHVKEVEELLKRLNS
DHD23	SKRKALEVSERVVRISEKVVRVLDESSDLLKKSYDDSDKFAELIDRHEEKIKKWKKLIKEWLEIIQRHKS
DHD23	SAEEFVKLSEEAVKRSKEILDIVRKQVKLVKAGVDKHEITDSLRKSEKLIEEHKELIKTHRDLLRREN
DHD24	SSTEILKRFKRALRESEKIVKHSRRVLKIIREVLKQKPTQAVHDLVRIIETQVKALEEQLKVLKRIVEALERQS
DHD24	DKQKEIKDILEKTRRIAEESRKIAEKFDEIIKRSTEGKIDESLTKELEELVKEVIKLSEDDARTSDDLVRKES
DHD26	DEDESIKLTRKSIEETRKSLKIIKEVVELIREVLKHIKDLDKEIFERIDKILDKYKKQVDTYDEILKEYEKKQR
DHD26	SELDEQKELIKKQEKLIEEQQRLLSKIRRMFKERVKDQELLREIQKVLKRSQEIVETSKKILDRSDKTTE
DHD28	DQKEINTRIVEKLERIFKKSKEIVRQSERVISTIEKKTEDERELDLLRRHVKIVREHLKLLEELLKIIKEVQKESE
DHD28	DTEELVKRLNELLKELSKLVKEFIKILETYRKDQTKDTSKISERVDRILKTYEDLLQKYKEILEKIEKQLS
DHD29	DYARLIDQAVEVTRKVVEVNVTVAEVNDKFAKHLGDEELRRVSEHLKEVSKDLQEVAKKSKDAARQVK

[0813]

DHD29	DVSKVAEEYLQISKTLVDISRTLLEISERLVRLVRTVADDRSEVKKAIEDSIEVLKTSEEVVRQIKRASDKLVKAIS
DHD31	DAKEIQRRVVEIQTEVVKLQKKAVDIIRKIIEAFNNSNIDQSLLEAAKEIVKEIDKLEKLTESLLEESKKLLKRSS
DHD31	SAEEVVKLAKIFLELLRESIKLLKRSVDLLRKSSDPSLDKSEAEKVSREIEKVSDTSLKLSKKALDVVKRALKVAS
DHD32	DEKDAARKARKVSEEAKEASKKIEKALEESKRILNTLKQKKDEQEVKVIKEHEDVLRQIEKIQKQVLEIQKEVAKLL ESLD
DHD32	SADDVARASEKVLRVARESAKAADKSLEVFKEVVKRGDKEAFLQVVKINEEVVKINITVIRILIEVSKTAT
DHD38	DEYVKETLKQLREAIASLREADKRITELVKEARKKPLSEAARKFAEAIVTHVKVVVEHVEVVLRHVEVLVEAKKNGV IDKSILDNALRIIENVIRLLSNVIRVVDEVLQDLD
DHD38	DASDVIRRIHELFEEVHRLIEAVHRAIEDVAKAAQKKGLDESAVEILAELSKELAKLSRRIAEISREIQKVVTDPDD KEAVERLKEIIKEIKKQLDELRDRLRKLQDLLYKLK
DHD60	SEDKAHHDIVRVLEELIKIHDEIMKISEEILKATSDSTATDETKEELKRRSKEAQKKSDTLVKIVKELEKESRKAQS
DHD60	DDEEKYRQIIREAQEISKTAKRILRDAQEISKRIRHQGVDRSEHQRLVDLLRELIKEHHKLLRQQEADTRND
DHD63	DRKDKARKASEKLEEVIQRWKTVADKWKKMVDLVSNGKLSQEEVARVTEELLKIQTELAKLLEEHAKVLQESAS
DHD63	SDEESIKTQSELIKTSEELLKDVKRIDEELQKLRDDPTLDESELKKRVKEWSDRVRKAKEISRKIQEIVKESKKRSS
DHD66	DKDEELRKVIEKYREMVKEYRKVIREYEEVIKSSKTIDKSSLISLSRKMVELSQRVIDVSDEVAKVLSRKQS
DHD66	TDEERLKKQTKELKEQTKQLEKQKDLLEKISNGEISKDEIQEIIKESKKIAKESQKALDSSRKALEEVS
DHD67	DEKEVSKEIIKVLKDIAKVQQKVIEVSQRLASVLRADDDNVVKRALEEYEKILEELRELNKEIEKLTDKYRKVTS
DHD67	DSDEQTKELEKLTELHKRHVEKLKKQTKESREVDSNKLWKSKDVKDKLSESEKELQKLSDQDKKAKDALESSRRKND
DHD69	DAEEQLKLLTKLLRHQQRLLQLIKESLKLIEKIDQSSQENQDEIRKWREVTKKLRELIKTSEKLVRELEKSYKKSS
DHD69	SLRDVVRRYQELVRRYTDELIKTLTEILKKYQKKGAEDKDASTELVKAVRTSLKLSKELLKLINSELLKEDS
DHD71	SKEELKRKLDELKKRSDTLKELSKKLKEISERNPDDKSVHRTIIRIHREFVKNHKEIVRVIEEIVSDKS
DHD71	SKQDEHDRLLKIHDKLVKQHDELLKLLTKLSRAGDSVTKKKLEEILRKLQEVSKQLEESLKDADKVSKDIN
DHD72	TVQSLLEQHVKIVKRSIEILERHTQILQDIARSQGVSKELEDVERQVKEYRKEVKKLEEDLRQLSRNSK
DHD72	SDSDRIEKLIRESTELLKEQQKLAKRSRELAETVESLPLTEEYLKQQREHQKKIEKLLKDSEKHLEELKRLVKSEK
DHD73	DSEKRIEDILRTDLELAKRDAELVKEHIKLVKRIDLSEELKKQVEDVEKESKKLEDSSEKLVQKVRKRSS
DHD73	DEEERAKDLRKYLEEQTQYYRTVTEHLRNLEKVVEELERRGKPSSELQQILERSQRIYKETTEIYDTSKKLIEELDK HHR
DHD148	PLEDILKRHLDKVRELVRLSEEVNKLAKEVLDILKDKRVDEKELDKVLKELEKVVEEYERAVKESRDLLRELRETTR
DHD148	DKERLLEITHERIQKLLDRNLEIIERLLRLLREARDIKDDDKLDKVIKRLKELSEESKDILDKIKELLKESEKELT

[0814]

DHD149	PEDEVIRVIEELLRIAAEVDEVHRRNVEVQEEASRVTDRERIERLNRESEELIKRSRELIEEQRKLIERLERLAT
DHD149	DLEELIKEYAEVVRRHHKAVRDLERLVRELANAKHASEELKRIATEILRIVKELIRVQERLIKLSEDSNEESR
DHD150	PTDEVIEVLKELLRIHRENLRVNEEIVEVNERASRVTDREELERLLRRSNELIKRSRELNEESKKLIEKLERLAT
DHD150	DNEEIIKEARRVVEEYKKAVDRLEELVRRAENAKHASEKELKDIVREILRISKELNKVSERLIELWERSQERAR
DHD151	PKEDIDRVSRELVRVHKELLEVLRKSTEIVEAVARNEKDERTIEEVLEEQERAVRKLEEVSKKHKEAVKRLK
DHD151	EIERLSEEIQKLSDRLIELIRRHSKVLEEIVRLLKHKDNDEREVRRLLKLLRDLTRRYEEVLRKVEEIVKRQEDESR
DHD152	PEEDILRLLRKLVEVDKELLEVVRESTEVVRIVARNEKDVETVERVLRKQEEVVRKYERVSRELEEAVRRLK
DHD152	ELKDLVEEIVKLSKENLKLWEDHSRVLEEIVRLLKHKDNDEREVRRLLKLLEDLTRRAEETSRRIEEIVKEAEDRAR
DHD153	DEERELREVLRKHHRVVREWTKVVEELKRVVELLIKRGETSEEDLIRVLKKLLEMDKRILEVNREVLRVLEKRLT
DHD153	SLEEIIEELVELVRRSVEIAKESDEVARRIVESEDKKKELIDTLRDLHREWQEVTKRAEELVREAEKEVR
DHD154	TAEELLEVHKKSDRVTKEHLRVSEEILKVVEVLTRGEVSSEVLKRVLRKLEELTDKLRRVTEEQRRVVEKLN
DHD154	DLEDLLRRIRRLVDEQRRLVEELERVSRRLEKAVRDNEDERELARLSREHSDIQDKHDKLAREILEVLKRLLERTE
DHD155	PEDDVVRIIKEDLESNREVLREQKEIHRILELVTRGEVSEEAIDRVLKRQEDLLKKQKESTDKARKVVEERR
DHD155	DEVRLITEWLKLSEESTRLLKELVELTRLIRNNVPNVEEILREHERISRELERLSRRLKDLADKLERTRR
DHD156	DEDEVVKVHEEHVKSHEEIHRSHEEVVRAAEEDKRDSRELRTLMEEHRKLLEENEKSIEEVKKIHERVKR
DHD156	KKEELIDISKEVLDLDDEINKISKEILELIKKLLRLKEEGREDKDKAREVKRRIRELHRRIQELNKRLRELHKRVQE TKR
DHD157	PEEDIARRVEDLLRKSEELIKESEKILKESKKLLDRNDSDKRVLETNLRLIDKHTKLLERNLELLEELLKLAEDVAK
DHD157	RFKDLSREYIEVVKRLLELSREALEVLREIKDTDKTDKKRIKELIDRLRKLIEEYKRIIDRLRKLSKDLEEEHR
DHD158	DEEELVKILKELQRLSEESLEINKRLVEILRLLRRGEVPKEEVEKKLREIKKEQEKLDREHEKIKKRIEEITK
DHD158	SLKEKILEIIERNMKLVELSNRSVEIVARILKGEKDDEETLERLLREWDKITRDYEEIIKESRKLVKELEEEAK
DHD159	SKTEILRKALEIHKEQIDIVRKLIELSEEVLKLVEESKEKNLEKLKRIDEETDRLLERLDELHKRLTELAERLK
DHD159	SDDEARKQLEEMKRRLREVEKKSKRVEERVRELERLVRENRERDEDRVLKTLEDLLRENEKLVRTIERHVREQRELSK EVK
DHD160	SEEELEKKADELRKLSEEWRKLQEEDKRLSEMVEKGELDLQEVDEHSLRVLERATEVHRTVDKVIEEILRTTN
DHD160	SEKERHRESQETQEEIRRTHEEIIRKLEEILRRAKAGELPEETLDRLRRIMERLKELSERLDDLVRKLRDDHRREQK
DHD161	SEKEILEELKRILKRVKDISDRLEELDKRTEEIARREPTKELVDELVKIHRDWLRLHEEILKLVDDALKKVEDATK
DHD161	DLRELLELQREASRLHREVKLLTELVKKLELIAKGEDIREEDLKRIKERLEEIKKRSKRIKEESDEIDKKTK
DHD162	SERELQREINKIVRRILEIHREVSELHQRAVKLIRENDNSEELEEISRRIEELSKELEKLVREHDEIVKTIE

[0815]

DHD162	SEREKLDRNDEELKEINKRVEEIKERSDRITEAIEKNERSEEEIRRLSREQNEALQRLLELHKKLVKLHRELLEDTR
DHD163	DKEDVIRVHDEQHKLIEEQLELTRRIAELVREIAKNTASEEEIKEMLKEIKRLDDRSREIQDRLQKLLEEIRRKTK
DHD163	TEEEIVELNKDIQRKSKEHIDLQNELVKKIERAIRENNITEELLEELERLLRESEKIVEEIRRITDKIRKDAK
DHD164	SEKEILERLLRLSKEQNEISEEIHRLTERLVELKRRKDDDERLKRILDRQKRLVERAREISKEYEDLIRKLE
DHD164	SMEELLRKNARLSRKQLKIIDEHLELSTKLTRGEAGDETLEEIERRSREMLEEQRRVDEESKRIREKLK
DHD165	SEEEIRDIVEKLLRTHEEVLKEIKKLLDDSERVRRRELDKKDLDRIQKEQRDIQEENKEKAKRFDELVKELKKAAK
DHD165	SEEHRRTMEKVEKEVRDIKRRSEEVKKKVKANTLSEEDLVRLLERLVEDHKRLQDLSQEIIERDEKATK
DHD166	DEDELAKEIEDVQRRNKESQEEHDKSVKKLEAAERGEIDEDSLLRVLEEDIKVLEKDIEVLERSIEVIEKAE
DHD166	SEKELIRRLLEQQRQHLRLSERLIELSRRLVEVVRKGKDNRDLLRELKKLSEEHKKHSKDDHEKVREIRERER
SYNZIP1	NLVAQLENEVASLENENETLKKKNLHKKDLIAYLEKEIANLRKKIEE
SYNZIP2	ARNAYLRKKIARLKKDNLQLERDEQNLEKIIANLRDEIARLENEVASHEQ
SYNZIP3	NEVTTLENDAAFIENENAYLEKEIARLRKEKAALRNRLAHKK
SYNZIP4	QKVAELKNRVAVKLNRNEQLKNKVEELKNRNAYLKNELATLENEVARLENDVAE
SYNZIP5	NTVKELKNYIQELEERNAELKNLKEHLKFAKAELEFELAAHKFE
SYNZIP6	QKVAQLKNRVAYKLKENAKLENIVARLENDNANLEKDIANLEKDIANLERDVAR
SYNZIP7	KEIEYLEKEIERLKDLREHLKQDNAAHRQELNALRLEEAKLEFILAHLLST
SYNZIP8	KEIANLEKEIASLEKKVAVLKQRNAAHKQEVAALRKEIAYVEDEIQYVEDE
SYNZIP9	QKVESLKQKIEELKQRKAQLKNDIANLEKEIAYAET
SYNZIP10	NLIATLRSTAAVLENENHVLEKEKEKLRKEKEQLLNKLEAYK
SYNZIP11	ELTDELKNKKEALRKDNAALLNELASLENEIANLEKEIAYFK
SYNZIP12	NEDLVLENRLAALRNENAALENDLARLEKEIAYLEKEIEREK
SYNZIP13	QKVEELKNKIAELENRNAVKKNRVAHLKQEIAYLKDELAAHEFE
SYNZIP14	NDLDAYEREAEKLEKKNEVLRNRLAALENELATLRQEVASMKQELQS
SYNZIP15	FENVTHEFILATLENENAKLRRLEAKLERELARLRNEVAWL
SYNZIP16	NILASLENKKEELKKLNAHLLKEIENLEKEIANLEKEIAYFK
SYNZIP17	NEKEELKSKKAELRNRIEQLKQKREQLKQKIANLRKEIEAYK
SYNZIP18	SIAATLENDLARLENENARLEKDIANLERDLAKLEREEAYF
SYNZIP19	NELESLENKKEELKNRNEELKQKREQLKQKLAALRNKLDAYKNRL

SYNZIP20	STVEELLRAIQELEKRNAELKNREELKNLVAHLRQELAAHKYE
SYNZIP21	NEVAQLENDVAVIENENAYLEKEIARLRKEIAALRDRLAHKK
SYNZIP22	KRIAYLRKKIAALKKDNANLEKDIANLENEIERLIKEIKTLENEVASHEQ

[0816]

[0817] 일부 구현예에서 본 명세서에 기재된 바와 같은 한 쌍의 이량체화 도메인은 서로 비공유적으로 결합한다.

[0818] 일부 구현예에서, 본 명세서에 기재된 바와 같은 한 쌍의 이량체화 도메인은 공유적으로 결합하여, 예를 들어 융합체(예를 들어, 본 명세서에 기재된 바와 같은, 예를 들어 인테인 매개 융합체)를 형성한다. 구현예에서, 한 쌍의 인테인 이량체화 도메인은 표 33의 단일 행에 열거된 바와 같은 사슬 A 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열) 및 사슬 B 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열)을 포함한다.

[0819] *유전자 변형 시스템을 위한 국재화 서열*

[0820] 특정 구현예에서, 유전자 편집기 시스템 RNA는 세포내 국재화 서열, 예를 들어 핵 국재화 서열(NLS)을 추가로 포함한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 서열번호 4000 및/또는 서열번호 4001에 포함된 바와 같은 NLS, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산

서열을 갖는 NLS를 포함한다.

[0821] 핵 국재화 서열은 RNA의 핵 내로의 유입을 촉진하는 RNA 서열일 수 있다. 특정 구현예에서, 핵 위치 신호는 주형 RNA 상에 위치한다. 특정 구현예에서, 유전자 변형 폴리펩티드는 제1 RNA 상에서 인코딩되며, 주형 RNA는 제2의 별개의 RNA며, 핵 위치 신호는 주형 RNA 상에 위치되고, 유전자 변형 폴리펩티드를 인코딩하는 RNA 상에서는 위치하지 않는다. 이론에 구속되고자 하지는 않지만, 일부 구현예에서 유전자 변형 폴리펩티드를 인코딩하는 RNA는 일차적으로 세포질로 표적화되어 그의 번역을 촉진하는 한편, 주형 RNA는 일차적으로 핵으로 표적화되어 그의 게놈 내로의 삽입을 촉진한다. 일부 구현예에서, 핵 위치 신호는 3' 말단, 5' 말단에, 또는 주형 RNA의 내부 영역에 존재한다. 일부 구현예에서, 핵 위치 신호는 이종성 서열의 3'(예를 들어, 직접적으로 이종성 서열의 3')거나, 또는 이종성 서열의 5'(예를 들어, 직접적으로 이종성 서열의 5')다. 일부 구현예에서, 핵 위치 신호는 주형 RNA의 5' UTR 외부 또는 3' UTR에 위치한다. 일부 구현예에서, 핵 위치 신호는 5' UTR과 3' UTR 사이에 위치되고, 여기서 선택적으로 핵 위치 신호는 이식유전자를 이용하여 전사되지 않는다(예를 들어, 핵 위치 신호는 항-센스 배향이거나, 전사 말단 신호 또는 폴리아데닐화 신호의 하류임). 일부 구현예에서, 핵 국재화 서열은 인트론의 내부에 위치한다. 일부 구현예에서, 복수의 동일 또는 상이한 핵 위치 신호는 RNA 내, 예를 들어 주형 RNA 내에 존재한다. 일부 구현예에서, 핵 위치 신호는 길이가 5, 10, 25, 50, 75, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900 또는 1000 bp 미만이다. 다양한 RNA 핵 국재화 서열이 사용될 수 있다. 예를 들어, 문헌[Lubelsky and Ulitsky, *Nature* 555 (107-111), 2018]은 RNA 국재화를 핵 내로 구동하는 RNA 서열을 기재한다. 일부 구현예에서, 핵 위치 신호는 SINE-유도된 핵 RNA 국재화(SIRLOIN) 신호이다. 일부 구현예에서, 핵 위치 신호는 핵-풍부화된 단백질에 결합한다. 일부 구현예에서, 핵 위치 신호는 HNRNPK 단백질에 결합한다. 일부 구현예에서, 핵 위치 신호는 피리미딘에 풍부하며, 예를 들어, C/T 풍부, C/U 풍부, C 풍부, T 풍부, 또는 U 풍부 영역이 있다. 일부 구현예에서, 핵 위치 신호는 긴 비-코딩 RNA로부터 유래된다. 일부 구현예에서, 핵 위치 신호는 MALATI 긴 비-코딩 RNA로부터 유래되거나, MALAT1의 600 뉴클레오티드 M 영역이다(문헌[Miyagawa et al., *RNA* 18, (738-751), 2012]에 기재됨). 일부 구현예에서, 핵 위치 신호는 BORG 긴 비-코딩 RNA로부터 유래되거나, AGCCC 모티프다(문헌[Zhang et al., *Molecular and Cellular Biology* 34, 2318-2329 (2014)]에 기재됨). 일부 구현예에서, 핵 국재화 서열은 문헌[Shukla et al., *The EMBO Journal* e98452 (2018)]에 기재되어 있다. 일부 구현예에서, 핵 위치 신호는 레트로바이러스로부터 유래된다.

[0822] 일부 구현예에서, 본 명세서에 기재된 폴리펩티드는 하나 이상(예를 들어, 2, 3, 4, 5)의 핵 표적화 서열, 예를 들어 핵 국재화 서열(NLS)를 포함한다. 일부 구현예에서, NLS는 이분절(bipartite) NLS이다. 일부 구현예에서, NLS는 NLS를 포함하는 단백질의 세포 핵 내로의 유입을 촉진한다. 일부 구현예에서, NLS는 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드의 N-말단에 융합된다. 일부 구현예에서, NLS는 유전자 변형 폴리펩티드의 C-말단에 융합된다. 일부 구현예에서, NLS는 Cas 도메인의 N-말단 또는 C-말단에 융합된다. 일부 구현예에서, 링커 서열은 NLS와 유전자 변형 폴리펩티드의 이웃 도메인 사이에 배치된다.

[0823] 일부 구현예에서, NLS는 아미노산 서열 MDSLLMNRRKFLYQFKNVRWAKGRRETYLC(서열번호 9), PKKRKVEGADKRTADGSEFESPKKKRKV(서열번호 10), RKSGKIAAIWKRPRKPKKKRKV(서열번호 11), KRTADGSEFESPKKKRKV(서열번호 12), KKTELQTTNAENKTKKL(서열번호 13), 또는 KRGINDRNFWRGENGRKTR(서열번호 14), KRPAATKKAGQAKKKK(서열번호 15), 또는 이의 기능성 단편 또는 변이체를 포함한다. 예시적인 NLS 서열은 PCT/EP2000/011690에도 기재되어 있으며, 이의 내용은 예시적인 핵 국재화 서열의 개시에 대하여 본 명세서에 참고로 포함된다. 일부 구현예에서, NLS는 표 11에 개시된 바와 같은 아미노산 서열을 포함한다. 이 표의 NLS는, 핵으로의 세포내 국재화를 개선하기 위해, 폴리펩티드에서 하나 이상의 위치에서 폴리펩티드 내 하나 이상의 복제, 예를 들어 N-말단 도메인 내, 펩티드 도메인 사이, C-말단 도메인 내, 또는 위치의 조합에서 NLS의 1, 2, 3 이상의 복제를 사용하여 이용될 수 있다. 다수의 독특한 서열들이 단일 폴리펩티드 내에서 사용될 수 있다. 서열은 예를 들어 염기성 아미노산의 1 또는 2개의 스트레치(stretch)를 갖는, 자연히 단일분절 또는 이분절일 수 있거나, 또는 키메라 이분절 서열로 사용될 수 있다. 서열 참조는, 세포내 국재화 예측 알고리즘을 이용하여 캐낸 서열에 대해 SeqNLS로서 나타낸 경우를 제외하고는, UniProt 수탁번호에 상응한다(문헌[Lin et al BMC Bioinformat 13:157 (2012)], 이는 그 전체가 본 명세서에 참고로 포함된다).

[0824] *표 11 유전자 변형 시스템에서 사용하기 위한 예시적인 핵 국재화 신호*

서열	서열 참조	서열번호
AHFKISGEKRPSTDPGKKAK NPKKKKKKDP	Q76IQ7	223
AHRAKKMSKTHA	P21827	224
ASPEYVNLPINGNG	SeqNLS	225
CTKRPRW	O88622, Q86W56, Q9QYM2, O02776	226
DKAKRVSRNKSEKKRR	O15516, Q5RAK8, Q91YB2, Q91YB0, Q8QGQ6, O08785, Q9WVS9, Q6YGZ4	227
EELRLKEELLKGIYA	Q9QY16, Q9UHL0, Q2TBP1, Q9QY15	228
EEQLRRRKNSRLNNTG	G5EFF5	229
EVLKVIRTGKRKKKAWKR MVTKVC	SeqNLS	230
HHHHHHHHHHHHQPH	Q63934, G3V7L5, Q12837	231
HKKKHPDASVNFSEFSK	P10103, Q4R844, P12682, B0CM99, A9RA84, Q6YKA4, P09429, P63159, Q08IE6, P63158, Q9YH06, B1MTB0	232
HKRTKK	Q2R2D5	233
IINGRKLKLKKSRRRSSQTS NNSFTSRRS	SeqNLS	234
KAEQERRK	Q8LH59	235
KEKRKRREELFIEQKKRK	SeqNLS	236
KKGKDEWFSRGKKP	P30999	237
KKGPSVQKRKKT	Q6ZN17	238

[0825]

KKKTVINDLLHYKKEK	SeqNLS, P32354	239
KKNGGKGKNKPSAKJKK	SeqNLS	240
KKPKWDDFKKKKK	Q15397, Q8BKS9, Q562C7	241
KKRKKD	SeqNLS, Q91Z62, Q1A730, Q969P5, Q2KHT6, Q9CPU7	242
KKRRKRRRK	SeqNLS	243
KKRRRRARK	Q9UMS6, D4A702, Q91YE8	244
KKSKRGR	Q9UBS0	245
KKSRKRGS	B4FG96	246
KKSTALSRELGKIMRRR	SeqNLS, P32354	247
KKSYQDPEIIAHSRPRK	Q9U7C9	248
KKTGKNRKLKSKRVKTR	Q9Z301, O54943, Q8K3T2	249
KKVSIAGQSGKLWRWKR	Q6YUL8	250
KKYENVVIKRSPRKRGRPRK	SeqNLS	251
KNKKRK	SeqNLS	252
KPKKKR	SeqNLS	253
KRAMKDDSHGNSTSPKRRK	Q0E671	254
KRANSNLVAAYEKAKKK	P23508	255
KRASEDTTSGSPPKKSSAGPKR	Q9BZZ5, Q5R644	256
KRFKRRWMVRKMKTKK	SeqNLS	257
KRGLNSSFETSPKKVK	Q8IV63	258
KRGNSSIGPNDLSKRKQRKK	SeqNLS	259
KRIHSVSLSQSQIDPSKKVKRAK	SeqNLS	260
KRKGKLKNKGSKRKK	O15381	261
KRRRRRRREKRKR	Q96GM8	262
KRSNDRTYSPEEEKQRRA	Q91ZF2	263
KRTVATNGDASGAHRAKKMSK	SeqNLS	264

[0826]

KRVYNKGEDEQEHLPKGKKR	SeqNLS	265
KSGKAPRRRAVSMDNSNK	Q9WVH4, O43524	266
KVNFLDMSLDDIIIYKELE	Q9P127	267
KVQHRIAKKTTRRRR	Q9DXE6	268
LSPSLSPL	Q9Y261, P32182, P35583	269
MDSLLMNRRKFLYQFKNVRWAKGRRETYLC	Q9GZX7	270
MPQNEYIELHRKRYGYRLDYHEKKRKKESREAHERSKKAKKMIGLKAKLYHK	SeqNLS	271
MVQLRPRASR	SeqNLS	272
NNKLLAKRRKGGASPKDDPMDDIK	Q965G5	273
NYKRPMDGTYGPPAKRHEGE	O14497, A2BH40	274
PDTKRAKLDSSETTMVKKK	SeqNLS	275
PEKRTKI	SeqNLS	276
PGGRGKKK	Q719N1, Q9UBP0, A2VDN5	277
PGKMDKGEHRQERRDRPY	Q01844, Q61545	278
PKKGDKYDKTD	Q45FA5	279
PKKKSRK	O35914, Q01954	280
PKKNKPE	Q22663	281
PKKRAKV	P04295, P89438	282
PKPKKLKVE	P55263, P55262, P55264, Q64640	283
PKRGRGR	Q9FYS5, Q43386	284
PKRRLVDDA	P0C797	285
PKRRRTY	SeqNLS	286
PLFKRR	A8X6H4, Q9TXJ0	287
PLRKAKR	Q86WB0, Q5R8V9	288
PPAKRKCIF	Q6AZ28, O75928, Q8C5D8	289
PPARRRRL	Q8NAG6	290

[0827]

PPKKKRKV	Q3L6L5, P03070, P14999, P03071	291
PPNKRMKVKH	Q8BN78	292
PPRIYPQLPSAPT	P0C799	293
PQRSPFPKSSVKR	SeqNLS	294
PRPRKVPR	P0C799	295
PRRRVQRKR	SeqNLS, Q5R448, Q5TAQ9	296
PRRVRLK	Q58DJ0, P56477, Q13568	297
PSRKRPR	Q62315, Q5F363, Q92833	298
PSSKKRKV	SeqNLS	299
PTKKRVK	P07664	300
QRPGPYDRP	SeqNLS	301
RGKGGKGLGKGGAKRHRK	SeqNLS	302
RKAGKGGGGHKTTKKRSAKDEKVP	B4FG96	303
RKIKLKRAK	A1L3G9	304
RKIKRKRAK	B9X187	305
RKKEAPGPREELRSRGR	O35126, P54258, Q5IS70, P54259	306
RKKRKGK	SeqNLS, Q29243, Q62165, Q28685, O18738, Q9TSZ6, Q14118	307
RKKRRQRRR	P04326, P69697, P69698, P05907, P20879, P04613, P19553, P0C1J9, P20893, P12506, P04612, Q73370, P0C1K0, P05906, P35965, P04609, P04610, P04614, P04608, P05905	308
RKKSIPLSIKNLKRKHKRKKNKITR	Q9C0C9	309
RKLVKPKNTKMKTKLRTNPY	Q14190	310
RKRLILSDKGQLDWKK	SeqNLS, Q91Z62, Q1A730, Q2KHT6, Q9CPU7	311
RKRLKSK	Q13309	312
RKRRVRDNM	Q8QPH4, Q809M7, A8C8X1, Q2VNC5, Q38SQ0, O89749, Q6DNQ9, Q809L9, Q0A429, Q20NV3, P16509, P16505, Q6DNQ5, P16506, Q6XT06, P26118,	313

[0828]

	Q2ICQ2, Q2RCG8, Q0A2D0, Q0A2H9, Q9IQ46, Q809M3, Q6J847, Q6J856, B4URE4, A4GCM7, Q0A440, P26120, P16511,	
RKRSPKDKKEKDLDGAGKRRKT	Q7RTP6	314
RKRTPRVDGQTGENDMNKRRRK	O94851	315
RLPVRRRRRR	P04499, P12541, P03269, P48313, P03270	316
RLRFRKPKSK	P69469	317
RQQRKR	Q14980	318
RRDLNSSFETSPKKVK	Q8K3G5	319
RRDRAKLR	Q9SLB8	320
RRGDGRRR	Q80WE1, Q5R9B4, Q06787, P35922	321
RRGRKRKAEKQ	Q812D1, Q5XXA9, Q99JF8, Q8MJG1, Q66T72, O75475	322
RRKKRR	Q0VD86, Q58DS6, Q5R6G2, Q9ERI5, Q6AYK2, Q6NYC1	323
RRKRSKSEDMDSVESKRRR	Q7TT18	324
RRKRSR	Q99PU7, D3ZHS6, Q92560, A2VDM8	325
RRPKGKTLQKRKPK	Q6ZN17	326
RRRGFERFGPDNMGRKRK	Q63014, Q9DBR0	327
RRRGKNKVAAQNCRK	SeqNLS	328
RRRKRR	Q5FVH8, Q6MZT1, Q08DH5, Q8BQP9	329
RRRQKQKGGASRRR	SeqNLS	330
RRRREGPRARRRR	P08313, P10231	331
RRTIRLKLVYDKCDRSCKIQKKNRNKCQYCRFHKCLSVGMSHNAIRFGRMPRSEKAKLKAE	SeqNLS	332
RRVPQRKEVSRCRKCRK	Q5RJN4, Q32L09, Q8CAK3, Q9NUL5	333
RVGGRRQAVECIEDLLNEPGQPLDLSCKRPRP	P03255	334

[0829]

RVVKLRIAP	P52639, Q8JMN0	335
RVVRRR	P70278	336
SKRKTKISRKTR	Q5RAY1, O00443	337
SYVKTVPNRTRTYIKL	P21935	338
TGKNEAKKRKIA	P52739, Q8K3J5, Q5RAU9	339
TLSPASSPSSVSCPVIPASTD ESPGSALNI	SeqNLS	340
VSKKQRTGKKIH	P52739, Q8K3J5, Q5RAU9	341
SPKKKRKVE		342
KRTAD GSEFE SPKKKRKVE		343
PAAKRVKLD		344
PKKKRKV		345
MDSLLMNRRKFLYQFKNVR WAKGRRETYLC		346
SPKKKRKVEAS		347
MAPKKKRKVGIHRGVP		348

[0830]

[0831] 일부 구현예에서, NLS는 이분절 NLS이다. 이분절 NLS는 스페이서 서열에 의해 분리된 2개의 염기성 아미노산 클러스터를 통상적으로 포함한다(이는, 예를 들어 약 10개의 아미노산 길이). 단일분절 NLS는 통상적으로 스페이서가 결여되어 있다. 이분절 NLS의 예는 뉴클레오플라스민(nucleoplasmin) NLS으로, 이는 서열 KR[PAATKKAGQA]KKKK(서열번호 15)를 갖고, 스페이서는 괄호표시되었다. 또 다른 예시적인 이분절 NLS는 서열 PKKKRKVEGADKRTADGSEFESPKKKRKV(서열번호 16)를 갖는다. 예시적인 NLS는 국제 특허 WO2020051561에 기재되어 있으며, 이는 핵 국재화 서열에 대한 그의 개시내용을 포함하여 그 전체가 본 명세서에 참고로 포함된다.

[0832] 특정 구현예에서, 유전자 편집기 시스템 폴리펩티드(예를 들어, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드)는 세포내 국재화 서열, 예를 들어 핵 국재화 서열 및/또는 핵소체 국재화 서열을 추가로 포함한다. 핵 국재화 서열 및/또는 핵소체 국재화 서열은 단백질의 핵 및/또는 핵소체 내로의 유입을 촉진하는 아미노산 서열일 수 있으며, 이는 이종성 서열의 게놈 내로의 통합을 촉진할 수 있다. 특정 구현예에서, 유전자 편집기 시스템 폴리펩티드(예를 들어, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드)는 핵소체 국재화 서열을 추가로 포함한다. 특정 구현예에서, 유전자 변형 폴리펩티드는 제1 RNA 상에서 인코딩되고, 주형 RNA는 제2의 별개의 RNA이고, 핵소체 국재화 신호는 유전자 변형 폴리펩티드를 인코딩하는 RNA 상에서 인코딩되고, 주형 RNA 상에서는 인코딩되지 않는다. 일부 구현예에서, 핵소체 국재화 신호는 폴리펩티드의 N-말단, C-말단에, 또는 내부 영역에 위치한다. 일부 구현예에서, 복수의 동일하거나 상이한 핵소체 국재화 신호들이 사용된다. 일부 구현예에서, 핵 위치 신호는 길이가 5, 10, 25, 50, 75, 또는 100개 미만의 아미노산이다. 다양한 폴리펩티드 핵소체 국재화 신호가 사용될 수 있다. 예를 들어, 문헌[Yang et al., *Journal of Biomedical Science* 22, 33 (2015)]은 핵소체 국재화 신호로서 또한 기능하는 핵 위치 신호를 기재한다. 일부 구현예에서, 핵소체 국재화 신호는 핵 위치 신호일 수도 있다. 일부 구현예에서, 핵소체 국재화 신호는 핵 위치 신호와 중복될 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 뻗어 있는 염기성 잔기들을 포함할 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 아르기닌 및 리신 잔기가 풍부할 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 핵소체에서 풍부화된 단백질로부터 유래될 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 리보솜성 RNA 유전자좌에서 풍부화된 단백질로부터 유래될 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 rRNA에 결합하는 단백질로부터 유래될 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 MSP58로부터 유래될 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 단일분절 모티프일 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 이분절 모티프일 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 다수의 단일분절 또는 이분절 모티프로 이루어질 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 단일분절 또는 이분절 모티프의 혼합물로 이루어질 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 이중 이분절 모티프일 수 있다. 일부 구현예에서, 핵소체 국재화 모티프는 KRASSQALGTIPKRRSSSRFIKRKK(서열번호 17)일 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 핵 인자-κB-유도

키나제로부터 유래될 수 있다. 일부 구현예에서, 핵소체 국재화 신호는 RKKRKKK 모티프(서열번호 18)일 수 있다(문헌[Birbach et al., Journal of Cell Science, 117 (3615-3624), 2004]에 기재됨).

[0833] *유전자 변형 폴리펩티드 및 시스템의 진화된 변이체*

[0834] 일부 구현예에서, 본 발명은 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드의 진화된 변이체를 제공한다. 진화된 변이체는 일부 구현예에서 참조 유전자 변형 폴리펩티드, 또는 그 안에 포함된 단편 또는 도메인 중 하나를 돌연변이시킴으로써 생산될 수 있다. 일부 구현예에서, 하나 이상의 도메인(예를 들어, 역전사효소 도메인)이 진화된다. 하나 이상의 그러한 진화된 변이체 도메인은, 일부 구현예에서 단독으로 또는 다른 도메인과 함께 진화될 수 있다. 진화된 변이체 도메인 또는 도메인들은, 일부 구현예에서 진화되지 않은 동족 구성요소(들) 또는 동족 구성요소(들)의 진화된 변이체와 조합될 수 있으며, 예를 들어 이는 병렬적 또는 연속적 방식으로 진화될 수 있다.

[0835] 일부 구현예에서, 참조 유전자 변형 폴리펩티드, 또는 이의 단편 또는 도메인을 돌연변이시키는 과정은 참조 유전자 변형 폴리펩티드 또는 이의 단편 또는 도메인을 돌연변이시키는 것을 포함한다. 구현예에서, 돌연변이유발은, 예를 들어 본 명세서에 기재된 바와 같이 연속적인 진화 방법(예를 들어, PACE) 또는 비연속적인 진화 방법(예를 들어, PANCE)을 포함한다. 일부 구현예에서, 진화된 유전자 변형 폴리펩티드 또는 이의 단편 또는 도메인은, 참조 유전자 변형 폴리펩티드의 아미노산 서열, 또는 이의 단편 또는 도메인에 비해 그의 아미노산 서열 내로 도입된 하나 이상의 아미노산 변이를 포함한다. 구현예에서, 아미노산 서열 변이는, 예를 들어 코딩 서열에서 임의의 특정 위치에서의 코돈의 변화, 하나 이상의 아미노산의 결실(예를 들어, 절단된 단백질), 하나 이상의 아미노산의 삽입, 또는 상기의 임의의 조합을 생성하는 유전자 변형 폴리펩티드를 인코딩하는 뉴클레오티드 서열에서의 변화 결과로서, 참조 유전자 변형 폴리펩티드의 아미노산 서열 내에 하나 이상의 돌연변이된 잔기(예를 들어, 보존적 치환, 비보존적 치환, 또는 이의 조합)를 포함할 수 있다. 진화된 변이체 유전자 변형 폴리펩티드는 유전자 변형 폴리펩티드의 하나 이상의 구성요소 또는 도메인에서의 변이체(예를 들어, 역전사효소 도메인 내로 도입된 변이체)를 포함할 수 있다.

[0836] 일부 양태에서, 본 개시내용은 유전자 변형 폴리펩티드, 시스템, 키트, 및 유전자 변형 폴리펩티드의 진화된 변이체를 사용 또는 포함하는 방법을 제공하며, 예를 들어 유전자 변형 폴리펩티드의 진화된 변이체 또는 PACE 또는 PANCE에 의해 생산되거나 생산가능한 유전자 변형 폴리펩티드를 이용한다. 구현예에서, 진화되지 않은 참조 유전자 변형 폴리펩티드는 본 명세서에 개시된 바와 같은 유전자 변형 폴리펩티드이다.

[0837] 본 명세서에 사용된 바와 같은 용어 "파지-보조된 지속적 진화(PACE)"는, 바이러스 벡터로서 파지를 이용하는 연속적 진화를 일반적으로 지칭한다. PACE 기술의 예는, 예를 들어 2009년 9월 8일 출원되고, 2010년 3월 11일 WO 2010/028347로 공개된, 국제 PCT 출원 PCT/US 2009/056194; 2011년 12월 22일 출원되고, 2012년 6월 28일 WO 2012/088381로 공개된, 국제 PCT 출원 PCT/US2011/066747; 2015년 5월 5일 발행된 미국 특허 9,023,594; 2017년 9월 26일 발행된 미국 특허 9,771,574; 2016년 7월 19일 발행된 미국 특허 9,394,537; 2015년 1월 20일 출원되고, 2015년 9월 11일 WO 2015/134121로 공개된, 국제 PCT 출원 PCT/US2015/012022; 2019년 1월 15일 발행된 미국 특허 10,179,911; 및 2016년 4월 15일 출원되고, 2016년 10월 20일 공개된 WO 2016/168631로 공개된 국제 PCT 출원 PCT/US2016/027795에 기재되어 있으며, 이들 각각의 전체 내용은 본 명세서에 참고로 포함된다.

[0838] 본 명세서에 사용된 바와 같은 용어 "파지-보조된 비연속적 진화(PANCE)"는 바이러스 벡터로서 파지를 이용하는 비연속적 진화를 일반적으로 지칭한다. PANCE 기술의 예는, 예를 들어 그 전체가 본 명세서에 참고로 포함된 문헌[Suzuki T. et al, Crystal structures reveal an elusive functional domain of pyrrolysyl-tRNA synthetase, Nat Chem Biol. 13(12): 1261-1266 (2017)]에 기재되어 있다. 간략하게, PANCE는, 신선한 숙주 세포(예를 들어, 이. 콜라이(E. coli) 세포)에 걸쳐, 진화될 관심 유전자를 함유하는 선택 파지(SP)의 연속 플라스크 수송을 사용한 신속한 생체 내 유도 진화(directed evolution)를 위한 기법이다. SP 내에 함유된 유전자가 연속적으로 진화하는 동안, 숙주 세포 내부의 유전자는 일정하게 유지될 수 있다. 파지 성장 후, 감염된 세포의 분취액은 숙주 이. 콜라이를 함유하는 연속된 플라스크를 형질주입시키는 데 사용될 수 있다. 이 과정은 원하는 표현형이 진화될 때까지, 예를 들어 많은 수송이 필요한 대로 반복 및/또는 계속될 수 있다.

[0839] PACE 및 PANCE를 유전자 변형 폴리펩티드에 적용하는 방법은 특히 상기 참고문헌들을 참조하여 당업자에 의해 쉽게 이해될 수 있다. 예를 들어 파지 입자를 사용하여, 예를 들어 숙주 세포의 집단에서 게놈-변형 단백질 또는 시스템의 연속 진화를 지시하기 위한 추가의 예시적인 방법은 유전자 변형 폴리펩티드의 진화된 변이체, 또는 이의 단편 또는 서브도메인을 생성하도록 적용될 수 있다. 그러한 방법의 비제한적인 예는 2009년 9월 8일 출원되고, 2010년 3월 11일 WO 2010/028347로 공개된, 국제 PCT 출원 PCT/US 2009/056194; 2011년 12월 22일

출원되고, 2012년 6월 28일 WO 2012/088381로 공개된, 국제 PCT 출원 PCT/US2011/066747; 2015년 5월 5일 발행된 미국 특허 9,023,594; 2017년 9월 26일 발행된 미국 특허 9,771,574; 2016년 7월 19일 발행된 미국 특허 9,394,537; 2015년 1월 20일 출원되고, 2015년 9월 11일 WO 2015/134121로 공개된, 국제 PCT 출원 PCT/US2015/012022; 2019년 1월 15일 발행된 미국 특허 10,179,911; 2019년 6월 14일 출원되고, 2019년 1월 31일 국제 특허 공개 WO 2019/023680으로 공개된 국제 출원 PCT/US2019/37216, 2016년 4월 15일 출원되고, 2016년 10월 20일 공개된 WO 2016/168631로 공개된 국제 PCT 출원 PCT/US2016/027795, 및 2019년 8월 23일 출원된 국제 특허 공개 PCT/US2019/47996에 기재되어 있으며, 이들 각각은 그 전체가 본 명세서에 참고로 포함된다.

[0840] 제한되지 않는 일부 예시적인 구현예에서, 진화된 변이체 유전자 변형 폴리펩티드, 이의 단편 또는 도메인의 진화 방법은: (a) 숙주 세포 집단을 관심 유전자(시작 유전자 변형 폴리펩티드, 또는 이의 단편 또는 도메인)를 포함하는 바이러스 벡터의 집단과 접촉시키는 단계를 포함하며, 여기서 (1) 숙주 세포는 바이러스 벡터에 의한 감염이 잘 되고; (2) 숙주 세포는 바이러스 입자의 생성에 필요한 바이러스 유전자를 발현하고; (3) 감염성 바이러스 입자의 생산에 필요한 적어도 하나의 바이러스 유전자의 발현이 관심 유전자의 기능에 의존하고; 및/또는 (4) 바이러스 벡터가 숙주 세포에서 단백질의 발현을 가능하게 하며, 이는 숙주 세포에 의해 바이러스 입자로 복제 및 패키징될 수 있다. 일부 구현예에서, 이 방법은 (b) 돌연변이율을 상승시키는 돌연변이를 가진 숙주 세포, 또는 이의 조합을 사용하여, (예를 들어, 돌연변이 플라스미드 또는 일부 게놈 변형 중 어느 하나, 예를 들어 프루핑-손상된 DNA 중합효소, SOS 유전자, 예컨대 UmuC, UmuD', 및/또는 RecA를 운반함으로써, 플라스미드-결합된 경우, 돌연변이는 유도성 프로모터의 제어 하에 있을 수 있음) 숙주 세포를 돌연변이 유발원과 접촉시키는 단계를 포함한다. 일부 구현예에서, 이 방법은 (c) 바이러스 입자의 바이러스 복제 및 생산을 가능하게 하는 조건 하에서 숙주 세포의 집단을 인큐베이션하는 단계를 포함하고, 여기서 숙주 세포는 숙주 세포 집단으로부터 제거되고, 신선하고 감염되지 않은 숙주 세포는 숙주 세포 집단 내로 도입되어, 이에 따라 숙주 세포의 집단을 보충하고 숙주 세포의 흐름을 생성한다. 일부 구현예에서, 세포는 관심 유전자가 돌연변이의 획득을 가능하게 하는 조건 하에서 인큐베이션된다. 일부 구현예에서, 이 방법은 (d) 바이러스 벡터의 돌연변이된 형태를 단리하고, 숙주 세포의 집단으로부터 진화된 유전자 생성물(예를 들어, 진화된 변이체 유전자 변형 폴리펩티드, 또는 이의 단편 또는 도메인)을 인코딩하는 단계를 추가로 포함한다.

[0841] 당업자는 상기 기재된 프레임워크 내에서 이용가능한 다양한 특징들을 이해할 것이다. 예를 들어 일부 구현예에서, 바이러스 벡터 또는 파지는 사상 파지(filamentous phage), 예를 들어 M13 파지, 예를 들어 M13 선택 파지이다. 특정 구현예에서, 감염성 바이러스 입자들의 생산에 필요한 유전자는 M13 유전자 III(gIII)이다. 구현예에서, 파지는 기능성 gIII이 결여될 수 있지만, 다르게는 gI, gII, gIV, gV, gVI, gVII, gVIII, gIX, 및 gX를 포함한다. 일부 구현예에서, 감염성 VSV 입자의 생성은 외피막(envelope) 단백질 VSV-G를 수반한다. 다양한 구현예는 레트로바이러스 벡터, 예를 들어 뮤린 백혈병 바이러스 벡터, 또는 렌티바이러스 벡터를 사용할 수 있다. 구현예에서, 레트로바이러스 벡터는, 예를 들어 바이러스의 천연 외피막 단백질에 대한 대체물로서 VSV-G 외피막 단백질로 효율적으로 패키징될 수 있다.

[0842] 일부 구현예에서, 숙주 세포는 적합한 횟수의 바이러스 생애 주기, 예를 들어 적어도 10, 적어도 20, 적어도 30, 적어도 40, 적어도 50, 적어도 100, 적어도 200, 적어도 300, 적어도 400, 적어도, 500, 적어도 600, 적어도 700, 적어도 800, 적어도 900, 적어도 1000, 적어도 1250, 적어도 1500, 적어도 1750, 적어도 2000, 적어도 2500, 적어도 3000, 적어도 4000, 적어도 5000, 적어도 7500, 적어도 10000회, 또는 그 이상의 연속된 바이러스 생애 주기에 따라 인큐베이션되며, 이는 M13 파지의 예시적이며 비제한적인 예에서 바이러스 생애 주기 당 10 내지 20분이다. 유사하게, 숙주 세포가 숙주 세포의 집단으로 남아 있는 시간을 조정하기 위해 조건이 조절될 수 있으며, 예를 들어 약 10, 약 11, 약 12, 약 13, 약 14, 약 15, 약 16, 약 17, 약 18, 약 19, 약 20, 약 21, 약 22, 약 23, 약 24, 약 25, 약 30, 약 35, 약 40, 약 45, 약 50, 약 55, 약 60, 약 70, 약 80, 약 90, 약 100, 약 120, 약 150, 또는 약 180분이다. 숙주 세포 집단은 부분적으로는 숙주 세포의 밀도, 또는 일부 구현예에서 유입물 내 숙주 세포 밀도에 의해 제어될 수 있으며, 예를 들어 10^3개 세포/ml, 약 10^4개 세포/ml, 약 10^5개 세포/ml, 약 $5 \cdot 10^5$개 세포/ml, 약 10^6개 세포/ml, 약 $5 \cdot 10^6$개 세포/ml, 약 10^7개 세포/ml, 약 $5 \cdot 10^7$개 세포/ml, 약 10^8개 세포/ml, 약 $5 \cdot 10^8$개 세포/ml, 약 10^9개 세포/ml, 약 $5 \cdot 10^9$개 세포/ml, 약 10^{10}개 세포/ml, 또는 약 $5 \cdot 10^{10}$개 세포/ml이다.

[0843] *인테인*

[0844] 일부 구현예에서, 하기에 더 상세히 기재되는 바와 같이, N-말단 부분의 C-말단 부분으로의 연결을 위해 인테인

-N(intN) 도메인은 본 명세서에 기재된 유전자 변형 폴리펩티드의 제1 도메인의 N-말단 부분에 융합될 수 있고, 인테인-C(intC) 도메인은 본 명세서에 기재된 유전자 변형 폴리펩티드의 제2 도메인의 C-말단 부분에 융합될 수 있으며, 이에 의해 제1 및 제2 도메인을 연결할 수 있다. 일부 구현예에서, 제1 및 제2 도메인은 DNA 결합 도메인, RNA 결합 도메인, RT 도메인, 및 엔도뉴클레아제 도메인으로부터 각각 독립적으로 선택된다.

[0845] 인테인은 자기-스플라이싱 단백질 인트론(예를 들어, 펩티드)으로서 발생할 수 있으며, 예를 들어 측접하는 N-말단과 C-말단 엑스테인(extein)(예를 들어, 연결될 단편)을 결찰한다. 인테인은 일부 경우에서, 그 자신을 절제하고 남아있는 단편(엑스테인)을 단백질 스플라이싱으로 알려진 과정에서 펩티드 결합을 이용하여 연결할 수 있는 단백질 단편을 포함한다. 인테인은 또한 "단백질 인트론"으로도 지칭된다. 그 자신을 절제하고 단백질의 나머지 부분을 연결하는 인테인의 과정은 본 명세서에서 "단백질 스플라이싱" 또는 "인테인-매개 단백질 스플라이싱"으로 명명된다.

[0846] 일부 구현예에서, 전구체 단백질의 인테인(인테인-매개 단백질 스플라이싱 전에 단백질을 함유하는 인테인)은 2개의 유전자로부터 유래한다. 그러한 인테인은 본 명세서에서 스플릿 인테인(예를 들어, 스플릿 인테인-N 및 스플릿 인테인-C)으로 지칭된다. 따라서, 인테인-기반 접근법은 제1 폴리펩티드 서열과 제2 폴리펩티드 서열을 함께 연결하는 데 사용될 수 있다. 예를 들어, 시아노박테리아에서, DNA 중합효소 III의 촉매 서브유닛인 DnaE는 2개의 별도의 유전자 dnaE-n 및 dnaE-c에 의해 인코딩된다. dnaE-n 유전자에 의해 인코딩되는 인테인-N-도메인은 제1 폴리펩티드 서열의 일부로서 위치할 때 제1 폴리펩티드 서열을 제2 폴리펩티드 서열과 연결할 수 있으며, 여기서 제2 폴리펩티드 서열은 dnaE-c 유전자에 의해 인코딩되는 것과 같은 인테인-C 도메인을 포함한다. 따라서, 일부 구현예에서, 단백질은 제1 및 제2 폴리펩티드 서열을 인코딩하는 핵산을 제공함으로써 제조될 수 있으며(예를 들어, 여기서 제1 핵산 분자는 제1 폴리펩티드 서열을 인코딩하고 제2 핵산 분자는 제2 폴리펩티드 서열을 인코딩함), 핵산은 제1 및 제2 폴리펩티드 서열의 생산을 가능하게 하고 인테인-기반 메커니즘을 통해 제1 폴리펩티드 서열의 제2 폴리펩티드 서열로의 연결을 가능하게 하는 조건 하에서 세포 내로 도입된다.

[0847] 이종성 단백질 단편을 연결하기 위한 인테인의 사용은, 예를 들어 문헌[Wood et al., J. Biol. Chem.289(21); 14512-9 (2014) (그 전체가 본 명세서에 참고로 포함됨)]에 기재되어 있다. 예를 들어, 별도의 단백질 단편에 융합될 때, 인테인 IntN 및 IntC는 서로 인식하고/하거나, 그들 자신을 스플라이스로 제거하고/하거나 그들이 융합된 단백질 단편의 측접하는 N-말단 및 C-말단 엑스테인을 동시에 결찰하여, 이에 의해 2개의 단백질 단편으로부터 전장 단백질을 재구성할 수 있다.

[0848] 일부 구현예에서, dnaE 인테인, Cfa-N(예를 들어, 스플릿 인테인-N) 및 Cfa-C(예를 들어, 스플릿 인테인-C) 인테인 쌍에 기반한 합성 인테인이 사용된다. 그러한 인테인의 예는, 예를 들어 문헌[Stevens et al., J Am Chem Soc. 2016 Feb. 24; 138(7):2162-5(그 전체가 본 명세서에 참고로 포함됨)]에 기재되어 있다. 본 개시내용에 따라 사용될 수 이는 인테인 쌍의 비제한적인 예는 다음을 포함한다: Cfa DnaE 인테인, Ssp GyrB 인테인, Ssp DnaX 인테인, Ter DnaE3 인테인, Ter ThyX 인테인, Rma DnaB 인테인 및 Cne Prp8 인테인(예를 들어, 본 명세서에 참고로 포함된 미국 특허 8,394,604에 기재된 바와 같음).

[0849] 스플릿 Cas9를 포함하는 일부 구현예에서, 인테인-N 및 인테인-C는 스플릿 Cas9의 N-말단 부분과 스플릿 Cas9의 C-말단 부분의 연결을 위해, 스플릿 Cas9의 N-말단 부분과 스플릿 Cas9의 C-말단 부분에 각각 융합될 수 있다. 예를 들어, 일부 구현예에서, 인테인-N은 스플릿 Cas9의 N-말단 부분의 C-말단에 융합되어, 즉 N-스플릿 Cas9의 N-말단 부분]-[인테인-N]~ C의 구조를 형성한다. 일부 구현예에서, 인테인-C는 스플릿 Cas9의 C-말단 부분의 N-말단에 융합되어, 즉 N-[인테인-C]~ [스플릿 Cas9의 C-말단 부분]-C의 구조를 형성한다. 인테인이 융합되는 단백질(예를 들어, 스플릿 Cas9)의 연결을 위한 인테인-매개 단백질 스플라이싱의 메커니즘은 문헌[Shah et al., Chem Sci. 2014; 5(1):446-461(본 명세서에 참고로 포함됨)]에 기재되어 있다. 인테인을 설계 및 사용하는 방법은 당업계에 알려져 있고, 예를 들어 WO2020051561, WO2014004336, WO2017132580, US20150344549, 및 US20180127780에 기재되어 있으며, 이들 각각은 그 전체가 본 명세서에 참고로 포함된다.

[0850] 일부 구현예에서, 분할은 둘 이상의 단편으로의 분열을 지칭한다. 일부 구현예에서, 스플릿 Cas9 단백질 또는 스플릿 Cas9는, 2개의 별도의 뉴클레오티드 서열에 의해 인코딩된 N-말단 단편 및 C-말단 단편으로서 제공되는 Cas9 단백질을 포함한다. Cas9 단백질의 N-말단 부분 및 C-말단 부분에 상응하는 폴리펩티드는 스플라이싱되어 재구성된 Cas9 단백질을 형성할 수 있다. 구현예에서, Cas9 단백질은 단백질의 정돈되지 않은(disorderd) 영역 내에서, 예를 들어 문헌[Nishimasu et al., Cell, Volume 156, Issue 5, pp. 935-949, 2014]에 기재된 바와 같이, 또는 문헌[Jiang et al. (2016) Science 351: 867-871 및 PDB file]에 기재된 바와 같이(이들 각각은 그 전체가 본 명세서에 참고로 포함됨), 2개의 단편으로 분열된다. 정돈되지 않은 영역은, 제한없이 X-선

결정분석, NMR 분광법, 전자 현미경(예를 들어, cryoEM) 및/또는 인실리코(*in silico*) 단백질 모델링을 포함하여 당업계에서 알려진 하나 이상의 단백질 구조 결정 기법에 의해 결정될 수 있다. 일부 구현예에서, 단백질은 임의의 C, T, A 또는 S에서, 예를 들어 아미노산 A292 내지 G364, F445 내지 K483, 또는 E565 내지 T637 간의 spCas9 영역 내에서, 또는 임의의 기타 Cas9, Cas9 변이체(예를 들어, nCas9, dCas9) 또는 다른 napDNAbp에서 상응하는 위치에서, 2개의 단편으로 분열된다. 일부 구현예에서, 단백질은 SpCas9 T310, T313, A456, S469, 또는 C574에서 2개의 단편으로 분열된다. 일부 구현예에서, 단백질을 2개의 단편으로 나누는 과정은 단백질 분할(spitting)로 지칭된다.

[0851] 일부 구현예에서, 단백질 단편은 길이가 약 2 내지 1000개 아미노산(예를 들어, 2 내지 10, 10 내지 50, 50 내지 100, 100 내지 200, 200 내지 300, 300 내지 400, 400 내지 500, 500 내지 600, 600 내지 700, 700 내지 800, 800 내지 900, 또는 900 내지 1000개 아미노산)의 범위이다. 일부 구현예에서, 단백질 단편은 길이가 약 5 내지 500개 아미노산(예를 들어, 5 내지 10, 10 내지 50, 50 내지 100, 100 내지 200, 200 내지 300, 300 내지 400, 또는 400 내지 500개 아미노산)의 범위이다. 일부 구현예에서, 단백질 단편은 길이가 약 20 내지 200개 아미노산(예를 들어, 20 내지 30, 30 내지 40, 40 내지 50, 50 내지 100, 또는 100 내지 200개 아미노산)의 범위이다.

[0852] 일부 구현예에서, 유전자 변형 폴리펩티드의 일부 또는 단편은 인테인에 융합된다. 뉴클레아제는 인테인의 N-말단 또는 C-말단에 융합될 수 있다. 일부 구현예에서, 융합 단백질의 일부 또는 단편은 인테인에 융합되고 AAV 캡시드 단백질에 융합된다. 인테인, 뉴클레아제 및 캡시드 단백질은 임의의 배열로(예를 들어, 뉴클레아제-인테인-캡시드, 인테인-뉴클레아제-캡시드, 캡시드-인테인-뉴클레아제 등) 함께 융합될 수 있다. 일부 구현예에서, 인테인의 N-말단은 융합 단백질의 C-말단에 융합되고, 인테인의 C-말단은 AAV 캡시드 단백질의 N-말단에 융합된다

[0853] 일부 구현예에서, 엔도뉴클레아제 도메인(예를 들어, 닉카제 Cas9 도메인)은 인테인-N에 융합되고, RT 도메인을 포함하는 폴리펩티드는 인테인-C에 융합된다.

[0854] 인테인-N 도메인 및 적합한 인테인 C-도메인의 예시적인 뉴클레오티드 및 아미노산 서열이 하기에 제공되어 있다:

[0855] DnaE 인테인-N DNA:

TGCCTGTCATACGAAACCGAGATACTGACAGTAGAATATGGCCTTCTGCCAATCGGGAAGAT
TGTGGAGAAACGGATAGAATGCACAGTTTACTCTGTCGATAACAATGGTAACATTTATACTC
AGCCAGTTGCCCAGTGGCACGACCGGGGAGAGCAGGAAGTATTCGAATACTGTCTGGAGGA
TGGAAGTCTCATTAGGGCCACTAAGGACCACAAATTTATGACAGTCGATGGCCAGATGCTGC
CTATAGACGAAATCTTTGAGCGAGAGTTGGACCTCATGCGAGTTGACAACCTTCCTAAT

[0856] (서열번호 29)

[0857] DnaE 인테인-N 단백질:

CLSYETEILTVEYGLLPIGKIVEKRIECTVYSVDNNGNIYTQPVAQWHDRGEQEVFEYCLEDGSLI

[0858] RATKDHKFMTVDGQMLPIDEIFERELDLMRVDNLPN (서열번호 30)

[0859] DnaE 인테인-C DNA:

ATGATCAAGATAGCTACAAGGAAGTATCTTGGCAAACAAAACGTTTATGATATTGGAGTCG

[0860] AAAGAGATCACAACTTTGCTCTGAAGAACGGATTCATAGCTTCTAAT (서열번호 31)

[0861] DnaE 인테인-C 단백질:

[0862] MIKIATRKYLGKQNVYDIGVERDHNFALKNGFIASN (서열번호 32)

[0863] Cfa-N DNA:

TGCCTGTCTTATGATACCGAGATACTTACCGTTGAATATGGCTTCTTGCCTATTGGAAAGATT
GTCGAAGAGAGAATTGAATGCACAGTATATACTGTAGACAAGAATGGTTTCGTTTACACACA
GCCCATTGCTCAATGGCACAATCGCGGCGAACAAGAAGTATTTGAGTACTGTCTCGAGGATG
GAAGCATCATACGAGCAACTAAAGATCATAAATTCATGACCACTGACGGGCAGATGTTGCC
AATAGATGAGATATTCGAGCGGGGCTTGGATCTCAAACAAGTGGATGGATTG CCA (서열번호

[0864] 33)

[0865] Cfa-N 단백질:

[0866] CLSYDTEILTVEYGFLPIGKIVEERIECTVYTVDKNGFVYTQPIAQWHNRGEQEVFEYCLEDGSIIR ATKDHKFMTTDGQMLPIDEIFERGLDLKQVDGLP (서열번호 34)

[0867] Cfa-C DNA:

[0868] ATGAAGAGGACTGCCGATGGATCAGAGTTTGAATCTCCCAAGAAGAAGAGGAAAGTAAAGA TAATATCTCGAAAAAGTCTTGGTACCCAAAATGTCTATGATATTGGAGTGGAGAAAGATCAC AACTTCCTTCTCAAGAACGGTCTCGTAGCCAGCAAC (서열번호 35)

[0869] Cfa-C 단백질:

[0870] MKRTADGSEFESPKKKRKVKIISRKSLGTQNVYDIGVEKDHNFLLKNGLVASN (서열번호 36)

[0871] 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드의 RBD는 인테인-기반 융합을 통해, 예를 들어 하기 표 33에 열거된 바와 같은 인테인 이량체화 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 인테인 이량체화 서열)을 통해 RT 도메인에 부착된다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드의 RBD는 인테인-기반 융합을 통해, 예를 들어 하기 표 33에 열거된 바와 같은 인테인 이량체화 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 인테인 이량체화 서열)을 통해 DBD(예를 들어, Cas 도메인, 예를 들어 Cas9 도메인, 예를 들어 nCas9 or dCas9 도메인)에 부착된다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드의 RT 도메인은 인테인-기반 융합을 통해, 예를 들어 하기 표 33에 열거된 바와 같은 인테인 이량체화 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 인테인 이량체화 서열)을 통해 DBD(예를 들어, Cas 도메인, 예를 들어 Cas9 도메인, 예를 들어 nCas9 or dCas9 도메인)에 부착된다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드의 DBD(예를 들어, Cas 도메인, 예를 들어 Cas9 도메인, 예를 들어 nCas9 또는 dCas9 도메인)는 인테인-기반 융합을 통하 RBD 및 RT 도메인에 부착된다. 구현예에서, DBD는 상이한 인테인 이량체화 서열, 예를 들어 하기 표 33에 열거된 바와 같은 인테인 이량체화 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 서열)을 통해 RBD 및 RT 도메인에 부착된다. 구현예에서, DBD는 동일한 인테인 이량체화 서열, 예를 들어 하기 표 33에 열거된 바와 같은 인테인 이량체화 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 아미노산 서열을 포함하는 서열)을 통해 RBD 및 RT 도메인에 부착된다. 일부 구현예에서, 서로 결합될 RBD 및 DBD의 인테인 이량체화 서열은 각각 하기 표 33의 단일 행에 열거된 바와 같은 사슬 A 서열 및 사슬 B 서열, 또는 각각 사슬 B 및 사슬 A 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열)을 포함한다. 일부 구현예에서, 서로 결합될 RBD 및 RT 도메인의 인테인 이량체화 서열은 각각 하기 표 33의 단일 행에 열거된 바와 같은 사슬 A 서열 및 사슬 B 서열, 또는 각각 사슬 B 및 사슬 A 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열)을 포함한다. 일부 구현예에서, 서로 결합될 RT 도메인 및 DBD의 인테인 이량체화 서열은 각각 하기 표 33의 단일 행에 열거된 바와 같은 사슬 A 서열 및 사슬 B 서열, 또는 각각 사슬 B 및 사슬 A 서열(또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 서열)을 포함한다.

[0872] 표 33. 예시적인 인테인 이량체화 서열

시스템	사슬 A 명칭	사슬 A 서열	예시적인 사슬 A 공급원	사슬 B 명칭	사슬 B 서열	예시적인 사슬 B 공급원
Sce-VMA	Sce-VMA-5'-v1	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPILYENDHFFDYMQK SKFHLTIEGPKVLAYLLGLWIGDGL SDRATFSVDSRDTSLMERVTEYAE KLNLCAEYKDRKEPQVAKTVNLYS KVVRG	snapgene	Sce-VMA-3'-v1	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQN	snapgene common features
Sce-VMA	Sce-VMA-5'-v1	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPILYENDHFFDYMQK SKFHLTIEGPKVLAYLLGLWIGDGL SDRATFSVDSRDTSLMERVTEYAE KLNLCAEYKDRKEPQVAKTVNLYS KVVRG	snapgene	Sce-VMA-3'-v2	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGERGNGSG	snapgene common features

[0873]

[0874]

Sce-VMA	Sce-VMA-5'-v1	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPILYENDHFFDYMQK SKFHLTIEGPKVLAYLLGLWIGDGL SDRATFSVDSRDTSLMERVTEYAE KLNLCAEYKDRKEPQVAKTVNLYS KVVRG	snapgene	Sce-VMA-3'-v3	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCTMTEKGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v1	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPILYENDHFFDYMQK SKFHLTIEGPKVLAYLLGLWIGDGL SDRATFSVDSRDTSLMERVTEYAE KLNLCAEYKDRKEPQVAKTVNLYS KVVRG	snapgene	Sce-VMA-3'-v4	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGEKSMGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v1	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPILYENDHFFDYMQK SKFHLTIEGPKVLAYLLGLWIGDGL	snapgene	Sce-VMA-3'-v5	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHN	https://www.nature.com/articles/nmeth.3585

[0875]

		SDRATFSVDSRDTSLMERVTEYAE KLNLCAEYKDRKEPQVAKTVNLYS KVVRG				
Sce-VMA	Sce-VMA-5'-v1	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPILYENDHFFDYMQK SKFHLTIEGPKVLAYLLGLWIGDGL SDRATFSVDSRDTSLMERVTEYAE KLNLCAEYKDRKEPQVAKTVNLYS KVVRG	snapgene	Sce-VMA-3'-v6	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE RGNGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v1	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPILYENDHFFDYMQK SKFHLTIEGPKVLAYLLGLWIGDGL SDRATFSVDSRDTSLMERVTEYAE KLNLCAEYKDRKEPQVAKTVNLYS KVVRG	snapgene	Sce-VMA-3'-v7	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCTM TEKGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v1	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR	snapgene	Sce-VMA-3'-v8	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE KSMGSG	https://www.nature.com/articles/nmeth.3585

[0876]

		KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPILYENDHFFDYMQK SKFHLTIEGPKVLAYLLGLWIGDGL SDRATFSVDSRDTSLMERVTEYAE KLNLCAEYKDRKEPQVAKTVNLYS KVVRG				
Sce-VMA	Sce-VMA-5'-v2	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG snapgene common features	snapgene	Sce-VMA-3'-v1	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQN	snapgene common features
Sce-VMA	Sce-VMA-5'-v2	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG snapgene common features	snapgene	Sce-VMA-3'-v2	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGERGNGSG	snapgene common features

[0877]

Sce-VMA	Sce-VMA-5'-v2	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG snapgene common features	snapgene	Sce-VMA-3'-v3	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCTMTEKGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v2	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG snapgene common features	snapgene	Sce-VMA-3'-v4	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGEKSMGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v2	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD	snapgene	Sce-VMA-3'-v5	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHN	https://www.nature.com/articles/nmeth.3585

[0878]

		LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG snapgene common features				
Sce-VMA	Sce-VMA-5'-v2	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG snapgene common features	snapgene	Sce-VMA-3'-v6	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE RGNGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v2	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG snapgene common features	snapgene	Sce-VMA-3'-v7	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCTM TEKGSG	https://www.nature.com/articles/nmeth.3585

[0879]

Sce-VMA	Sce-VMA-5'-v2	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG snapgene common features	snapgene	Sce-VMA-3'-v8	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE KSMGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v3	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPILYEND HFFDYMQKSKFHLTIEGPKVLAYLL GLWIGDGLSDRATFSVDSRDTSLM ERVTEYAEKLNLCAEYKDRKEPQV AKTVNLYSKVVRG	snapgene	Sce-VMA-3'-v1	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQN	snapgene common features
Sce-VMA	Sce-VMA-5'-v3	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPILYEND	snapgene	Sce-VMA-3'-v2	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA	snapgene common features

[0880]

		HFFDYMQKSKFHLTIEGPKVLAYLL GLWIGDGLSDRATFSVDSRDTSLM ERVTEYAEKLNLCAEYKDRKEPQV AKTVNLYSKVVRG				RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGERGNGSG	
Sce-VMA	Sce-VMA-5'-v3	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPILYEND HFFDYMQKSKFHLTIEGPKVLAYLL GLWIGDGLSDRATFSVDSRDTSLM ERVTEYAEKLNLCAEYKDRKEPQV AKTVNLYSKVVRG	snapgene	Sce-VMA-3'-v3	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCTMTEKGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v3	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPILYEND HFFDYMQKSKFHLTIEGPKVLAYLL GLWIGDGLSDRATFSVDSRDTSLM ERVTEYAEKLNLCAEYKDRKEPQV AKTVNLYSKVVRG	snapgene	Sce-VMA-3'-v4	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGEKSMGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v3	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK	snapgene	Sce-VMA-3'-v5	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHN	https://www.nature.com/articles/nmeth.3585

[0881]

		APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPILYEND HFFDYMQKSKFHLTIEGPKVLAYLL GLWIGDGLSDRATFSVDSRDTSLM ERVTEYAEKLNLCAEYKDRKEPQV AKTVNLYSKVVRG				
Sce-VMA	Sce-VMA-5'-v3	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPILYEND HFFDYMQKSKFHLTIEGPKVLAYLL GLWIGDGLSDRATFSVDSRDTSLM ERVTEYAEKLNLCAEYKDRKEPQV AKTVNLYSKVVRG	snapgene	Sce-VMA-3'-v6	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE RGNGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v3	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPILYEND HFFDYMQKSKFHLTIEGPKVLAYLL GLWIGDGLSDRATFSVDSRDTSLM ERVTEYAEKLNLCAEYKDRKEPQV AKTVNLYSKVVRG	snapgene	Sce-VMA-3'-v7	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCTM TEKGSG	https://www.nature.com/articles/nmeth.3585

[0882]

Sce-VMA	Sce-VMA-5'-v3	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPILYEND HFFDYMQKSKFHLTIEGPKVLAYLL GLWIGDGLSDRATFSVDSRDTSLM ERVTEYAEKLNLCAEYKDRKEPQV AKTVNLYSKVVRG	snapgene	Sce-VMA-3'-v8	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE KSMGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v4	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG	snapgene	Sce-VMA-3'-v1	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQN	snapgene common features
Sce-VMA	Sce-VMA-5'-v4	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG	snapgene	Sce-VMA-3'-v2	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD	snapgene common features

[0883]

		LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG			YYGITLSDDSDHQFLLGS QVVVQNCGERGNGSG	
Sce-VMA	Sce-VMA-5'-v4	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG	snapgene	Sce-VMA-3'-v3	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCTMTEKGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v4	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG	snapgene	Sce-VMA-3'-v4	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGEKSMGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v4	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA	snapgene	Sce-VMA-3'-v5	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHN	https://www.nature.com/articles/nmeth.3585

[0884]

		NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG				
Sce-VMA	Sce-VMA-5'-v4	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG	snapgene	Sce-VMA-3'-v6	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE RGNGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v4	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG	snapgene	Sce-VMA-3'-v7	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCTM TEKGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v4	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR	snapgene	Sce-VMA-3'-v8	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS	https://www.nature.com/articl

[0885]

		EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPILYENDH FFDYMQKSKFHLTIEGPKVLAYLLG LWIGDGLSDRATFSVDSRDTSLME RVTEYAEKLNLCAEYKDRKEPQVA KTVNLYSKVVRG				DHQFLLANQVVVHNCGE KSMGSG	es/nmeth.3585
Sce-VMA	Sce-VMA-5'-v5	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v1	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQN	snapgene common features
Sce-VMA	Sce-VMA-5'-v5	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v2	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGERGNGSG	snapgene common features

[0886]

Sce-VMA	Sce-VMA-5'-v5	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v3	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCTMTEKGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v5	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v4	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGEKSMGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v5	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v5	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHN	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v5	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF	https://www.nature.com/articles/	Sce-VMA-3'-v6	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS	https://www.nature.com/articl

		TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPI	es/nmeth. 3585		DHQFLLANQVVVHNCGE RGNGSG	es/nmeth. 3585
Sce- VMA	Sce- VMA-5'- v5	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPI	https://w ww.nature .com/articl es/nmeth. 3585	Sce- VMA-3'- v7	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCTM TEKGSG	https://w ww.nature .com/articl es/nmeth. 3585
Sce- VMA	Sce- VMA-5'- v5	CFAKGTNVLMADGSIECIENIEVGN KVMGKDGRPREVIKLPRGRETMYS VVQKSQHRAHKSDSSREVPELLKF TCNATHELVVRTPRSVRRLSRTIKG VEYFEVITFEMGQKKAPDGRIVELV KEVSKSYPISEGPERANELVESYR KASNKAYFEWTIEARDLSLLGSHV RKATYQTYAPI	https://w ww.nature .com/articl es/nmeth. 3585	Sce- VMA-3'- v8	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE KSMGSG	https://w ww.nature .com/articl es/nmeth. 3585
Sce- VMA	Sce- VMA-5'- v6	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://w ww.nature .com/articl es/nmeth. 3585	Sce- VMA-3'- v1	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQN	snapgene common features

[0887]

[0888]

Sce-VMA	Sce-VMA-5'-v6	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v2	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGERGNGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v6	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v3	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCTMTEKGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v6	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v4	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGEKSMGSG	snapgene common features

[0889]

Sce-VMA	Sce-VMA-5'-v6	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v5	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHN	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v6	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v6	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE RGNGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v6	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v7	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCTM TEKGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v6	GGIIYVGCFAKGTNVLMADGSIECI ENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v8	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE KSMGSG	https://www.nature.com/articles/nmeth.3585

[0890]

Sce-VMA	Sce-VMA-5'-v7	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v1	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQN	snapgene common features
Sce-VMA	Sce-VMA-5'-v7	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v2	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGERGNGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v7	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v3	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCTMTEKGSG	snapgene common features

[0891]

Sce-VMA	Sce-VMA-5'-v7	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v4	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGEKSMGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v7	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v5	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHN	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v7	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v6	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE RGNGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v7	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v7	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCTM TEKGSG	https://www.nature.com/articles/nmeth.3585

[0892]

		ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPI				
Sce-VMA	Sce-VMA-5'-v7	GGVVLEKGCFAKGTNVLMADGSIE CIENIEVGNKVMGKDGRPREVIKLP RGRETMYSVVQKSQHRAHKSDSS REVPELLKFTCNATHELVVRTPRS VRRLSRTIKGVEYFEVITFEMGQKK APDGRIVELVKEVSKSYPISEGPER ANELVESYRKASNKAYFEWTIEAR DLSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v8		https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v8	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v1	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQN	snapgene common features
Sce-VMA	Sce-VMA-5'-v8	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v2	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGERGNGSG	snapgene common features

[0893]

Sce-VMA	Sce-VMA-5'-v8	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v3	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCTMTEKGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v8	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v4	GHGGIRNNLNTENPLWD AIVGLGFLKDGVKNIPSF LSTDNIGTRETFLAGLIDS DGYVTDEHGIKATIKTIHT SVRDGLVSLARSLGLVV SVNAEPAKVDMNVTKHK ISYAIYMSGGDVLLNVLS KCAGSKKFRPAPAAAFA RECRGFYFELQELKEDD YYGITLSDDSDHQFLLGS QVVVQNCGEKSMGSG	snapgene common features
Sce-VMA	Sce-VMA-5'-v8	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v5	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHN	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v8	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v6	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS	https://www.nature.com/articl

Sce-VMA		EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585		DHQFLLANQVVVHNCGE RGNGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v8	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v7	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCTM TEKGSG	https://www.nature.com/articles/nmeth.3585
Sce-VMA	Sce-VMA-5'-v8	GGFQTVGCFAKGTNVLMADGSIEC IENIEVGNKVMGKDGRPREVIKLPR GRETMYSVVQKSQHRAHKSDSSR EVPELLKFTCNATHELVVRTPRSV RRLSRTIKGVEYFEVITFEMGQKKA PDGRIVELVKEVSKSYPISEGPERA NELVESYRKASNKAYFEWTIEARD LSLLGSHVRKATYQTYAPI	https://www.nature.com/articles/nmeth.3585	Sce-VMA-3'-v8	VLLNVLSKCAGSKKFRP APAAAFARECRGFYFEL QELKEDDYYGITLSDDS DHQFLLANQVVVHNCGE KSMGSG	https://www.nature.com/articles/nmeth.3585

[0894]

[0895] *추가의 도메인*

[0896] 유전자 변형 폴리펩티드는 표적 DNA 서열 및 주형 핵산(예를 들어, 주형 RNA)에 결합하고, 표적 부위에 닉을 생성하고, 주형을 DNA 내로 작성하여(예를 들어, 역전사), 표적 부위의 변형을 생성할 수 있다. 일부 구현예에서, 추가의 도메인이 과정의 효율을 향상시키기 위해 폴리펩티드에 첨가될 수 있다. 일부 구현예에서, 유전자 변형 폴리펩티드는 역전사된 DNA를 표적 부위의 DNA에 연결시키기 위한 추가의 DNA 결찰 도메인을 함유할 수 있다. 일부 구현예에서, 폴리펩티드는 이종성 RNA-결합 도메인을 포함할 수 있다. 일부 구현예에서, 폴리펩티드는 5'에서 3' 엑소뉴클레아제 활성을 갖는 도메인을 포함할 수 있다(예를 들어, 여기서 5'에서 3' 엑소뉴클레아제 활성은, 예를 들어 원래의 게놈 서열에 대한 변경을 위해, 표적 부위의 변경의 복구를 증가시킴). 일부 구현예에서, 폴리펩티드는 3'에서 5' 엑소뉴클레아제 활성, 예를 들어 교정 활성을 갖는 도메인을 포함할 수 있다. 일부 구현예에서, 라이팅 도메인, 예를 들어 RT 도메인은 3'에서 5' 엑소뉴클레아제 활성, 예를 들어 교정 활성을 갖는다.

[0897] ***주형 핵산***

[0898] 본 명세서에 기재된 유전자 변형 시스템은 주형 핵산 서열을 이용하여 숙주 표적 DNA 부위를 변형시킬 수 있다. 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 표적-프라이밍된 역전사(TPRT)에 의해 RNA 서열 주형을 숙주 표적 DNA 부위로 전사한다. RNA 서열 주형의 숙주 게놈 내로의 역전사를 통해 DNA 서열(들)을 변형시킴으로써, 유전자 변형 시스템은(예를 들어, CRISPR 시스템과 달리) 숙주 세포 내로 도입되는 외인성 DNA 서열에 대한 필요 없이 목적 서열을 표적 게놈 내로 삽입할 뿐만 아니라, 외인성 DNA 삽입 단계를 제거할 수 있다. 유전자 변형 시스템은 표적 게놈으로부터의 서열을 결실 또는 목적 서열을 이용하여 치환을 도입할 수도 있다. 따라서, 유전자 변형 시스템은 목적 서열, 예를 들어 이종성 유전자 코딩 및/또는 기능 정보를 포함하는 서열을 함유하는 맞춤화된 RNA 서열 주형의 이용을 위한 플랫폼을 제공한다.

[0899] 일부 구현예에서, 주형 핵산은 유전자 변형 폴리펩티드에 결합하는 하나 이상의 서열(예를 들어, 2개의 서열)을 포함한다.

[0900] 일부 구현예에서, 주형 RNA는 표 S4에 열거된 바와 같은 핵산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 주형 RNA는 표 S4에 열거된 바와 같은 주형 서열의 5' 말단 블록 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 주형 RNA는 표 S4에 열거된 바와 같은 주형 서열의 PBS 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 주형 RNA는 표 S4에 열거된 바와 같은 주형 서열의 링커 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 주형 RNA는 표 S4에 열거된 바와 같은 주형 서열의 하나 이상(예를 들어, 1, 2, 3, 또는 4개)의 RRS 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 주형 RNA는 표 S4에 열거된 바와 같은 주형 서열의 3' 말단 블록 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 주형 RNA는 (예를 들어, 5'에서 3' 순서로) 표 S4에 열거된 바와 같은 5' 말단 블록 서열, PBS 서열, 하나 이상의 RRS 서열, 및 3' 말단 블록 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다.

[0901] 일부 구현예에서, 본 명세서에 기재된 시스템 또는 방법은 단일 주형 핵산(예를 들어, 주형 RNA)을 포함한다. 일부 구현예에서, 본 명세서에 기재된 시스템 또는 방법은 복수의 주형 핵산(예를 들어, 주형 RNA)을 포함한다. 예를 들어, 본 명세서에 기재된 시스템은 (예를 들어, 5'에서 3'으로) 유전자 변형 폴리펩티드(예를 들어, DNA-결합 도메인 및/또는 엔도뉴클레아제 도메인, 예를 들어 gRNA)에 결합하는 서열 및 표적 부위(예를 들어, 표적 게놈 내 부위의 제2 가닥)에 결합하는 서열을 포함하는 제1 RNA, 및 선택적으로 유전자 변형 폴리펩티드에 결합하는(예를 들어, RT 도메인에 특이적으로 결합하는) 서열, 이종 목적 서열, 및 PBS 서열을 포함하는 제2 RNA(예를 들어, 주형 RNA)를 포함한다. 일부 구현예에서, 시스템이 복수의 핵산을 포함하는 경우, 각각의 핵산은 접합 도메인을 포함한다. 일부 구현예에서, 접합 도메인은, 예를 들어 상보적 서열의 혼성화에 의해 핵산 분자의 회합을 가능하게 한다. 예를 들어, 일부 구현예에서, 제1 RNA는 제1 접합 도메인을 포함하고 제2 RNA는 제2 접합 도메인을 포함하며, 제1 및 제2 접합 도메인은, 예를 들어 엄격한 조건 하에서 서로 혼성화할 수 있다. 일부 구현예에서, 혼성화를 위한 엄격한 조건은 4× 염화나트륨/나트륨 시트레이트(SSC) 중, 약 65C에서의 혼성화에 이어 약 65C에서 1xSSC 내 세척을 포함한다.

[0902] 일부 구현예에서, 주형 핵산은 RNA를 포함한다. 일부 구현예에서, 주형 핵산은 DNA(예를 들어, 단일 가닥 또는 이중 가닥 DNA)를 포함한다.

[0903] 일부 구현예에서, 주형 핵산은 표적 핵산에 대해 상동성을 갖는 하나 이상(예를 들어, 2개)의 상동성 도메인을 포함한다. 일부 구현예에서, 상동성 도메인은 길이가 약 10 내지 20, 20 내지 50, 또는 50 내지 100개 뉴클레오티드이다.

[0904] 일부 구현예에서, 주형 RNA는, 예를 들어 유전자 변형 폴리펩티드를 관심 표적 부위로 지시하기 위한 gRNA 서열을 포함할 수 있다. 일부 구현예에서, 주형 RNA는 (예를 들어, 5'에서 3'으로) (i) 선택적으로 표적 부위에 결합하는 gRNA 스페이서(예를 들어, 표적 게놈 내 부위의 제2 가닥), (ii) 선택적으로 본 명세서에 기재된 폴리펩티드에 결합하는 gRNA 스캐폴드(예를 들어, 유전자 변형 폴리펩티드 또는 Cas 폴리펩티드), (iii) 돌연변이 영역을 포함하는 이종 목적 서열(선택적으로 이종 목적 서열은 5'에서 3'으로 제1 상동성 영역, 돌연변이 영역,

및 제2 상동성 영역을 포함함), 및 (iv) 3' 표적 상동성 도메인을 포함하는 프라이머 결합 부위(PBS) 서열을 포함한다.

[0905] 본 명세서에 기재된 게놈 편집 시스템의 주형 핵산(예를 들어, 주형 RNA) 구성요소는 통상적으로 시스템의 유전자 변형 폴리펩티드에 결합할 수 있다. 일부 구현예에서 주형 핵산(예를 들어, 주형 RNA)은 유전자 변형 폴리펩티드에 결합할 수 있는 3' 영역을 갖는다. 결합 영역, 예를 들어 3' 영역은 시스템의 유전자 변형 폴리펩티드에 결합할 수 있는, 예를 들어 적어도 1, 2 또는 3개의 헤어핀 루프를 갖는, 구조화된 RNA 영역일 수 있다. 결합 영역은 주형 핵산(예를 들어, 주형 RNA)을 임의의 폴리펩티드 모듈과 회합시킬 수 있다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)의 결합 영역은 폴리펩티드에서 RNA-결합 도메인과 회합할 수 있다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)의 결합 영역은 유전자 변형 폴리펩티드의 역전사 도메인과 회합할 수 있다(예를 들어, RT 도메인에 특이적으로 결합할 수 있음). 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 폴리펩티드의 DNA 결합 도메인, 예를 들어 Cas9-유래 DNA 결합 도메인과 회합하는 gRNA와 회합할 수 있다. 일부 구현예에서, 결합 영역은 또한 DNA 표적 인식, 예를 들어 표적 DNA 서열과 혼성화하고 폴리펩티드, 예를 들어 Cas9 도메인에 결합하는 gRNA를 제공할 수 있다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 폴리펩티드의 다중 구성요소, 예를 들어 DNA 결합 도메인 및 역전사 도메인과 회합할 수 있다.

[0906] 일부 구현예에서 주형 RNA는 3' 말단에서 폴리-A 테일을 갖는다. 일부 구현예에서 주형 RNA는 3' 말단에서 폴리-A 테일을 갖지 않는다.

[0907] 일부 구현예에서, 주형 RNA는 표적 세포의 게놈 DNA에서(예를 들어, 대상체에서, 예를 들어 생체 외 또는 생체 내에서, 예를 들어 표적 조직 또는 기관에서) 주어진 돌연변이를 교정하도록 맞춤화될 수 있다. 예를 들어, 돌연변이는 야생형 서열에 비해 질환-연관 돌연변이일 수 있다. 이론에 구속되고자 함이 없이, 임의의 주어진 표적 부위 및 편집물은 다양한 편집 효율 및 충실도를 초래할 유전자 변형 시스템에 사용하기 위한 다수의 가능한 주형 RNA 분자를 가질 것이다. 이러한 스크리닝 부담을 부분적으로 감소시키기 위해, 경험적 파라미터의 세트는 주형 RNA 또는 이의 일부의 최적의 초기 인실리코 설계를 보장하는 데 도움이 된다. 비제한적인 예시적 예로서, 선택된 돌연변이에 대해 다음 설계 파라미터가 이용될 수 있다. 일부 구현예에서, 설계는 표적 영역의 역할을 하는 돌연변이의 양쪽에서 서열에 측접하는 대략 500 bp(예를 들어, 최대 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, or 700 bp, 및 선택적으로 적어도 20, 30, 40, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 또는 650 bp)를 획득함으로써 개시된다. 일부 구현예에서, 주형 핵산은 gRNA를 포함한다. 일부 구현예에서, gRNA는 표적 부위에 결합하는 서열(예를 들어, CRISPR 스페이서)을 포함한다. 일부 구현예에서, 주형 핵산을 표적 영역으로 표적화하는 데 사용하기 위한 표적 부위에 결합하는 서열(예를 들어, CRISPR 스페이서)은 사용되는 특정 유전자 변형 폴리펩티드(예를 들어, CRISPR/Cas 도메인을 포함하는, 예를 들어 엔도뉴클레아제 도메인 또는 라이팅 도메인)를 고려함으로써 선택된다(예를 들어, Cas9의 경우, 20개 뉴클레오티드 gRNA 결합 영역의 바로 3'에 있는 NGG의 프로토스페이서-인접 모티프(PAM)). 일부 구현예에서, CRISPR 스페이서는 먼저 PAM이 유전자 변형 시스템 유도 편집에 의해 붕괴될 것인지 여부에 의해 순위를 매김으로써 선택된다. 일부 구현예에서, PAM의 붕괴는 편집 효율성을 증가시킬 것이다. 일부 구현예에서, PAM은 유전자 변형 동안 표적 부위에서 침묵 돌연변이(예를 들어, 존재하는 경우, 표적 핵산 서열에 의해 인코딩된 아미노산 잔기를 변경하지 않는 돌연변이)를 (예를 들어, 게놈 DNA에서 표적 부위에 또 다른 변형의 일부 또는 첨가로서) 도입함에 의해서도 붕괴될 수 있다. 일부 구현예에서, CRISPR 스페이서는 원하는 편집 위치에 대한 그들의 상응하는 게놈 부위 인접에 의해 서열의 순위를 매김으로써 선택된다. 일부 구현예에서, gRNA는 gRNA 스캐폴드를 포함한다. 일부 구현예에서, 사용된 gRNA 스캐폴드는 표준 스캐폴드(예를 들어, Cas9의 경우, 5'-GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC-3')일 수 있거나, 하나 이상의 뉴클레오티드 치환을 함유할 수 있다. 일부 구현예에서, 이종 목적 서열은 유전자 변형에 의해 표적 부위 내로 작성될 수 있는 임의의 삽입, 치환 또는 결실을 제외하고, 적어도 90% 동일성, 예를 들어, 적어도 90%, 95%, 98%, 99%, 또는 100% 동일성을 갖거나, 또는 제1 가닥 닉의 3'(예를 들어, 제1 가닥 닉의 바로 3' 또는 제1 가닥 닉의 3'에서 최대 1, 2, 3, 4, 또는 5개의 뉴클레오티드)에서 표적 부위에 대해 동일성이 없는 1, 2, 3, 4, 또는 5개 이하의 위치를 포함한다. 일부 구현예에서, 3' 표적 상동성 도메인은 적어도 90% 동일성, 예를 들어, 적어도 90%, 95%, 98%, 99%, 또는 100% 동일성을 함유하거나, 또는 제1 가닥 닉의 5'(예를 들어, 제1 가닥 닉의 바로 5' 또는 제1 가닥 닉의 3'에서 최대 1, 2, 3, 4, 또는 5개의 뉴클레오티드)에서 표적 부위에 대해 동일성이 없는 1, 2, 3, 4, 또는 5개 이하의 위치를 포함한다.

[0908] 일부 구현예에서, 주형 핵산은 주형 RNA이다. 일부 구현예에서, 주형 RNA는 하나 이상의 변형된 뉴클레오티드를 포함한다. 예를 들어, 일부 구현예에서, 주형 RNA는 하나 이상의 데옥시리보뉴클레오티드를 포함한다. 일부 구

현예에서, 주형 RNA의 영역은 DNA 뉴클레오티드에 의해 대체되어, 예를 들어 분자의 안정성을 향상시킨다. 예를 들어, 주형의 3' 말단은 DNA 뉴클레오티드를 포함할 수 있는 한편, 주형의 나머지는 역전사될 수 있는 RNA 뉴클레오티드를 포함한다. 예를 들어 일부 구현예에서, 이종 목적 서열은 주로 또는 전체적으로 RNA 뉴클레오티드(예를 들어, 적어도 90%, 95%, 98%, 또는 99% RNA 뉴클레오티드)로 구성된다. 일부 구현예에서, PBS 서열은 주로 또는 전체적으로 DNA 뉴클레오티드(예를 들어, 적어도 90%, 95%, 98%, 또는 99% DNA 뉴클레오티드)로 구성된다. 다른 구현예에서, 게놈 내로 작성하기 위한 이종 목적 서열은 DNA 뉴클레오티드를 포함할 수 있다. 일부 구현예에서, 주형 내 DNA 뉴클레오티드는 DNA-의존성 DNA 중합효소 활성이 가능한 도메인에 의해 게놈 내로 카피된다. 일부 구현예에서, DNA-의존성 DNA 중합효소 활성은 폴리펩티드 내 DNA 중합효소 도메인에 의해 제공된다. 일부 구현예에서, DNA-의존성 DNA 중합효소 활성은 또한 DNA-의존성 DNA 중합, 예를 들어 제2 가닥 합성이 가능한 역전사효소 도메인에 의해 제공된다. 일부 구현예에서, 주형 분자는 DNA 뉴클레오티드만으로 구성된다.

[0909] 일부 구현예에서, 본 명세서에 기재된 시스템은 본 명세서에 기재된 주형 RNA의 서열을 함께 포함하는 2개의 핵산을 포함한다. 일부 구현예에서, 2개의 핵산은 서로 비공유적으로 회합되어 있으며, 예를 들어 (예를 들어, 염기쌍 형성을 통해) 서로 직접적으로 회합되어 있거나, 하나 이상의 추가 분자를 포함하는 복합체의 일부로서 간접적으로 회합되어 있다.

[0910] 본 명세서에 기재된 주형 RNA는 5'에서 3'으로 (1) gRNA 스페이서; (2) gRNA 스캐폴드; (3) 이종 목적 서열 (4) 프라이머 결합 부위(PBS) 서열을 포함할 수 있다. 이들 구성요소 각각은 이제 보다 상세히 기재된다.

[0911] *gRNA 스페이서 및 gRNA 스캐폴드*

[0912] 본 명세서에 기재된 주형 RNA는 유전자 변형 시스템을 표적 핵산으로 지시하는 gRNA 스페이서, 및 주형 RNA와 유전자 변형 폴리펩티드의 Cas 도메인의 회합을 촉진하는 gRNA 스캐폴드를 포함할 수 있다. 본 명세서에 기재된 시스템은 또한 주형 핵산의 일부가 아닌 gRNA를 포함할 수 있다. 예를 들어, gRNA 스페이서 및 gRNA 스캐폴드를 포함하지만, 이종 목적 서열 또는 PBS 서열을 포함하지 않는 gRNA는, 예를 들어 표적 핵산의 풀림을 촉진하거나 숙주 세포에 의한 원하는 편집의 MMR 역전을 감소시키거나(예를 들어, 본 명세서의 말단 블록 서열 및 추가 가이드 RNA 섹션에 기재된 바와 같음), 예를 들어 본 명세서에서 "제2 가닥 닉 생성"라는 표제의 섹션에 기재된 바와 같이 제2 가닥 닉 생성을 유도하는 데 사용될 수 있다.

[0913] 일부 구현예에서, gRNA는 CRISPR-연관 단백질 결합에 참여하는 스캐폴드 서열, 및 게놈 표적을 위한 사용자-정의된 약 20개의 뉴클레오티드 표적화 서열로 구성되는 짧은 합성 RNA이다. 완전한 gRNA의 구조는 문헌[Nishimasu et al. Cell 156, P935-949 (2014)]에 의해 설명되었다. gRNA(단일-가이드 RNA의 경우 sgRNA로도 지칭됨)는 인공 테트라루프에 의해 연결된 crRNA- 및 tracrRNA-유래 서열로 이루어진다. crRNA 서열은 가이드(20 nt) 및 반복(12 nt) 영역으로 나뉠 수 있는 반면, tracrRNA 서열은 항-반복(14 nt) 및 3개의 tracrRNA 스템 루프로 나뉠 수 있다(Nishimasu et al. Cell 156, P935-949 (2014)). 실제로, 가이드 RNA 서열은 17 내지 24개의 뉴클레오티드(예를 들어, 19, 20, 또는 21개의 뉴클레오티드) 길이를 갖고 표적화된 핵산 서열에 상보적이 되도록 일반적으로 설계된다. 효과적인 가이드 RNA의 설계에 사용하기 위한 맞춤형 gRNA 생성기 및 알고리즘이 상업적으로 구매 가능하다. 일부 구현예에서, gRNA는 천연 CRISPR 시스템 유래의 2개의 RNA 구성요소, 예를 들어 crRNA 및 tracrRNA를 포함한다. 당업계에 잘 알려진 바와 같이, gRNA는 또한 (뉴클레아제에 결합하기 위한) tracrRNA 및 (편집/결합을 위해 뉴클레아제를 표적화된 서열로 가이드하는) 적어도 하나의 crRNA 둘 다로부터의 서열을 함유하는 키메라 단일 가이드 RNA(sgRNA)를 포함할 수 있다. 화학적으로 변형된 sgRNA가 또한 CRISPR-연관 단백질과 함께 사용하기에 효과적인 것으로 입증되었으며; 예를 들어, 문헌[Hendel et al. (2015) Nature Biotechnol., 985 - 991]을 참조한다. 일부 구현예에서, gRNA 스페이서는 표적 유전자와 연관된 DNA 서열에 상보적인 핵산 서열을 포함한다.

[0914] gRNA를 포함하는 주형 핵산, 예를 들어 주형 RNA의 영역은 표적 DNA에 결합된 gRNA의 덜 감긴(underwound) 리본-유사 구조를 채용한다(예를 들어, 문헌[Mulepati et al. Science 19 Sep 2014:Vol. 345, Issue 6203, pp. 1479-1484]에 기재된 바와 같음) 이론에 구속되고자 함이 없이, 이러한 비-정규 구조는 RNA-DNA 하이브리드 중 매 6 번째 뉴클레오티드의 회전에 의해 용이하게 되는 것으로 생각된다. 따라서, 일부 구현예에서, gRNA를 포함하는 주형 핵산, 예를 들어 주형 RNA의 영역은 어느 정도의 간격, 예를 들어 매 6 번째 염기에서 표적 부위와의 증가된 미스매칭을 용인할 수 있다. 일부 구현예에서, 표적 부위에 대해 상동성을 포함하는 gRNA를 포함하는 주형 핵산, 예를 들어 주형 RNA의 영역은 규칙적인 간격, 예를 들어 매 6 번째 염기에서 동요 위치를 가질 수 있으며, 이는 표적 부위와의 염기쌍을 필요로 하지 않는다.

[0915] 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은, 예를 들어 유전자 변형 폴리펩티드의 Cas9도메인(표 8)에

적절한 길이의 gRNA 스페이서 서열을 포함하는, 예를 들어 5' 말단에서, 표적 부위에 대하여 적어도 80%, 85%, 90%, 95%, 99%, 또는 100% 상동성의 적어도 15, 16, 17, 18, 19, 20, 21, 22, 23, 또는 24개 염기를 갖는다.

[0916] 표 12는 유전자 변형을 위해 표 8에 열거된 Cas 변이체에 적용하기 위하여 gRNA 및/또는 주형 RNA를 설계하기 위한 구성요소를 정의하는 파라미터를 제공한다. 절단 부위는 검증되거나 예측된 프로토스페이서 인접 모티프(PAM) 요건, (PAM 부위의 가장 상류 염기에 대해) 절단 부위의 검증되거나 예측된 위치를 나타낸다. 주어진 효소에 대한 gRNA는 crRNA, 테트라루프, 및 tracrRNA 서열을 연쇄화(concatenating)시키고, 표적 부위에서 프로토스페이서와 일치하는 스페이서(최소) 및 스페이서(최대) 내 길이의 5' 스페이서를 추가로 첨가함으로써 조립될 수 있다. 또한, 표적에서 ssDNA의 예측된 위치는, 표적 프라이밍된 역전사를 개시하기 위해, 닉의 바로 5'에서 서열에 어닐링할 수 있는 주형 RNA의 PBS 서열을 설계하는데 중요하다. 일부 구현예에서, 본 명세서에 기재된 gRNA 스캐폴드는 5'에서 3' 방향으로, 표 12의 crRNA, 표 12의 동일한 행의 테트라루프, 및 표 12의 동일한 행의 tracrRNA를 포함하는 핵산 서열, 또는 이에 대하여 적어도 70%, 80%, 85%, 90%, 95%, 또는 99% 동일성을 갖는 서열을 포함한다. 일부 구현예에서, 스캐폴드를 포함하는 gRNA 또는 주형 RNA는 표 12의 동일한 행에 나타낸 스페이서(최소) 및 스페이서(최대) 내의 길이를 갖는 gRNA 스페이서를 추가로 포함한다. 일부 구현예에서, 표 12에 따른 서열을 갖는 gRNA 또는 주형 RNA는 유전자 변형 폴리펩티드를 추가로 포함하는 시스템에 포함되며, 여기서 유전자 변형 폴리펩티드는 표 12의 동일한 행에 기재된 Cas 도메인을 포함한다.

[0917] *표 12. 유전자 변형 시스템에서 표 8에 열거된 Cas 변이체에 적용하기 위하여 gRNA 및/또는 주형 RNA를 설계하기 위한 구성요소를 정의하는 파라미터*

변이체	PAM(들)	절단	단계	스페이서 (최소)	스페이서 (최대)	crRNA	테트라루프	tracrRNA
Nme2Cas9	NNNNCC	-3	1	22	24	GTTGTAGC TCCCTTTCT CATTTCG	GAAA	CGAAATGAGAACCGTTGCTACAATAAGGC CGTCTGAAAAGATGTGCCGCAACGCTCTG CCCCTTAAAGCTTCTGCTTTAAGGGGCATC GTTTA
PpnCas9	NNNNRTT		1	21	24	GTTGTAGC TCCCTTTTT CATTTCGC	GAAA	GCGAAATGAAAAACGTTGTTACAATAAGA GATGAATTTCTCGCAAAGCTCTGCCTCTTG AAATTTCGGTTTCAAGAGGCATC
SauCas9	NNGRR;N NGRRT	-3	1	21	23	GTTTTAGT ACTCTG	GAAA	CAGAATCTACTAAAACAAGGCAAAATGCC GTGTTTATCTCGTCAACTTGTTGGCGAGA
SauCas9-KKH	NNNRR;N NNRRT	-3	1	21	21	GTTTTAGT ACTCTG	GAAA	CAGAATCTACTAAAACAAGGCAAAATGCC GTGTTTATCTCGTCAACTTGTTGGCGAGA
SauriCas9	NNGG	-3	1	21	21	GTTTTAGT ACTCTG	GAAA	CAGAATCTACTAAAACAAGGCAAAATGCC GTGTTTATCTCGTCAACTTGTTGGCGAGA
SauriCas9-KKH	NNRG	-3	1	21	21	GTTTTAGT ACTCTG	GAAA	CAGAATCTACTAAAACAAGGCAAAATGCC GTGTTTATCTCGTCAACTTGTTGGCGAGA
ScaCas9-Sc++	NNG	-3	1	20	20	GTTTTAGA GCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTA TCAACTTGAAAAAGTGGCACCGAGTCGGT GC
SpyCas9	NGG	-3	1	20	20	GTTTTAGA GCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTA TCAACTTGAAAAAGTGGCACCGAGTCGGT GC

[0918]

SpyCas9-NG	NG (NGG=NGA=NGT>NGC)	-3	1	20	20	GTTTTAGAGCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-SpRY	NRN>NYN	-3	1	20	20	GTTTTAGAGCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
St1Cas9	NNAGAAW>NNAGGAW=NNGGAAW	-3	1	20	20	GTCTTTGTACTCTG	GTAC	CAGAAGCTACAAAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTT
BlatCas9	NNNNCNAA>NNNNCNDO>NNNNC	-3	1	19	23	GCTATAGTTCCTTACT	GAAA	GGTAAGTTGCTATAGTAAGGGCAACAGACCCGAGGCGTTGGGGATCGCCTAGCCCGTGTTTACGGGCTCTCCCCATATTCAAAATAATGACAGACGAGCACCTTGGAGCATTTATCTCCGAGGTGCT
cCas9-v16	NNVACT;NNVATGM;NNVATT;NNVGCT;NNVGTG;NNVGTT	-3	2	21	21	GTCTTAGTACTCTG	GAAA	CAGAATCTACTAAGACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA
cCas9-v17	NNVRRN	-3	2	21	21	GTCTTAGTACTCTG	GAAA	CAGAATCTACTAAGACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA
cCas9-v21	NNVACT;NNVATGM;NNVATT;NNVGCT;NNVGTG;NNVGTT	-3	2	21	21	GTCTTAGTACTCTG	GAAA	CAGAATCTACTAAGACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA
cCas9-v42	NNVRRN	-3	2	21	21	GTCTTAGTACTCTG	GAAA	CAGAATCTACTAAGACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA
CdiCas9	NNRHHHY;NNRAAAY		2	22	22	ACTGGGGGTTCAG	GAAA	CTGAACCTCAGTAAGCATTGGCTCGTTTCCAATGTTGATTGCTCCGCCGGTGCTCCTTATTTTTAAGGGCGCCGGC
CjeCas9	NNNNRYAC	-3	2	21	23	GTTTTAGTCCCT	GAAA	AGGGACTAAAATAAAGAGTTTGCGGGACTCTGCGGGGTTACAATCCCCTAAAACCGC
GeoCas9	NNNNCRAA		2	21	23	GTCATAGTTCCCCTGA	GAAA	TCAGGGTTACTATGATAAGGGCTTTCTGCCTAAGGCAGACTGACCCGCGGCGTTGGGGATCGCCTGTCGCCCGCTTTTGGCGGGCATTCCCCATCCTT
iSpyMacCas9	NAAN	-3	2	19	21	GTTTTAGAGCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
NmeCas9	NNNNGAYT;NNNNGYTT;NNNNGAYA;NNNNGTCT	-3	2	20	24	GTTGTAGCTCCCTTTCTCATTTCG	GAAA	CGAAATGAGAACCGTTGCTACAATAAGGCCGTCTGAAAAGATGTGCCGCAACGCTCTGCCCCTTAAAGCTTCTGCTTTAAGGGGCATCGTTTA

[0919]

ScaCas9	NNG	-3	2	20	20	GTTTTAGAGCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
ScaCas9-HiFi-Sc++	NNG	-3	2	20	20	GTTTTAGAGCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-3var-NRRH	NRRH	-3	2	20	20	GTTTAAGAGCTATGCTG	GAAA	CAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-3var-NRTH	NRTH	-3	2	20	20	GTTTAAGAGCTATGCTG	GAAA	CAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-3var-NRCH	NRCH	-3	2	20	20	GTTTAAGAGCTATGCTG	GAAA	CAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-HF1	NGG	-3	2	20	20	GTTTTAGAGCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-QQR1	NAAG	-3	2	20	20	GTTTTAGAGCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-SpG	NGN	-3	2	20	20	GTTTTAGAGCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-VQR	NGAN	-3	2	20	20	GTTTTAGAGCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-VRER	NGCG	-3	2	20	20	GTTTTAGAGCTA	GAAA	TAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-xCas	NG;GAA;GAT	-3	2	20	20	GTTTAAGAGCTATGCTG	GAAA	CAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SpyCas9-xCas-NG	NG	-3	2	20	20	GTTTAAGAGCTATGCTG	GAAA	CAGCATAGCAAGTTTAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
St1Cas9-CNRZ1066	NNACAA	-3	2	20	20	GTCTTTGTACTCTG	GTAC	CAGAAGCTACAAAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTT
St1Cas9-LMG1831	NNGCAA	-3	2	20	20	GTCTTTGTACTCTG	GTAC	CAGAAGCTACAAAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTT
St1Cas9-MTH17CL396	NNAAAA	-3	2	20	20	GTCTTTGTACTCTG	GTAC	CAGAAGCTACAAAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTT
St1Cas9-TH1477	NNGAAA	-3	2	20	20	GTCTTTGTACTCTG	GTAC	CAGAAGCTACAAAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTT

[0920]

sRGN3.1	NNGG		1	21	23	GTTTTAGTACTCTG	GAAA	CAGAATCTACTGAAACAAGACAATATGTCGTGTTTATCCCATCAATTTATTGGTGGGATTTT
sRGN3.3	NNGG		1	21	23	GTTTTAGTACTCTG	GAAA	CAGAATCTACTGAAACAAGACAATATGTCGTGTTTATCCCATCAATTTATTGGTGGGATTTT

[0921]

[0922] 본 명세서에서, RNA 서열(예를 들어, 주형 RNA 서열)이 티민(T)을 포함하는 특정 서열(예를 들어, 표 12의 서열 또는 이의 일부)을 포함한다고 말할 때, 물론 RNA 서열은 T 대신에 우라실(U)을 포함할 수 있다(그리고 흔히 포함한다)고 이해된다. 예를 들어, RNA 서열은 표 12의 서열에서 T로 나타낸 모든 위치에서 U를 포함할 수 있다. 더 구체적으로, 본 개시내용은 표 12의 모든 gRNA 스캐폴드 서열에 따른 RNA 서열을 제공하며, 여기서 RNA 서열은 표 12의 서열에서 각각의 T 대신에 U를 갖는다. 추가로, 말단 U 및 T는 선택적으로 tracrRNA 서열에 추가되거나 제거될 수 있고 RNA로서 제공되는 경우 변형되거나 변형되지 않을 수 있다. 예에 의해 구속되고자 함이 없

이, 표 12에 예시된 것에 대해 대안적인 gRNA 스캐폴드 서열의 형태는 또한 표 8에 예시된 다양한 Cas9 효소 또는 이의 유도체, 예를 들어 뉴클레오티드 첨가, 치환, 또는 결실이 있는 대체 gRNA 스캐폴드 서열, 예를 들어 스템-루프 구조가 추가되거나 제거된 서열과 함께 기능할 수 있다. 본 명세서에서 gRNA 스캐폴드 서열은 주어진 시스템, Cas-RT 융합 폴리펩티드, 적응증, 표적 돌연변이, 주형 RNA, 또는 전달 비히클에 대해 유사하게 최적화될 수 있는 유전자 변형 시스템의 구성요소를 나타내는 것으로 고려된다.

[0923] *RNA 결합 도메인 동원 부위(RRS)*

[0924] 일부 구현예에서, 본 명세서에 기재된 주형 RNA는 본 명세서에 기재된 바와 같은 RBD에 결합할 수 있는 RNA 결합 도메인(RBD) 동원 부위(RRS)를 포함한다. 일부 구현예에서, RRS는 본 명세서에 기재된 바와 같은 유전자 변형 폴리펩티드 또는 복합체의 RBD에 결합한다. 일부 구현예에서, RRS는 주형 RNA의 5' 말단에 위치한다. 일부 구현예에서, RRS는 주형 RNA 5' 말단의 5, 10, 15, 20, 25, 또는 30개 뉴클레오티드 내에 위치한다. 일부 구현예에서, RRS는 하나 이상(예를 들어, 1 또는 2개)의 스템-루프 서열을 포함한다.

[0925] 일부 구현예에서, 주형 핵산은 복수의 RRS 서열(예를 들어, 복수의 동일한 RRS 서열, 또는 복수의 상이한 RRS 서열)을 포함한다. 일부 구현예에서, RRS 서열은 적어도 2, 3, 4, 5, 6, 7, 8, 9, 또는 10회 반복된다. 일부 구현예에서, 복수의 RRS 서열은 하나 이상의 링커 서열에 의해 분리된다. 일부 구현예에서, 복수의 RRS 서열은 서로 인접하게(예를 들어, 개재 링커 서열 없이) 위치한다.

[0926] 일부 구현예에서, RRS는 PBS와 이종 목적 서열 사이에 위치하지 않는다. 일부 구현예에서, RRS는 PBS와 이종 목적 서열에 위치한다.

[0927] 일부 구현예에서, RRS는 표 40에 열거된 바와 같은 RRS의 핵산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, RRS는 표 40에 열거된 바와 같은 RRS의 핵산 서열, 또는 이와 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개 이하의 뉴클레오티드 차이를 갖는 핵산 서열을 포함한다. 본 명세서에서, RNA 서열(예를 들어, RRS)이 티민(T)을 포함하는 특정 서열(예를 들어, 표 40의 서열 또는 이의 일부)을 포함한다고 말할 때, 물론 RNA 서열은 T 대신에 우라실(U)을 포함할 수 있다(그리고 흔히 포함한다)고 이해된다. 예를 들어, RNA 서열은 표 40의 서열에서 T로 나타낸 모든 위치에서 U를 포함할 수 있다. 더 구체적으로, 본 개시내용은 표 40의 모든 RRS 서열에 따른 RNA 서열을 제공하며, 여기서 RNA 서열은 표 40의 서열에서 각각의 T 대신에 U를 갖는다.

[0928] **표 40. 예시적인 RNA 결합 도메인 동원 부위(RRS)**

RBP 인식 부위 (RRS)	RBP 결합 파트너	서열(5'에서 3')
MS2	MCP	gcACATGAGGATCACCCATGTgc
PP7	PCP	caTAAGGAGTTTATATGGAAACCCTTAtg
com	Com	CTGAATGCCTGCGAGCATC
LS4-1	LS4	GGCAGAGAAAGGCCATACAATCATTGGCCTTGTGAGGCC GTGTGTCTTCCAGTGGC
LS12-1	LS12	GGCAGAGAAAGGCCATACAATCATTGGCTTTTCCATGACG CCAGTTCCAGTGGC
BoxB	람다 N(1-22)	GGGCCCTGAAGAAGGGCCC
Kt	L7Ae	GGATCCGTGATCGGAAACGTGAGATCC
CS1	LS4	GGTGGCAGAGAAAGGCGAAAGCCTTGTGAGGCCATCAA
CS2	LS12	GGATGCAGAGAACGAAAGTTCCATGACGCATCCAA

[0929]

[0930] *말단 블록 서열*

[0931] 일부 구현예에서, 본 명세서에 기재된 바와 같은 주형 RNA는 하나 이상의 말단 블록 서열을 포함한다. 일부 경우에, 본 명세서에 기재된 바와 같은 말단 블록 서열 또는 말단 보호 서열은 주형 RNA를 엑소뉴클레아제 분해로부터 보호할 수 있다(예를 들어, 주형 RNA의 엑소뉴클레아제 분해를 말단 블록 서열이 결여된 다른 유사한 주형 RNA에 비해 적어도 25%, 50%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 감소시킴). 일부 경우에, 본 명세서에 기재된 바와 같은 말단 블록 서열 또는 말단 보호 서열은 역전사효소 반응을 종결시키는 역할을 할 수 있다. 일부 구현예에서, 말단 블록 서열은 (예를 들어, 닉이 있는 표적 핵산 가닥과 쌍을 형성하는) 5' 프로-스페이서에 인접하거나, 이의 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 또는 30개 뉴클레오티드

내에 위치한다. 구현예에서, 5' 프로-스페이서 서열은 닉이 있는 표적 핵산 가닥에 대해 100% 상보성을 갖고/갖거나 Cas 도메인(예를 들어, Cas9 도메인, 예를 들어 nCas9)에 의해 닉 생성 활성을 지시한다. 구현예에서, 5' 프로-스페이서 서열은 표적 핵산 가닥에 대해 17개 이하의 상보성 뉴클레오티드(예를 들어, 약 5, 10, 11, 12, 13, 14, 15, 16, 또는 17개의 상보성 뉴클레오티드)를 갖고, 예를 들어 닉 생성 없이 표적 핵산의 풀림을 촉진한다. 일부 구현예에서, 말단 블록 서열(예를 들어, 5' 말단 블록 서열)은 본 명세서에 기재된 바와 같은 gRNA 스페이서(예를 들어, 프로-스페이서)를 포함한다. 일부 구현예에서, 말단 블록 서열(예를 들어, 5' 말단 블록 서열)은 본 명세서에 기재된 바와 같은 gRNA 스캐폴드를 포함한다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 프로-스페이서는 완전한 닉 생성에 충분한 길이를 갖지 않거나, 제한된 닉 생성에 적합한 길이를 갖는다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 gRNA 스페이서는 완전한 닉 생성에 적합한 길이를 갖는다.

[0932] 일부 구현예에서, 말단 블록 서열은 표 41에 열거된 바와 같은 말단 블록 서열의 핵산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열, 또는 이의 역상보체를 포함한다. 일부 구현예에서, 말단 블록 서열은 표 41에 열거된 바와 같은 말단 블록 서열의 핵산 서열, 또는 이와 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개 이하의 뉴클레오티드 차이를 갖는 핵산 서열, 또는 이의 역상보체를 포함한다. 본 명세서에서, RNA 서열(예를 들어, 말단 블록 서열)이 티민(T)을 포함하는 특정 서열(예를 들어, 표 41의 서열 또는 이의 일부)을 포함한다고 말할 때, 물론 RNA 서열은 T 대신에 우라실(U)을 포함할 수 있다(그리고 흔히 포함한다)고 이해된다. 예를 들어, RNA 서열은 표 41의 서열에서 T로 나타낸 모든 위치에서 U를 포함할 수 있다. 더 구체적으로, 본 개시내용은 표 41의 모든 말단 블록 서열에 따른 RNA 서열을 제공하며, 여기서 RNA 서열은 표 41의 서열에서 각각의 T 대신에 U를 갖는다.

[0933] **표 41. 예시적인 말단 블록 서열**

말단 블록	서열(5'에서 3')
G-사중체(G-quadruplex)	GGTGGTGGTGG
Tinoco 헤어핀	GGACTTCGGTCC
GC-Geo 헤어핀	CTCATAGTTCCCCTGAGAAATCAGGGTTACTATGAG
Nme2Cas9 스캐폴드	GTTGTAGCTCCCTTTCTCATTTCGGAAACGAAATGAGAACCGTTGCTAC AATAAGGCCGTCTGAAAAGATGTGCCGCAACGCTCTGCCCCTTAAAGC TTCTGCTTTAAGGGGCATCGTTTA
Nme2Cas9 스페이서+스캐폴드	**CAGTACATGACCTTACGGGA**GTTGTAGCTCCCTTTCTCATTTCG GAAACGAAATGAGAACCGTTGCTACAATAAGGCCGTCTGAAAAGATGT GCCGCAACGCTCTGCCCCTTAAAGCTTCTGCTTTAAGGGGCATCGTTTA
Nme2Cas9 **16 nt** 스페이서+스캐폴드	**ACATGACCTTACGGGA**GTTGTAGCTCCCTTTCTCATTTCGGAAAC GAAATGAGAACCGTTGCTACAATAAGGCCGTCTGAAAAGATGTGCCGC AACGCTCTGCCCCTTAAAGCTTCTGCTTTAAGGGGCATCGTTTA
BlatCas9 스캐폴드	GCTATAGTTCCTTACTGAAAGGTAAGTTGCTATAGTAAGGGCAACAGA CCCGAGGCGTTGGGGATCGCCTAGCCCGTGTTTACGGGCTCTCCCCAT ATTCAAAATAATGACAGACGAGCACCTTGGAGCATTTATCTCCGAGGT GCT
GeoCas9	GTCATAGTTCCCCTGAGAAATCAGGGTTACTATGATAAGGGCTTTCTGC CTAAGGCAGACTGACCCGCGGCGTTGGGGATCGCCTGTCGCCCGCTTT TGGCGGGCATTCCCCATCCTT
PpnCas9 스캐폴드	GTTGTAGCTCCCTTTTTCATTTCGCGAAAGCGAAATGAAAAACGTTGTT ACAATAAGAGATGAATTTCTCGCAAAGCTCTGCCTCTTGAAATTTCGGT TTCAAGAGGCATC
CdiCas9스캐폴드	ACTGGGGTTCAGGAAACTGAACCTCAGTAAGCATTGGCTCGTTTCCAAT GTTGATTGCTCCGCCGGTGCTCCTTATTTTTAAGGGCGCCGGC
SpyCas9+ 헤어핀 스캐폴드	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCA ACTTGAAAAAGTGGCACCG**GGACTTCGGTCC**CGGTGC
St1Cas9 스캐폴드	GTCTTTGTACTCTGGTACCAGAAGCTACAAAGATAAGGCTTCATGCCGA AATCAACACCCTGTCATTTTATGGCAGGGTGTTTT
cCas9-v16 스캐폴드	GTCTTAGTACTCTGGAAACAGAATCTACTAAGACAAGGCAAAATGCCG TGTTTATCTCGTCAACTTGTTGGCGAGA
SpyCas9-3var-NRRH 스캐폴드	GTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATAAGGCTAG TCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC
SauCas9 스캐폴드	GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAAAATGCCG TGTTTATCTCGTCAACTTGTTGGCGAGA
CjeCas9 스캐폴드	GTTTTAGTCCCTGAAAAGGGACTAAAATAAAGAGTTTGCGGGACTCTG CGGGGTTACAATCCCCTAAAACCGC
SpyCas9 스캐폴드	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCA ACTTGAAAAAGTGGCACCGAGTCGGTGC

[0934]

[0935] 일부 구현예에서, 말단 블록은, 예를 들어 본 명세서에 기재된 바와 같은 프로-스페이서 서열(예를 들어, 5' 프로토스페이서 서열)을 포함한다. 특정 구현예에서, 프로-스페이서 서열은 표적 핵산 가득에 대해 17개 이하의 상보성 뉴클레오티드(예를 들어, 약 17, 18, 19, 20, 21, 22, 또는 23개의 상보성 뉴클레오티드)를 갖는다. 특정 구현예에서, 프로-스페이서 서열은 표적 핵산의 풀림 및 닉 생성을 촉진한다.

[0936] *이종 목적 서열*

[0937] 본 명세서에 기재된 주형 RNA는 유전자 변형 폴리펩티드가 역전사를 위한 주형으로서 사용하여, 원하는 서열을 표적 핵산에 작성할 수 있는 이종 목적 서열을 포함할 수 있다. 일부 구현예에서, 이종 목적 서열은 5'에서 3'으로 사후 편집 상동성 영역, 돌연변이 영역, 및 사전 편집 상동성 영역을 포함한다. 이론에 구속되고자 함이 없이, 주형 RNA에 대해 역전사를 수행하는 RT는 먼저 사전 편집 상동성 영역, 그 돌연변이 영역, 그 다음 사후 편집 상동성 영역을 역전사하여, 이에 의해 양쪽에 상동성 영역이 있는 원하는 돌연변이를 포함하는 DNA 가닥을

형성한다.

[0938] 일부 구현예에서, 이종 목적 서열은 길이가 적어도 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 120, 140, 160, 180, 200, 500, 또는 1,000개의 뉴클레오티드(nt), 또는 길이가 적어도 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 7.5, 8, 8.5, 9, 9.5, 또는 10 킬로베이스이다. 일부 구현예에서, 이종 목적 서열은 길이가 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 120, 140, 160, 180, 200, 500, 1,000, 또는 2000개 이하의 뉴클레오티드(nt), 또는 길이가 단지 20, 15, 10, 9, 8, 7, 6, 5, 4, 또는 3 킬로베이스 이하이다. 일부 구현예에서, 이종 목적 서열은 길이가 30 내지 1000, 40 내지 1000, 50 내지 1000, 60 내지 1000, 70 내지 1000, 74 내지 1000, 75 내지 1000, 76 내지 1000, 77 내지 1000, 78 내지 1000, 79 내지 1000, 80 내지 1000, 85 내지 1000, 90 내지 1000, 100 내지 1000, 120 내지 1000, 140 내지 1000, 160 내지 1000, 180 내지 1000, 200 내지 1000, 500 내지 1000, 30 내지 500, 40 내지 500, 50 내지 500, 60 내지 500, 70 내지 500, 74 내지 500, 75 내지 500, 76 내지 500, 77 내지 500, 78 내지 500, 79 내지 500, 80 내지 500, 85 내지 500, 90 내지 500, 100 내지 500, 120 내지 500, 140 내지 500, 160 내지 500, 180 내지 500, 200 내지 500, 30 내지 200, 40 내지 200, 50 내지 200, 60 내지 200, 70 내지 200, 74 내지 200, 75 내지 200, 76 내지 200, 77 내지 200, 78 내지 200, 79 내지 200, 80 내지 200, 85 내지 200, 90 내지 200, 100 내지 200, 120 내지 200, 140 내지 200, 160 내지 200, 180 내지 200, 30 내지 100, 40 내지 100, 50 내지 100, 60 내지 100, 70 내지 100, 74 내지 100, 75 내지 100, 76 내지 100, 77 내지 100, 78 내지 100, 79 내지 100, 80 내지 100, 85 내지 100, 또는 90 내지 100개의 뉴클레오티드(nt), 또는 길이가 1 내지 20, 1 내지 15, 1 내지 10, 1 내지 9, 1 내지 8, 1 내지 7, 1 내지 6, 1 내지 5, 1 내지 4, 1 내지 3, 1 내지 2, 2 내지 20, 2 내지 15, 2 내지 10, 2 내지 9, 2 내지 8, 2 내지 7, 2 내지 6, 2 내지 5, 2 내지 4, 2 내지 3, 3 내지 20, 3 내지 15, 3 내지 10, 3 내지 9, 3 내지 8, 3 내지 7, 3 내지 6, 3 내지 5, 3 내지 4, 4 내지 20, 4 내지 15, 4 내지 10, 4 내지 9, 4 내지 8, 4 내지 7, 4 내지 6, 4 내지 5, 5 내지 20, 5 내지 15, 5 내지 10, 5 내지 9, 5 내지 8, 5 내지 7, 5 내지 6, 6 내지 20, 6 내지 15, 6 내지 10, 6 내지 9, 6 내지 8, 6 내지 7, 7 내지 20, 7 내지 15, 7 내지 10, 7 내지 9, 7 내지 8, 8 내지 20, 8 내지 15, 8 내지 10, 8 내지 9, 9 내지 20, 9 내지 15, 9 내지 10, 10 내지 15, 10 내지 20, 또는 15 내지 20 킬로베이스이다. 일부 구현예에서, 이종 목적 서열은 길이가 10 내지 100, 10 내지 90, 10 내지 80, 10 내지 70, 10 내지 60, 10 내지 50, 10 내지 40, 10 내지 30, 또는 10 내지 20개의 nt, 예를 들어 길이가 10 내지 80, 10 내지 50, 또는 10 내지 20개의 nt, 예를 들어 길이 약 10 내지 20개의 nt이다. 일부 구현예에서, 이종 목적 서열은 길이가 8 내지 30, 9 내지 25, 10 내지 20, 11 내지 16, 또는 12 내지 15개의 뉴클레오티드이고, 예를 들어 길이가 11 내지 16개의 nt이다. 이론에 구속되고자 함이 없이, 일부 구현예에서, 더 큰 삽입 크기, 더 큰 편집 영역(예를 들어, 표적 영역에서 제1 편집/치환과 제2 편집/치환 사이의 거리), 및 더 많은 수의 원하는 편집(예를 들어, 이종 목적 서열과 표적 게놈의 불일치)은 더 긴 최적의 이종 목적 서열을 생성할 수 있다.

[0939] 특정 구현예에서, 주형 핵산은, 예를 들어 이종성 코딩 영역을 게놈에 도입; 엑손 구조/대안 스플라이싱에 작용 또는 유발, 예를 들어 하나 이상의 엑손의 엑손 스키핑(exon skipping)을 야기; 내인성 유전자의 붕괴를 유발, 예를 들어 유전자 녹아웃을 형성; 내인성 유전자의 전사 활성화를 유발; 내인성 DNA의 후성유전학적 조절을 유발; 하나 이상의 작동가능하게 연결된 유전자의 상향-조절을 유발, 예를 들어 유전자 활성화 또는 과발현을 야기; 하나 이상의 작동가능하게 연결된 유전자의 하향-조절을 유발, 예를 들어 유전자 녹다운을 형성 등에 의해, 숙주 게놈 기능을 변경 또는 특정하는 서열을 함유하도록 확인, 설계, 조작, 및 구축될 수 있는 맞춤화된 RNA 서열 주형을 포함한다. 특정 구현예에서, 맞춤화된 RNA 서열 주형은 엑손 및/또는 이식유전자를 코딩하는 서열을 함유하고, 전사 인자 활성제, 억제자, 인핸서 등 및 이의 조합에 대한 결합 부위를 제공하도록 조작될 수 있다. 일부 구현예에서, 맞춤화된 주형은 표적 부위에 작동가능하게 연결된 내인성 RNA 전사체 또는 내인성 단백질에 발현될 핵산 또는 펩티드 태그를 인코딩하도록 조작될 수 있다. 다른 구현예에서, 코딩 서열은 스플라이스 공여 부위, 스플라이스 수여 부위, 또는 폴리-A 테일을 이용하여 추가로 맞춤화될 수 있다.

[0940] 시스템의 주형 핵산(예를 들어, 주형 RNA)은 통상적으로 표적 DNA 내로 원하는 서열을 작성하기 위한 목적 서열(예를 들어, 이종 목적 서열)을 포함한다. 목적 서열은 코딩 또는 비-코딩일 수 있다. 주형 핵산(예를 들어, 주형 RNA)은 표적 DNA 유전자좌에서 삽입, 돌연변이, 또는 결실을 초래하도록 설계될 수 있다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)는 표적 DNA에서 삽입을 유발하도록 설계될 수 있다. 예를 들어, 주형 핵산(예

를 들어, 주형 RNA)은 이종성 서열을 함유할 수 있으며, 여기서 역전사는 표적 DNA 내로 이종성 서열의 삽입을 초래할 것이다. 다른 구현예에서, RNA 주형은 표적 DNA 내로 결실을 도입하도록 설계될 수 있다. 예를 들어, 주형 핵산(예를 들어, 주형 RNA)은 원하는 결실의 상류 및 하류에서 표적 DNA와 일치할 수 있고, 여기서 역전사는 개재 서열 없이, 예를 들어 개재 서열의 결실을 유발하면서, 주형 핵산(예를 들어, 주형 RNA)로부터 상류 및 하류 서열의 카피를 초래할 것이다. 다른 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 표적 DNA 내로 편집물을 도입하도록 설계될 수 있다. 예를 들어, 주형 RNA는 하나 이상의 뉴클레오티드를 제외하고 표적 DNA 서열과 일치할 수 있으며, 여기서 역전사는 이들 편집물의 표적 DNA 내로의 카피를 초래하여, 돌연변이, 예를 들어 전이 또는 전환 돌연변이를 생성할 것이다.

[0941] 일부 구현예에서, 목적 서열의 표적 부위 내로의 작성은 뉴클레오티드의 치환을 초래하며, 예를 들어 여기서 목적 서열의 전체 길이는 하나 이상의 미스매치 염기가 있는 표적 부위의 일치 길이에 상응한다. 일부 구현예에서, 이종 목적 서열은 서열 변경이 조합, 예를 들어 동시 추가 및 결실, 추가 및 치환, 또는 결실 및 치환이 발생할 수 있도록 설계될 수 있다.

[0942] 일부 구현예에서, 이종 목적 서열은 오픈 리딩 프레임 또는 오픈 리딩 프레임의 단편을 함유할 수 있다. 일부 구현예에서, 이종 목적 서열은 코작 서열을 갖는다. 일부 구현예에서, 이종 목적 서열은 내부 리보솜 도입 부위를 갖는다. 일부 구현예에서, 이종 목적 서열은 T2A 또는 P2A 부위와 같은 자기-절단 펩티드를 갖는다. 일부 구현예에서, 이종 목적 서열은 출발 코돈을 갖는다. 일부 구현예에서, 주형 RNA는 스플라이스 수여 부위를 갖는다. 일부 구현예에서, 주형 RNA는 스플라이스 공여 부위를 갖는다. 예시적인 스플라이스 수여 및 스플라이스 공여 부위는, 그 전체가 본 명세서에 참고로 포함된 WO2016044416에 기재되어 있다. 예시적인 스플라이스 수여 부위 서열은 기술 분야의 당업자에게 알려져 있다. 일부 구현예에서, 주형 RNA는 정지 코돈의 하류에서 마이크로RNA 결합 부위를 갖는다. 일부 구현예에서, 주형 RNA는 오픈 판독 프레임의 정지 코돈의 하류에서 폴리A 테일을 갖는다. 일부 구현예에서, 주형 RNA는 하나 이상의 엑손을 포함한다. 일부 구현예에서, 주형 RNA는 하나 이상의 인트론을 포함한다. 일부 구현예에서, 주형 RNA는 진핵성 전사 종결자를 포함한다. 일부 구현예에서, 주형 RNA는 향상된 번역 요소 또는 번역 향상 요소를 포함한다. 일부 구현예에서, RNA는 인간 T-세포 백혈병 바이러스(HTLV-1) R 영역을 포함한다. 일부 구현예에서, RNA는 B형 간염 바이러스(HPRE) 또는 우드척(Woodchuck) 간염 바이러스(WPRE)의 것과 같은 핵 수출을 향상하는 번역후 조절 요소를 포함한다.

[0943] 일부 구현예에서, 이종 목적 서열은 비-코딩 서열을 함유할 수 있다. 예를 들어, 주형 핵산(예를 들어, 주형 RNA)은 조절 요소, 예를 들어 프로모터 또는 인핸서 서열 또는 miRNA 결합 부위를 포함할 수 있다. 일부 구현예에서, 표적 부위에서 목적 서열의 통합은 내인성 유전자의 상향조절을 초래할 것이다. 일부 구현예에서, 표적 부위에서 목적 서열의 통합은 내인성 유전자의 하향조절을 초래할 것이다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 조직 특이적 프로모터 또는 인핸서를 포함하며, 이들 각각은 단방향 또는 양방향일 수 있다. 일부 구현예에서 프로모터는 RNA 중합효소 I 프로모터, RNA 중합효소 II 프로모터, 또는 RNA 중합효소 III 프로모터이다. 일부 구현예에서 프로모터는 TATA 요소를 포함한다. 일부 구현예에서 프로모터는 B 인식 요소를 포함한다. 일부 구현예에서 프로모터는 전사 인자에 대한 하나 이상의 결합 부위를 갖는다.

[0944] 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 후성유전학적 변형을 조정하는 부위를 포함한다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 염색질 절연체를 포함한다. 예를 들어, 주형 핵산(예를 들어, 주형 RNA)은 CTCF 부위 또는 DNA 메틸화를 위해 표적화된 부위를 포함한다.

[0945] 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 이펙터 서열에 작동가능하게 연결된 적어도 하나의 조절 영역으로 구성된 유전자 발현 단위를 포함한다. 이펙터 서열은, RNA로 전사되는 서열(예를 들어, 코딩 서열 또는 비-코딩 서열, 예컨대 마이크로RNA를 인코딩하는 서열)일 수 있다.

[0946] 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)의 이종 목적 서열은 내인성 인트론에서 표적 게놈 내로 삽입된다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)의 이종 목적 서열은 표적 게놈 내로 삽입되고, 이에 의해 새로운 엑손으로 작용한다. 일부 구현예에서, 이종 목적 서열의 표적 게놈 내로의 삽입은 천연 엑손의 대체 또는 천연 엑손의 스키핑을 초래한다.

[0947] 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)의 이종 목적 서열은 게놈 세이프 하버 부위, 예컨대 AAVS1, CCR5, ROSA26, 또는 알부민 유전자좌에서 표적 게놈 내로 삽입된다. 일부 구현예에서, 유전자 변형은 T-세포 수용체 α 상수(*TRAC*) 유전자좌 내로 CAR를 통합하는 데 사용된다(Eyquem et al Nature 543, 113-117 (2017)). 일부 구현예에서, 유전자 변형 시스템은 T-세포 수용체 β 상수(*TRBC*) 유전자좌 내로 CAR를 통합하는 데 사용된다. 많은 다른 세이프 하버부는 컴퓨터 계산 접근에 의해 확인되었으며(Pellenz et al Hum Gen Ther 30, 814-

828 (2019)), 유전자 변형 시스템-매개 통합에 사용될 수 있다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)의 이종 목적 서열은 유전자간 또는 유전자내 영역에서 게놈에 첨가된다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)의 이종 목적 서열은 내인성 활성 유전자의 0.1 kb, 0.25 kb, 0.5 kb, 0.75, kb, 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 7.5 kb, 10 kb, 15 kb, 20 kb, 25 kb, 50, 75 kb, 또는 100 kb 내 5' 또는 3'에서 게놈에 첨가된다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)의 이종 목적 서열은 내인성 프로모터 또는 인핸서의 0.1 kb, 0.25 kb, 0.5 kb, 0.75, kb, 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 7.5 kb, 10 kb, 15 kb, 20 kb, 25 kb, 50, 75 kb, 또는 100 kb 내 5' 또는 3'에서 게놈에 첨가된다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)의 이종 목적 서열은, 예를 들어 50 내지 50,000 염기쌍(예를 들어, 50 내지 40,000 bp, 500 내지 30,000 bp, 500 내지 20,000 bp, 100 내지 15,000 bp, 500 내지 10,000 bp, 50 내지 10,000 bp, 50 내지 5,000 bp일 수 있다.

[0948] 주형 핵산(예를 들어, 주형 RNA)은 표적 DNA 유전자좌에서 삽입, 돌연변이, 또는 결실을 초래하도록 설계될 수 있다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)는 표적 DNA에서 삽입을 유발하도록 설계될 수 있다. 예를 들어, 주형 핵산(예를 들어, 주형 RNA)은 이종 목적 서열을 함유할 수 있으며, 여기서 역전사는 표적 DNA 내로 이종 목적 서열의 삽입을 초래할 것이다. 다른 구현예에서, RNA 주형은 표적 DNA 내로 결실을 작성하도록 설계될 수 있다. 예를 들어, 주형 핵산(예를 들어, 주형 RNA)은 원하는 결실의 상류 및 하류에서 표적 DNA와 일치할 수 있고, 여기서 역전사는 개재 서열 없이, 예를 들어 개재 서열의 결실을 유발하면서, 주형 핵산(예를 들어, 주형 RNA)로부터 상류 및 하류 서열의 카피를 초래할 것이다. 다른 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 표적 DNA 내로 편집물을 작성하도록 설계될 수 있다. 예를 들어, 주형 RNA는 하나 이상의 뉴클레오티드를 제외하고 표적 DNA 서열과 일치할 수 있으며, 여기서 역전사는 이들 편집물의 표적 DNA 내로의 카피를 초래하여, 돌연변이, 예를 들어 전이 또는 전환 돌연변이를 생성할 것이다.

[0949] 일부 구현예에서, 사전 편집 상동성 도메인은 표적 핵산 분자에 포함된 핵산 서열과 100% 서열 동일성을 갖는 핵산 서열을 포함한다.

[0950] 일부 구현예에서, 사후 편집 상동성 도메인은 표적 핵산 분자에 포함된 핵산 서열과 100% 서열 동일성을 갖는 핵산 서열을 포함한다.

[0951] 일부 구현예에서, 상동성 도메인(예를 들어, 사전 편집 상동성 도메인)은 하기 표 38에 열거된 바와 같은 상동성 1 서열의 핵산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 상동성 도메인(예를 들어, 사전 편집 상동성 도메인)은 하기 표 38에 열거된 바와 같은 상동성 1 서열의 핵산 서열, 또는 이에 대하여 1, 2, 3, 4, 또는 5개 이하의 뉴클레오티드 차이를 갖는 핵산 서열을 포함한다. 일부 구현예에서, 상동성 도메인은 0 내지 30개 뉴클레오티드(예를 들어, 약 0 내지 10, 10 내지 20, 또는 20 내지 30개 뉴클레오티드)의 길이를 갖는다. 본 명세서에서, RNA 서열(예를 들어, 상동성 도메인 서열)이 티민(T)을 포함하는 특정 서열(예를 들어, 표 38의 서열 또는 이의 일부)을 포함한다고 말할 때, 물론 RNA 서열은 T 대신에 우라실(U)을 포함할 수 있다(그리고 흔히 포함한다)고 이해된다. 예를 들어, RNA 서열은 표 38의 서열에서 T로 나타낸 모든 위치에서 U를 포함할 수 있다. 더 구체적으로, 본 개시내용은 표 38의 모든 상동성 도메인 서열에 따른 RNA 서열을 제공하며, 여기서 RNA 서열은 표 38의 서열에서 각각의 T 대신에 U를 갖는다. 특정 구현예에서, 상동성 도메인은 0 내지 5, 5 내지 10, 10 내지 15, 15 내지 20, 20 내지 25, 25 내지 30, 30 내지 35, 35 내지 40, 40 내지 45, 또는 45 내지 50개 뉴클레오티드의 길이를 갖는다. 특정 구현예에서, 상동성 도메인은 50 내지 100, 100 내지 150, 150 내지 200, 200 내지 250, 250 내지 300, 300 내지 350, 350 내지 400, 400 내지 450, 또는 450 내지 550개 뉴클레오티드의 길이를 갖는다. 특정 구현예에서, 상동성 도메인은 약 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 또는 500개 뉴클레오티드의 길이를 갖는다.

[0952] 표 38. 예시적인 상동성 1 서열

리포터	편집 유형	편집 서열(5'에서 3')	상동성 1 길이	상동성 1 서열 (5'에서 3')
BFP에서 GFP	SNP	GT	3 nt	ACG
	SNP	GT	3 nt	ACG
	SNP	GT	3 nt	ACG
	SNP	GT	3 nt	ACG
	SNP	GT	3 nt	ACG
	SNP	GT	3 nt	ACG
	SNP	GT	3 nt	ACG
	SNP	GT	3 nt	ACG
	SNP	GT	3 nt	ACG
250 bp GFP 삽입	250 bp 삽입	AGAATTTTGTAATACGACTCACTATAGGGCGGCCGGGAATTCGC CACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGT GCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAG TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCA AGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTG CCCTGGCCCACCCTCGTGACCACCCTGACGTACG	0 nt	
	250 bp 삽입	CAGAATTTTGTAATACGACTCACTATAGGGCGGCCGGGAATTCG CCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGT GCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAG TTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCA AGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTG CCCTGGCCCACCCTCGTGACCACCCTGACGTAC	1 nt	G
	250 bp 삽입	TCAGAATTTTGTAATACGACTCACTATAGGGCGGCCGGGAATTC GCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTG GTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACA AGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGG	2 nt	CG

[0953]

[0954]

		CAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCG TGCCCTGGCCCACCCTCGTGACCACCCTGACGTA		
	250 bp 삽입	GTCAGAATTTTGTAATACGACTCACTATAGGGCGGCCGGGAATT CGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTG GTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACA AGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGG CAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCG TGCCCTGGCCCACCCTCGTGACCACCCTGACGT	3 nt	ACG
	250 bp 삽입	CGTCAGAATTTTGTAATACGACTCACTATAGGGCGGCCGGGAAT TCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGT GGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCAC AAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTACG GCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCC GTGCCCTGGCCCACCCTCGTGACCACCCTGACG	4 nt	TACG
	250 bp 삽입	CCGTCAGAATTTTGTAATACGACTCACTATAGGGCGGCCGGGAA TTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGG TGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCA CAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTAC GGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCC CGTGCCCTGGCCCACCCTCGTGACCACCCTGAC	5 nt	GTACG
	250 bp 삽입	ACCGTCAGAATTTTGTAATACGACTCACTATAGGGCGGCCGGGA ATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGG GTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCC ACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCTA CGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGC CCGTGCCCTGGCCCACCCTCGTGACCACCCTGA	6 nt	CGTACG
	250 bp 삽입	AACCGTCAGAATTTTGTAATACGACTCACTATAGGGCGGCCGGG AATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGG GGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGC CACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACCT ACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTG CCCGTGCCCTGGCCCACCCTCGTGACCACCCTG	7 nt	ACGTACG

[0955]

	250 bp 삽입	GAACCGTCAGAATTTTGTAATACGACTCACTATAGGGCGGCCGG GAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCG GGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGG CCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCACC TACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCT GCCCGTGCCCTGGCCCACCCTCGTGACCACCCT	8 nt	GACGTACG
	250 bp 삽입	TGAACCGTCAGAATTTTGTAATACGACTCACTATAGGGCGGCCG GGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACC GGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACG GCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCAC CTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGC TGCCCGTGCCCTGGCCCACCCTCGTGACCACCC	9 nt	TGACGTACG
	250 bp 삽입	GTGAACCGTCAGAATTTTGTAATACGACTCACTATAGGGCGGCC GGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCAC CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAAC GGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCCA CCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAG CTGCCCGTGCCCTGGCCCACCCTCGTGACCACC	10 nt	CTGACGTACG
	250 bp 삽입	AGTGAACCGTCAGAATTTTGTAATACGACTCACTATAGGGCGGC CGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCA CCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAA CGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGCC ACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAA GCTGCCCGTGCCCTGGCCCACCCTCGTGACCAC	11 nt	CCTGACGTACG
	250 bp 삽입	TAGTGAACCGTCAGAATTTTGTAATACGACTCACTATAGGGCGG CCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTC ACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAA ACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGATGC CACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCA AGCTGCCCGTGCCCTGGCCCACCCTCGTGACCA	12 nt	CCCTGACGTACG
	250 bp	TTAGTGAACCGTCAGAATTTTGTAATACGACTCACTATAGGGCG GCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTT CACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTA	13 nt	ACCCTGACGTACG

[0956]

	삽입	AACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGAT GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGG CAAGCTGCCCGTGCCCTGGCCCACCCTCGTGACC		CACCCTGACGTACG
	250 bp 삽입	TTTAGTGAACCGTCAGAATTTTGTAATACGACTCACTATAGGGC GGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGT TCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGT AAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGCGAT GCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGG CAAGCTGCCCGTGCCCTGGCCCACCCTCGTGAC	14 nt	
	250 bp 삽입	GTTTAGTGAACCGTCAGAATTTTGTAATACGACTCACTATAGGG CGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCT GTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGAC GTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCAC CGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTGA	15 nt	CCACCCTGACGTAC G
	250 bp 삽입	CGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCACTATAGG GCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGC TGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGA CGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCAC CGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGTG	16 nt	ACCACCCTGACGTA CG
	250 bp 삽입	TCGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCACTATAGG GCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGC TGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGA CGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGGC GATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCAC CGGCAAGCTGCCCGTGCCCTGGCCCACCCTCGT	17 nt	GACCACCCTGACGT ACG
	250 bp 삽입	CTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCACTATAG GGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAG CTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCG ACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGG CGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCA CCGGCAAGCTGCCCGTGCCCTGGCCCACCCTCG	18 nt	TGACCACCCTGACG TACG

[0957]

	250 bp 삽입	GCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCACTATA GGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGA GCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGC GACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGG GCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACC ACCGGCAAGCTGCCCGTGCCCTGGCCCACCCTC	19 nt	GTGACCACCCTGAC GTACG
	250 bp 삽입	AGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCACTAT AGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGG AGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGG CGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAG GGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCAC CACCGGCAAGCTGCCCGTGCCCTGGCCCACCCT	20 nt	CGTGACCACCCTGA CGTACG
mCherry 삽입	>750 bp 삽입	[mCherry- 발현 카세트]	0 nt		
	>750 bp 삽입	[mCherry- 발현 카세트]	1 nt	G	
	>750 bp 삽입	[mCherry- 발현 카세트]	2 nt	CG	
	>750 bp 삽입	[mCherry- 발현 카세트]	3 nt	ACG	
	>750 bp 삽입	[mCherry- 발현 카세트]	4 nt	TACG	

[0958]

>750 bp 삽입	[mCherry- 발현 카세트]	5 nt	GTACG
>750 bp 삽입	[mCherry- 발현 카세트]	6 nt	CGTACG
>750 bp 삽입	[mCherry- 발현 카세트]	7 nt	ACGTACG
>750 bp 삽입	[mCherry- 발현 카세트]	8 nt	GACGTACG
>750 bp 삽입	[mCherry- 발현 카세트]	9 nt	TGACGTACG
>750 bp 삽입	[mCherry- 발현 카세트]	10 nt	CTGACGTACG
>750 bp 삽입	[mCherry- 발현 카세트]	11 nt	CCTGACGTACG
>750 bp 삽입	[mCherry- 발현 카세트]	12 nt	CCCTGACGTACG

	>750 bp 삽입	[mCherry- 발현 카세트]	13 nt	ACCCTGACGTACG
	>750 bp 삽입	[mCherry- 발현 카세트]	14 nt	CACCCTGACGTACG
	>750 bp 삽입	[mCherry- 발현 카세트]	15 nt	CCACCCTGACGTAC G
	>750 bp 삽입	[mCherry- 발현 카세트]	16 nt	ACCACCCTGACGTA CG
	>750 bp 삽입	[mCherry- 발현 카세트]	17 nt	GACCACCCTGACGT ACG
	>750 bp 삽입	[mCherry- 발현 카세트]	18 nt	TGACCACCCTGACG TACG
	>750 bp 삽입	[mCherry- 발현 카세트]	19 nt	GTGACCACCCTGAC GTACG
	>750 bp 삽입	[mCherry- 발현 카세트]	20 nt	CGTGACCACCCTGA CGTACG

[0959]

[0960] 일부 구현예에서, 상동성 도메인(예를 들어, 사전 편집 상동성 도메인)은 하기 표 39에 열거된 바와 같은 상동성 2 서열의 핵산 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 상동성 도메인(예를 들어, 사전 편집 상동성 도메인)은 하기 표 39에 열거된 바와 같은 상동성 2 서열의 핵산 서열, 또는 이에 대하여 1, 2, 3, 4, 또는 5개 이하의 뉴클레오티드 차이를 갖는 핵산 서열을 포함한다. 일부 구현예에서, 상동성 도메인은 0 내지 1000개 뉴클레오티드(예를 들어, 약 0 내지 5, 5 내지 10, 10 내지 50, 50 내지 100, 100 내지 500, 또는 500 내지 1000개 뉴클레오티드)의 길이를 갖는다. 본 명세서에서, RNA 서열(예를 들어, 상동성 도메인 서열)이 티민(T)을 포함하는 특정 서열(예를 들어, 표 39의 서열 또는 이의 일부)을 포함한다고 말할 때, 물론 RNA 서열은 T 대신에 우라실(U)을 포함할 수 있다(그리고 흔히 포함한다)고 이해된다. 예를 들어, RNA 서열은 표 39의 서열에서 T로 나타낸 모든 위치에서 U를 포함할 수 있다. 더 구체적으로, 본 개시내용은 표 39의 모든 상동성 도메인 서열에 따른 RNA 서열을 제공하며, 여기서 RNA 서열은 표 39의 서열에서 각각의 T 대신에 U를 갖는다.

[0961] 표 39. 예시적인 상동성 2 서열

리포터	상동성 2 길이	상동성 2 서열(5'에서 3')	상동성 1 쌍
BFP에서 GFP	8 nt	ACCCTGAC	
	11 nt	ACCACCCTGAC	
	12 nt	GACCACCCTGAC	
	13 nt	TGACCACCCTGAC	
	14 nt	GTGACCACCCTGAC	
	16 nt	TCGTGACCACCCTGAC	
	20 nt	ACCCTCGTGACCACCCTGAC	
	24 nt	GCCCACCCTCGTGACCACCCTGAC	
	25 nt	GGCCCACCCTCGTGACCACCCTGAC	
250 bp GFP 삽입	500 nt	CCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAAC GCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC ATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT ATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC CATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	499 nt	CCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACG CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACA TCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTA TGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACC ATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[0962]

[0963]

	498 nt	CGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGC CAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACAT CAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT GCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCA TGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG TAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGA GCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	497 nt	GCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCC AATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC AAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT GCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCA TGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG TAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGA GCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	496 nt	CCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCA ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCA AGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGC CCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG GTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCT CCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA TAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1
	495 nt	CTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAA TAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA GTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCC CAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG GTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCT CCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA TAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1
	494 nt	TGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAAT AGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAG	0 nt 상동성 1

[0964]

		TGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC AGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT GATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATA ACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	0 nt 상동성 1
493 nt	GGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAAT AGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAG TGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC AGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT GATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATA ACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	0 nt 상동성 1
492 nt	GCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATA GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGT GTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA GTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT GATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATA ACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	0 nt 상동성 1
491 nt	CTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAG GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTG TATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAG TACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGA TGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCAC CCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAAC CCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTT TAGTGAACCGTC	0 nt 상동성 1
490 nt	TGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGG GACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT ATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGT ACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGAT	

[0965]

		GCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACC CCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACC CCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT AGTGAACCGTC	
	489 nt	GACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGG ACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT CATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTAC ATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATG CGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCC CATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCC CGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTA GTGAACCGTC	0 nt 상동성 1
	488 nt	ACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGA CTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC ATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTAC ATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATG CGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCC CATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCC CGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTA GTGAACCGTC	0 nt 상동성 1
	487 nt	CCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGAC TTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCA TATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACAT GACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCA TTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCG CCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGT GAACCGTC	0 nt 상동성 1
	486 nt	CGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTT TCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATA TGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATG ACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG TTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC	0 nt 상동성 1

[0966]

		CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	
	485 nt	GCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTT CCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT GCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGA CCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGT TTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	0 nt 상동성 1
	484 nt	CCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTC CATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT GCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGA CCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGT TTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	0 nt 상동성 1
	483 nt	CCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCC ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATG CCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGAC CTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTT TGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGA CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	482 nt	CAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCA TTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGC CAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT TACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT GGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGAC GTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1

[0967]

	481 nt	AACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCAT TGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC AAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT ACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTG GCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACG TCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGT TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCG TC
	480 nt	ACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATT GACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCA AGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA CGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG CAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGT CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTT GACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGT C
	479 nt	CGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTG ACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAA GTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTAC GGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC AGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTC AATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTG ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC
	478 nt	GACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGA CGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAG TCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACG GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCA GTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCA ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGA CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC
	477 nt	ACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGAC GTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGT CCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACG GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCA

[0968]

		GTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCA ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGA CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	476 nt	CCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACG TCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTC CGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGG GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAG TACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAA TGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	475 nt	CCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGT CAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCC GCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGG ACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT ACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAAT GGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	474 nt	CCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTC AATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCG CCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGA CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGG GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	473 nt	CCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCA ATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGAC TTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGG GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	472 nt	CGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAA TGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCC CCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTT	0 nt 상동성 1

[0969]

		TCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACAC CAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGA GTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	471 nt	GCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAAT GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCC CCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTT CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACC AATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAAT GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	470 nt	CCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATG GGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCC CTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTC CTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCA ATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGT TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	469 nt	CCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGG GTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCT ATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCT ACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAAT GGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGG GCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	468 nt	CATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGG TGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTA TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTA CTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATG GGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTG TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGG CGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	467 nt	ATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGT GGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTAT	0 nt 상동성 1

[0970]

		TGACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTAC TTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGG GCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGC GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC
466 nt	TTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTG GAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATT GACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACT TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
465 nt	TGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTG GAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATT GACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACT TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
464 nt	GACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGG AGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTG ACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTT GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
463 nt	ACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGA GTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGA CGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTG GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGC GTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTT GGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGT AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[0971]

	462 nt	CGTCAATAAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAG TATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGAC GTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGG CAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGT GGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTG GCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTA GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	461 nt	GTCAATAAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGT ATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACG TCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGC AGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	460 nt	TCAATAAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTA TTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGT CAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCA GTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	459 nt	CAATAAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTAT TTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTC AATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCA GTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	458 nt	AATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTT ACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAA TGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTA CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATA GCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA	0 nt 상동성 1

		AAATCAACGGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	457 nt	ATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTA CGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAAT GACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTA CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATA GCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	456 nt	TAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTAC GGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATG ACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACA TCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAG CGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAA AATCAACGGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	455 nt	AATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTAC GGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATG ACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACA TCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAG CGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAA AATCAACGGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	454 nt	ATGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACG GTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGA CGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACAT CTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGC GGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA ATCAACGGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTA CGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	453 nt	TGACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGG TAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGAC GGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATC TACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCG	0 nt 상동성 1

[0972]

	GTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAA TCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTAC GGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
452 nt	GACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGT AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACG GTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCT ACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGG TTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAAT CAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
451 nt	ACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTA AACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGG TAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTA CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGT TTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC AACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
450 nt	CGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAA ACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGT AAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTAC GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTT TGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCA ACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
449 nt	GTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAA CTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTA AATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACG TATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTT GACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAA CGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
448 nt	TATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAA TGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTA	0 nt 상동성 1

[0973]

		TTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGA CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG GGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	447 nt	ATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAA TGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTA TTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGA CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG GGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	446 nt	TGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTG CCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAAT GGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTAT TAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGAC TCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	445 nt	GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATG GCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATT AGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACT CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	444 nt	TTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCC ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGG CCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTA GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTC ACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGG ACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	443 nt	TCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCC ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGG	0 nt 상동성 1

[0974]

		CCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTA GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTC ACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGG ACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	442 nt	CCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCA CTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT CATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCAC GGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGAC TTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGG TCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 삼동성 1
	441 nt	CCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCAC TTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCC GCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTC ATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACG GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTT TCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC TATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 삼동성 1
	440 nt	CATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACT TGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCC GCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTC ATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACG GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTT TCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC TATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 삼동성 1
	439 nt	ATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTT GGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG CCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT CGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGG GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 삼동성 1

[0975]

	438 nt	TAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTG GCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC CTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC GCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGG GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	437 nt	AGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGG CAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCG CTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA AAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	436 nt	GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGC AGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCG CTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA AAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	435 nt	TAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCA GTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTG GCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCT ATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAA ATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA AGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	434 nt	AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAG TACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGC ATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT TACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTC CAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAAT	0 nt 상동성 1

[0976]

		GTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTCGTTTAGTGAACCGTC	
	433 nt	ACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGT ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAAATGGCCCGCCTGGC ATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT TACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTC CAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAAT GTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	432 nt	CGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA CATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAAATGGCCCGCCTGGCAT TATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTA CCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCA AGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT CGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCA GAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	431 nt	GCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTAC ATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAAATGGCCCGCCTGGCATT ATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC CATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	430 nt	CCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACA TCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAAATGGCCCGCCTGGCATTA TGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACC ATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	429 nt	CAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACAT CAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT GCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCA TGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG	0 nt 상동성 1

[0977]

		TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG TAATAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGA GCTCGTTTAGTGAACCGTC	
	428 nt	AATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATC AAGTGTATCATATGCCAAGTCCGCCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT GCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCA TGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG TAATAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGA GCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	427 nt	ATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCA AGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGC CCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG GTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCT CCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA TAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1
	426 nt	TAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAA GTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCC CAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG GTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCT CCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA TAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1
	425 nt	AGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAG TGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC AGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT GATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATA ACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	
	424 nt	GGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGT GTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA GTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT	0 nt 상동성 1

[0978]

		GATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATA ACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	
	423 nt	GGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTG TATCATATGCCAAGTCCGCCCCCCTATTGACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAG TACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGA TGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCAC CCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAAC CCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTT TAGTGAACCGTC	0 nt 상동성 1
	422 nt	GACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGT ATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGT ACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGAT GCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACC CCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACC CCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT AGTGAACCGTC	0 nt 상동성 1
	421 nt	ACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTAT CATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTAC ATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATG CGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCC CATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCC CGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTA GTGAACCGTC	0 nt 상동성 1
	420 nt	CTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATC ATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTAC ATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATG CGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCC CATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCC CGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTA GTGAACCGTC	
	419 nt	TTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCA TATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACAT	0 nt 상동성 1

[0979]

		GACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCA TTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCG CCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGT GAACCGTC	
	418 nt	TTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCAT ATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACAT GACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCA TTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCG CCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGT GAACCGTC	0 nt 상동성 1
	417 nt	TCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATA TGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATG ACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG TTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	0 nt 상동성 1
	416 nt	CCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT GCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGA CCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGT TTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	0 nt 상동성 1
	415 nt	CATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATAT GCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGA CCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGT TTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	0 nt 상동성 1

[0980]

	414 nt	ATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATG CCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGAC CTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTT TGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGA CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	413 nt	TTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGC CAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT TACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT GGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGAC GTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	412 nt	TGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCC AAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT ACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTG GCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACG TCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGT TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCG TC	0 nt 상동성 1
	411 nt	GACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCA AGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA CGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG CAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGT CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTT GACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGT C	0 nt 상동성 1
	410 nt	ACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAA GTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTAC GGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC AGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTC AATGGGAGTTTGTTTTGGCACCAAAATCAACGGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTG ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[0981]

409 nt	CGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAG TCCGCCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACG GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCA GTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCA ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGA CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 삼동성 1
408 nt	GTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGT CCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACG GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCA GTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCA ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGA CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 삼동성 1
407 nt	TCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTC CGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGG GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAG TACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAA TGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 삼동성 1
406 nt	CAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCC GCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGG ACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT ACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAAT GGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 삼동성 1
405 nt	AATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCG CCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGGA CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGG GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 삼동성 1
404 nt	ATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGC CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGAC TTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGG	0 nt 삼동성 1

[0982]

	GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
403 nt	TGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCC CCCTATTGACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTT TCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACAC CAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGA GTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
402 nt	GGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCC CCTATTGACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTT CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACC AATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAAT GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
401 nt	GGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCC CTATTGACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTC CTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCA ATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGT TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
400 nt	GTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCT ATTGACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCT ACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAAT GGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGG GCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
399 nt	TGGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTA TTGACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTA CTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATG GGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTG TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGG CGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
398 nt	GGAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTAT TGACGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTAC	0 nt 상동성 1

[0983]

		TTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGG GCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGGC GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	397 nt	GAGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATT GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACT TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	396 nt	AGTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTG ACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTT GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	395 nt	GTATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGA CGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTG GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGC GTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTT GGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGGCGGT AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	394 nt	TATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGAC GTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGG CAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGT GGATAGCGGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTG GCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGGCGGTA GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	393 nt	ATTTACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACG TCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGC AGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[0984]

	392 nt	TTTACGGTAAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGT CAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCA GTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	391 nt	TTACGGTAAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTC AATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCA GTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	390 nt	TACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCA ATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGT ACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGAT AGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC AAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGT GTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	389 nt	ACGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAA TGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTA CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATA GCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	388 nt	CGGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAAT GACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTA CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATA GCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	387 nt	GGTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATG ACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACA TCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAG CGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAA	0 nt 상동성 1

[0985]

		AATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	386 nt	GTAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGA CGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACAT CTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGC GGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA ATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTA CGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	385 nt	TAAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGAC GGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATC TACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCG GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAA TCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTAC GGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	384 nt	AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACG GTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCT ACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGG TTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAAT CAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	383 nt	AACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGG TAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTA CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGT TTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC AACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	382 nt	ACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGT AAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTAC GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTT TGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCA ACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	381 nt	CTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTA AATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACG	0 nt 상동성 1

[0986]

		TATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTT GACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAA CGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	380 nt	TGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAA ATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGT ATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTG ACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAAC GGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTG GGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	379 nt	GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAA TGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTA TTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGA CTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG GGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	378 nt	CCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAAT GGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTAT TAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGAC TCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	377 nt	CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATG GCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATT AGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACT CACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	376 nt	CACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATG GCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATT AGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACT CACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[0987]

	375 nt	ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGG CCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTA GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTC ACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGG ACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	374 nt	CTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCC CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT CATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCAC GGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGAC TTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGG TCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	373 nt	TTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCC GCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTC ATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACG GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTT TCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC TATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	372 nt	TGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCC GCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTC ATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACG GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTT TCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC TATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	371 nt	GGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCG CCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT CGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGG GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	370 nt	GCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC CTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC GCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGG GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC	0 nt 상동성 1

[0988]

	369 nt	CAAAATGTCGTAATAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	
		CAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCG CTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA AAATGTCGTAATAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	368 nt	AGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCG CTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA AAATGTCGTAATAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	367 nt	GTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTG GCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCT ATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAA ATGTCGTAATAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA AGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	366 nt	TACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGC ATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT TACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTC CAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAAT GTCGTAATAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	365 nt	ACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGC ATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT TACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTC CAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAAT GTCGTAATAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	364 nt	CATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCAT TATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTA	0 nt 상동성 1

[0989]

		CCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCA AGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT CGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCA GAGCTCGTTTAGTGAACCGTC	
	363 nt	ATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATT ATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC CATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	362 nt	TCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTA TGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACC ATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	361 nt	CAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT GCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCA TGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG TAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGA GCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	360 nt	AAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTAT GCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCA TGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG TAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGA GCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	359 nt	AGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGC CCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG GTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCT CCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA TAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1

[0990]

	358 nt	GTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCC CAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG GTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCT CCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA TAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1
	357 nt	TGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCC AGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT GATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATA ACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	0 nt 상동성 1
	356 nt	GTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCA GTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT GATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATA ACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	0 nt 상동성 1
	355 nt	TATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAG TACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGA TGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCAC CCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAAC CCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTT TAGTGAACCGTC	0 nt 상동성 1
	354 nt	ATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGT ACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGAT GCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACC CCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACC CCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT AGTGAACCGTC	0 nt 상동성 1
	353 nt	TCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTA CATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGAT GCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACC CCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACC	0 nt 상동성 1

[0991]

	352 nt	CCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT AGTGAACCGTC	
		CATATGCCAAGTCCGCCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTAC ATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATG CGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCC CATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCC CGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTA GTGAACCGTC	0 nt 상동성 1
	351 nt	ATATGCCAAGTCCGCCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTAC ATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATG CGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCC CATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCC CGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTA GTGAACCGTC	0 nt 상동성 1
	350 nt	TATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACAT GACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCA TTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCG CCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGT GAACCGTC	0 nt 상동성 1
	349 nt	ATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACAT GACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCA TTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCG CCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGT GAACCGTC	0 nt 상동성 1
	348 nt	TGCCAAGTCCGCCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATG ACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG TTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	
	347 nt	GCCAAGTCCGCCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGA CCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGT	0 nt 상동성 1

[0992]

		TTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	
	346 nt	CCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGAC CTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTT TGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGA CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	345 nt	CAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCT TACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT GGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGAC GTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	344 nt	AAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTT ACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTG GCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACG TCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGT TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCG TC	0 nt 상동성 1
	343 nt	AGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTA CGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG CAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGT CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTT GACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGT C	0 nt 상동성 1
	342 nt	GTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTAC GGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC AGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTC AATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTG ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[0993]

	341 nt	TCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACG GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCA GTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCA ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGA CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	340 nt	CCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACG GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCA GTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCA ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGA CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	339 nt	CGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGG GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAG TACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAA TGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	338 nt	GCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGG ACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT ACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAAT GGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	337 nt	CCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGA CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGG GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	336 nt	CCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGAC TTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGG GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	335 nt	CCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTT TCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACAC CAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGA	0 nt 상동성 1

[0994]

		GTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	334 nt	CCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTT CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACC AATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAAT GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	333 nt	CTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTC CTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCA ATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGT TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	332 nt	TATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCC TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAA TGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTT TGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	331 nt	ATTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCT ACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAAT GGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGG GCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	330 nt	TTGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTA CTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATG GGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTG TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGG CGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	329 nt	TGACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTAC TTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGG GCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGC GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	328 nt	GACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACT TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG	0 nt 상동성 1

[0995]

		CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	327 nt	ACGTCAATGACGGTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTT GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG CGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	326 nt	CGTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTG GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGC GTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTT GGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGT AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	325 nt	GTCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGG CAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGT GGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTG GCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTA GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	324 nt	TCAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGC AGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	323 nt	CAATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCA GTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	322 nt	AATGACGGTAAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCA GTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[0996]

	321 nt	ATGACGGTAAAATGGCCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGT ACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGAT AGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC AAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGT GTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	320 nt	TGACGGTAAAATGGCCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTA CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATA GCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	319 nt	GACGGTAAAATGGCCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTA CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATA GCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	318 nt	ACGGTAAAATGGCCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACA TCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAG CGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAA AATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	317 nt	CGGTAAAATGGCCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACAT CTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGC GGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA ATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTA CGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	316 nt	GGTAAATGGCCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATC TACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCG GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAA TCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTAC GGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	315 nt	GTAAAATGGCCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCT ACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGG TTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAAT	0 nt 상동성 1

[0997]

		CAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	314 nt	TAAATGGCCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTA CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGT TTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC AACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	313 nt	AAATGGCCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTAC GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTT TGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCA ACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	312 nt	AATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACG TATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTT GACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAA CGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	311 nt	ATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGT ATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTG ACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAAC GGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTG GGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	310 nt	TGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTA TTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGA CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG GGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	309 nt	GGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTAT TAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGAC TCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	308 nt	GCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATT AGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACT	0 nt 상동성 1

[0998]

		CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	307 nt	CCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTA GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTC ACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGG ACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	306 nt	CCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAG TCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCA CGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGA CTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG GTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	305 nt	CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT CATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCAC GGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGAC TTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGG TCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	304 nt	GCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTC ATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACG GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTT TCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC TATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	303 nt	CCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCAT CGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGG GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	302 nt	CTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC GCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGG GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[0999]

	301 nt	TGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC GCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGG GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	300 nt	GGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCG CTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA AAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	299 nt	GCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCT ATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAA ATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA AGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	298 nt	CATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTA TTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTT CCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAA TGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA GCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	297 nt	ATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT TACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTC CAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAAT GTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	296 nt	TTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATT ACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCC AAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG TCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC AGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	295 nt	TATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTA CCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCA AGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT	0 nt 상동성 1

[1000]

		CGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCA GAGCTCGTTTAGTGAACCGTC	
	294 nt	ATGCCCAGTACATGACCTTACGGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC CATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	293 nt	TGCCCAGTACATGACCTTACGGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACC ATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	292 nt	GCCCAGTACATGACCTTACGGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCA TGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG TAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGA GCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	291 nt	CCCAGTACATGACCTTACGGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCAT GGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGT CTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT AATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAG CTCGTTTAGTGAACCGTC	0 nt 상동성 1
	290 nt	CCAGTACATGACCTTACGGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG GTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCT CCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA TAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1
	289 nt	CAGTACATGACCTTACGGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATG GTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCT CCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA TAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1
	288 nt	AGTACATGACCTTACGGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT GATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC	0 nt 상동성 1

[1001]

	287 nt	GTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGT GATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATA ACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	0 nt 상동성 1
	286 nt	TACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGA TGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCAC CCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAAC CCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTT TAGTGAACCGTC	0 nt 상동성 1
	285 nt	ACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGAT GCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACC CCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACC CCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT AGTGAACCGTC	0 nt 상동성 1
	284 nt	CATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGAT GCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACC CCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACC CCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT AGTGAACCGTC	0 nt 상동성 1
	283 nt	ATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATG CGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCC CATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCC CGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTA GTGAACCGTC	0 nt 상동성 1
	282 nt	TGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGC GGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCC ATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCC GCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAG TGAACCGTC	0 nt 상동성 1

[1002]

	281 nt	GACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCG GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCA TTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCG CCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGT GAACCGTC	0 nt 상동성 1
	280 nt	ACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGG TTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	0 nt 상동성 1
	279 nt	CCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGT TTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	0 nt 상동성 1
	278 nt	CTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTT TGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGA CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	277 nt	TTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT GGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGAC GTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	276 nt	TACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTT GGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGAC GTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	275 nt	ACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTG GCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACG TCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGT	0 nt 상동성 1

[1003]

	274 nt	TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCG TC CGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGG CAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACACCCATTGACGT CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTT GACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGT C	0 nt 상동성 1
	273 nt	GGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGC AGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACACCCATTGACGTC AATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTG ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	272 nt	GGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCA GTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACACCCATTGACGTCA ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGA CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	271 nt	GACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAG TACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACACCCATTGACGTCAA TGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	270 nt	ACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT ACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACACCCATTGACGTCAAT GGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	269 nt	CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACACCCATTGACGTCAATGG GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	268 nt	TTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACACCCATTGACGTCAATGG GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	267 nt	TTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACA CCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACACCCATTGACGTCAATGGG	0 nt 상동성 1

[1004]

	AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAA ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
266 nt	TCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACAC CAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGA GTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
265 nt	CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACC AATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAAT GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
264 nt	CTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCA ATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGT TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
263 nt	TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAA TGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTT TGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
262 nt	ACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAAT GGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGG GCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
261 nt	CTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATG GGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTG TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGG CGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
260 nt	TTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGG GCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGC GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
259 nt	TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG CGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1005]

	258 nt	GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG CGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	257 nt	GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGC GTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTT GGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGT AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	256 nt	CAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGT GGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTG GCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTA GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	255 nt	AGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	254 nt	GTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	253 nt	TACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGA TAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCAC CAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCG TGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	252 nt	ACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGAT AGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC AAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGT GTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	251 nt	CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATA GCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	250 nt	ATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATA GCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA	0 nt 상동성 1

[1006]

		AAATCAACGGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	
	249 nt	TCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAG CGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAA AATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	248 nt	CTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGC GGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA ATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTA CGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	247 nt	TACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCG GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAA TCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTAC GGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	246 nt	ACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGG TTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAAT CAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	245 nt	CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGT TTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC AACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	244 nt	GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTT TGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCA ACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	243 nt	TATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTT GACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAA CGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	242 nt	ATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTG ACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAAC GGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTG GGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1007]

	241 nt	TTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGA CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG GGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	240 nt	TAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGAC TCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	239 nt	AGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACT CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	238 nt	GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTC ACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGG ACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	237 nt	TCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCA CGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGA CTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG GTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	236 nt	CATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCAC GGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGAC TTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGG TCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	235 nt	ATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACG GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTT TCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC TATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	234 nt	TCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGG GGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTT CCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCT ATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	233 nt	CGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGG GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC	0 nt 상동성 1

[1008]

		CAAAATGTCGTAATAACCCCGCCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	
	232 nt	GCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGG GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	231 nt	CTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA AAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	230 nt	TATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGAT TTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAA AATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATAT AAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	229 nt	ATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATT TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAA ATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA AGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	228 nt	TTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTT CCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAA TGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA GCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	227 nt	TACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTC CAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAAT GTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	226 nt	ACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCC AAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG TCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC AGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	225 nt	CCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCA AGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT CGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCA GAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1009]

	224 nt	CATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	223 nt	ATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAA GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	222 nt	TGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAG TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG TAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGA GCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	221 nt	GGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGT CTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT AATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAG CTCGTTTAGTGAACCGTC	0 nt 상동성 1
	220 nt	GTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCT CCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA TAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1
	219 nt	TGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTC CACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAAT AACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTC GTTTAGTGAACCGTC	0 nt 상동성 1
	218 nt	GATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCC ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATA ACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	0 nt 상동성 1
	217 nt	ATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCA CCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAA CCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGT TTAGTGAACCGTC	0 nt 상동성 1
	216 nt	TGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCAC CCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAAC	0 nt 상동성 1

[1010]

		CCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTT TAGTGAACCGTC	
	215 nt	GCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACC CCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACC CCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT AGTGAACCGTC	0 nt 상동성 1
	214 nt	CGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCC CATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCC CGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTA GTGAACCGTC	0 nt 상동성 1
	213 nt	GGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCC ATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCC GCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAG TGAACCGTC	0 nt 상동성 1
	212 nt	GTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCA TTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCG CCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGT GAACCGTC	0 nt 상동성 1
	211 nt	TTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	0 nt 상동성 1
	210 nt	TTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATT GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	0 nt 상동성 1
	209 nt	TTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTG ACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCC CGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAA CCGTC	0 nt 상동성 1
	208 nt	TGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGA CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1

[1011]

	207 nt	GGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGAC GTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	206 nt	GCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACG TCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGT TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCG TC	0 nt 상동성 1
	205 nt	CAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGT CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTT GACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGT C	0 nt 상동성 1
	204 nt	AGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTC AATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTG ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	203 nt	GTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCA ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGA CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	202 nt	TACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAA TGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	201 nt	ACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAAT GGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	200 nt	CACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATG GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGC AAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	199 nt	ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGG GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	198 nt	CCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGG AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAA ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1012]

	197 nt	CAATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGA GTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	196 nt	AATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAAT GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	195 nt	ATGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGT TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	194 nt	TGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTT TGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	193 nt	GGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGG GCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	192 nt	GGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTG TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGG CGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	191 nt	GCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGC GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	190 nt	CGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	189 nt	GTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTT GGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGT AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	188 nt	TGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTG GCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTA GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	187 nt	GGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTG GCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTA GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1013]

	186 nt	GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	185 nt	ATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCA CCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGC GTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	184 nt	TAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCAC CAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCG TGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	183 nt	AGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC AAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGT GTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	182 nt	GCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	181 nt	CGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAA AATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	180 nt	GGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAA ATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTA CGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	179 nt	GTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAA TCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTAC GGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	178 nt	TTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAAT CAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	177 nt	TTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC AACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	176 nt	TGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCA ACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1014]

	175 nt	GACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAA CGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	174 nt	ACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAAC GGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTG GGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	173 nt	CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG GGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	172 nt	TCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	171 nt	CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	170 nt	ACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGG ACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	169 nt	CGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGA CTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG GTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	168 nt	GGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGAC TTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGG TCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	167 nt	GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTT TCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC TATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	166 nt	GGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTT CCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCT ATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	165 nt	GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1015]

	164 nt	ATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCC AAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTAT ATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	163 nt	TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA AAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	162 nt	TTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAA AATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATAT AAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	161 nt	TCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAA ATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA AGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	160 nt	CCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAA TGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA GCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	159 nt	CAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAAT GTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	158 nt	AAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATG TCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC AGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	157 nt	AGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT CGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCA GAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	156 nt	GTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTC GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	155 nt	TCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCG TAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGA GCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	154 nt	CTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGT AATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAG CTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1016]

	153 nt	TCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTA ATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGC TCGTTTAGTGAACCGTC	0 nt 상동성 1
	152 nt	CCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAA TAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1
	151 nt	CACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAAT AACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTC GTTTAGTGAACCGTC	0 nt 상동성 1
	150 nt	ACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATA ACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	0 nt 상동성 1
	149 nt	CCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAA CCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGT TTAGTGAACCGTC	0 nt 상동성 1
	148 nt	CCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAAC CCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTT TAGTGAACCGTC	0 nt 상동성 1
	147 nt	CCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACC CCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT AGTGAACCGTC	0 nt 상동성 1
	146 nt	CATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCC CGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTA GTGAACCGTC	0 nt 상동성 1
	145 nt	ATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCC GCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAG TGAACCGTC	0 nt 상동성 1
	144 nt	TTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCG CCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGT GAACCGTC	0 nt 상동성 1
	143 nt	TGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCG CCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGT GAACCGTC	0 nt 상동성 1

[1017]

	142 nt	GACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGC CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTG AACCGTC	0 nt 상동성 1
	141 nt	ACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCC CGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAA CCGTC	0 nt 상동성 1
	140 nt	CGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	139 nt	GTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCC GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAAC CGTC	0 nt 상동성 1
	138 nt	TCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGT TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCG TC	0 nt 상동성 1
	137 nt	CAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTT GACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGT C	0 nt 상동성 1
	136 nt	AATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTG ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	135 nt	ATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGA CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	134 nt	TGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	133 nt	GGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	132 nt	GGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGC AAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	131 nt	GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	130 nt	AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAA ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	129 nt	GTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1018]

	128 nt	TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAAT GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	127 nt	TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	126 nt	TGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	125 nt	GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGG GCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	124 nt	TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGG CGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	123 nt	TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGC GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	122 nt	TTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCG GTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	121 nt	TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	120 nt	GGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGT AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	119 nt	GCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTA GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	118 nt	CACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAG GCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	117 nt	ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	116 nt	CCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGC GTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	115 nt	CAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCG TGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	114 nt	AAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGT GTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	113 nt	AAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	112 nt	AATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGT ACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1019]

	111 nt	ATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTA CGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	110 nt	TCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTAC GGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	109 nt	CAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	108 nt	AACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	107 nt	ACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	106 nt	CGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	105 nt	GGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTG GGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	104 nt	GGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	103 nt	GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	102 nt	ACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	101 nt	CTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAG GTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	100 nt	TTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGG TCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	99 nt	TTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGT CTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	98 nt	TCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC TATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	97 nt	CCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCT ATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	96 nt	CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	95 nt	AAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTAT ATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1020]

	94 nt	AAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	93 nt	AATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATAT AAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	92 nt	ATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATA AGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	91 nt	TGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAA GCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	90 nt	GTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	89 nt	TCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGC AGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	88 nt	CGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCA GAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	87 nt	GTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAG AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	86 nt	TAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGA GCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	85 nt	AATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAG CTCGTTTAGTGAACCGTC	0 nt 상동성 1
	84 nt	ATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGC TCGTTTAGTGAACCGTC	0 nt 상동성 1
	83 nt	TAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCT CGTTTAGTGAACCGTC	0 nt 상동성 1
	82 nt	AACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTC GTTTAGTGAACCGTC	0 nt 상동성 1
	81 nt	ACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCG TTTAGTGAACCGTC	0 nt 상동성 1
	80 nt	CCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGT TTAGTGAACCGTC	0 nt 상동성 1
	79 nt	CCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTT TAGTGAACCGTC	0 nt 상동성 1
	78 nt	CCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTT AGTGAACCGTC	0 nt 상동성 1

[1021]

	77 nt	CGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	76 nt	GCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	75 nt	CCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	74 nt	CCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	73 nt	CCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	72 nt	CGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	71 nt	GTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	70 nt	TTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	69 nt	TGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	68 nt	GACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	67 nt	ACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	66 nt	CGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	65 nt	GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	64 nt	CAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	63 nt	AAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	62 nt	AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	61 nt	ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	60 nt	TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	59 nt	GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	58 nt	GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	57 nt	GCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1022]

	56 nt	CGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	55 nt	GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	54 nt	GTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	53 nt	TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	52 nt	AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	51 nt	GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	50 nt	GCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	49 nt	CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	48 nt	GTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	47 nt	TGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	46 nt	GTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	45 nt	TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	44 nt	ACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	43 nt	CGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	42 nt	GGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	41 nt	GTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	40 nt	TGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	39 nt	GGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	38 nt	GGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	37 nt	GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	36 nt	AGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	35 nt	GGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	34 nt	GTCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	33 nt	TCTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	32 nt	CTATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	31 nt	TATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	30 nt	ATATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	29 nt	TATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1

[1023]

	28 nt	ATAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	27 nt	TAAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	26 nt	AAGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	25 nt	AGCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	24 nt	GCAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	23 nt	CAGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	22 nt	AGAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	21 nt	GAGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	20 nt	AGCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	19 nt	GCTCGTTTAGTGAACCGTC	0 nt 상동성 1
	18 nt	CTCGTTTAGTGAACCGTC	0 nt 상동성 1
	17 nt	TCGTTTAGTGAACCGTC	0 nt 상동성 1
	16 nt	CGTTTAGTGAACCGTC	0 nt 상동성 1
	15 nt	GTTTAGTGAACCGTC	0 nt 상동성 1
	14 nt	TTTAGTGAACCGTC	0 nt 상동성 1
	13 nt	TTAGTGAACCGTC	0 nt 상동성 1
	12 nt	TAGTGAACCGTC	0 nt 상동성 1
	11 nt	AGTGAACCGTC	0 nt 상동성 1
	10 nt	GTGAACCGTC	0 nt 상동성 1
	9 nt	TGAACCGTC	0 nt 상동성 1
	8 nt	GAACCGTC	0 nt 상동성 1
	7 nt	AACCGTC	0 nt 상동성 1
	6 nt	ACCGTC	0 nt 상동성 1
	5 nt	CCGTC	0 nt 상동성 1
	4 nt	CGTC	0 nt 상동성 1
	3 nt	GTC	0 nt 상동성 1
	2 nt	TC	0 nt 상동성 1
	1 nt	C	0 nt 상동성 1

[1024]

	0 nt		0 nt 삼동성 1
	(0-500 nt)	GCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGTAA CGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAGTA CATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGCAT TATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTA CCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCA AGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGT CGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCA GAGCTCGTTTAGTGAACCGT	1 nt 삼동성 1
	(0-500 nt)	GGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAG TACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGC ATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT TACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTC CAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAAT GTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTCGTTTAGTGAACCG	2 nt 삼동성 1
	(0-500 nt)	TGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATAGT AACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGCAG TACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCTGGC ATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTAT TACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTC CAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAAT GTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAG CAGAGCTCGTTTAGTGAACC	3 nt 삼동성 1
	(0-500 nt)	ATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGC AGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCG CTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA AAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGAAC	4 nt 삼동성 1

[1025]

	(0-500 nt)	AATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCATA GTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGGC AGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCG CTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA AAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGAA	5 nt 상동성 1
	(0-500 nt)	AAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCAT AGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTGG CAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGCCT GGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCG CTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGA TTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCA AAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATA TAAGCAGAGCTCGTTTAGTGA	6 nt 상동성 1
	(0-500 nt)	TAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCCCA TAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTTG GCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCCGC CTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATC GCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACGGG GATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTC CAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTA TATAAGCAGAGCTCGTTTAGTG	7 nt 상동성 1
	(0-500 nt)	GTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTCC CATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACT TGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCC GCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTC ATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACG GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTT TCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC TATATAAGCAGAGCTCGTTTAGT	8 nt 상동성 1
	(0-500 nt)	GGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTTC CCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCAC	9 nt 상동성 1

[1026]

	TTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCC GCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGTC ATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCACG GGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGACTT TCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTC TATATAAGCAGAGCTCGTTTAG	10 nt 삼동성 1
(0-500 nt)	CGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGTT CCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCA CTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGGCCC CGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTAGT CATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTCAC GGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGGAC TTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGAGG TCTATATAAGCAGAGCTCGTTTA	11 nt 삼동성 1
(0-500 nt)	ACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATGT TCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCC ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGG CCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTA GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTC ACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGG ACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA GGTCTATATAAGCAGAGCTCGTTT	12 nt 삼동성 1
(0-500 nt)	TACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTATG TTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCC ACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATGG CCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATTA GTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACTC ACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGGG ACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGGA GGTCTATATAAGCAGAGCTCGTT	13 nt 삼동성 1
(0-500 nt)	TTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTAT GTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGC CCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAATG GCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTATT	

[1027]

		AGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGACT CACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCGT	14 nt 상동성 1
	(0-500 nt)	CTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGTA TGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTG CCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAAT GGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTAT TAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGAC TCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACGG GACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG AGGTCTATATAAGCAGAGCTCG	15 nt 상동성 1
	(0-500 nt)	ACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGT ATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAA TGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTA TTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGA CTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG GGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCTC	16 nt 상동성 1
	(0-500 nt)	AACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACG TATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTAAA TGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACGTA TTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTTGA CTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG GGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCT	17 nt 상동성 1
	(0-500 nt)	TAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGAC GTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAA CTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGTA AATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTACG TATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTTT GACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAA	

[1028]

		CGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAGC	
	(0-500 nt)	ATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGA CGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAA ACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGGT AAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTAC GTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGTT TGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCA ACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACGGT GGGAGGTCTATATAAGCAGAG	18 nt 상동성 1
	(0-500 nt)	CATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATG ACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTA AACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACGG TAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCTA CGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGGT TTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATC AACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAGA	19 nt 상동성 1
	(0-500 nt)	ACATAACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAAT GACGTATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGT AAACTGCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTCCGCCCCCTATTGACGTCAATGACG GTAAATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTACGGGACTTTCCTACTTGGCAGTACATCT ACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATAGCGG TTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAAT CAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTGTACG GTGGGAGGTCTATATAAGCAG	20 nt 상동성 1
mCherry 삽입	(0-500 nt)	ATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATA GCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCACT ATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCC CATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGG	0 nt 상동성 1

[1029]

		CGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGC CCACCCTCGTGACCACCCTGACGTACG	
	(0-500 nt)	CATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGATA GCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCA AAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGTG TACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCACT ATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCC CATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGGG CGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGGC CCACCCTCGTGACCACCCTGACGTAC	1 nt 삼동정 1
	(5-500 nt)	ACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGAT AGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACC AAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCGT GTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCAC TATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGC CCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAGG GCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTGG CCCACCCTCGTGACCACCCTGACGTA	2 nt 삼동정 1
	(5-500 nt)	TACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTGGA TAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCAC CAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGGCG TGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCA CTATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTG CCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGAG GGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCTG GCCCACCCTCGTGACCACCCTGACGT	3 nt 삼동정 1
	(5-500 nt)	GTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACT CACTATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGA GGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCT GGGCCCACCCTCGTGACCACCCTGACG	4 nt 삼동정 1

[1030]

	(5-500 nt)	AGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGTG GATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGC ACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTAGG CGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACT CACTATAGGGCGGCCGGGAATTCGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGG TGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGCGA GGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCCCT GGCCCACCCTCGTGACCACCCTGAC
	(5-500 nt)	CAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGCGT GGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTG GCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGTA GGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGA CTCACTATAGGGCGGCCGGGAATTCGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGT GGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGC GAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCC CTGGCCCACCCTCGTGACCACCCTGA
	(5-500 nt)	GCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGGC GTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTT GGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGGT AGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACG ACTCACTATAGGGCGGCCGGGAATTCGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGT GGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGGGC GAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGTGCC CTGGCCCACCCTCGTGACCACCCTG
	(5-500 nt)	GGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG CGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATAC GACTCACTATAGGGCGGCCGGGAATTCGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGG GTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAGG GCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCGT GCCCTGGCCCACCCTCGTGACCACCCT
	(5-500 nt)	TGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGGG CGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTT

5 nt 상동성 1	
6 nt 상동성 1	
7 nt 상동성 1	
8 nt 상동성 1	
9 nt 상동성 1	

[1031]

		TGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGCGG TAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATAC GACTCACTATAGGGCGGCCGGGAATTCGTCGACTGGATCCCCGGGTACCGAGCTCGAATTCACTGGCCG GTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCACTGGCCGTCGTTTTACAACGTCGTGACTG GGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAG CGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCT GTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCC CTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCT CTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGAT TAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCC ACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTG ATTTATAAGGGATTTTGCCGATTTCGGCCTATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGC GAATTTTAACAAAATATTAACGTTTACAATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCC TATTTGTTTATTTTTCTAAATACATTCAAATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTT CAATAATATTGAAAAAGGAAGAGTATGAGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCG GCATTTTGCCTTCCTGTTTTTGCTCACCCAGAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTG GGAAGACTTTAAATCGTACGTGTTCCTGTCTAAAACCCCTCTTGTCGATTCTTACAACACTTTTTCGGA TCTTTCCACGATCGGTGCAGATTCCAAACAGGAATCTATCGCGCACGCTTCGAAAAGGCCATCCGTCAG GATGGCCTTCCCCATTATGATTCTTCTCGCTTCCGCATCGTGCGTCGCTTCGCTATTACGCCAGCTGGC GAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAA AACGACGGCCAGTGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGA ATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTG CAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAG TTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGG TTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCC TTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATT TAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCC CTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAAC TGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCCTAT TGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACAATT TCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCAAAT ATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATGAGT ATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCACCCA GAAACGCTGGTGAAAGTAAAAGATGCTGAAGATCAGTTGGGAAGACTTTAAATCGTACGTGTTCCTGTC TAAACAATTTTCACAGGTAGGTGCGGCGAGCCGCTTCGCCATTACGCCAGCTGGCGAAAGGGGGATGTG CTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTG CCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCACTGGCCGTC GTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCT TTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAA TGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGA CCGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTT CCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGA TTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCG CCCTGATAGACGGTTTTTCGCCCTTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAA ACTGGAACAACACTCAACCCTATCTCGGTCTATTCTTTTGATTTATAAGGGATTTTGCCGATTTCGGCC TATTGGTTAAAAAATGAGCTGATTTAACAAAAATTTAACGCGAATTTTAACAAAATATTAACGTTTACA ATTTCAGGTGGCACTTTTCGGGGAAATGTGCGCGGAACCCCTATTTGTTTATTTTTCTAAATACATTCA AATATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAGAGTATG AGTATTCAACATTTCCGTGTCGCCCTTATTCCCTTTTTTGCGGCATTTTGCCTTCCTGTTTTTGCTCAC CCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACG ACGTTGTAAAACGACGGCCAGTGCCAAGCTTGCATGCCTGCAGGTCGACTCTAGAGGATCCCCGGGTA CCGAGCTCGAATTCACTGGCCGTCGTTTTACAACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTT AATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCC TTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGG GTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTC TTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGG GTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGG GCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGAACGCAAATGGGCGGTAGGCGTGTACGGTGG GAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCACTATAGGGCGG CCGGGAATTCGTCGACTGGATCCCCGGGTACCGAGCTCGAATTCACTGGCCGTCGTTTTACAACGTCGT GACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCAGCTGGCG TAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAATGGGACG CGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCC AGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCG TCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAA AACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCCTTTGACG TTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGAACGCAAATGGGCGGTAGGCGTGTA CGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAATACGACTCACTATA GGGCGGCCGGGAATTCGTCGACTGGATCCCCGGGTACCGAGCTCGAATTCACTGGCCGTCGTTTTACA ACGTCGTGACTGGGAAAACCCTGGCGTTACCCAACTTAATCGCCTTGCAGCACATCCCCCTTTCGCCA GCTGGCGTAATAGCGAAGAGGCCCGCACCGATCGCCCTTCCCAACAGTTGCGCAGCCTGAATGGCGAA TGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTAC ACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCT TTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAGTGCTTTACGGCACCTCGAC CCCAAAAAACTTGATTAGGGTGATGGTTCACGTAGTGGGCCATCGCCCTGATAGACGGTTTTTCGCCC TTTGACGTTGGAGTCCACGTTCTTTAATAGTGGACTCTTGTTCCAAACTGAACGCAAATGGGCGG	
	(5-500 nt)	TTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATGG GCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGT TTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGGC GGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAAT ACGACTCACTATAGGGCGGCCGGGAATTCGTCGACTGGATCCCCGGGTACCGAGCTCGAATTCACTGG GGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGAG GGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCCG TGCCCTGGCCCACCCTCGTGACCACC	
	(5-500 nt)	CTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAATG GGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTG TTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGGG CGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGTAA TACGACTCACTATAGGGCGGCCGGGAATTCGTCGACTGGATCCCCGGGTACCGAGCTCGAATTCACCG GGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGCGA GGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGCCC GTGCCCTGGCCCACCCTCGTGACCAC	11 nt 삼통성 1
	(5-500 nt)	ACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAAT GGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTT GTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATGG GCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTGT AATACGACTCACTATAGGGCGGCCGGGAATTCGTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCAC CGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGGC GAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTGC CCGTGCCCTGGCCCACCCTCGTGACCA	12 nt 삼통성 1
	(5-500 nt)	TACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCAA TGGGCGTGGATAGCGGTTTGACTCACGGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTT TGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTG	13 nt 삼통성 1

[1032]

		TAATACGACTCACTATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCA CCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTG CCCGTGCCCTGGCCCACCCTCGTGACC	
	(5-500 nt)	CTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACCA ATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGT TTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAATG GGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTTTG TAATACGACTCACTATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTCA CCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCGG CGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCTG CCCGTGCCCTGGCCCACCCTCGTGAC	14 nt 삼동성 1
	(5-500 nt)	CCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACACC AATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAG TTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAAT GGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTTT GTAATACGACTCACTATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTTC ACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCCG GCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGCT GCCCGTGCCCTGGCCCACCCTCGTGA	15 nt 삼동성 1
	(5-500 nt)	TCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACAC CAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGA GTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAAA TGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATTT TGTAATACGACTCACTATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGTT CACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTCC GGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAGC TGCCCGTGCCCTGGCCCACCCTCGTG	16 nt 삼동성 1
	(5-500 nt)	TTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACA CCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGG AGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCAA ATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAATT TTGTAATACGACTCACTATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTGT TCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGTC	17 nt 삼동성 1

		CGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAAG
	(5-500 nt)	CTGCCCGTGCCCTGGCCCACCCTCGT
		TTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC
		ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGG
		GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA
		AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAAT
		TTTGTAATACGACTCACTATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTG
		TTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
		CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAA
		GCTGCCCGTGCCCTGGCCCACCCTCG
	(5-500 nt)	CTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTAC
		ACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGG
		GAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGACGCA
		AATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAGAAT
		TTTGTAATACGACTCACTATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAGCTG
		TTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCGTGT
		CCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGGCAA
		GCTGCCCGTGCCCTGGCCCACCCTC
	(5-500 nt)	ACTTTCCTACTTGGCAGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGT
		ACACCAATGGGCGTGGATAGCGGTTTGACTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAAT
		GGGAGTTTGTTTTGGCACCAAAATCAACGGGACTTTCCAAAATGTCGTAATAACCCCGCCCCGTTGAC
		GCAAATGGGCGGTAGGCGTGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGTGAACCGTCAG
		AATTTTGTAATACGACTCACTATAGGGCGGCCGGGAATTCGCCACCATGGTGAGCAAGGGCGAGGAG
		CTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGGACGGCGACGTAAACGGCCACAAGTTCAGCG
		TGTCCGGCGAGGGCGAGGGCGATGCCACCTACGGCAAGCTGACCCTGAAGTTCATCTGCACCACCGG
		CAAGCTGCCCGTGCCCTGGCCCACCCT

[1033]

[1034] *PBS 서열*

[1035] 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 PBS 서열을 포함한다. 일부 구현예에서, PBS 서열은 이종 목적 서열의 3'에 배치되고, 본 명세서에 기재된 시스템에 의해 변형될 부위에 인접한 서열에 상보적이거나, 시스템/유전자 변형 폴리펩티드에 의해 변형될 부위에 인접한 서열에 상보적인 서열에 대하여 1, 2, 3, 4, 또는 5개 이하의 미스매치를 포함한다. 일부 구현예에서, PBS 서열은 표적 핵산 분자 내 닉 부위의 1, 2, 3, 4, 5, 6, 7, 8, 9, 또는 10개 뉴클레오티드 내에 결합한다. 일부 구현예에서, 표적 핵산 분자에 대한 PBS 서열의 결합은, 예를 들어 표적-프라이밍된 역전사(TPRT)에 대한 프라이머로서 작용하는 3' 상동성 도메인을 이용하여, TPRT의 개시를 허용한다. 일부 구현예에서, PBS 서열은 길이가 3 내지 5, 5 내지 10, 10 내지 30, 10 내지 25, 10 내지 20, 10 내지 19, 10 내지 18, 10 내지 17, 10 내지 16, 10 내지 15, 10 내지 14, 10 내지 13, 10 내지 12, 10 내지 11, 11 내지 30, 11 내지 25, 11 내지 20, 11 내지 19, 11 내지 18, 11 내지 17, 11 내지 16, 11 내지 15, 11 내지 14, 11 내지 13, 11 내지 12, 12 내지 30, 12 내지 25, 12 내지 20, 12 내지 19, 12 내지 18, 12

내지 17, 12 내지 16, 12 내지 15, 12 내지 14, 12 내지 13, 13 내지 30, 13 내지 25, 13 내지 20, 13 내지 19, 13 내지 18, 13 내지 17, 13 내지 16, 13 내지 15, 13 내지 14, 14 내지 30, 14 내지 25, 14 내지 20, 14 내지 19, 14 내지 18, 14 내지 17, 14 내지 16, 14 내지 15, 15 내지 30, 15 내지 25, 15 내지 20, 15 내지 19, 15 내지 18, 15 내지 17, 15 내지 16, 16 내지 30, 16 내지 25, 16 내지 20, 16 내지 19, 16 내지 18, 16 내지 17, 17 내지 30, 17 내지 25, 17 내지 20, 17 내지 19, 17 내지 18, 18 내지 30, 18 내지 25, 18 내지 20, 18 내지 19, 19 내지 30, 19 내지 25, 19 내지 20, 20 내지 30, 20 내지 25, 또는 25 내지 30개의 뉴클레오티드, 예를 들어 길이가 10 내지 17, 12 내지 16, 또는 12 내지 14개의 뉴클레오티드이다. 일부 구현예에서, PBS 서열은 길이가 5 내지 20, 8 내지 16, 8 내지 14, 8 내지 13, 9 내지 13, 9 내지 12, 또는 10 내지 12개 뉴클레오티드, 예를 들어 길이가 9 내지 12개 뉴클레오티드이다.

[1036] 주형 핵산(예를 들어, 주형 RNA)은 표적 DNA에 대해 어느 정도 상동성을 가질 수 있다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA) PBS 서열 도메인은, 표적 DNA가 주형 핵산(예를 들어, 주형 RNA)의 역전사를 프라이밍하도록 위치하도록, 표적 DNA에 대한 어닐링 영역의 역할을 할 수 있다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 RNA의 3' 말단에서 표적 DNA에 대해 정확히 상동성인 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 175, 200개 또는 그 이상의 염기를 갖는다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)은 주형 핵산(예를 들어, 주형 RNA)의 5' 말단에서, 표적 DNA에 대해 적어도 50%, 60%, 70%, 80%, 85%, 90%, 95%, 97%, 98%, 99% 또는 100% 상동성인 적어도 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 20, 25, 30, 35, 40, 45, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 175, 200개 또는 그 이상의 염기를 갖는다.

[1037] 일부 구현예에서, PBS 서열은 표 37에 열거된 바와 같은 핵산 서열, 또는 이에 대하여 적어도 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, PBS 서열은 표 37에 열거된 바와 같은 핵산 서열, 또는 이에 대하여 1, 2, 3, 4, 또는 5개 이하의 뉴클레오티드 차이를 갖는 핵산 서열을 포함한다. 본 명세서에서, RNA 서열(예를 들어, PBS서열 내)이 티민(T)을 포함하는 특정 서열(예를 들어, 표 37의 서열 또는 이의 일부)을 포함한다고 말할 때, 물론 RNA 서열은 T 대신에 우라실(U)을 포함할 수 있다(그리고 흔히 포함한다)고 이해된다. 예를 들어, RNA 서열은 표 37의 서열에서 T로 나타낸 모든 위치에서 U를 포함할 수 있다. 더 구체적으로, 본 개시내용은 표 37의 모든 PBS 서열에 따른 RNA 서열을 제공하며, 여기서 RNA 서열은 표 37의 서열에서 각각의 T 대신에 U를 갖는다. 특정 구현예에서, PBS는 1 내지 3, 3 내지 5, 5 내지 8, 8 내지 10, 10 내지 12, 12 내지 15, 15 내지 17, 17 내지 20, 20 내지 25, 25 내지 30, 또는 30 내지 40개 뉴클레오티드의 길이를 갖는다. 특정 구현예에서, PBS는 약 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 또는 25개 뉴클레오티드의 길이를 갖는다.

[1038] **표 37. 예시적인 PBS 서열.** 하기 서열은 5'에서 3' 방향으로 열거되어 있다.

PBS 길이	서열(5'에서 3')
1 nt	G
2 nt	GC
3 nt	GCG
4 nt	GCGT
5 nt	GCGTG
6 nt	GCGTGC
7 nt	GCGTGCA
8 nt	GCGTGCAG
9 nt	GCGTGCAGT
10 nt	GCGTGCAGTG
11 nt	GCGTGCAGTGC
12 nt	GCGTGCAGTGCT
13 nt	GCGTGCAGTGCTT
14 nt	GCGTGCAGTGCTTC
15 nt	GCGTGCAGTGCTTCG
16 nt	GCGTGCAGTGCTTCGG
17 nt	GCGTGCAGTGCTTCGGC

[1039]

[1040] ***유도성 활성을 갖는 gRNA***

[1041] 일부 구현예에서, 본 명세서에 기재된 gRNA(예를 들어, 제2 가닥 닉 생성에 사용된 주형 RNA 또는 gRNA의 일부인 gRNA)는 유도성 활성을 갖는다. 유도성 활성은 (gRNA에 추가하여) 차단 도메인을 추가로 포함하는 주형 핵산, 예를 들어 주형 RNA에 의해 달성될 수 있으며, 여기서 차단 도메인의 일부 또는 전체 서열은 gRNA의 일부 또는 전체에 대해 적어도 부분적으로 상보적이다. 따라서 차단 도메인은 gRNA의 일부 또는 전체에 혼성화 또는 실질적으로 혼성화할 수 있다. 일부 구현예에서, gRNA가 차단 도메인이 gRNA에 혼성화되거나 또는 실질적으로 혼성화되는 제1 입체형태, 및 차단 도메인이 gRNA에 혼성화되지 않거나 또는 실질적으로 혼성화되지 않는 제2 입체형태를 채용할 수 있도록, 차단 도메인 및 유도성 활성 gRNA는 주형 핵산, 예를 들어 주형 RNA 상에 배치된다. 일부 구현예에서, 제1 입체형태에서, gRNA는 유전자 변형 폴리펩티드(예를 들어, 주형 핵산 결합 도메인, DNA 결합 도메인, 또는 엔도뉴클레아제 도메인(예를 들어, CRISPR/Cas 단백질))에 결합할 수 없거나 또는 차단 도메인이 결여된 다른 유사한 주형 RNA와 비교하여 실질적으로 감소된 친화도로 결합한다. 일부 구현예에서, 제2 입체형태에서, gRNA는 유전자 변형 폴리펩티드(예를 들어, 주형 핵산 결합 도메인, DNA 결합 도메인, 또는 엔도뉴클레아제 도메인(예를 들어, CRISPR/Cas 단백질))에 결합할 수 있다. 일부 구현예에서, gRNA가 제1 또는 제2 입체형태인지 여부는 DNA 결합 또는 (예를 들어, CRISPR/Cas 단백질 중 유전자 변형 폴리펩티드가 포함하는) 유전자 변형 폴리펩티드의 엔도뉴클레아제 활성이 활성인지 여부에 영향을 미칠 수 있다.

[1042] 일부 구현예에서, 제2 닉을 조정하는 gRNA는 유도성 활성을 갖는다. 일부 구현예에서, 제2 닉을 조정하는 gRNA는 주형이 역전사 후에 유도된다. 일부 구현예에서, gRNA의 차단 도메인으로의 혼성화는 오프너(opener) 분자를 사용하여 붕괴될 수 있다. 일부 구현예에서, 오프너 분자는 gRNA 또는 차단 도메인의 일부 또는 전체에 결합하는 작용제를 포함하고, gRNA의 차단 도메인에 대한 혼성화를 억제한다. 일부 구현예에서, 오프너 분자는, 예를 들어 gRNA, 차단 도메인, 또는 둘 다에 부분적으로 또는 전체적으로 상보적인 서열을 포함하는 핵산을 포함한다. 적절한 오프너 분자를 선택 또는 설계함으로써, 오프너 분자의 제공은, CRISPR/Cas 단백질과 회합하고 CRISPR/Cas 단백질의 회합된 기능(예를 들어, DNA 결합 및/또는 엔도뉴클레아제 활성)을 제공하도록, gRNA의 입체형태에서의 변화를 촉진할 수 있다. 이론에 구속되고자 함이 없이, 선택된 시간 및/또는 위치에서 오프너 분자의 제공은 gRNA, CRISPR/Cas 단백질, 또는 이를 포함하는 유전자 변형 시스템의 활성의 공간적 및 시간적 제어를 가능하게 할 수 있다. 일부 구현예에서, 오프너 분자는 유전자 변형 폴리펩티드 및 또는 주형 핵산을 포함하는 세포에 대해 외인성이다. 일부 구현예에서, 오프너 분자는 (예를 들어, 유전자 변형 폴리펩티드 및 또는 gRNA를 포함하는 주형 핵산 및 차단 도메인을 포함하는 세포에 대해 내인성인) 내인성 작용제를 포함한다. 예를

들어, 유도성 gRNA, 차단 도메인, 및 오프너 분자는, 오프너 분자가 표적 세포 또는 조직에서 발현된 내인성 작용제이도록 선택될 수 있고, 예를 들어 이에 의해 표적 세포 또는 조직에서 유전자 변형 시스템의 활성을 보장한다. 추가의 예로서, 유도성 gRNA, 차단 도메인, 및 오프너 분자는, 오프너 분자가 하나 이상의 비-표적 세포 또는 조직에서 부재하거나 실질적으로 발현되지 않도록 선택될 수 있고, 예를 들어 이에 의해 유전자 변형 시스템의 활성이 하나 이상의 비-표적 세포 또는 조직에서 발생하지 않거나 실질적으로 발생하지 않거나, 또는 표적 세포 또는 조직과 비교하여 감소된 수준으로 발생하는 것을 보장한다. 예시적인 차단 도메인, 오프너 분자, 및 이의 용도는 PCT 출원 공개 WO2020044039A1에 기재되어 있으며, 이는 그 전체가 본 명세서에 참고로 포함된다. 일부 구현예에서, 주형 핵산, 예를 들어 주형 RNA는, 예를 들어 역전사효소 또는 RNA 결합 도메인, 및 gRNA에 의해, 유전자 변형 폴리펩티드의 하나 이상의 구성요소에 의한 결합을 위한 하나 이상의 서열 또는 구조를 포함할 수 있다. 일부 구현예에서, gRNA는 유전자 변형 폴리펩티드의 주형 핵산 결합 도메인(예를 들어, RNA 결합 도메인)과의 상호작용을 용이하게 한다. 일부 구현예에서, gRNA는, 예를 들어 표적 세포 게놈에서 유전자 변형 폴리펩티드를 매칭 표적에 지시한다.

[1043] *추가 가이드 RNA*

[1044] 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 시스템은, 예를 들어 표적 핵산의 풀림 또는 닉 생성(예를 들어, 주형 RNA의 PBS 서열에 의해 인식되는 반대 가닥의 닉 생성)을 위한 추가 가이드 RNA(gRNA)를 포함한다. 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 시스템은 1, 2, 3, 4, 또는 5개의 추가적인 별개의 gRNA를 포함한다. 특정 구현예에서, 추가 가이드 RNA는 주형 RNA와 별도의 분자이다. 다른 구현예에서, 추가 가이드 RNA는 (예를 들어, 주형 RNA의 3' 말단에서) 주형 RNA에 부착되거나 이에 통합된다. 구현예에서, 추가 가이드 RNA는 링커 영역에 의해 주형 RNA의 나머지 부분에 부착된다. 일부 구현예에서, 추가 가이드 RNA는, 예를 들어 표 44에 언급된 바와 같이 스템-루프 서열을 포함한다.

[1045] 일부 구현예에서, 추가 gRNA는, 예를 들어 주형 RNA의 말단 블록 서열(예를 들어, 본 명세서에 기재된 바와 같음)에 부착된, 프로-스페이서를 포함한다. 일부 구현예에서, 프로-스페이서는 17개 뉴클레오티드 이상(예를 들어, 적어도 약 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 또는 40개 뉴클레오티드)이다. 구현예에서, 프로-스페이서는 Cas 도메인(예를 들어, Cas9 도메인, 예를 들어 nCas9 도메인)에 의해 표적 핵산의 닉 생성을 지시한다. 다른 구현예에서, 프로-스페이서는 길이가 17개 뉴클레오티드 이하(예를 들어, 약 5, 10, 11, 12, 13, 14, 15, 16, 또는 17개 뉴클레오티드 길이)이다. 구현예에서, 17개 뉴클레오티드 이하 길이의 프로-스페이서는 (예를 들어, Cas 도메인, 예를 들어 Cas9 도메인, 예를 들어 dCas9 도메인에 의해) 표적 핵산의 풀림을 지시하지만, Cas 도메인(예를 들어, Cas9 도메인, 예를 들어 nCas9 도메인)에 의해 표적 핵산의 닉 생성을 지시하지 않는다.

[1046] 일부 구현예에서, 추가 gRNA는 표 44에 열거된 바와 같은 스캐폴드 서열, 또는 이에 대하여 적어도 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99% 동일성을 갖는 핵산 서열을 포함한다. 일부 구현예에서, 추가 gRNA는 표 44에 열거된 바와 같은 스캐폴드 서열, 또는 이와 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 40, 또는 50개 이하의 뉴클레오티드 서열을 갖는 핵산 서열을 포함한다.

[1047] 본 명세서에서, RNA 서열(예를 들어, gRNA 서열)이 티민(T)을 포함하는 특정 서열(예를 들어, 표 44의 서열 또는 이의 일부)을 포함한다고 말할 때, 물론 RNA 서열은 T 대신에 우라실(U)을 포함할 수 있다(그리고 흔히 포함한다)고 이해된다. 예를 들어, RNA 서열은 표 44의 서열에서 T로 나타낸 모든 위치에서 U를 포함할 수 있다. 더 구체적으로, 본 개시내용은 표 44의 모든 gRNA 서열에 따른 RNA 서열을 제공하며, 여기서 RNA 서열은 표 44의 서열에서 각각의 T 대신에 U를 갖는다.

[1048] **표 44. 예시적인 추가 가이드 RNA(gRNA) 서열**

Cas9 변이체	프로-스페이서 서열(5'에서 3')	스캐폴드 서열(5'에서 3')
Nme2Cas9 스캐폴드		GTTGTAGCTCCCTTTCTCATTTCGGAAACGAAATGAGAACC GTTGCTACAATAAGGCCGTCTGAAAAGATGTGCCGCAACG CTCTGCCCCTTAAAGCTTCTGCTTTAAGGGGCATCGTTTA
Nme2Cas9 **스페이서+** 스캐폴드	**CAGTACATGACCTT ACGGGA**	GTTGTAGCTCCCTTTCTCATTTCGGAAACGAAATGAGAACC GTTGCTACAATAAGGCCGTCTGAAAAGATGTGCCGCAACG CTCTGCCCCTTAAAGCTTCTGCTTTAAGGGGCATCGTTTA
Nme2Cas9 **16 nt 스페이서+** 스캐폴드	**ACATGACCTTACGG GA**	GTTGTAGCTCCCTTTCTCATTTCGGAAACGAAATGAGAACC GTTGCTACAATAAGGCCGTCTGAAAAGATGTGCCGCAACG CTCTGCCCCTTAAAGCTTCTGCTTTAAGGGGCATCGTTTA
BlatCas9 스캐폴드		GCTATAGTTCCTTACTGAAAGGTAAGTTGCTATAGTAAGGG CAACAGACCCGAGGCGTTGGGGATCGCCTAGCCCGTGTTT ACGGGCTCTCCCCATATTCAAAATAATGACAGACGAGCACC TTGGAGCATTTATCTCCGAGGTGCT
GeoCas9		GTCATAGTTCCCCTGAGAAATCAGGGTTACTATGATAAGGG CTTTCTGCCTAAGGCAGACTGACCCGCGGCGTTGGGGATC GCCTGTCGCCCGCTTTTGGCGGGCATTCCCCATCCTT
PpnCas9 스캐폴드		GTTGTAGCTCCCTTTTTCATTTCGCGAAAGCGAAATGAAAA ACGTTGTTACAATAAGAGATGAATTTCTCGCAAAGCTCTGC CTCTTGAAATTTCGGTTTCAAGAGGCATC

[1049]

CdiCas9 스캐폴드		ACTGGGGTTCAGGAAACTGAACCTCAGTAAGCATTGGCTC GTTTCCAATGTTGATTGCTCCGCCGGTGCTCCTTATTTTTAA GGGCGCCGGC
SpyCas9+ **헤어핀** 스캐폴드		GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCC GTTATCAACTTGAAAAAGTGGCACCG**GGACTTCGGTC CCGGTGC**
St1Cas9 스캐폴드		GTCTTTGTACTCTGGTACCAGAAGCTACAAAGATAAGGCTT CATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTT TT
cCas9-v16 스캐폴드		GTCTTAGTACTCTGGAAACAGAATCTACTAAGACAAGGCAA AATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA
SpyCas9-3var-NRRH 스캐폴드		GTTTAAGAGCTATGCTGGAAACAGCATAGCAAGTTTAAATA AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTC GGTGC
SauCas9 스캐폴드		GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCAA AATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA
CjeCas9 스캐폴드		GTTTTAGTCCCTGAAAAGGGACTAAAATAAAGAGTTTGCG GGACTCTGCGGGGTTACAATCCCCTAAAACCGC
SpyCas9 스캐폴드		GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCC GTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC

[1050]

[1051] ***유전자 변형 시스템의 환형 RNA 및 리보자임***

[1052] 표적 세포 내에서 제형화, 전달, 또는 유전자 변형 반응 동안 환형 및/또는 선형 RNA 상태를 이용하는 것은 유용할 수 있는 것으로 고려된다. 따라서 본 명세서에 기재된 임의의 양태의 일부 구현예에서, 유전자 변형 시스템은 하나 이상의 환형 RNA(circRNA)를 포함한다. 본 명세서에 기재된 임의의 양태의 일부 구현예에서, 유전자 변형 시스템은 하나 이상의 선형 RNA를 포함한다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 핵산(예를 들어, 주형 핵산, 유전자 변형 폴리펩티드를 인코딩하는 핵산 분자, 또는 둘 다)은 circRNA이다. 일부 구현예에서, 환형 RNA 분자는 유전자 변형 폴리펩티드를 인코딩한다. 일부 구현예에서, 유전자 변형 폴리펩티드를 인코딩하는 circRNA 분자는 숙주 세포로 전달된다. 일부 구현예에서, 환형 RNA 분자는, 예를 들어 본 명세서에 기재된 바와 같은 재조합효소를 인코딩한다. 일부 구현예에서, 재조합효소를 인코딩하는 circRNA 분자는 숙주 세포

로 전달된다. 일부 구현예에서, 유전자 변형 폴리펩티드를 인코딩하는 circRNA 분자는 (예를 들어, 숙주 세포 내에서, 예를 들어 숙주 세포의 핵에서) 번역 전에 선형화된다.

[1053] 환형 RNA(circRNA)는 세포 내에서 자연 발생하는 것으로 발견되었으며, 인간 세포에서 비-코딩 및 단백질 코딩 역할을 둘 다 포함하여, 다양한 기능을 갖는 것으로 발견되어 왔다. circRNA는 자기-스플라이싱 인트론을 RNA 분자(또는 RNA 분자를 인코딩하는 DNA)로 혼입시킴으로써 조작되어 RNA의 환형화를 초래할 수 있으며, 그 조작된 cirRNA는 향상된 단백질 생산 및 안정성을 가질 수 있는 것으로 나타난다(Wesselhoeft et al. *Nature Communications* 2018). 일부 구현예에서, 유전자 변형 폴리펩티드는 circRNA로서 인코딩된다. 일부 구현예에서, 주형 핵산은 DNA, 예컨대 dsDNA 또는 ssDNA이다. 특정 구현예에서, circDNA는 주형 RNA를 포함한다.

[1054] 일부 구현예에서, circRNA는 하나 이상의 리보자임 서열을 포함한다. 일부 구현예에서, 리보자임 서열은 예를 들어 숙주 세포에서 자가절단에 의해 활성화되며, 예를 들어 이에 의해 circRNA의 선형화를 초래한다. 일부 구현예에서, 리보자임은 마그네슘 농도가, 예를 들어 숙주 세포 내에서, 절단에 충분한 수준으로 도달하는 경우 활성화된다. 일부 구현예에서, circRNA는 숙주 세포로의 전달 전에 낮은 마그네슘 환경으로 유지된다. 일부 구현예에서, 리보자임은 단백질-반응성 리보자임이다. 일부 구현예에서, 리보자임은 핵산-반응성 리보자임이다. 일부 구현예에서, circRNA는 절단 부위를 포함한다. 일부 구현예에서, circRNA는 제2 절단 부위를 포함한다.

[1055] 일부 구현예에서, circRNA는 표적 세포의 핵에서 선형화된다. 일부 구현예에서, 세포의 핵에서 circRNA의 선형화는, 예를 들어 절단 사건을 활성화시키기 위해, 세포의 핵에 존재하는 구성요소를 포함한다. 일부 구현예에서, 핵 요소, 예를 들어 핵 단백질, 예를 들어 게놈-상호작용 단백질, 예를 들어 후성유전학적 변형자, 예를 들어 EZH2에 반응성인 리보자임, 예를 들어 B2 또는 ALU 요소로부터의 리보자임이, 예를 들어 유전자 변형 시스템의 circRNA 내로 포함된다. 일부 구현예에서, circRNA의 핵 국재화는 리보자임의 자가촉매 활성 및 circRNA의 선형화의 증가를 초래한다.

[1056] 일부 구현예에서, 리보자임은 유전자 변형 시스템의 다른 구성요소 중 하나 이상에 대해 이종성이다. 일부 구현예에서, 유도성 리보자임(예를 들어, 본 명세서에 기재된 바와 같은 circRNA)은, 예를 들어 단백질 리간드-반응성 압타머 설계를 이용함으로써, 합성적으로 형성된다. 담배 둥근무늬 바이러스 해머헤드 리보자임의 위성 RNA를 MS2 외피 단백질 압타머와 함께 이용하기 위한 시스템이 문헌[Kennedy et al. Nucleic Acids Res 42(19):12306-12321 (2014)(그 전체가 본 명세서에 참고로 포함됨)]에 기재되었으며, MS2 외피 단백질의 존재하에 리보자임 활성의 활성화를 초래한다. 구현예에서, 그러한 시스템은 세포질 또는 핵에 국재화된 단백질 리간드에 반응한다. 일부 구현예에서, 단백질 리간드는 MS2가 아니다. 리간드를 표적화하기 위해 RNA 압타머를 생성하기 위한 방법은, 예를 들어 지수적 농축에 의한 리간드의 계통적 진화(systematic evolution of ligands by exponential enrichment; SELEX)를 기반으로 기재되었으며(Tuerk and Gold, Science 249(4968):505-510 (1990); Ellington and Szostak, Nature 346(6287):818-822 (1990); 이들 각각의 방법은 본 명세서에 참고로 포함됨], 일부 경우에서, 인실리코 설계의 도움을 받는다(Bell et al. PNAS 117(15):8486-8493, 이 방법은 본 명세서에 참고로 포함됨). 따라서 일부 구현예에서, 표적 리간드에 대한 압타머가 생성되고, 예를 들어 단백질 리간드 존재 하에서, 예를 들어 리보자임-매개 절단 및 circRNA 선형화를 촉발하기 위해 합성 리보자임 시스템 내로 혼입된다. 일부 구현예에서, circRNA 선형화는, 예를 들어 세포질 내 리간드와 회합하는 압타머를 사용하여, 세포질 내에서 촉발된다. 일부 구현예에서, circRNA 선형화는, 예를 들어 핵 내 리간드와 회합하는 압타머를 사용하여, 핵 내에서 촉발된다. 구현예에서, 핵 내 리간드는 후성유전학적 변형자 또는 전사 인자를 포함한다. 일부 구현예에서, 선형화를 촉발하는 리간드는 표적-외 세포보다 표적상 세포에서 더 높은 수준으로 존재한다.

[1057] 핵산-반응성 리보자임 시스템이 circRNA 선형화를 위해 이용될 수 있다는 것이 추가로 고려된다. 예를 들어, 리보자임 활성화를 촉발하기 위해 정의된 표적 핵산 분자를 감지하는 바이오센서는, 예를 들어 문헌[Penchovsky, Biotechnology Advances 32(5):1015-1027 (2014)(본 명세서에 참고로 포함됨)]에 기재되어 있다. 이들 방법에 의해, 리보자임은 비활성 상태로 자연적으로 접히고, 정의된 표적 핵산 분자(예를 들어, RNA 분자) 존재 하에서만 활성화된다. 일부 구현예에서, 유전자 변형 시스템의 circRNA는 정의된 표적 핵산, 예를 들어 RNA, 예를 들어 mRNA, miRNA, 가이드 RNA, gRNA, sgRNA, ncRNA, lncRNA, tRNA, snRNA, 또는 mtRNA 존재 하에서 활성화되는 핵산-반응성 리보자임을 포함한다. 일부 구현예에서, 선형화를 촉발하는 핵산은 표적-외 세포보다 표적상 세포에서 더 높은 수준으로 존재한다.

[1058] 본 명세서의 임의의 양태의 일부 구현예에서, 유전자 변형 시스템은 관심 표적 조직 또는 표적 세포에 대한 유도성 특이성을 갖는 하나 이상의 리보자임, 예를 들어 관심 표적 조직 또는 표적 세포에서 더 높은 수준으로 존

재하는 리간드 또는 핵산에 의해 활성화되는 리보자임을 포함한다. 일부 구현예에서, 유전자 변형 시스템은 세포내 구획, 예를 들어, 핵, 핵소체, 세포질, 또는 미토콘드리아에 대해 유도성 특이성을 갖는 리보자임을 포함한다. 일부 구현예에서, 표적 세포내 구획에서 더 높은 수준으로 존재하는 리간드 또는 핵산에 의해 활성화되는 리보자임. 일부 구현예에서, 유전자 변형 시스템의 RNA 구성요소는, 예를 들어 선형화에 의해 활성화되는 circRNA로서 제공된다. 일부 구현예에서, 유전자 변형 폴리펩티드를 인코딩하는 circRNA의 선형화는 분자를 번역에 대해 활성화시킨다. 일부 구현예에서, 유전자 변형 시스템의 circRNA 구성요소를 활성화시키는 신호는, 예를 들어 그 시스템이 이들 세포에서 특이적으로 활성화되도록, 표적상 세포 또는 조직에서 더 높은 수준으로 존재한다.

[1059] 일부 구현예에서, 유전자 변형 시스템의 RNA 구성요소는 선형화에 의해 비활성화되는 circRNA로서 제공된다. 일부 구현예에서, 유전자 변형 폴리펩티드를 인코딩하는 circRNA는 절단 및 분해에 의해 비활성화된다. 일부 구현예에서, 유전자 변형 폴리펩티드를 인코딩하는 circRNA는 폴리펩티드의 코딩 서열로부터 번역 신호를 분리하는 절단에 의해 비활성화된다. 일부 구현예에서, 유전자 변형 시스템의 circRNA 구성요소를 비활성화시키는 신호는, 그 시스템이 이들 세포에서 특이적으로 비활성화되도록, 표적-외 세포 또는 조직에서 더 높은 수준으로 존재한다.

[1060] ***표적 핵산 부위***

[1061] 일부 구현예에서, 유전자 변형 후, 편집된 서열 주위의 표적 부위는, 예를 들어 문헌[Karst et al. (2020) *bioRxiv* doi.org/10.1101/645903(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, 예를 들어 표적 부위의 긴-판독 암플리콘 시퀀싱에 의해 결정 시, 예를 들어 편집 사건의 약 50% 또는 10% 미만으로 제한된 수의 삽입 또는 결실을 함유한다. 일부 구현예에서, 표적 부위는, 예를 들어 문헌[Karst et al. *bioRxiv* doi.org/10.1101/645903 (2020)(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, 예를 들어 표적 부위의 긴-판독 암플리콘 시퀀싱에 의해 결정 시, 다수 연속 편집 사건, 예를 들어 헤드-투-테일(head-to-tail) 또는 헤드-투-헤드 중복을 나타내지 않는다. 일부 구현예에서, 표적 부위는 주형 RNA에 상응하는 통합된 서열을 함유한다. 일부 구현예에서, 표적 부위는, 예를 들어 문헌[Karst et al. *bioRxiv* doi.org/10.1101/645903 (2020) (그 전체가 본 명세서에 참고로 포함됨).]에 기재된 바와 같이, 예를 들어 표적 부위의 긴-판독 암플리콘 시퀀싱에 의해 결정 시, 약 1% 또는 10% 초과의 사건에서 내인성 RNA로부터 생성되는 삽입물을 함유하지 않는다. 일부 구현예에서, 표적 부위는 주형 RNA에 상응하는 통합된 서열을 함유한다.

[1062] 본 발명의 특정 양태에서, 유전자 변형 시스템에 의해 통합된 숙주 DNA-결합 부위는 유전자 내, 인트론 내, 엑손 내, 임의의 유전자의 코딩 영역 외부인 ORF, 유전자의 조절 영역 내, 또는 유전자의 조절 영역 외부일 수 있다. 다른 양태에서, 폴리펩티드는 하나 이상의 숙주 DNA 서열에 결합할 수 있다.

[1063] 일부 구현예에서, 유전자 변형 시스템은 다수의 대립유전자에서 표적 유전자좌를 편집하는 데 사용된다. 일부 구현예에서, 유전자 변형 시스템은 특이적 대립유전자를 편집하도록 설계된다. 예를 들어, 유전자 변형 폴리펩티드는 하나의 대립유전자에만 존재하는 특이적 서열로 지시될 수 있으며, 예를 들어 표적 대립유전자, 예를 들어 gRNA 또는 어닐링 도메인에 대해서는 상동성을 갖지만, 제2의 동족 대립유전자에 대해서는 상동성이 없는 주형 RNA를 포함한다. 일부 구현예에서, 유전자 변형 시스템은 일배체형-특이적 대립유전자를 변경시킬 수 있다. 일부 구현예에서, 특이적 대립유전자를 표적화하는 유전자 변형 시스템은, 예를 들어 표적 대립유전자에 대해 적어도 2, 4, 6, 8, 또는 10배 선호를 갖는 대립 유전자를 우선적으로 표적화한다.

[1064] ***제2 가닥 닉 생성***

[1065] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템은 (예를 들어, 유전자 변형 폴리펩티드에서) 제1 가닥에 닉을 생성하는 닉카제 활성, 및 (예를 들어, 유전자 변형 폴리펩티드와 별개의 폴리펩티드에서) 표적 DNA의 제2 가닥에 닉을 생성하는 닉카제 활성을 포함한다. 본 명세서에 논의된 바와 같이, 이론에 구속되고자 함이 없이, 표적 부위 DNA의 제1 가닥의 닉 생성은 주형 RNA의 서열, 예를 들어 이종 목적 서열을 역전사하기 위해 RT 도메인에 의해 사용될 수 있는 3' OH를 제공하는 것으로 생각된다. 이론에 구속되고자 함이 없이, 제2 가닥에 추가의 닉을 도입하는 것은 세포 DNA 복구 기구가 원래 게놈 서열보다 이종성 목적 서열 기반 서열을 더 빈번하게 채용하도록 편향시킬 수 있는 것으로 생각된다. 일부 구현예에서, 제2 가닥에 대한 추가의 닉은 제1 가닥에 대한 닉과 동일한 엔도뉴클레아제 도메인(예를 들어, 닉카제 도메인)에 의해 만들어진다. 일부 구현예에서, 동일한 유전자 변형 폴리펩티드는 제1 가닥에 대한 닉과 제2 가닥에 대한 닉 둘 다를 수행한다. 일부 구현예에서, 유전자 변형 폴리펩티드는 CRISPR/Cas 도메인을 포함하고, 제2 가닥에 대한 추가의 닉은, 예를 들어 CRISPR/Cas 도메인이 제2 가닥에 닉을 생성하도록 지시하는 제2 gRNA를 포함하는, 추가의 핵산에 의해 지

시된다. 다른 구현예에서, 추가의 제2 가닥 닉은 제1 가닥에 대한 닉 외에 상이한 엔도뉴클레아제 도메인(예를 들어, 닉카제 도메인)에 의해 만들어진다. 일부 구현예에서, 상기 상이한 엔도뉴클레아제 도메인은 유전자 변형 폴리펩티드와 별도의, 추가의 폴리펩티드(예를 들어, 본 발명의 시스템은 추가의 폴리펩티드를 추가로 포함함)에 위치한다. 일부 구현예에서, 추가의 폴리펩티드는 본 명세서에 기재된 엔도뉴클레아제 도메인(예를 들어, 닉카제 도메인)을 포함한다. 일부 구현예에서, 추가의 폴리펩티드는, 예를 들어 본 명세서에 기재된 바와 같은, DNA 결합 도메인을 포함한다.

[1066] 본 명세서에서, 제2 가닥 닉이 일어나는 위치는, 제1 가닥 닉에 상대적으로, 원하는 유전자 변형 DNA 변형의 수득, 원하지 않는 이중 가닥 파손(DSB)의 발생, 원하지 않는 삽입의 발생, 또는 원하지 않는 결실의 발생 중 하나 이상에 영향을 미칠 수 있는 정도인 것으로 고려된다. 이론에 구속되고자 함이 없이, 제2 가닥 닉 생성은 두 가지 일반적 배향, 즉 내측 및 외측 닉으로 발생할 수 있다.

[1067] 일부 구현예에서, 내측 닉 배향에서, RT 도메인은 (예를 들어, 주형 RNA(예를 들어, 이종 목적 서열)를 사용하여) 제2 가닥 닉에서 멀어지며 중합한다. 일부 구현예에서, 내측 닉 배향에서, 제1 가닥에 대한 닉의 위치 및 제2 가닥에 대한 닉의 위치는(예를 들어, 두 닉 모두가 CRISPR/Cas 도메인을 포함하는 폴리펩티드(예를 들어, 유전자 변형 폴리펩티드)에 의해 만들어진다는 시나리오에서) 제1 PAM 부위 및 제2 PAM 부위 사이에 위치한다. 외부에 2개의 PAM이 있고 내부에 2개의 닉이 있는 경우, 이러한 내측 닉 배향은 또한 "PAM-아웃(PAM-out)"으로 지칭될 수 있다. 일부 구현예에서, 내측 닉 배향에서, 제1 가닥에 대한 닉의 위치 및 제2 가닥에 대한 닉의 위치는 폴리펩티드와 추가의 폴리펩티드가 표적 DNA에 결합하는 부위 사이에 있다. 일부 구현예에서, 내측 닉 배향에서, 제2 가닥에 대한 닉의 위치는 폴리펩티드와 추가의 폴리펩티드의 결합 부위 사이에 위치하고, 제1 가닥에 대한 닉은 또한 폴리펩티드와 추가의 폴리펩티드의 결합 부위 사이에 위치한다. 일부 구현예에서, 내측 닉 배향에서, 제1 가닥에 대한 닉의 위치 및 제2 가닥에 대한 닉의 위치는 PAM 부위와 표적 부위로부터 떨어진 제2 폴리펩티드의 결합 부위 사이에 위치한다.

[1068] 내측 닉 배향을 제공하는 유전자 변형 시스템의 예는 CRISPR/Cas 도메인을 포함하는 유전자 변형 폴리펩티드, 제1 가닥 상의 표적 부위 DNA의 닉 생성을 지시하는 gRNA를 포함하는 주형 RNA, 및 제1 닉의 위치로부터 멀리 떨어진 부위에서 닉 생성을 지시하는 추가의 gRNA를 포함하는 추가의 핵산을 포함하고, 여기서 제1 닉의 위치 및 제2 닉의 위치는 두 gRNA가 유전자 변형 폴리펩티드에 지시하는 부위 중의 PAM 부위 사이에 있다. 추가의 예로서, 내측 닉 배향을 제공하는 또 다른 유전자 변형 시스템은, 징크 핑거 분자 및 제1 닉카제 도메인을 포함하는 유전자 변형 폴리펩티드(여기서, 징크 핑거 분자는, 제1 닉카제 도메인이 표적 부위의 제1 가닥에 닉을 생성하도록 지시하는 방식으로 표적 DNA에 결합함); CRISPR/Cas 도메인을 포함하는 추가의 폴리펩티드, 및 그 추가의 폴리펩티드가 제2 가닥 상에서 표적 부위 DNA로부터 떨어진 부위에 닉을 생성하도록 지시하는 gRNA를 포함하는 추가의 핵산을 포함하며, 여기서 제1 닉의 위치 및 제2 닉의 위치는 PAM 부위와 징크 핑거 분자가 결합하는 부위 사이에 있다. 추가의 예로서, 내측 닉 배향을 제공하는 또 다른 유전자 변형 시스템은, 징크 핑거 분자 및 제1 닉카제 도메인을 포함하는 유전자 변형 폴리펩티드(여기서, 징크 핑거 분자는, 제1 닉카제 도메인이 표적 부위의 제1 가닥에 닉을 생성하도록 지시하는 방식으로 표적 DNA에 결합함); TAL 이펙터 분자 및 제2 닉카제 도메인을 포함하는 추가의 폴리펩티드를 포함하고, 여기서 TAL 이펙터 분자는 추가의 폴리펩티드가 제2 가닥에 닉을 생성하도록 지시하는 방식으로 표적 부위로부터 떨어진 부위에 결합하고, 여기서 제1 닉의 위치와 제2 닉의 위치는 TAL 이펙터 분자가 결합하는 부위와 징크 핑거 분자가 결합하는 부위 사이에 있다.

[1069] 일부 구현예에서, 외측 닉 배향에서, RT 도메인은 (예를 들어, 주형 RNA(예를 들어, 이종 목적 서열)를 사용하여) 제2 가닥 닉을 향해 중합한다. 일부 구현예에서, 제1 및 제2 닉 모두가 CRISPR/Cas 도메인을 포함하는 폴리펩티드(예를 들어, 유전자 변형 폴리펩티드)에 의해 만들어지는 경우, 제1 PAM 부위와 제2 PAM 부위는 제1 가닥에 대한 닉의 위치와 제2 가닥에 대한 닉의 위치 사이에 위치한다. 내부에 2개의 PAM이 있고 외부에 2개의 닉이 있는 경우, 이러한 외측 닉 배향은 또한 "PAM-인(PAM-in)"으로 지칭될 수 있다. 일부 구현예에서, 외측 닉 배향에서, 폴리펩티드(예를 들어, 유전자 변형 폴리펩티드)와 추가의 폴리펩티드는 제1 가닥에 대한 닉의 위치와 제2 가닥에 대한 닉의 위치 사이의 표적 DNA 상의 부위에 결합한다. 일부 구현예에서, 외측 닉 배향에서, 제2 가닥에 대한 닉의 위치는 제1 가닥에 대한 닉의 위치에 대하여 폴리펩티드 및 추가의 폴리펩티드의 결합 부위의 반대쪽에 위치한다. 일부 구현예에서, 외측 배향에서, PAM 부위와 표적 부위로부터 떨어진 제2 폴리펩티드의 결합 부위는 제1 가닥에 대한 닉의 위치와 제2 가닥에 대한 닉의 위치 사이에 위치한다.

[1070] 외측 닉 배향을 제공하는 유전자 변형 시스템의 예는 CRISPR/Cas 도메인을 포함하는 유전자 변형 폴리펩티드, 제1 가닥 상의 표적 부위 DNA의 닉 생성을 지시하는 gRNA를 포함하는 주형 RNA, 및 제1 닉의 위치로부터 멀리 떨어진 부위에서 닉 생성을 지시하는 추가의 gRNA를 포함하는 추가의 핵산을 포함하고, 여기서 제1 닉의 위치

및 제2 닉의 위치는 두 gRNA가 유전자 변형 폴리펩티드에 지시하는 부위 중의 PAM 부위의 외측에 있다(즉, PAM 부위는 제1 닉의 위치와 제2 닉의 위치 사이에 있음). 추가의 예로서, 외측 닉 배향을 제공하는 또 다른 유전자 변형 시스템은, 징크 핑거 분자 및 제1 닉카제 도메인을 포함하는 유전자 변형 폴리펩티드(여기서, 징크 핑거 분자는, 제1 닉카제 도메인이 표적 부위의 제1 가닥에 닉을 생성하도록 지시하는 방식으로 표적 DNA에 결합함); CRISPR/Cas 도메인을 포함하는 추가의 폴리펩티드, 및 그 추가의 폴리펩티드가 제2 가닥 상에서 표적 부위 DNA로부터 떨어진 부위를 닉 생성하도록 지시하는 gRNA를 포함하는 추가의 핵산을 포함하며, 여기서 제1 닉의 위치 및 제2 닉의 위치는 PAM 부위와 징크 핑거 분자가 결합하는 부위의 외측에 있다(즉, PAM 부위와 징크 핑거 분자가 결합하는 부위는 제1 닉의 위치와 제2 닉의 위치 사이에 있음). 추가의 예로서, 외측 닉 배향을 제공하는 또 다른 유전자 변형 시스템은, 징크 핑거 분자 및 제1 닉카제 도메인을 포함하는 유전자 변형 폴리펩티드(여기서, 징크 핑거 분자는, 제1 닉카제 도메인이 표적 부위의 제1 가닥에 닉을 생성하도록 지시하는 방식으로 표적 DNA에 결합함); TAL 이펙터 분자 및 제2 닉카제 도메인을 포함하는 추가의 폴리펩티드를 포함하고, 여기서 TAL 이펙터 분자는 추가의 폴리펩티드가 제2 가닥에 닉을 생성하도록 지시하는 방식으로 표적 부위로부터 떨어진 부위에 결합하고, 여기서 제1 닉의 위치와 제2 닉의 위치는 TAL 이펙터 분자가 결합하는 부위와 징크 핑거 분자가 결합하는 부위의 외측에 있다(즉, TAL 이펙터 분자가 결합하는 부위와 징크 핑거 분자가 결합하는 부위는 제1 닉의 위치와 제2 닉의 위치 사이에 있음).

[1071] 이론에 구속되고자 함이 없이, 제2 가닥 닉이 제공되는 유전자 변형 시스템의 경우, 일부 구현예에서 외측 닉 배향이 바람직한 것으로 생각된다. 본 명세서에 기재된 바와 같이, 내측 닉은 외측 닉 배향보다 더 많은 수의 이중 가닥 파손(DSB)를 생성할 수 있다. DSB는 세포의 핵에서 DSB 복구 경로에 의해 인식될 수 있으며, 이는 원하지 않는 삽입 및 결실을 생성한다. 외측 닉 배향은 DSB 형성의 감소된 위험, 및 상응하는 더 적은 양의 원하지 않는 삽입 및 결실을 제공할 수 있다. 일부 구현예에서, 원하지 않는 삽입 및 결실은 이종 목적 서열에 의해 인코딩되지 않는 삽입 및 결실, 예를 들어 이종 목적 서열에 의해 인코딩된 변형에 관계없는 이중 가닥 파손 복구 경로에 의해 생성되는 삽입 및 결실이다. 일부 구현예에서, 원하는 유전자 변형은 이종 목적 서열에 의해 인코딩되는(예를 들어, 표적 부위 내로 이종 목적 서열을 유전자 변형 작성함으로써 달성되는) 표적 DNA에 대한 변화(예를 들어, 치환, 삽입, 또는 결실)를 포함한다. 일부 구현예에서, 제1 가닥 닉 및 제2 가닥 닉은 외측 배향이다.

[1072] 추가적으로, 제1 가닥 닉과 제2 가닥 닉 사이의 거리는, 원하는 유전자 변형 시스템 DNA 변형의 수득, 원하지 않는 이중 가닥 파손(DSB)의 발생, 원하지 않는 삽입의 발생, 또는 원하지 않는 결실의 발생 중 하나 이상에 영향을 미칠 수 있는 정도이다. 이론에 구속되고자 함이 없이, 제2 가닥 닉 이득인, 이종 목적 서열의 표적 DNA로의 혼입을 향한 DNA 복구의 편향은, 제1 가닥과 제2 가닥 닉 사이의 거리가 감소됨에 따라 증가하는 것으로 생각된다. 그러나, 또한 DSB 형성의 위험은 제1 가닥 닉과 제2 가닥 닉 사이의 거리가 감소됨에 따라 증가하는 것으로 생각된다. 따라서, 원하지 않는 삽입 및/또는 결실은 제1 가닥 닉과 제2 가닥 닉 사이의 거리가 감소함에 따라 증가할 수 있는 것으로 생각된다. 일부 구현예에서, 제1 가닥 닉과 제2 가닥 닉 사이의 거리는 이종 목적 서열의 표적 DNA 내로의 혼입을 향한 DNA 복구의 편향의 이득과 DSB 형성 및 원하지 않는 결실 및/또는 삽입의 위험 간에 균형을 맞추도록 선택된다. 일부 구현예에서, 제1 가닥 닉과 제2 가닥 닉이 적어도 역치 거리로 떨어져 있는 시스템은, 제1 닉과 제2 닉이 역치 거리 보다 적게 떨어져 있는 다른 유사한 내측 닉 배향 시스템에 상대적으로, 원하는 유전자 변형 시스템 변형 결과의 증가된 수준, 원하지 않는 결실의 감소된 수준, 및/또는 원하지 않는 삽입의 감소된 수준을 갖는다. 일부 구현예에서, 역치 거리(들)가 하기 제공된다.

[1073] 일부 구현예에서, 제1 닉과 제2 닉은 적어도 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 또는 200개 뉴클레오티드 떨어져 있다. 일부 구현예에서, 제1 닉과 제2 닉은 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 또는 250개 이하 뉴클레오티드 떨어져 있다. 일부 구현예에서, 제1 닉과 제2 닉은 20 내지 200, 30 내지 200, 40 내지 200, 50 내지 200, 60 내지 200, 70 내지 200, 80 내지 200, 90 내지 200, 100 내지 200, 110 내지 200, 120 내지 200, 130 내지 200, 140 내지 200, 150 내지 200, 160 내지 200, 170 내지 200, 180 내지 200, 190 내지 200, 20 내지 190, 30 내지 190, 40 내지 190, 50 내지 190, 60 내지 190, 70 내지 190, 80 내지 190, 90 내지 190, 100 내지 190, 110 내지 190, 120 내지 190, 130 내지 190, 140 내지 190, 150 내지 190, 160 내지 190, 170 내지 190, 180 내지 190, 20 내지 180, 30 내지 180, 40 내지 180, 50 내지 180, 60 내지 180, 70 내지 180, 80 내지 180, 90 내지 180, 100 내지 180, 110 내지 180, 120 내지 180, 130 내지 180, 140 내지 180, 150 내지 180, 160 내지 180, 170 내지 180, 20 내지 170, 30 내지 170, 40 내지 170, 50 내지 170, 60 내지 170, 70 내지 170, 80 내지 170, 90 내지 170, 100 내지 170, 110 내지 170, 120 내지 170, 130 내지 170, 140 내지 170, 150 내지 170, 160 내지 170, 20 내지

160, 30 내지 160, 40 내지 160, 50 내지 160, 60 내지 160, 70 내지 160, 80 내지 160, 90 내지 160, 100 내지 160, 110 내지 160, 120 내지 160, 130 내지 160, 140 내지 160, 150 내지 160, 20 내지 150, 30 내지 150, 40 내지 150, 50 내지 150, 60 내지 150, 70 내지 150, 80 내지 150, 90 내지 150, 100 내지 150, 110 내지 150, 120 내지 150, 130 내지 150, 140 내지 150, 20 내지 140, 30 내지 140, 40 내지 140, 50 내지 140, 60 내지 140, 70 내지 140, 80 내지 140, 90 내지 140, 100 내지 140, 110 내지 140, 120 내지 140, 130 내지 140, 20 내지 130, 30 내지 130, 40 내지 130, 50 내지 130, 60 내지 130, 70 내지 130, 80 내지 130, 90 내지 130, 100 내지 130, 110 내지 130, 120 내지 130, 20 내지 120, 30 내지 120, 40 내지 120, 50 내지 120, 60 내지 120, 70 내지 120, 80 내지 120, 90 내지 120, 100 내지 120, 110 내지 120, 20 내지 110, 30 내지 110, 40 내지 110, 50 내지 110, 60 내지 110, 70 내지 110, 80 내지 110, 90 내지 110, 100 내지 110, 20 내지 100, 30 내지 100, 40 내지 100, 50 내지 100, 60 내지 100, 70 내지 100, 80 내지 100, 90 내지 100, 20 내지 90, 30 내지 90, 40 내지 90, 50 내지 90, 60 내지 90, 70 내지 90, 80 내지 90, 20 내지 80, 30 내지 80, 40 내지 80, 50 내지 80, 60 내지 80, 70 내지 80, 20 내지 70, 30 내지 70, 40 내지 70, 50 내지 70, 60 내지 70, 20 내지 60, 30 내지 60, 40 내지 60, 50 내지 60, 20 내지 50, 30 내지 50, 40 내지 50, 20 내지 40, 30 내지 40, 또는 20 내지 30개 뉴클레오티드 떨어져 있다. 일부 구현예에서, 제1 가닥 닉 및 제2 가닥 닉은 40 내지 100개 뉴클레오티드 떨어져 있다.

[1074] 이론에 구속되고자 함이 없이, 제2 가닥 닉이 제공되고, 내측 닉 배향이 선택된 유전자 변형 시스템의 경우, 제1 가닥 닉과 제2 가닥 닉 사이의 거리를 증가시키는 것이 바람직할 수 있는 것으로 생각된다. 본 명세서에 기재된 바와 같이, 내측 닉 배향은 외측 닉 배향보다 많은 수의 DSB를 생성할 수 있고, 외측 닉 배향보다 더 많은 양의 원하지 않는 삽입 및 결실을 생성할 수 있지만, 닉 사이의 거리를 증가시키는 것은, DSB, 원하지 않는 결실, 및/또는 원하지 않는 삽입의 증가를 완화할 수 있다. 일부 구현예에서, 제1 닉과 제2 닉이 적어도 역치 거리로 떨어져 있는 내측 닉 배향은, 제1 닉과 제2 닉이 역치 거리 보다 적게 떨어져 있는 다른 유사한 내측 닉 배향 시스템에 상대적으로, 원하는 유전자 변형 시스템 변형 결과의 증가된 수준, 원하지 않는 결실의 감소된 수준, 및/또는 원하지 않는 삽입의 감소된 수준을 갖는다. 일부 구현예에서 역치 거리가 하기 제공된다.

[1075] 일부 구현예에서, 제1 가닥 닉 및 제2 가닥 닉은 내측 배향이다. 일부 구현예에서, 제1 가닥 닉 및 제2 가닥 닉은 내측 배향이고 제1 가닥 닉 및 제2 가닥 닉은 적어도 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 220, 240, 260, 280, 300, 350, 400, 450, 또는 500개 뉴클레오티드, 적어도 100개 뉴클레오티드 떨어져 있다(및 선택적으로 500, 400, 300, 200, 190, 180, 170, 160, 150, 140, 130, 또는 120개 뉴클레오티드 떨어져 있음). 일부 구현예에서, 제1 가닥 닉 및 제2 가닥 닉은 내측 배향이고 제1 가닥 닉 및 제2 가닥 닉은 100 내지 200, 110 내지 200, 120 내지 200, 130 내지 200, 140 내지 200, 150 내지 200, 160 내지 200, 170 내지 200, 180 내지 200, 190 내지 200, 100 내지 190, 110 내지 190, 120 내지 190, 130 내지 190, 140 내지 190, 150 내지 190, 160 내지 190, 170 내지 190, 180 내지 190, 100 내지 180, 110 내지 180, 120 내지 180, 130 내지 180, 140 내지 180, 150 내지 180, 160 내지 180, 170 내지 180, 100 내지 170, 110 내지 170, 120 내지 170, 130 내지 170, 140 내지 170, 150 내지 170, 160 내지 170, 100 내지 160, 110 내지 160, 120 내지 160, 130 내지 160, 140 내지 160, 150 내지 160, 100 내지 150, 110 내지 150, 120 내지 150, 130 내지 150, 140 내지 150, 100 내지 140, 110 내지 140, 120 내지 140, 130 내지 140, 100 내지 130, 110 내지 130, 120 내지 130, 100 내지 120, 110 내지 120, 또는 100 내지 110개 뉴클레오티드 떨어져 있다.

[1076] ***화학적으로 변형된 핵산 및 핵산 말단 특징부***

[1077] 본 명세서에 기재된 핵산(예를 들어, 주형 핵산, 예를 들어 주형 RNA; 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산(예를 들어, mRNA); 또는 gRNA)은 비변형 또는 변형된 핵염기를 포함할 수 있다. 자연 발생 RNA는 4개의 염기성 리보뉴클레오티드, 즉 ATP, CTP, UTP 및 GTP로부터 합성되지만, 번역후 변형된 뉴클레오티드를 함유할 수 있다. 또한, 대략 100개의 다양한 뉴클레오시드 변형이 RNA에서 확인되었다(Rozenski, J, Crain, P, and McCloskey, J. (1999). The RNA Modification Database: 1999 update. Nucl Acids Res 27: 196-197). RNA는 또한 자연에서 발생하지 않는 완전히 합성인 뉴클레오티드를 포함할 수 있다.

[1078] 일부 구현예에서, 화학적 변형은 WO/2016/183482, 미국 특허 공개 20090286852, 국제 출원 WO/2012/019168, WO/2012/045075, WO/2012/135805, WO/2012/158736, WO/2013/039857, WO/2013/039861, WO/2013/052523, WO/2013/090648, WO/2013/096709, WO/2013/101690, WO/2013/106496, WO/2013/130161, WO/2013/151669, WO/2013/151736, WO/2013/151672, WO/2013/151664, WO/2013/151665, WO/2013/151668, WO/2013/151671, WO/2013/151667, WO/2013/151670, WO/2013/151666, WO/2013/151663, WO/2014/028429, WO/2014/081507, WO/2014/093924, WO/2014/093574, WO/2014/113089, WO/2014/144711, WO/2014/144767, WO/2014/144039,

WO/2014/152540, WO/2014/152030, WO/2014/152031, WO/2014/152027, WO/2014/152211, WO/2014/158795, WO/2014/159813, WO/2014/164253, WO/2015/006747, WO/2015/034928, WO/2015/034925, WO/2015/038892, WO/2015/048744, WO/2015/051214, WO/2015/051173, WO/2015/051169, WO/2015/058069, WO/2015/085318, WO/2015/089511, WO/2015/105926, WO/2015/164674, WO/2015/196130, WO/2015/196128, WO/2015/196118, WO/2016/011226, WO/2016/011222, WO/2016/011306, WO/2016/014846, WO/2016/022914, WO/2016/036902, WO/2016/077125, 또는 WO/2016/077123에 제공된 것이며, 이들 각각은 그 전체가 본 명세서에 참고로 포함된다. 화학적으로 변형된 뉴클레오티드의 폴리뉴클레오티드로의 혼입은, 뉴클레오티드에서 변형의 위치에 따라, 핵염기, 백본 또는 이들 둘 다 내로 포함되는 변형을 초래할 수 있음이 이해된다. 일부 구현예에서, 백본 변형은 EP 2813570에 제공된 것이며, 이는 그 전체가 본 명세서에 참고로 포함된다. 일부 구현예에서, 변형된 캡은 미국 특허 공개 20050287539에 제공된 것이며, 이는 그 전체가 본 명세서에 참고로 포함된다.

[1079] 일부 구현예에서, 화학적으로 변형된 핵산(예를 들어 RNA, 예를 들어 mRNA)은 ARCA: 항-역전 캡 유사체(m27.3'-OGP3G), GP3G(비메틸화 캡 유사체), m7GP3G(모노메틸화 캡 유사체), m32.2.7GP3G(트리메틸화 캡 유사체), m5CTP(5'-메틸-시티딘 트리포스페이트), m6ATP(N6-메틸-아데노신-5'-트리포스페이트), s2UTP(2-티오-우리딘 트리포스페이트), 및 Ψ(슈도우리딘 트리포스페이트) 중 하나 이상을 포함한다.

[1080] 일부 구현예에서, 화학적으로 변형된 핵산은 5' 캡, 예를 들어 7-메틸구아노신 캡(예를 들어, O-Me-m7G 캡); 초메틸화 캡 유사체; NAD+-유래 캡 유사체(예를 들어, 문헌[Kiledjian, Trends in Cell Biology 28, 454-464 (2018)]에 기재된 바와 같음); 또는 변형된, 예를 들어 비오티닐화된 캡 유사체(예를 들어, 문헌[Bednarek et al., Phil Trans R Soc B 373, 20180167 (2018)]에 기재된 바와 같음)를 포함한다.

[1081] 일부 구현예에서, 화학적으로 변형된 핵산은 다음 중 하나 이상으로부터 선택되는 3' 특징부를 포함한다: 폴리A 테일; 쌍을 형성하지 않은 5개 뉴클레오티드에 의해 측접된 16-뉴클레오티드의 긴 스템-루프 구조(예를 들어, 문헌[Mannironi et al., Nucleic Acid Research 17, 9113-9126 (1989)]에 기재된 바와 같음); 삼중-나선형 구조(예를 들어, 문헌[Brown et al., PNAS 109, 19202-19207 (2012)]에 기재된 바와 같음); tRNA, Y RNA, 또는 볼트 RNA 구조(예를 들어, 문헌[Labno et al., Biochemica et Biophysica Acta 1863, 3125-3147 (2016)]에 기재된 바와 같음); 하나 이상의 데옥시리보뉴클레오티드 트리포스페이트(dNTP), 2'O-메틸화 NTP, 또는 포스포로티오에이트-NTP의 혼입; 단일 뉴클레오티드 화학적 변형(예를 들어, 3' 말단 리보스의 반응성 알데히드로의 산화에 이어 알데히드-반응성 변형된 뉴클레오티드의 접합); 또는 또 다른 핵산 분자에 대한 화학적 결찰.

[1082] 일부 구현예에서, 핵산(예를 들어, 주형 핵산)은, 예를 들어 디하이드로우리딘, 이노신, 7-메틸구아노신, 5-메틸시티딘(5mC), 5' 포스페이트 리보티미딘, 2'-O-메틸 리보티미딘, 2'-O-에틸 리보티미딘, 2'-플루오로 리보티미딘, C-5 프로피닐-데옥시시티딘(pdC), C-5 프로피닐-데옥시우리딘(pdU), C-5 프로피닐-시티딘(pC), C-5 프로피닐-우리딘(pU), 5-메틸 시티딘, 5-메틸 우리딘, 5-메틸 데옥시시티딘, 5-메틸 데옥시우리딘 메톡시, 2,6-디아미노퓨린, 5'-디메톡시트리틸-N4-에틸-2'-데옥시시티딘, C-5 프로피닐-f-시티딘(pfC), C-5 프로피닐-f-우리딘(pfU), 5-메틸 f-시티딘, 5-메틸 f-우리딘, C-5 프로피닐-m-시티딘(pmC), C-5 프로피닐-f-우리딘(pmU), 5-메틸 m-시티딘, 5-메틸 m-우리딘, LNA(잠금 핵산), MGB(마이너 그루브 바인더) 슈도우리딘(Ψ), 1-N-메틸슈도우리딘(1-Me-Ψ), 또는 5-메톡시우리딘(5-MO-U)으로부터 선택되는, 하나 이상의 변형된 뉴클레오티드를 포함한다.

[1083] 일부 구현예에서, 핵산은 백본 변형, 예를 들어 백본에서 당 또는 포스페이트 기에 대한 변형을 포함한다. 일부 구현예에서, 핵산은 핵염기 변형을 포함한다.

[1084] 일부 구현예에서, 핵산은 표 13의 하나 이상의 화학적으로 변형된 뉴클레오티드, 표 14의 하나 이상의 화학적 백본 변형, 표 15의 하나 이상의 화학적으로 변형된 캡을 포함한다. 예를 들어 일부 구현예에서, 핵산은 둘 이상(예를 들어, 3, 4, 5, 6, 7, 8, 9, 또는 10개 또는 이상)의 다양한 유형의 화학적 변형을 포함한다. 예로서, 핵산은, 예를 들어 본 명세서에 기재된 바와 같은, 예를 들어 표 13의 둘 이상(예를 들어, 3, 4, 5, 6, 7, 8, 9, 또는 10개 또는 그 이상)의 다양한 유형의 변형된 핵염기를 포함할 수 있다. 대안적으로 또는 조합하여, 핵산은, 예를 들어 본 명세서에 기재된 바와 같은, 예를 들어 표 14의 둘 이상(예를 들어, 3, 4, 5, 6, 7, 8, 9, 또는 10개 또는 그 이상)의 다양한 유형의 백본 변형을 포함할 수 있다. 대안적으로 또는 조합하여, 핵산은, 예를 들어 본 명세서에 기재된 바와 같은, 예를 들어 표 15의 하나 이상의 변형된 캡을 포함할 수 있다. 예를 들어 일부 구현예에서, 핵산은 하나 이상의 유형의 변형된 핵염기 및 하나 이상의 유형의 백본 변형; 하나 이상의 유형의 변형된 핵염기 및 하나 이상의 변형된 캡; 하나 이상의 유형의 변형된 캡 및 하나 이상의 유형의 백본 변형; 또는 하나 이상의 유형의 변형된 핵염기, 하나 이상의 유형의 백본 변형, 및 하나 이상의 유형의 변형된 캡을 포함한다.

[1085] 일부 구현예에서, 핵산은 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000개, 또는 그 이상)의 변형된 핵염기를 포함한다. 일부 구현예에서, 핵산의 모든 핵염기는 변형된다. 일부 구현예에서, 핵산은 백본의 하나 이상(예를 들어, 2, 3, 4, 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 700, 800, 900, 1000개, 또는 그 이상)의 위치에서 변형된다. 일부 구현예에서, 핵산의 모든 백본 위치는 변형된다.

[1086] **표 13. 변형된 뉴클레오티드**

5-아자-우리딘	N2-메틸-6-티오-구아노신
2-티오-5-아자-미딘	N2,N2-디메틸-6-티오-구아노신
2-티오우리딘	피리딘-4-온 리보뉴클레오시드
4-티오-슈도우리딘	2-티오-5-아자-우리딘
2-티오-슈도우리딘	2-티오미딘
5-하이드록시우리딘	4-티오-슈도미딘
3-메틸우리딘	2-티오-슈도와이딘
5-카르복시메틸-우리딘	3-메틸미딘
1-카르복시메틸-슈도우리딘	1-프로피닐-슈도미딘
5-프로피닐-우리딘	1-메틸-1-데아자-슈도미딘
1-프로피닐-슈도우리딘	2-티오-1-메틸-1-데아자-슈도우리딘
5-타우리노메틸우리딘	4-메톡시-슈도미딘
1-타우리노메틸-슈도우리딘	5'-O-(1-티오포스페이트)-아데노신
5-타우리노메틸-2-티오-우리딘	5'-O-(1-티오포스페이트)-시티딘
1-타우리노메틸-4-티오-우리딘	5'-O-(1-티오포스페이트)-구아노신
5-메틸-우리딘	5'-O-(1-티오포스페이트)-우리딘
1-메틸-슈도우리딘	5'-O-(1-티오포스페이트)-슈도우리딘
4-티오-1-메틸-슈도우리딘	2'-O-메틸-아데노신
2-티오-1-메틸-슈도우리딘	2'-O-메틸-시티딘
1-메틸-1-데아자-슈도우리딘	2'-O-메틸-구아노신
2-티오-1-메틸-1-데아자-슈도미딘	2'-O-메틸-우리딘
디하이드로우리딘	2'-O-메틸-슈도우리딘
디하이드로슈도우리딘	2'-O-메틸-이노신
2-티오-디하이드로미딘	2-메틸아데노신
2-티오-디하이드로슈도우리딘	2-메틸티오-N6-메틸아데노신
2-메톡시우리딘	2-메틸티오-N6 이소펜테닐아데노신
2-메톡시-4-티오-우리딘	2-메틸티오-N6-(시스-하이드록시이소펜테닐)아데노신
4-메톡시-슈도우리딘	

[1087]

4-메톡시-2-티오-슈도우리딘	N6-메틸-N6-트레오닐카르바모일아데노신
5-아자-시티딘	N6-하이드록시노르발릴카르바모일아데노신
슈도이소시티딘	2-메틸티오-N6-하이드록시노르발릴
3-메틸-시티딘	카르바모일아데노신
N4-아세틸시티딘	2'-O-리보실아데노신(포스페이트)
5-포르밀시티딘	1,2'-O-디메틸이노신
N4-메틸시티딘	5,2'-O-디메틸시티딘
5-하이드록시메틸시티딘	N4-아세틸-2'-O-메틸시티딘
1-메틸-슈도이소시티딘	리시딘
피롤로-시티딘	7-메틸구아노신
피롤로-슈도이소시티딘	N2,2'-O-디메틸구아노신
2-티오-시티딘	N2,N2,2'-O-트리메틸구아노신
2-티오-5-메틸-시티딘	2'-O-리보실구아노신(포스페이트)
4-티오-슈도이소시티딘	와이부토신
4-티오-1-메틸-슈도이소시티딘	퍼옥시와이부토신
4-티오-1-메틸-1-데아자-슈도이소시티딘	하이드록시와이부토신
1-메틸-1-데아자-슈도이소시티딘	과소변형된 하이드록시와이부토신
제불라린	메틸와이오신
5-아자-제불라린	케오신
5-메틸-제불라린	에톡시케오신
5-아자-2-티오-제불라린	갈락토실-케오신
2-티오-제불라린	만노실-케오신
2-메톡시-시티딘	7-시아노-7-데아자구아노신
2-메톡시-5-메틸-시티딘	7-아미노메틸-7-데아자구아노신
4-메톡시-슈도이소시티딘	아르케오신(archaeosine)
4-메톡시-1-메틸-슈도이소시티딘	5,2'-O-디메틸우리딘
2-아미노퓨린	4-티오우리딘
2,6-디아미노퓨린	5-메틸-2-티오우리딘
7-데아자-아데닌	2-티오-2'-O-메틸우리딘
7-데아자-8-아자-아데닌	3-(3-아미노-3-카르복시프로필)우리딘
7-데아자-2-아미노퓨린	5-메톡시우리딘
7-데아자-8-아자-2-아미노퓨린	우리딘 5-옥시아세트산
7-데아자-2,6- 디아미노퓨린	우리딘 5-옥시아세트산 메틸 에스테르
7-데아자-8-아자-2,6-디아미노퓨린	5-(카르복시하이드록시메틸)우리딘)
1-메틸아데노신	5-(카르복시하이드록시메틸)우리딘 메틸 에스테르
N6-이소펜테닐아데노신	5-메톡시카르보닐메틸우리딘
N6-(시스-하이드록시이소펜테닐)아데노신	5-메톡시카르보닐메틸-2'-O-메틸우리딘
2-메틸티오-N6-(시스-하이드록시이소펜테닐)	5-메톡시카르보닐메틸-2-티오우리딘
아데노신	5-아미노메틸-2-티오우리딘
N6-글리시닐카르바모일아데노신	5-메틸아미노메틸우리딘
N6-트레오닐카르바모일아데노신	5-메틸아미노메틸-2-티오우리딘
2-메틸티오-N6-트레오닐 카르바모일아데노신	5-메틸아미노메틸-2-셀레노우리딘
N6,N6-디메틸아데노신	5-카르바모일메틸우리딘
7-메틸아데닌	5-카르바모일메틸-2'-O-메틸우리딘
2-메틸티오-아데닌	5-카르복시메틸아미노메틸우리딘
2-메톡시-아데닌	5-카르복시메틸아미노메틸-2'-O-메틸우리딘
이노신	5-카르복시메틸아미노메틸-2-티오우리딘
1-메틸-이노신	N4,2'-O-디메틸시티딘
와이오신	5-카르복시메틸우리딘
와이부토신	N6,2'-O-디메틸아데노신
7-데아자-구아노신	N,N6,O-2'-트리메틸아데노신

[1088]

7-데아자-8-아자-구아노신	N2,7-디메틸구아노신
6-티오-구아노신	N2,N2,7-트리메틸구아노신
6-티오-7-데아자-구아노신	3,2'-O-디메틸우리딘
6-티오-7-데아자-8-아자-구아노신	5-메틸디하이드로우리딘
7-메틸-구아노신	5-포르밀-2'-O-메틸시티딘
6-티오-7-메틸-구아노신	1,2'-O-디메틸구아노신
7-메틸이노신	4-데메틸와이오신
6-메톡시-구아노신	이소와이오신
1-메틸구아노신	N6-아세틸아데노신
N2-메틸구아노신	
N2,N2-디메틸구아노신	
8-옥소-구아노신	
7-메틸-8-옥소-구아노신	
1-메틸-6-티오-구아노신	

[1089]

[1090] **표 14. 백본 변형**

2'-O-메틸백본
펩티드 핵산(PNA) 백본
포스포로티오에이트 백본
모르폴리노 백본
카르바메이트 백본
실록산 백본
설파이드 백본
설폭사이드 백본
설폰 백본
포름아세틸 백본
티오포름아세틸 백본
메틸렌포름아세틸 백본
리보아세틸 백본
알켄 함유 백본
설파메이트 백본
설포네이트 백본
설폰아미드 백본
메틸렌이미노 백본
메틸렌하이드라지노 백본
아미드 백본

[1091]

[1092] **표 15. 변형된 캡**

m7GpppA
m7GpppC
m2,7GpppG
m2,2,7GpppG
m7Gpppm7G
m7,2'OmeGpppG
m72'dGpppG
m7,3'OmeGpppG
m7,3'dGpppG
GppppG
m7GppppG
m7GppppA
m7GppppC
m2,7GppppG
m2,2,7GppppG
m7Gppppm7G
m7,2'OmeGppppG
m72'dGppppG
m7,3'OmeGppppG
m7,3'dGppppG

[1093]

[1094] 유전자 변형 시스템의 주형을 포함하는 뉴클레오티드는 천연 또는 변형된 염기, 또는 이의 조합일 수 있다. 예를 들어, 주형은 슈도우리딘, 디하이드로우리딘, 이노신, 7-메틸구아노신, 또는 다른 변형된 염기를 함유할 수 있다. 일부 구현예에서, 주형은 잠금 핵산 뉴클레오티드를 함유할 수 있다. 일부 구현예에서, 주형에서 사용된 변형된 염기는 주형의 역전사를 억제하지 않는다. 일부 구현예에서, 주형에서 사용된 변형된 염기는 역전사, 예를 들어 특이성 또는 정확도를 개선시킬 수 있다.

[1095] 일부 구현예에서, 시스템의 RNA 구성요소(예를 들어, 주형 RNA 또는 gRNA)는 하나 이상의 뉴클레오티드 변형을 포함한다. 일부 구현예에서, gRNA의 변형 패턴은, 미변형된 또는 말단-변형된 가이드와 비교하여 생체 내 활성에 현저한 영향을 줄 수 있다(예를 들어, 문헌[Finn et al. *Cell Rep* 22(9):2227-2235 (2018)]의 도 1D에 나타낸 바와 같음; 그 전체가 본 명세서에 참고로 포함됨). 이론에 구속되고자 함이 없이, 이 과정은 적어도 부분적으로는 변형에 의해 부여된 RNA의 안정화로 인한 것일 수 있다. 그러한 변형의 비제한적인 예는 2'-O-메틸(2'-O-Me), 2'-O-(2-메톡시에틸)(2'-O-MOE), 2'- 플루오로(2'-F), 뉴클레오티드들 사이의 포스포로티오에이트(PS) 결합, G-C 치환, 및 뉴클레오티드들 사이의 역전된 무염기 연결 및 이들의 균등물을 포함할 수 있다.

[1096] 일부 구현예에서, 주형 RNA(예를 들어, 표적 부위에 결합하는 이의 부분에서) 또는 가이드 RNA는 5' 말단 영역을 포함한다. 일부 구현예에서, 주형 RNA 또는 가이드 RNA는 5' 말단 영역을 포함하지 않는다. 일부 구현예에서, 5' 말단 영역은, 예를 들어 문헌[Briner AE et al, Molecular Cell 56: 333-339 (2014)(그 전체가

본 명세서에 참고로 포함됨; 본 명세서에서, 예를 들어 모든 가이드 RNA에 대해 적용가능함)]에 sgRNA에 대하여 기재된 바와 같은, gRNA 스페이서 영역을 포함한다. 일부 구현예에서, 5' 말단 영역은 5' 말단 변형을 포함한다. 일부 구현예에서, 스페이서 영역을 갖는 또는 갖지 않는 5' 말단 영역은 crRNA, trRNA, sgRNA 및/또는 dgRNA와 회합될 수 있다. gRNA 스페이서 영역은, 일부 경우에서 가이드 영역, 가이드 도메인, 또는 표적화 도메인을 포함할 수 있다.

[1097] 일부 구현예에서, 주형 RNA(예를 들어, 표적 부위에 결합하는 이의 일부에서) 또는 본 명세서에 기재된 가이드 RNA는, 그 전체가 본 명세서에 참고로 포함된 WO2018107028A1의 표 4에 나타낸 임의의 서열을 포함한다. 일부 구현예에서, 서열이 가이드 및/또는 스페이서 영역을 나타내는 경우, 조성물은 이 영역을 포함하거나 포함하지 않을 수 있다. 일부 구현예에서, 가이드 RNA는, 예를 들어 서열번호에 의해 확인되는 바와 같이, WO2018107028A1의 표 4에 나타낸 임의의 서열의 하나 이상의 변형을 포함한다. 구현예에서, 뉴클레오티드는 동일하거나 상이할 수 있고/있거나, 나타낸 변형 패턴은 WO2018107028A1의 표 4에 나타낸 바와 같은 가이드 서열의 변형 패턴과 동일하거나 유사할 수 있다. 일부 구현예에서, 변형 패턴은 gRNA 또는 gRNA의 영역(예를 들어 5' 말단 영역, 하부 스템 영역, 벌지 영역, 상부 스템 영역, 연계 영역, 헤어핀 1 영역, 헤어핀 2 영역, 3' 말단 영역)의 변형의 상대적 위치 및 동일성을 포함한다. 일부 구현예에서, 변형 패턴은 WO2018107028A1의 표 4의 서열 열에 나타낸 서열들 중 어느 하나의 및/또는 서열의 하나 이상의 영역에 걸친 변형의 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100%를 함유한다. 일부 구현예에서, 변형 패턴은 WO2018107028A1의 표 4의 서열 열에 나타낸 서열들 중 어느 하나의 변형 패턴에 대하여 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일하다. 일부 구현예에서, 변형 패턴은 예를 들어 5' 말단 영역, 하부 스템 영역, 벌지 영역, 상부 스템 영역, 연계 영역, 헤어핀 1 영역, 헤어핀 2 영역, 및/또는 3' 말단 영역에서, WO2018107028A1의 표 4에 나타낸 서열의 하나 이상의 영역에 대해 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일하다. 일부 구현예에서, 변형 패턴은 5' 말단 영역에 걸쳐 서열의 변형 패턴에 대해 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99%, 또는 100% 동일하다. 일부 구현예에서, 변형 패턴은 하부 스템에 걸쳐 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일하다. 일부 구현예에서, 변형 패턴은 벌지에 걸쳐 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일하다. 일부 구현예에서, 변형 패턴은 상부 스템에 걸쳐 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일하다. 일부 구현예에서, 변형 패턴은 연계에 걸쳐 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일하다. 일부 구현예에서, 변형 패턴은 헤어핀 1에 걸쳐 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일하다. 일부 구현예에서, 변형 패턴은 헤어핀 2에 걸쳐 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일하다. 일부 구현예에서, 변형 패턴은 3' 말단에 걸쳐 적어도 50%, 55%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 동일하다. 일부 구현예에서, 변형 패턴은 WO2018107028A1의 표 4에 서열, 또는 그러한 서열의 영역(예를 들어, 5' 말단, 하부 스템, 벌지, 상부 스템, 연계, 헤어핀 1 영역, 헤어핀 2 영역, 3' 말단 영역)의 변형 패턴과, 예를 들어 0, 1, 2, 3, 4, 5, 6개, 또는 그 이상의 뉴클레오티드에서 상이하다. 일부 구현예에서, gRNA는 WO2018107028A1의 표 4의 서열의 변형과, 예를 들어 0, 1, 2, 3, 4, 5, 6개, 또는 그 이상의 뉴클레오티드에서 상이하다. 일부 구현예에서, gRNA는 WO2018107028A1의 표 4의 서열의 영역(예를 들어, 5' 말단 영역, 하부 스템, 벌지, 상부 스템, 연계, 헤어핀 1 영역, 헤어핀 2 영역, 3' 말단)의 변형과, 예를 들어 0, 1, 2, 3, 4, 5, 6개, 또는 그 이상의 뉴클레오티드에서 상이한 변형을 포함한다.

[1098] 일부 구현예에서, 주형 RNA(예를 들어, 표적 부위에 결합하는 이의 부분에서) 또는 gRNA는 2'-O-메틸(2'-O-Me) 변형된 뉴클레오티드를 포함한다. 일부 구현예에서, gRNA는 2'-O-(2-메톡시 에틸)(2'-O-moe) 변형된 뉴클레오티드를 포함한다. 일부 구현예에서, gRNA는 2'-플루오로(2'-F) 변형된 뉴클레오티드를 포함한다. 일부 구현예에서, gRNA는 뉴클레오티드들 사이에 포스포로티오에이트(PS) 결합을 포함한다. 일부 구현예에서, gRNA는 5' 말단 변형, 3' 말단 변형, 또는 5' 및 3' 말단 변형을 포함한다. 일부 구현예에서, 5' 말단 변형은 뉴클레오티드들 사이에 포스포로티오에이트(PS) 결합을 포함한다. 일부 구현예에서, 5' 말단 변형은 2'-O-메틸(2'-O-Me), 2'-O-(2-메톡시 에틸)(2'-O-MOE), 및/또는 2'-플루오로(2'-F) 변형된 뉴클레오티드를 포함한다. 일부 구현예에서, 5' 말단 변형은 적어도 하나의 포스포로티오에이트(PS) 결합, 및 2'-O-메틸(2'-O- Me), 2'-O-(2-메톡시 에틸)(2'-O-MOE), 및/또는 2'-플루오로(2'-F) 변형 뉴클레오티드 중 하나 이상을 포함한다. 말단 변형은 포스포로티오에이트(PS), 2'-O-메틸(2'-O-Me), 2'-O-(2- 메톡시에틸)(2'-O-MOE), 및/또는 2'-플루오로(2'-F) 변형을 포함할 수 있다. 균등한 말단 변형은 본 명세서에 기재된 구현예들에 의해서도 포함된다. 일부 구현예에서, 주형 RNA 또는 gRNA는 주형 RNA 또는 gRNA의 하나 이상의 영역의 변형과 조합된 말단 변형을 포함한다. RNA, 예를

들어 gRNA 및 이의 제형을 보호하기 위한 추가의 예시적인 변형 및 방법은 WO2018126176A1에 기재된 바와 같으며, 이는 그 전체가 본 명세서에 참고로 포함된다.

[1099] 일부 구현예에서, 구조-가이드된 계통적 접근법들이 예를 들어 문헌[Mir et al. *Nat Commun* 9:2641 (2018)(그 전체가 본 명세서에 참고로 포함됨)]에 기재된 바와 같이, 주형 RNA 또는 가이드 RNA에 변형(예를 들어, 2'-OMe-RNA, 2'-F-RNA, 및 PS 변형)을 도입하는 데 사용된다. 일부 구현예에서, 2'-F-RNA의 도입은, 예를 들어 C3'-엔도 당 퍼커링(puckering)을 최소로 간섭하면서, RNA:RNA 또는 RNA:DNA 듀플렉스(duplex)의 열 및 뉴클레아제 안정성을 증가시킨다. 일부 구현예에서, 2'-F는 2'-OH가 RNA:DNA 듀플렉스 안정성에 중요한 위치에서, 2'-OMe보다 더 양호하게 용인될 수 있다. 일부 구현예에서, crRNA는, 예를 들어 그 전체가 본 명세서에 참고로 포함된 문헌[Mir et al. *Nat Commun* 9:2641 (2018)]의 보충 표 1에 기재된 바와 같이, Cas9 활성을 감소시키지 않는 하나 이상의 변형, 예를 들어 C10, C20, 또는 C21(완전히 변형됨)을 포함한다. 일부 구현예에서, tracrRNA는 Cas9 활성을 감소시키지 않는 하나 이상의 변형, 예를 들어 문헌Mir et al. *Nat Commun* 9:2641 (2018)의 보충 표 1]의 T2, T6, T7, 또는 T8(완전히 변형됨)을 포함한다. 일부 구현예에서, crRNA는 (예를 들어, 본 명세서에 기재된 바와 같은) 하나 이상의 변형, 예를 들어 C20 및 T2를 포함하는 tracrRNA와 쌍을 형성할 수 있는 하나 이상의 변형을 포함한다. 일부 구현예에서, gRNA는 예를 들어 crRNA 및 tracrRNA의 키메라를 포함한다(예를 들어, 문헌[Jinek et al. *Science* 337(6096):816-821 (2012)]). 구현예에서, crRNA 및 tracrRNA로부터의 변형은, 예를 들어 향상된 안정성을 갖는 변형된 gRNA를 생산하도록 단일-가이드 키메라 상에서 맵핑된다.

[1100] 일부 구현예에서, gRNA 분자는 자연 발생 구조적 구성요소, 예를 들어 헤어핀의 부가 또는 공제에 의해 변형될 수 있다. 일부 구현예에서, gRNA는, 예를 들어 그 전체가 본 명세서에 참고로 포함된 WO2018106727에 기재된 바와 같은, 결실된 하나 이상의 3' 헤어핀 요소를 갖는 gRNA를 포함할 수 있다. 일부 구현예에서, gRNA는 부가된 헤어핀 구조, 예를 들어 스페이서 영역에서 부가된 헤어핀 구조를 함유할 수 있으며, 이는 문헌[Kocak et al. *Nat Biotechnol* 37(6):657-666 (2019)]의 교시에서 CRISPR-Cas 시스템의 특이성을 증가시키는 것으로 나타났다. 단축된 gRNA 및 생체 활성을 개선시키는 특이적 변형의 예를 포함하는 추가의 변형들은, 그 전체가 본 명세서에 참고로 포함된 US20190316121에서 확인할 수 있다.

[1101] 일부 구현예에서, 구조-가이드된 계통적 접근법(예를 들어, 문헌[Mir et al. *Nat Commun* 9:2641 (2018)]에 기재된 바와 같음; 그 전체가 본 명세서에 참고로 포함됨)이 주형 RNA에 대한 변형을 찾기 위해 이용된다. 구현예에서, 변형은 주형 RNA의 가이드 영역의 포함 또는 배제로 확인된다. 일부 구현예에서, 주형 RNA에 결합된 폴리펩티드의 구조는, 이후에 예를 들어 RNA와 폴리펩티드의 회합을 붕괴할 위험이 더 적은, 변형에 대해 선택될 수 있는 RNA의 비-단백질 접촉된 뉴클레오티드를 결정하는 데 사용된다. 주형 RNA에서 제2 구조는 또한, 예를 들어 변형, 예를 들어 헤어핀, 스템 및/또는 벌지를 선택하기 위한 2차 구조를 결정하기 위해, 소프트웨어 도구, 예를 들어 rrna.urmc.rochester.edu/RNAstructureWeb에서 입수가능한 RNAstructure 도구에 의해 인실리코로 예측될 수도 있다(Bellaousov et al. *Nucleic Acids Res* 41:W471-W474 (2013); 그 전체가 본 명세서에 참고로 포함됨).

[1102] ***조성물 및 시스템의 생산***

[1103] 당업자에 의해 이해되는 바와 같이, 핵산 작제물 및 단백질 또는 폴리펩티드를 설계 및 구축하는 방법(예컨대, 본 명세서에 기재된 시스템, 작제물 및 폴리펩티드)은 당업계에서 일상적이다. 일반적으로, 재조합 방법이 사용될 수 있다. 일반적으로 문헌[Smales & James (Eds.), *Therapeutic Proteins: Methods and Protocols* (Methods in Molecular Biology), Humana Press (2005); 및 Crommelin, Sindelar & Meibohm (Eds.), *Pharmaceutical Biotechnology: Fundamentals and Applications*, Springer (2013)]을 참조한다. 핵산 조성물의 설계, 제조, 평가, 정제 및 조작 방법은 문헌[Green and Sambrook (Eds.), *Molecular Cloning: A Laboratory Manual* (Fourth Edition), Cold Spring Harbor Laboratory Press (2012)]에 기재되어 있다.

[1104] 본 개시내용은 부분적으로는 핵산, 예를 들어 본 명세서에 기재된 유전자 변형 폴리펩티드를 인코딩하는, 벡터, 본 명세서에 기재된 주형 핵산, 또는 둘 다를 제공한다. 일부 구현예에서, 벡터는 선택적 마커, 예를 들어 항생제 내성 마커를 포함한다. 일부 구현예에서, 항생제 내성 마커는 카나마이신 내성 마커이다. 일부 구현예에서, 항생제 내성 마커는 베타-락탐 항생제에 대한 내성을 부여하지 않는다. 일부 구현예에서, 벡터는 암피실린 내성 마커를 포함하지 않는다. 일부 구현예에서, 벡터는 카나마이신 내성 마커를 포함하고, 암피실린 내성 마커는 포함하지 않는다. 일부 구현예에서, 유전자 변형 폴리펩티드를 인코딩하는 벡터는 (예를 들어, 표적 세포, 조직, 기관 또는 대상체에의 투여시) 표적 세포 게놈 내로 통합된다. 일부 구현예에서, 유전자 변형 폴리펩티드를 인코딩하는 벡터는 (예를 들어, 표적 세포, 조직, 기관 또는 대상체에의 투여시) 표적 세포 게놈 내로 통합되지

않는다. 일부 구현예에서, 주형 핵산(예를 들어, 주형 RNA)을 인코딩하는 벡터는 (예를 들어, 표적 세포, 조직, 기관 또는 대상체에의 투여시) 표적 세포 게놈 내로 통합되지 않는다. 일부 구현예에서, 벡터가 표적 세포 게놈에서 표적 부위 내로 통합되는 경우, 선택적 마커는 게놈 내로 통합되지 않는다. 일부 구현예에서, 벡터가 표적 세포 게놈에서 표적 부위 내로 통합되는 경우, 벡터 유지에 연관된 유전자 또는 서열(예를 들어, 플라스미드 유지 유전자)은 게놈 내로 통합되지 않는다. 일부 구현예에서, 벡터가 표적 세포 게놈에서 표적 부위 내로 통합되는 경우, 수송 조절 서열(예를 들어, AAV로부터의, 예를 들어 역 말단 반복부)은 게놈 내로 통합되지 않는다. 일부 구현예에서, 표적 세포, 조직, 기관, 또는 대상체에 대한 (예를 들어, 본 명세서에 기재된 유전자 변형 폴리펩티드, 본 명세서에 기재된 주형 핵산, 또는 둘 다를 인코딩하는) 벡터의 투여는 상기 표적 세포, 조직, 기관 또는 대상체의 게놈(들)에서 하나 이상의 표적 부위 내로 벡터의 부분의 통합을 초래한다. 일부 구현예에서, 통합 물질을 포함하는 표적 부위의 99, 95, 90, 80, 70, 60, 50, 40, 30, 20, 10, 5, 4, 3, 2, 또는 1% 미만(예를 들어, 표적 부위 없음)은 선택 마커(예를 들어, 항생체 내성 유전자), 수송 조절 서열(예를 들어, AAV 유래의, 예를 들어 역 말단 반복부), 또는 벡터 유래의 둘 다를 포함한다.

[1105] 재조합 단백질은 또한 곤충 세포, 효모, 박테리아 또는 기타 세포를 사용하여 적절한 프로모터 제어 하에 생산될 수도 있지만, 본 명세서에 기재된 치료적 약학 단백질 또는 폴리펩티드를 생산하기 위한 예시적인 방법은 포유동물 세포에서의 발현을 포함한다. 포유동물 발현 벡터는 복제 기점, 적합한 프로모터, 및 기타 5' 또는 3' 측접하는 비-전사 서열과 같은 비-전사 요소들, 및 필요한 리보솜 결합 부위, 폴리아데닐화 부위, 스플라이스 공여체 및 수용체 부위, 및 종료 서열과 같은 5' 또는 3' 비번역 서열을 포함할 수 있다. SV40 바이러스 게놈, 예를 들어 SV40 기원으로부터 유래된 DNA 서열, 초기 프로모터, 스플라이스 및 폴리아데닐화 부위는 이종 DNA 서열의 발현에 필요한 기타 유전자 요소를 제공하는 데 사용될 수 있다. 박테리아, 진균, 효모, 및 포유동물 세포 숙주와의 사용을 위한 적절한 클로닝과 발현 벡터가 문헌[Green & Sambrook, *Molecular Cloning: A Laboratory Manual* (Fourth Edition), Cold Spring Harbor Laboratory Press (2012)]에 기재되어 있다.

[1106] 다양한 포유동물 세포 배양 시스템이 재조합 단백질 발현 및 제조를 위해 사용될 수 있다. 포유동물 발현 시스템의 예는 CHO, COS, HEK293, HeLA, 및 BHK 세포주를 포함한다. 단백질 치료제의 생산을 위한 숙주 세포 배양의 과정은 문헌[Zhou and Kantardjieff (Eds.), *Mammalian Cell Cultures for Biologics Manufacturing (Advances in Biochemical Engineering/Biotechnology)*, Springer (2014)]에 기재되어 있다. 본 명세서에 기재된 조성물은 재조합 단백질을 인코딩하는, 벡터, 예컨대 바이러스 벡터, 예를 들어 렌티바이러스 벡터를 포함할 수 있다. 일부 구현예에서, 벡터, 예를 들어 바이러스 벡터는 재조합 단백질을 인코딩하는 핵산을 포함할 수 있다.

[1107] 단백질 치료제의 정제는 문헌[Franks, Protein Biotechnology: *Isolation, Characterization, and Stabilization*, Humana Press (2013); and in Cutler, *Protein Purification Protocols (Methods in Molecular Biology)*, Humana Press (2010)]에 기재되어 있다.

[1108] 본 개시내용은 또한 유전자 변형 폴리펩티드 및/또는 게놈 표적 부위에 대한 특이성을 갖는 주형 핵산 분자(예를 들어, 주형 RNA)의 생산을 위한 조성물 및 방법을 제공한다. 일 양태에서, 방법은 상류 상동성 절편, 이종 목적 서열 절편, 유전자 변형 폴리펩티드 결합 모티프, 및 gRNA 절편을 포함하여 RNA 절편의 생산을 포함한다.

[1109] ***치료 적용***

[1110] 본 명세서에 기재된 바와 같은 유전자 변형 시스템은 세포(예를 들어, 동물 세포, 식물 세포, 또는 진균 세포)를 변형시키는 데 사용될 수 있다. 본 명세서에 기재된 바와 같은 유전자 변형 시스템은 포유동물 세포(예를 들어, 인간 세포)를 변형시키는 데 사용될 수 있다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 시스템은 가축 동물(예를 들어, 소, 말, 양, 염소, 돼지 라마, 알파카, 낙타, 야크, 닭, 오리, 거위, 또는 타조) 유래의 세포를 변형시키는 데 사용될 수 있다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 유전자 변형 시스템은, 예를 들어 동물 세포, 예를 들어 포유동물 세포(예를 들어, 인간 세포), 식물 세포, 또는 진균 세포를 변형시키는 데, 실험실 도구 또는 연구 도구로 사용될 수 있거나, 실험실 방법 또는 연구 방법에서 사용될 수 있다.

[1111] 코딩 유전자를 RNA 서열 주형 내로 통합시킴으로써, 유전자 변형 시스템은, 예를 들어 기능 손실 돌연변이를 갖는 개인에서 치료용 이식유전자의 발현을 제공함으로써, 기능 수득 돌연변이를 정상의 이식유전자로 대체함으로써, 기능 수득 돌연변이 발현을 제거하기 위한 조절 서열을 제공함으로써, 및/또는 작동가능하게 연결된 유전자, 이식유전자 및 이의 시스템의 발현을 제어함으로써 치료적 필요를 다룰 수 있다. 특정 구현예에서, RNA 서열 주형은 숙주 세포의 치료적 필요에 특이적인 프로모터 영역, 예를 들어 조직 특이적 프로모터 또는 인핸서

를 인코딩한다. 또 다른 구현예에서, 프로모터는 코딩 서열에 작동가능하게 연결될 수 있다.

[1112] 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 표적 유전자의 발현(예를 들어, 전사 또는 번역)을 증가시키거나 감소시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 프로모터 또는 인핸서, 예를 들어 전사 인자에 결합하는 서열에서 서열을 변경, 추가, 또는 결실시킴으로써 표적 유전자의 발현(예를 들어, 전사 또는 번역)을 증가시키거나 감소시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 표적 유전자의 번역을 변경시키고(예를 들어, 아미노산 서열을 변경함), 시작 또는 정지 코돈을 삽입 또는 결실시키고, 유전자의 번역 프레임을 변경 또는 고정시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은, 예를 들어 스플라이스 어셉터 수여 또는 공여 부위를 삽입, 결실, 또는 변경시킴으로써, 표적 유전자의 스플라이싱을 변경시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 전사체 또는 단백질 반감기를 변경시킨다. 일부 구현예에서, 삽입, 결실, 치환, 또는 이의 조합은 표적 유전자, 예를 들어 표적 유전자에 의해 인코딩된 단백질의 활성을 변경, 증가, 감소시킨다.

[1113] *보상 편집*

[1114] 일부 구현예에서, 본 명세서에 제공된 시스템 또는 방법은 보상 편집을 도입하는 데 사용될 수 있다. 일부 구현예에서, 보상 편집은 질환 또는 장애와 연관된 유전자의 위치에 존재하며, 이는 질환-유발 돌연변이의 위치와 상이하다. 일부 구현예에서, 보상 돌연변이는 원인성 돌연변이를 함유하는 유전자 내에는 없다. 일부 구현예에서, 보상 편집은 질환-유발 돌연변이를 무효화할 수 있거나 이를 보상할 수 있다. 일부 구현예에서, 보상 편집은 질환-유발 돌연변이의 돌연변이체 효과를 억제 또는 역전시키기 위해 본 명세서에 제공된 시스템 또는 방법에 의해 도입될 수 있다.

[1115] *조절 편집*

[1116] 일부 구현예에서, 본 명세서에 제공된 시스템 또는 방법은 조절 편집을 도입하는 데 사용될 수 있다. 일부 구현예에서, 조절 편집은 유전자의 조절 서열, 예를 들어 유전자 프로모터, 유전자 인핸서, 유전자 억제자, 또는 유전자 스플라이싱을 조절하는 서열에 도입된다. 일부 구현예에서, 조절 편집은 표적 유전자의 발현 수준을 증가 또는 감소시킨다. 일부 구현예에서, 표적 유전자는 질환-유발 돌연변이를 함유하는 유전자와 동일하다. 일부 구현예에서, 표적 유전자는 질환-유발 돌연변이를 함유하는 유전자와 상이하다.

[1117] *반복부 확장 질환*

[1118] 일부 구현예에서, 본 명세서에 제공된 시스템 또는 방법은 반복부 확장 질환을 치료하는 데 사용될 수 있다. 일부 구현예에서, 예를 들어 유전자 변형 폴리펩티드를 포함하는, 본 명세서에 제공된 시스템 또는 방법은 맞춤화된 RNA 주형에 따라 유전자좌에서 반복부의 수를 재설정함으로써 반복부 확장 질환을 치료하는 데 사용될 수 있다.

[1119] ***투여 및 전달***

[1120] 본 명세서에 기재된 조성물 및 시스템은 시험과 내 또는 생체 내에서 사용될 수 있다. 일부 구현예에서, 시스템 또는 시스템의 구성요소는, 예를 들어 시험관 내 또는 생체 내에서, 세포(예를 들어, 포유동물 세포, 예를 들어 인간 세포)에 전달된다. 일부 구현예에서, 세포는 진핵 세포, 예를 들어 다세포 유기체, 예를 들어 동물, 예를 들어 포유동물(예를 들어, 인간, 돼지, 소), 조류(예를 들어, 가금류, 예컨대 닭, 칠면조, 또는 오리), 또는 어류의 세포이다. 일부 구현예에서, 세포는 비-인간 동물 세포(예를 들어, 실험실 동물, 가축 동물, 또는 반려 동물)이다. 일부 구현예에서, 세포는 줄기 세포(예를 들어, 조혈 줄기 세포), 섬유아세포, 또는 T 세포이다. 일부 구현예에서, 세포는 면역 세포, 예를 들어 T 세포(예를 들어, Treg, CD4, CD8, γδ 또는 기억 T 세포), B 세포(예를 들어, 기억 B 세포 또는 형질 세포), 또는 NK 세포이다. 일부 구현예에서, 세포는 비-분열 세포, 예를 들어 비-분열 섬유아세포 또는 비-분열 T 세포이다. 일부 구현예에서, 세포는 HSC이고 p53은 상향조절되지 않거나, 예를 들어 PCT/US2019/048607의 실시예 30에 기재된 방법에 따라 결정 시 10%, 5%, 2%, 또는 1% 미만으로 상향조절된다. 당업자는 유전자 변형 시스템의 구성요소가 폴리펩티드, 핵산(예를 들어, DNA, RNA), 및 이들의 조합의 형태로 전달될 수 있음을 이해할 것이다.

[1121] 일 구현예에서 시스템 및/또는 시스템의 구성요소는 핵산으로 전달된다. 예를 들어, 유전자 변형 폴리펩티드는 폴리펩티드를 인코딩하는 DNA 또는 RNA의 형태로 전달될 수 있고, 주형 RNA는 RNA 또는 RNA로 전사될 그의 상보적인 DNA의 형태로 전달될 수 있다. 일부 구현예에서, 시스템 또는 시스템의 구성요소는 1, 2, 3, 4개, 또는 그 이상의 별개의 핵산 분자로 전달된다. 일부 구현예에서 시스템 또는 시스템의 구성요소는 DNA와 RNA의 조합으로서 전달된다. 일부 구현예에서 시스템 또는 시스템의 구성요소는 DNA와 단백질의 조합으로서 전달된다. 일부 구

현예에서 시스템 또는 시스템의 구성요소는 RNA와 단백질의 조합으로서 전달된다. 일부 구현예에서 유전자 변형 폴리펩티드는 단백질로서 전달된다.

[1122] 일부 구현예에서 시스템 또는 시스템의 구성요소는 벡터를 사용하여 세포, 예를 들어 포유동물 세포 또는 인간 세포로 전달된다. 벡터는, 예를 들어 플라스미드 또는 바이러스일 수 있다. 일부 구현예에서, 전달은 생체 내, 시험관 내, 생체 외 또는 제 위치에서 이루어진다. 일부 구현예에서 바이러스는 아데노 관련 바이러스(AAV), 렌티바이러스, 또는 아데노바이러스이다. 일부 구현예에서, 시스템 또는 시스템의 구성성분은 바이러스-유사 입자 또는 바이로솜을 이용하여 세포에 전달된다. 일부 구현예에서 전달은 하나 초과의 바이러스, 바이러스-유사 입자 또는 바이로솜을 사용한다.

[1123] 일 구현예에서, 본 명세서에 기재된 조성물 및 시스템은 리포좀 또는 다른 유사한 소포 내에 제형화될 수 있다. 리포좀은 내부 수성 구획을 둘러싸고 있는 단층 또는 다층 지질 이중층 및 상대적으로 불침투성인 외부 친유성 인지질 이중층으로 구성된 구형의 소포 구조이다. 리포좀은 음이온성, 중성 또는 양이온성일 수 있다. 리포좀은 생체적합성이고, 비독성이며, 친수성 및 친유성 약물 분자 둘 모두를 전달하고, 혈장 효소에 의한 분해로부터 그들의 카고를 보호하고, 생물학적 막 및 혈액-뇌 장벽(BBB)을 가로질러 그들의 로드를 수송할 수 있다(예를 들어, 검토를 위해 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조).

[1124] 소포는 몇몇 상이한 유형의 지질로부터 제조될 수 있지만; 약물 담체로서 리포좀을 생성하는 데 인지질이 가장 흔히 사용된다. 다층 소포 지질을 제조하기 위한 방법이 당업계에 알려져 있다(예를 들어, 미국 특허 6,693,086을 참조하며, 다층 소포 지질 제제에 관한 이의 교시내용은 본 명세서에 참고로 포함됨). 지질 필름이 수용액과 혼합될 때 소포 형성이 자발적일 수 있지만, 그것은 또한 균질기, 초음파처리기, 또는 압출 장치를 사용함으로써 전단 형태로 힘을 인가함으로써 촉진될 수 있다(예를 들어, 검토를 위해 문헌[Spuch and Navarro, Journal of Drug Delivery, vol. 2011, Article ID 469679, 12 pages, 2011. doi:10.1155/2011/469679] 참조). 압출된 지질은 압출된 지질 제제에 관한 교시가 본 명세서에 참고로 포함되는 문헌[Templeton et al., Nature Biotech, 15:647-652, 1997]에 기재된 바와 같이 감소하는 크기의 필터를 통해 압출시킴으로써 제조될 수 있다.

[1125] 다양한 나노입자, 예컨대 리포좀, 지질 나노입자, 양이온성 지질 나노입자, 이온화성 지질 나노입자, 중합체성 나노입자, 금 나노입자, 덴드리머(dendrimer), 사이클로덱스트린 나노입자, 미셀 또는 상기의 조합이 전달을 위해 사용될 수 있다.

[1126] 지질 나노입자는 본 명세서에 기재된 약제학적 조성물을 위한 생체적합성 및 생분해성 전달 시스템을 제공하는 담체의 예이다. 나노구조화된 지질 담체(NLC)는 고체 지질 나노입자(SLN)의 특성을 보유하고, 약물 안정성 및 로딩 능력을 개선시키고, 약물 누출을 방지하는 변형된 고체 지질 나노입자(SLN)이다. 중합체 나노입자(PNP)는 약물 전달의 중요한 구성성분이다. 이들 나노입자는 특정 표적에 대한 약물 전달을 효과적으로 유도하고, 약물 안정성 및 약물의 제어 방출을 개선시킬 수 있다. 리포좀과 중합체를 조합한 담체의 유형인 지질-중합체 나노입자(PLN)가 또한 사용될 수 있다. 이들 나노입자는 PNP 및 리포좀의 상보적 이점을 가진다. PLN은 코어-셸 구조로 구성되며; 중합체 코어는 안정한 구조를 제공하고, 인지질 셸은 양호한 생체 적합성을 제공한다. 이와 같이, 2가지 구성성분은 약물 캡슐화 효율을 증가시키고, 표면 변형을 용이하게 하고, 수용성 약물의 누출을 방지한다. 검토를 위하여, 예를 들어, 문헌[Li et al. 2017, Nanomaterials 7, 122; doi:10.3390/nano7060122]을 참조한다.

[1127] 엑소좀도 또한, 본 명세서에 기재된 조성물 및 시스템을 위한 약물 운반 비히클로서 사용될 수 있다. 검토를 위해, 문헌[Ha et al. July 2016. Acta Pharmaceutica Sinica B. Volume 6, Issue 4, Pages 287-296; https://doi.org/10.1016/j.apsb.2016.02.001]을 참고한다.

[1128] 푸소좀은 표적 세포와 반응 및 융합하고, 따라서 다양한 분자에 대한 전달 비히클로 사용될 수 있다. 이들은 일반적으로 내강 또는 동공을 둘러싼 양친매성 지질의 이중층 및 양친매성 지질 이중층과 상호작용하는 푸소젠(fusogen)으로 이루어진다. 푸소젠 구성요소는 융합 및 페이로드 전달에 대한 표적 세포 특이성을 부여하여, 프로그램가능한 세포 특이성을 갖는 전달 비히클의 형성을 가능하게 하기 위해 조작가능한 것으로 나타났다(예를 들어, 특허출원 WO2020014209 참조, 푸소좀 설계, 제조, 및 용법에 관한 이의 교시 내용은 본 명세서에 참고로 포함됨).

[1129] 일부 구현예에서, 유전자 변형 시스템의 단백질 구성요소(들)는 주형 핵산(예를 들어, 주형 RNA)과 미리 회합될 수 있다. 예를 들어, 일부 구현예에서, 유전자 변형 폴리펩티드는 먼저 주형 핵산(예를 들어, 주형 RNA)과 조합

되어, 리보핵단백질(RNP) 복합체를 형성할 수 있다. 일부 구현예에서, RNP는, 예를 들어 형질감염, 뉴클레오펙션, 바이러스, 소포, LNP, 엑소좀, 푸소좀을 통해 세포로 전달될 수 있다.

[1130] 유전자 변형 시스템은 세포, 조직 및 다세포 유기체로 도입될 수 있다. 일부 구현예에서, 시스템 또는 시스템의 구성요소는 기계적 수단 또는 물리적 수단을 통해 세포로 전달된다.

[1131] 단백질 치료제의 제형은 문헌[Meyer (Ed.), *Therapeutic Protein Drug Products: Practical Approaches to formulation in the Laboratory, Manufacturing, and the Clinic*, Woodhead Publishing Series (2012)]에 기재되어 있다.

[1132] *조직 특이적 활성/투여*

[1133] 일부 구현예에서, 본 명세서에 기재된 시스템은 하나 이상의 특징부(예를 들어, 프로모터 또는 마이크로RNA 결합 부위)를 사용하여 표적외 세포 또는 조직에서 활성을 제한할 수 있다.

[1134] 일부 구현예에서, 본 명세서에 기재된 핵산(예를 들어, 주형 RNA 또는 주형 RNA를 인코딩하는 DNA)은 프로모터 서열, 예를 들어 조직 특이적 프로모터 서열을 포함한다. 일부 구현예에서, 조직-특이적 프로모터는 유전자 변형 시스템의 표적-세포 특이성을 증가시키는 데 사용된다. 예를 들어, 프로모터는 표적 세포 유형에서 활성이지만 비-표적 세포 유형에서 활성이 아닌(또는 더 낮은 수준으로 활성인) 것에 기초하여 선택될 수 있다. 따라서, 프로모터가 비-표적 세포의 게놈 내로 통합되더라도, 통합된 유전자의 발현을 구동하지 않을 것이다(또는 단지 낮은 수준의 발현만 구동할 것임). 주형 RNA에 조직-특이적 프로모터 서열을 갖는 시스템은 또한, 예를 들어 본 명세서에 기재된 바와 같은, 예를 들어 주형 RNA 또는 유전자 변형 단백질을 인코딩하는 핵산에서, 마이크로RNA 결합 부위와 조합하여 사용될 수 있다. 주형 RNA에 조질-특이적 프로모터 서열을 갖는 시스템은 또한 조직-특이적 프로모터에 의해 구동되는, 유전자 변형 폴리펩티드를 인코딩하는 DNA와 조합하여 사용되어, 예를 들어 비-표적 세포보다 표적 세포에서 유전자 변형 단백질의 더 높은 수준을 달성할 수 있다. 일부 구현예에서, 예를 들어 간 적응증의 경우, 조직-특이적 프로모터는 본 명세서에 참고로 포함된 WO2020014209의표 3으로부터 선택된다.

[1135] 일부 구현예에서, 본 명세서에 기재된 핵산(예를 들어, 주형 RNA 또는 주형 RNA를 인코딩하는 DNA)은 마이크로RNA 결합 부위를 포함한다. 일부 구현예에서, 마이크로RNA 결합 부위는 유전자 변형 시스템의 표적-세포 특이성을 증가시키는 데 사용된다. 예를 들어, 마이크로RNA 결합 부위는 비-표적 세포 유형에 존재하지만, 표적 세포 유형에는 존재하지 않는(또는 비-표적 세포에 비해 감소된 수준으로 존재하는) miRNA에 의해 인식된다는 것에 기초하여 선택될 수 있다. 따라서, 주형 RNA가 비-표적 세포에 존재하는 경우, miRNA에 의해 결합될 수 있고, 주형 RNA가 표적 세포에 존재하는 경우, 이는 miRNA에 의해 결합되지 않을 수 있다(또는 비-표적 세포에 대해 감소된 수준으로 결합한다). 이론에 구속되고자 하지는 않지만, miRNA의 주형 RNA에 대한 결합은 이의 활성을 간섭할 수 있고, 예를 들어 이종성 목적 서열의 게놈 내로의 삽입을 간섭할 수 있다. 따라서, 시스템은 비-표적 세포의 게놈을 편집하는 것보다 더 효율적으로 표적 세포의 게놈을 편집할 수 있을 것이며, 예를 들어 이종 목적 서열은 비-표적 세포의 게놈 내 보다 더 효율적으로 표적 세포의 게놈 내로의 삽입될 수 있거나, 삽입 또는 결실은 비-표적 세포보다 표적 세포에서 더 효율적으로 생산된다. 주형 RNA(또는 이를 인코딩하는 DNA)에서 마이크로RNA 결합 부위를 갖는 시스템은 유전자 변형 폴리펩티드를 인코딩하는 핵산과 조합되어 사용될 수도 있으며, 여기서 유전자 변형 폴리펩티드의 발현은, 예를 들어 본 명세서에 기재된 바와 같은 제2의 마이크로RNA 결합 부위에 의해 제어된다. 일부 구현예에서, 예를 들어 간 적응증의 경우, miRNA는 본 명세서에 참고로 포함된 WO2020014209의 표 4로부터 선택된다.

[1136] 일부 구현예에서, 주형 RNA는 마이크로RNA 서열, siRNA 서열, 가이드 RNA 서열, 또는 피위(piwi) RNA 서열을 포함한다.

[1137] *프로모터*

[1138] 일부 구현예에서, 하나 이상의 프로모터 또는 인핸서 요소는, 이종 목적 서열의 발현을 제어하는 유전자 변형 단백질을 인코딩하는 핵산 또는 주형 핵산에 작동가능하게 연결된다. 특정 구현예에서, 하나 이상의 프로모터 또는 인핸서 요소는 세포 유형 또는 조직 특이적 요소를 포함한다. 일부 구현예에서, 프로모터 또는 인핸서는 이종 목적 서열의 발현을 자연적으로 제어하는 프로모터 또는 인핸서와 동일하거나 이로부터 유래된다. 예를 들어, 오르니틴 트랜스카르보밀라제(transcarbomylase) 프로모터 및 인핸서는, 오르니틴 트랜스카르보밀라제 결핍을 교정하기 위해 본 발명에 의해 제공된 시스템 또는 방법에서 오르니틴 트랜스르카보밀라제 유전자의 발현을 제어하는 데 사용될 수 있다. 일부 구현예에서, 프로모터는 표 16 또는 17 또는 이의 기능성 단편 또는 변이체

의 프로모터이다.

[1139] 상업적으로 구매가능한 예시적인 조직 특이적 프로모터는, 예를 들어 인터넷 주소(예를 들어, www.invivogen.com/tissue-specific-promoters)에서 찾을 수 있다. 일부 구현예에서, 프로모터는 천연 프로모터 또는 최소 프로모터로, 예를 들어 이는 소정의 유전자의 5' 영역으로부터 단일 단편으로 이루어진다. 일부 구현예에서, 천연 프로모터는 코어 프로모터 및 그의 천연 5' UTR을 포함한다. 일부 구현예에서, 5' UTR은 인트론을 포함한다. 다른 구현예에서, 이들은 복합 프로모터를 포함하며, 이는 상이한 기원의 프로모터 요소들을 조합하거나 원위 인핸서를 동일한 기원의 최소 프로모터와 조립함으로써 생성되었다.

[1140] 예시적인 세포 또는 조직 특이적 프로모터가 하기 표에 제공되며, 이를 인코딩하는 예시적인 핵산 서열은 당업계에 알려져 있고, 다양한 공급원, 예컨대 RefSeq를 포함하여 NCBI 데이터베이스, 및 진핵성 프로모터(//epd.epfl.ch//index.php) 데이터베이스를 이용하여 쉽게 접근될 수 있다.

[1141] *표 16. 예시적인 세포 또는 조직-특이적 프로모터*

프로모터	**표적 세포**
B29 프로모터	B 세포
CD14 프로모터	단구 세포
CD43 프로모터	백혈구 및 혈소판
CD45 프로모터	조혈세포
CD68 프로모터	대식세포
데스민(Desmin) 프로모터	근육 세포
엘라스타제-1 프로모터	췌장 선방 세포
엔도글린(Endoglin) 프로모터	내피세포
피브로넥틴 프로모터	분화 세포, 치유 조직
Flt-1 프로모터	내피세포
GFAP 프로모터	성상세포
GPIIB 프로모터	거핵세포
ICAM-2 프로모터	내피세포
INF-Beta 프로모터	조혈세포
Mb 프로모터	근육 세포
Nphs1 프로모터	족세포
OG-2 프로모터	조골세포, 치아모세포(Odonblast)
SP-B 프로모터	폐
Syn1 프로모터	뉴런
WASP 프로모터	조혈세포
SV40/bAlb 프로모터	간
SV40/bAlb 프로모터	간
SV40/Cd3 프로모터	백혈구 및 혈소판
SV40/CD45 프로모터	조혈세포
NSE/RU5' 프로모터	성숙 뉴런

[1142]

[1143] *표 17. 추가의 예시적인 세포 또는 조직-특이적 프로모터*

[1144]

프로모터	유전자 설명	유전자 특이성
APOA2	아포지질단백질 A-II	간세포(간 전구세포 유래)
SERPINA1 (hAAT)	세르핀 펩티다제 억제제, 클레이드 A (알파-1 항프로테나제, 항트립신), 구성원 1 (알파 1 항-트립신으로도 명명됨)	간세포 (최종적인 내배엽 단계로부터)
CYP3A	시토크롬 P450, 패밀리 3, 서브패밀리 A, 폴리펩티드	성숙한 간세포
MIR122	마이크로RNA 122	간세포 (초기 배아 간 세포 유래) 및 내배엽
췌장 특이적 프로모터		
프로모터	**유전자 설명**	**유전자 특이성**
INS	인슐린	췌장 베타 세포 (최종 내배엽 단계로부터)
IRS2	인슐린 수용체 기질 2	췌장 베타 세포
Pdx1	췌장 및 십이지장 호메오박스 1	췌장 (최종 내배엽 단계로부터)
Alx3	아리스탈레스-유사(Aristaless-like) 호메오박스 3	췌장 베타 세포 (최종 내배엽 단계로부터)
Ppy	췌장 폴리펩티드	PP 췌장 세포 (감마 세포)
심장 특이적 프로모터		
프로모터	**유전자 설명**	**유전자 특이성**

[1145]

Myh6(aMHC)	미오신, 중쇄 6, 심근, 알파	심장 근육세포의 후기 분화 마커(심방 특이성)
MYL2 (MLC-2v)	미오신, 경쇄 2, 조절, 심장, 느림	심장 근육세포의 후기 분화 마커(심실 특이성)
ITNNl3 (cTnl)	트로포닌 I 형 3(심장)	심근세포 (미성숙 상태로부터)
ITNNl3 (cTnl)	트로포닌 I 형 3(심장)	심근세포 (미성숙 상태로부터)
NPPA(ANF)	나트륨뇨 펩티드 전구체 A(심방 나트륨뇨 인자로도 명명됨)	성인 세포에서 심방 특이성
Slc8a1(Ncx1)	용질 담체 패밀리 8(나트륨/칼슘 교환기), 구성원 1	조기 발생 단계로부터의 심근세포
CNS 특이적 프로모터		
프로모터	**유전자 설명**	**유전자 특이성**
SYN1(hSyn)	시냅신 I	뉴런
GFAP	아교세포섬유 산성 단백질	성상세포
INA	인터넥신 뉴런성 중간체 필라멘트 단백질, 알파(a-인터넥신)	신경전구세포(neuroprogenitor)
NES	네스틴	신경전구세포 및 외배엽
MOBP	미엘린-연관 희소돌기아교세포 염기성 단백질	희소돌기아교세포
MBP	미엘린 염기성 단백질	희소돌기아교세포
TH	티로신 하이드록실라제	도파민작용성 뉴런
FOXA2 (HNF3 베타)	포크헤드(Forkhead) 박스 A2	도파민작용성 뉴런(내배엽의 마커로서도 사용됨)
피부 특이적 프로모터		
프로모터	**유전자 설명**	**유전자 특이성**
FLG	필라그린	과립층 유래의 각질형성세포
K14	케라틴 14	과립 및 기저 층으로부터의 각질형성세포

[1146]

TGM3	트랜스글루타미나제 3	과립층 유래의 각질형성세포
면역 세포 특이적 프로모터		
프로모터	**유전자 설명**	**유전자 특이성**
ITGAM (CD11B)	인테그린, 알파 M(보체 구성요소 3 수용체 3 서브유닛)	단핵구, 대식세포, 과립구, 자연 살해 세포
요생식기 세포 특이적 프로모터		
프로모터	**유전자 설명**	**유전자 특이성**
Pbsn	프로바신(Probasin)	전립선 상피
Upk2	유로플라킨 2	방광
Sbp	스페르민 결합 단백질	전립선
Fer1l4	Fer-1-유사 4	방광
내피세포 특이적 프로모터		
프로모터	**유전자 설명**	**유전자 특이성**
ENG	엔도글린	내피세포
다능성 및 배아 세포 특이적 프로모터		
프로모터	**유전자 설명**	**유전자 특이성**
Oct4 (POU5F1)	POU 클래스 5 호메오박스 1	다능성 세포 (생식 세포, ES 세포, iPS 세포)
NANOG	나노그(Nanog) 호메오박스	다능성 세포 (ES 세포, iPS 세포)
합성 Oct4	Oct-4 코어 인핸서 요소에 기반한 합성 프로모터	다능성 세포(ES 세포, iPS 세포)
T 단미증	단미증	중배엽
NES	네스틴	신경전구세포 및 외배엽
SOX17	SRY(성 결정 영역 Y)-박스 17	내배엽

[1147]

FOXA2 (HNFJ 베타)	포크헤드 박스 A2	내배엽(도파민작용성 뉴런의 마커로서도 사용됨)
MIR122	마이크로RNA 122	내배엽 및 간세포 (초기단계 배아 간 세포로부터~

[1148]

[1149] 이용되는 숙주/벡터 시스템에 따라, 구성 및 유도성 프로모터, 전사 인핸서 요소, 전사 종결자 등을 포함하는 많은 임의의 적합한 전사 및 번역 제어 요소가 발현 벡터 상에서 사용될 수 있다(예를 들어, 문헌[Bitter et al. (1987) *Methods in Enzymology,* 153:516-544] 참조; 그 전체가 참조로서 여기 포함됨).

[1150] 일부 구현예에서, 유전자 변형 단백질 또는 주형 핵산을 인코딩하는 핵산은 제어 요소, 예를 들어 프로모터와 같은 전사 제어 요소에 작동가능하게 연결된다. 전사 제어 요소는 일부 구현예에서 진핵 세포, 예를 들어 포유동물 세포; 또는 원핵 세포(예를 들어, 박테리아 또는 고세포) 중 어느 하나에서 기능성일 수 있다. 일부 구현예에서, 폴리펩티드를 인코딩하는 뉴클레오티드 서열은 다중 제어 요소들에 작동가능하게 연결되며, 예를 들어 이는 원핵 및 진핵 세포 모두에서 폴리펩티드를 인코딩하는 뉴클레오티드 서열의 발현을 가능하게 한다.

[1151] 예시적인 목적을 위해, 공간적으로 제한된 프로모터들의 예는 뉴런-특이적 프로모터, 지방세포-특이적 프로모터, 심근세포-특이적 프로모터, 평활근-특이적 프로모터, 광수용체-특이적 프로모터 등을 포함하지만 이에 제한되지 않는다. 뉴런-특이적인 공간적으로 제한된 프로모터는, 뉴런-특이적 에놀라제(NSE) 프로모터(예를 들어, EMBL HSENO2, X51956 참조); 방향족 아미노산 데카르복실라제(AADC) 프로모터, 신경미세섬유 프로모터(예를 들어, GenBank HUMNFL, L04147, 참조); 시냅신 프로모터(예를 들어, GenBank HUMSYNIB, M55301 참조); thy-1 프로모터(예를 들어, 문헌[Chen et al. (1987) Cell 51:7-19; 및 Llewellyn, et al. (2010) Nat. Med. 16(10):1161-1166] 참조); 세로토닌 수용체 프로모터(예를 들어, GenBank S62283, 참조); 티로신 하이드록실라제 프로모터(TH)(예를 들어, 문헌[Oh et al. (2009) Gene Ther 16:437; Sasaoka et al. (1992) Mol. Brain Res. 16:274; Boundy et al. (1998) J. Neurosci. 18:9989; 및 Kaneda et al. (1991) Neuron 6:583-594)] 참조); GnRH 프로모터(예를 들어, 문헌[Radovick et al. (1991) Proc. Natl. Acad. Sci. USA 88:3402-3406] 참조); L7 프로모터(예를 들어, 문헌[Oberdick et al. (1990) Science 248:223-226] 참조); DNMT 프로모터(예를 들어, 문헌[Bartge et al. (1988) Proc. Natl. Acad. Sci. USA 85:3648-3652] 참조); 엔케팔린(enkephalin) 프로모터(예를 들어, 문헌[Comb et al. (1988) EMBO J. 17:3793-3805] 참조); 미엘린 염기성 단백질(MBP) 프로모터; Ca2+-칼모듈린-의존성 단백질 키나제 II-알파(CamKII α) 프로모터(예를 들어, 문헌[Mayford et al. (1996) Proc. Natl. Acad. Sci. USA 93:13250; 및 Casanova et al. (2001) Genesis 31:37] 참조); CMV 인핸서/혈소판-유래 성장 인자- β 프로모터(예를 들어, 문헌[Liu et al. (2004) Gene Therapy 11:52-60] 참조); 등을 포함하지만 이에 제한되지 않는다.

[1152] 지방세포-특이적인 공간적으로 제한된 프로모터는 aP2 유전자 프로모터/인핸서, 예를 들어 인간 aP2 유전자의 -5.4 kb 내지 +21 bp의 영역(예를 들어, 문헌[Tozzo et al. (1997) Endocrinol. 138:1604; Ross et al. (1990) Proc. Natl. Acad. Sci. USA 87:9590; 및 Pavjani et al. (2005) Nat. Med. 11:797)] 참조); 포도당 수송체-4(GLUT4) 프로모터(예를 들어, 문헌[Knight et al. (2003) Proc. Natl. Acad. Sci. USA 100:14725)] 참조); 지방산 자리옮김효소(FAT/CD36) 프로모터(예를 들어, 문헌[Kuriki et al. (2002) Biol. Pharm. Bull. 25:1476; 및 Sato et al. (2002) J. Biol. Chem. 277:15703)] 참조); 스테아로일-CoA 불포화효소-1(SCD1) 프로모터(문헌[Tabor et al. (1999) J. Biol. Chem. 274:20603] 참조); 렙틴 프로모터(예를 들어, 문헌[Mason et al. (1998) Endocrinol. 139:1013; 및 Chen et al. (1999) Biochem. Biophys. Res. Comm. 262:187)] 참조); 아디포넥틴 프로모터(예를 들어, 문헌[Kita et al. (2005) Biochem. Biophys. Res. Comm. 331:484; 및 Chakrabarti (2010) Endocrinol. 151:2408] 참조); 아디프신 프로모터(예를 들어, 문헌[Platt et al. (1989) Proc. Natl. Acad. Sci. USA 86:7490] 참조); 레시스틴 프로모터(예를 들어, 문헌[Seo et al. (2003) Molec. Endocrinol. 17:1522] 참조); 등을 포함하지만 이에 제한되지 않는다.

[1153] 심근세포-특이적인 공간적으로 제한된 프로모터는 다음 유전자들로부터 유래된 제어 서열: 미오신 경쇄-2, α-미오신 중쇄, AE3, 심장 트로포닌 C, 심장 액틴 등을 포함하지만 이에 제한되지 않는다 (Franz et al. (1997) Cardiovasc. Res. 35:560-566; Robbins et al. (1995) Ann. N.Y. Acad. Sci. 752:492-505; Linn et al. (1995) Circ. Res. 76:584-591; Parmacek et al. (1994) Mol. Cell. Biol. 14:1870-1885; Hunter et al. (1993) Hypertension 22:608-617; 및 Sartorelli et al. (1992) Proc. Natl. Acad. Sci. USA 89:4047-4051).

[1154] 평활근-특이적인 공간적으로 제한된 프로모터는 SM22α 프로모터(예를 들어, 문헌[Akyurek et al. (2000) Mol. Med. 6:983]; 및 미국 특허 7,169,874 참조); 스무드린(smoothelin) 프로모터(예를 들어, WO 2001/018048 참조); α-평활근 액틴 프로모터; 등을 포함하지만 이에 제한되지 않는다. 예를 들어, 그 안에 2개의 CArG 요소가 놓여있는, SM22α 프로모터의 0.4 kb 영역은 혈관 평활근 세포-특이적 발현을 매개하는 것으로 나타났다(예를 들어, 문헌[Kim, et al. (1997) Mol. Cell. Biol. 17, 2266-2278; Li, et al., (1996) J. Cell Biol. 132, 849-859; 및 Moessler, et al. (1996) Development 122, 2415-2425] 참조).

[1155] 광수용체-특이적인 공간적으로 제한된 프로모터는 로돕신 프로모터; 로돕신 키나제 프로모터(Young et al. (2003) Ophthalmol. Vis. Sci. 44:4076); 베타 포스포디에스테라제 유전자 프로모터(Nicoud et al. (2007) J. Gene Med. 9:1015); 망막세포 변성증 유전자 프로모터(Nicoud et al. (2007) 상동); 광수용체간(interphotoreceptor) 레티노이드-결합 단백질(IRBP) 유전자 인핸서(Nicoud et al. (2007) 상동); IRBP 유전자 프로모터(Yokoyama et al. (1992) Exp Eye Res. 55:225) 등을 포함하지만 이에 제한되지 않는다.

[1156] 일부 구현예에서, 유전자 변형 시스템, 예를 들어 유전자 변형 폴리펩티드를 인코딩하는 DNA, 주형 RNA를 인코딩하는 DNA, 또는 이종 목적 서열을 인코딩하는 DNA 또는 RNA는 하나 이상의 요소가 조직-특이적 프로모터, 예를 들어 T-세포에서 활성인 프로모터에 작동가능하게 연결되도록 설계된다. 추가 구현예에서, T-세포 활성 프로모터는 다른 세포 유형, 예를 들어 B-세포, NK 세포에서 비활성이다. 일부 구현예에서, T-세포 활성 프로모터는

T-세포 수용체의 구성요소, 예를 들어 TRAC, TRBC, TRGC, TRDC를 인코딩하는 유전자에 대한 프로모터로부터 유래된다. 일부 구현예에서, T-세포 활성 프로모터는 T-세포-특이적 분화 단백질 클러스터의 구성요소, 예를 들어 CD3, 예를 들어 CD3D, CD3E, CD3G, CD3Z를 인코딩하는 유전자에 대한 프로모터로부터 유래된다. 일부 구현예에서, 유전자 변형 시스템의 T-세포-특이적 프로모터는 세포 유형에 걸쳐 공적으로 입수가능한 유전자 발현 데이터를 비교하고, T-세포에서 향상된 발현을 갖는 유전자로부터 프로모터를 선택함으로써 발견된다. 일부 구현예에서, 프로모터는 원하는 발현 폭에 따라, 예를 들어 T-세포에서만 활성인 프로모터, NK 세포에서만 활성인 프로모터, T-세포 및 NK 세포 둘 다에서 활성인 프로모터를 선택할 수 있다.

[1157] 당업계에 알려진 세포-특이적 프로모터는 유전자 변형 단백질의 발현을, 예를 들어 본 명세서에 기재된 바와 같이, 지시하는 데 사용될 수 있다. 비제한적인 예시적인 포유동물 세포-특이적 프로모터는 세포-특이적 방식으로 Cre 재조합효소를 발현하는 마우스에서 특징화 및 사용된다. 특정 비제한적인 예시적인 포유동물 세포-특이적 프로모터는, 본 명세서에 참고로 포함된 US9845481의 표 1에 열거되었다.

[1158] 일부 구현예에서, 본 명세서에 기재된 벡터는 발현 카세트를 포함한다. 통상적으로, 발현 카세트는 프로모터 서열에 작동가능하게 연결된 본 발명의 핵산 분자를 포함한다. 예를 들어, 그 코딩 서열의 발현에 영향을 미치는 경우, 프로모터는 코딩 서열과 작동가능하게 연결되었다(예를 들어, 코딩 서열이 프로모터의 전사 제어 하에 존재한다). 서열 인코딩은 센스 또는 안티센스 배향으로 조절 서열에 작동가능하게 연결될 수 있다. 특정 구현예에서, 프로모터는 이종성 프로모터이다. 특정 구현예에서, 발현 카세트는 추가의 요소, 예를 들어 인트론, 인핸서, 폴리아데닐화 부위, 우드척 반응 요소(WRE) 및/또는 인코딩 서열의 발현 수준에 영향을 미치는 것으로 알려진 기타 요소를 포함할 수 있다. 프로모터는 코딩 서열 또는 기능성 RNA의 발현을 통상적으로 제어한다. 특정 구현예에서, 프로모터 서열은 근위 및 더 원위의 상류 요소를 포함하며, 인핸서 요소를 추가로 포함할 수 있다. 인핸서는 통상적으로 프로모터 활성을 자극할 수 있고, 프로모터의 선천적인 요소 또는 프로모터의 수준 또는 조직-특이성을 향상하기 위해 삽입된 이종성 요소일 수 있다. 특정 구현예에서, 프로모터는 그 전체가 자연의 유전자로부터 유래된다. 특정 구현예에서, 프로모터는 상이한 자연 발생 프로모터로부터 유래된 상이한 요소로 구성된다. 특정 구현예에서, 프로모터는 합성 뉴클레오티드 서열을 포함한다. 상이한 프로모터는 상이한 조직 또는 세포 유형에서 또는 발달의 상이한 단계에서 또는 상이한 환경 조건이나 약물 또는 전사 보조-인자의 존재 또는 부재에 반응하여 유전자의 발현을 지시할 것임은 당업자는 이해할 것이다. 편재성, 세포 유형-특이적, 조직-특이적, 발달 단계-특이적, 및 조건적 프로모터, 예를 들어 약물-반응성 프로모터(예를 들어, 테트라사이클린-반응성 프로모터)는 당업자에게 잘 알려져 있다. 프로모터의 예는 포스포글리세레이트 키나제(PKG) 프로모터, CAG(CMV 인핸서, 닭 베타 액틴 프로모터(CBA) 및 토끼 베타 글로빈 인트론의 복합체), NSE(뉴런 특이적 엔올라제), 시냅신 또는 NeuN 프로모터, SV40 초기 프로모터, 마우스 유방 종양 바이러스 LTR 프로모터; 아데노바이러스 주요 후기 프로모터(Ad MLP); 단순포진 바이러스(HSV) 프로모터, 거대세포바이러스(CMV) 프로모터, 예컨대 CMV 급 초기 프로모터 영역(CMVIE), SFFV 프로모터, 라우스 육종 바이러스(RSV) 프로모터, 합성 프로모터, 하이브리드 프로모터 등을 포함하지만 이에 제한되지 않는다. 기타 프로모터는 인간 기원 또는 쥐를 포함한 기타 종일 수 있다. 통상의 프로모터는 예를 들어 인간 거대세포바이러스(CMV) 급 초기 유전자 프로모터, SV40 초기 프로모터, 라우스 육종 바이러스 긴 말단 반복부, [베타]- 액틴, 쥐 인슐린 프로모터, 포스포글리세레이트 키나제 인간 알파-1 항트립신(hAAT) 프로모터, 트랜스타이레틴(transthyretin) 프로모터, TBG 프로모터 및 기타 간-특이적 프로모터, 데스민 프로모터 및 유사한 근육-특이적 프로모터, EF1 -알파 프로모터, 다수-조직 특이성을 갖는 하이브리드 프로모터, 시냅신과 같은 뉴런에 특이적인 프로모터 및 글리세르알데히드-3-포스페이트 탈수소화효소 프로모터를 포함하고, 이들 모두는 당업자에게 잘 알려지고 쉽게 입수가능하며, 관심 서열의 코딩 서열의 높은 수준의 발현을 수득하는 데 사용될 수 있다. 추가적으로, 쥐 메탈로티오네인(metallothionein) 유전자와 같은 비-바이러스 유전자로부터 유래된 서열은 본 발명에서의 이용을 또한 찾을 것이다. 그러한 프로모터 서열은 예를 들어 스트라타진(Stratagene)(San Diego, CA 소재)으로부터 상업적으로 구매가능하다. 추가의 예시적인 프로모터 서열은 예를 들어 WO2018213786A1(그 전체가 본 명세서에 참고로 포함됨)에 기재되어 있다.

[1159] 일부 구현예에서, 아포지질단백질 E 인핸서(ApoE) 또는 이의 기능성 단편은 예를 들어 간에서의 발현을 구동하는 데 사용된다. 일부 구현예에서, ApoE 인핸서의 2개의 복제물 또는 이의 기능성 단편이 사용된다. 일부 구현예에서, ApoE 인핸서 또는 이의 기능성 단편은 프로모터, 예를 들어 인간 알파-1 항트립신(hAAT) 프로모터와 조합되어 사용된다.

[1160] 일부 구현예에서, 조절 서열은 조직-특이적 유전자 발현능을 부여한다. 일부 경우들에서, 조직-특이적 조절 서열은 조직 특이적 방식으로 전사를 유도하는 조직-특이적 전사 인자에 결합한다. 다양한 조직-특이적 조절 서열

(예를 들어, 프로모터, 인핸서 등)은 당업계에 알려져 있다. 예시적인 조직-특이적 조절 서열은 다음의 조직-특이적 프로모터: 간-특이적 티록신 결합 글로불린(TBG) 프로모터, 인슐린 프로모터, 글루카곤 프로모터, 소마토스타틴 프로모터, 췌장 폴리펩티드(PPY) 프로모터, 시냅신-1(Syn) 프로모터, 크레아틴 키나제(MCK) 프로모터, 포유동물 데스민(DES) 프로모터, α-미오신 중쇄(a-MHC) 프로모터, 또는 심장 트로포닌 T(cTnT) 프로모터를 포함하지만 이에 제한되지 않는다. 기타 예시적인 프로모터는 베타-액틴 프로모터, B형 간염 바이러스 코어 프로모터(Sandig et al., Gene Ther., 3:1002-9 (1996)); 알파-페토단백질(AFP) 프로모터(Arbuthnot et al., Hum. Gene Ther., 7:1503-14 (1996)), 뼈 오스테오칼신 프로모터(Stein et al., Mol. Biol. Rep., 24:185-96 (1997)); 뼈 시알로단백질 프로모터(Chen et al., J. Bone Miner. Res., 11:654-64 (1996)), CD2 프로모터(Hansal et al., J. Immunol., 161:1063-8 (1998)); 면역글로불린 중쇄 프로모터; T 세포 수용체 α-사슬 프로모터, 뉴런성 예컨대 뉴런-특이적 엔올라제(NSE) 프로모터(Andersen et al., Cell. Mol. Neurobiol., 13:503-15 (1993)), 신경미세섬유 경쇄 유전자 프로모터(Piccioli et al., Proc. Natl. Acad. Sci. USA, 88:5611-5 (1991)), 및 뉴런-특이적 vgf 유전자 프로모터(Piccioli et al., Neuron, 15:373-84 (1995)), 및 기타를 포함한다. 추가의 예시적인 프로모터 서열은 예를 들어 미국 특허 10300146(그 전체가 본 명세서에 참고로 포함됨)에 기재되어 있다. 일부 구현예에서, 조직-특이적 조절 요소, 예를 들어 조직-특이적 프로모터는, 예를 들어 RNA-seq 또는 단백질 발현 데이터, 또는 이의 조합에 의해 측정됨에 따라, 소정의 조직에서 고도로 발현되는 유전자에 작동가능하게 연결되는 것으로 알려진 것으로부터 선택된다. 발현에 의해 조직 특이성을 분석하기 위한 방법은, 그 전체가 본 명세서에 참고로 포함된, 문헌[Fagerberg et al. Mol Cell Proteomics 13(2):397-406 (2014)]에 교시되어 있다.

[1161] 일부 구현예에서, 본 명세서에 기재된 벡터는 멀티시스트로닉(multicistronic) 발현 작제물이다. 멀티시스트로닉 발현 작제물은, 예를 들어 제1 프로모터와 제1 인코딩 핵산 서열을 포함하는 제1 발현 카세트, 및 예를 들어 제2 프로모터와 제2 인코딩 핵산 서열을 포함하는 제2 발현 카세트를 보유하는 작제물을 포함한다. 그러한 멀티시스트로닉 발현 작제물은, 일부 경우에서 헤어핀 RNA와 같이 비번역된 유전자 생성물을, 폴리펩티드, 예를 들어 유전자 변형 폴리펩티드 및 유전자 변형 주형과 함께 전달하는데 특히 유용하다. 일부 구현예에서, 멀티시스트로닉 발현 작제물은, 예를 들어 프로모터 간섭 또는 인접한 비혼화성 핵산 요소 존재로 인해, 하나 이상의 포함된 이식 유전자의 감소된 발현 수준을 나타낼 수 있다. 멀티시스트로닉 발현 작제물이 바이러스 벡터의 일부인 경우, 자기-상보적 핵산 서열의 존재는 일부 경우에서 바이러스 복제 또는 패키징에 필요한 구조의 형성을 간섭할 수 있다.

[1162] 일부 구현예에서, 서열은 헤어핀을 갖는 RNA를 인코딩한다. 일부 구현예에서, 헤어핀 RNA는 가이드 RNA, 주형 RNA, shRNA, 또는 마이크로RNA이다. 일부 구현예에서, 제1 프로모터는 RNA 중합효소 I 프로모터이다. 일부 구현예에서, 제1 프로모터는 RNA 중합효소 II 프로모터이다. 일부 구현예에서, 제2 프로모터는 RNA 중합효소 III 프로모터이다. 일부 구현예에서, 제2 프로모터는 U6 또는 H1 프로모터이다.

[1163] 이론에 구속되고자 함이 없이, 멀티시스트로닉 발현 작제물은 단지 하나의 시스트론을 함유하는 발현 시스템과 비교하여 최적의 발현 수준을 달성할 수 없다. 둘 이상의 프로모터 요소를 포함하는 멀티시스트로닉 발현 작제물로 달성된 보다 적은 발현 수준에 대해 제안된 이유들 중 하나는 프로모터 간섭의 현상이다(예를 들어, 문헌[Curtin J A, Dane A P, Swanson A, Alexander I E, Ginn S L. *Bidirectional promoter interference between two widely used internal heterologous promoters in a late-generation lentiviral construct*. Gene Ther. 2008 March; 15(5):384-90; 및 Martin-Duque P, Jezzard S, Kaftansis L, Vassaux G. *Direct comparison of the insulating properties of two genetic elements in an adenoviral vector containing two different expression cassettes*. Hum Gene Ther. 2004 October; 15(10):995-1002] 참조; 프로모터 간섭 현상의 개시를 위해, 두 참고 모두 본 명세서에 참고로 포함됨). 일부 구현예에서, 프로모터 간섭의 문제는, 예를 들어 내부 리보솜 도입 부위에 의해 분리된 다중 인코딩 핵산 서열의 전사를 구동하는 단지 하나의 프로모터를 포함하는 멀티시스트로닉 발현 작제물을 생산함으로써, 또는 그 자신의 프로모터를 포함하는 시스트론을 전사 절연체 요소로 구분함으로써 극복될 수 있다. 일부 구현예에서, 다중 시스트론의 단일-프로모터 구동 발현은 시스트론의 균일하지 않은 발현 수준을 생성할 수 있다. 일부 구현예에서, 프로모터는 효율적으로 분리될 수 없고, 분리 요소는 일부 유전자 전이(gene transfer) 벡터, 예를 들어 일부 레트로바이러스 벡터들과 양립성이 아닐 수 있다.

[1164] *마이크로RNA*

[1165] 마이크로RNA(miRNA) 및 기타 간섭 핵산은 표적 RNA 전사체 절단/분해 또는 표적 메신저 RNA(mRNA)의 번역 저해를 통해 유전자 발현을 일반적으로 조절한다. miRNA는 일부 경우에서, 통상적으로 최종 19 내지 25의 비번역된 RNA 생성물로서 자연적으로 발현될 수 있다. miRNA는 표적 mRNA의 3' 비번역된 영역(UTR)과 서열-특이적 상호작

용을 통한 그의 활성을 일반적으로 나타낸다. 이들 내인성적으로 발현된 miRNA는 miRNA 듀플렉스로, 나아가 성숙한 단일 가닥 miRNA 분자로 순차적으로 가공되는 헤어핀 전구체를 형성할 수 있다. 이러한 성숙한 miRNA는 일반적으로 다중단백질 복합체, miRISC를 가이드하며, 이는 성숙한 miRNA에 대한 그들의 상보성에 근거하여 표적 mRNA의 표적 3' UTR 영역을 확인한다. 유용한 이식유전자 생성물은 연결된 폴리펩티드의 발현을 조절하는, 예를 들어 miRNA 또는 miRNA 결합 부위를 포함할 수 있다. miRNA 유전자의 비제한적인 열거; 이들 유전자 및 이들의 동족체는 예를 들어 참조로서 본 명세서에 포함된 US10300146, 22:25-25:48에 열거된 바와 같은 방법에서, 이식유전자로서 또는 소간섭 핵산에 대한 표적으로서 유용하다(예를 들어, miRNA 스폰지, 안티센스 올리고뉴클레오티드). 일부 구현예에서, 하나 이상의 상기 miRNA에 대한 하나 이상의 결합 부위는 이식유전자, 예를 들어 rAAV 벡터에 의해 전달된 이식유전자 내에 포함되어, 예를 들어 이식유전자를 보유하는 동물의 하나 이상의 조직에서 그 이식유전자의 발현을 억제한다. 일부 구현예에서, 결합 부위는 조직 특이적 방식으로 이식유전자의 발현을 제어하도록 선택될 수 있다. 예를 들어, 간-특이적 miR-122에 대한 결합 부위는 간에서 그 이식유전자의 발현을 저해하기 위해 이식유전자 내로 포함될 수 있다. 추가의 예시적인 miRNA 서열은, 예를 들어 미국 특허 10,300,146(그 전체가 본 명세서에 참고로 포함됨)에 기재되어 있다.

[1166] miR 억제제 또는 miRNA 억제제는, 일반적으로 miRNA 발현 및/또는 가공을 차단하는 작용제이다. 그러한 작용물의 예는 마이크로RNA 안타고니스트, 마이크로RNA 특이적 안티센스, 마이크로RNA 스폰지 및 마이크로RNA 올리고뉴클레오티드(이중 가닥, 헤어핀, 짧은 올리고뉴클레오티드)를 포함하지만 이에 제한되지 않으며 이는 드로샤(Drosha) 복합체와의 miRNA 상호작용을 억제한다. 마이크로RNA 억제제, 예를 들어 miRNA 스폰지는 이식유전자로부터의 세포 내에서 발현될 수 있다(예를 들어, 문헌[Ebert, M. S. Nature Methods, Epub Aug. 12, 2007]에 기재된 바와 같음; 이는 그 전체가 본 명세서에 참고로 포함됨). 일부 구현예에서, 마이크로RNA 스폰지, 또는 기타 miR 억제제가 AAV와 사용된다. 마이크로RNA 스폰지는 상보적인 칠량체 시드(seed) 서열을 통해 miRNA를 전반적으로 특이하게 억제한다. 일부 구현예에서, miRNA의 전체 과는 단일 스폰지 서열을 사용하여 침묵화될 수 있다. miRNA 기능을 세포에서 침묵화하기 위한 기타 방법(miRNA 표적의 탈억제)는 당업자에게 명백할 것이다.

[1167] 일부 구현예에서, 유전자 변형 시스템, 주형 RNA, 또는 본 명세서에 기재된 폴리펩티드는 표적 조직, 예를 들어 제1 조직에 투여되거나 표적 조직에서 활성이다(예를 들어, 표적 조직에서 더 활성임). 일부 구현예에서, 유전자 변형 시스템, 주형 RNA, 또는 폴리펩티드는 비-표적 조직에 투여되지 않거나 비-표적 조직에서 덜 활성이다(예를 들어, 비-표적 조직에서 활성이 아님). 일부 구현예에서, 본 명세서에 기재된 유전자 변형 시스템, 주형 RNA, 또는 폴리펩티드는 표적 조직, 예를 들어 제1 조직 내에서 DNA를 변형시키는 데(예를 들어, 비-표적 조직에서 DNA를 변형시키지 않는 데) 유용하다.

[1168] 일부 구현예에서, 유전자 변형 시스템은 (a) 본 명세서에 기재된 폴리펩티드 또는 이를 인코딩하는 핵산, (b) 본 명세서에 기재된 주형 핵산(예를 들어, 주형 RNA), 및 (c) 표적 조직에 특이적인 하나 이상의 제1 조직-특이적 발현-제어 서열을 포함하고, 여기서 표적 조직에 특이적인 하나 이상의 제1 조직-특이적 발현-제어 서열은 (a), (b), 또는 (a) 및 (b)과 작동적으로 연관되며, 여기서 (a)에 회합된 경우, (a)는 폴리펩티드를 인코딩하는 핵산을 포함한다.

[1169] 일부 구현예에서, (b)에서 핵산은 RNA를 포함한다.

[1170] 일부 구현예에서, (b)에서 핵산은 DNA를 포함한다.

[1171] 일부 구현예에서, (b)에서 핵산은: (i) 단일 가닥이거나, 단일 가닥 절편을 포함하며, 예를 들어 단일 가닥 DNA이거나, 단일 가닥 절편 및 하나 이상의 이중 가닥 절편을 포함하고; (ii) 역 말단 반복부를 갖거나; 또는 (iii) (i)과 (ii) 둘 다이다.

[1172] 일부 구현예에서, (b)에서 핵산은 이중 가닥이거나 또는 이중 가닥 절편을 포함한다.

[1173] 일부 구현예에서, (a)는 폴리펩티드를 인코딩하는 핵산을 포함한다.

[1174] 일부 구현예에서, (a)에서 핵산은 RNA를 포함한다.

[1175] 일부 구현예에서, (a)에서 핵산은 DNA를 포함한다.

[1176] 일부 구현예에서, (a)에서 핵산은: (i) 단일 가닥이거나, 단일 가닥 절편을 포함하며, 예를 들어 단일 가닥 DNA이거나, 단일 가닥 절편 및 하나 이상의 이중 가닥 절편을 포함; (ii) 역 말단 반복부를 가짐; 또는 (iii) (i)과 (ii) 둘 다이다.

[1177] 일부 구현예에서, (a)에서 핵산은 이중 가닥이거나 또는 이중 가닥 절편을 포함한다.

[1178] 일부 구현예에서, (a), (b), 또는 (a) 및 (b)에서 핵산은 선형이다.

[1179] 일부 구현예에서, (a), (b), 또는 (a) 및 (b)에서 핵산은 환형, 예를 들어, 플라스미드 또는 미니서클이다.

[1180] 일부 구현예에서, 이종성 목적 서열은 제1 프로모터와 작동적으로 연관된다.

[1181] 일부 구현예에서, 하나 이상의 제1 조직-특이적 발현-제어 서열은 조직 특이적 프로모터를 포함한다.

[1182] 일부 구현예에서, 조직-특이적 프로모터는: (i) 이종성 목적서열, (ii) 레트로바이러스 RT를 인코딩하는 핵산, 또는 (iii) (i) 및 (ii)와 작동적으로 연관된 제1 프로모터를 포함한다.

[1183] 일부 구현예에서, 하나 이상의 제1 조직-특이적 발현-제어 서열은: (i) 이종성 목적서열, (ii) 레트로바이러스 RT 도메인을 인코딩하는 핵산, 또는 (iii) (i) 및 (ii)와 작동적으로 연관된 조직-특이적 마이크로RNA 인식 서열을 포함한다.

[1184] 일부 구현예에서, 시스템은 조직-특이적 프로모터를 포함하고, 시스템은 하나 이상의 조직-특이적 마이크로RNA 인식 서열을 추가로 포함하고, 여기서: (i) 조직-특이적 프로모터는: (I) 이종성 목적 서열, (II) 레트로바이러스 RT 도메인을 인코딩하는 핵산, 또는 (III) (I) 및 (II)와 작동적으로 연관되며; 및/또는 (ii) 하나 이상의 조직-특이적 마이크로RNA 인식 서열은: (I) 이종성 목적 서열, (II) 레트로바이러스 RT를 인코딩하는 핵산, 또는 (III) (I) 및 (II)와 작동적으로 연관된다.

[1185] 일부 구현예에서, (a)는 폴리펩티드를 인코딩하는 핵산을 포함하며, 핵산은 폴리펩티드를 인코딩하는 핵산과 작동적으로 연관된 프로모터를 포함한다.

[1186] 일부 구현예에서, 폴리펩티드를 인코딩하는 핵산은, 폴리펩티드 코딩 서열과 작동적으로 연관된 표적 조직에 대해 특이적인 하나 이상의 제2의 조직-특이적 발현-제어 서열을 포함한다.

[1187] 일부 구현예에서, 하나 이상의 제2 조직-특이적 발현-제어 서열은 조직 특이적 프로모터를 포함한다.

[1188] 일부 구현예에서, 조직-특이적 프로모터는 폴리펩티드를 인코딩하는 핵산과 작동적으로 연관된 프로모터이다.

[1189] 일부 구현예에서, 하나 이상의 제2 조직-특이적 발현-제어 서열은 조직-특이적 마이크로RNA 인식 서열을 포함한다.

[1190] 일부 구현예에서, 폴리펩티드를 인코딩하는 핵산과 작동적으로 연관된 프로모터는 조직-특이적 프로모터로, 시스템은 하나 이상의 조직-특이적 마이크로RNA 인식 서열을 추가로 포함한다.

[1191] 일부 구현예에서, 본 발명에 의해 제공된 시스템의 핵산 구성요소인, (예를 들어, 폴리펩티드를 인코딩하거나, 이종성 목적 서열을 포함하는) 서열은 단백질 발현 수준을 변형시키는 비번역 영역에 의해 측접된다. 다양한 5' 및 3' UTR은 단백질 발현에 영향을 미칠 수 있다. 예를 들어, 일부 구현예에서, 코딩 서열은 RNA 안정성 또는 단백질 번역을 변형시키는 5' UTR이 선행할 수 있다. 일부 구현예에서, 서열은 RNA 안정성 또는 단백질 번역을 변형시키는 3' UTR이 뒤따를 수 있다. 일부 구현예에서, 서열은 5' UTR이 선행하고 3' UTR이 뒤따를 수 있으며 이는 RNA 안정성 또는 번역을 변형시킨다. 일부 구현예에서, 보체 인자 3의 5' 및/또는 3' UTR(C3)(cactcctccccatcctctccctctgtccctctgtccctctgaccctgcactgtcccagcacc) 또는 오로소뮤코이드(orosomucoid) 1(ORM1)(caggacacagccttggatcaggacagagacttgggggccatcctgcccctccaacccgacatgtgtacctcagctttttccctcacttgcatcaataaagcttctgtgtttggaacagctaa)로부터 선택될 수 있다(Asrani et al. RNA Biology 2018). 특정 구현예에서, 5' UTR은 C3으로부터의 5'UTR이고, 3' UTR은 ORM1로부터의 3' UTR이다. 특정 구현예에서, 단백질 발현을 위한 5' UTR 및 3' UTR, 예를 들어 유전자 변형 폴리펩티드 또는 이종성 목적 서열을 위한 mRNA(또는 그 RNA를 인코딩하는 DNA)는 최적화된 발현 서열을 포함한다. 일부 구현예에서, 5' UTR은 GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGAGCCACC 및/또는 UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGA)을 포함하는 3' UTR을 포함하며, 이는 예를 들어, 문헌[Richner et al. *Cell* 168(6): P1114-1125 (2017)]에 기재된 바와 같으며, 이의 서열은 본 명세서에 참고로 포함된다.

[1192] 일부 구현예에서, 5' 및/또는 3' UTR은 단백질 발현을 향상시키기 위해 선택될 수 있다. 일부 구현예에서, 5' 및/또는 3' UTR은 과잉생산 저해가 최소화되도록 단백질 발현을 변형시키기 위해 선택될 수 있다. 일부 구현예

에서, UTR은 코딩 서열 주위에 존재하며, 예를 들어 코딩 서열 외부에 존재하며, 다른 구현예는 코딩 서열에 인접하고, 일부 구현예에서는 추가의 조절 요소들(예를 들어, miRNA 결합 부위, cis-조절 부위)이 UTR에 포함된다.

[1193] 일부 구현예에서, 유전자 변형 시스템의 오픈 판독 프레임, 예를 들어 유전자 변형 폴리펩티드를 인코딩하는 mRNA(또는 mRNA를 인코딩하는 DNA)의 ORF 또는 이종성 목적 서열의 mRNA(또는 mRNA를 인코딩하는 DNA)의 하나 이상의 ORF는, 그의 발현을 향상시키는 5' 및/또는 3' 비번역 영역(UTR)에 의해 측접된다. 일부 구현예에서, 시스템의 mRNA 구성요소(또는 DNA 구성요소로부터 생산된 전사체)의 5' UTR은 서열 5'-GGGAAAUAAGAGAGAAAAGAAGAGUAAGAAGAAAUAUAAGAGCCACC-3'을 포함한다. 일부 구현예에서, 시스템의 mRNA 구성요소(또는 DNA 구성요소로부터 생산된 전사체)의 3' UTR은 서열 5'-UGAUAAUAGGCUGGAGCCUCGGUGGCCAUGCUUCUUGCCCCUUGGGCCUCCCCCCAGCCCCUCCUCCCCUUCCUGCACCCGUACCCCCGUGGUCUUUGAAUAAAGUCUGA-3'을 포함한다. 5' UTR 및 3' UTR의 이러한 조합은 문헌[Richner et al. Cell 168(6): P1114-1125 (2017)]에 의해, 작동가능하게 연결된 ORF의 바람직한 발현을 생성하는 것으로 나타났으며, 여기 포함된 교시 및 서열은 여기에 참조로서 포함된다. 일부 구현예에서, 본 명세서에 기재된 시스템은 전사체를 인코딩하는 DNA를 포함하고, 여기서 DNA는, 상응하는 5' UTR 및 3' UTR 서열을 포함하고, 상기 열거된 서열에서 U는 T로 치환된다. 일부 구현예에서, 시스템의 RNA 구성요소를 생산하기 위해 사용된 DNA 벡터는 시험관 내 전사를 개시하기 위해 5' UTR의 프로모터 상류, 예를 들어 T7, T3, 또는 SP6 프로모터를 더 포함한다. 상기 5' UTR은 GGG로 시작하며, 이는 T7 RNA 중합효소를 이용하여 전사를 최적화하기 위한 적합한 출발이다. 전사 수준을 조절하고, 대안적인 5' UTR을 적합화하기 위해 전사 개시 부위를 변경하기 위해, 문헌[Davidson et al. *Pac Symp Biocomput* 433-443 (2010)]의 교시는 T7 프로모터 변이체 및 그의 발견 방법을 설명하며, 이는 이들 두 특성 모두를 충족한다.

[1194] *바이러스 벡터 및 이의 구성요소*

[1195] 바이러스는, 본 명세서에 기재된 바와 같은 관련 효소 또는 도메인, 예를 들어 본 명세서에 사용된 중합효소 또는 중합효소 기능, 예를 들어 DNA-의존성 DNA 중합효소, RNA-의존성 RNA 중합효소, RNA-의존성 DNA 중합효소, DNA-의존성 RNA 중합효소, 역전사효소의 공급원에 더하여, 본 명세서에 기재된 시스템을 위한 전달 비히클의 유용한 공급원이다. 일부 효소들, 예를 들어 역전사효소는 다중 활성을 가질 수 있으며, 예를 들어 RNA-의존성 DNA 중합 및 DNA-의존성 DNA 중합 둘 모두, 예를 들어 제1 및 제2 가닥 합성이 가능하다. 일부 구현예에서, 유전자 변형 전달 시스템 또는 이의 구성요소들의 공급원으로 사용되는 바이러스는 문헌[Baltimore *Bacteriol Rev* 35(3):235-241 (1971)]에 의해 기재된 바와 같은 그룹으로부터 선택될 수 있다.

[1196] 일부 구현예에서, 바이러스는 그룹 I 바이러스로부터 선택되고, 예를 들어 DNA 바이러스이며, dsDNA를 비리온 내로 패키징한다. 일부 구현예에서, 그룹 I 바이러스는 예를 들어, 아데노바이러스, 헤르페스바이러스, 폭스바이러스(Poxviruse)로부터 선택된다.

[1197] 일부 구현예에서, 바이러스는 그룹 II 바이러스로부터 선택되고, 예를 들어 DNA 바이러스이며 ssDNA를 비리온 내로 패키징한다. 일부 구현예에서, 그룹 II 바이러스는 예를 들어 파보바이러스(Parvovirus)로부터 선택된다. 일부 구현예에서, 파보바이러스는 디펜도파보바이러스(dependoparvovirus), 예를 들어 아데노-연관 바이러스(AAV)이다.

[1198] 일부 구현예에서, 바이러스는 그룹 III 바이러스로부터 선택되고, 예를 들어 RNA 바이러스이며 ssDNA를 비리온 내로 패키징한다. 일부 구현예에서, 그룹 III 바이러스는, 예를 들어 레오바이러스(Reoviruse)로부터 선택된다. 일부 구현예에서, 그러한 비리온에 함유된 dsRNA 중 하나 또는 두 가닥 모두는 숙주 세포 내로의 형질도입시 mRNA로서 직접 제공될 수 있는 코딩 분자이며, 예를 들어 숙주 세포 내로 형질도입시 임의의 개입하는 핵산 복제 또는 중합 단계를 필요로 함이 없이 단백질로 직접 번역될 수 있다.

[1199] 일부 구현예에서, 바이러스는 그룹 IV 바이러스로부터 선택되고, 예를 들어 RNA 바이러스이며, ssRNA(+)를 비리온 내로 패키징한다. 일부 구현예에서, 그룹 IV 바이러스는, 예를 들어 코로나바이러스, 피코르나바이러스(Picornavirus), 토가바이러스(Togavirus)로부터 선택된다. 일부 구현예에서, 그러한 비리온 내에 함유된 ssRNA(+)는 숙주 세포 내로의 형질도입시 mRNA로서 직접 제공될 수 있는 코딩 분자이며, 예를 들어 숙주 세포 내로 형질도입시 임의의 개입하는 핵산 복제 또는 중합 단계를 필요로 함이 없이 단백질로 직접 번역될 수 있다.

[1200] 일부 구현예에서, 바이러스는 그룹 V 바이러스로부터 선택되고, 예를 들어 RNA 바이러스이며, ssRNA(-)를 비리

온 내로 패키징한다. 일부 구현예에서, 그룹 V 바이러스는, 예를 들어 오르토믹소바이러스(Orthomyxoviruses), 랍도바이러스(Rhabdoviruses)로부터 선택된다. 일부 구현예에서, ssRNA(-) 게놈을 갖는 RNA 바이러스는 바이러스 게놈과 함께 숙주 세포로 형질도입되는 비리온 내부의 효소, 예를 들어 RNA-의존성 RNA 중합효소 또한 운반하며, 이는 ssRNA(-)를 숙주에 의해 직접 번역될 수 있는 ssRNA(+)로 카피할 수 있다.

[1201] 일부 구현예에서, 바이러스는 그룹 VI 바이러스로부터 선택되고, 예를 들어 레트로바이러스이며, ssRNA(+)를 비리온 내로 패키징한다. 일부 구현예에서, 그룹 VI 바이러스는, 예를 들어 레트로바이러스로부터 선택된다. 일부 구현예에서, 레트로바이러스는 렌티바이러스, 예를 들어 HIV-1, HIV-2, SIV, BIV이다. 일부 구현예에서, 레트로바이러스는 스푸마바이러스(spumavirus), 예를 들어 포미 바이러스, 예를 들어, HFV, SFV, BFV이다. 일부 구현예에서, 그러한 비리온 내에 함유된 ssRNA(+)는 숙주 세포 내로의 형질도입시 mRNA로서 직접 제공될 수 있는 코딩 분자이며, 예를 들어 숙주 세포 내로 형질도입시 임의의 개입하는 핵산 복제 또는 중합 단계를 필요로 함이 없이 단백질로 직접 번역될 수 있다. 일부 구현예에서, ssRNA(+)는 먼저 역전사되고, 복제되어 dsDNA 게놈 중간체를 생성하고, 이로부터 mRNA가 숙주 세포 내에서 전사될 수 있다. 일부 구현예에서, ssRNA(+) 게놈을 갖는 RNA 바이러스는 바이러스 게놈을 이용하여 숙주 세포로 형질도입되는 비리온 내 효소, 예를 들어 RNA-의존성 DNA 중합효소도 운반하며, 이는 mRNA로 전사될 수 있고 숙주에 의해 번역될 수 있는 dsDNA 내로 ssRNA(+)를 카피할 수 있다. 일부 구현예에서, 그룹 VI 레트로바이러스로부터의 역전사효소는 유전자 변형 폴리펩티드의 역전사효소 도메인으로서 포함된다.

[1202] 일부 구현예에서, 바이러스는 그룹 VII 바이러스로부터 선택되고, 예를 들어 레트로바이러스이며, dsRNA를 비리온 내로 패키징한다. 일부 구현예에서, 그룹 VII 바이러스는, 예를 들어 헤파드나바이러스(Hepadnavirus)로부터 선택된다. 일부 구현예에서, 그러한 비리온에 함유된 dsRNA 중 하나 또는 두 가닥 모두는 숙주 세포 내로의 형질도입시 mRNA로서 직접 제공될 수 있는 코딩 분자이며, 예를 들어 숙주 세포 내로 형질도입시 임의의 개입하는 핵산 복제 또는 중합 단계를 필요로 함이 없이 단백질로 직접 번역될 수 있다. 일부 구현예에서, 그러한 비리온 내에 함유된 dsRNA의 하나 또는 두 가닥 모두는 먼저 역전사되고, 복제되어 dsDNA 게놈 중간체를 생성하며, 이로부터 mRNA가 숙주 세포 내에서 전사될 수 있다. 일부 구현예에서, dsRNA 게놈을 갖는 RNA 바이러스는 바이러스 게놈을 이용하여 숙주 세포로 형질도입되는 비리온 내 효소, 예를 들어 RNA-의존성 DNA 중합효소도 운반하며, 이는 mRNA로 전사될 수 있고 숙주에 의해 번역될 수 있는 dsDNA 내로 dsRNA를 카피할 수 있다. 일부 구현예에서, 그룹 VII 레트로바이러스로부터의 역전사효소는 유전자 변형 폴리펩티드의 역전사효소 도메인으로서 포함된다.

[1203] 일부 구현예에서, 본 발명에서 핵산을 전달하기 위해 사용된 비리온은 유전자 변형의 과정에 관여하는 효소들 또한 운반할 수 있다. 예를 들어, 레트로바이러스 비리온은 핵산과 함께 숙주 세포 내로 전달된 역전사효소 도메인을 함유할 수 있다. 일부 구현예에서, RNA 주형은 비리온 내에서 유전자 변형 폴리펩티드와, 둘 모두가 바이러스 입자로부터 핵산의 형질도입시 표적 세포에 함께 전달되도록 회합될 수 있다. 일부 구현예에서, 비리온 내 핵산은 DNA를 포함할 수 있으며, 예를 들어 선형 ssDNA, 선형 dsDNA, 환형 ssDNA, 환형 dsDNA, 미니서클 DNA, dbDNA, ceDNA를 포함할 수 있다. 일부 구현예에서, 비리온 내 핵산은 RNA, 예를 들어 RNA, 예를 들어, 선형 ssRNA, 선형 dsRNA, 환형 ssRNA, 환형 dsRNA를 포함할 수 있다. 일부 구현예에서, 바이러스 게놈은 숙주 세포 내로의 형질도입시 환형화될 수 있으며, 예를 들어 ssRNA 분자는 공유 결합되어 환형 ssRNA를 형성할 수 있고, 선형 dsRNA 분자는 공유 결합되어 환형 dsDNA 또는 하나 이상의 환형 ssRNA를 형성할 수 있다. 일부 구현예에서, 바이러스 게놈은 숙주 세포에서 회전 환 복제에 의해 복제될 수 있다. 일부 구현예에서, 바이러스 게놈은 단일 핵산 분자를 포함할 수 있고, 예를 들어 비-절편화된 게놈을 포함한다. 일부 구현예에서, 바이러스 게놈은 둘 이상의 핵산 분자를 포함할 수 있으며, 예를 들어 절편화된 게놈을 포함한다. 일부 구현예에서, 비리온 내 핵산은 하나 이상의 단백질과 회합될 수 있다. 일부 구현예에서, 비리온 내 하나 이상의 단백질은 형질도입시 숙주 세포로 전달될 수 있다. 일부 구현예에서, 천연 바이러스는, 비리온 패키징 신호를 표적 핵산에의 첨가함으로써 핵산 전달을 위해 조정될 수 있으며, 여기서 숙주 세포는 패키징 신호를 함유한 표적 핵산을 패키징하기 위해 사용된다.

[1204] 일부 구현예에서, 전달 비히클로 사용된 비리온은 공생하는 인간 바이러스를 포함할 수 있다. 일부 구현예에서, 전달 비히클로 사용된 비리온은 아넬로바이러스를 포함할 수 있으며, 이의 사용은 WO2018232017A1에 설명되고, 이는 그 전체가 본 명세서에 참고로 포함된다.

[1205] *AAV 투여*

[1206] 일부 구현예에서, 아데노-연관 바이러스(AAV)는 본 명세서에 기재된 시스템, 주형 핵산, 및/또는 폴리펩티드와

함께 사용된다. 일부 구현예에서, AAV는 본 명세서에 기재된 시스템, 주형 핵산, 및/또는 폴리펩티드를 전달, 투여 또는 패키징하기 위해 사용된다. 일부 구현예에서, AAV는 재조합 AAV(rAAV)이다.

[1207] 일부 구현예에서, 시스템은 (a) 본 명세서에 기재된 폴리펩티드 또는 이를 인코딩하는 핵산, (b) 본 명세서에 기재된 주형 핵산(예를 들어, 주형 RNA), 및 (c) 표적 조직에 특이적인 하나 이상의 제1 조직-특이적 발현-제어 서열을 포함하고, 여기서 표적 조직에 특이적인 하나 이상의 제1 조직-특이적 발현-제어 서열은 (a), (b), 또는 (a) 및 (b)과 작동적으로 연관되며, 여기서 (a)에 회합된 경우, (a)는 폴리펩티드를 인코딩하는 핵산을 포함한다.

[1208] 일부 구현예에서, 본 명세서에 기재된 시스템은 제1의 재조합 아데노-연관 바이러스(rAAV) 캡시드 단백질을 더 포함하고; 여기서(a) 또는 (b) 중 적어도 하나는 제1의 rAAV 캡시드 단백질과 회합되고, 여기서(a) 또는 (b) 중 적어도 하나는 AAV 역 말단 반복부(ITR)에 의해 측접된다.

[1209] 일부 구현예에서, (a) 및 (b)는 제1 rAAV 캡시드 단백질과 회합된다.

[1210] 일부 구현예에서, (a) 및 (b)는 단일 핵산 상에 존재한다.

[1211] 일부 구현예에서, 시스템은 제2의 rAAV 캡시드 단백질을 더 포함하고, 여기서(a) 또는 (b) 중 적어도 하나는 제2의 rAAV 캡시드 단백질과 회합되고, 여기서 제2의 rAAV 캡시드 단백질과 회합된 (a) 또는 (b) 중 적어도 하나는 제1의 rAAV 캡시드 단백질과 회합된 (a) 또는 (b) 중 적어도 하나와 상이하다.

[1212] 일부 구현예에서, 제1 또는 제2의 rAAV 캡시드 단백질과 회합된 (a) 또는 (b) 중 적어도 하나는 제1 또는 제2 rAAV 캡시드 단백질의 내부에 분산되며, 제1 또는 제2 rAAV 캡시드 단백질은 AAV 캡시드 입자의 형태로 존재한다.

[1213] 일부 구현예에서, 시스템은 나노입자를 더 포함하며, 여기서 나노입자는 (a) 또는 (b) 중 적어도 하나와 회합된다.

[1214] 일부 구현예에서, (a) 및 (b)는 각각: a) 제1 rAAV 캡시드 단백질 및 제2 rAAV 캡시드 단백질; b) 나노입자 및 제1 rAAV 캡시드 단백질; c) 제1 rAAV 캡시드 단백질; d) 제1 아데노바이러스 캡시드 단백질; e) 제1 나노입자 및 제2 나노입자; 또는 f) 제1 나노입자와 회합된다.

[1215] 바이러스 벡터는, 예를 들어 본 발명에 의해 제공된 방법에서의 이용을 위해, 본 발명에 의해 제공된 시스템 전체 또는 부분을 전달하는데 유용하다. 상이한 바이러스로부터 유래된 시스템은 폴리펩티드, 또는 핵산의 전달을 위해 사용되어왔으며; 예를 들어 인테그라제-결핍 렌티바이러스, 아데노바이러스, 아데노-연관 바이러스(AAV), 단순포진 바이러스, 및 배큘로바이러스(문헌[Hodge et al. Hum Gene Ther 2017; Narayanavari et al. Crit Rev Biochem Mol Biol 2017; Boehme et al. Curr Gene Ther 2015]에서 검토됨).

[1216] 아데노바이러스는 잘 정의된 생물학, 유전적 안정성, 높은 형질도입 효율, 및 거대규모 생산의 용이성을 제공하는 유전자 전달 비히클로 사용되어온 흔한 바이러스다(예를 들어, 문헌[Lee et al. Genes & Diseases 2017]에 의해 검토됨, 참고). 이들은 선형 dsDNA 게놈을 갖고, 조직 및 세포 향성(tropism)에서 상이한 다양한 혈청형으로 나타난다. 수령자 세포에서 감염성 바이러스의 복제를 방지하기 위해, 패키징을 위해 사용된 아데노바이러스 게놈은 일부 또는 모든 내인성 바이러스 단백질이 결실되며, 이는 바이러스 생산 세포에 인 트랜스(in trans) 제공된다. 이는 게놈을 헬퍼-의존성으로 만들며, 이는 소위 헬퍼 기능에 의해 제공되는 상실 구성요소들의 존재하에서만 이들이 바이러스 입자들 내로 복제 및 패키징될 수 있음을 의미한다. 모든 바이러스 ORF가 제거된 헬퍼-의존성 아데노바이러스 시스템은 약 37 kb에 달하는 외래 DNA 패키징에 적합할 수 있다(Parks et al. J Virol 1997). 일부 구현예에서, 아데노바이러스 벡터는 유전자 변형 시스템의 폴리펩티드 또는 주형 구성요소에 상응하는 DNA를 전달하기 위해 사용되거나, 또는 이들 모두는 별개의 또는 동일한 아데노바이러스 벡터 상에 함유된다. 일부 구현예에서, 아데노바이러스는 자기-패키징을 할 수 없는 헬퍼-의존성 아데노바이러스(HD-AdV)이다. 일부 구현예에서, 아데노바이러스는 아데노바이러스 입자 내로의 패키징에 필요한 서열 구성요소를 유지하는 한편, 결실된 내인성 ORF의 모든 또는 실질적인 부분들을 가졌던 고용량 아데노바이러스(HC-AdV)이다. 이 유형의 벡터에 대해, 게놈 패키징에 필요한 아데노바이러스 서열만이 비코딩 서열이다: 양 말단 모두에서 역 말단 반복부(ITR) 및 5'-말단에서 패키징 신호(Jager et al. Nat Protoc 2009). 일부 구현예에서, 아데노바이러스 게놈은 최적 생산 및 안정성을 위한 최소 게놈 크기를 충족시키기 위한 스터퍼 DNA를 또한 포함한다(예를 들어, 문헌[Hausl et al. Mol Ther 2010] 참조). 일부 구현예에서, 아데노바이러스는 유전자 변형 시스템을 간으로 전달하는 데 사용된다.

[1217] 일부 구현예에서, 아데노바이러스는 유전자 변형 시스템을 HSCs, 예를 들어, HDAd5/35++로 전달하기 위해 사용된다. HDAd5/35++는, 벡터를 간으로부터 탈표적화하는 변형된 혈청형 35를 갖는 아데노바이러스이다(Wang et al. Blood Adv 2019). 일부 구현예에서, 유전자 변형 시스템을 HSC로 전달하는 아데노바이러스는 원시적 HSC, 예를 CD46 상에서 특이적으로 발현되는 수용체를 이용한다.

[1218] 아데노-연관 바이러스(AAV)는 파보비리데(parvoviridae) 과에 속하며, 더 구체적으로는 디펜도파보바이러스 속을 구성한다. AAV 게놈은 선형 단일 가닥 DNA 분자로 구성되며, 이는 대략 4.7 킬로베이스(kb)를 함유하고, 비-구조적 Rep(복제) 및 구조적 Cap(캡시드) 단백질을 인코딩하는 2개의 주요한 오픈 판독 프레임(ORF)로 이루어진다. 캡 유전자 내에서 제2의 ORF는 조립-활성화 단백질(AAP)을 인코딩하는 것으로 확인되었다. AAV 코딩 영역에 측접하는 DNA는, 대략 145개 뉴클레오티드 길이인 2개의 시스-작용 역 말단 반복부로, DNA 복제의 프라이머로서 기능하는 에너지적으로 안정한 헤어핀 구조로 접힐 수 있는 단속된 회문 서열을 갖는다. DNA 복제에서 이들의 역할에 더하여, ITR 서열은 바이러스 DNA의 세포 게놈 내로의 통합, 숙주 게놈 또는 플라스미드로부터 구제, 및 바이러스 핵산의 성숙한 비리온 내로의 캡슐화에 관여되는 것으로 나타났다(Muzyczka, (1992) Curr. Top. Micro. Immunol. 158:97-129). 일부 구현예에서, 하나 이상의 유전자 변형 핵산 구성요소는, 바이러스 패키징을 위해, AAV로부터 유래된 ITR에 의해 측접된다. 예를 들어, WO2019113310을 참조한다.

[1219] 일부 구현예에서, 유전자 변형 시스템의 하나 이상의 구성요소는 적어도 하나의 AAV 벡터를 통해 운반된다. 일부 구현예에서, 적어도 하나의 AAV 벡터는 특정 세포, 조직, 유기체에 대한 향성에 대해 선택된다. 일부 구현예에서, AAV 벡터는 슈도타입(pseudotype)화, 예를 들어 AAV2/8되며, 여기서 캡시드 단백질이 AAV8로부터의 것으로 대체된 것을 제외한 AAV2는 작제물의 설계를 설명한다. 기재된 임의의 벡터는 슈도타입 유도체일 수 있는 것으로 이해되며, 여기서 AAV 게놈을 패키징하기 위해 사용된 캡시드 단백질은 상이한 AAV 혈청형의 것으로부터 유래된다. 벡터 선택에 제한을 두고자 하지 않지만, 예시적인 AAV 혈청형의 목록은 표 18에서 찾을 수 있다. 일부 구현예에서, 유전자 변형을 위해 사용되는 AAV는, 문헌에서 예증되어 온 바와 같이, 새로운 세포 또는 조직 향성에 대해 진전될 수 있다(예를 들어, 문헌[Davidsson et al. Proc Natl Acad Sci USA 2019]).

[1220] 일부 구현예에서, AAV 전달 벡터는 2개의 AAV 역 말단 반복부(ITR) 및 관심 뉴클레오티드 서열(예를 들어, 유전자 변형 폴리펩티드를 코딩하는 서열 또는 DNA 주형, 또는 이들 모두)을 갖는 벡터로, 상기 ITR 각각은 단속된(또는 비인접) 회문 서열, 즉 3개의 절편: 5'→판독시에는 동일하지만, 서로에 대하여 반대로 위치할 때 혼성화되는 제1 절편 및 마지막 절편, 및 동일한 절편들을 구분하는 상이한 절편으로 구성된 서열을 갖는다. 예를 들어, WO2012123430을 참조한다.

[1221] 통상적으로, 캡시드가 있는 AAV 비리온은 rAAV 또는 scAAV 게놈, Rep 단백질, 및 Cap 단백질을 인코딩하는 플라스미드 또는 플라스미드들을 도입함으로써 생산된다(Grimm et al, 1998). 이들 헬퍼 플라스미드들의 인 트랜스(in trans) 도입시, AAV 게놈은 숙주 게놈으로부터 "구제"되고(즉, 방출 및 순차적으로 회수됨), 감염성 AAV를 생산하기 위해 추가로 단백질막으로 둘러싸인다(encapsidated). 일부 구현예에서, 하나 이상의 유전자 변형 핵산은 ITR-측접된 핵산을 헬퍼 기능과 함께 패키징 세포 내로 도입함으로써 AAV 입자 내로 패키징된다.

[1222] 일부 구현예에서, AAV 게놈은 소위 자기-상보적인 게놈으로도 명명되며(scAAV로도 지칭됨), ITR들 사이에 위치한 서열은, 원하는 핵산 서열의 역전 보체에 추가하여, 원하는 핵산 서열(예를 들어, 유전자 변형 폴리펩티드 또는 주형, 또는 둘 모두를 인코딩하는 DNA) 둘 모두를 함유하여, 이들 두 구성요소들이 접혀 자기-하이브리드 형성될 수 있도록 한다. 일부 구현예에서, 자기-상보적인 모듈은 DNA가 둘로 되접히는 것을 가능하게 하는 개재 서열에 의해 분리되어, 예를 들어 스템-루프를 형성한다. scAAV는 dsDNA를 형성하기 위해 ITR 프라이밍 및 제2 가닥 합성에 우선 의존하기보다는, 핵으로의 도입시에 전사에 대한 준비가 된 장점을 갖는다. 일부 구현예에서, 유전자 변형 구성요소들 중 하나 이상은 scAAV로서 설계되며, 여기서 AAV ITR 사이의 서열은 자기-혼성화하여 dsDNA를 생성할 수 있는 2개의 역 상보화 모듈을 함유한다.

[1223] 일부 구현예에서, 세포로 전달된(예를 들어, 폴리펩티드를 인코딩하는, 또는 주형, 또는 둘 모두) 핵산은 폐쇄된-말단, 선형 듀플렉스 DNA(CELiD DNA 또는 ceDNA)이다. 일부 구현예에서, ceDNA는 AAV 게놈의 복제 형태로부터 유래된다(Li et al. PLoS One 2013). 일부 구현예에서, 핵산(예를 들어, 폴리펩티드를 인코딩하는, 또는 주형 DNA, 또는 이들 모두)은 ITR, 예를 들어 AAV ITR에 의해 측접되며, 여기서 적어도 하나의 ITR은 말단 분해 부위 및 반복 단백질 결합 부위를 포함한다(때때로 복제 단백질 결합 부위로서도 지칭됨). 일부 구현예에서, ITR은 아데노-연관 바이러스, 예를 들어, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, 또는 이의 조합으로부터 유래된다. 일부 구현예에서, ITR은 대칭적이다. 일부 구현예에서, ITR은 비대칭적이다. 일부 구현예에서, 적어도 하나의 Rep 단백질은 작제물의 복제를 가능하게 하기 위해 제공된다.

일부 구현예에서, 적어도 하나의 Rep 단백질은 아데노-연관 바이러스, 예를 들어, AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, 또는 이의 조합으로부터 유래된다. 일부 구현예에서, ceDNA는 생산 세포에 (i) ITR, 예를 들어 AAV ITR에 의해 측접된 DNA, 및 (ii) ITR-의존성 복제에 필요한 구성 요소들, 예를 들어 AAV 단백질 Rep78 및 Rep52(또는 단백질을 인코딩하는 핵산)을 갖는 제공함으로써 생성된다. 일부 구현예에서, ceDNA는 어떤 캡시드 단백질도 없으며, 예를 들어 감염성 AAV 입자 내로 패키징되지 않는다. 일부 구현예에서, ceDNA는 LNP로 제형화된다(예를 들어, WO2019051289A1 참조).

[1224] 일부 구현예에서, ceDNA 벡터는 2개의 자가 상보적인 서열로 이루어지며, 예를 들어 비대칭 또는 대칭 또는 실질적으로 대칭인 여기 정의된 것과 같은 ITR, 상기 발현 카세트에 측접되고, 여기서 ceDNA 벡터는 캡시드 단백질과 회합되지 않는다. 일부 구현예에서, ceDNA 벡터는 AAV 게놈 내에서 발견되는 2개의 자가 상보적인 서열을 포함하며, 여기서 적어도 하나의 ITR은 작동적 Rep-결합 요소(RBE)(종종 여기서 "RBS"로 지칭되기도 함) 및 AAV의 말단 분해 부위(trs) 또는 RBE의 기능적 변이체를 포함한다. 예를 들어, WO2019113310을 참조한다.

[1225] 일부 구현예에서, AAV 게놈은 4개의 복제 단백질 및 3개의 캡시드 단백질을 인코딩하는 2개의 유전자를 각각 포함한다. 일부 구현예에서, 이 유전자들은 145-bp 역 말단 반복부(ITR)에 의해 어느 하나의 측면 상에서 측접된다. 일부 구현예에서, 비리온은 1:1:10 비로 생산된 3개 이하의 캡시드 단백질(Vp1, Vp2, 및/또는 Vp3)을 포함한다. 일부 구현예에서, 캡시드 단백질은 동일한 오픈 판독 프레임 및/또는 대체 스플라이싱(Vp1) 및 대안적인 번역 개시 부위(각각 Vp2 및 Vp3)로부터 생산된다. Vp3은 비리온에서 가장 풍부한 서브유닛이며, 바이러스의 향성을 규정하는 세포 표면에서 수용체 인식에 참여한다. 일부 구현예에서, Vp1은 포스포리파제 도메인, 예를 들어 Vp1의 N-말단에서 바이러스 감염성에의 기능을 포함한다.

[1226] 일부 구현예에서, 바이러스 벡터의 패키징 용량은 벡터 내로 패키징될 수 있는 유전자 변형 시스템의 크기를 제한한다. 예를 들어, AAV의 패키징 용량은 약 4.5 kb(예를 들어, 약 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 또는 6.0 kb)일 수 있으며, 예를 들어 하나 또는 2개의 역 말단 반복부(ITR), 예를 들어, 145 염기 ITR을 포함한다.

[1227] 일부 구현예에서, 재조합 AAV(rAAV)는 시스-작용 145-bp ITRs 측접하는 벡터 이식유전자 카세트를 포함하며, 이는 예를 들어 외래 DNA의 패키징을 위해 4.5 kb까지 제공한다. 감염에 이어, rAAV는 일부 경우에서, 본 발명의 융합 단백질을 발현하고, 환형 헤드-투-테일 콘카티머(concatemer)에서 에피좀적으로 존재함으로써 숙주 게놈 내로의 통합 없이 유지된다. rAAV는 예를 들어 시험관 내 및 생체 내에서 사용될 수 있다. 일부 구현예에서, AAV-매개 유전자 전달은 유전자의 코딩 서열의 길이가 야생형 AAV 게놈보다 동일하거나 더 클 것을 요구한다.

[1228] 이 크기를 초과하는 유전자의 AAV 전달 및/또는 거대한 생리학적 조절 요소의 사용은, 예를 들어 둘 이상의 단편으로 전달되도록 단백질(들)을 나눔으로써 달성될 수 있다. 일부 구현예에서, N-말단 단편은 인테인-N 서열에 융합된다. 일부 구현예에서, C-말단 단편은 인테인-C 서열에 융합된다. 구현예에서, 단편은 둘 이상의 AAV 벡터 내로 패키징된다.

[1229] 일부 구현예에서, 이중 AAV 벡터는 거대 이식유전자 발현 카세트를 2개의 별도의 절반(5 및 3 말단, 또는 헤드 및 테일)으로 분할시켜 제조되며, 예를 들어 여기서 카세트의 각각의 절반은 단일 AAV 벡터(<5 kb) 내로 패키징된다. 전장 이식유전자 발현 카세트의 재조립은, 이중 AAV 벡터 둘 모두에 의해 이후 일부 구현예에서 동일한 세포의 공동 감염시에 달성될 수 있다. 일부 구현예에서, 공동 감염은 다음 중 하나 이상에 따른다: (1) 5 및 3 게놈 사이의 상동성 재조합(HR)(이중 AAV 중복 벡터); (2) ITR-매개된 테일-투-헤드 콘카티머화(이중 AAV 트랜스-스플라이싱 벡터); 및/또는 (3) 이들 2개의 메커니즘의 조합(이중 AAV 하이브리드 벡터). 일부 구현예에서, 생체 내에서 이중 AAV 벡터의 사용은 전장 단백질의 발현을 생성한다. 일부 구현예에서, 이중 AAV 벡터 플랫폼의 이용은 크기가 약 4.0, 4.1, 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 또는 5.0 kb보다 더 큰 이식 유전자에 대한 효율적이고 독자생존가능한 유전자 전이 전략을 나타낸다. 일부 구현예에서, AAV 벡터는 예를 들어 핵산 및 펩티드의 시험관 내 생산에 있어서 표적 핵산을 이용하여 세포를 형질도입하는 데 사용될 수도 있다. 일부 구현예에서, AAV 벡터는 생체 내 및 생체 외 유전자 치료 절차에 사용될 수 있다(예를 들어, 문헌[West et al., Virology 160:38-47 (1987); 미국 특허 4,797,368; WO 93/24641; Kotin, Human Gene Therapy 5:793-801 (1994); Muzyczka, J. Clin. Invest.94:1351 (1994)] 참조; 이들 각각은 그 전체가 본 명세서에 참고로 포함됨). 재조합 AAV 벡터의 구축은 문헌[미국 특허 5,173,414; Tratschin et al., Mol. Cell. Biol.5:3251-3260 (1985); Tratschin, et al., Mol. Cell. Biol.4:2072-2081 (1984); Hermonat & Muzyczka, PNAS 81:6466-6470 (1984); 및 Samulski et al., J. Virol.63:03822-3828 (1989)](그 전체가 본 명세서에 참고로 포함됨)을 포함한 많은 간행물에 기재되어 있다.

[1230] 일부 구현예에서, 본 명세서에 기재된 유전자 변형 폴리펩티드(예를 들어, 하나 이상의 가이드 핵산 존재 또는

없음)는, AAV, 렌티바이러스, 아데노바이로스 또는 기타 플라스미드 또는 바이러스 벡터 유형을 이용하여, 특히 예를 들어 미국 특허 8,454,972(아데노바이러스용 제형, 용량), 미국 특허 8,404,658(AAV용 제형, 용량) 및 미국 특허 5,846,946(DNA 플라스미드용 제형, 용량), 및 렌티바이러스, AAV 및 아데노바이러스에 대한 임상 시험 및 간행물들로부터의 제형 및 용량을 사용하여 전달될 수 있다. 예를 들어, AAV의 경우, 투여 경로, 제형 및 용량은 미국 특허 8,454,972 및 AAV에 관한 임상 시험에서 기재된 것과 같을 수 있다. 아데노바이러스의 경우, 투여 경로, 제형 및 용량은 미국 특허 8,404,658 및 아데노바이러스에 관한 임상 시험에서 기재된 것과 같을 수 있다. 플라스미드 전달의 경우, 투여 경로, 제형 및 용량은 미국 특허 5,846,946 및 플라스미드를 포함하는 임상 연구에 기재된 것과 같을 수 있다. 용량은 평균 70 kg의 개인(예를 들어, 성인 남성)을 기준으로 하거나 외삽될 수 있으며, 환자, 대상, 상이한 체중의 포유동물 및 종에 대해 조절될 수 있다. 투여 빈도는 의약 또는 수의학 전문가(의사, 수의사)의 영역에 있으며, 환자 또는 대상의 연령, 성별, 일반 건강 상태 및 다루어지는 특정 상태 또는 증상을 포함한 일반적인 인자들에 따라 달라진다. 일부 구현예에서, 바이러스 벡터는 관심 조직 내로 주사될 수 있다. 세포 유형 특이적 유전자 변형에 대해, 유전자 변형 폴리펩티드 및 선택적인 가이드 핵산의 발현은, 일부 구현예에서 세포-유형 특이적 프로모터에 의해 구동될 수 있다.

[1231] 일부 구현예에서, AAV는, 예를 들어 면역 반응을 활성화할 수 있는 세포 입자들의 초원심분리를 필요로 하지 않는 정제 방법으로 인해 낮은 독성을 가능하게 한다. 일부 구현예에서, AAV는 예를 들어 숙주 게놈 내로 실질적으로 통합되지 않기 때문에, 삽입성 돌연변이 생성 유발의 낮은 가능성을 가능하게 한다.

[1232] 일부 구현예에서, AAV는 약 4.4, 4.5, 4.6, 4.7, 또는 4.75 kb의 패키징 한계를 갖는다. 일부 구현예에서, 유전자 변형 폴리펩티드-인코딩 서열, 프로모터, 및 전사 종결자는 단일 바이러스 벡터 내로 맞추어질 수 있다. SpCas9(4.1 kb)는 일부 경우에 AAV 내로 패키징하기 어려울 수 있다. 따라서, 일부 구현예에서, 다른 유전자 변형 폴리펩티드 코딩 서열 또는 염기 편집기보다 길이가 더 짧은 유전자 변형 폴리펩티드 코딩 서열이 사용된다. 일부 구현예에서, 유전자 변형 폴리펩티드 인코딩 서열은 약 4.5 kb, 4.4 kb, 4.3 kb, 4.2 kb, 4.1 kb, 4 kb, 3.9 kb, 3.8 kb, 3.7 kb, 3.6 kb, 3.5 kb, 3.4 kb, 3.3 kb, 3.2 kb, 3.1 kb, 3 kb, 2.9 kb, 2.8 kb, 2.7 kb, 2.6 kb, 2.5 kb, 2 kb, 또는 1.5 kb 미만이다.

[1233] AAV는 AAV1, AAV2, AAV5 또는 이들의 임의의 조합일 수 있다. 일부 구현예에서, AAV의 유형은 표적화되는 세포에 대하여 선택된다; 예를 들어, AAV 혈청형 1, 2, 5 또는 하이브리드 캡시드 AAV1, AAV2, AAV5 또는 이들의 임의의 조합은 뇌 또는 뉴런 세포를 표적화하기 위해 선택될 수 있거나; 또는 AAV4는 심장 조직을 표적화하기 위해 선택될 수 있다. 일부 구현예에서, AAV8은 간으로의 전달을 위해 선택된다. 이들 세포와 관련하여 예시적인 AAV 혈청형은 예를 들어 문헌[Grimm, D. et al, J. Virol.82: 5887-5911 (2008)]에 기재되어 있다(그 전체가 본 명세서에 참고로 포함됨). 일부 구현예에서, AAV는 혈청형, 하위유형, 및 자연 발생 AAV 및 재조합 AAV를 지칭한다. AAV는 바이러스 그 자체 또는 그의 유도체를 지칭하는 데 사용될 수 있다. 일부 구현예에서, AAV는 AAV1, AAV2, AAV3, AAV3B, AAV4, AAV5, AAV6, AAV6.2, AAV7, AAVrh.64R1, AAVhu.37, AAVrh.8, AAVrh.32.33, AAV8, AAV9, AAV-DJ, AAV2/8, AAVrh10, AAVLK03, AV10, AAV11, AAV 12, rh10, 및 이들의 하이브리드, 조류 AAV, 소 AAV, 개 AAV, 말 AAV, 영장류 AAV, 비영장류 AAV, 및 양 AAV를 포함한다. AAV의 다양한 혈청형의 게놈 서열 및 천연 말단 반복부(TR)의 서열, Rep 단백질, 및 캡시드 서브유닛이 당업계에 알려져 있다. 그러한 서열은 문헌 또는 GeneBank와 같은 공공 데이터베이스에서 찾을 수 있다. 추가의 예시적인 AAV 혈청형이 표 18에 열거되어 있다.

[1234] **표 18. 예시적인 AAV 혈청형.**

표적 조직	비히클	참조
간	AAV (AAV8[1], AAVrh.8[1], AAVhu.37[1], AAV2/8, AAV2/rh10[2], AAV9, AAV2, NP40[3], NP59[2,3], AAV3B[5], AAV-DJ[4], AAV-LK01[4], AAV-LK02[4], AAV-LK03[4], AAV-LK19[4], AAV5[7] 아데노바이러스 (Ad5, HC-AdV[6])	1. Wang *et al.*, *Mol. Ther.* **18**, 118–25 (2010) 2. Ginn *et al.*, *JHEP Reports*, 100065 (2019) 3. Paulk *et al.*, *Mol. Ther.* **26**, 289–303 (2018). 4. L. Lisowski *et al.*, *Nature*. **506**, 382–6 (2014). 5. L. Wang *et al.*, *Mol. Ther.* **23**, 1877–87 (2015). 6. Hausl *Mol Ther* (2010) 7. Davidoff *et al.*, *Mol. Ther.* 11, 875-88 (2005)
폐	AAV (AAV4, AAV5, AAV6[1], AAV9, H22[2]) 아데노바이러스 (Ad5, Ad3, Ad21, Ad14)[3]	1. Duncan *et al.*, *Mol Ther Methods Clin Dev* (2018) 2. Cooney *et al.*, Am J Respir Cell Mol Biol (2019) 3. Li *et al.*, *Mol Ther Methods Clin Dev* (2019)
피부	AAV (AAV6[1], AAV-LK19[2])	1. Petek et al., *Mol. Ther.* (2010) 2. L. Lisowski *et al.*, *Nature*. **506**, 382–6 (2014).
HSCs	아데노바이러스 (HDAd5/35^{++})	Wang et al. *Blood Adv* (2019)

[1235]

[1236] 일부 구현예에서, 약제학적 조성물(예를 들어, 본 명세서에 기재된 바와 같은 AAV를 포함하는)은 10% 미만의 중공 캡시드(empty capsid), 8% 미만의 중공 캡시드, 7% 미만의 중공 캡시드, 5% 미만의 중공 캡시드, 3% 미만의 중공 캡시드, 또는 1 % 미만의 중공 캡시드를 갖는다. 일부 구현예에서, 약제학적 조성물은 약 5% 미만의 중공 캡시드를 갖는다. 일부 구현예에서, 중공 캡시드의 수는 검출 한계 아래이다. 일부 구현예에서, 예를 들어 중공 캡시드는 예를 들어 적거나 실질적으로 치료 이익이 없으면서 불리한 반응(예를 들어, 면역 반응, 염증 반응, 간 반응 및/또는 심장 반응)을 생성할 수 있기 때문에, 약제학적 조성물은 적은 양의 중공 캡시드를 갖는 것이 유리하다.

[1237] 일부 구현예에서, 약제학적 조성물 중 잔류 숙주 세포 단백질(rHCP)은, 1 x 10^{13} vg/ml 당 100 ng/ml 이하의 rHCP, 예를 들어, 1 x 10^{13} vg/ml 당 40 ng/ml 이하의 rHCP 또는 1 x 10^{13} vg/ml 당 1 내지 50 ng/ml rHCP이다. 일부 구현예에서, 약제학적 조성물은 1 x 10^{13} vg 당 10 ng 미만의 rHCP 또는 1 x 10^{13} vg 당 5 ng 미만의 rHCP, 1 x 10^{13} vg 당 4 ng 미만의 rHCP, 또는 1 x 10^{13} vg 당 3 ng 미만의 rHCP, 또는 이들 사이의 임의의 농도를 포함한다. 일부 구현예에서, 약제학적 조성물 중 잔류 숙주 세포 DNA(hcDNA)는 1 x 10^{13} vg/ml 당 5 x 10^{6} pg/ml

이하의 hcDNA, 1 x 10^{13} vg/ml 당 1.2 x 10^{6} pg/ml 이하의 hcDNA, 또는 1 x 10^{13} vg/ml당 1x 10^{5} pg/ml hcDNA이다. 일부 구현예에서, 상기 약제학적 조성물 중 잔류 숙주 세포는 DNA는 1 x 10^{13} vg 당 5.0 x 10^{5} pg 미만, 1 x 10^{13} vg 당 2.0 x 10^{5} pg 미만, 1 x 10^{13} vg 당 1.1 x 10^{5} pg 미만, 1 x 10^{13} vg 당 1.0 x 10^{5} pg 미만의 hcDNA, 1 x 10^{13} vg 당 0.9 x 10^{5} pg 미만의 hcDNA, 1 x 10^{13} vg 당 0.8 x 10^{5} pg 미만의 hcDNA, 또는 이들 사이의 임의의 농도이다.

[1238] 일부 구현예에서, 약제학적 조성물 중 잔류 플라스미드 DNA는 1.0 x 10^{13} vg/ml 당 1.7 x 10^{5} pg/ml 이하, 또는 1 x 1.0 x 10^{13} vg/ml 당 1 x 10^{5} pg/ml, 또는 1.0 x 10^{13} vg/ml 당 1.7 x 10^{6} pg/ml이다. 일부 구현예에서, 약제학적 조성물 중 잔류 DNA 플라스미드는 1.0 x 10^{13} vg을 기준으로 10.0 x 10^{5} pg 미만, 1.0 x 10^{13} vg을 기준으로 8.0 x 10^{5} pg 미만 또는 1.0 x 10^{13} vg을 기준으로 6.8 x 10^{5} pg 미만이다. 구현예에서, 약제학적 조성물은 1.0 x 10^{13} vg 당 0.5 ng 미만, 1.0 x 10^{13} vg 당 0.3 ng 미만, 1.0 x 10^{13} vg 당 0.22 ng 미만 또는 1.0 x 10^{13} vg 당 0.2 ng 미만 또는 임의의 중간 농도의 소 혈청 알부민(BSA)을 포함한다. 구현예에서, 약제학적 조성물 중 벤조나제는 1.0 x 10^{13} vg을 기준으로 0.2 ng 미만, 1.0 x 10^{13} vg을 기준으로 0.1 ng 미만, 1.0 x 10^{13} vg을 기준으로 0.09 ng 미만, 1.0 x 10^{13} vg을 기준으로 0.08 ng 미만 또는 임의의 중간 농도를 포함한다. 구현예에서, 약제학적 조성물 중 폴록사머 188은 약 10 내지 150 ppm, 약 15 내지 100 ppm 또는 약 20 내지 80 ppm이다. 구현예에서, 약제학적 조성물 중 세슘은 50 pg/g(ppm) 미만, 30 pg / g(ppm) 미만 또는 20 pg / g(ppm) 미만 또는 임의의 중간 농도이다.

[1239] 구현예에서, 약학 조성물은 예를 들어 SDS-PAGE에 의해 결정 시, 10% 미만, 8% 미만, 7% 미만, 6% 미만, 5% 미만, 4% 미만, 3% 미만, 2% 미만 또는 이들 사이의 임의의 백분율의 총 불순물을 포함한다. 구현예에서, 예를 들어 SDS-PAGE에 의해 결정 시, 총 순도는 90% 초과, 92% 초과, 93% 초과, 94% 초과, 95% 초과, 96% 초과, 97% 초과, 98% 초과 또는 이들 사이의 임의의 농도이다. 구현예에서, 예를 들어 SDS-PAGE에 의해 측정될 때, 하나가 아닌 명명되지 않은 관련 불순물은 5% 초과, 4% 초과, 3% 초과, 2% 초과, 또는 이들 사이의 임의의 백분율이다. 구현예에서, 약학 조성물은 총 캡시드에 대하여 충전된 캡시드의 백분율로(예를 들어 분석용 초원심분리에 의해 측정 시 피크 1 + 피크 2) 85% 초과, 86% 초과, 87% 초과, 88% 초과, 89% 초과, 90% 초과, 91% 초과, 91.9% 초과, 92% 초과, 93% 초과, 또는 이들 사이의 임의의 백분율이다. 약학 조성물의 구현예에서, 분석용 초원심분리에 의해 피크 1로 측정된 충전된 캡시드의 백분율은 20 내지 80%, 25 내지 75%, 30 내지 75%, 35 내지 75%, 또는 37.4 내지 70.3%이다. 약학 조성물의 구현예들에서, 분석용 초원심분리에 의해 피크 2로 측정된 충전된 캡시드의 백분율은 20 내지 80%, 20 내지 70%, 22 내지 65%, 24 내지 62%, 또는 24.9 내지 60.1%다.

[1240] 일 구현예에서, 약학 조성물은 1.0 내지 5.0 x 10^{13} vg/mL, 1.2 내지 3.0 x 10^{13} vg/mL 또는 1.7 내지 2.3 x 10^{13} vg/ml의 유전체 역가를 포함한다. 일 구현예에서, 약학 조성물은 5 CFU/mL 미만, 4 CFU/mL 미만, 3 CFU/mL 미만, 2 CFU/mL 미만 또는 1 CFU/mL 미만 또는 임의의 중간 생략산의 생물학적 부하를 나타낸다. 구현예에서, USP, 예를 들어 USP <85>(그 전체가 참고로서 포함됨)에 따른 내독소의 양은 1.0 EU/mL 미만, 0.8 EU/mL 미만, 또는 0.75 EU/mL 미만이다. 구현예에서, USP, 예를 들어 USP <785>(그 전체가 참고로서 포함됨)에 따른 약학 조성물의 삼투압농도는 350 내지 450 mOsm/kg, 370 내지 440 mOsm/kg 또는 390 내지 430 mOsm/kg이다. 구현예에서, 약학 조성물은 용기 당 25 ㎛ 초과인 입자 1200개 미만, 용기 당 25 ㎛ 초과인 입자 1000개 미만, 용기 당 25 ㎛ 초과인 입자 500개 미만, 또는 임의의 중간체 값을 함유한다. 구현예에서, 약학 조성물은 용기 당 10 ㎛ 초과인 입자 10,000개 미만, 용기 당 10 ㎛ 초과인 입자 8000개 미만, 용기 당 10 pm 초과인 입자 600개 미만을 함유한다.

[1241] 일 구현예에서, 약학 조성물은 0.5 내지 5.0 x 10^{13} vg/mL, 1.0 내지 4.0 x 10^{13} vg / mL, 1.5 내지 3.0 x 10^{13} vg/ml 또는 1.7 내지 2.3 x 10^{13} vg/ml의 유전체 역가를 갖는다. 일 구현예에서, 본 명세서에 기재된 약학 조성물은 다음 중 하나 이상을 포함한다: 1.0 x 10^{13} vg 당 약 0.09 ng 미만의 벤조나제, 약 30 pg/g(ppm) 미만의 세슘, 약 20 내지 80 ppm의 폴록사머 188, 1.0 x 10^{13} vg당 약 0.22 ng 미만의 BSA, 1.0 x 10^{13} vg 당 약 6.8 x 10^{5} pg 미만의 잔류 DNA 플라스미드, 1.0 x 10^{13} vg 당 약 1.1 x 10^{5} pg 미만의 잔류 hcDNA, 1.0 x 10^{13} vg 당

약 4 ng 미만의 rHCP, pH 7.7 내지 8.3, 약 390 내지 430 mOsm/kg, 용기 당 > 25 ㎛ 크기의 입자 약 600개 미만, 용기 당 > 10 ㎛ 크기의 입자 약 6000개 미만, 약 1.7 x 10^{13} 내지 2.3 x 10^{13} vg/mL의 유전체 역가, 1.0 x 10^{13} vg 당 약 3.9 x 10^{8} 내지 8.4 x 10^{10} IU의 감염 역가, 1.0 x 10^{13} vg당 약 100 내지 300 pg의 총 단백질, A7SMA 마우스에서 7.5 x 10^{13} vg/kg 용량의 바이러스 벡터를 이용하여 >24일의 평균 생존, 시험관 내 세포 기반 분석을 기준으로 약 70 내지 130%의 상대 및/또는 약 5% 의 중공 캡시드. 다양한 구현예에서, 본 명세서에 기재된 약학 조성물은 여기 논의된 임의의 바이러스 입자를 포함하고, 참조 표준의 ±20% 사이, ±15% 사이, ±10% 사이 또는 ±5% 내의 역가를 유지한다. 일부 구현예에서, 역가는 적합한 시험관 내 세포 분석 또는 생체 내 동물 모델을 사용하여 측정된다.

[1242] AAV 입자의 제조, 특성화 및 투약하는 추가의 방법은, 그 전체가 본 명세서에 참고로 포함된 WO2019094253에 교시되어 있다.

[1243] 본 발명과 합치되는 이용될 수 있는 추가의 rAAV 작제물은 문헌[Wang et al 2019]에 기재되어 있는 것들을 포함하며, 이는 그의 표 1을 포함하여, 그 전체가 본 명세서에 참고로 포함된 //doi.org/10.1038/s41573-019-0012-9에서 입수가능하다.

[1244] *지질 나노입자*

[1245] 본 명세서에 제공된 방법 및 시스템은, 임의의 적합한 담체 또는 전달 양식을 이용할 수 있으며, 이는 특정 구현예에서 지질 나노입자(LNP)를 포함한다. 지질 나노입자는, 일부 구현예에서 하나 이상의 이온성 지질, 예컨대 하나 이상의 이온성 지질, 예컨대 비-양이온성 지질(예를 들어, 중성 또는 음이온, 또는 쯔비터이온성 지질); 하나 이상의 접합된 지질(예컨대 PEG-접합된 지질 또는 그 전체가 본 명세서에 참고로 포함된 WO2019217941의 표 5에 기재된 중합체에 접합된 지질); 하나 이상의 스테롤(예를 들어, 콜레스테롤); 및, 선택적으로, 하나 이상의 표적화 분자(예를 들어, 접합된 수용체, 수용체 리간드, 항체); 또는 상기의 조합을 포함한다.

[1246] 나노입자 형성에서 사용될 수 있는 지질(예를 들어, 지질 나노입자)은, 예를 들어, 잠고문헌으로 포함된 WO2019217941의 표 4에 기재된 것들을 포함하며 - 예를 들어 지질 함유 나노입자는 WO2019217941의 표 4의 하나 이상의 지질을 포함한다. 지질 나노입자는 중합체, 예컨대 참고로서 포함된 WO2019217941의 표 5에 기재된 중합체와 같은 추가의 요소들을 포함할 수 있다.

[1247] 일부 구현예에서, 존재하는 경우, 접합된 지질은 PEG-디아실글리세롤(DAG)(예컨대, 1-(모노메톡시-폴리에틸렌글리콜)-2,3-디미리스토일글리세롤(PEG-DMG)), PEG-디알킬옥시프로필(DAA), PEG-인지질, PEG-세라마이드(Cer), 페길화 포스파티딜에탄올아민(PEG-PE), PEG 숙시네이트 디아실글리세롤(PEGS-DAG)(예컨대 4-O-(2',3'-디(테트라데카노일옥시)프로필-1-O-(w-메톡시(폴리에톡시)에틸) 부탄디오에이트(PEG-S-DMG)), PEG 디알콕시프로필카르밤, N- (카르보닐-메톡시폴리에틸렌글리콜 2000)-1,2-디스테아로일-sn-글리세로-3-포스포에탄올아민 나트륨염, WO2019051289(참고로 포함됨)의 표 2에 기재된 것들, 및 상기의 조합들 중 하나 이상을 포함할 수 있다.

[1248] 일부 구현예에서, 지질 나노입자 내에 포함될 수 있는 스테롤은 하나 이상의 콜레스테롤 또는 콜레스테롤 유도체, 예컨대 참고로서 포함된 WO2009/127060 또는 US2010/0130588에 있는 것들을 포함한다. 추가의 예시적인 스테롤은 파이토스테롤을 포함하며, 본 명세서에 참고로 포함된 문헌[Eygeris et al (2020), dx.doi.org/10.1021/acs.nanolett.0c01386]에 기재된 것들을 포함한다.

[1249] 일부 구현예에서, 지질 입자는 이온화 지질, 비-양이온성 지질, 입자들의 집합을 저해하는 접합된 지질, 및 스테롤을 포함한다. 이들 구성요소들의 양은 독립적으로, 그리고 원하는 특성을 달성하기 위해 변화될 수 있다. 예를 들어, 일부 구현예에서, 지질 나노입자는 이온화 지질을 총 지질의 약 20 몰% 내지 약 90 몰%(다른 구현예에서 이는 지질 나노입자에 존재하는 총 지질의 20 내지 70%(몰), 30 내지 60%(몰) 또는 40 내지 50%(몰); 약 50 몰% 내지 약 90 몰%), 비-양이온성 지질을 총 지질의 약 5 몰% 내지 약 30 몰%, 접합된 지질을 총 지질의 약 0.5 몰% 내지 약 20 몰%의 양으로, 및 스테롤을 총 지질의 약 20 몰% 내지 약 50 몰%의 양으로 포함한다. 총 지질 대 핵산(예를 들어, 유전자 변형 폴리펩티드를 인코딩하는 핵산 또는 주형 핵산)의 비는 원하는 바에 따라 변화될 수 있다. 예를 들어, 총 지질 대 핵산(질량 또는 중량) 비는 약 10:1 내지 약 30:1이다.

[1250] 일부 구현예에서, 이온화 지질은 양이온성 지질, 이온화성 양이온성 지질, 예를 들어 pH에 따라 양의 하전된 형태 또는 중성 형태로 존재할 수 있는 양이온성 지질, 또는 쉽게 양자화될 수 있는 아민-함유 지질일 수 있다. 일부 구현예에서, 양이온성 지질은, 예를 들어 생리학적 조건 하에서 양으로 하전될 수 있는 지질이다. 예시적인 양이온성 지질은 양의 하전을 갖는 하나 이상의 아민 기(들)을 포함한다. 일부 구현예에서, 지질 입자는 제

형 내에서 하나 이상의 중성 지질, 이온화성 아민-함유 지질, 생물분해성 알킨 지질, 스테로이드, 다중불포화 지질을 포함하는 인지질, 구조 지질(예를 들어 스테롤), PEG, 콜레스테롤 및 중합체 접합된 지질과 함께 양이온성 지질을 포함한다. 일부 구현예에서, 양이온성 지질은 이온화 양이온성 지질일 수 있다. 본 명세서에 개시된 바와 같은 예시적인 양이온성 지질은 6.0에 걸쳐 효과적인 pKa를 가질 수 있다. 구현예에서, 지질 나노입자는, 제1 양이온성 지질보다는, (예를 들어, 제1의 유효 pKa보다 더 큰) 상이한 유효한 pKa를 갖는 제2의 양이온성 지질을 포함할 수 있다. 지질 나노입자는 40 내지 60 몰%의 양이온성 지질, 중성 지질, 스테로이드, 중합체 접합된 지질, 및 치료제, 예를 들어 본 명세서에 기재된 핵산(예를 들어, RNA)(예를 들어, 주형 핵산 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산)을, 지질 나노입자 내에 캡슐화되거나 또는 이와 회합되어 포함할 수 있다. 일부 구현예에서, 핵산은 양이온성 지질과 공동 제형화된다. 핵산은 LNP, 예를 들어 양이온성 지질을 포함하는 LNP의 표현에 흡착될 수 있다. 일부 구현예에서, 핵산은 LNP, 예를 들어 양이온성 지질을 포함하는 LNP 내에 캡슐화될 수 있다. 일부 구현예에서, 지질 나노입자는 예를 들어 표적화제로 코팅된 표적화 모이어티를 포함할 수 있다. 구현예에서, LNP 제형은 생물분해성이다. 일부 구현예에서, 본 명세서에 기재된 하나 이상의 지질을 포함하는 지질 나노입자, 예를 들어 화학식 i, 화학식 ii, 화학식 ii, 화학식 vii 및/또는 화학식 ix는 적어도 1%, 적어도 5%, 적어도 10%, 적어도 20%, 적어도 30%, 적어도 40%, 적어도 50%, 적어도 60%, 적어도 70%, 적어도 80%, 적어도 90%, 적어도 92%, 적어도 95%, 적어도 97%, 적어도 98% 또는 100%의 RNA 분자, 예를 들어 주형 RNA 및/또는 유전자 변형 폴리펩티드를 인코딩하는 mRNA를 캡슐화한다.

[1251] 일부 구현예에서, 지질 대 핵산 비(질량/질량 비; w/w 비)는 약 1:1 내지 약 25:1, 약 10:1 내지 약 14:1, 약 3:1 내지 약 15: 1, 약 4:1 내지 약 10:1, 약 5:1 내지 약 9:1, 또는 약 6:1 내지 약 9:1의 범위일 수 있다. 지질 및 핵산의 양은 원하는 N/P 비, 예를 들어 3, 4, 5, 6, 7, 8, 9, 10 이상의 N/P 비를 제공하도록 조절될 수 있다. 일반적으로, 지질 나노입자 제형의 전체 지질 함량은 약 5 mg/ml 내지 약 30 mg/mL의 범위일 수 있다.

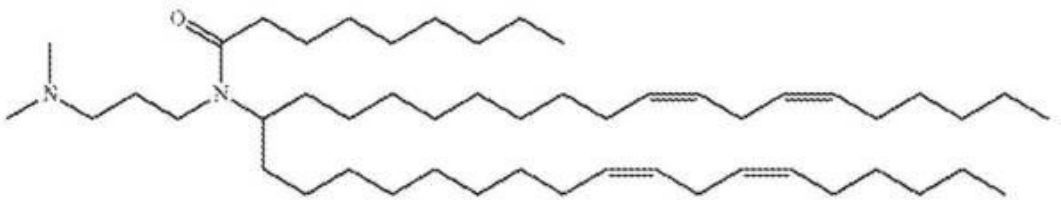
[1252] 지질 입자 제형에서 사용될 수 있는 예시적인 이온화 지질은, 제한 없이 본 명세서에 참조로 포함된, WO2019051289의 표 1에 열거된 것들을 포함한다. 추가의 예시적인 지질은, 제한 없이 다음 화학식 중 하나 이상을 포함한다: US2016/0311759의 X; US20150376115 또는 US2016/0376224의 I; US20160151284의 I, II 또는 III; US20170210967의 I, IA, II, 또는 IIA; US20150140070의 I-c; US2013/0178541의 A; US2013/0303587 또는 US2013/0123338의 I; US2015/0141678의 I; US2015/0239926의 II, III, IV, 또는 V; US2017/0119904의 I; WO2017/117528의 I 또는 II; US2012/0149894의 A; US2015/0057373의 A; WO2013/116126의 A; US2013/0090372의 A; US2013/0274523의 A; US2013/0274504의 A; US2013/0053572의 A; WO2013/016058의 A; WO2012/162210의 A; US2008/042973의 I; US2012/01287670의 I, II, III, 또는 IV; US2014/0200257의 I 또는 II; US2015/0203446의 I, II, 또는 III; US2015/0005363의 I 또는 III; US2014/0308304의 I, IA, IB, IC, ID, II, IIA, IIB, IIC, IID, 또는 III-XXIV; US2013/0338210의; WO2009/132131의 I, II, III, 또는 IV; US2012/01011478의 A; US2012/0027796의 I 또는 XXXV; US2012/0058144의 XIV 또는 XVII; US2013/0323269의; US2011/0117125의 I; US2011/0256175의 I, II, 또는 III; US2012/0202871의 I, II, III, IV, V, VI, VII, VIII, IX, X, XI, XII; US2011/0076335의 I, II, III, IV, V, VI, VII, VIII, X, XII, XIII, XIV, XV, 또는 XVI; US2006/008378의 I 또는 II; US2013/0123338의 I; US2015/0064242의 I 또는 X-A-Y-Z; US2013/0022649의 XVI, XVII, 또는 XVIII; US2013/0116307의 I, II, 또는 III; US2013/0116307의 I, II, 또는 III; US2010/0062967의 I 또는 II; US2013/0189351의 I-X; US2014/0039032의 I; US2018/0028664의 V; US2016/0317458의 I; US2013/0195920의 I; US10,221,127의 5, 6, 또는 10; WO2018/081480의 III-3; WO2020/081938의 I-5 또는 I-8; US9,867,888의 18 또는 25; US2019/0136231의 A; WO2020/219876의 II; US2012/0027803의 1; US2019/0240349의 OF-02; US10,086,013의 23; 문헌 [Miao et al (2020)]의 cKK-E12/A6; WO2010/053572의 C12-200; 문헌 [Dahlman et al (2017)]의 7C1; 문헌 [Whitehead et al]의 304-013 또는 503-013; US9,708,628의 TS-P4C2; WO2020/106946의 I; WO2020/106946의 I.

[1253] 일부 구현예에서, 이온화 지질은 MC3(6Z,9Z,28Z,3 1Z)-헵타트리아콘타-6,9,28,3 1-테트라엔-19-일-4-(디메틸아미노) 부타노에이트(DLin-MC3-DMA 또는 MC3), 예를 들어 이는 WO2019051289A9(그 전체가 본 명세서에 참고로 포함됨)의 실시예 9에 기재된 바와 같다. 일부 구현예에서, 이온화 지질은 예를 들어 WO2019051289A9(그 전체가 본 명세서에 참고로 포함됨)의 실시예 10에 기재된 바와 같은 지질 ATX-002이다. 일부 구현예에서, 이온화 지질은 (13Z,16Z)-A,A-디메틸-3- 노닐도코사-13, 16-디엔-1-아민(화합물 32)로, 이는 예를 들어 WO2019051289A9(그 전체가 본 명세서에 참고로 포함됨)의 실시예 11에 기재된 바와 같다. 일부 구현예에서, 이온화 지질은 화합물 6 또는 화합물 22로, 이는 예를 들어 WO2019051289A9(그 전체가 본 명세서에 참고로 포함됨)의 실시예12에 기재된 바와 같다. 일부 구현예에서, 이온화 지질은 헵타데칸-9-일 8-((2-히드록시에틸)(6-옥소-6-(운데실옥시)헥실)아미노)옥타노에이트(SM-102); 예를 들어, US9,867,888(그 전

체가 본 명세서에 참고로 포함됨)의 실시예 1에 기재된 바와 같다. 일부 구현예에서, 이온화 지질은 9Z,12Z)-3-((4,4-비스(옥틸옥시)부타노일)옥시)-2-((((3-(디에틸아미노)프로폭시)카르보닐)옥시)메틸)프로필 옥타데카-9,12-디에노에이트(LP01)이며, 예를 들어 WO2015/095340(그 전체가 본 명세서에 참고로 포함됨)의 실시예 13에서 합성되는 바와 같다. 일부 구현예에서, 이온화 지질은 디((Z)-비-2-엔-1-일) 9-((4-디메틸아미노)부타노일)옥시)헵타데칸디오에이트(L319)며, 예를 들어 US2012/0027803(그 전체가 본 명세서에 참고로 포함됨)의 실시예 7, 8 또는 9에서 합성되는 바와 같다. 일부 구현예에서, 이온화 지질은 1,1'-((2-(4-(2-((2-(비스(2-히드록시도데실)아미노)에틸)(2-히드록시도데실)아미노)에틸)피페라진-1-일)에틸)아잔디일)비스(도데칸-2-올)(C12-200)며, 예를 들어 WO2010/053572(그 전체가 본 명세서에 참고로 포함됨)의 실시예 14 및 16에서 합성되는 바와 같다. 일부 구현예에서, 이온화 지질은; 이미다졸 콜레스테롤 에스테르(ICE) 지질(3S, 10R, 13R, 17R)-10, 13-디메틸-17- ((R)-6-메틸헵탄-2-일)-2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17-테트라데카하이드로-1H-시클로펜타[a]페난트렌-3-일 3-(1H-이미다졸-4-일)프로파노에이트, 예를 들어 WO2020/106946(그 전체가 본 명세서에 참고로 포함됨)의 구조식 (I)이다.

[1254] 본 명세서에 기재된 조성물의 전달을 위해 지질 나노입자를 형성하기 위해(예를 들어, 다른 지질 성분과 조합되어) 사용될 수 있는 지질 화합물의 일부 비제한적인 예, 예를 들어 본 명세서에 기재된 핵산(예를 들어, RNA)(예를 들어, 주형 핵산 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산)은 다음을 포함한다:

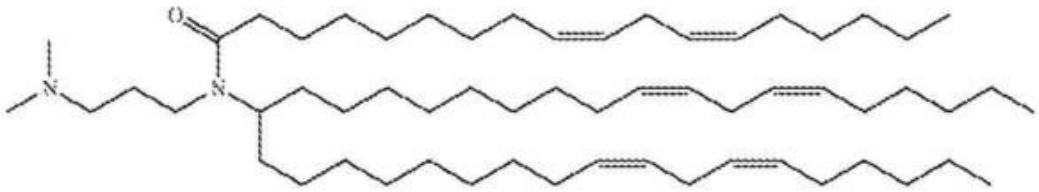
[1255] [화학식 i]



[1256]

[1257] 일부 구현예에서, 화학식 i를 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 간 및/또는 간세포에 전달하는 데 사용된다.

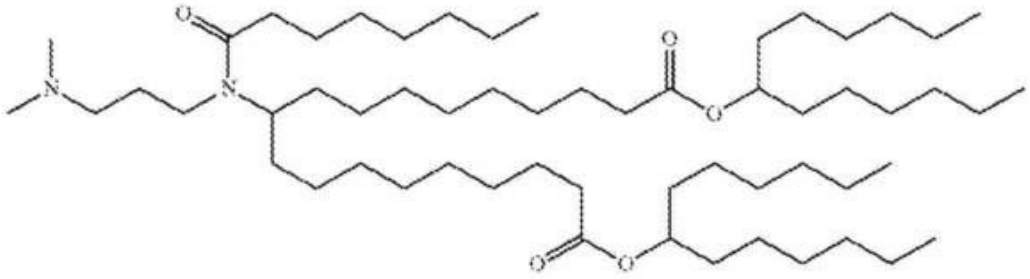
[1258] [화학식 ii]



[1259]

[1260] 일부 구현예에서, 화학식 ii를 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 간 및/또는 간세포에 전달하는 데 사용된다.

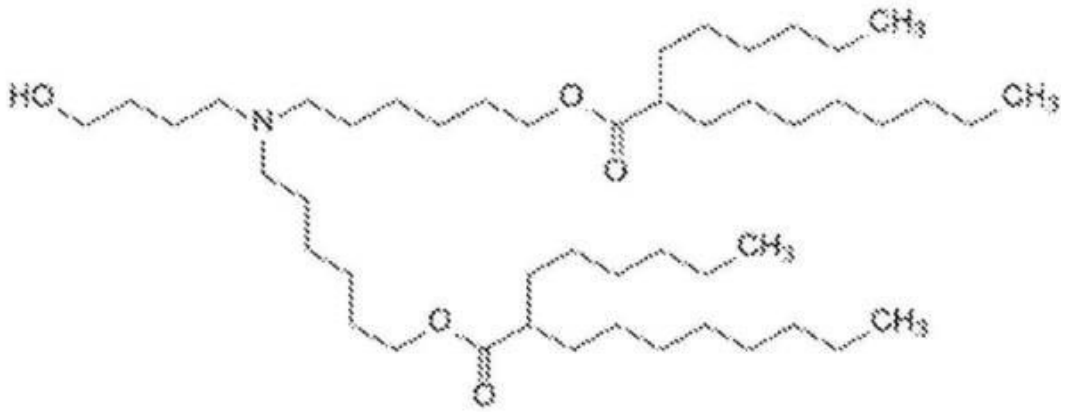
[1261] [화학식 iii]



[1262]

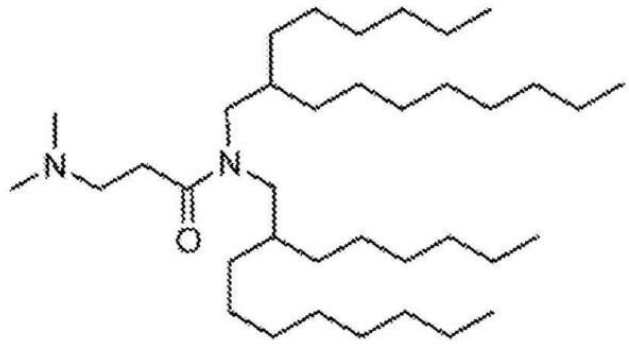
[1263] 일부 구현예에서, 화학식 iii을 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 간 및/또는 간세포에 전달하는 데 사용된다.

[1264] [화학식 iv]



[1265]

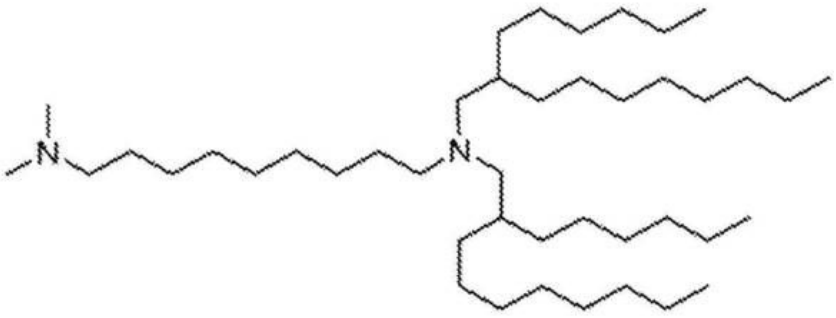
[1266] [화학식 v]



[1267]

[1268] 일부 구현예에서, 화학식 v를 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 간 및/또는 간세포에 전달하는 데 사용된다.

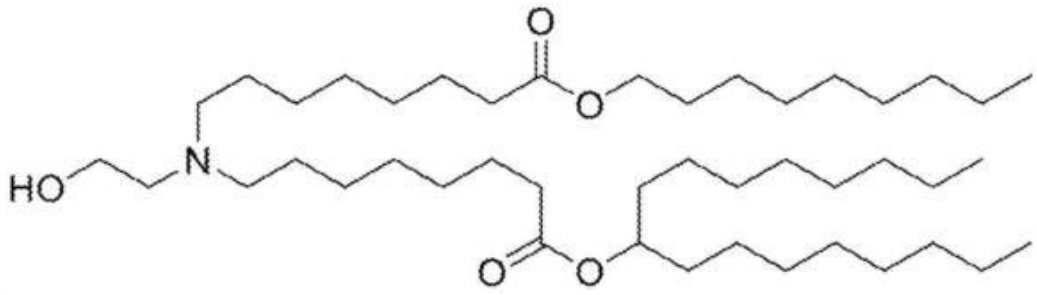
[1269] [화학식 vi]



[1270]

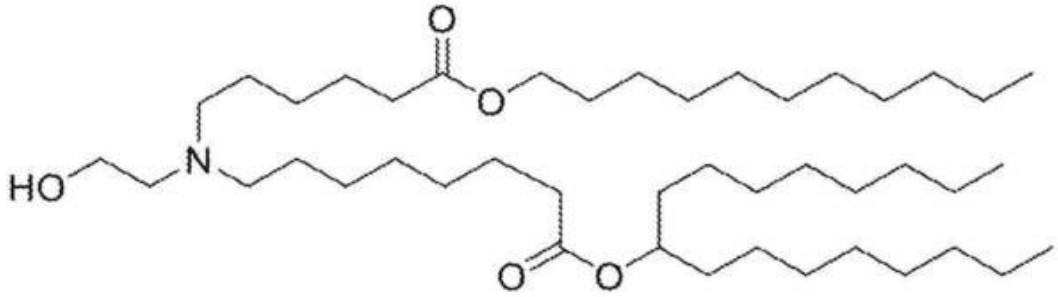
[1271] 일부 구현예에서, 화학식 vi을 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 간 및/또는 간세포에 전달하는 데 사용된다.

[1272] [화학식 vii]



[1273]

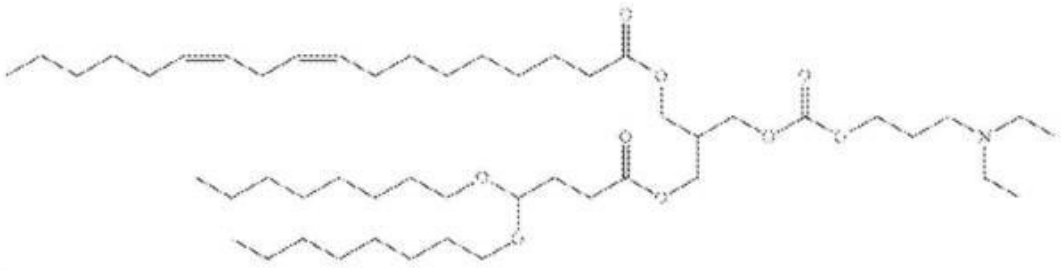
[1274] [화학식 viii]



[1275]

[1276] 일부 구현예에서, 화학식 viii을 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 간 및/또는 간세포에 전달하는 데 사용된다.

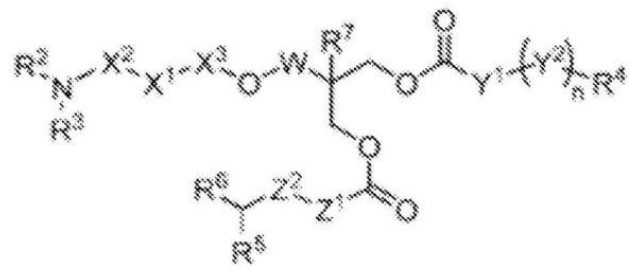
[1277] [화학식 ix]



[1278]

[1279] 일부 구현예에서, 화학식 ix을 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 간 및/또는 간세포에 전달하는 데 사용된다.

[1280] [화학식 x]

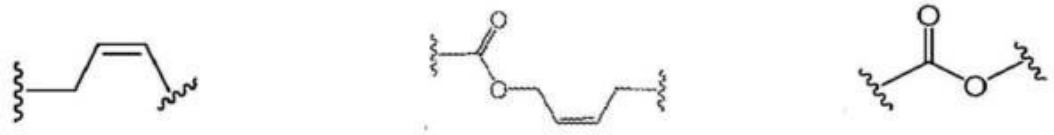


[1281]

[1282] 여기서

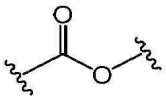
[1283] X^1는 O, NR^1, 또는 직접 결합이고, X^2는 C2-5 알킬렌이고, X^3은 C(=O) 또는 직접 결합이고, R^1은 H 또는 Me이고, R^3은 Ci-3 알킬이고, R^2는 Ci-3 알킬이거나, R^2는 이것이 부착된 질소 원자 및 X^2의 1 내지 3개의 탄소 원자와

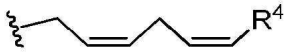
함께 4-, 5-, 또는 6-원 고리를 형성하거나, 또는 X^1은 NR^1이고, R^1 및 R^2는 이들이 부착된 질소 원자와 함께 5- 또는 6-원 고리를 형성하거나, 또는 R^2는 R^3 및 이들이 부착된 질소 원자와 함께 5-, 6-, 또는 7-원 고리를 형성하고, Y^1은 C2-12 알킬렌이고, Y^2는 다음에서 선택되고,

[1284] 

[1285] (어느 하나의 배향), (어느 하나의 배향), (어느 하나의 배향),

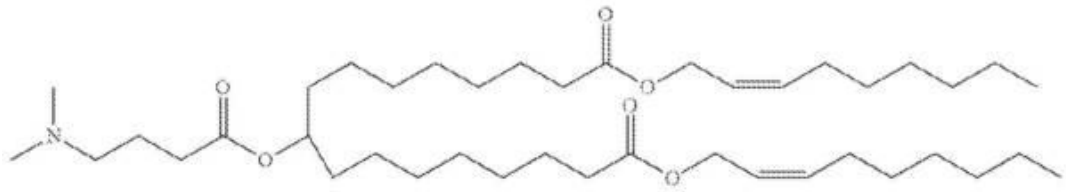
[1286] n은 0 내지 3이고, R^4는 Ci-15 알킬이고, Z^1은 Ci-6 알킬렌 또는 직접 결합이고,

[1287] Z는  (어느 하나의 배향) 또는 부재하며, 단, Z^1이 직접 결합인 경우, Z^2는 부재하고;

[1288] R^5는 C5-9 알킬 또는 C6-10 알콕시이고, R^6은 C5-9 알킬 또는 C6-10 알콕시이고, W는 메틸렌 또는 직접 결합이며, R^7은 H 또는 이의 염, 단, R^3 및 R^2는 C2 알킬이고, X^1은 O이고, X^2는 선형 C3 알킬렌이고, X^3은 C(=O)이고, Y^1은 선형 Ce 알킬렌이고, $(Y^2)n-R^4$는  이고, R^4는 선형 C5 알킬이고, Z^1은 C2 알킬렌이고, Z^2는 부재하고, W는 메틸렌이고, R^7은 H이며, R^5 및 R^6은 Cx 알콕시가 아니다.

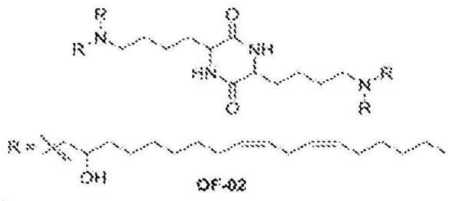
[1289] 일부 구현예에서, 화학식 xii을 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 간 및/또는 간세포에 전달하는 데 사용된다.

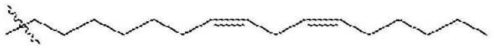
[1290] [화학식 xi]

[1291] 

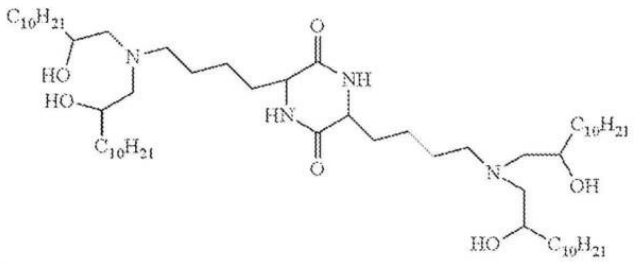
[1292] 일부 구현예에서, 화학식 xi를 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 간 및/또는 간세포에 전달하는 데 사용된다.

[1293] [화학식 xii]

[1294] 

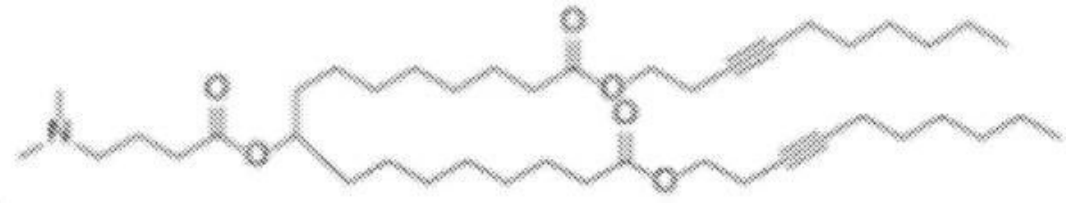
[1295] (여기서, R= )

[1296] [화학식 xiii]



[1297]

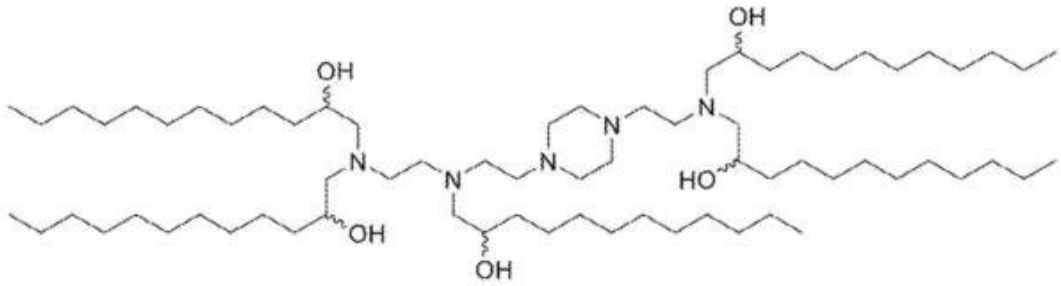
[1298] [화학식 xiv]



[1299]

[1300] 일부 구현예에서 LNP는 화학식 xiii의 화합물 및 화학식 xiv의 화합물을 포함한다.

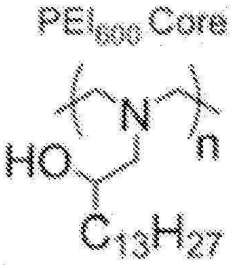
[1301] [화학식 xv]



[1302]

[1303] 일부 구현예에서, 화학식 xv를 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 간 및/또는 간세포에 전달하는 데 사용된다.

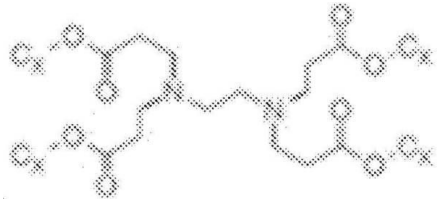
[1304] [화학식 xvi]



[1305]

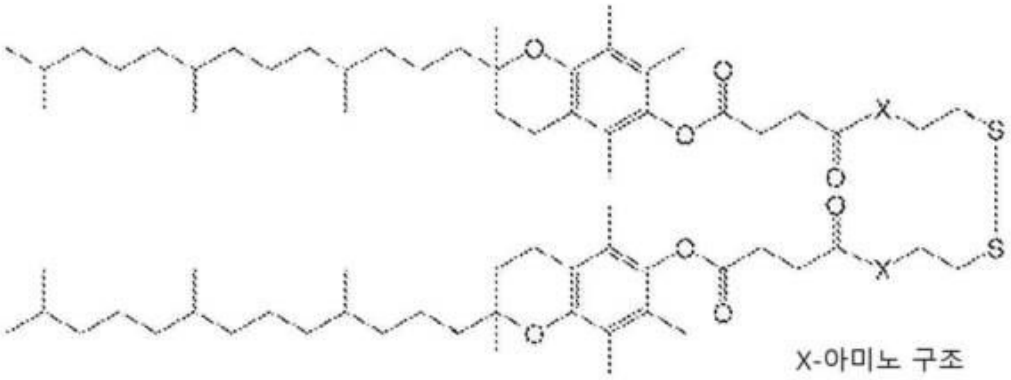
[1306] 일부 구현예에서, 화학식 xvi의 제형을 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 폐 내피 세포로 전달하는 데 사용된다.

[1307] [화학식 xvii]



[1308]

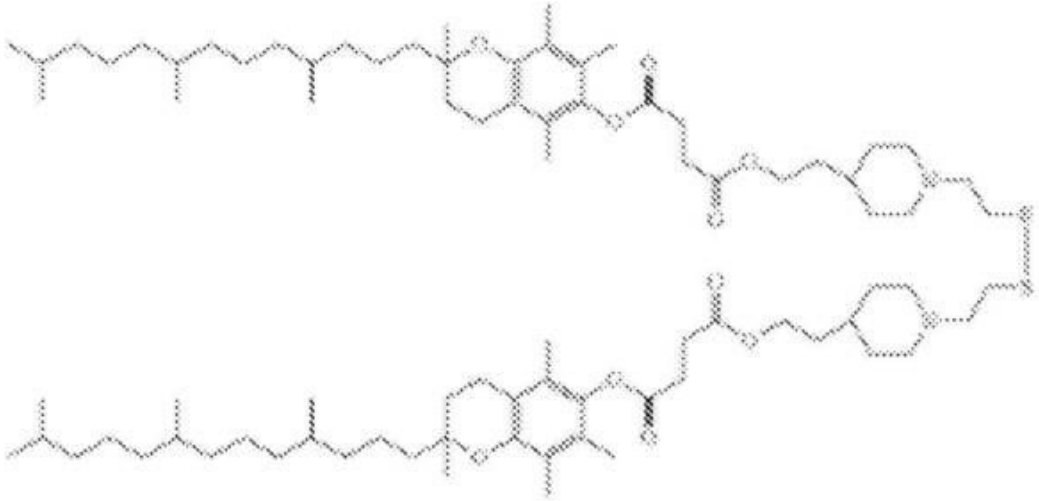
[1309] [화학식 xviii(a)]



[1310]

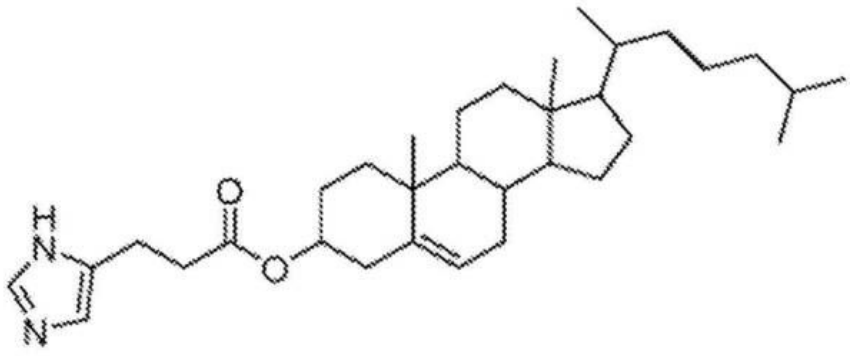
[1311] (여기서, X=)

[1312] [화학식 xviii(b)]



[1313]

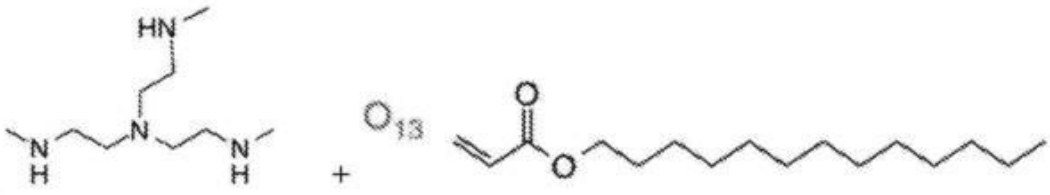
[1314] [화학식 xix]



[1315]

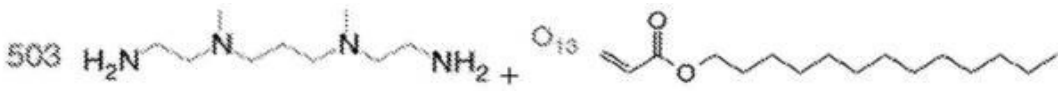
[1316] 일부 구현예에서, 본 명세서에 기재된 조성물의 전달을 위해 지질 나노입자를 형성하는 데 사용되는 지질 화합물, 예를 들어 본 명세서에 기재된 핵산(예를 들어, RNA)(예를 들어, 주형 핵산 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산)은 다음 반응들 중 어느 하나에 의해 제조된다:

[1317] [화학식 xx(a)]



[1318]

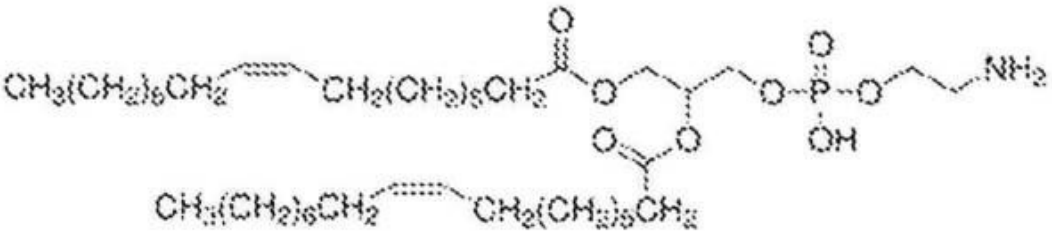
[1319] [화학식 xx(b)]



[1320]

[1321] 예시적인 비-양이온성 지질은 디스테아로일-sn-글리세로-포스포에탄올아민, 디스테아로릴포스파티딜콜린(DSPC), 디올레일포스파티딜콜린(DOPC), 디팔미토일포스파티딜콜린(DPPC), 디올레일포스파티딜글리세롤(DOPG), 디팔미토일포스파티딜글리세롤(DPPG), 디올레일포스파티딜에탄올아민(DOPE), 1,2-디올레일-sn-글리세로-3-포스포에탄올아민(DOPE), 팔미토일올레일포스파티딜콜린(POPC), 팔미토일올레일포스파티딜에탄올아민(POPE), 디올레일포스파티딜에탄올아민 4-(N-말레이미도메틸)-시클로헥산-1 - 카르복실레이트(DOPE-mal), 디팔미토일 포스파티딜 에탄올아민(DPPE), 디미리스토일포스포에탄올아민(DMPE), 디스테아로일-포스파티딜-에탄올아민(DSPE), 모노메틸-포스파티딜에탄올아민(예컨대 16-O-모노메틸 PE), 디메틸-포스파티딜에탄올아민(예컨대 16-O-디메틸 PE), 18-1-트랜스 PE, 1-스테아로일-2-올레일-포스파티딜에탄올아민(SOPE), 수소화 콩 포스파티딜콜린(HSPC), 달걀 포스파티딜콜린(EPC), 디올레일포스파티딜세린(DOPS), 스핑고미엘린(SM), 디미리스토일 포스파티딜콜린(DMPC), 디미리스토일포스파티딜글리세롤(DMPG), 디스테아로일포스파티딜글리세롤(DSPG), 디에루코일포스파티딜콜린(DEPC), 팔미토일올레일포스파티딜글리세롤(POPG), 디엘라이도일-포스파티딜에탄올아민(DEPE), 레시틴, 포스파티딜에탄올아민, 리소레시틴, 리소포스파티딜에탄올아민, 포스파티딜세린, 포스파티딜이노시톨, 스핑고미엘린, 달걀 스핑고미엘린(ESM), 세팔린, 카르디올리핀, 포스파티드산, 세레브로사이드, 디세틸포스페이트, 리소포스파티딜콜린, 디리올레일포스파티딜콜린, 또는 이들의 혼합물을 포함하지만 이에 제한되지 않는다. 기타 디아실포스파티딜콜린 및 디아실포스파티딜에탄올아민 인지질이 또한 사용될 수 있는 것으로 이해된다. 이들 지질에서 아실 기는 바람직하게는 C10-C24의 탄소 사슬을 갖는 지방산, 예를 들어 라우로일, 미리스토일, 팔미토일, 스테아로일 또는 올레일로부터 유도된 아실 기이다. 추가의 예시적인 지질은 특정 구현예에서, 본 명세서에 참고로 포함된 문헌[Kim et al. (2020) dx.doi.org/10.1021/acs.nanolett.0c01386]을 제한없이 포함한다. 그러한 지질은 일부 구현예에서, mRNA와 함께 간 형질주입을 개선하는 것으로 밝혀진 식물 지질(예를 들어, DGTS)을 포함한다. 일부 구현예에서, 비-양이온성 지질은 다음 구조를 가질 수 있다:

[1322] [화학식 xxi]



[1323]

[1324] 지질 나노입자에서의 이용에 적합한 비-양이온성 지질의 기타 예는, 예를 들어, 스테아릴아민, 도데실아민(dodeeylamine), 헥사데실아민, 아세틸 팔미테이트, 글리세롤 리시놀레이트, 헥사데실 스테아레이트, 이소프로필 미리스테이트, 양쪽성 아크릴성 중합체, 트리에탄올아민-라우릴 설페이트, 알킬-아릴 설페이트 폴리에틸옥실레이트화 지방산 아미드, 디옥타데실 디메틸 암모늄 브로마이드, 세라마이드, 스핑고미엘린 등과 같은 비-인지질을 제한없이 포함한다. 기타 비-양이온성 지질은, 그 내용 전체가 본 명세서에 참고로 포함된, WO2017/099823 또는 미국 특허 공개 US2018/0028664에 기재되어 있다.

[1325] 일부 구현예에서, 비-양이온성 지질은 올레산 또는, 그 전체가 본 명세서에 참고로 포함된, US2018/0028664의

화학식 I, II 또는 IV의 화합물이다. 비-양이온성 지질은, 예를 들어 지질 나노입자에 존재하는 총 지질 중 0 내지 30%(몰)를 포함할 수 있다. 일부 구현예에서, 비-양이온성 지질 함량은 지질 나노입자에 존재하는 총 지질의 5 내지 20%(몰) 또는 10 내지 15%(몰)이다. 일 구현예에서, 이온화 가능한 지질 대 중성 지질의 몰 비율은 약 2:1 내지 약 8:1(예를 들어, 약 2:1, 3:1, 4:1, 5:1, 6:1, 7: 1 또는 8:1)의 범위이다.

[1326] 일부 구현예에서, 지질 나노입자는 임의의 인지질을 포함하지 않는다.

[1327] 일부 양태에서, 지질 나노입자 막 온전성을 제공하기 위해, 스테롤과 같은 구성요소를 추가로 포함할 수 있다. 지질 나노입자에서 사용될 수 있는 일 예시적인 스테롤은 콜레스테롤 및 이의 유도체이다. 콜레스테롤 유도체의 비제한적인 예는, 극성 유사체들, 예컨대 5a-콜레스탄올(choiestanol), 53-코프로스탄올, 콜레스테릴-(2'-하이드록시)-에틸 에테르, 콜레스테릴-(4'-하이드록시)-부틸 에테르, 및 6-케토콜레스탄올; 비-극성 유사체들, 예컨대 5a-콜레스탄, 콜레스테논, 5a-콜레스타논, 5p-콜레스타논, 및 콜레스테릴 데카노에이트; 및 이들의 혼합물을 포함한다. 일부 구현예에서, 콜레스테롤 유도체는 극성 유사체, 예를 들어 콜레스테릴-(4 '-하이드록시)-부틸 에테르이다. 예시적인 콜레스테롤 유도체는 PCT 공개 WO2009/127060 및 미국 특허 공개 US2010/0130588에 기재되어 있으며, 이들 각각은 그 전체가 본 명세서에 참고로 포함된다.

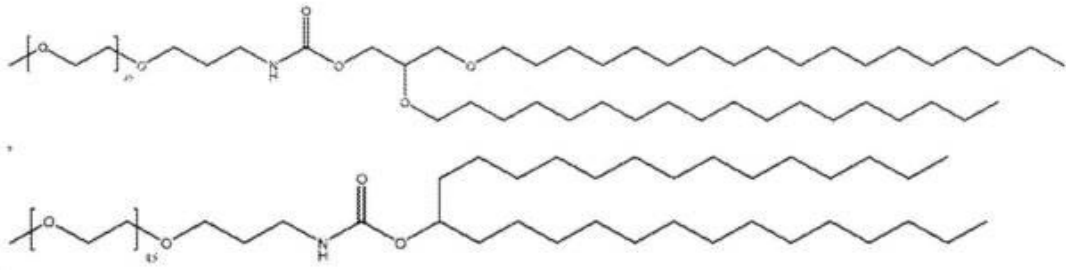
[1328] 일부 구현예에서, 막 온전성을 제공하는 구성요소, 예컨대 스테롤은 지질 나노입자에 존재하는 총 지질의 0 내지 50%(몰)(예를 들어, 0 내지 10%, 10 내지 20%, 20 내지 30%, 30 내지 40%, 또는 40 내지 50%)를 포함할 수 있다. 일부 구현예에서, 그러한 구성요소는 지질 나노입자에 존재하는 총 지질 함량의 20 내지 50%(몰) 30 내지 40%(몰)이다.

[1329] 일부 구현예에서, 지질 나노입자는 폴리에틸렌 글리콜(PEG) 또는 접합된 지질(conjugated lipid) 분자를 포함할 수 있다. 일반적으로, 이들은 지질 나노입자의 집합을 저해하는 데 사용되고/되거나 입체 안정화를 제공한다. 예시적인 접합된 지질은 이에 제한되지는 않지만, PEG-지질 접합물, 폴리옥사졸린(POZ)-지질 접합물, 폴리아미드-지질 접합물(예컨대 ATTA-지질 접합물), 양이온성-중합체 지질(CPL) 접합물, 및 이들의 혼합물을 포함한다. 일부 구현예에서, 접합된 지질 분자는 PEG-지질 접합물, 예를 들어(메톡시 폴리에틸렌글리콜)-접합된 지질이다.

[1330] 예시적인 PEG-지질 접합물은 PEG-디아실글리세롤(DAG)(예컨대, 1-(모노메톡시-폴리에틸렌글리콜)-2,3-디미리스토일글리세롤(PEG-DMG)), PEG-디알킬옥시프로필(DAA), PEG-인지질, PEG-세라마이드(Cer), 페길화 포스파티딜에탄올아민(PEG-PE), 1,2-디미리스토일-sn-글리세롤, 메톡시폴리에틸렌 글리콜(DMG-PEG-2K), PEG 석시네이트 디아실글리세롤(PEGS-DAG)(예컨대 4-0-(2',3'-디(테트라데카노일옥시)프로필-1-0-(w-메톡시(폴리에톡시)에틸) 부탄디오에이트(PEG-S-DMG)), PEG 디알콕시프로필카르밤, N-(카르보닐-메톡시폴리에틸렌 글리콜 2000)-1,2-디스테아릴-sn-글리세로-3-포스포에탄올아민 나트륨 염, 또는 이들의 혼합물을 포함하지만 이에 제한되지 않는다. 추가의 예시적인 PEG-지질 접합물은, 예를 들어 US5,885,613, US6,287,591, US2003/0077829, US2003/0077829, US2005/0175682, US2008/0020058, US2011/0117125, US2010/0130588, US2016/0376224, US2017/0119904, 및 US/099823에 기재되어 있으며, 이들 모두의 내용은 그 전체가 본 명세서에 참고로 포함된다.

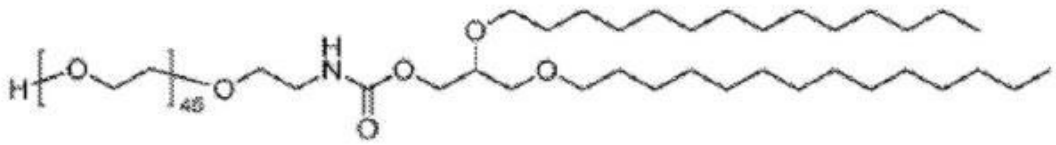
[1331] 일부 구현예에서, PEG-지질은, 그 내용 전체가 본 명세서에 참고로 포함된 US2018/0028664의 화학식 III, III-a-I, III-a-2, III-b-1, III-b-2, 또는 V의 화합물이다. 일부 구현예에서, PEG-지질은, 이들의 내용 전체가 본 명세서에 참고로 포함된, US20150376115 또는 US2016/0376224의 화학식 II이다. 일부 구현예에서, PEG-DAA 접합물은, 예를 들어 PEG-디라우릴옥시프로필, PEG-디미리스틸옥시프로필, PEG-디팔미틸옥시프로필, 또는 PEG-디스테아릴옥시프로필일 수 있다. PEG-지질은 PEG-DMG, PEG-디라우릴글리세롤, PEG-디팔미토일글리세롤, PEG-디스테릴글리세롤, PEG-디라우릴글리카미드, PEG-디미리스틸글리카미드, PEG-디팔미토일글리카미드, PEG-디스테릴글리카미드, PEG-콜레스테롤(1-[8'-(콜레스트-5-엔-3[베타]-옥시)카르복스아미드-3',6'-디옥시옥타닐] 카르바모일-[오메가]-메틸-폴리(에틸렌 글리콜), PEG-DMB(3,4-디테트라데옥시벤질-[오메가]-메틸-폴리(에틸렌 글리콜) 에테르), 및 1,2-디미리스토일-sn-글리세로-3-포스포에탄올아민-N-[메톡시(폴리에틸렌 글리콜)-2000] 중 하나 이상일 수 있다. 일부 구현예에서, PEG-지질은 PEG-DMG 1,2-디미리스토일-sn-글리세로-3-포스포에탄올아민-N-[메톡시(폴리에틸렌글리콜))-2000]을 포함한다. 일부 구현예에서, PEG-지질은 다음으로부터 선택되는 구조를 포함한다:

[1332] [화학식 xxii],



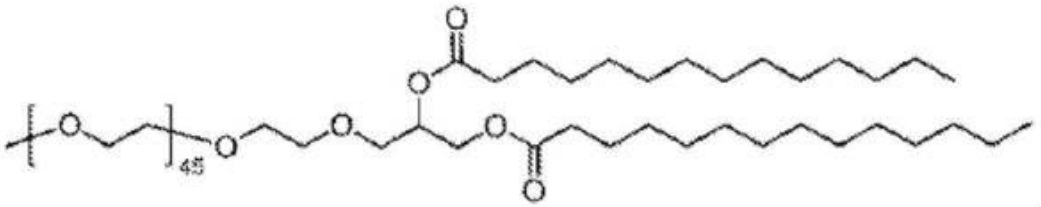
[1333]

[1334] [화학식 xxiii]



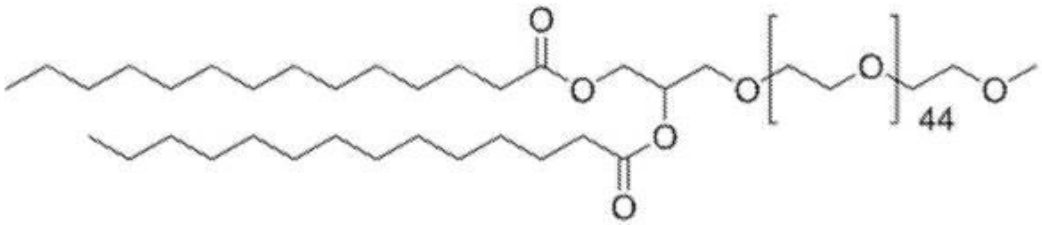
[1335]

[1336] [화학식 xxiv]



[1337] , 및

[1338] [화학식 xxv]



[1339] .

[1340] 일부 구현예에서, PEG 외의 분자와 접합된 지질은 PEG-지질 대신 사용될 수도 있다. 예를 들어, 폴리옥사졸린(POZ)-지질 접합물, 폴리아미드-지질 접합물(예컨대 ATTA-지질 접합물), 및 양이온성-중합체 지질(GPL) 접합물은 PEG-지질 대신 또는 이에 추가하여 사용될 수 있다.

[1341] 예시적인 접합된 지질, 즉 PEG-지질, (POZ)-지질 접합물, ATTA-지질 접합물 및 양이온성 중합체-지질은 WO2019051289A9 및 WO2020106946A1의 표 2에 열거된 PCT 및 LIS에 기재되었으며, 이들 내용 모두는 그 전체가 본 명세서에 참고로 포함된다.

[1342] 일부 구현예에서, LNP는 화학식 xix의 화합물, 화학식 xxi의 화합물 및 화학식 xxv의 화합물을 포함한다. 일부 구현예에서, 화학식 xix, 화학식 xxi 및 화학식 xxv의 제형을 포함하는 LNP는 본 명세서에 기재된 유전자 변형 조성물을 폐 또는 허파 세포로 전달하는 데 사용된다.

[1343] 일부 구현예에서, 모든 지질 나노입자는 화학식 i, 화학식 ii, 화학식 iii, 화학식 vii, 및 화학식 ix로부터 선택되는 하나 이상의 양이온성 지질을 포함할 수 있다. 일부 구현예에서, LNP는 하나 이상의 중성 지질, 예를 들어 DSPC, DPPC, DMPC, DOPC, POPC, DOPE, SM, 스테로이드, 예를 들어 콜레스테롤, 및/또는 하나 이상의 중합체 접합 지질, 예를 들어 페길화 지질, 예를 들어 PEG-DAG, PEG-PE, PEG-S-DAG, PEG-cer 또는 PEG 디알콕시프로필 카르바메이트를 추가로 포함할 수 있다.

[1344] 일부 구현예에서, PEG 또는 접합된 지질은 지질 나노입자에 존재하는 총 지질의 0 내지 20%(몰)을 포함할 수 있

다. 일부 구현예에서, PEG 또는 접합된 지질 함량은 지질 나노입자에 존재하는 총 지질의 0.5 내지 10% 또는 2 내지 5%(몰)이다. 이온성 지질, 비양이온성 지질, 스테롤 및 PEG/접합된 지질의 몰 비는 필요에 따라 변화될 수 있다. 예를 들어, 지질 입자는 조성물의 몰 기준 또는 총 중량 기준으로 30 내지 70%의 이온화 지질, 조성물의 몰 기준 또는 총 중량 기준으로 0 내지 60%의 콜레스테롤, 조성물의 몰 기준 또는 총 중량 기준으로 0 내지 30%의 비-양이온성 지질, 및 조성물의 몰 기준 또는 총 중량 기준으로 1 내지 10%의 접합된 지질을 포함할 수 있다. 바람직하게는, 조성물은 조성물의 몰 기준 또는 총 중량 기준으로 30 내지 40%의 이온화 지질, 조성물의 몰 기준 또는 총 중량 기준으로 40 내지 50%의 콜레스테롤, 및 조성물의 몰 기준 또는 총 중량 기준으로 10 내지 20%의 비-양이온성 지질을 포함한다. 일부 기타 구현예들에서, 조성물은 조성물의 몰 기준 또는 총 중량 기준으로 50 내지 75% 이온화 지질, 조성물의 몰 기준 또는 총 중량 기준으로 20 내지 40%의 콜레스테롤, 및 조성물의 몰 기준 또는 총 중량 기준으로 5 내지 10%의 비-양이온성 지질 및 조성물의 몰 기준 또는 총 중량 기준으로 1 내지 10%의 접합된 지질이다. 조성물은 조성물의 몰 기준 또는 총 중량 기준으로 60 내지 70%의 이온화 지질, 조성물의 몰 기준 또는 총 중량 기준으로 25 내지 35%의 콜레스테롤, 및 조성물의 몰 기준 또는 총 중량 기준으로 5 내지 10%의 비-양이온성 지질을 포함할 수 있다. 조성물은 조성물의 몰 기준 또는 총 중량 기준으로 90%의 이온화 지질 및 조성물의 몰 기준 또는 총 중량 기준으로 2 내지 15%의 비-양이온성 지질도 함유할 수 있다. 제형은, 예를 들어 조성물의 몰 기준 또는 총 중량 기준으로 8 내지 30%의 이온화 지질, 조성물의 몰 기준 또는 총 중량 기준으로 5 내지 30%의 비-양이온성 지질, 및 조성물의 몰 기준 또는 총 중량 기준으로 0~20%의 콜레스테롤; 조성물의 몰 기준 또는 총 중량 기준으로 4 내지 25%의 이온화 지질, 조성물의 몰 기준 또는 총 중량 기준으로 4 내지 25%의 비-양이온성 지질, 조성물의 몰 기준 또는 총 중량 기준으로 2 내지 25%의 콜레스테롤, 조성물의 몰 기준 또는 총 중량 기준으로 10 내지 35%의 접합 지질, 및 조성물의 몰 기준 또는 총 중량 기준으로 5%의 콜레스테롤; 또는 조성물의 몰 기준 또는 총 중량 기준으로 2 내지 30%의 이온화 지질, 조성물의 몰 기준 또는 총 중량 기준으로 2 내지 30%의 비-양이온성 지질, 조성물의 몰 기준 또는 총 중량 기준으로 1 내지 15%의 콜레스테롤, 조성물의 몰 기준 또는 총 중량 기준으로 2 내지 35%의 접합 지질, 및 조성물의 몰 기준 또는 총 중량 기준으로 1 내지 20%의 콜레스테롤; 또는 심지어는 조성물의 몰 기준 또는 총 중량 기준으로 90% 이하의 이온화 지질 및 조성물의 몰 기준 또는 총 중량 기준으로 2 내지 10%의 비-양이온성 지질, 또는 조성물의 몰 기준 또는 총 중량 기준으로 심지어는 100%의 양이온성 지질을 포함하는 지질 나노입자 제형일 수도 있다. 일부 구현예에서, 지질 입자 제제는 50: 10:38.5: 1.5 몰 비의 이온화 지질, 인지질, 콜레스테롤 및 PEG-일화 지질을 포함한다. 일부 다른 구현예에서, 지질 입자 제형은 60:38.5: 1.5 몰 비의 이온화 지질, 콜레스테롤 및 PEG화 지질을 포함한다.

[1345] 일부 구현예에서, 지질 입자는 이온화 지질, 비-양이온성 지질(예를 들어, 인지질), 스테롤(예를 들어, 콜레스테롤) 및 PEG화 지질을 포함하며, 여기서 지질의 몰 비는 이온화 지질의 경우 40~60을 목표로, 20~70 몰 백분율의 범위이며, 비-양이온성 지질의 몰 백분율은 0 내지 15를 목표로, 0 내지 30의 범위이며, 스테롤의 몰 백분율은 30 내지 50을 목표로, 20 내지 70의 범위이며, PEG화 지질의 몰 백분율은 2 내지 5를 목표로, 1 내지 6의 범위다.

[1346] 일부 구현예에서, 지질 입자는 50: 10:38.5: 1.5 몰 비의 이온화 지질 / 비-양이온성 지질/ 스테롤/ 접합된 지질을 포함한다.

[1347] 일 양태에서, 본 개시내용은 인지질, 레시틴, 포스파티딜콜린 및 포스파티딜에탄올아민을 포함하는 지질 나노입자 제형을 제공한다.

[1348] 일부 구현예에서, 하나 이상의 추가 화합물이 포함될 수도 있다. 그러한 화합물은 별도로 투여될 수 있거나, 또는 추가의 화합물은 본 발명의 지질 나노입자에 포함될 수 있다. 즉, 지질 나노입자는, 핵산에 더하여 기타 화합물 또는 첫 번째와 다른, 적어도 제2의 핵산을 함유할 수 있다. 제한없이, 기타 추가의 화합물들은 작은 또는 큰 유기 또는 무기 분자, 단당류, 이당류, 삼당류, 올리고당, 다당류, 펩티드, 단백질, 펩티드 유사체 및 이의 유도체, 펩티도미메틱스(peptidomimetics), 핵산, 핵산 유사체 및 이의 유도체, 생물학적 재료로 제조된 추출물, 또는 이의 임의의 조합으로 이루어진 군으로부터 선택될 수 있다.

[1349] 일부 구현예에서, 지질 나노입자(또는 지질 나노입자를 포함하는 제형)는 반응성 불순물(예를 들어, 알데히드 또는 케톤)을 결여하거나, 또는 예측된 수준 미만의 반응성 불순물(예를 들어, 알데히드 또는 케톤)을 포함한다. 이론에 구속되고자 하지는 않지만, 일부 구현예에서 지질 시약이 지질 나노입자 제형을 만드는 데 사용되며, 지질 시약은 오염성 반응성 불순물(예를 들어, 알데히드 또는 케톤)을 포함할 수 있다. 지질 시약은 예측된 수준 미만의 반응성 불순물(예를 들어, 알데히드 또는 케톤)을 갖는 것에 기초한 제조를 위해 선택될 수 있다. 이론에 구속되고자 함이 없이, 일부 구현예에서 알데히드는 RNA의 변형 및 손상, 예를 들어 염기들 간의

가교결합 및/또는 지질의 RNA로의 공유결합 접합(예를 들어, 지질-RNA 부가체를 형성)을 일으킬 수 있다. 이는 일부 경우에서, 역전사효소 반응의 실패 및/또는 부적절한 염기들의, 예를 들어 병소(들)의 부위(들)에서의 혼입, 예를 들어 새롭게 합성된 표적 DNA에서의 돌연변이를 야기할 수 있다.

[1350] 일부 구현예에서, 지질 나노입자 제형은 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 총 반응성 불순물(예를 들어, 알데히드) 함량을 포함하는 지질 시약을 사용하여 생산된다. 일부 구현예에서, 지질 나노입자 제형은 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 임의의 단일 반응성 불순물(예를 들어, 알데히드) 종을 포함하는 지질 시약을 사용하여 생산된다. 일부 구현예에서, 지질 나노입자 제형은, (i) 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 총 반응성 불순물(예를 들어, 알데히드) 함량; 및 (ii) 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 임의의 단일 반응성 불순물(예를 들어, 알데히드) 종을 포함하는 지질 시약을 사용하여 생산된다. 일부 구현예에서, 지질 나노입자 제형은 복수의 지질 시약을 사용하여 생산되며, 그 다수 중 각각의 지질 시약은 이 문단에서 설명된 하나 이상의 기준을 독립적으로 충족한다. 일부 구현예에서, 복수의 각각의 지질 시약은 동일한 기준, 예를 들어 이 문단의 기준을 충족한다.

[1351] 일부 구현예에서, 지질 나노입자 제형은 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 총 반응성 불순물(예를 들어, 알데히드) 함량을 포함한다. 일부 구현예에서, 지질 나노입자 제형은 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 임의의 단일 반응성 불순물(예를 들어, 알데히드) 종을 포함한다. 일부 구현예에서, 지질 나노입자 제형은, (i) 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 총 반응성 불순물(예를 들어, 알데히드) 함량; 및 (ii) 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 임의의 단일 반응성 불순물(예를 들어, 알데히드) 종을 포함한다.

[1352] 일부 구현예에서, 본 명세서에 기재된 바와 같은 지질 나노입자 또는 그의 제형에 사용된 지질 시약 중 하나 이상, 또는 선택적으로 모두는 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 총 반응성 불순물(예를 들어, 알데히드) 함량을 포함한다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 지질 나노입자 또는 그의 제형에 사용된 지질 시약 중 하나 이상, 또는 선택적으로 모두는 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 임의의 단일 반응성 불순물(예를 들어, 알데히드) 종을 포함한다. 일부 구현예에서, 본 명세서에 기재된 바와 같은 지질 나노입자 또는 그의 제형에 사용된 지질 시약 중 하나 이상, 또는 선택적으로 모두는, (i) 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 총 반응성 불순물(예를 들어, 알데히드) 함량; 및 (ii) 5%, 4%, 3%, 2%, 1%, 0.9%, 0.8%, 0.7%, 0.6%, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만의 임의의 단일 반응성 불순물(예를 들어, 알데히드) 종을 포함한다.

[1353] 일부 구현예에서, 총 알데히드 함량 및/또는 임의의 단일 반응성 불순물(예를 들어, 알데히드) 종의 양은, 예를 들어 PCT/US21/20948의 실시예 40에서 기재된 방법에 따라, 예를 들어 직렬 질량분석법(MS/MS)와 커플링된, 액체 크로마토그래피에 의해 결정된다. 일부 구현예에서, 반응성 불순물(예를 들어, 알데히드) 함량 및/또는 반응성 불순물(예를 들어, 알데히드) 종의 양은, 예를 들어 지질 시약에서 반응성 불순물(예를 들어 알데히드)의 존재와 회합된 핵산 분자(예를 들어 본 명세서에 기재된 바와 같은, 예를 들어 RNA 분자)의 하나 이상의 화학적 변형을 검출함으로써 결정된다. 일부 구현예에서, 반응성 불순물(예를 들어, 알데히드) 함량 및/또는 반응성 불순물(예를 들어, 알데히드) 종의 양은, 예를 들어 PCT/US21/20948의 실시예 41에 기재된 방법에 따라, 예를 들어 지질 시약에서, 반응성 불순물(예를 들어, 알데히드)의 존재와 회합된 뉴클레오티드 또는 뉴클레오시드(예를 들어, 리보뉴클레오티드 또는 리보뉴클레오시드, 예를 들어 본 명세서에 기재된 바와 같이, 예를 들어 주형 핵산에 포함된 또는 그로부터 단리된)의 하나 이상의 화학적 변형을 검출함으로써 결정된다. 구현예에서, 핵산 분자, 뉴클레오티드, 또는 뉴클레오시드의 화학적 변형은, 예를 들어 PCT/US21/20948의 실시예 41에 기재된 방법에 따라, 예를 들어 LC-MS/MS 분석을 사용하여, 하나 이상의 변형된 뉴클레오티드 또는 뉴클레오시드의 존재를 결정함으로써 검출된다.

[1354] 일부 구현예에서, 본 명세서에 기재된(예를 들어, 주형 핵산 또는 유전자 변형 폴리펩티드를 인코딩하는 핵산) 핵산(예를 들어, RNA은 알데히드 변형을 포함하지 않거나, 또는 알데히드 변형의 예비선택된 양 미만을 포함한다. 일부 구현예에서, 평균적으로, 핵산은 1000 뉴클레오티드 당 50, 20, 10, 5, 2, 또는 1 미만의 알데히드 변형을 갖고, 예를 들어 여기서 2개의 뉴클레오티드의 단일 가교결합은 단일 알데히드 변형이다. 일부 구현예에서, 알데히드 변형은 RNA 부가체(예를 들어, 지질-RNA 부가체)이다. 일부 구현예에서, 알데히드-변형된

뉴클레오티드는 염기들 간에 가교결합한다. 일부 구현예에서, 본 명세서에 기재된 핵산(예를 들어, RNA)은 뉴클레오티드 간의 50, 20, 10, 5, 2, 또는 1 미만의 가교결합을 포함한다.

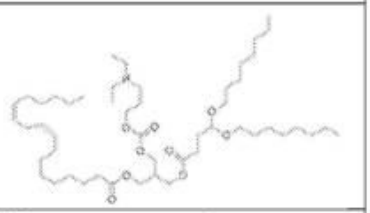
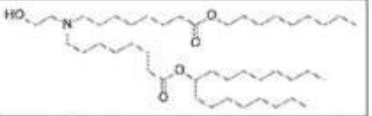
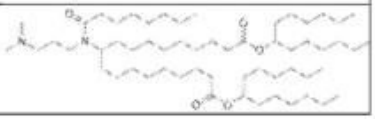
[1355] 일부 구현예에서, LNP는 표적화 도메인의 첨가에 의해 특이적 조직으로 향한다. 예를 들어, 생물학적 리간드는 LNP의 표면 상에 디스플레이되어 동족 수용체를 디스플레이는 세포와의 상호작용을 향상할 수 있어서, 이에 따라 세포가 수용체를 발현하는 조직과의 회합 및 조직으로의 카고 전달을 구동한다. 일부 구현예에서, 생물학적 리간드는 간으로의 전달을 구동하는 리간드, 예를 들어 GalNAc를 디스플레이하는 LNP일 수 있으며, 이는 아시알로당단백질 수용체(ASGPR)를 디스플레이하는 간세포로 핵산 카고의 전달을 생성한다. 문헌[Akinc et al. *Mol Ther* 18(7):1357-1364 (2010)]의 연구는, 관찰가능한 LNP 카고 효과에 대한 ASGPR에 따라 달라지는 LNP를 산출하기 위한 PEG-지질(GalNAc-PEG-DSG)로의 3가 GalNAc 리간드의 접합을 교시한다(예를 들어, 이의 도 6 참조). 예를 들어 폴레이트, 트랜스페린 또는 항체를 포함하는, 기타 리간드를 디스플레이하는 LNP 제형은 그 전체가 본 명세서에 참고로 포함된 WO2017223135에서 그 안에 사용된 참조, 즉 문헌[Kolhatkar et al., Curr Drug Discov Technol. 2011 8:197-206; Musacchio and Torchilin, Front Biosci. 2011 16:1388-1412; Yu et al., Mol Membr Biol. 2010 27:286-298; Patil et al., Crit Rev Ther Drug Carrier Syst. 2008 25:1-61 ; Benoit et al., Biomacromolecules. 2011 12:2708-2714; Zhao et al., Expert Opin Drug Deliv. 2008 5:309-319; Akinc et al., Mol Ther. 2010 18:1357-1364; Srinivasan et al., Methods Mol Biol. 2012 820:105-116; Ben-Arie et al., Methods Mol Biol. 2012 757:497-507; Peer 2010 J Control Release. 20:63-68; Peer et al., Proc Natl Acad Sci U S A. 2007 104:4095-4100; Kim et al., Methods Mol Biol. 2011 721:339-353; Subramanya et al., Mol Ther. 2010 18:2028-2037; Song et al., Nat Biotechnol. 2005 23:709-717; Peer et al., Science. 2008 319:627-630; 및 Peer and Lieberman, Gene Ther. 2011 18:1127-1133]에 추가하여 논의된다.

[1356] 일부 구현예에서, LNP는, 전통적인 성분들, 예컨대 이온화 양이온성 지질, 양친매성 인지질, 콜레스테롤 및 폴리(에틸렌 글리콜)(PEG) 지질을 포함하는 제제에 선택적 기관 표적화(SORT) 분자의 첨가에 의해 조직-특이적 활성에 대해 선택된다. 문헌[Cheng et al. Nat Nanotechnol 15(4):313-320 (2020)]의 교시는 보충적 "SORT" 구성요소의 첨가는 생체 내 RNA 전달 프로파일을 정확하게 변경하고, SORT 분자의 백분율 및 생물물리학 특성의 함수로서, 조직-특이적(예를 들어, 폐, 간, 비장) 유전자 전달 및 편집을 매개한다는 것을 예증한다.

[1357] 일부 구현예에서, LNP는 생물분해성, 이온화 지질을 포함한다. 일부 구현예에서, LNP는 3-((4,4-비스(옥틸옥시)부타노일)옥시)-2-((((3-(디에틸아미노)프로폭시)카르보닐)옥시)메틸)프로필(9Z,12Z)-옥타데카-9,12-디에노에이트) 로도 지칭되는 (9Z,12Z)-3-((4,4-비스(옥틸옥시)부타노일)옥시)-2-((((3-(디에틸아미노)프로폭시)카르보닐)옥시)메틸)프로필 옥타데카-9,12-디에노에이트, 또는 또 다른 이온화 지질을 포함한다. 예를 들어, WO2019/067992, WO/2017/173054, WO2015/095340, 및 WO2014/136086뿐만 아니라, 그 안에 제공된 참고문헌의 지질을 참조한다. 일부 구현예에서, LNP 지질의 맥락에서 양이온성 및 이온화라는 용어는 상호 교환가능하며, 예를 들어 여기서 이온화 지질은 pH에 따라 달라지는 양이온성이다.

[1358] 일부 구현예에서, 본 명세서에 기재된 LNP는 표 19에 기재된 지질을 포함한다.

[1359]

지질 ID	화학명	분자량	구조
LIPIDV003	(9Z,12Z)-3-((4,4-비스(옥틸옥시)부타노일)옥시)-2-((((3-(디에틸아미노)프로폭시)카르보닐)옥시)메틸)프로필 옥타데카-9, 12-디에노에이트	852.29	
LIPIDV004	헵타데칸-9-일 8-((2-하이드록시에틸)(8-(노닐옥시)-8-옥소옥틸)아미노)옥타노에이트	710.18	
LIPIDV005		919.56	

[1360] 일부 구현예에서, 유전자 변형 시스템의 다수의 구성요소는 단일 LNP 제형으로서 제조될 수 있으며, 예를 들어 LNP 제형은 유전자 변형 폴리펩티드 및 RNA 주형을 인코딩하는 mRNA를 포함한다. 핵산 구성요소의 비는 치료제

의 특성을 최대화하기 위해 변화될 수 있다. 일부 구현예에서, RNA 주형 대 유전자 변형 폴리펩티드를 인코딩하는 mRNA의 비는 몰 비로, 약 1:1 내지 100:1, 예를 들어, 약 1:1 내지 20:1, 약 20:1 내지 40:1, 약 40:1 내지 60:1, 약 60:1 내지 80:1, 또는 약 80:1 내지 100:1이다. 다른 구현예에서, 다수의 핵산의 시스템은 별도의 제형, 예를 들어 주형 RNA를 포함하는 하나의 LNP 제형 및 유전자 변형 폴리펩티드를 인코딩하는 mRNAf를 포함하는 제2의 LNP 제형에 의해 제조될 수 있다. 일부 구현예에서, 시스템은 LNP로 제조된 2개가 넘는 핵산 구성요소를 포함할 수 있다. 일부 구현예에서, 시스템은 단백질, 예를 들어 유전자 변형 폴리펩티드, 및 적어도 하나의 LNP 제형으로 제조된 주형 RNA를 포함할 수 있다.

[1361] 일부 구현예에서, LNP 제형의 평균 LNP 직경은, 예를 들어 동적 광산란(DLS)에 의해 측정하여 10 범위의 nm 및 100 범위의 nm 사이일 수 있다. 일부 구현예에서, LNP 제형의 평균 LNP 직경은 약 40 nm 내지 약 150 nm, 예컨대 약 40 nm, 45 nm, 50 nm, 55 nm, 60 nm, 65 nm, 70 nm, 75 nm, 80 nm, 85 nm, 90 nm, 95 nm, 100 nm, 105 nm, 110 nm, 115 nm, 120 nm, 125 nm, 130 nm, 135 nm, 140 nm, 145 nm, 또는 150 nm일 수 있다. 일부 구현예에서, LNP 제형의 평균 LNP 직경은 약 50 nm 내지 약 100 nm, 약 50 nm 내지 약 90 nm, 약 50 nm 내지 약 80 nm, 약 50 nm 내지 약 70 nm, 약 50 nm 내지 약 60 nm, 약 60 nm 내지 약 100 nm, 약 60 nm 내지 약 90 nm, 약 60 nm 내지 약 80 nm, 약 60 nm 내지 약 70 nm, 약 70 nm 내지 약 100 nm, 약 70 nm 내지 약 90 nm, 약 70 nm 내지 약 80 nm, 약 80 nm 내지 약 100 nm, 약 80 nm 내지 약 90 nm, 또는 약 90 nm 내지 약 100 nm일 수 있다. 일부 구현예에서, LNP 제형의 평균 LNP 직경은 약 70 nm 내지 약 100 nm일 수 있다. 특정 구현예에서, LNP 제형의 평균 LNP 직경은 약 80 nm일 수 있다. 일부 구현예에서, LNP 제형의 평균 LNP 직경은 약 100 nm일 수 있다. 일부 구현예에서, LNP 제형의 평균 LNP 직경은 약 1 mm 내지 약 500 mm, 약 5 mm 내지 약 200 mm, 약 10 mm 내지 약 100 mm, 약 20 mm 내지 약 80 mm, 약 25 mm 내지 약 60 mm, 약 30 mm 내지 약 55 mm, 약 35 mm 내지 약 50 mm, 또는 약 38 mm 내지 약 42 mm의 범위이다.

[1362] LNP는 일부 경우에서 상대적으로 균질할 수 있다. 다분산성 지수는 LNP의 균질성, 예를 들어 지질 나노입자의 입자 크기 분포를 나타내는 데 사용될 수 있다. 작은(예를 들어, 0.3 미만) 다분산성 지수는 일반적으로 좁은 입자 크기 분포를 나타낸다. LNP는 약 0 내지 약 0.25, 예컨대 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.11, 0.12, 0.13, 0.14, 0.15, 0.16, 0.17, 0.18, 0.19, 0.20, 0.21, 0.22, 0.23, 0.24, 또는 0.25의 다분산성 지수를 가질 수 있다. 일부 구현예에서, LNP의 다분산성 지수는 약 0.10 내지 약 0.20일 수 있다.

[1363] LNP의 제타 전위는 조성물의 동전위를 나타내는 데 사용될 수 있다. 일부 구현예에서, 제타 전위는 LNP의 표면 하전을 기재할 수 있다. 상대적으로 낮은 양 또는 음의 하전을 갖는 지질 나노입자가 일반적으로 바람직한데, 이는 고하전 종이 많을수록 신체에서 세포, 조직 및 기타 요소와 바람직하지 않게 상호작용할 수 있다. 일부 구현예에서, LNP의 제타 전위는 약 -10 mV 내지 약 +20 mV, 약 -10 mV 내지 약 +15 mV, 약 -10 mV 내지 약 +10 mV, 약 -10 mV 내지 약 +5 mV, 약 -10 mV 내지 약 0 mV, 약 -10 mV 내지 약 -5 mV, 약 -5 mV 내지 약 +20 mV, 약 -5 mV 내지 약 +15 mV, 약 -5 mV 내지 약 +10 mV, 약 -5 mV 내지 약 +5 mV, 약 -5 mV 내지 약 0 mV, 약 0 mV 내지 약 +20 mV, 약 0 mV 내지 약 +15 mV, 약 0 mV 내지 약 +10 mV, 약 0 mV 내지 약 +5 mV, 약 +5 mV 내지 약 +20 mV, 약 +5 mV 내지 약 +15 mV, 또는 약 +5 mV 내지 약 +10 mV일 수 있다.

[1364] 단백질 및/또는 핵산, 예를 들어 유전자 변형 폴리펩티드 또는 폴리펩티드를 인코딩하는 mRNA의 캡슐화의 효율은, 제공된 최초 양에 비해, 캡슐화되거나 다르게는, 제조 후 LNP와 회합된, 단백질 및/또는 핵산의 양을 기재한다. 캡슐화 효율은 높은 것(예를 들어, 100%에 가까운 것)이 바람직하다. 캡슐화 효율은, 예를 들어 하나 이상의 유기 용매 또는 세제로 지질 나노입자를 분해하기 전과 후에 지질 나노입자를 함유하는 용액 중의 단백질 또는 핵산의 양을 비교함으로써, 측정될 수 있다. 음이온 교환 수지는 용액 중 자유 단백질 또는 핵산(예를 들어, RNA)의 양을 측정하는 데 사용될 수 있다. 형광은 용액 중 자유 단백질 및/또는 핵산(예를 들어, RNA)의 양을 측정하는 데 사용될 수 있다. 본 명세서에 기재된 지질 나노입자의 경우, 단백질 및/또는 핵산의 캡슐화 효율은 적어도 50%, 예를 들어 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, 또는 100%일 수 있다. 일부 구현예에서, 캡슐화 효율은 적어도 80%일 수 있다. 일부 구현예에서, 캡슐화 효율은 적어도 90%일 수 있다. 일부 구현예에서, 캡슐화 효율은 적어도 95%일 수 있다.

[1365] LNP는 하나 이상의 코팅을 선택적으로 포함할 수 있다. 일부 구현예에서, LNP는 코팅을 갖는 캡슐, 필름 또는 정제로 제형화될 수 있다. 본 명세서에 기재된 조성물을 포함하는 캡슐, 필름, 또는 정제는 임의의 유용한 크기, 인장 강도, 경도 또는 밀도를 가질 수 있다.

[1366] LNP의 추가의 예시적인 지질, 제형, 방법, 및 특성화는, WO2020061457에 의해 교시되며, 이는 그 전체가 본 명

세서에 참고로 포함된다.

[1367] 일부 구현예에서, 시험관 내 또는 생체 외 세포 리포펙션은 Lipofectamine MessengerMax(Thermo Fisher) 또는 TransIT-mRNA 형질주입 시약(Mirus Bio)을 사용하여 수행된다. 특정 구현예에서, LNP는 GenVoy_ILM 이온화 지질 혼합물(Precision Nanosystems)을 사용하여 제형화된다. 특정 구현예에서, LNP는 2,2-디리놀레일-4-디메틸아미노에틸-[1,3]-디옥솔란(DLin-KC2-MA) 또는 디리놀레일메틸-4-디메틸아미노부티레이트(DLin-MC3-DMA 또는 MC3)을 사용하여 제형화되며, 상기 제형 및 이의 생체 내 사용은, 그 전체가 본 명세서에 참고로 포함된, 문헌[Jayaraman et al. Angew Chem Int Ed Engl 51(34):8529-8533 (2012)]에 교시된다.

[1368] CRISPR-Cas 시스템, 예를 들어 Cas9-gRNA RNP, gRNA, Cas9 mRNA의 전달에 최적화된 LNP 제형은, 둘 다 참고로 포함된, WO2019067992 및 WO2019067910에 기재되어 있다.

[1369] 핵산의 전달에 유용한 추가적인 특이적 LNP 제형은, 둘 다가 참고로 포함된 US8158601 및 US8168775에 기재되어 있으며, 이는 ONPATTRO라는 명칭 하에 판매되는, 파티시란(patisiran)에서 사용된 제형을 포함한다.

[1370] 유전자 변형 LNP의 예시적인 투약량은 약 0.1, 0.25, 0.3, 0.5, 1, 2, 3, 4, 5, 6, 8, 10, 또는 100 mg/kg(RNA)을 포함할 수 있다. 시스템의 하나 이상의 구성요소를 인코딩하는 핵산을 포함하는 AAV의 예시적인 투여량은 약 10^{11}, 10^{12}, 10^{13}, 및 10^{14} vg/kg의 MOI를 포함할 수 있다.

[1371] ***키트, 제조 물품, 및 약제학적 조성물***

[1372] 일 양태에서 본 개시내용은, 예를 들어 본 명세서에 기재된 바와 같은, 유전자 변형 폴리펩티드 또는 유전자 변형 시스템을 포함하는 키트를 제공한다. 일부 구현예에서, 키트는 유전자 변형 폴리펩티드(폴리펩티드를 인코딩하는 핵산) 및 주형 RNA(또는 상기 주형 RNA를 인코딩하는 DNA)를 포함한다. 일부 구현예에서, 키트는 시스템을 세포 내로 도입하기 위한 시약, 예를 들어 형질감염 시약, LNP 등을 추가로 포함한다. 일부 구현예에서, 키트는 본 명세서에 기재된 임의의 방법에 적합하다. 일부 구현예에서, 키트는, 예를 들어 제조 물품 내에 배치된, 하나 이상의 요소, 조성물(예를 들어, 약제학적 조성물), 유전자 변형 폴리펩티드, 및/또는 유전자 변형 시스템, 또는 이의 기능성 단편 또는 구성요소를 포함한다. 일부 구현예에서, 키트는 이의 사용을 위한 지침을 포함한다.

[1373] 일 양태에서, 본 개시내용은 제조 물품을 제공하며, 예를 들어 그 안에 본 명세서에 기재된 바와 같은 키트, 또는 이의 구성요소가 배치된다.

[1374] 일 양태에서 본 개시내용은, 예를 들어 본 명세서에 기재된 바와 같은, 유전자 변형 폴리펩티드 또는 유전자 변형 시스템을 포함하는 약제학적 조성물을 제공한다. 일부 구현예에서, 약제학적 조성물은 약제학적으로 허용 가능한 담체 또는 부형제를 추가로 포함한다. 일부 구현예에서, 약제학적 조성물은 주형 RNA 및/또는 폴리펩티드를 인코딩하는 RNA를 포함한다. 구현예에서, 약제학적 조성물은 다음 특성 중 하나 이상(예를 들어, 1, 2, 3, 또는 4개)을 갖는다:

[1375] (a) 예를 들어 몰 기준으로, 주형 RNA 및/또는 폴리펩티드를 인코딩하는 RNA에 대하여, 1% 미만(예를 들어, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만)의 DNA 주형;

[1376] (b) 예를 들어 몰 기준으로, 주형 RNA 및/또는 폴리펩티드를 인코딩하는 RNA에 대하여, 1% 미만(예를 들어, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만)의 캡핑되지 않은 RNA;

[1377] (c) 예를 들어 몰 기준으로, 주형 RNA 및/또는 폴리펩티드를 인코딩하는 RNA에 대하여, 1% 미만(예를 들어, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만)의 부분 길이 RNA;

[1378] (d) 반응되지 않은 캡 디뉴클레오티드가 실질적으로 결여됨.

[1379] i) ***화학, 제조, 및 제어(CMC)***

[1380] 단백질 치료제의 정제는, 예를 들어 문헌[Franks, Protein Biotechnology: *Isolation, Characterization, and Stabilization*, Humana Press (2013); 및 Cutler, *Protein Purification Protocols (Methods in Molecular Biology)*, Humana Press (2010)]에 기재되어 있다.

[1381] 일부 구현예에서, 유전자 변형 시스템, 폴리펩티드, 및/또는 주형 핵산(예를 들어, 주형 RNA)은 특정 품질 표준에 따른다. 일부 구현예에서, 본 명세서에 기재된 방법에 의해 생산된 유전자 변형 시스템, 폴리펩티드, 및/또는 주형 핵산(예를 들어, 주형 RNA)은 특정 품질 표준에 따른다. 따라서, 본 개시내용은, 일부 양태에서 특정

품질 표준에 따르는 유전자 변형 시스템, 폴리펩티드, 및/또는 주형 핵산(예를 들어, 주형 RNA)의 제조 방법에 관한 것이며, 예를 들어 여기서 상기 품질 표준이 분석된다. 본 개시내용은, 일부 양태에서 유전자 변형 시스템, 폴리펩티드, 및/또는 주형 핵산(예를 들어, 주형 RNA)에서 상기 품질 표준의 분석 방법에도 관한 것이다. 일부 구현예에서, 품질 표준은 다음 중 하나 이상(예를 들어, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 또는 12개)을 포함하지만 이에 제한되지 않는다:

[1382] (i) 주형 RNA의 길이, 예를 들어 주형 RNA가 상기 참조 길이 초과 또는 참조 길이 범위 내의 길이를 갖는지 여부, 예를 들어 존재하는 주형 RNA의 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99%는 100, 125, 150, 175, 또는 200개 뉴클레오티드 길이 초과인지 여부;

[1383] (ii) 주형 RNA 상에서 폴리A 테일의 존재, 부재, 및/또는 길이, 예를 들어 존재하는 주형 RNA의 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99%가 폴리A 테일(예를 들어, 길이가 적어도 5, 10, 20, 30, 50, 70, 100개 뉴클레오티드인 폴리A 테일)을 함유하는지 여부;

[1384] (iii) 주형 RNA 상의 5' 캡의 존재, 부재, 및/또는 유형, 예를 들어 존재하는 주형 RNA의 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99%가 5' 캡을 함유하는지 여부, 예를 들어 그 캡이 7-메틸구아노신 캡, 예를 들어 O-Me-m7G 캡인지 여부;

[1385] (iv) 주형 RNA에서 (예를 들어, 슈도우리딘, 디하이드로우리딘, 이노신, 7-메틸구아노신, 1-N-메틸슈도우리딘(1-Me-Ψ), 5-메톡시우리딘(5-MO-U), 5-메틸시티딘(5mC), 또는 잠금 뉴클레오티드로부터 선택되는) 하나 이상의 변형된 뉴클레오티드의 존재, 부재, 및/또는 유형, 예를 들어 존재하는 주형 RNA의 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99%가 하나 이상의 변형된 뉴클레오티드를 함유하는지 여부;

[1386] (v) (예를 들어, 시간 경과에 따라 및/또는 미리 선택된 조건 하에서) 주형 RNA의 안정성, 예를 들어 안정성 시험 후 주형 RNA의 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99%가 온전한 상태(예를 들어, 100, 125, 150, 175, 또는 200개 초과의 뉴클레오티드 길이)로 남아 있는지 여부;

[1387] (vi) DNA 변형 시스템에서 주형 RNA의 역가, 예를 들어 주형 RNA를 포함하는 시스템이 역가에 대해 분석된 후 표적 부위의 적어도 1%가 변형되는지 여부;

[1388] (vii) 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드의 길이, 예를 들어 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드가 참조 길이 초과 또는 참조 길이 범위 내의 길이를 갖는지 여부, 예를 들어 존재하는 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드의 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99%가 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1600, 1700, 1800, 1900, 또는 2000개 아미노산 길이를 초과하는지(및 선택적으로, 2500, 2000, 1500, 1400, 1300, 1200, 1100, 1000, 900, 800, 700, 또는 600개 아미노산 길이 이하) 여부;

[1389] (viii) 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드 상의 번역후 변형의 존재, 부재 및/또는 유형, 예를 들어 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드의 적어도 80, 85, 90, 95, 96, 97, 98, 또는 99%가 포스포릴화, 메틸화, 아세틸화, 미리스토일화, 팔미토일화, 이소프레닐화, 글리피아티온 또는 리포일화 또는 이들의 임의의 조합을 함유하는지 여부;

[1390] (ix) 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드에서 하나 이상의 인공, 합성 또는 비-천연 아미노산(예를 들어, 오르니틴, β-알라닌, GABA, δ아미노레불린산, PABA, D-아미노산(예를 들어, D-알라닌 또는 D-글루타메이트), 아미노이소부티르산, 데하이드로알라닌, 시스타티온, 란티오닌, 드젠콜산, 디아미노피멜산, 호모알라닌, 노르발린, 노르류신, 호모노르류신, 호모세린, O-메틸-호모세린 및 O-에틸-호모세린, 에티오닌, 셀레노시스테인, 셀레노호모시스테인, 셀레노메티오닌, 셀레노에티오닌, 텔루로시스테인, 또는 텔루로메티오닌으로부터 선택됨)의 존재, 부재, 및/또는 유형, 예를 들어 존재하는 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드의 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99%가 하나 이상의 인공, 합성, 또는 비-정규 아미노산을 함유하는지 여부;

[1391] (x) (예를 들어, 시간 경과에 따라 및/또는 미리 선택된 조건 하에서) 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드의 안정성, 예를 들어 안정성 시험 후 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드의 적어도 80%, 85%, 90%, 95%, 96%, 97%, 98%, 또는 99%가 온전한 상태(예를 들어, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, 1500, 1600, 1700, 1800, 1900, 또는 2000개 초과의 아미노산 길이(및 선택적으로, 2500, 2000, 1500, 1400, 1300, 1200, 1100, 1000, 900, 800,

700, 또는 600개 이하의 아미노산 길이))로 남아 있는지 여부;

[1392] (xi) DNA 변형 시스템에서 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드의 역가, 예를 들어 폴리펩티드, 제1 폴리펩티드, 또는 제2 폴리펩티드를 포함하는 시스템이 역가에 대해 분석된 후 표적 부위의 적어도 1%가 변형되는지 여부; 또는

[1393] (xii) 하나 이상의 발열원, 바이러스, 진균류, 박테리아 병원체, 또는 숙주 세포 단백질의 존재, 부재, 및/또는 수준, 예를 들어 시스템에 발열원, 바이러스, 진균류, 박테리아 병원체, 또는 숙주 세포 단백질 오염이 없거나 실질적으로 없는지 여부.

[1394] 일부 구현예에서, 본 명세서에 기재된 시스템 또는 약제학적 조성물에는 내독소가 없다.

[1395] 일부 구현예에서, 하나 이상의 발열원, 바이러스, 진균류, 박테리아 병원체, 및/또는 숙주 세포 단백질의 존재, 부재, 및/또는 수준이 결정된다. 구현예에서, 시스템에 발열원, 바이러스, 진균류, 박테리아 병원체, 및/또는 숙주 세포 단백질 오염이 없거나 실질적으로 없는지 여부가 결정된다.

[1396] 일부 구현예에서, 본 명세서에 기재된 바와 같은 약제학적 조성물 또는 시스템은 다음 특성 중 하나 이상(예를 들어, 1, 2, 3, 또는 4개)을 갖는다:

[1397] (a) 예를 들어 몰 기준으로, 주형 RNA 및/또는 폴리펩티드를 인코딩하는 RNA에 대하여, 1% 미만(예를 들어, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만)의 DNA 주형;

[1398] (b) 예를 들어 몰 기준으로, 주형 RNA 및/또는 폴리펩티드를 인코딩하는 RNA에 대하여, 1% 미만(예를 들어, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만)의 캡핑되지 않은 RNA;

[1399] (c) 예를 들어 몰 기준으로, 주형 RNA 및/또는 폴리펩티드를 인코딩하는 RNA에 대하여, 1% 미만(예를 들어, 0.5%, 0.4%, 0.3%, 0.2%, 또는 0.1% 미만)의 부분 길이 RNA;

[1400] (d) 반응되지 않은 캡 디뉴클레오티드가 실질적으로 결여됨.

[1401] **실시예**

[1402] **실시예 1: RNA-기반 유전자 변형을 위한 RNA 주형의 트랜스-동원**

[1403] 이 실시예는 인간 게놈 내 서열의 표적화된 편집을 위한 예시적인 3-구성요소 RNA 유전자 변형 시스템의 사용을 설명한다(도 1). 더 구체적으로, 이 예시적인 시스템에서, (1) 유전자 변형 폴리펩티드는 (2) 가이드 RNA 분자 및 (3) 트랜스-주형 RNA 분자에 결합하여, 함께 유전자 변형 복합체를 형성하며, 이는 표적 유전자좌에 결합하여 닉을 생성하고 게놈 DNA의 하나의 가닥에서 원하는 편집물의 주형 통합을 수행한다. 숙주 복구 경로는 제2 가닥에 편집물을 통합하고 DNA 닉을 결찰시킨다. 이 실시예는 인간 세포에서의 편집 활성에 대한 다양한 유전자 변형 시스템 구성 및 트랜스-주형 RNA 동원 상호작용의 효능을 평가하는 수단으로 시험관 내 유전자 변형을 위해 포유동물 세포에 3-구성요소 시스템을 도입하는 것을 추가로 설명한다.

[1404] <u>1) 유전자 변형 폴리펩티드:</u>

[1405] 이러한 예시적인 시스템의 유전자 변형 폴리펩티드는 다음을 포함한다:

[1406] 1. 하나의 엔도뉴클레아제 도메인이 비활성화된 Cas-뉴클레아제(예를 들어, Spy N863A Cas9),

[1407] 2. 역전사효소(RT), 및

[1408] 3. RNA 결합 단백질(RBP; 표 31에서와 같음). RBP는 개재 펩티드 링커 서열과 함께 또는 없이 1 내지 5회 반복될 수 있다.

[1409] 이들 3개 도메인은 (1) 표 32에서와 같이 펩티드 링커(도 2a), 또는 (2) 표 33에서와 같이 인테인 쌍을 사용하는 공유 결합에 의해, 또는 (3) (a) 표 34의 화학적 결합 또는 특정 광 파장에 의해 유도되거나, (b) 표 35 및 36의 리간드가 없는 상호작용에서 본질적으로 발생하는(도 2b), 두 융합 도메인(FD)의 이량체화에 의해 하나로 합쳐진다. FD는 일부 경우에 개재 펩티드 링커 서열과 함께 또는 없이 1 내지 30회 반복될 수 있다.

[1410] 3개의 도메인은 다음 구성 중 임의의 것으로 조립될 수 있으며, 여기서 하기 표시된 링커는 선택적으로 인테인- 또는 FD-쌍으로 대체될 수 있다:

[1411] · Cas9 - 링커 - RT - 링커 - $(RBP)_n$

[1412] · RT - 링커 - Cas9 - 링커 - $(RBP)_n$

[1413] · $(RBP)_n$ - 링커 - Cas9 - 링커 - RT

[1414] · $(RBP)_n$ - 링커 - RT - 링커 - Cas9

[1415] · Cas9 - 링커 $(RBP)_n$ - 링커 - RT

[1416] · RT - 링커 - $(RBP)_n$ - 링커 - Cas9

[1417] · Cas9 - 링커 - RT

[1418] 2) 가이드 RNA:

[1419] 가이드 RNA(gRNA) 분자(5'-GCCGAAGCACTGCACGCCGTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC -3'; 서열번호 16,701)는 Cas9에 결합하고 표적 유전자좌에 상보적인 프로-스페이서 서열을 함유하여, 표적 게놈 유전자좌로 유전자 변형 복합체의 국재화를 초래한다.

[1420] 3) 트랜스-주형 RNA:

[1421] 예시적인 트랜스-주형 RNA(ttRNA) 분자는 하기와 같이(그리고 도 3a 내지 3b에 도시된 바와 같이) 3'에서 5'으로 **PBS-주형** 영역을 함유한다:

[1422] 1. 닉이 생긴 DNA 가닥과 염기쌍을 형성하여 닉이 생긴 DNA의 프라이머 연장을 가능하게 하는 1 내지 17 nt 프라이머 결합 부위(PBS; 표 37에서와 같음), 바로 이어서

[1423] 2. 표적 유전자좌와 상동성인 0 내지 20 nt 영역(상동성 1; 표 38의 "상동성 1"에서와 같음)

[1424] 3. 게놈에 대한 원하는 변형 - 삽입, 결실 또는 치환(편집; 표 38의 "편집"에서와 같음)

[1425] 4. 표적 유전자좌와 상동성인 0 내지 500 nt 영역(상동성 2; 표 39의 "상동성 2"에서와 같음).

[1426] 예시적인 ttRNA는 추가로 다음을 포함한다:

[1427] 5. RBP 동원 부위(**RRS**; 표 40에서와 같음). RRS는 개재 RNA 링커 서열과 함께 또는 없이 1 내지 5회, 예를 들어 1 내지 4회 반복될 수 있다.

[1428] 예시적인 ttRNA는 또한 선택적으로 다음을 포함할 수 있다:

[1429] 6. 엑소뉴클레아제-매개 분해로부터 ttRNA 분자를 보호하는 말단-보호 RNA 2차 구조(**말단 블록**; 표 41에서와 같음).

[1430] ttRNA의 이러한 구성요소는 하기 구성 중 임의의 것으로 조립될 수 있으며, 여기서 RRS는 최대 20 nt의 RNA 링커 서열(가장 긴 서열은 5'-ACTAACATACAACTAACATA-3'임)에 의해 PBS-주형에 연결될 수 있다:

[1431] · $(RRS)_n$ - 링커 - PBS-주형

[1432] · $(RRS)_n$ - 링커 - PBS-주형 - 말단 블록

[1433] · PBS-주형 - 링커 - $(RRS)_n$

[1434] · 말단 블록 - PBS-주형 - 링커 - $(RRS)_n$

[1435] · 말단 블록 - $(RRS)_n$ - 링커 - PBS-주형

[1436] · 말단 블록 - $(RRS)_n$ - 링커 - PBS-주형 - 말단 블록

[1437] · PBS-주형 - 링커 - $(RRS)_n$ - 말단 블록

[1438] · 말단 블록 - PBS-주형 - 링커 - $(RRS)_n$ - 말단 블록

[1439] 3-구성요소 트랜스-유전자 변형 시스템 구성의 평가:

[1440] 방법 1:

[1441] 트랜스-유전자 변형 복합체의 다양한 구성의 게놈-편집 능력을 결정하기 위해, 다음 중 하나(표 42에서와 같음)를 발현하는 게놈 랜딩 패드를 보유하는 포유동물 HEK293T 또는 U2OS 세포주를 생성한다:

[1442] i. BFP-GFP SNP 리포터 라인: 단일 염기쌍 치환을 통해 GFP-발현 세포주로 전환되는 BFP-발현 세포주;

[1443] ii. 250 bp GFP 삽입 리포터 라인: GFP-발현이 결실된 코딩 서열의 250 bp 삽입을 통해 복원되는 비-형광 부분-결실-GFP-세포주; 또는

[1444] iii. mCherry 삽입 리포터 라인: mCherry 코딩 서열을 도입하는 >750 bp 삽입에 의한 GFP 코딩 서열의 중단을 통해 GFP-발현이 손실되고 mCherry 발현을 얻는 GFP-발현 세포주.

[1445] 유전자 변형 폴리펩티드는 DNA 플라스미드 또는 mRNA의 형질감염에 의해, 또는 DNA 플라스미드를 렌티바이러스에 패키징함으로써 이들 세포에 도입된다. 그 다음 유전자 변형 폴리펩티드를 함유하는 세포는, 다음을 하도록 함께 설계된 ttRNA-발현 DNA 플라스미드 또는 ncRNA, 및 gRNA-발현 DNA 플라스미드 또는 ncRNA로 형질감염/전기천공된다:

[1446] i. BFP-GFP SNP 리포터 라인: BFP를 GFP로 전환;

[1447] ii. 250 bp GFP 삽입 리포터 라인: GFP 코딩 서열을 복원; 또는

[1448] iii. mCherry 삽입 리포터 라인: mCherry 코딩 서열을 이용하여 GFP 코딩 서열을 중단.

[1449] 트랜스-유전자 변형 시스템 구성의 게놈-편집 능력을 평가하기 위해, ttRNA 및 gRNA의 형질감염/전기천공 4 내지 10일 후 유세포 분석법으로 세포를 분석한다. 편집의 충실도는 ttRNA 및 gRNA의 형질감염/전기천공 후 3 내지 10일 후에 게놈 DNA를 수집하여 평가된다. 표적 유전자좌에서 의도된 돌연변이 대 의도하지 않은 돌연변이의 빈도는 암플리콘 시퀀싱에 의해 분석된다.

[1450] 방법 2:

[1451] 3-구성요소 트랜스-유전자 변형 시스템 구성의 평가:

[1452] 트랜스-유전자 변형 복합체의 다양한 구성의 게놈-편집 능력을 결정하기 위해, 다음 중 하나(표 42에서와 같음)를 발현하는 게놈 랜딩 패드를 보유하는 포유동물 HEK293T 또는 U2OS 세포주를 생성한다:

[1453] i. BFP-GFP SNP 리포터 라인: 단일 염기쌍 치환을 통해 GFP-발현 세포주로 전환되는 BFP-발현 세포주;

[1454] ii. 150 또는 250 bp GFP 삽입 리포터 라인: GFP-발현이 결실된 코딩 서열의 150 또는 250 bp 삽입을 통해 복원되는 비-형광 부분-결실-GFP-세포주; 또는

[1455] iii. mKate2 또는 mCherry 삽입 리포터 라인: mKate2 또는 mCherry 코딩 서열을 도입하는 >700(예를 들어, >700 또는 >750) bp 삽입에 의한 GFP 코딩 서열의 중단을 통해 GFP-발현이 손실되고 mKate2 또는 mCherry 발현을 얻는 GFP-발현 세포주.

[1456] 유전자 변형 폴리펩티드는 DNA 플라스미드 또는 mRNA의 형질감염에 의해, 또는 DNA 플라스미드를 렌티바이러스에 패키징함으로써 이들 세포에 도입된다. 세포는 다음을 하도록 함께 설계된 ttRNA-발현 DNA 플라스미드 또는 ncRNA, 및 gRNA-발현 DNA 플라스미드 또는 ncRNA로 추가로 공동 형질감염/전기천공된다:

[1457] i. BFP-GFP SNP 리포터 라인: BFP를 GFP로 전환;

[1458] ii. 150 또는 250 bp GFP 삽입 리포터 라인: GFP 코딩 서열을 복원; 또는

[1459] iii. mKate2 또는 mCherry 삽입 리포터 라인: mKate2 또는 mCherry 코딩 서열을 이용하여 GFP 코딩 서열을 중단.

[1460] 트랜스-유전자 변형 시스템 구성의 게놈-편집 능력을 평가하기 위해, ttRNA 및 gRNA의 형질감염/전기천공 4 내지 10일 후 유세포 분석법으로 세포를 분석한다. 편집의 충실도는 ttRNA 및 gRNA의 형질감염/전기천공 후 3 내지 10일 후에 게놈 DNA를 수집하여 평가된다. 표적 유전자좌에서 의도된 돌연변이 대 의도하지 않은 돌연변이의 빈도는 암플리콘 시퀀싱에 의해 분석된다.

[1461] **실시예 2: Cas9-복합체에 의한 제2 DNA 가닥의 결합을 통한 RNA-기반 유전자 변형을 위한 RNA 주형의 트랜스-동원**

[1462] 이 실시예는 인간 게놈 내 서열의 표적화된 편집을 위한 4 내지 5개 구성요소 RNA-기반 유전자 변형 시스템의 사용을 설명한다(도 4a 내지 4b). 더 구체적으로, (1) 유전자 변형 폴리펩티드는 (2) 가이드 RNA 분자 및 (3) 트랜스-주형 RNA 분자에 결합하여, 함께 유전자 변형 복합체를 형성하며, 이는 표적 유전자좌에 결합하여 닉을 생성하고, 게놈 DNA의 하나의 가닥에서 원하는 편집물의 주형 통합을 수행한다. 추가의 (4) Cas9(예를 들어, dCas9 또는 nCas9 도메인)(도 4a) 또는 제2 유전자 변형 폴리펩티드(도 4b) 분자의 일부인 그러한 도메인은 트랜스-주형 RNA에 의해(예를 들어, 5' 프로-스페이서를 함유하는 말단 블록에 의해), 또는 (5) 이의 종-일치 가이드 RNA에 의해 닉이 생긴 DNA로 동원된다. 제2 Cas9-복합체는 버블을 확장하고 대규모 삽입을 촉진할 수 있으며, 또한 제2 가닥에 닉을 도입하여 제2 가닥 합성을 개시하고/하거나 편집된 가닥이 카피되고 유지되어야 하는 세포의 내인성 복구 시스템(예를 들어, 미스매치 복구 시스템)에 신호를 전달하는 데 사용될 수 있다. 이 실시예는 인간 세포에서의 편집 활성에 대한 다양한 유전자 변형 시스템 구성 및 트랜스-주형 RNA 동원 상호작용의 효능을 평가하는 수단으로 시험관 내 유전자 변형을 위해 포유동물 세포에 4 내지 5개 구성요소 시스템을 도입하는 것을 추가로 설명한다.

[1463] 1) 유전자 변형 폴리펩티드:

[1464] 유전자 변형 폴리펩티드는 다음을 포함한다:

[1465] 1. 하나의 엔도뉴클레아제 도메인이 비활성화된 Cas-뉴클레아제(예를 들어, Spy N863A Cas9),

[1466] 2. 역전사효소(RT), 및

[1467] 3. RNA 결합 단백질(RBP; 표 31에서와 같음). RBP는 개재 펩티드 링커 서열과 함께 또는 없이 1 내지 5회 반복될 수 있다.

[1468] 이들 3개 도메인은 (1) 표 32에서와 같이 펩티드 링커, 또는 (2) 표 33에서와 같이 인테인 쌍을 사용하는 공유 결합에 의해, 또는 (a) 표 34의 화학적 결합 또는 특정 광 파장에 의해 유도되거나, (b) 표 35 및 36의 리간드가 없는 상호작용에서 본질적으로 발생하는, 두 융합 도메인(FD)의 이량체화에 의해 하나로 합쳐진다. FD는 일부 경우에 개재 펩티드 링커 서열과 함께 또는 없이 1 내지 30회 반복될 수 있다.

[1469] 3개의 도메인은 다음 구성 중 임의의 것으로 조립될 수 있으며, 여기서 하기 표시된 링커는 일부 경우에 인테인- 또는 FD-쌍으로 대체될 수 있다.:

[1470] · Cas9 - 링커 - RT - 링커 - $(RBP)_n$

[1471] · RT - 링커 - Cas9 - 링커 - $(RBP)_n$

[1472] · $(RBP)_n$ - 링커 - Cas9 - 링커 - RT

[1473] · $(RBP)_n$ - 링커 - RT - 링커 - Cas9

[1474] · Cas9 - 링커 $(RBP)_n$ - 링커 - RT

[1475] · RT - 링커 - $(RBP)_n$ - 링커 - Cas9

[1476] · Cas9 - 링커 - RT

[1477] 2) 가이드 RNA:

[1478] 가이드 RNA(gRNA) 분자(5'-GCCGAAGCACTGCACGCCGTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC -3'; 서열번호 16,701)는 Cas9에 결합하고 표적 유전자좌에 상보적인 프로-스페이서 서열을 함유하여, 표적 게놈 유전자좌로 유전자 변형 복합체의 국재화를 초래한다.

[1479] 3) 트랜스-주형 RNA:

[1480] 트랜스-주형 RNA(ttRNA) 분자는 하기와 같이 3'에서 5'으로 **PBS-주형** 영역을 함유한다:

[1481] 1. 닉이 생긴 DNA 가닥과 염기쌍을 형성하여 닉이 생긴 DNA의 프라이머 연장을 가능하게 하는 1 내지 17

nt 프라이머 결합 부위(PBS; 표 37에서와 같음), 바로 이어서

[1482] 2. 표적 유전자좌와 상동성인 0 내지 20 nt 영역(상동성 1; 표 38의 "상동성 1"에서와 같음)

[1483] 3. 게놈에 대한 원하는 변형 - 삽입, 결실 또는 치환(편집; 표 38의 "편집"에서와 같음)

[1484] 4. 표적 유전자좌와 상동성인 0 내지 500 nt 영역(상동성 2; 표 39의 "상동성 2"에서와 같음).

[1485] 예시적인 ttRNA는 추가로 다음을 포함한다:

[1486] 5. RBP 동원 부위(**RRS**; 표 40에서와 같음). RRS는 개재 RNA 링커 서열과 함께 또는 없이 1 내지 5회(예를 들어, 1 내지 4회) 반복될 수 있다.

[1487] 예시적인 ttRNA는 또한 선택적으로 다음을 포함할 수 있다:

[1488] 6. 엑소뉴클레아제-매개 분해로부터 ttRNA 분자를 보호하는 말단-보호 RNA 2차 구조(**말단 블록**; 표 41에서와 같음).

[1489] ttRNA의 이러한 구성요소는 하기 구성 중 임의의 것으로 조립될 수 있으며, 여기서 RRS는 최대 20 nt의 RNA 링커 서열(가장 긴 서열은 5'-ACTAACATACAACTAACATA-3'임; 서열번호 16,702)에 의해 PBS-주형에 연결될 수 있다:

[1490] · $(RRS)_n$ - 링커 - PBS-주형

[1491] · $(RRS)_n$ - 링커 - PBS-주형 - 말단 블록

[1492] · PBS-주형 - 링커 - $(RRS)_n$

[1493] · 말단 블록 - PBS-주형 - 링커 - $(RRS)_n$

[1494] · 말단 블록 - $(RRS)_n$ - 링커 - PBS-주형

[1495] · 말단 블록 - $(RRS)_n$ - 링커 - PBS-주형 - 말단 블록

[1496] · PBS-주형 - 링커 - $(RRS)_n$ - 말단 블록

[1497] · 말단 블록 - PBS-주형 - 링커 - $(RRS)_n$ - 말단 블록

[1498] 4) 추가의 (제2) Cas9:

[1499] ttRNA에 인코딩되거나(도 4b), 별도의 가이드 RNA로서 도입되는(도 4a; 또한 하기 5 참조) gRNA 스캐폴드와 쌍을 형성함으로써 닉이 생긴 DNA 가닥에 결합하는, 추가의 닉카제(예를 들어, Spy N863A Cas9, 또는 다른 오르소고날 Cas9) 또는 촉매적으로 비활성인("사멸") Cas9. 전체 스페이서와 쌍을 형성한 닉카제 Cas9는 제2의 닉이 생기지 않은 가닥에 닉을 도입하는 데 사용된다. 대안적으로, 표적에 대해 17 nt 미만의 상보성을 갖는 프로-스페이서를 갖는 gRNA-2와 쌍을 형성하는 사멸 Cas9 또는 닉카제는 제2 가닥에 닉을 생성하지 않으면서, 긴 RT 반응을 위해 DNA 버블을 개방한다. 추가의 닉카제 또는 촉매적으로 비활성인 Cas9는 별도의 도메인(예를 들어, 도 4a) 또는 추가 (제2) 유전자 변형 폴리펩티드의 일부(도 4b)일 수 있다.

[1500] 5) 선택적인 제2 Cas9-쌍 가이드RNA:

[1501] 표 44에서와 같이 닉이 생긴 가닥에 상보적인 프로-스페이서 서열을 갖는 제2 가이드 RNA(gRNA-2) 분자는 DNA 버블 연장 또는 제2 가닥 닉 생성을 위한, 닉이 생긴 DNA 가닥에 제2 Cas9를 동원하도록 첨가될 수 있다. gRNA-2 프로-스페이서는 표적 가닥에 완전히 상보적이어서 제2 Cas9의 닉카제 활성을 자극할 수 있거나, 제2 Cas9에 의한 닉 생성을 억제하기 위해 17개 이하의 상보성 뉴클레오티드를 함유할 수 있다.

[1502] 4 내지 5 구성요소 트랜스-유전자 변형 시스템 구성의 평가:

[1503] 방법 1:

[1504] 트랜스-유전자 변형 시스템 복합체의 다양한 구성의 게놈-편집 능력을 결정하기 위해, 다음 중 하나(표 42에서와 같음)를 발현하는 게놈 랜딩 패드를 보유하는 포유동물 HEK293T 또는 U2OS 세포주를 생성한다:

[1505] i. BFP-GFP SNP 리포터 라인: 단일 염기쌍 치환을 통해 GFP-발현 세포주로 전환되는 BFP-발현 세포주;

[1506] ii. 250 bp GFP 삽입 리포터 라인: GFP-발현이 결실된 코딩 서열의 250 bp 삽입을 통해 복원되는 비-형광 부분-결실-GFP-세포주; 또는

[1507] iii. mCherry 삽입 리포터 라인: mCherry 코딩 서열을 도입하는 >750 bp 삽입에 의한 GFP 코딩 서열의 중단을 통해 GFP-발현이 손실되고 mCherry 발현을 얻는 GFP-발현 세포주.

[1508] 유전자 변형 폴리펩티드는 DNA 플라스미드 또는 mRNA의 형질감염에 의해, 또는 DNA 플라스미드를 렌티바이러스에 패키징함으로써 이들 세포에 도입된다. 세포는 다음을 하도록 함께 설계된 1) ttRNA-발현 DNA 플라스미드 또는 ncRNA, 2) gRNA-발현 DNA 플라스미드 또는 ncRNA로 추가로 공동 형질감염/전기천공된다:

[1509] i. BFP-GFP SNP 리포터 라인: BFP를 GFP로 전환;

[1510] ii. 250 bp GFP 삽입 리포터 라인: GFP 코딩 서열을 복원; 또는

[1511] iii. mCherry 삽입 리포터 라인: mCherry 코딩 서열을 이용하여 GFP 코딩 서열을 중단,

[1512] 3) 제2 Cas9를 발현하는 DNA 플라스미드 또는 mRNA, 및 선택적으로 4) gRNA-2-발현 DNA 플라스미드 또는 ncRNA와 함께.

[1513] 트랜스-유전자 변형 시스템 구성의 게놈-편집 능력을 평가하기 위해, ttRNA 및 gRNA의 형질감염/전기천공 4 내지 10일 후 유세포 분석법으로 세포를 분석한다. 편집의 충실도는 ttRNA 및 gRNA의 형질감염/전기천공 후 3 내지 10일 후에 게놈 DNA를 수집하여 평가된다. 표적 유전자좌에서 의도된 돌연변이 대 의도하지 않은 돌연변이의 빈도는 암플리콘 시퀀싱에 의해 분석된다.

[1514] 방법 2:

[1515] 트랜스-유전자 변형 시스템 복합체의 다양한 구성의 게놈-편집 능력을 결정하기 위해, 다음 중 하나(표 42에서와 같음)를 발현하는 게놈 랜딩 패드를 보유하는 포유동물 HEK293T 또는 U2OS 세포주를 생성한다:

[1516] i. BFP-GFP SNP 리포터 라인: 단일 염기쌍 치환을 통해 GFP-발현 세포주로 전환되는 BFP-발현 세포주;

[1517] ii. 150 또는 250 bp GFP 삽입 리포터 라인: GFP-발현이 결실된 코딩 서열의 150 또는 250 bp 삽입을 통해 복원되는 비-형광 부분-결실-GFP-세포주; 또는

[1518] iii. mKate2 또는 mCherry 삽입 리포터 라인: mKate2 또는 mCherry 코딩 서열을 도입하는 >700(예를 들어, >750) bp 삽입에 의한 GFP 코딩 서열의 중단을 통해 GFP-발현이 손실되고 mKate2 또는 mCherry 발현을 얻는 GFP-발현 세포주.

[1519] 유전자 변형 폴리펩티드는 DNA 플라스미드 또는 mRNA의 형질감염에 의해, 또는 DNA 플라스미드를 렌티바이러스에 패키징함으로써 이들 세포에 도입된다. 세포는 다음을 하도록 함께 설계된 1) ttRNA-발현 DNA 플라스미드 또는 ncRNA, 2) gRNA-발현 DNA 플라스미드 또는 ncRNA로 추가로 공동 형질감염/전기천공된다:

[1520] iv. BFP-GFP SNP 리포터 라인: BFP를 GFP로 전환;

[1521] v. 150 또는 250 bp GFP 삽입 리포터 라인: GFP 코딩 서열을 복원; 또는

[1522] vi. mKate2 또는 mCherry 삽입 리포터 라인: mKate2 또는 mCherry 코딩 서열을 이용하여 GFP 코딩 서열을 중단,

[1523] 3) 제2 Cas9를 발현하는 DNA 플라스미드 또는 mRNA, 및 선택적으로 4) gRNA-2-발현 DNA 플라스미드 또는 ncRNA와 함께.

[1524] 트랜스-유전자 변형 시스템 구성의 게놈-편집 능력을 평가하기 위해, ttRNA 및 gRNA의 형질감염/전기천공 4 내지 10일 후 유세포 분석법으로 세포를 분석한다. 편집의 충실도는 ttRNA 및 gRNA의 형질감염/전기천공 후 3 내지 10일 후에 게놈 DNA를 수집하여 평가된다. 표적 유전자좌에서 의도된 돌연변이 대 의도하지 않은 돌연변이의 빈도는 암플리콘 시퀀싱에 의해 분석된다.

[1525] **실시예 3: 역전사효소의 제2 Cas9-매개 동원**

[1526] 이 실시예는 인간 게놈 내 서열의 표적화된 편집을 위한 4 또는 5개 구성요소 RNA 유전자 변형 시스템의 사용을 설명한다(도 5a 내지 5b). 더 구체적으로, (1) 내지 (2) 스플릿 유전자 변형 폴리펩티드는 (3) 내지 (4) 1 또는 2개 가이드 RNA 분자 및 (5) 트랜스-주형 RNA 분자에 결합하여, 함께 유전자 변형 복합체를 형성하며, 이는 표

적 유전자좌에 결합하여 닉을 생성하고 게놈 DNA의 하나의 가닥에서 원하는 편집물의 주형 통합을 수행한다. 스플릿 유전자 변형 폴리펩티드의 제1 구성요소는 RNA 결합 단백질에 융합된 닉카제 Cas9로 구성되며, 이는 트랜스-주형 RNA 분자에 결합하여 동원한다. 제2 구성요소는 역전사효소 도메인에 융합된 제2 Cas9로 구성되며, 이는 닉이 생성된 DNA 가닥에 원하는 편집물의 주형 통합을 수행한다. Cas9-RT는 트랜스-주형 RNA에 의해, 또는 별도의 가이드 RNA 분자에 의해 표적 유전자좌로 동원된다. 제2 Cas9-복합체의 존재는 버블을 확장하고 대규모 삽입을 촉진하는 역할을 할 수 있으며, 또한 제2 가닥에 닉을 도입하여 제2 가닥 합성을 개시하고/하거나 편집된 가닥이 카피되고 유지되어야 하는 세포의 내인성 미스매치 복구 시스템에 신호를 전달하는 데 사용될 수 있다. 이 실시예는 인간 세포에서의 편집 활성에 대한 다양한 유전자 변형 시스템 구성 및 트랜스-주형 RNA 동원 상호작용의 효능을 평가하는 수단으로 시험관 내 유전자 변형을 위해 포유동물 세포에 4 내지 5개 구성요소 시스템을 도입하는 것을 추가로 설명한다.

[1527] 1 내지 2) 유전자 변형 폴리펩티드:

[1528] 유전자 변형 효소는 2개의 폴리펩티드 구성요소, 즉 Cas9-RNA 결합 단백질(RBP) 구성요소, 및 Cas9-역전사효소(RT) 구성요소를 포함한다. RBP는 일부 경우에 개재 펩티드 링커 서열과 함께 또는 없이 1 내지 5회 반복될 수 있다.

[1529] 이들 2개 구성요소 내의 이들 도메인은 (1) 표 32에서와 같이 펩티드 링커, 또는 (2) 표 33에서와 같이 인테인 쌍을 사용하는 공유 결합에 의해, 또는 , 또는 (a) 표 34의 화학적 결합 또는 특정 광 파장에 의해 유도되거나, (b) 표 35 및 36의 리간드가 없는 상호작용에서 본질적으로 발생하는, 두 융합 도메인(FD)의 이량체화에 의해 하나로 합쳐진다. FD는 일부 경우에 개재 펩티드 링커 서열과 함께 또는 없이 1 내지 30회 반복될 수 있다.

[1530] Cas9-RBP 구성요소는 하기 구성 중 하나로 조립될 수 있으며, 여기서 하기 표시된 링커는 일부 경우에 인테인- 또는 FD-쌍으로 대체될 수 있다.:

[1531] · Cas9 - 링커 - $(RBP)_n$

[1532] · $(RBP)_n$ - 링커 - Cas9

[1533] 제2 Cas9-RT 구성요소는 하기 구성 중 하나로 조립될 수 있으며, 여기서 하기 표시된 링커는 일부 경우에 인테인- 또는 FD-쌍으로 대체될 수 있다.:

[1534] · Cas9 - 링커 - RT

[1535] · RT - 링커 - Cas9

[1536] 3 내지 4) 가이드 RNA:

[1537] 제1 가이드 RNA(gRNA) 분자(5'-GCCGAAGCACTGCACGCCGTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC -3'; 서열번호 16,703)는 Cas9에 결합하고 표적 유전자좌에 상보적인 프로-스페이서 서열을 함유하여, 표적 게놈 유전자좌로 Cas9-RBP 구성요소의 국재화를 초래한다.

[1538] 선택적인 제2 가이드 RNA(gRNA-2) 분자는 Cas9-RT가 트랜스-주형 RNA의 5'-말단에서 가이드 RNA 구조에 의해 제2 가닥으로 동원되지 않는 경우 추가될 수 있다.

[1539] 5) 트랜스-주형 RNA:

[1540] 예시적인 트랜스-주형 RNA(ttRNA) 분자는 하기와 같이 3'에서 5'으로 **PBS-주형** 영역을 함유한다:

[1541] 1. 닉이 생긴 DNA 가닥과 염기쌍을 형성하여 닉이 생긴 DNA의 프라이머 연장을 가능하게 하는 1 내지 17 nt 프라이머 결합 부위(PBS; 표 37에서와 같음), 바로 이어서

[1542] 2. 표적 유전자좌와 상동성인 0 내지 20 nt 영역(상동성 1; 표 38의 "상동성 1"에서와 같음)

[1543] 3. 게놈에 대한 원하는 변형 - 삽입, 결실 또는 치환(편집; 표 38의 "편집"에서와 같음)

[1544] 4. 표적 유전자좌와 상동성인 0 내지 500 nt 영역(상동성 2; 표 39의 "상동성 2"에서와 같음).

[1545] 예시적인 ttRNA는 추가로 다음을 포함한다:

[1546] 5. RBP 동원 부위(**RRS**; 표 40에서와 같음). RRS는 개재 RNA 링커 서열과 함께 또는 없이 1 내지 5회 반복

될 수 있다.

[1547] 예시적인 ttRNA는 또한 선택적으로 다음을 포함할 수 있다:

[1548] 6. 엑소뉴클레아제-매개 분해로부터 ttRNA 분자를 보호하는 말단-보호 RNA 2차 구조(**말단 블록**; 표 41에서와 같음).

[1549] ttRNA의 이러한 구성요소는 하기 구성 중 임의의 것으로 조립될 수 있으며, 여기서 RRS는 최대 20 nt의 RNA 링커 서열(가장 긴 서열은 5'-ACTAACATACAACTAACATA-3'임)에 의해 PBS-주형에 연결될 수 있다:

[1550] · $(RRS)_n$ - 링커 - PBS-주형

[1551] · $(RRS)_n$ - 링커 - PBS-주형 - 말단 블록

[1552] · PBS-주형 - 링커 - $(RRS)_n$

[1553] · 말단 블록 - PBS-주형 - 링커 - $(RRS)_n$

[1554] · 말단 블록 - $(RRS)_n$ - 링커 - PBS-주형

[1555] · 말단 블록 - $(RRS)_n$ - 링커 - PBS-주형 - 말단 블록

[1556] · PBS-주형 - 링커 - $(RRS)_n$ - 말단 블록

[1557] · 말단 블록 - PBS-주형 - 링커 - $(RRS)_n$ - 말단 블록

[1558] 4 내지 5 구성요소 트랜스-유전자 변형 시스템 구성의 평가:

[1559] 트랜스-유전자 변형 복합체의 다양한 구성의 게놈-편집 능력을 결정하기 위해, 다음 중 하나(표 42에서와 같음)를 발현하는 게놈 랜딩 패드를 보유하는 포유동물 HEK293T 또는 U2OS 세포주를 생성한다:

[1560] i. BFP-GFP SNP 리포터 라인: 단일 염기쌍 치환을 통해 GFP-발현 세포주로 전환되는 BFP-발현 세포주;

[1561] ii. 250 bp GFP 삽입 리포터 라인: GFP-발현이 결실된 코딩 서열의 250 bp 삽입을 통해 복원되는 비-형광 부분-결실-GFP-세포주; 또는

[1562] iii. mCherry 삽입 리포터 라인: mCherry 코딩 서열을 도입하는 >750 bp 삽입에 의한 GFP 코딩 서열의 중단을 통해 GFP-발현이 손실되고 mCherry 발현을 얻는 GFP-발현 세포주.

[1563] 유전자 변형 폴리펩티드는 DNA 플라스미드 또는 mRNA의 형질감염에 의해, 또는 DNA 플라스미드를 렌티바이러스에 패키징함으로써 이들 세포에 도입된다. 그 다음 유전자 변형 폴리펩티드를 함유하는 세포는, 다음을 하도록 함께 설계된 1) ttRNA-발현 DNA 플라스미드 또는 ncRNA, 2) gRNA-발현 DNA 플라스미드 또는 ncRNA, 및 선택적으로 3) gRNA-2-발현 DNA 플라스미드 또는 ncRNA로 형질감염/전기천공된다:

[1564] vii. BFP-GFP SNP 리포터 라인: BFP를 GFP로 전환;

[1565] viii. 250 bp GFP 삽입 리포터 라인: GFP 코딩 서열을 복원; 또는

[1566] ix. mCherry 삽입 리포터 라인: mCherry 코딩 서열을 이용하여 GFP 코딩 서열을 중단.

[1567] 트랜스-유전자 변형 시스템 구성의 게놈-편집 능력을 평가하기 위해, ttRNA 및 gRNA의 형질감염/전기천공 4 내지 10일 후 유세포 분석법으로 세포를 분석한다. 편집의 충실도는 ttRNA 및 gRNA의 형질감염/전기천공 후 3 내지 10일 후에 게놈 DNA를 수집하여 평가된다. 표적 유전자좌에서 의도된 돌연변이 대 의도하지 않은 돌연변이의 빈도는 암플리콘 시퀀싱에 의해 분석된다.

[1568] **실시예 4: 인간 세포에서 GFP/BFP 분석을 사용한 유전자 변형 폴리펩티드의 활성 정량화**

[1569] *방법 1*

[1570] 이 실시예는 유전자 변형 폴리펩티드 및 주형 RNA를 함유하는 유전자 변형 시스템의 사용을 설명한다. 이 실시예에서, 주형 RNA는 다음을 함유한다:

[1571] (1) gRNA 스페이서;

[1572] (2) gRNA 스캐폴드;

[1573] (3) 이종 목적 서열; 및

[1574] (4) 프라이머 결합 부위(PBS) 서열.

[1575] 더 구체적으로, 주형 RNA는 다음 서열을 포함한다:

[1576] GCCGAAGCACTGCACGCCGTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGT CCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCACCCTGACGTACGGCGTGCAG TGCTT (서열번호 16,704).

[1577] 주어진 유전자 변형 폴리펩티드(예를 들어, 본 명세서에 기재된 것) 및 주형 RNA를 포함하는 유전자 변형 시스템을 HEK293T BFP-발현 세포주에 형질감염시킨다. 유전자 변형 폴리펩티드 및 주형 RNA를 DNA 형식의 뉴클레오펙션에 의해 전달한다. 구체적으로, 800 ng의 유전자 변형 폴리펩티드 플라스미드 DNA를 200 ng의 주형 RNA와 플라스미드 형식으로 결합한다. 플라스미드 DNA 형식의 변형 폴리펩티드와 주형 RNA를 250,000 HEK293T BFP-발현 세포를 함유하는 25 μL SF 버퍼에 첨가하고, 세포를 프로그램 DS-150을 사용하여 뉴클레오펙션시킨다. 뉴클레오펙션 후, 세포를 37℃, 5% CO_2에서 3일 동안 성장시킨 후, 세포 용해 및 게놈 DNA 추출을 수행한다. 유전자 편집 활성을 분석하기 위해, BFP 유전자좌에 측접하는 프라이머를 사용하여 유전자좌 전체에 걸쳐 증폭시킬 수 있다. Illumina MiSeq를 사용하여 짧은 판독물 시퀀싱을 통해 암플리콘을 분석한다. BFP 유전자 서열의 GFP 유전자로의 전환은 편집이 성공적임을 나타낸다. 일부 구현예에서, 분석은 샘플 중 BFP 유전자 카피의 적어도 10%, 20%, 30%, 40%, 50%, 60%, 또는 70%가 GFP 유전자로 전환된다는 것을 나타낼 것이다.

[1578] *방법 2*

[1579] 이 실시예는 유전자 변형 폴리펩티드 및 주형 RNA를 함유하는 유전자 변형 시스템의 사용을 설명한다. 이 실시예에서, 주형 RNA는 다음을 함유한다:

[1580] (1) gRNA 스페이서;

[1581] (2) gRNA 스캐폴드;

[1582] (3) 이종 목적 서열; 및

[1583] (4) 프라이머 결합 부위(PBS) 서열.

[1584] 더 구체적으로, 주형 RNA는 다음 서열을 포함한다:

[1585] GCCGAAGCACTGCACGCCGTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGT CCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCACCCTGACGTACGGCGTGCAG TGCTT (서열번호 16,705).

[1586] 주어진 유전자 변형 폴리펩티드(예를 들어, 본 명세서에 기재된 것), gRNA, 및 주형 RNA를 포함하는 유전자 변형 시스템을 HEK293T/U2OS BFP-발현 세포주에 형질감염/전기천공시킨다. 유전자 변형 폴리펩티드 및 주형 RNA를 DNA 형식의 뉴클레오펙션에 의해 전달한다. 구체적으로, 300 내지 800 ng의 유전자 변형 폴리펩티드 플라스미드 DNA를 50 내지 200 ng의 주형 RNA 및 50 내지 200 ng의 gRNA와 플라스미드 형식으로 결합한다. 플라스미드 DNA 형식의 변형 폴리펩티드, 주형 RNA 및 gRNA를 250,000 HEK293T/U2OS BFP-발현 세포를 함유하는 25 μL SF/SE 버퍼에 첨가하고, 세포를 프로그램 DS-150/DN-100을 사용하여 뉴클레오펙션시킨다. 뉴클레오펙션 후, 세포를 37℃, 5% CO_2에서 적어도 3일 동안 성장시킨 후, 세포 용해 및 게놈 DNA 추출을 수행한다. 유전자 편집 활성을 분석하기 위해, BFP 유전자좌에 측접하는 프라이머를 사용하여 유전자좌 전체에 걸쳐 증폭시킬 수 있다. Illumina MiSeq를 사용하여 짧은 판독물 시퀀싱을 통해 암플리콘을 분석한다. BFP 유전자 서열의 GFP 유전자로의 전환은 편집이 성공적임을 나타낸다. 일부 구현예에서, 분석은 샘플 중 BFP 유전자 카피의 적어도 10%, 20%, 30%, 40%, 50%, 60%, 또는 70%가 GFP 유전자로 전환된다는 것을 나타낼 것이다.

[1587] **실시예 5: HEK293T & U2OS 세포에서 풀링된 스크리닝에 의한 재작성(rewriter) 폴리펩티드 선택**

[1588] 이 실시예는 인간 게놈 내 코딩 서열의 표적화된 편집을 위한 RNA 유전자 변형 시스템의 사용을 설명한다. 더 구체적으로, 이 실시예는, 예를 들어 풀링된 스크리닝 접근법에 의해 인간 세포에서 활성을 편집하기 위해 새로

운 유전자 변형 폴리펩티드를 평가하는 수단으로서, 유전자 변형 후보의 라이브러리로 HEK293T 및 U2OS 세포를 감염시킨 후 세포에서 시험관 내 유전자 변형을 위한 주형 가이드 RNA(tgRNA)를 형질감염시키는 것을 설명한다.

[1589] 본 명세서에서 분석된 유전자 변형 폴리펩티드 라이브러리 후보는 각각 다음을 포함한다: 1) 하나의 엔도뉴클레아제 활성 부위를 비활성화시키는 N863A 돌연변이를 함유하는 S. 피오게네스(Spy) Cas9 닉카제; 2) **표 10**에 기재된 122개의 펩티드 링커 중 하나; 및 3) 레트로바이러스 기원의 **표 6**의 역전사효소(RT) 도메인. 이용된 특정 레트로바이러스 RT 도메인이 단량체로 기능할 것으로 예상되는 경우 이를 선택하였다. 각각의 선택된 RT 도메인에 대해, 야생형 서열뿐만 아니라, 주요 야생형 서열에 점 돌연변이가 설치된 버전을 시험하였다. 특히, 야생형이거나 다양한 돌연변이를 함유하는 143개의 RT 도메인을 시험하였다. 총 17,446개의 Cas-링커-RT 유전자 변형 폴리펩티드를 시험하였다.

[1590] 본 명세서에 기재된 시스템은 다음을 포함하는 2-구성요소 시스템이다: 1) 렌티바이러스 카세트 내에 인간 코돈-최적화 유전자 변형 폴리펩티드 라이브러리 후보를 인코딩하는 발현 플라스미드, 및 2) Cas에 의해 인식되고 관심 게놈 유전자좌로 국재화하며, 또한 U6 프로모터에 의해 구동되는 RT 도메인에 의해 원하는 편집물을 게놈으로 역전사하는 주형을 제공하는 비-코딩 tgRNA 서열을 발현하는 tgRNA 발현 플라스미드. 렌티바이러스 카세트는 (i) 포유동물 세포에서의 발현을 위한 CMV 프로모터; (ii) 나타낸 바와 같은 유전자 변형 폴리펩티드 라이브러리 후보; (iii) 자가-절단 T2A 폴리펩티드; (iv) 포유동물 세포에서 선택을 가능하게 하는 퓨로마이신 내성 유전자; 및 (v) 폴리A 테일 종결 신호를 포함한다.

[1591] 유전자 변형 폴리펩티드 라이브러리 후보를 발현하는 세포의 풀을 준비하기 위해, HEK293T 또는 U2OS 세포를 유전자 변형 후보 플라스미드 라이브러리의 풀링된 렌티바이러스 제제를 이용하여 형질도입시켰다. HEK293 Lenti-X 세포를 렌티바이러스 플라스미드 형질감염 전에 15 cm 플레이트(약 $12x10^6$개 세포)에 시딩하였다. Lentiviral Packaging Mix(Biosettia, 27 ug) 및 유전자 변형 후보 라이브러리에 대한 플라스미드 DNA(27 ug)를 사용한 렌티바이러스 플라스미드 형질감염을 제조업체의 프로토콜에 따라 Lipofectamine 2000 및 Opti-MEM 배지를 사용하여 다음 날 수행하였다. 다음 날 전체 배지 교체에 의해 세포외 DNA를 제거하였고 바이러스-함유 배지를 48시간 후에 수확하였다. 렌티바이러스 배지를 Lenti-X Concentrator(TaKaRa Biosciences)를 사용하여 농축시켰고 5 mL의 렌티베아러스 분취액을 제조하였으며 -80° C에서 보관하였다. 렌티바이러스 적정은 퓨로마이신 선택 후 콜로니 형성 단위를 계수함으로써 수행하였다. BFP-발현 게놈 랜딩 패드를 보유하는 HEK293T 또는 U2OS 세포를 배양 플레이트에 $6x10^7$개 세포로 시딩하였고 0.3의 감염다중도(MOI)로 형질도입하여 세포당 다중 감염을 최소화하였다. 감염된 세포의 선택을 가능하게 하기 위해 감염 후 48시간 후에 퓨로마이신(2.5 ug/mL)을 첨가하였다. 세포를 적어도 7일 동안 퓨로마이신 선택 하에 유지시킨 다음, tgRNA 전기천공을 위해 규모를 확대시켰다.

[1592] 분석에서 유전자 변형 라이브러리 후보의 게놈-편집 능력을 결정하기 위해, 그 다음 감염된 BFP-발현 HEK293T 또는 U2OS 세포를 라이브러리 후보당 >1000x 커버리지에 대해 충분한 세포 수에서 BFP를 GFP로 전환시키도록 설계된, 200 ng의 tgRNA(g4 또는 g10) 플라스미드를 이용하여 250,000개 세포/웰을 전기천공함으로써 형질감염시켰다.

[1593] g4 tgRNA(5'에서 3')는 다음과 같다: 20개 뉴클레오티드 스페이서 영역(GCCGAAGCACTGCACGCCGT), 스캐폴드 영역(GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC), BFP를 GFP로 변형시키는 단일 염기 치환을 인코딩하는 주형 영역(**굵은 글씨**) 및 기능성 유전자 변형 반응의 완료 시 유전자 변형 폴리펩티드의 재결합을 방지하는 SpyCas9 PAM에 동의어 점 돌연변이(NGG에서 NCG)를 도입한 PAM 비활성화(밑줄)(ACCCTGAC**G**TACG), 및 13 뉴클레오티드 PBS(GCGTGCAGTGCTT).

[1594] 유사하게, g10 tgRNA(5'에서 3')는 다음과 같다: 20개 뉴클레오티드 스페이서 영역(AGAAGTCGTGCTGCTTCATG), 스캐폴드 영역(GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC), BFP를 GFP로 변형시키는 단일 염기 치환을 인코딩하는 주형 영역(**굵은 글씨**) 및 기능성 유전자 변형 반응의 완료 시 유전자 변형 폴리펩티드의 재결합을 방지하는 SpyCas9 PAM에 동의어 점 돌연변이(NGG에서 NGA)를 도입한 PAM 비활성화(밑줄)(ACCCTGACC**T**ACGGCGTGCAGTGCTTCGGCCGCTACCCCGA**T**CACAT), 및 13 뉴클레오티드 PBS(GAAGCAGCACGAC).

[1595] 분석에서 다양한 작제물의 게놈 편집 능력을 분석하기 위해, 전기천공 후 6 내지 7일 후에 GFP 발현을 위해 형광 활성화 세포 분류(FACS)에 의해 세포를 분류하였다. 세포를 편집되지 않은(BFP+) 세포, 편집된 (GFP+) 세포 및 불완전한 편집(BFP-, GFP-) 세포의 별개의 집단으로 분류하고 수확하였다. 분석 동안에 농축을 결정하기 위해 분류되지 않은 세포 샘플을 또한 입력 집단으로 수확하였다.

[1596] 어떤 유전자 변형 라이브러리 후보가 분석에서 게놈 편집 능력을 갖는지를 결정하기 위해, 분류된 세포 집단과

분류되지 않은 세포 집단에서 게놈 DNA(gDNA)를 수확하고, 각각의 집단에서 유전자 변형 라이브러리 후보를 시퀀싱함으로써 분석하였다. 간략하게, 유전자 변형 서열을 렌티바이러스 카세트에 특이적인 프라이머를 사용하여 게놈으로부터 증폭시켰고, 게놈 DNA를 희석하기 위해 PCR의 두 번째 라운드에서 증폭시킨 다음, 제조업체의 프로토콜에 따라 Oxford Nanopore Sequencing Technology를 사용하여 시퀀싱하였다.

[1597] 시퀀싱 판독물의 품질 관리 후, 적어도 1500개 및 3200개 이하의 뉴클레오티드의 판독물을 유전자 변형 폴리펩티드 라이브러리 서열에 맵핑하였고 라이브러리 서열과 최소 약 80% 일치하는 판독물을 주어진 후보에 대해 성공적으로 정렬된 것으로 간주하였다. 분석에서 유전자 편집을 수행할 수 있는 유전자 변형 후보를 확인하기 위해, 편집된 집단에 있는 각각의 라이브러리 후보의 판독 수를 초기의 분류되지 않은 집단에서 판독 수와 비교하였다. 이러한 풀링된 스크린의 목적을 위해, 게놈-편집 능력이 있는 유전자 변형 후보를 분류되지 않은(입력) 세포에 비해 전환된(GFP+) 집단이 농축된 후보로 선택하였으며, 여기서 농축은 참조(요소 ID 번호: 17380)의 농축 수준 이상인 것으로 결정되었다.

[1598] 다수의 유전자 변형 폴리펩티드 후보는 GFP+ 세포 집단이 농축된 것으로 결정되었다. 예를 들어, 시험한 17,446개 후보 중, 3,300개 초과가 표 1에 나타낸 유사한 실험 조건(g4 tgRNA를 사용한 HEK293T; g10 tgRNA를 사용한 HEK293T 세포; 또는 g4 tgRNA를 사용한 U2OS 세포) 하에서 참조와 적어도 동등한 GFP+ 분류 집단(분류되지 않은 집단에 비해)의 농축을 나타내었다. 17,446개의 후보를 또한 g10 tgRNA를 사용한 U2OS 세포에서 시험하였지만, 풀링된 스크린은 실험 조건 하에서 분류되지 않은(입력) 세포에 비해 전환된(GFP+) 집단이 농축된 후보를 산출하지 않았으며; 이러한 결과를 설명하기 위해 추가 조사가 필요하다.

[1599] **표 1.** 스크린에서 순위가 매겨진 성능에 따라 정렬된 링커 및 RT 서열의 조합. 이 표의 각각의 RT의 아미노산 서열은 표 6에 제공되어 있다.

[1600]

링커 아미노산 서열	링커 서열 번호	RT 도메인 명칭
EAAAKGSS	16,901	PERV_Q4VFZ2_3mutA_WS
EAAAKEAAAKEAAAKEAAAK	16,902	MLVMS_P03355_PLV919
PAPEAAAK	16,903	MLVFF_P26809_3mutA
EAAAKPAPGGG	16,904	MLVFF_P26809_3mutA
GSSGSSGSSGSSGSSGSS	16,905	PERV_Q4VFZ2_3mut
PAPGGGEAAAK	16,906	MLVAV_P03356_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	16,907	MLVMS_P03355_PLV919
GSSEAAAK	16,908	MLVFF_P26809_3mutA
EAAAKPAPGGS	16,909	MLVFF_P26809_3mutA
GGSGGSGGSGGSGGSGGS	16,910	MLVFF_P26809_3mutA

AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAK A	16,911	XMRV6_A1Z651_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAK A	16,912	PERV_Q4VFZ2_3mutA_WS
EAAAKEAAAKEAAAK	16,913	MLVFF_P26809_3mutA
PAPEAAAKGSS	16,914	MLVFF_P26809_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAK A	16,915	PERV_Q4VFZ2_3mutA_WS
EAAAKEAAAKEAAAK	16,916	PERV_Q4VFZ2_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAK A	16,917	AVIRE_P03360_3mutA
PAPAPAPAPAP	16,918	MLVCB_P08361_3mutA
PAPAPAPAPAP	16,919	MLVFF_P26809_3mutA
EAAAKGGSPAP	16,920	PERV_Q4VFZ2_3mutA_WS
PAP	16,921	MLVMS_P03355_PLV919
PAPGGGGSS	16,922	WMSV_P03359_3mutA
SGSETPGTSESATPES	16,923	MLVFF_P26809_3mutA
PAPEAAAKGSS	16,924	XMRV6_A1Z651_3mutA
EAAAKGGSGGG	16,925	MLVMS_P03355_PLV919
GGGGSGGGGS	16,926	MLVFF_P26809_3mutA
GGGPAPGSS	16,927	MLVAV_P03356_3mutA
GGSGGSGGSGGSGGSGGS	16,928	XMRV6_A1Z651_3mut
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	16,929	MLVCB_P08361_3mutA
GSSPAP	16,930	AVIRE_P03360_3mutA
EAAAKGSSPAP	16,931	MLVFF_P26809_3mutA
GSSGGGEAAAK	16,932	MLVFF_P26809_3mutA
GGSGGSGGSGGSGGSGGS	16,933	MLVMS_P03355_3mutA_WS
PAPAPAPAP	16,934	MLVFF_P26809_3mutA
EAAAKEAAAKEAAAKEAAAK	16,935	XMRV6_A1Z651_3mutA
EAAAKGGSPAP	16,936	MLVMS_P03355_3mutA_WS
PAPGGSEAAAK	16,937	AVIRE_P03360_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	16,938	AVIRE_P03360_3mutA
EAAAKGGGGSEAAAK	16,939	MLVCB_P08361_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAK A	16,940	WMSV_P03359_3mutA
GSS	16,941	MLVMS_P03355_PLV919
GSSGSSGSSGSS	16,942	MLVMS_P03355_PLV919
GSSPAPEAAAK	16,943	XMRV6_A1Z651_3mutA
GGSPAPEAAAK	16,944	MLVFF_P26809_3mutA
GGGEAAAKGGS	16,945	MLVFF_P26809_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	16,946	PERV_Q4VFZ2_3mutA_WS
GGGGGGGG	16,947	PERV_Q4VFZ2_3mut

[1601]

GGGPAP	16,948	MLVCB_P08361_3mutA
PAPAPAPAPAPAP	16,949	MLVCB_P08361_3mutA
GGSGGSGGSGGSGGSGGS	16,950	MLVCB_P08361_3mutA
PAP	16,951	MLVMS_P03355_3mutA_WS
GGSGGSGGSGGSGGSGGS	16,952	PERV_Q4VFZ2_3mutA_WS
PAPAPAPAPAPAP	16,953	MLVMS_P03355_PLV919
EAAAKPAPGSS	16,954	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAKEAAAK	16,955	MLVMS_P03355_3mutA_WS
EAAAKGGS	16,956	MLVMS_P03355_3mutA_WS
GGGGSEAAAKGGGGS	16,957	MLVFF_P26809_3mutA
EAAAKPAPGSS	16,958	MLVFF_P26809_3mutA
GGGGSGGGGSGGGGSGGGGS	16,959	MLVMS_P03355_PLV919
EAAAKGGGGGS	16,960	MLVMS_P03355_PLV919
GGSPAP	16,961	XMRV6_A1Z651_3mutA
EAAAKGGGPAP	16,962	MLVMS_P03355_PLV919
EAAAKEAAAKEAAAKEAAAKEAAAK	16,963	MLVFF_P26809_3mutA
PAP	16,964	MLVCB_P08361_3mutA
EAAAK	16,965	XMRV6_A1Z651_3mutA
GGSGSSPAP	16,966	PERV_Q4VFZ2_3mutA_WS
GSSGSSGSSGSSGSSGSS	16,967	MLVMS_P03355_PLV919
GSSEAAAKGGG	16,968	MLVAV_P03356_3mutA
GGGEAAAKGGS	16,969	XMRV6_A1Z651_3mutA
EAAAKGGGGSEAAAK	16,970	MLVAV_P03356_3mutA
GGGGSGGGGSGGGGS	16,971	MLVFF_P26809_3mutA
GGGGSGGGGSGGGGSGGGGS	16,972	AVIRE_P03360_3mutA
SGSETPGTSESATPES	16,973	AVIRE_P03360_3mutA
GGGEAAAKPAP	16,974	MLVFF_P26809_3mutA
EAAAKGSSGGG	16,975	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAK	16,976	WMSV_P03359_3mut
GGSGGSGGSGGS	16,977	XMRV6_A1Z651_3mutA
GGSEAAAKPAP	16,978	MLVFF_P26809_3mutA
EAAAKGSSGGG	16,979	XMRV6_A1Z651_3mutA
GGGGS	16,980	MLVFF_P26809_3mutA
GGGEAAAKGSS	16,981	MLVMS_P03355_PLV919
PAPAPAPAPAPAP	16,982	MLVAV_P03356_3mutA
GGGGSGGGGSGGGGSGGGGS	16,983	MLVCB_P08361_3mutA
GGGEAAAKGSS	16,984	MLVCB_P08361_3mutA
PAPGGSGSS	16,985	MLVFF_P26809_3mutA
GSAGSAAGSGEF	16,986	MLVCB_P08361_3mutA
PAPGGSEAAAK	16,987	MLVMS_P03355_3mutA_WS
GGSGSS	16,988	XMRV6_A1Z651_3mutA

[1602]

PAPGGGGSS	16,989	MLVMS_P03355_PLV919
GSSGSSGSS	16,990	XMRV6_A1Z651_3mut
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	16,991	MLVMS_P03355_3mutA_WS
EAAAK	16,992	MLVMS_P03355_PLV919
GSSGSSGSSGSS	16,993	MLVFF_P26809_3mutA
PAPGGGGSS	16,994	MLVCB_P08361_3mutA
GGGEAAAKGGS	16,995	MLVCB_P08361_3mutA
PAPGGGEAAAK	16,996	MLVMS_P03355_PLV919
GGGGGSPAP	16,997	XMRV6_A1Z651_3mutA
EAAAKGGS	16,998	XMRV6_A1Z651_3mutA
EAAAKGSSPAP	16,999	XMRV6_A1Z651_3mut
PAPEAAAK	17,000	MLVAV_P03356_3mutA
GGSGGSGGSGGS	17,001	MLVMS_P03355_3mutA_WS
GGGPAPGGS	17,002	MLVMS_P03355_PLV919
GSSGSSGSSGSS	17,003	PERV_Q4VFZ2_3mutA_WS
EAAAKPAPGGS	17,004	MLVCB_P08361_3mutA
GSSGSS	17,005	MLVFF_P26809_3mutA
EAAAKEAAAKEAAAKEAAAK	17,006	MLVCB_P08361_3mutA
EAAAKEAAAKEAAAKEAAAK	17,007	FLV_P10273_3mutA
GSS	17,008	MLVFF_P26809_3mutA
EAAAKEAAAK	17,009	MLVMS_P03355_3mutA_WS
PAPEAAAKGGG	17,010	MLVAV_P03356_3mutA
GGSGSSEAAAK	17,011	MLVFF_P26809_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,012	PERV_Q4VFZ2
GSSEAAAKPAP	17,013	AVIRE_P03360_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,014	MLVCB_P08361_3mutA
EAAAKGGG	17,015	MLVFF_P26809_3mutA
GSSPAPGGG	17,016	MLVCB_P08361_3mutA
GGGPAPGSS	17,017	MLVMS_P03355_PLV919
GGGGGS	17,018	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,019	PERV_Q4VFZ2_3mut
GGGGSGGGGSGGGGSGGGGSGGGGS	17,020	WMSV_P03359_3mutA
EAAAKEAAAKEAAAK	17,021	PERV_Q4VFZ2_3mut
PAPAPAPAP	17,022	MLVCB_P08361_3mutA
GSSGSSGSSGSSGSS	17,023	PERV_Q4VFZ2_3mut
GGGGSSEAAAK	17,024	MLVMS_P03355_3mutA_WS
GGSGGSGGSGGS	17,025	MLVCB_P08361_3mutA
PAPEAAAKGGS	17,026	MLVCB_P08361_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,027	MLVCB_P08361_3mutA
EAAAKGGGGSEAAAK	17,028	MLVMS_P03355_PLV919
EAAAKGGGGSEAAAK	17,029	MLVMS_P03355_3mutA_WS

[1603]

EAAAKGGGPAP	17,030	XMRV6_A1Z651_3mut
EAAAKEAAAKEAAAKEAAAKEAAAK	17,031	MLVMS_P03355_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,032	FLV_P10273_3mutA
GGSEAAAKGGG	17,033	MLVMS_P03355_3mutA_WS
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,034	KORV_Q9TTC1-Pro_3mutA
GGGPAPGGS	17,035	MLVCB_P08361_3mutA
PAPAPAPAPAPAP	17,036	XMRV6_A1Z651_3mutA
GGSGSSGGG	17,037	XMRV6_A1Z651_3mutA
GGSGSSGGG	17,038	MLVCB_P08361_3mutA
GGGEAAAKGGS	17,039	MLVMS_P03355_3mutA_WS
EAAAK	17,040	MLVCB_P08361_3mutA
GGSPAPGSS	17,041	MLVMS_P03355_3mutA_WS
GGGGSSEAAAK	17,042	PERV_Q4VFZ2_3mut
PAPAPAPAPAP	17,043	MLVBM_Q7SVK7_3mut
EAAAKEAAAKEAAAKEAAAK	17,044	MLVAV_P03356_3mutA
GGGGGSGSS	17,045	MLVCB_P08361_3mutA
EAAAKGSSPAP	17,046	MLVMS_P03355_3mutA_WS
PAPAPAPAPAPAP	17,047	MLVMS_P03355_3mutA_WS
GSSGGGGGS	17,048	MLVMS_P03355_3mutA_WS
PAPGSSGGG	17,049	MLVMS_P03355_PLV919
GGSGGGPAP	17,050	MLVCB_P08361_3mutA
GGGGGGG	17,051	MLVCB_P08361_3mutA
GSSGSSGSSGSSGSSGSS	17,052	MLVCB_P08361_3mutA
GGGPAPGGS	17,053	MLVFF_P26809_3mutA
EAAAKGGSGGG	17,054	PERV_Q4VFZ2_3mut
EAAAKGGGGSS	17,055	MLVMS_P03355_3mutA_WS
GSSGSSGSSGSSGSSGSS	17,056	MLVMS_P03355_3mut
GGGGSGGGGSGGGGSGGGGS	17,057	MLVBM_Q7SVK7_3mutA_WS
PAPAPAPAPAP	17,058	MLVMS_P03355_PLV919
GGGEAAAKGGS	17,059	MLVMS_P03355_PLV919
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,060	MLVMS_P03355_3mut
GSAGSAAGSGEF	17,061	MLVMS_P03355_3mutA_WS
GSSGSSGSSGSSGSS	17,062	MLVFF_P26809_3mutA
EAAAKGGSGSS	17,063	MLVFF_P26809_3mutA
PAPGGG	17,064	MLVFF_P26809_3mutA
GGGPAPGSS	17,065	XMRV6_A1Z651_3mutA
PAPEAAAKGGS	17,066	AVIRE_P03360_3mutA
PAPGGGEAAAK	17,067	MLVFF_P26809_3mut
GGGGSSEAAAK	17,068	MLVCB_P08361_3mutA

[1604]

EAAAK	17,069	MLVMS_P03355_PLV919
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,070	BAEVM_P10272_3mutA
GGSGGGEAAAK	17,071	MLVMS_P03355_PLV919
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,072	MLVFF_P26809_3mutA
GSSPAPGGS	17,073	XMRV6_A1Z651_3mutA
GGSGGGPAP	17,074	MLVMS_P03355_PLV919
EAAAK	17,075	AVIRE_P03360_3mutA
GSS	17,076	XMRV6_A1Z651_3mutA
GGSGGSGGS	17,077	MLVFF_P26809_3mutA
EAAAKEAAAKEAAAKEAAAK	17,078	AVIRE_P03360_3mut
PAPEAAAKGGG	17,079	PERV_Q4VFZ2_3mutA_WS
GGGGGSEAAAK	17,080	BAEVM_P10272_3mutA
GGSGSSGGG	17,081	MLVMS_P03355_3mutA_WS
GGGGGGG	17,082	MLVMS_P03355_3mutA_WS
GSSEAAAKPAP	17,083	PERV_Q4VFZ2_3mut
GGGGGSEAAAK	17,084	WMSV_P03359_3mut
GGGGSGGGGSGGGGSGGGGSGGGGS	17,085	MLVFF_P26809_3mut
GGGEAAAKGGS	17,086	AVIRE_P03360_3mutA
GGSPAPGGG	17,087	AVIRE_P03360_3mutA
GSAGSAAGSGEF	17,088	MLVAV_P03356_3mutA
EAAAK	17,089	MLVAV_P03356_3mutA
EAAAKPAPGSS	17,090	WMSV_P03359_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,091	PERV_Q4VFZ2_3mutA_WS
GGSEAAAKPAP	17,092	MLVCB_P08361_3mutA
PAPAPAPAPAPAP	17,093	MLVBM_Q7SVK7_3mutA_WS
GGSPAPGGG	17,094	MLVMS_P03355_3mutA_WS
GGSEAAAKGGG	17,095	MLVMS_P03355_3mut
GGSGGSGGSGGS	17,096	MLVFF_P26809_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,097	MLVFF_P26809_3mutA
GGG	17,098	AVIRE_P03360_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,099	PERV_Q4VFZ2_3mut
GGSGGSGGSGGS	17,100	MLVMS_P03355_3mutA_WS
GGGEAAAK	17,101	MLVCB_P08361_3mutA
GSSGSSGSSGSSGSSGSS	17,102	MLVMS_P03355_3mutA_WS
GSSGGGPAP	17,103	MLVMS_P03355_3mutA_WS
GSSEAAAKPAP	17,104	MLVFF_P26809_3mutA
EAAAKEAAAK	17,105	MLVMS_P03355_PLV919
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,106	MLVCB_P08361_3mut
GGGGGG	17,107	MLVMS_P03355_3mutA_WS

[1605]

GGSGSSGGG	17,108	MLVFF_P26809_3mutA
GSSGGGEAAAK	17,109	PERV_Q4VFZ2_3mutA_WS
PAPAPAPAPAP	17,110	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,111	SFV3L_P27401_2mut
EAAAKGGSGGG	17,112	BAEVM_P10272_3mutA
GGGGSSPAP	17,113	PERV_Q4VFZ2_3mutA_WS
GGGEAAAKPAP	17,114	MLVMS_P03355_PLV919
GGSGGGPAP	17,115	BAEVM_P10272_3mutA
PAPGSSGGS	17,116	MLVMS_P03355_PLV919
GGSGGGPAP	17,117	MLVMS_P03355_3mutA_WS
EAAAKGGSPAP	17,118	PERV_Q4VFZ2_3mutA_WS
EAAAKGGSGGG	17,119	MLVMS_P03355_3mutA_WS
PAPGSSGGG	17,120	MLVFF_P26809_3mutA
GSSEAAAKGGS	17,121	MLVFF_P26809_3mutA
PAPGSSEAAAK	17,122	MLVFF_P26809_3mutA
EAAAKGSSPAP	17,123	KORV_Q9TTC1-Pro_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,124	MLVBM_Q7SVK7_3mutA_WS
PAPGSSEAAAK	17,125	MLVMS_P03355_PLV919
EAAAKGSSGGG	17,126	MLVMS_P03355_3mutA_WS
EAAAKGGGGGS	17,127	AVIRE_P03360_3mutA
EAAAKEAAAKEAAAK	17,128	MLVMS_P03355_PLV919
PAPAPAPAPAPAP	17,129	MLVFF_P26809_3mutA
GGGGSGGGGSGGGGS	17,130	MLVCB_P08361_3mutA
PAPGGSEAAAK	17,131	MLVCB_P08361_3mutA
PAPGSSEAAAK	17,132	MLVBM_Q7SVK7_3mutA_WS
PAPEAAAKGSS	17,133	AVIRE_P03360_3mutA
GGSPAPGSS	17,134	WMSV_P03359_3mutA
PAPGGSGGG	17,135	MLVMS_P03355_PLV919
EAAAKGGSGSS	17,136	MLVMS_P03355_3mutA_WS
GGSGGG	17,137	MLVFF_P26809_3mutA
GGSEAAAKGSS	17,138	KORV_Q9TTC1_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,139	MLVCB_P08361_3mutA
PAPAPAPAPAPAP	17,140	PERV_Q4VFZ2_3mutA_WS
PAPEAAAK	17,141	MLVMS_P03355_3mutA_WS
GGSEAAAKGGG	17,142	MLVMS_P03355_PLV919
GSSPAP	17,143	MLVMS_P03355_3mutA_WS
GGGGSS	17,144	MLVMS_P03355_PLV919
GGGEAAAKPAP	17,145	AVIRE_P03360_3mutA
EAAAKPAPGGS	17,146	MLVAV_P03356_3mutA

[1606]

EAAAKGGGPAP	17,147	MLVAV_P03356_3mutA
PAPGGSEAAAK	17,148	BAEVM_P10272_3mutA
PAPGGSGSS	17,149	MLVMS_P03355_3mutA_WS
PAPGGSGSS	17,150	AVIRE_P03360_3mutA
GGSGGGPAP	17,151	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAKEAAAK	17,152	BAEVM_P10272_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGS	17,153	MLVMS_P03355_PLV919
GGGGSSPAP	17,154	MLVCB_P08361_3mutA
GSSGGGPAP	17,155	MLVFF_P26809_3mutA
GGGGSSGGS	17,156	MLVMS_P03355_PLV919
GGSGGG	17,157	MLVCB_P08361_3mutA
GSSGGGGGS	17,158	MLVMS_P03355_PLV919
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,159	XMRV6_A1Z651_3mutA
GGGGGSGSS	17,160	KORV_Q9TTC1_3mut
GGGEAAAKGGS	17,161	BAEVM_P10272_3mutA
GGSGGG	17,162	BAEVM_P10272_3mutA
PAPAPAP	17,163	KORV_Q9TTC1-Pro_3mut
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,164	SFV3L_P27401_2mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,165	MLVBM_Q7SVK7_3mutA_WS
GSSGSSGSSGSSGSS	17,166	MLVMS_P03355_3mutA_WS
GSSGGGEAAAK	17,167	MLVMS_P03355_3mutA_WS
GSSGGSEAAAK	17,168	MLVFF_P26809_3mutA
PAP	17,169	MLVMS_P03355_PLV919
EAAAKGGGGSEAAAK	17,170	MLVBM_Q7SVK7_3mutA_WS
PAPAP	17,171	AVIRE_P03360_3mutA
PAP	17,172	MLVFF_P26809_3mutA
GSSGGG	17,173	MLVMS_P03355_3mut
GSSPAPGGS	17,174	MLVFF_P26809_3mutA
PAPAPAPAP	17,175	XMRV6_A1Z651_3mutA
EAAAKGSSGGS	17,176	PERV_Q4VFZ2_3mut
PAPEAAAKGGG	17,177	KORV_Q9TTC1-Pro_3mutA
PAPGGS	17,178	MLVCB_P08361_3mutA
EAAAKGGG	17,179	MLVCB_P08361_3mutA
GSSEAAAKPAP	17,180	MLVMS_P03355_PLV919
PAPGGS	17,181	MLVFF_P26809_3mutA
EAAAKGGS	17,182	MLVCB_P08361_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,183	FLV_P10273_3mutA
PAPGGSEAAAK	17,184	MLVAV_P03356_3mutA
GSS	17,185	MLVCB_P08361_3mutA

[1607]

[1608]

GSSGSSGSSGSS	17,186	AVIRE_P03360_3mutA
GSSGSSGSS	17,187	MLVFF_P26809_3mutA
GSSGGG	17,188	MLVMS_P03355_PLV919
EAAAK	17,189	MLVFF_P26809_3mutA
GGSPAPEAAAK	17,190	MLVCB_P08361_3mutA
GGSGSS	17,191	MLVCB_P08361_3mutA
GSSPAPGGG	17,192	MLVMS_P03355_PLV919
EAAAKEAAAKEAAAKEAAAKEAAAK	17,193	MLVAV_P03356_3mutA
EAAAKGSSPAP	17,194	FLV_P10273_3mutA
GGGGSS	17,195	XMRV6_A1Z651_3mutA
GGSPAPGSS	17,196	MLVMS_P03355_PLV919
EAAAKEAAAKEAAAKEAAAKEAAAK	17,197	MLVMS_P03355_3mutA_WS
PAPEAAAKGGG	17,198	FLV_P10273_3mutA
EAAAKPAPGGS	17,199	XMRV6_A1Z651_3mut
PAPAP	17,200	BAEVM_P10272_3mutA
EAAAKEAAAKEAAAKEAAAK	17,201	MLVMS_P03355_PLV919
GSSPAPGGG	17,202	MLVMS_P03355_PLV919
EAAAKGGGPAP	17,203	KORV_Q9TTC1_3mutA
PAPEAAAK	17,204	MLVMS_P03355_PLV919
PAPGGGEAAAK	17,205	PERV_Q4VFZ2_3mutA_WS
EAAAKGSSGGS	17,206	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAK	17,207	MLVMS_P03355_PLV919
GSSEAAAK	17,208	MLVMS_P03355_3mutA_WS
GSSGSSGSSGSS	17,209	MLVMS_P03355_3mutA_WS
GGGGSGGGGSGGGGSGGGGS	17,210	MLVMS_P03355_3mutA_WS
EAAAKGGGGSEAAAK	17,211	MLVMS_P03355_3mut
GGS	17,212	MLVCB_P08361_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,213	XMRV6_A1Z651_3mutA
GGSGSSPAP	17,214	MLVCB_P08361_3mutA
GGGGSGGGGSGGGGS	17,215	XMRV6_A1Z651_3mutA
PAPAPAPAPAP	17,216	BAEVM_P10272_3mutA
PAPAPAPAPAP	17,217	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAKEAAAK	17,218	MLVBM_Q7SVK7_3mut
GGGGSGGGGSGGGGSGGGGSGGGGS	17,219	BAEVM_P10272_3mutA
GGSGGSGGS	17,220	MLVMS_P03355_3mutA_WS
EAAAKPAPGSS	17,221	MLVMS_P03355_PLV919
GSS	17,222	MLVMS_P03355_3mutA_WS
PAPEAAAKGGS	17,223	MLVMS_P03355_3mutA_WS
GGGPAPGGS	17,224	MLVMS_P03355_3mutA_WS
EAAAKGGGGSS	17,225	MLVAV_P03356_3mutA
GSSGSSGSSGSSGSS	17,226	MLVFF_P26809_3mut

SGSETPGTSESATPES	17,227	PERV_Q4VFZ2_3mut
GGSEAAAKGGG	17,228	MLVMS_P03355_3mut
GSSGSSGSSGSSGSSGSS	17,229	AVIRE_P03360_3mutA
PAPAPAPAPAPAP	17,230	AVIRE_P03360_3mut
GGSGGS	17,231	XMRV6_A1Z651_3mutA
PAPGSSEAAAK	17,232	MLVCB_P08361_3mut
GGSPAPEAAAK	17,233	PERV_Q4VFZ2_3mut
EAAAKGGGGGS	17,234	MLVCB_P08361_3mutA
GGSGGSGGSGGS	17,235	MLVMS_P03355_PLV919
GGGGSSEAAAK	17,236	MLVMS_P03355_PLV919
GSSEAAAKGGG	17,237	MLVFF_P26809_3mutA
PAPGGS	17,238	MLVMS_P03355_3mutA_WS
EAAAKGGSGGG	17,239	MLVCB_P08361_3mutA
EAAAKGGG	17,240	PERV_Q4VFZ2_3mut
PAPGGS	17,241	XMRV6_A1Z651_3mutA
GSSPAPGGG	17,242	XMRV6_A1Z651_3mutA
PAPEAAAKGGG	17,243	MLVMS_P03355_3mutA_WS
GSSEAAAKGGG	17,244	PERV_Q4VFZ2_3mutA_WS
PAPGGSEAAAK	17,245	XMRV6_A1Z651_3mutA
GGGGGS	17,246	MLVMS_P03355_3mutA_WS
GGSPAPEAAAK	17,247	MLVMS_P03355_3mutA_WS
GGGPAP	17,248	MLVFF_P26809_3mutA
PAPGSSGGG	17,249	XMRV6_A1Z651_3mutA
PAPGSSGGG	17,250	MLVBM_Q7SVK7_3mutA_WS
GGGEAAAKGSS	17,251	MLVMS_P03355_3mutA_WS
GSSEAAAKGGS	17,252	MLVCB_P08361_3mutA
PAPGGSGSS	17,253	MLVCB_P08361_3mutA
EAAAKGGGGSEAAAK	17,254	BAEVM_P10272_3mutA
PAPAPAP	17,255	PERV_Q4VFZ2_3mutA_WS
GGGGGG	17,256	MLVAV_P03356_3mutA
GSSPAPEAAAK	17,257	MLVCB_P08361_3mutA
GGSGGSGGS	17,258	MLVMS_P03355_3mutA_WS
GSSGSSGSSGSSGSS	17,259	XMRV6_A1Z651_3mut
GGGPAPGGS	17,260	XMRV6_A1Z651_3mutA
GGGPAPEAAAK	17,261	BAEVM_P10272_3mutA
GGSGGG	17,262	AVIRE_P03360_3mutA
SGSETPGTSESATPES	17,263	PERV_Q4VFZ2_3mutA_WS
EAAAKGSSPAP	17,264	MLVMS_P03355_PLV919
GSSEAAAK	17,265	XMRV6_A1Z651_3mut
GSSGGSGGG	17,266	MLVFF_P26809_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,267	WMSV_P03359_3mutA

[1609]

[1610]

GGGGSEAAAKGGGGS	17,268	MLVMS_P03355_PLV919
PAPGGGGSS	17,269	MLVMS_P03355_3mutA_WS
SGSETPGTSESATPES	17,270	MLVMS_P03355_3mutA_WS
GGSPAPEAAAK	17,271	KORV_Q9TTC1-Pro_3mutA
GSSEAAAKGGG	17,272	MLVMS_P03355_3mutA_WS
GSSEAAAK	17,273	WMSV_P03359_3mutA
GGGGSEAAAKGGGGS	17,274	AVIRE_P03360_3mutA
GSS	17,275	WMSV_P03359_3mutA
PAPGGSEAAAK	17,276	MLVFF_P26809_3mutA
GGGGS	17,277	MLVMS_P03355_3mutA_WS
GGGPAP	17,278	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,279	MLVMS_P03355_3mutA_WS
EAAAKPAPGSS	17,280	PERV_Q4VFZ2_3mut
EAAAKPAPGSS	17,281	MLVCB_P08361_3mutA
GGGGGG	17,282	WMSV_P03359_3mutA
EAAAKPAPGGS	17,283	MLVMS_P03355_PLV919
PAPGGGEAAAK	17,284	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAKEAAAKEAAAK	17,285	AVIRE_P03360_3mutA
GSSEAAAKPAP	17,286	XMRV6_A1Z651_3mutA
PAPGGSEAAAK	17,287	MLVBM_Q7SVK7_3mutA_WS
PAPGSS	17,288	MLVCB_P08361_3mutA
EAAAKGGG	17,289	MLVMS_P03355_3mutA_WS
EAAAKPAP	17,290	MLVCB_P08361_3mutA
PAPEAAAKGGS	17,291	MLVBM_Q7SVK7_3mutA_WS
GGSPAPGGG	17,292	MLVCB_P08361_3mutA
PAPGGSGSS	17,293	WMSV_P03359_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,294	MLVMS_P03355_PLV919
GGSGGGPAP	17,295	MLVMS_P03355_PLV919
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,296	MLVMS_P03355
PAPEAAAKGSS	17,297	MLVCB_P08361_3mutA
EAAAKGSS	17,298	MLVMS_P03355_3mutA_WS
GGSGGS	17,299	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAK	17,300	BAEVM_P10272_3mutA
GGGGSEAAAKGGGGS	17,301	FLV_P10273_3mutA
GGSEAAAKGGG	17,302	MLVCB_P08361_3mutA
GSSGSSGSSGSSGSS	17,303	BAEVM_P10272_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,304	MLVFF_P26809_3mutA
EAAAKGGG	17,305	PERV_Q4VFZ2_3mut
GGGGGSEAAAK	17,306	MLVCB_P08361_3mutA

EAAAKPAPGGS	17,307	MLVMS_P03355_3mutA_WS
GGGGGSGSS	17,308	XMRV6_A1Z651_3mutA
PAPGSSEAAAK	17,309	MLVMS_P03355_3mutA_WS
GSSEAAAKPAP	17,310	MLVCB_P08361_3mutA
EAAAKGSSPAP	17,311	MLVAV_P03356_3mutA
GGGPAPGGS	17,312	WMSV_P03359_3mutA
GGSPAP	17,313	MLVMS_P03355_3mutA_WS
GGSEAAAKGGG	17,314	MLVMS_P03355_3mutA_WS
GGGGGGGG	17,315	MLVFF_P26809_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,316	MLVMS_P03355_3mutA_WS
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,317	MLVBM_Q7SVK7_3mutA_WS
GSSPAPGGG	17,318	MLVAV_P03356_3mutA
GGGGGG	17,319	AVIRE_P03360_3mutA
GSSGGS	17,320	MLVMS_P03355_3mutA_WS
GGSPAPGSS	17,321	MLVFF_P26809_3mutA
PAPEAAAKGGG	17,322	PERV_Q4VFZ2_3mut
EAAAKGGGPAP	17,323	MLVFF_P26809_3mutA
GGGEAAAKGGS	17,324	MLVMS_P03355_PLV919
GGSGSSPAP	17,325	MLVFF_P26809_3mutA
SGSETPGTSESATPES	17,326	WMSV_P03359_3mutA
PAPGGSEAAAK	17,327	MLVBM_Q7SVK7_3mutA_WS
GGSGGG	17,328	MLVMS_P03355_PLV919
GGGGSSPAP	17,329	PERV_Q4VFZ2_3mut
GGGEAAAKGSS	17,330	MLVAV_P03356_3mutA
PAPAPAPAPAPAP	17,331	MLVMS_P03355_3mutA_WS
EAAAKGGGGSEAAAK	17,332	PERV_Q4VFZ2
EAAAKEAAAKEAAAKEAAAKEAAAK	17,333	MLVMS_P03355_PLV919
GGGGGSEAAAK	17,334	PERV_Q4VFZ2_3mut
PAPGSSEAAAK	17,335	MLVCB_P08361_3mutA
GSAGSAAGSGEF	17,336	PERV_Q4VFZ2_3mutA_WS
EAAAKGGGGSEAAAK	17,337	MLVFF_P26809_3mutA
GGSPAPGGG	17,338	PERV_Q4VFZ2_3mutA_WS
GSSEAAAKGGG	17,339	AVIRE_P03360_3mutA
GGGEAAAKPAP	17,340	MLVMS_P03355_3mutA_WS
GGGPAP	17,341	AVIRE_P03360_3mutA
GGSEAAAK	17,342	MLVCB_P08361_3mutA
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,343	PERV_Q4VFZ2_3mut
EAAAKPAPGGS	17,344	MLVBM_Q7SVK7_3mutA_WS

[1611]

AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,345	XMRV6_A1Z651_3mut
GGGGGGGG	17,346	MLVCB_P08361_3mutA
PAPGSS	17,347	PERV_Q4VFZ2_3mut
EAAAK	17,348	PERV_Q4VFZ2_3mut
GSAGSAAGSGEF	17,349	MLVMS_P03355_3mutA_WS
PAPGGGEAAAK	17,350	PERV_Q4VFZ2_3mut
EAAAKGSSGGS	17,351	MLVFF_P26809_3mut
GGGGSEAAAKGGGGS	17,352	BAEVM_P10272_3mutA
GGGGSGGGGSGGGGS	17,353	MLVMS_P03355_PLV919
EAAAKGGGGSEAAAK	17,354	BAEVM_P10272_3mut
PAPGGGEAAAK	17,355	MLVMS_P03355_3mutA_WS
GGSEAAAKPAP	17,356	MLVMS_P03355_3mutA_WS
PAPAP	17,357	MLVCB_P08361_3mutA
PAPAP	17,358	MLVFF_P26809_3mutA
GGSPAP	17,359	AVIRE_P03360_3mutA
EAAAKGSSGGS	17,360	MLVCB_P08361_3mutA
PAPGSSGGS	17,361	AVIRE_P03360_3mutA
EAAAKGGGGSEAAAK	17,362	XMRV6_A1Z651_3mutA
PAPAPAP	17,363	BAEVM_P10272_3mutA
GGSGGSGGSGGSGGSGGS	17,364	MLVMS_P03355_PLV919
GGGGGSGSS	17,365	MLVMS_P03355_PLV919
PAPGSSEAAAK	17,366	XMRV6_A1Z651_3mut
GGSEAAAKPAP	17,367	XMRV6_A1Z651_3mutA
EAAAKEAAAKEAAAKEAAAK	17,368	XMRV6_A1Z651_3mut
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,369	WMSV_P03359_3mut
GGSGGGEAAAK	17,370	XMRV6_A1Z651_3mutA
GGGEAAAK	17,371	XMRV6_A1Z651_3mutA
GGGGSGGGGSGGGGS	17,372	MLVMS_P03355_3mutA_WS
GGSGGSGGSGGSGGS	17,373	MLVFF_P26809_3mutA
GSSGGGGGS	17,374	MLVMS_P03355_3mut
PAPGGSEAAAK	17,375	MLVMS_P03355_3mutA_WS
GSSGGSPAP	17,376	MLVMS_P03355_3mutA_WS
SGSETPGTSESATPES	17,377	XMRV6_A1Z651_3mutA
GGGGSGGGGS	17,378	MLVMS_P03355_PLV919
PAPAPAPAPAP	17,379	MLVMS_P03355_3mut
GSSGSS	17,380	XMRV6_A1Z651_3mutA
GSSEAAAKPAP	17,381	PERV_Q4VFZ2_3mut
GGSGSSGGG	17,382	MLVMS_P03355_3mutA_WS
EAAAKEAAAK	17,383	MLVCB_P08361_3mutA
GSSGSSGSSGSS	17,384	MLVMS_P03355_3mutA_WS

[1612]

GSSPAPGGG	17,385	PERV_Q4VFZ2_3mutA_WS
EAAAKEAAAKEAAAK	17,386	MLVMS_P03355_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,387	SFV1_P23074_2mutA
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,388	MLVMS_P03355_PLV919
GSAGSAAGSGEF	17,389	MLVMS_P03355_PLV919
PAPGSSEAAAK	17,390	MLVMS_P03355_3mutA_WS
GGSEAAAK	17,391	MLVMS_P03355_3mutA_WS
GSSGSSGSSGSSGSS	17,392	PERV_Q4VFZ2_3mutA_WS
GGSEAAAKPAP	17,393	PERV_Q4VFZ2_3mutA_WS
GGSGGSGGS	17,394	MLVCB_P08361_3mutA
EAAAKGGSGSS	17,395	MLVCB_P08361_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGS	17,396	FLV_P10273_3mutA
EAAAKEAAAKEAAAKEAAAK	17,397	MLVBM_Q7SVK7_3mutA_WS
GGSGSSPAP	17,398	BAEVM_P10272_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,399	XMRV6_A1Z651_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGS	17,400	MLVBM_Q7SVK7_3mutA_WS
GGSGSS	17,401	WMSV_P03359_3mutA
PAPEAAAK	17,402	MLVCB_P08361_3mutA
EAAAKPAP	17,403	BAEVM_P10272_3mutA
GSSPAP	17,404	PERV_Q4VFZ2_3mutA_WS
GGGPAP	17,405	PERV_Q4VFZ2_3mutA_WS
EAAAKGGSGSS	17,406	MLVMS_P03355_3mutA_WS
EAAAKGGGGSEAAAK	17,407	AVIRE_P03360_3mutA
GGSGGG	17,408	KORV_Q9TTC1-Pro_3mutA
GSSPAP	17,409	MLVFF_P26809_3mutA
GGSGSSEAAAK	17,410	BAEVM_P10272_3mutA
PAPGSSGGS	17,411	BAEVM_P10272_3mutA
GGGGGG	17,412	MLVFF_P26809_3mutA
PAPGGSEAAAK	17,413	MLVMS_P03355_PLV919
PAPGGS	17,414	MLVMS_P03355_PLV919
GGSGGSGGSGGS	17,415	BAEVM_P10272_3mutA
GSSPAP	17,416	MLVCB_P08361_3mutA
PAPAPAPAP	17,417	MLVMS_P03355_3mutA_WS
GGGGGG	17,418	MLVCB_P08361_3mutA
GSSGSSGSSGSSGSSGSS	17,419	KORV_Q9TTC1-Pro_3mutA
GSSEAAAKGGS	17,420	BAEVM_P10272_3mutA
GGSEAAAK	17,421	FLV_P10273_3mutA
GGSGGSGGSGGSGGS	17,422	KORV_Q9TTC1-Pro_3mutA
GSSPAPEAAAK	17,423	PERV_Q4VFZ2_3mut

[1613]

GSSGSSGSSGSSGSS	17,424	XMRV6_A1Z651_3mutA
EAAAKPAPGGS	17,425	MLVMS_P03355_3mut
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,426	FLV_P10273_3mut
GGSPAPEAAAK	17,427	XMRV6_A1Z651_3mut
EAAAKGGSGGG	17,428	MLVFF_P26809_3mutA
EAAAKEAAAKEAAAKEAAAK	17,429	MLVFF_P26809_3mutA
GSSPAP	17,430	WMSV_P03359_3mutA
PAPAPAPAP	17,431	MLVAV_P03356_3mutA
PAPGGSEAAAK	17,432	KORV_Q9TTC1_3mut
GGSGSSEAAAK	17,433	MLVBM_Q7SVK7_3mutA_WS
GSSGGG	17,434	MLVCB_P08361_3mutA
GGGEAAAKGSS	17,435	PERV_Q4VFZ2_3mut
PAPGGSGGG	17,436	MLVFF_P26809_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,437	FFV_O93209
PAPGGGGSS	17,438	MLVMS_P03355_3mutA_WS
EAAAKGGS	17,439	MLVAV_P03356_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,440	MLVBM_Q7SVK7_3mutA_WS
GGSGGSGGS	17,441	WMSV_P03359_3mutA
PAPAP	17,442	MLVMS_P03355_3mutA_WS
GSSGGGEAAAK	17,443	MLVAV_P03356_3mutA
GGGGSSEAAAK	17,444	MLVFF_P26809_3mutA
EAAAKGSSGGS	17,445	MLVMS_P03355_PLV919
EAAAKGGGGSEAAAK	17,446	MLVMS_P03355_3mutA_WS
GGGGGGGG	17,447	MLVMS_P03355_PLV919
GSSGSSGSS	17,448	MLVMS_P03355_PLV919
GGGEAAAKPAP	17,449	PERV_Q4VFZ2_3mutA_WS
GGGGGSGSS	17,450	MLVMS_P03355_3mutA_WS
GGGGGGG	17,451	MLVMS_P03355_PLV919
GGS	17,452	MLVMS_P03355_PLV919
GSSGGG	17,453	MLVMS_P03355_3mutA_WS
EAAAKGGSGSS	17,454	PERV_Q4VFZ2_3mutA_WS
PAPGSSEAAAK	17,455	MLVMS_P03355_PLV919
GSSEAAAKPAP	17,456	MLVMS_P03355_PLV919
GGSPAPGSS	17,457	BAEVM_P10272_3mutA
GSAGSAAGSGEF	17,458	MLVCB_P08361_3mut
GGSPAPGGG	17,459	PERV_Q4VFZ2_3mut
GGGGSGGGGSGGGGSGGGGS	17,460	MLVMS_P03355_3mut
GSSGSSGSS	17,461	PERV_Q4VFZ2_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,462	PERV_Q4VFZ2_3mut

[1614]

GGGGSEAAAKGGGGS	17,463	MLVCB_P08361_3mutA
GGSEAAAKGSS	17,464	MLVAV_P03356_3mutA
EAAAKGGGGSEAAAK	17,465	MLVCB_P08361_3mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,466	XMRV6_A1Z651_3mutA
PAPGGGEAAAK	17,467	MLVMS_P03355_3mutA_WS
GSSGGGEAAAK	17,468	PERV_Q4VFZ2_3mutA_WS
GSSGSS	17,469	MLVCB_P08361_3mut
PAPAPAPAPAPAP	17,470	PERV_Q4VFZ2_3mut
GGSPAPGGG	17,471	MLVFF_P26809_3mutA
GGSGGSGGSGGSGGS	17,472	MLVCB_P08361_3mutA
EAAAKEAAAK	17,473	MLVFF_P26809_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,474	GALV_P21414_3mut
PAPAPAPAPAPAP	17,475	WMSV_P03359_3mutA
GGGEAAAKGGS	17,476	KORV_Q9TTC1_3mutA
EAAAKGGGPAP	17,477	KORV_Q9TTC1_3mut
PAPEAAAKGSS	17,478	MLVBM_Q7SVK7_3mutA_WS
PAPEAAAKGSS	17,479	FLV_P10273_3mutA
PAPGGSEAAAK	17,480	MLVMS_P03355_3mut
GSSPAPGGG	17,481	BAEVM_P10272_3mutA
GGGEAAAKPAP	17,482	KORV_Q9TTC1-Pro_3mutA
GGGGSGGGGS	17,483	MLVMS_P03355_PLV919
GGGEAAAKGSS	17,484	MLVFF_P26809_3mutA
PAPGGGGSS	17,485	MLVBM_Q7SVK7_3mutA_WS
GSSEAAAK	17,486	BAEVM_P10272_3mutA
GGGGGGGG	17,487	MLVMS_P03355_PLV919
PAPGSSGGS	17,488	MLVAV_P03356_3mutA
GGGGSGGGGSGGGGSGGGGS	17,489	BAEVM_P10272_3mutA
PAP	17,490	MLVMS_P03355_3mut
EAAAKGSSPAP	17,491	XMRV6_A1Z651_3mutA
PAPEAAAKGGS	17,492	MLVFF_P26809_3mutA
GSSGGGEAAAK	17,493	BAEVM_P10272_3mutA
PAPAPAP	17,494	MLVMS_P03355_3mutA_WS
GGSEAAAKGGG	17,495	MLVMS_P03355_PLV919
GSSEAAAK	17,496	PERV_Q4VFZ2_3mut
GGGG	17,497	MLVMS_P03355_3mutA_WS
GGGGGS	17,498	MLVMS_P03355_3mut
GGGGSSEAAAK	17,499	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,500	SFV3L_P27401-Pro_2mutA
GGSEAAAKGSS	17,501	MLVMS_P03355_3mutA_WS

[1615]

PAPGSSGGS	17,502	XMRV6_A1Z651_3mutA
GGSPAP	17,503	MLVMS_P03355_3mutA_WS
GGGGSSEAAAK	17,504	BAEVM_P10272_3mut
GGSGGSGGSGGS	17,505	AVIRE_P03360_3mutA
PAPGSSGGS	17,506	MLVFF_P26809_3mutA
GSSPAPGGG	17,507	MLVMS_P03355_3mutA_WS
GGGGGGG	17,508	MLVMS_P03355_3mutA_WS
EAAAKGGGGGS	17,509	MLVMS_P03355_3mutA_WS
EAAAKGGSGGG	17,510	MLVMS_P03355_PLV919
GGGGSSEAAAK	17,511	XMRV6_A1Z651_3mutA
GGGGSEAAAKGGGGS	17,512	MLVBM_Q7SVK7_3mutA_WS
GSSGSS	17,513	MLVMS_P03355_PLV919
GGSGGG	17,514	MLVMS_P03355_PLV919
PAPEAAAKGGG	17,515	AVIRE_P03360_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,516	FOAMV_P14350-Pro_2mutA
GGGGGSGSS	17,517	PERV_Q4VFZ2_3mut
GSSGSSGSSGSSGSS	17,518	KORV_Q9TTC1-Pro_3mut
GGGGSEAAAKGGGGS	17,519	MLVMS_P03355_3mutA_WS
GGGGGSPAP	17,520	FLV_P10273_3mut
GGGEAAAK	17,521	MLVMS_P03355_3mutA_WS
GGSGGSGGSGGS	17,522	FLV_P10273_3mutA
GGG	17,523	MLVMS_P03355_PLV919
GGSPAPEAAAK	17,524	BAEVM_P10272_3mutA
EAAAKEAAAK	17,525	FLV_P10273_3mutA
GGGEAAAKPAP	17,526	BAEVM_P10272_3mutA
GGGEAAAKGGS	17,527	PERV_Q4VFZ2_3mut
GGSGGSGGS	17,528	PERV_Q4VFZ2_3mut
EAAAKGGGPAP	17,529	XMRV6_A1Z651_3mutA
EAAAK	17,530	MLVBM_Q7SVK7_3mutA_WS
PAPEAAAKGGG	17,531	PERV_Q4VFZ2_3mut
EAAAKGSS	17,532	MLVCB_P08361_3mutA
GGSEAAAKGGG	17,533	MLVBM_Q7SVK7_3mutA_WS
GGGGSGGGGSGGGGSGGGGS	17,534	XMRV6_A1Z651_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGS	17,535	BAEVM_P10272_3mut
GGGGSSPAP	17,536	PERV_Q4VFZ2_3mutA_WS
GGSGGSGGSGGSGGSGGS	17,537	PERV_Q4VFZ2_3mut
GGGEAAAKPAP	17,538	PERV_Q4VFZ2_3mut
EAAAKEAAAK	17,539	BAEVM_P10272_3mutA

[1616]

GGSGSSEAAAK	17,540	XMRV6_A1Z651_3mutA
PAPEAAAKGSS	17,541	WMSV_P03359_3mutA
PAPAPAPAPAP	17,542	XMRV6_A1Z651_3mutA
GSSGGGEAAAK	17,543	MLVMS_P03355_PLV919
GSSPAPGGG	17,544	MLVFF_P26809_3mutA
GGSPAPEAAAK	17,545	MLVFF_P26809_3mut
PAPGGSEAAAK	17,546	PERV_Q4VFZ2_3mut
GGGGSS	17,547	MLVFF_P26809_3mutA
GGSGSSGGG	17,548	BAEVM_P10272_3mutA
GSSGGGEAAAK	17,549	MLVMS_P03355_3mutA_WS
EAAAKGGS	17,550	MLVBM_Q7SVK7_3mutA_WS
GGGPAPGGS	17,551	MLVMS_P03355_PLV919
EAAAKEAAAK	17,552	MLVMS_P03355_PLV919
GSSGSSGSS	17,553	MLVMS_P03355_PLV919
GGGEAAAKPAP	17,554	MLVAV_P03356_3mutA
SGSETPGTSESATPES	17,555	FLV_P10273_3mutA
PAPAPAPAPAP	17,556	KORV_Q9TTC1-Pro_3mut
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,557	BAEVM_P10272_3mutA
PAPGSSGGG	17,558	MLVMS_P03355_3mutA_WS
GSSGGGEAAAK	17,559	XMRV6_A1Z651_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGS	17,560	XMRV6_A1Z651_3mutA
GGGGSSPAP	17,561	MLVFF_P26809_3mutA
GGSGGGPAP	17,562	PERV_Q4VFZ2_3mutA_WS
GSS	17,563	PERV_Q4VFZ2_3mut
EAAAKGSSPAP	17,564	MLVMS_P03355_3mut
EAAAKGGG	17,565	XMRV6_A1Z651_3mutA
GSSGSSGSSGSS	17,566	WMSV_P03359_3mutA
PAPEAAAKGSS	17,567	MLVMS_P03355_PLV919
GSSEAAAK	17,568	AVIRE_P03360_3mutA
EAAAKGGSGSS	17,569	AVIRE_P03360_3mutA
GSSEAAAK	17,570	MLVMS_P03355_3mut
GGSGSSEAAAK	17,571	MLVMS_P03355_PLV919
GGSEAAAKGGG	17,572	MLVFF_P26809_3mutA
GGGGSGGGGSGGGGSGGGGS	17,573	MLVAV_P03356_3mutA
PAPAPAPAPAPAP	17,574	MLVFF_P26809_3mut
EAAAKPAPGSS	17,575	KORV_Q9TTC1-Pro_3mut
PAPGSSEAAAK	17,576	MLVAV_P03356_3mutA
GGGGSSPAP	17,577	WMSV_P03359_3mutA
EAAAKGGGGGS	17,578	MLVMS_P03355_3mutA_WS
GGGEAAAKGGS	17,579	MLVMS_P03355_3mut

[1617]

GGSGSSGGG	17,580	MLVMS_P03355_3mut
GGGPAPGGS	17,581	MLVAV_P03356_3mutA
PAPGGGGGS	17,582	MLVMS_P03355_PLV919
GGGPAPGSS	17,583	PERV_Q4VFZ2_3mut
GGGGGGG	17,584	MLVFF_P26809_3mutA
GGSGGGGSS	17,585	MLVCB_P08361_3mutA
GGGGGG	17,586	FLV_P10273_3mutA
GGSEAAAKGSS	17,587	PERV_Q4VFZ2_3mut
GGSPAPGGG	17,588	BAEVM_P10272_3mutA
GGSPAPGSS	17,589	AVIRE_P03360_3mutA
GGSGGSGGSGGS	17,590	KORV_Q9TTC1_3mut
EAAAKEAAAKEAAAKEAAAKEAAAK	17,591	MLVBM_Q7SVK7_3mut
PAPGSSGGS	17,592	XMRV6_A1Z651_3mut
EAAAKGGGGSS	17,593	PERV_Q4VFZ2_3mutA_WS
GGSGGSGGSGGSGGS	17,594	PERV_Q4VFZ2_3mutA_WS
PAPGGSGGG	17,595	MLVMS_P03355_PLV919
PAPGSSGGG	17,596	PERV_Q4VFZ2_3mutA_WS
GSSGSS	17,597	BAEVM_P10272_3mutA
EAAAKGSS	17,598	MLVFF_P26809_3mutA
GGGPAP	17,599	MLVMS_P03355_PLV919
EAAAKGGGGGS	17,600	MLVFF_P26809_3mutA
EAAAKGGSPAP	17,601	MLVBM_Q7SVK7_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,602	WMSV_P03359_3mutA
GSSPAPGGG	17,603	MLVBM_Q7SVK7_3mutA_WS
GGGEAAAKGSS	17,604	AVIRE_P03360_3mutA
GGGGSSEAAAK	17,605	AVIRE_P03360_3mutA
GGGGGGGG	17,606	PERV_Q4VFZ2_3mutA_WS
PAPGSSEAAAK	17,607	BAEVM_P10272_3mutA
EAAAKGSS	17,608	MLVFF_P26809_3mut
GSSEAAAKGGG	17,609	MLVCB_P08361_3mutA
GGSEAAAK	17,610	MLVBM_Q7SVK7_3mutA_WS
GSSEAAAKGGG	17,611	PERV_Q4VFZ2_3mutA_WS
PAPGGSGGG	17,612	WMSV_P03359_3mutA
GSSGGSGGG	17,613	MLVCB_P08361_3mutA
EAAAKGSSGGG	17,614	FLV_P10273_3mutA
GSSEAAAK	17,615	MLVCB_P08361_3mutA
GSSGGGEAAAK	17,616	MLVMS_P03355_3mut
GGGGSGGGGS	17,617	MLVCB_P08361_3mutA

[1618]

EAAAKGGGGSEAAAK	17,618	MLVBM_Q7SVK7_3mutA_WS
EAAAKGGG	17,619	PERV_Q4VFZ2_3mutA_WS
EAAAKGGSPAP	17,620	MLVMS_P03355_PLV919
GGGPAPGGS	17,621	AVIRE_P03360_3mutA
GSSEAAAK	17,622	MLVBM_Q7SVK7_3mutA_WS
GSSGGGEAAAK	17,623	PERV_Q4VFZ2_3mut
SGSETPGTSESATPES	17,624	MLVMS_P03355_PLV919
GGSGSSPAP	17,625	MLVMS_P03355_3mut
GGGGGG	17,626	MLVBM_Q7SVK7_3mutA_WS
GGSPAPGGG	17,627	XMRV6_A1Z651_3mutA
GGSGSS	17,628	PERV_Q4VFZ2_3mutA_WS
PAP	17,629	MLVBM_Q7SVK7_3mutA_WS
EAAAKPAPGSS	17,630	MLVMS_P03355_PLV919
EAAAKGGG	17,631	MLVMS_P03355_3mut
GSSEAAAKPAP	17,632	PERV_Q4VFZ2_3mutA_WS
GGGGSS	17,633	MLVMS_P03355_3mutA_WS
GGSGSSEAAAK	17,634	PERV_Q4VFZ2_3mut
GGGGSS	17,635	BAEVM_P10272_3mutA
PAPAP	17,636	MLVFF_P26809_3mut
PAPEAAAKGGG	17,637	BAEVM_P10272_3mutA
EAAAKGGS	17,638	MLVMS_P03355_PLV919
PAPAPAPAPAPAP	17,639	PERV_Q4VFZ2_3mutA_WS
GGGGGSEAAAK	17,640	MLVMS_P03355_3mut
PAPGGS	17,641	PERV_Q4VFZ2_3mut
GGGGSS	17,642	MLVCB_P08361_3mutA
GGGGS	17,643	MLVAV_P03356_3mutA
GSSPAPEAAAK	17,644	MLVMS_P03355_PLV919
GGGGSSGGS	17,645	MLVFF_P26809_3mutA
PAPEAAAKGSS	17,646	MLVMS_P03355_PLV919
GGSGSSEAAAK	17,647	MLVMS_P03355_3mutA_WS
EAAAKGGG	17,648	MLVAV_P03356_3mutA
PAPGSSEAAAK	17,649	FLV_P10273_3mutA
EAAAKGSSGGG	17,650	MLVCB_P08361_3mutA
PAPEAAAK	17,651	KORV_Q9TTC1-Pro_3mutA
GGSPAPEAAAK	17,652	KORV_Q9TTC1-Pro_3mut
GGSGGSGGSGGSGGSGGS	17,653	MLVAV_P03356_3mutA
GSSEAAAKPAP	17,654	MLVBM_Q7SVK7_3mutA_WS

[1619]

AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,655	KORV_Q9TTC1-Pro_3mutA
GSSGGGEAAAK	17,656	XMRV6_A1Z651_3mut
PAPGGSGGG	17,657	AVIRE_P03360_3mutA
PAPGGSEAAAK	17,658	PERV_Q4VFZ2_3mutA_WS
GGGGS	17,659	MLVMS_P03355_3mutA_WS
GGGGSGGGGSGGGGS	17,660	MLVBM_Q7SVK7_3mutA_WS
PAPAPAPAPAP	17,661	PERV_Q4VFZ2_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAK	17,662	MLVMS_P03355_3mut
GSSGGSEAAAK	17,663	MLVMS_P03355_3mutA_WS
GGSGGSGGSGGS	17,664	WMSV_P03359_3mutA
EAAAKGSSGGG	17,665	WMSV_P03359_3mutA
EAAAKGGG	17,666	PERV_Q4VFZ2_3mutA_WS
SGSETPGTSESATPES	17,667	PERV_Q4VFZ2_3mut
PAPGSSGGS	17,668	MLVMS_P03355_3mutA_WS
PAPEAAAKGSS	17,669	PERV_Q4VFZ2_3mut
PAPEAAAK	17,670	AVIRE_P03360_3mutA
GSSEAAAKGGG	17,671	BAEVM_P10272_3mutA
GSSPAP	17,672	MLVAV_P03356_3mutA
EAAAKEAAAKEAAAKEAAAK	17,673	MLVFF_P26809_3mut
PAPGGSGSS	17,674	MLVAV_P03356_3mutA
GGGGSGGGGSGGGGS	17,675	PERV_Q4VFZ2_3mutA_WS
GSSGGSEAAAK	17,676	MLVCB_P08361_3mutA
EAAAKGGS	17,677	KORV_Q9TTC1-Pro_3mutA
EAAAKGGS	17,678	MLVFF_P26809_3mutA
GGSPAP	17,679	MLVMS_P03355_PLV919
GGSGSS	17,680	MLVMS_P03355_PLV919
SGSETPGTSESATPES	17,681	WMSV_P03359_3mut
GGGGGGG	17,682	WMSV_P03359_3mut
GGSPAPGSS	17,683	MLVCB_P08361_3mutA
GGGGSSGGS	17,684	WMSV_P03359_3mut
PAPGGS	17,685	MLVMS_P03355_PLV919
PAPGSSGGS	17,686	MLVCB_P08361_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,687	MLVFF_P26809_3mut
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,688	PERV_Q4VFZ2_3mut
GGSGGSGGSGGSGGS	17,689	BAEVM_P10272_3mutA
GSSEAAAK	17,690	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAKEAAAK	17,691	KORV_Q9TTC1-Pro_3mutA
GGSGGSGGSGGSGGS	17,692	MLVMS_P03355_3mut
PAPAPAPAPAPAP	17,693	MLVMS_P03355_3mut
GGSPAPEAAAK	17,694	MLVMS_P03355_PLV919

[1620]

EAAAK	17,695	WMSV_P03359_3mutA
EAAAKGSSGGS	17,696	MLVBM_Q7SVK7_3mutA_WS
GGSGGGGSS	17,697	MLVMS_P03355_3mutA_WS
GGGEAAAKPAP	17,698	MLVMS_P03355_3mut
EAAAKGGSGGG	17,699	XMRV6_A1Z651_3mutA
GGGGGSEAAAK	17,700	KORV_Q9TTC1-Pro_3mutA
GGGGGG	17,701	BAEVM_P10272_3mutA
GGGGGG	17,702	MLVMS_P03355_3mut
GGGGGGG	17,703	MLVBM_Q7SVK7_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,704	AVIRE_P03360
PAPGSSGGS	17,705	PERV_Q4VFZ2_3mut
GGGGGS	17,706	XMRV6_A1Z651_3mut
EAAAKPAP	17,707	XMRV6_A1Z651_3mutA
GGG	17,708	MLVMS_P03355_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,709	FLV_P10273_3mut
EAAAKGSSPAP	17,710	MLVMS_P03355_3mut
SGSETPGTSESATPES	17,711	BAEVM_P10272_3mutA
GGSPAPEAAAK	17,712	MLVMS_P03355_3mut
GSSGSSGSSGSS	17,713	MLVAV_P03356_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,714	MLVMS_P03355_3mut
GGSPAP	17,715	MLVCB_P08361_3mutA
GGGGGSEAAAK	17,716	MLVMS_P03355_3mutA_WS
GGGGG	17,717	MLVFF_P26809_3mutA
GSSEAAAK	17,718	MLVAV_P03356_3mutA
GGS	17,719	BAEVM_P10272_3mut
EAAAKGGSPAP	17,720	MLVCB_P08361_3mutA
PAPAPAPAP	17,721	FLV_P10273_3mutA
PAPGGGEAAAK	17,722	MLVCB_P08361_3mutA
GGGGSSEAAAK	17,723	MLVMS_P03355_3mutA_WS
GGGGG	17,724	PERV_Q4VFZ2_3mutA_WS
GGSGGSGGSGGSGGSGGS	17,725	PERV_Q4VFZ2_3mut
GGGGG	17,726	MLVMS_P03355_3mut
PAPEAAAKGGG	17,727	MLVBM_Q7SVK7_3mutA_WS
GSSGGGPAP	17,728	XMRV6_A1Z651_3mutA
GSSGSSGSSGSSGSSGSS	17,729	PERV_Q4VFZ2_3mutA_WS
EAAAKGGSPAP	17,730	PERV_Q4VFZ2_3mut
GSSGGSEAAAK	17,731	MLVMS_P03355_PLV919

[1621]

GSS	17,732	PERV_Q4VFZ2_3mut
EAAAKGGS	17,733	WMSV_P03359_3mutA
GGGGGSPAP	17,734	PERV_Q4VFZ2_3mutA_WS
EAAAKGSS	17,735	MLVMS_P03355_PLV919
EAAAKGGGGSS	17,736	KORV_Q9TTC1-Pro_3mutA
PAPGSSGGG	17,737	PERV_Q4VFZ2_3mut
GGGGSSEAAAK	17,738	MLVFF_P26809_3mut
PAPAPAP	17,739	MLVMS_P03355_3mut
GSSGGSEAAAK	17,740	XMRV6_A1Z651_3mut
PAPEAAAKGSS	17,741	MLVMS_P03355_3mutA_WS
GGSGGSGGSGGSGGS	17,742	MLVMS_P03355_3mutA_WS
GGSGSSPAP	17,743	XMRV6_A1Z651_3mutA
GGGGSSPAP	17,744	MLVMS_P03355_PLV919
GGGGS	17,745	MLVCB_P08361_3mutA
EAAAKEAAAKEAAAKEAAAK	17,746	PERV_Q4VFZ2_3mutA_WS
EAAAKEAAAK	17,747	KORV_Q9TTC1_3mutA
PAPGGGEAAAK	17,748	BAEVM_P10272_3mutA
GSSGGSEAAAK	17,749	XMRV6_A1Z651_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,750	FLV_P10273_3mut
GSSEAAAKPAP	17,751	MLVMS_P03355_3mutA_WS
EAAAKPAPGSS	17,752	PERV_Q4VFZ2_3mutA_WS
GSSGGSPAP	17,753	XMRV6_A1Z651_3mutA
GSSEAAAKGGG	17,754	PERV_Q4VFZ2_3mut
GGGEAAAKGGS	17,755	WMSV_P03359_3mutA
GSSEAAAKGGG	17,756	MLVFF_P26809_3mut
PAPAPAP	17,757	KORV_Q9TTC1-Pro_3mutA
EAAAKGGSPAP	17,758	MLVMS_P03355_3mutA_WS
PAPGGSEAAAK	17,759	PERV_Q4VFZ2_3mut
GGGGS	17,760	MLVBM_Q7SVK7_3mutA_WS
EAAAKGSSGGG	17,761	KORV_Q9TTC1_3mut
EAAAKGGGPAP	17,762	MLVCB_P08361_3mutA
EAAAKGSS	17,763	BAEVM_P10272_3mutA
GGSPAPGGG	17,764	MLVBM_Q7SVK7_3mutA_WS
GGGGSEAAAKGGGGS	17,765	MLVMS_P03355_3mutA_WS
GGGEAAAKGGS	17,766	PERV_Q4VFZ2_3mutA_WS
EAAAKGGGGSS	17,767	MLVMS_P03355_3mutA_WS
EAAAKGGGPAP	17,768	MLVFF_P26809_3mut
GSSPAP	17,769	PERV_Q4VFZ2_3mutA_WS
EAAAKGGS	17,770	MLVMS_P03355_3mut
GGGGSS	17,771	KORV_Q9TTC1-Pro_3mutA

[1622]

[1623]

EAAAKGSSPAP	17,772	MLVMS_P03355_3mutA_WS
GGGPAP	17,773	PERV_Q4VFZ2_3mut
EAAAKGSSGGS	17,774	XMRV6_A1Z651_3mutA
PAPGGG	17,775	MLVAV_P03356_3mutA
GSSPAPEAAAK	17,776	BAEVM_P10272_3mutA
GGGPAP	17,777	MLVBM_Q7SVK7_3mutA_WS
GSSGGGGGS	17,778	AVIRE_P03360_3mutA
SGSETPGTSESATPES	17,779	MLVMS_P03355_PLV919
GGGPAP	17,780	MLVFF_P26809_3mut
EAAAKGGGGSS	17,781	XMRV6_A1Z651_3mutA
GGGGSSPAP	17,782	XMRV6_A1Z651_3mut
GGGGSEAAAKGGGGS	17,783	MLVMS_P03355_3mut
GSSPAP	17,784	MLVBM_Q7SVK7_3mutA_WS
GGSGSSEAAAK	17,785	FLV_P10273_3mutA
SGSETPGTSESATPES	17,786	MLVBM_Q7SVK7_3mutA_WS
PAPGGG	17,787	AVIRE_P03360_3mutA
GGGEAAAKPAP	17,788	MLVMS_P03355_3mutA_WS
EAAAKGGSGSS	17,789	PERV_Q4VFZ2_3mut
GGSPAPGGG	17,790	MLVAV_P03356_3mutA
PAPGGSGSS	17,791	BAEVM_P10272_3mutA
GSSGGSPAP	17,792	MLVFF_P26809_3mutA
EAAAKGSSGGG	17,793	PERV_Q4VFZ2_3mut
GGGGSGGGGS	17,794	PERV_Q4VFZ2_3mutA_WS
GSSGGGGGS	17,795	BAEVM_P10272_3mutA
GGGGSSGGS	17,796	MLVBM_Q7SVK7_3mutA_WS
EAAAKGGS	17,797	PERV_Q4VFZ2_3mutA_WS
GSSGSSGSSGSS	17,798	MLVMS_P03355_3mut
GGS	17,799	MLVMS_P03355_3mutA_WS
GSSGGSEAAAK	17,800	MLVBM_Q7SVK7_3mutA_WS
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,801	XMRV6_A1Z651
GGGGG	17,802	FLV_P10273_3mutA
PAPEAAAKGSS	17,803	PERV_Q4VFZ2_3mut
GGGGGG	17,804	WMSV_P03359_3mut
EAAAKGGG	17,805	BAEVM_P10272_3mutA
GGGGSS	17,806	MLVMS_P03355_3mutA_WS
GSSGGGEAAAK	17,807	KORV_Q9TTC1_3mut
GGSGSS	17,808	AVIRE_P03360_3mutA

[1624]

EAAAKPAP	17,809	MLVMS_P03355_3mut
EAAAKEAAAKEAAAK	17,810	FLV_P10273_3mutA
GGGG	17,811	XMRV6_A1Z651_3mutA
GSSPAPGGS	17,812	BAEVM_P10272_3mutA
GSSGGGGGS	17,813	MLVFF_P26809_3mutA
GGGGSSGGS	17,814	MLVAV_P03356_3mutA
GGS	17,815	PERV_Q4VFZ2_3mut
GGGGG	17,816	WMSV_P03359_3mutA
GSSGSSGSSGSSGSSGSS	17,817	FLV_P10273_3mutA
PAPGGGGSS	17,818	MLVAV_P03356_3mutA
GGGGGGGG	17,819	BAEVM_P10272_3mutA
SGSETPGTSESATPES	17,820	MLVCB_P08361_3mutA
PAPGGG	17,821	BAEVM_P10272_3mutA
GSSGSSGSS	17,822	MLVCB_P08361_3mutA
GGSGSS	17,823	MLVMS_P03355_3mutA_WS
EAAAKGGGGSEAAAK	17,824	WMSV_P03359_3mutA
GGGGGGGG	17,825	FLV_P10273_3mutA
GSSGSS	17,826	MLVMS_P03355_3mutA_WS
PAPEAAAKGGS	17,827	XMRV6_A1Z651_3mutA
EAAAKEAAAK	17,828	MLVMS_P03355_3mut
GGGGSGGGGSGGGGS	17,829	BAEVM_P10272_3mutA
EAAAKGSSPAP	17,830	MLVMS_P03355_PLV919
GGGGSSEAAAK	17,831	MLVMS_P03355_3mut
GGGGSSEAAAK	17,832	BAEVM_P10272_3mutA
PAPGGSGSS	17,833	PERV_Q4VFZ2_3mut
GGSGGGEAAAK	17,834	MLVFF_P26809_3mut
PAPEAAAKGGS	17,835	PERV_Q4VFZ2_3mut
GGGPAPGSS	17,836	AVIRE_P03360_3mut
PAPGGSGGG	17,837	PERV_Q4VFZ2_3mutA_WS
GGGGGGGG	17,838	PERV_Q4VFZ2_3mutA_WS
GSSEAAAK	17,839	MLVMS_P03355_3mutA_WS
GGGGSGGGGSGGGGS	17,840	PERV_Q4VFZ2_3mutA_WS
EAAAKGGS	17,841	MLVMS_P03355_3mut
GGGGGSGSS	17,842	MLVCB_P08361_3mut
GGGPAP	17,843	KORV_Q9TTC1-Pro_3mutA
EAAAKPAPGGG	17,844	MLVCB_P08361_3mut
GSSGGSPAP	17,845	MLVCB_P08361_3mutA
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,846	MLVMS_P03355_3mut
PAPAPAPAP	17,847	MLVMS_P03355_3mut
GSSGGS	17,848	XMRV6_A1Z651_3mutA
GSSEAAAKGGG	17,849	MLVMS_P03355_3mut

GGSGSSPAP	17,850	MLVMS_P03355_3mutA_WS
GSSEAAAKGGS	17,851	MLVMS_P03355_PLV919
EAAAKEAAAKEAAAKEAAAKEAAAK	17,852	BAEVM_P10272_3mut
PAPGGGGSS	17,853	KORV_Q9TTC1_3mutA
EAAAKGSS	17,854	MLVMS_P03355_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,855	FFV_O93209_2mut
GGSGGSGGSGGSGGSGGS	17,856	BAEVM_P10272_3mutA
GGGGGG	17,857	MLVMS_P03355_PLV919
PAPEAAAK	17,858	BAEVM_P10272_3mutA
GGSGSSEAAAK	17,859	MLVAV_P03356_3mutA
GGG	17,860	MLVCB_P08361_3mutA
GGGGG	17,861	MLVCB_P08361_3mutA
GGSGGSGGSGGS	17,862	KORV_Q9TTC1-Pro_3mutA
GSSGSSGSSGSSGSSGSS	17,863	XMRV6_A1Z651_3mutA
GSSEAAAKPAP	17,864	FLV_P10273_3mutA
GGGEAAAKPAP	17,865	MLVCB_P08361_3mutA
GSSGSSGSS	17,866	MLVMS_P03355_3mutA_WS
PAPAPAPAP	17,867	MLVMS_P03355_PLV919
EAAAKGGG	17,868	MLVMS_P03355_PLV919
PAPAPAPAPAPAP	17,869	FLV_P10273_3mutA
EAAAKGGSGSS	17,870	MLVMS_P03355_3mut
GGGGGG	17,871	PERV_Q4VFZ2_3mutA_WS
PAPGGG	17,872	MLVCB_P08361_3mutA
GGGGGSGSS	17,873	KORV_Q9TTC1_3mutA
GGGGSGGGGSGGGGSGGGGS	17,874	XMRV6_A1Z651_3mut
GGSGGSGGS	17,875	KORV_Q9TTC1-Pro_3mutA
EAAAKPAPGGG	17,876	MLVMS_P03355_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,877	XMRV6_A1Z651
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,878	FLV_P10273_3mutA
EAAAKGGGGSEAAAK	17,879	PERV_Q4VFZ2_3mutA_WS
GGGPAPGSS	17,880	AVIRE_P03360_3mutA
GGGGG	17,881	MLVMS_P03355_3mutA_WS
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,882	MLVMS_P03355_3mut
GGGGSGGGGS	17,883	MLVMS_P03355_3mutA_WS
EAAAKGGSPAP	17,884	XMRV6_A1Z651_3mutA
EAAAKGSSPAP	17,885	AVIRE_P03360_3mutA
PAPGGSGSS	17,886	KORV_Q9TTC1-Pro_3mutA
GSS	17,887	MLVBM_Q7SVK7_3mutA_WS
GSS	17,888	WMSV_P03359_3mut

[1625]

GGGPAPGSS	17,889	MLVFF_P26809_3mutA
EAAAKPAP	17,890	MLVMS_P03355_3mut
GSSPAPEAAAK	17,891	FLV_P10273_3mutA
GGSPAPGSS	17,892	MLVBM_Q7SVK7_3mutA_WS
GGGGGSEAAAK	17,893	XMRV6_A1Z651_3mut
PAPEAAAKGGG	17,894	WMSV_P03359_3mutA
PAPGGG	17,895	PERV_Q4VFZ2_3mut
GGSPAPEAAAK	17,896	WMSV_P03359_3mutA
GGSGGGGSS	17,897	PERV_Q4VFZ2_3mut
EAAAKGGGGSS	17,898	PERV_Q4VFZ2_3mut
EAAAKGGSPAP	17,899	AVIRE_P03360_3mut
GGSGGGGSS	17,900	WMSV_P03359_3mutA
PAPGSSEAAAK	17,901	MLVFF_P26809_3mut
GSSEAAAK	17,902	MLVMS_P03355_PLV919
GSAGSAAGSGEF	17,903	AVIRE_P03360_3mutA
EAAAKGGSGSS	17,904	MLVMS_P03355_3mut
GGSEAAAKPAP	17,905	MLVMS_P03355_PLV919
GGGGSGGGGSGGGGSGGGGSGGGGS	17,906	MLVFF_P26809_3mutA
PAPGSSEAAAK	17,907	PERV_Q4VFZ2_3mutA_WS
GGGGSSPAP	17,908	MLVMS_P03355_3mutA_WS
PAPAPAP	17,909	MLVCB_P08361_3mutA
EAAAKPAPGGG	17,910	MLVBM_Q7SVK7_3mutA_WS
GGGPAPGSS	17,911	BAEVM_P10272_3mutA
PAP	17,912	MLVMS_P03355_3mutA_WS
PAPGGSGGG	17,913	MLVMS_P03355_3mutA_WS
GGSGGSGGSGGSGGS	17,914	MLVBM_Q7SVK7_3mutA_WS
PAPAPAPAP	17,915	XMRV6_A1Z651_3mut
GSSPAPGGG	17,916	MLVMS_P03355_3mutA_WS
GSSPAPGGG	17,917	MLVMS_P03355_3mut
PAPGGG	17,918	MLVMS_P03355_PLV919
GGGEAAAKGSS	17,919	WMSV_P03359_3mut
EAAAKGSS	17,920	KORV_Q9TTC1-Pro_3mutA
EAAAKGGS	17,921	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAKEAAAKEAAAK	17,922	PERV_Q4VFZ2_3mut
PAPEAAAKGGG	17,923	MLVMS_P03355_PLV919
EAAAKGSSGGG	17,924	MLVFF_P26809_3mut
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,925	PERV_Q4VFZ2
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,926	MLVAV_P03356_3mutA

[1626]

GSSGGSGGG	17,927	MLVFF_P26809_3mut
GSSGSSGSSGSS	17,928	PERV_Q4VFZ2_3mutA_WS
GGSPAPGGG	17,929	MLVMS_P03355_PLV919
GSS	17,930	BAEVM_P10272_3mut
GGGPAPGSS	17,931	MLVMS_P03355_3mutA_WS
GGGGSS	17,932	KORV_Q9TTC1_3mutA
GSSGGSGGG	17,933	BAEVM_P10272_3mutA
EAAAKEAAAKEAAAK	17,934	MLVCB_P08361_3mutA
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,935	FLV_P10273_3mutA
PAPGGGGGS	17,936	PERV_Q4VFZ2_3mut
PAPAPAPAPAP	17,937	KORV_Q9TTC1-Pro_3mutA
EAAAK	17,938	MLVMS_P03355_3mutA_WS
GGG	17,939	MLVCB_P08361_3mut
GGSEAAAKGGG	17,940	BAEVM_P10272_3mutA
GGGGGSGSS	17,941	MLVAV_P03356_3mutA
EAAAKGSSPAP	17,942	MLVBM_Q7SVK7_3mutA_WS
GGSGGSGGS	17,943	XMRV6_A1Z651_3mut
EAAAKPAPGGG	17,944	KORV_Q9TTC1-Pro_3mutA
GGGPAPEAAAK	17,945	FLV_P10273_3mutA
GGSPAPEAAAK	17,946	MLVMS_P03355_3mutA_WS
GGSGGSGGSGGSGGS	17,947	MLVFF_P26809_3mut
EAAAKGGSGSS	17,948	MLVMS_P03355_PLV919
GGGEAAAKGGS	17,949	MLVBM_Q7SVK7_3mutA_WS
PAPAPAPAP	17,950	BAEVM_P10272_3mutA
EAAAKEAAAKEAAAKEAAAK	17,951	MLVMS_P03355_3mut
EAAAKPAP	17,952	XMRV6_A1Z651_3mut
EAAAKEAAAK	17,953	MLVBM_Q7SVK7_3mutA_WS
EAAAKGGG	17,954	BAEVM_P10272_3mut
EAAAKGSS	17,955	MLVAV_P03356_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,956	MLVFF_P26809_3mut
GGGPAPGSS	17,957	PERV_Q4VFZ2_3mutA_WS
GGGG	17,958	PERV_Q4VFZ2_3mut
EAAAKGGSGSS	17,959	MLVMS_P03355_PLV919
GGGGSGGGGSGGGGS	17,960	MLVMS_P03355_3mutA_WS
EAAAK	17,961	MLVMS_P03355_3mutA_WS
GGGGSS	17,962	PERV_Q4VFZ2
PAPEAAAKGGS	17,963	MLVCB_P08361_3mut
GSS	17,964	MLVMS_P03355_3mut
GSAGSAAGSGEF	17,965	MLVFF_P26809_3mutA

[1627]

EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,966	KORV_Q9TTC1-Pro_3mut
GGGGSGGGGS	17,967	AVIRE_P03360_3mutA
EAAAK	17,968	MLVMS_P03355_3mut
GGGPAPGGS	17,969	PERV_Q4VFZ2_3mut
GGGGSGGGGSGGGGS	17,970	MLVMS_P03355_PLV919
PAPGGG	17,971	MLVMS_P03355_3mutA_WS
GGGEAAAKPAP	17,972	PERV_Q4VFZ2_3mutA_WS
EAAAKPAPGSS	17,973	KORV_Q9TTC1-Pro_3mutA
PAPGSS	17,974	KORV_Q9TTC1_3mutA
GSAGSAAGSGEF	17,975	PERV_Q4VFZ2_3mut
PAPGGGGSS	17,976	KORV_Q9TTC1-Pro_3mutA
GSSGGGEAAAK	17,977	MLVCB_P08361_3mutA
GSS	17,978	AVIRE_P03360_3mutA
GSSGSSGSSGSS	17,979	XMRV6_A1Z651_3mutA
PAPEAAAKGGG	17,980	MLVMS_P03355_PLV919
GGGPAPEAAAK	17,981	MLVCB_P08361_3mutA
PAPGGGGGS	17,982	MLVCB_P08361_3mutA
EAAAKEAAAKEAAAKEAAAK	17,983	PERV_Q4VFZ2_3mutA_WS
GGGGGSPAP	17,984	MLVFF_P26809_3mutA
GSSGSSGSSGSSGSS	17,985	PERV_Q4VFZ2
GSSPAPEAAAK	17,986	MLVMS_P03355_PLV919
GSSGSSGSSGSSGSSGSS	17,987	MLVBM_Q7SVK7_3mutA_WS
GSSGSSGSSGSSGSSGSS	17,988	MLVMS_P03355_3mutA_WS
GGSPAPEAAAK	17,989	MLVAV_P03356_3mutA
GSSGGG	17,990	BAEVM_P10272_3mut
EAAAKGSSGGS	17,991	KORV_Q9TTC1-Pro_3mutA
GGSGSSEAAAK	17,992	MLVMS_P03355_3mutA_WS
GGGPAPEAAAK	17,993	MLVFF_P26809_3mutA
GGGPAPGGS	17,994	MLVMS_P03355_3mutA_WS
GGGGG	17,995	MLVMS_P03355_PLV919
GGGEAAAKPAP	17,996	MLVBM_Q7SVK7_3mutA_WS
GGGGSGGGGS	17,997	WMSV_P03359_3mut
GGGPAPEAAAK	17,998	PERV_Q4VFZ2_3mut
GGSGSSEAAAK	17,999	MLVMS_P03355_PLV919
EAAAKGGGPAP	18,000	MLVMS_P03355_3mutA_WS
GSSGSSGSSGSSGSS	18,001	KORV_Q9TTC1-Pro_3mutA
PAPAP	18,002	WMSV_P03359_3mutA
GGSPAPGSS	18,003	MLVAV_P03356_3mutA
GGSGGGPAP	18,004	MLVMS_P03355_3mut
GGSPAP	18,005	MLVMS_P03355_PLV919

[1628]

EAAAKGGSPAP	18,006	PERV_Q4VFZ2_3mut
GSSPAPGGG	18,007	KORV_Q9TTC1-Pro_3mutA
GSAGSAAGSGEF	18,008	MLVMS_P03355_3mut
GGSPAP	18,009	PERV_Q4VFZ2_3mut
GSSGSS	18,010	KORV_Q9TTC1-Pro_3mut
GGGPAPGSS	18,011	MLVMS_P03355_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	18,012	FOAMV_P14350
PAPGSSGGG	18,013	MLVMS_P03355_PLV919
GGSEAAAKPAP	18,014	BAEVM_P10272_3mutA
GGGGGS	18,015	MLVCB_P08361_3mutA
PAPEAAAKGGS	18,016	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	18,017	BAEVM_P10272_3mutA
GGSEAAAK	18,018	BAEVM_P10272_3mutA
GSSPAPEAAAK	18,019	MLVMS_P03355_3mutA_WS
PAPGGG	18,020	WMSV_P03359_3mut
EAAAKPAP	18,021	PERV_Q4VFZ2_3mut
GSSGSSGSSGSSGSS	18,022	WMSV_P03359_3mut
PAPGGG	18,023	MLVBM_Q7SVK7_3mutA_WS
GGSGGGEAAAK	18,024	BAEVM_P10272_3mutA
PAPGGS	18,025	MLVMS_P03355_3mut
GGSGGSGGSGGS	18,026	MLVBM_Q7SVK7_3mutA_WS
EAAAKEAAAKEAAAKEAAAK	18,027	PERV_Q4VFZ2_3mut
GGSEAAAKGGG	18,028	WMSV_P03359_3mutA
GGGPAP	18,029	BAEVM_P10272_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	18,030	XMRV6_A1Z651_3mut
GGSPAPGSS	18,031	KORV_Q9TTC1_3mut
GGGPAPGSS	18,032	MLVMS_P03355_3mut
GGGGSSGGS	18,033	BAEVM_P10272_3mutA
GGGEAAAKGSS	18,034	KORV_Q9TTC1-Pro_3mutA
PAPAP	18,035	MLVBM_Q7SVK7_3mutA_WS
GGSPAPGGG	18,036	PERV_Q4VFZ2_3mut
PAPGSS	18,037	PERV_Q4VFZ2_3mutA_WS
GSSGGSPAP	18,038	MLVBM_Q7SVK7_3mutA_WS
EAAAKGGGGSEAAAK	18,039	PERV_Q4VFZ2_3mut
GSSEAAAKGGS	18,040	KORV_Q9TTC1-Pro_3mut
PAPAPAPAP	18,041	KORV_Q9TTC1-Pro_3mutA
GGSEAAAKPAP	18,042	WMSV_P03359_3mutA

[1629]

PAPGGS	18,043	FLV_P10273_3mutA
EAAAKGGGPAP	18,044	PERV_Q4VFZ2_3mut
GGSGSSGGG	18,045	AVIRE_P03360_3mutA
EAAAKGGSGSS	18,046	BAEVM_P10272_3mutA
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	18,047	MLVCB_P08361_3mutA
GSSEAAAKGGS	18,048	XMRV6_A1Z651_3mutA
GGGGG	18,049	BAEVM_P10272_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	18,050	SFV3L_P27401_2mutA
GGGEAAAKGSS	18,051	MLVMS_P03355_PLV919
EAAAKGGGGSEAAAK	18,052	KORV_Q9TTC1_3mutA
EAAAKGGG	18,053	AVIRE_P03360_3mut
GGSGGG	18,054	MLVMS_P03355_3mutA_WS
GGSGSSGGG	18,055	MLVMS_P03355_PLV919
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	18,056	KORV_Q9TTC1_3mut
GGGGSEAAAKGGGGS	18,057	KORV_Q9TTC1_3mutA
PAPAPAPAPAP	18,058	FLV_P10273_3mutA
GGS	18,059	MLVBM_Q7SVK7_3mutA_WS
GGGGGSEAAAK	18,060	MLVBM_Q7SVK7_3mutA_WS
GSSGSSGSSGSSGSS	18,061	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAK	18,062	MLVMS_P03355_3mut
GGSGSSGGG	18,063	PERV_Q4VFZ2_3mut
PAP	18,064	MLVFF_P26809_3mut
GSSPAPEAAAK	18,065	MLVAV_P03356_3mutA
EAAAKGGGGSS	18,066	MLVMS_P03355_3mut
GGGEAAAKGGS	18,067	XMRV6_A1Z651_3mut
GGSGGGPAP	18,068	MLVBM_Q7SVK7_3mutA_WS
GSAGSAAGSGEF	18,069	BAEVM_P10272_3mutA
GSSEAAAK	18,070	MLVCB_P08361_3mut
PAPGSS	18,071	MLVMS_P03355_3mut
EAAAKEAAAKEAAAK	18,072	MLVAV_P03356_3mutA
GSAGSAAGSGEF	18,073	XMRV6_A1Z651_3mutA
GSSGSSGSSGSS	18,074	BAEVM_P10272_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	18,075	KORV_Q9TTC1-Pro_3mut
GGGGSSEAAAK	18,076	WMSV_P03359_3mut
GSSGGGEAAAK	18,077	MLVBM_Q7SVK7_3mutA_WS
EAAAKPAP	18,078	MLVFF_P26809_3mutA
GGSPAPGGG	18,079	KORV_Q9TTC1_3mutA

[1630]

[1631]

PAPEAAAK	18,080	FLV_P10273_3mutA
GSSGSSGSS	18,081	MLVBM_Q7SVK7_3mutA_WS
GSSGGGEAAAK	18,082	FLV_P10273_3mutA
GGSPAP	18,083	MLVBM_Q7SVK7_3mutA_WS
GSAGSAAGSGEF	18,084	KORV_Q9TTC1-Pro_3mutA
PAPGGSEAAAK	18,085	MLVMS_P03355_PLV919
GGSPAPEAAAK	18,086	MLVBM_Q7SVK7_3mutA_WS
GGGGGSPAP	18,087	MLVBM_Q7SVK7_3mutA_WS
EAAAKGSSPAP	18,088	WMSV_P03359_3mut
EAAAKGGGPAP	18,089	MLVBM_Q7SVK7_3mutA_WS
PAPGSS	18,090	KORV_Q9TTC1-Pro_3mutA
GGSGSSGGG	18,091	BAEVM_P10272_3mut
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	16,901	FFV_O93209-Pro_2mut
GGSGGSGGSGGSGGSGGS	16,902	WMSV_P03359_3mutA
GGSGGSGGS	16,903	PERV_Q4VFZ2_3mutA_WS
GGGGG	16,904	PERV_Q4VFZ2_3mutA_WS
GGGPAP	16,905	FLV_P10273_3mutA
PAPGGSGGG	16,906	XMRV6_A1Z651_3mutA
GGGGSEAAAKGGGGS	16,907	XMRV6_A1Z651_3mut
EAAAKGSSGGG	16,908	KORV_Q9TTC1-Pro_3mutA
GSSGGSEAAAK	16,909	WMSV_P03359_3mut
EAAAKGGSGSS	16,910	PERV_Q4VFZ2_3mut
PAPAPAPAPAP	16,911	PERV_Q4VFZ2_3mut
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	16,912	MLVMS_P03355_3mutA_WS
GGGGGGG	16,913	KORV_Q9TTC1_3mutA
EAAAK	16,914	KORV_Q9TTC1-Pro_3mutA
GGGEAAAKGGS	16,915	KORV_Q9TTC1-Pro_3mutA
GGGEAAAKGGS	16,916	PERV_Q4VFZ2_3mutA_WS
GGGGGSPAP	16,917	XMRV6_A1Z651_3mut
GGGGSGGGGSGGGGSGGGGS	16,918	MLVFF_P26809_3mut
GGGGGGG	16,919	MLVFF_P26809_3mut
PAPAPAPAPAPAP	16,920	AVIRE_P03360_3mutA
GSSPAPGGG	16,921	FLV_P10273_3mutA
GGGGGSPAP	16,922	MLVMS_P03355_3mutA_WS
GGGGSGGGGSGGGGS	16,923	MLVMS_P03355_3mut
GGGGSGGGGSGGGGS	16,924	KORV_Q9TTC1_3mut
GSSEAAAKGGS	16,925	MLVAV_P03356_3mutA

GSSGSSGSSGSSGSS	16,926	MLVMS_P03355_3mut
EAAAKGGGGGS	16,927	PERV_Q4VFZ2_3mutA_WS
GSSGGGGGS	16,928	PERV_Q4VFZ2_3mut
GGGEAAAKPAP	16,929	MLVMS_P03355_3mut
GSSGGSPAP	16,930	PERV_Q4VFZ2_3mutA_WS
GSSGGGPAP	16,931	BAEVM_P10272_3mutA
GGGGGSGSS	16,932	MLVMS_P03355_PLV919
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	16,933	BAEVM_P10272_3mut
PAPEAAAK	16,934	MLVMS_P03355_3mut
GGGGSGGGGSGGGGS	16,935	FLV_P10273_3mutA
GGSGSSGGG	16,936	WMSV_P03359_3mutA
EAAAKGGS	16,937	PERV_Q4VFZ2_3mut
EAAAKGSSPAP	16,938	MLVCB_P08361_3mut
EAAAKGGSGSS	16,939	WMSV_P03359_3mutA
GSSGSS	16,940	PERV_Q4VFZ2_3mutA_WS
PAPAPAPAP	16,941	MLVMS_P03355_PLV919
GGSGGG	16,942	PERV_Q4VFZ2_3mutA_WS
GSS	16,943	MLVBM_Q75VK7_3mutA_WS
PAP	16,944	KORV_Q9TTC1-Pro_3mutA
GGSGSSEAAAK	16,945	MLVFF_P26809_3mut
PAPEAAAKGSS	16,946	KORV_Q9TTC1-Pro_3mutA
GGSGGS	16,947	MLVCB_P08361_3mutA
GGGGGGG	16,948	PERV_Q4VFZ2_3mutA_WS
GGSPAPEAAAK	16,949	MLVBM_Q75VK7_3mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	16,950	KORV_Q9TTC1_3mutA
GGSPAP	16,951	MLVMS_P03355_3mut
GGSEAAAKGGG	16,952	PERV_Q4VFZ2_3mut
GGGGSGGGGS	16,953	FLV_P10273_3mutA
GGGEAAAK	16,954	BAEVM_P10272_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	16,955	SFV3L_P27401_2mut
GGSEAAAKPAP	16,956	KORV_Q9TTC1-Pro_3mutA
GSSGGGEAAAK	16,957	MLVMS_P03355_PLV919
GGGGGSEAAAK	16,958	MLVMS_P03355_PLV919
EAAAKGGSGGG	16,959	MLVMS_P03355_3mutA_WS
GGGGSSPAP	16,960	MLVAV_P03356_3mutA
EAAAKEAAAK	16,961	MLVMS_P03355_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	16,962	SFV3L_P27401_2mut
GSSGSSGSSGSSGSS	16,963	MLVMS_P03355_PLV919
GSSGGG	16,964	KORV_Q9TTC1-Pro_3mutA

[1632]

[1633]

GSSGGS	16,965	MLVFF_P26809_3mutA
GGGGSGGGGS	16,966	XMRV6_A1Z651_3mutA
PAPGSS	16,967	MLVBM_Q7SVK7_3mutA_WS
GGGPAPEAAAK	16,968	XMRV6_A1Z651_3mutA
EAAAKGGS	16,969	MLVFF_P26809_3mut
GSS	16,970	KORV_Q9TTC1_3mutA
GGGG	16,971	PERV_Q4VFZ2_3mut
GGGGGSEAAAK	16,972	AVIRE_P03360_3mutA
GSSGSSGSSGSSGSS	16,973	MLVMS_P03355_PLV919
PAPGGSGGG	16,974	PERV_Q4VFZ2_3mut
GGGPAP	16,975	PERV_Q4VFZ2_3mut
GGGPAPEAAAK	16,976	AVIRE_P03360_3mutA
GGGEAAAK	16,977	MLVCB_P08361_3mut
GGG	16,978	MLVFF_P26809_3mutA
EAAAKPAPGSS	16,979	XMRV6_A1Z651_3mutA
GGSGSSEAAAK	16,980	PERV_Q4VFZ2_3mutA_WS
EAAAKGSS	16,981	MLVMS_P03355_3mut
GGSGSSEAAAK	16,982	BAEVM_P10272_3mut
GGSGGG	16,983	MLVBM_Q7SVK7_3mutA_WS
GGGPAP	16,984	MLVMS_P03355_PLV919
GGSPAPGGG	16,985	PERV_Q4VFZ2_3mutA_WS
GGGGGSEAAAK	16,986	MLVFF_P26809_3mutA
EAAAKGSSGGS	16,987	MLVBM_Q7SVK7_3mut
PAPAP	16,988	XMRV6_A1Z651_3mut
GSSPAPGGS	16,989	MLVBM_Q7SVK7_3mutA_WS
GSSEAAAKGGG	16,990	WMSV_P03359_3mutA
EAAAKGGGGGS	16,991	PERV_Q4VFZ2_3mut
GSSGSSGSSGSSGSS	16,992	MLVCB_P08361_3mutA
EAAAKGGGGSS	16,993	PERV_Q4VFZ2_3mut
EAAAKGSS	16,994	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	16,995	AVIRE_P03360_3mutA
EAAAKGGS	16,996	MLVCB_P08361_3mut
GSSGGSEAAAK	16,997	MLVAV_P03356_3mutA
EAAAKPAPGGS	16,998	PERV_Q4VFZ2_3mut
GGSGGS	16,999	MLVAV_P03356_3mutA
EAAAKGSSGGG	17,000	AVIRE_P03360_3mutA
GGSGGSGGSGGS	17,001	PERV_Q4VFZ2_3mut
GGGGGGGG	17,002	KORV_Q9TTC1_3mutA
GGSGSSEAAAK	17,003	MLVCB_P08361_3mutA

EAAAKGGG	17,004	MLVBM_Q7SVK7_3mutA_WS
GGGGSGGGGSGGGGS	17,005	MLVCB_P08361_3mut
GGSGGSGGSGGS	17,006	PERV_Q4VFZ2_3mutA_WS
PAPAPAPAPAP	17,007	WMSV_P03359_3mut
EAAAKEAAAKEAAAKEAAAK	17,008	PERV_Q4VFZ2_3mut
GGSGGSGGS	17,009	XMRV6_A1Z651_3mutA
PAPGGGGSS	17,010	BAEVM_P10272_3mutA
GSSEAAAKGGS	17,011	MLVCB_P08361_3mut
GSSGGGPAP	17,012	MLVCB_P08361_3mutA
GGSGSS	17,013	MLVBM_Q7SVK7_3mutA_WS
GGGGGSEAAAK	17,014	MLVAV_P03356_3mutA
GSSEAAAK	17,015	PERV_Q4VFZ2_3mutA_WS
GGGGGSGSS	17,016	MLVBM_Q7SVK7_3mutA_WS
EAAAKGGSGSS	17,017	MLVFF_P26809_3mut
PAP	17,018	FLV_P10273_3mutA
GGGGG	17,019	MLVMS_P03355_3mutA_WS
EAAAK	17,020	PERV_Q4VFZ2_3mut
GSS	17,021	FLV_P10273_3mutA
PAPAPAPAPAPAP	17,022	KORV_Q9TTC1-Pro_3mutA
EAAAKEAAAKEAAAKEAAAK	17,023	MLVCB_P08361_3mut
EAAAKGGGGSEAAAK	17,024	XMRV6_A1Z651_3mut
PAPGGSGGG	17,025	MLVBM_Q7SVK7_3mutA_WS
GGSGGGPAP	17,026	WMSV_P03359_3mutA
GGGGSSEAAAK	17,027	MLVBM_Q7SVK7_3mutA_WS
PAPGGGGSS	17,028	MLVCB_P08361_3mut
GGSGGSGGSGGS	17,029	PERV_Q4VFZ2_3mutA_WS
PAPGGSGGG	17,030	MLVMS_P03355_3mutA_WS
GSSPAPGGS	17,031	MLVCB_P08361_3mutA
GSSGSSGSS	17,032	MLVFF_P26809_3mut
PAPGGGGGS	17,033	MLVBM_Q7SVK7_3mutA_WS
GSSPAP	17,034	PERV_Q4VFZ2_3mut
GGSGGG	17,035	KORV_Q9TTC1-Pro_3mut
EAAAKGGGGSEAAAK	17,036	PERV_Q4VFZ2_3mutA_WS
GGSPAPEAAAK	17,037	PERV_Q4VFZ2_3mutA_WS
EAAAKPAP	17,038	BAEVM_P10272_3mut
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,039	MLVMS_P03355_3mut
EAAAKGGGGSS	17,040	MLVFF_P26809_3mut

[1634]

EAAAKEAAAK	17,041	MLVCB_P08361_3mut
GSSEAAAKGGS	17,042	PERV_Q4VFZ2_3mut
GGSPAP	17,043	KORV_Q9TTC1-Pro_3mutA
EAAAKEAAAKEAAAKEAAAK	17,044	MLVMS_P03355_3mutA_WS
GSSGSSGSSGSSGSS	17,045	BAEVM_P10272_3mut
PAPEAAAK	17,046	MLVMS_P03355_3mut
GSSGGSPAP	17,047	PERV_Q4VFZ2
GGGPAPGGS	17,048	BAEVM_P10272_3mutA
EAAAKPAPGGS	17,049	MLVMS_P03355_PLV919
GGGGSGGGGS	17,050	PERV_Q4VFZ2
GGGEAAAK	17,051	KORV_Q9TTC1-Pro_3mut
EAAAKGGGGGS	17,052	FLV_P10273_3mutA
GGSPAPGSS	17,053	MLVMS_P03355_3mut
GSSPAPEAAAK	17,054	MLVMS_P03355_3mutA_WS
GSAGSAAGSGEF	17,055	MLVBM_Q7SVK7_3mutA_WS
EAAAK	17,056	BAEVM_P10272_3mutA
EAAAKGGGGSS	17,057	BAEVM_P10272_3mutA
GGG	17,058	WMSV_P03359_3mut
GGSGSSPAP	17,059	BAEVM_P10272_3mut
GGSEAAAKPAP	17,060	MLVBM_Q7SVK7_3mutA_WS
EAAAKGGSGSS	17,061	MLVCB_P08361_3mut
PAPGSS	17,062	MLVAV_P03356_3mutA
PAPEAAAKGGG	17,063	MLVCB_P08361_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,064	FOAMV_P14350-Pro_2mut
GSSGSSGSS	17,065	PERV_Q4VFZ2_3mut
PAPGGG	17,066	MLVMS_P03355_3mut
PAPGGS	17,067	PERV_Q4VFZ2_3mut
GSSGGG	17,068	MLVMS_P03355_PLV919
GSSGSSGSSGSSGSSGSS	17,069	WMSV_P03359_3mut
PAP	17,070	AVIRE_P03360_3mutA
EAAAKGSSPAP	17,071	MLVBM_Q7SVK7_3mutA_WS
GSSGSSGSSGSS	17,072	MLVMS_P03355_PLV919
GGGGSGGGGSGGGGSGGGGSGGGGS	17,073	AVIRE_P03360
GGGGS	17,074	PERV_Q4VFZ2_3mut
EAAAKGSSGGG	17,075	MLVBM_Q7SVK7_3mutA_WS
GGGGGG	17,076	KORV_Q9TTC1-Pro_3mut
GGSGSSEAAAK	17,077	PERV_Q4VFZ2_3mut

[1635]

GSSPAPEAAAK	17,078	MLVBM_Q7SVK7_3mutA_WS
GGGGSGGGGS	17,079	MLVBM_Q7SVK7_3mutA_WS
GSSGGGGGS	17,080	MLVAV_P03356_3mutA
GSAGSAAGSGEF	17,081	WMSV_P03359_3mutA
GGGEAAAKGSS	17,082	BAEVM_P10272_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,083	FFV_O93209-Pro_2mut
PAPGGSGGG	17,084	MLVCB_P08361_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,085	SFV3L_P27401_2mut
GGSGSSPAP	17,086	MLVMS_P03355_PLV919
GGGGGG	17,087	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAKEAAAKEAAAK	17,088	PERV_Q4VFZ2_3mut
EAAAKGSSPAP	17,089	MLVFF_P26809_3mut
GGGPAPGGS	17,090	MLVBM_Q7SVK7_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,091	SFV3L_P27401
PAP	17,092	PERV_Q4VFZ2_3mut
EAAAKGGS	17,093	MLVMS_P03355_PLV919
GSSGGSEAAAK	17,094	WMSV_P03359_3mutA
GGSGSSEAAAK	17,095	KORV_Q9TTC1-Pro_3mutA
EAAAKEAAAKEAAAK	17,096	PERV_Q4VFZ2
GGSGGGEAAAK	17,097	MLVMS_P03355_3mutA_WS
GGGGSGGGGSGGGGSGGGGS	17,098	BAEVM_P10272_3mut
EAAAKGSS	17,099	XMRV6_A1Z651_3mutA
GSSGGGGGS	17,100	WMSV_P03359_3mutA
GSSGSSGSSGSSGSSGSS	17,101	MLVFF_P26809_3mutA
GGSGSS	17,102	MLVAV_P03356_3mutA
EAAAKGGGGSEAAAK	17,103	MLVMS_P03355_PLV919
EAAAKGGGPAP	17,104	PERV_Q4VFZ2
GGSEAAAKGGG	17,105	MLVAV_P03356_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,106	MLVBM_Q7SVK7_3mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,107	KORV_Q9TTC1-Pro_3mutA
GSSPAPEAAAK	17,108	MLVFF_P26809_3mutA
GGGGSEAAAKGGGGS	17,109	PERV_Q4VFZ2_3mut
GSSGSSGSSGSS	17,110	PERV_Q4VFZ2_3mut
GGSEAAAK	17,111	MLVFF_P26809_3mutA
GGGGGGGG	17,112	MLVMS_P03355_3mut
GSSGGG	17,113	XMRV6_A1Z651_3mutA
EAAAKGGS	17,114	BAEVM_P10272_3mutA

[1636]

GGGGS	17,115	BAEVM_P10272_3mutA
GGSEAAAKGGG	17,116	KORV_Q9TTC1-Pro_3mutA
GGSGSSGGG	17,117	KORV_Q9TTC1_3mutA
GGSGSSEAAAK	17,118	WMSV_P03359_3mut
EAAAKGGSGSS	17,119	MLVBM_Q7SVK7_3mutA_WS
GGS	17,120	BAEVM_P10272_3mutA
GGGPAPGSS	17,121	WMSV_P03359_3mutA
GSSGSSGSSGSSGSS	17,122	AVIRE_P03360_3mut
GGGEAAAKPAP	17,123	XMRV6_A1Z651_3mut
GSSGGG	17,124	MLVFF_P26809_3mutA
GGSPAPGSS	17,125	PERV_Q4VFZ2_3mut
PAPGGS	17,126	MLVCB_P08361_3mut
PAPAPAPAPAP	17,127	KORV_Q9TTC1_3mutA
GSSGGS	17,128	MLVCB_P08361_3mutA
GSSGGSEAAAK	17,129	PERV_Q4VFZ2_3mut
EAAAKGSSGGS	17,130	MLVMS_P03355_PLV919
EAAAKGGG	17,131	WMSV_P03359_3mut
PAPGGGGGS	17,132	BAEVM_P10272_3mutA
GGGGSEAAAKGGGGS	17,133	WMSV_P03359_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,134	MLVMS_P03355_3mutA_WS
GGS	17,135	KORV_Q9TTC1-Pro_3mutA
GSSGGSPAP	17,136	BAEVM_P10272_3mutA
GGG	17,137	MLVMS_P03355_PLV919
PAPGSS	17,138	KORV_Q9TTC1-Pro_3mut
GGSEAAAKGGG	17,139	FLV_P10273_3mutA
GGSEAAAKPAP	17,140	PERV_Q4VFZ2_3mutA_WS
GGGGSSPAP	17,141	XMRV6_A1Z651_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,142	PERV_Q4VFZ2_3mutA_WS
GGGG	17,143	PERV_Q4VFZ2_3mutA_WS
GGSEAAAKPAP	17,144	MLVMS_P03355_3mut
PAPGSSGGG	17,145	MLVMS_P03355_3mutA_WS
PAPEAAAKGGS	17,146	AVIRE_P03360_3mut
GGGGSSPAP	17,147	MLVMS_P03355_3mutA_WS
GGGGSGGGGSGGGGSGGGGS	17,148	PERV_Q4VFZ2_3mut
GGGEAAAK	17,149	MLVMS_P03355_3mut
GGGGSS	17,150	MLVFF_P26809_3mut
GGSPAPGSS	17,151	XMRV6_A1Z651_3mut
GGGGS	17,152	KORV_Q9TTC1-Pro_3mutA
EAAAKGSSGGS	17,153	FLV_P10273_3mutA
GSS	17,154	MLVMS_P03355_PLV919
GGGG	17,155	MLVMS_P03355_PLV919

[1637]

[1638]

GSSGGS	17,156	MLVMS_P03355_PLV919
GGSGGSGGSGGS	17,157	MLVMS_P03355_3mut
PAPEAAAKGGS	17,158	MLVMS_P03355_3mut
EAAAKGSSGGG	17,159	BAEVM_P10272_3mutA
GSSEAAAK	17,160	KORV_Q9TTC1-Pro_3mutA
GSAGSAAGSGEF	17,161	KORV_Q9TTC1_3mutA
GGGGGSEAAAK	17,162	MLVCB_P08361_3mut
GGGG	17,163	WMSV_P03359_3mut
GGGGSSEAAAK	17,164	MLVMS_P03355_PLV919
PAPGGG	17,165	WMSV_P03359_3mutA
EAAAKGGSGGG	17,166	MLVAV_P03356_3mutA
GGGPAPGGS	17,167	MLVMS_P03355_3mut
EAAAKPAP	17,168	PERV_Q4VFZ2_3mutA_WS
GSSGSSGSS	17,169	KORV_Q9TTC1-Pro_3mutA
GSSPAPGGS	17,170	XMRV6_A1Z651_3mut
GGGGGSPAP	17,171	BAEVM_P10272_3mutA
GGSGSSGGG	17,172	PERV_Q4VFZ2_3mutA_WS
GGGEAAAKGSS	17,173	AVIRE_P03360_3mut
GSSEAAAK	17,174	FLV_P10273_3mutA
EAAAK	17,175	MLVMS_P03355_3mut
EAAAKGGSGSS	17,176	WMSV_P03359_3mut
GSSEAAAKGGG	17,177	PERV_Q4VFZ2_3mut
PAPGSSGGG	17,178	BAEVM_P10272_3mutA
EAAAKGGGGGS	17,179	MLVMS_P03355_3mut
GGSEAAAKPAP	17,180	AVIRE_P03360_3mut
GGGPAPGGS	17,181	XMRV6_A1Z651_3mut
GGGGS	17,182	KORV_Q9TTC1_3mutA
GGSGGSGGSGGSGGS	17,183	XMRV6_A1Z651_3mut
GGGPAP	17,184	KORV_Q9TTC1-Pro_3mut
EAAAKPAP	17,185	MLVBM_Q7SVK7_3mutA_WS
GGSEAAAK	17,186	MLVMS_P03355_PLV919
GSSEAAAKPAP	17,187	KORV_Q9TTC1-Pro_3mutA
GGSGSS	17,188	MLVMS_P03355_3mut
EAAAKPAPGGG	17,189	PERV_Q4VFZ2_3mut
GGSPAPEAAAK	17,190	KORV_Q9TTC1_3mutA
GGSEAAAKGGG	17,191	AVIRE_P03360_3mutA
GGGGSEAAAKGGGGS	17,192	MLVMS_P03355_PLV919
GSSGGGEAAAK	17,193	KORV_Q9TTC1-Pro_3mutA
EAAAKGGGPAP	17,194	WMSV_P03359_3mut
GSSPAP	17,195	XMRV6_A1Z651_3mutA

AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,196	SFV3L_P27401-Pro
GGSEAAAKGSS	17,197	MLVMS_P03355_PLV919
GSSGGSEAAAK	17,198	KORV_Q9TTC1-Pro_3mutA
GGSEAAAKGSS	17,199	KORV_Q9TTC1-Pro_3mutA
EAAAKGGG	17,200	AVIRE_P03360_3mutA
GSSGGSEAAAK	17,201	BAEVM_P10272_3mutA
GGGGSEAAAKGGGGS	17,202	KORV_Q9TTC1-Pro_3mut
PAPGSSEAAAK	17,203	MLVMS_P03355_3mut
PAPEAAAK	17,204	WMSV_P03359_3mut
PAPGGSGSS	17,205	PERV_Q4VFZ2_3mutA_WS
PAPGSS	17,206	BAEVM_P10272_3mut
PAPGGGGGS	17,207	MLVMS_P03355_3mut
EAAAKPAPGSS	17,208	MLVBM_Q75VK7_3mutA_WS
GSSPAPGGS	17,209	MLVMS_P03355_PLV919
GGSGSSEAAAK	17,210	MLVMS_P03355_3mut
GGGGGG	17,211	KORV_Q9TTC1-Pro_3mutA
EAAAKEAAAKEAAAKEAAAK	17,212	MLVBM_Q75VK7_3mut
GGSPAPGSS	17,213	MLVMS_P03355_PLV919
PAPAPAPAPAP	17,214	MLVCB_P08361_3mut
GGSGSSPAP	17,215	WMSV_P03359_3mutA
EAAAKGGSGGG	17,216	PERV_Q4VFZ2_3mutA_WS
GSSGSSGSSGSSGSS	17,217	PERV_Q4VFZ2_3mut
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,218	KORV_Q9TTC1_3mutA
GSSGGGEAAAK	17,219	WMSV_P03359_3mutA
GSSGGSEAAAK	17,220	FLV_P10273_3mutA
GGGGGGGG	17,221	PERV_Q4VFZ2_3mut
PAPGGSEAAAK	17,222	FLV_P10273_3mutA
GGGGSSPAP	17,223	BAEVM_P10272_3mutA
PAPAPAPAP	17,224	WMSV_P03359_3mut
GGSEAAAKPAP	17,225	PERV_Q4VFZ2_3mut
PAPGGSGGG	17,226	BAEVM_P10272_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,227	MLVMS_P03355_3mut
GGGGSGGGGSGGGGS	17,228	PERV_Q4VFZ2_3mut
GGSGGGPAP	17,229	PERV_Q4VFZ2_3mut
GGGPAPEAAAK	17,230	MLVFF_P26809_3mut
GGGGGSGSS	17,231	MLVMS_P03355_3mutA_WS
GSS	17,232	MLVCB_P08361_3mut
GGGGGSPAP	17,233	MLVMS_P03355_PLV919
GGSPAP	17,234	MLVAV_P03356_3mutA

[1639]

GGGPAPGGS	17,235	KORV_Q9TTC1-Pro_3mutA
PAPGSSGGG	17,236	FLV_P10273_3mutA
PAPGSSGGG	17,237	WMSV_P03359_3mutA
PAPGGS	17,238	MLVBM_Q7SVK7_3mutA_WS
GGGEAAAKGSS	17,239	PERV_Q4VFZ2_3mutA_WS
GGSEAAAKGSS	17,240	MLVBM_Q7SVK7_3mutA_WS
PAPGGSEAAAK	17,241	MLVCB_P08361_3mut
GGSEAAAKGGG	17,242	XMRV6_A1Z651_3mutA
GGSGGGGSS	17,243	WMSV_P03359_3mut
GGGEAAAKPAP	17,244	KORV_Q9TTC1_3mutA
EAAAKGSS	17,245	KORV_Q9TTC1-Pro_3mut
PAPEAAAKGSS	17,246	MLVFF_P26809_3mut
GSAGSAAGSGEF	17,247	PERV_Q4VFZ2_3mut
EAAAKGGGGGS	17,248	WMSV_P03359_3mut
EAAAKGSSPAP	17,249	WMSV_P03359_3mutA
GGGGSEAAAKGGGGS	17,250	XMRV6_A1Z651_3mutA
GSSEAAAKPAP	17,251	SFV3L_P27401-Pro_2mutA
GGGGGG	17,252	PERV_Q4VFZ2_3mutA_WS
PAPGGS	17,253	BAEVM_P10272_3mut
PAP	17,254	AVIRE_P03360_3mut
PAPAPAP	17,255	MLVBM_Q7SVK7_3mutA_WS
GGGG	17,256	PERV_Q4VFZ2_3mutA_WS
GSSGGSEAAAK	17,257	MLVBM_Q7SVK7_3mut
GGSGGGGSS	17,258	MLVFF_P26809_3mut
GGGGSSGGS	17,259	AVIRE_P03360_3mutA
GSSPAPGGG	17,260	PERV_Q4VFZ2_3mutA_WS
GGSEAAAKPAP	17,261	MLVMS_P03355_PLV919
PAP	17,262	KORV_Q9TTC1-Pro_3mut
GSSGGS	17,263	PERV_Q4VFZ2_3mut
GGGGG	17,264	PERV_Q4VFZ2_3mut
GSSGGGPAP	17,265	FLV_P10273_3mutA
GSSEAAAKGGG	17,266	KORV_Q9TTC1-Pro_3mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,267	MLVCB_P08361_3mut
GGSEAAAKPAP	17,268	MLVCB_P08361_3mut
PAPAPAPAPAPAP	17,269	BAEVM_P10272_3mutA
GGGGSEAAAKGGGGS	17,270	MLVMS_P03355_3mut
EAAAKPAPGSS	17,271	MLVMS_P03355_3mut
GSSGSSGSSGSSGSS	17,272	MLVBM_Q7SVK7_3mutA_WS

[1640]

PAPEAAAKGSS	17,273	MLVAV_P03356_3mut
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,274	AVIRE_P03360_3mut
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,275	PERV_Q4VFZ2_3mut
GGSEAAAKGGG	17,276	PERV_Q4VFZ2_3mutA_WS
GGSGGGGSS	17,277	MLVFF_P26809_3mutA
PAPEAAAKGSS	17,278	MLVCB_P08361_3mut
GGG	17,279	PERV_Q4VFZ2_3mutA_WS
GGSGGGEAAAK	17,280	MLVMS_P03355_3mut
EAAAKGGGGSS	17,281	WMSV_P03359_3mut
GSSPAPGGG	17,282	WMSV_P03359_3mutA
EAAAKGSSGGG	17,283	PERV_Q4VFZ2_3mut
GGSGGGEAAAK	17,284	PERV_Q4VFZ2_3mutA_WS
GGSGGSGGSGGSGGS	17,285	PERV_Q4VFZ2_3mutA_WS
EAAAKPAPGGS	17,286	PERV_Q4VFZ2_3mutA_WS
GGGGGSEAAAK	17,287	PERV_Q4VFZ2_3mutA_WS
GSSPAP	17,288	MLVFF_P26809_3mut
GGGEAAAKPAP	17,289	AVIRE_P03360_3mut
GSSGGSEAAAK	17,290	MLVMS_P03355_PLV919
EAAAKPAPGGS	17,291	WMSV_P03359_3mutA
PAPGGG	17,292	KORV_Q9TTC1_3mutA
EAAAKGSSPAP	17,293	KORV_Q9TTC1-Pro_3mut
GSSPAPEAAAK	17,294	MLVFF_P26809_3mut
GGSGGGEAAAK	17,295	MLVFF_P26809_3mutA
GSSGSSGSS	17,296	WMSV_P03359_3mutA
EAAAKGGS	17,297	BAEVM_P10272_3mut
EAAAKPAPGGS	17,298	KORV_Q9TTC1_3mutA
EAAAKPAPGGS	17,299	BAEVM_P10272_3mutA
GSSGGGGGS	17,300	PERV_Q4VFZ2_3mut
PAPGGGGSS	17,301	PERV_Q4VFZ2_3mut
GSSGSSGSS	17,302	WMSV_P03359_3mut
EAAAKEAAAKEAAAKEAAAK	17,303	WMSV_P03359_3mut
GGS	17,304	AVIRE_P03360_3mut
EAAAKPAPGSS	17,305	MLVFF_P26809_3mut
EAAAKGGG	17,306	KORV_Q9TTC1_3mut
PAPGSSEAAAK	17,307	MLVMS_P03355_3mut
PAPGSSGGS	17,308	MLVMS_P03355_PLV919
GSSPAPEAAAK	17,309	MLVMS_P03355_3mut
GSSGSSGSSGSSGSSGSS	17,310	WMSV_P03359_3mutA
GGGGS	17,311	BAEVM_P10272_3mut
GSSPAP	17,312	MLVMS_P03355_3mut

[1641]

EAAAKGGGGSEAAAK	17,313	KORV_Q9TTC1-Pro_3mutA
EAAAKEAAAK	17,314	WMSV_P03359_3mutA
GGGGSSGGS	17,315	MLVCB_P08361_3mutA
PAPGGSEAAAK	17,316	BAEVM_P10272_3mut
EAAAKGGSPAP	17,317	MLVFF_P26809_3mut
GSSGGSGGG	17,318	MLVBM_Q7SVK7_3mutA_WS
GSSGGS	17,319	PERV_Q4VFZ2_3mut
PAPGGSGSS	17,320	PERV_Q4VFZ2_3mutA_WS
EAAAKGGSGSS	17,321	KORV_Q9TTC1-Pro_3mutA
PAPAP	17,322	MLVCB_P08361_3mut
EAAAKGSSPAP	17,323	PERV_Q4VFZ2_3mutA_WS
EAAAKPAPGGG	17,324	MLVMS_P03355_PLV919
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,325	MLVBM_Q7SVK7_3mut
EAAAKGGGGSS	17,326	MLVMS_P03355_PLV919
PAPEAAAK	17,327	PERV_Q4VFZ2_3mut
EAAAKPAPGSS	17,328	BAEVM_P10272_3mutA
GGSPAP	17,329	PERV_Q4VFZ2_3mutA_WS
GGSGGS	17,330	BAEVM_P10272_3mutA
PAPEAAAKGSS	17,331	KORV_Q9TTC1_3mut
PAPGSS	17,332	MLVMS_P03355_PLV919
PAPAPAPAPAP	17,333	MLVAV_P03356_3mutA
GGG	17,334	XMRV6_A1Z651_3mutA
GGGPAP	17,335	PERV_Q4VFZ2_3mutA_WS
GSSPAPEAAAK	17,336	KORV_Q9TTC1_3mutA
PAP	17,337	BAEVM_P10272_3mutA
GGSPAP	17,338	BAEVM_P10272_3mutA
PAPEAAAKGGS	17,339	MLVMS_P03355_PLV919
PAPGSSGGS	17,340	PERV_Q4VFZ2_3mutA_WS
PAPAPAPAPAPAP	17,341	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAK	17,342	MLVCB_P08361_3mut
GGSGGSGGSGGSGGS	17,343	MLVMS_P03355_PLV919
EAAAKPAPGGS	17,344	MLVMS_P03355_3mut
GGSGGS	17,345	MLVMS_P03355_PLV919
EAAAKPAP	17,346	MLVMS_P03355_3mutA_WS
GGSEAAAK	17,347	XMRV6_A1Z651_3mutA
GGSGGG	17,348	KORV_Q9TTC1_3mut
GGSGGGEAAAK	17,349	PERV_Q4VFZ2_3mut
PAPEAAAKGGG	17,350	AVIRE_P03360
PAPAP	17,351	PERV_Q4VFZ2_3mut
GSS	17,352	KORV_Q9TTC1-Pro_3mutA
EAAAKGSSGGG	17,353	MLVAV_P03356_3mutA

[1642]

[1643]

GGSPAPGSS	17,354	MLVBM_Q7SVK7_3mutA_WS
PAPEAAAK	17,355	MLVAV_P03356_3mut
EAAAKGGSPAP	17,356	BAEVM_P10272_3mutA
PAPAPAPAP	17,357	WMSV_P03359_3mutA
PAPGGSEAAAK	17,358	MLVMS_P03355_3mut
GGSGGSGGSGGS	17,359	WMSV_P03359_3mut
GGGGGSGSS	17,360	XMRV6_A1Z651_3mut
PAPGGSGGG	17,361	KORV_Q9TTC1_3mutA
GGS	17,362	MLVMS_P03355_3mut
EAAAK	17,363	WMSV_P03359_3mut
GGGEAAAKGSS	17,364	MLVBM_Q7SVK7_3mutA_WS
GGSPAPGSS	17,365	MLVCB_P08361_3mut
GGSEAAAKPAP	17,366	PERV_Q4VFZ2_3mut
GGGGSGGGGSGGGGSGGGGSGGGGS	17,367	MLVCB_P08361_3mutA
GGSGSS	17,368	BAEVM_P10272_3mutA
GGGEAAAKGSS	17,369	WMSV_P03359_3mutA
EAAAKGGSPAP	17,370	WMSV_P03359_3mut
GSSPAPEAAAK	17,371	MLVMS_P03355_3mut
GGSGGSGGSGGS	17,372	MLVMS_P03355_PLV919
GSSPAPEAAAK	17,373	WMSV_P03359_3mut
GSSGSSGSSGSS	17,374	PERV_Q4VFZ2
GGSGSSEAAAK	17,375	WMSV_P03359_3mutA
GGSGGG	17,376	MLVFF_P26809_3mut
GGSPAPGGG	17,377	MLVFF_P26809_3mut
GGSGGSGGS	17,378	BAEVM_P10272_3mutA
GGGGSSEAAAK	17,379	MLVBM_Q7SVK7_3mut
GGSPAPGSS	17,380	MLVMS_P03355_3mut
EAAAKPAPGSS	17,381	AVIRE_P03360_3mut
GGGGSSGGS	17,382	FLV_P10273_3mutA
GGSPAPEAAAK	17,383	PERV_Q4VFZ2_3mut
GGSEAAAK	17,384	MLVMS_P03355_3mutA_WS
GSSGSSGSSGSS	17,385	MLVCB_P08361_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,386	MLVMS_P03355_PLV919
GGGGG	17,387	PERV_Q4VFZ2_3mut
GGSEAAAKGSS	17,388	MLVCB_P08361_3mutA
GSSGGG	17,389	MLVBM_Q7SVK7_3mutA_WS
PAPGSSGGG	17,390	KORV_Q9TTC1-Pro_3mutA
GGSGGS	17,391	BAEVM_P10272_3mut

EAAAKGGGGGS	17,392	MLVBM_Q7SVK7_3mutA_WS
GGSGSSPAP	17,393	MLVCB_P08361_3mut
PAPGSSGGG	17,394	KORV_Q9TTC1
PAPGGSGGG	17,395	MLVMS_P03355_3mut
GGGG	17,396	WMSV_P03359_3mutA
EAAAKGGSPAP	17,397	MLVCB_P08361_3mut
GSSGSS	17,398	FLV_P10273_3mutA
GGSEAAAKPAP	17,399	SFV3L_P27401_2mut
EAAAKGSSGGS	17,400	MLVAV_P03356_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,401	MLVAV_P03356_3mutA
EAAAKGGSGSS	17,402	PERV_Q4VFZ2_3mutA_WS
GGGGG	17,403	MLVCB_P08361_3mut
GGGEAAAK	17,404	BAEVM_P10272_3mut
GGSGGSGGSGGS	17,405	MLVCB_P08361_3mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,406	PERV_Q4VFZ2
PAPAPAPAPAP	17,407	MLVMS_P03355_3mutA_WS
EAAAKEAAAK	17,408	XMRV6_A1Z651_3mut
GSSGGSEAAAK	17,409	PERV_Q4VFZ2_3mutA_WS
PAPGGSEAAAK	17,410	KORV_Q9TTC1-Pro_3mutA
EAAAKGGGPAP	17,411	MLVBM_Q7SVK7_3mutA_WS
PAPGGSGSS	17,412	PERV_Q4VFZ2
SGSETPGTSESATPES	17,413	MLVMS_P03355_3mut
GGSGGS	17,414	MLVMS_P03355_PLV919
EAAAKGGS	17,415	FLV_P10273_3mut
GGSPAPGSS	17,416	MLVMS_P03355_3mutA_WS
EAAAKEAAAKEAAAKEAAAK	17,417	FFV_O93209_2mut
GSSGGSGGG	17,418	MLVMS_P03355_3mutA_WS
PAPGSSEAAAK	17,419	WMSV_P03359_3mut
PAPAPAPAPAPAP	17,420	KORV_Q9TTC1_3mutA
GGGGSS	17,421	BAEVM_P10272_3mut
GGGGSEAAAKGGGGS	17,422	AVIRE_P03360_3mut
GSSPAPEAAAK	17,423	KORV_Q9TTC1-Pro_3mutA
PAPEAAAKGGG	17,424	MLVBM_Q7SVK7_3mut
EAAAKEAAAK	17,425	WMSV_P03359_3mut
EAAAK	17,426	SFV3L_P27401-Pro_2mutA
GSSGGSGGG	17,427	XMRV6_A1Z651_3mutA
GGGEAAAKPAP	17,428	WMSV_P03359_3mutA
GGSGGS	17,429	MLVFF_P26809_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,430	FOAMV_P14350_2mutA

[1644]

GGGGG	17,431	MLVAV_P03356_3mutA
GSSGGSEAAAK	17,432	BAEVM_P10272_3mut
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,433	SFV1_P23074
GGSGGGPAP	17,434	MLVCB_P08361_3mut
GGSGSS	17,435	PERV_Q4VFZ2_3mut
SGSETPGTSESATPES	17,436	MLVFF_P26809_3mut
EAAAKGGSPAP	17,437	MLVMS_P03355_3mut
PAPAP	17,438	PERV_Q4VFZ2_3mut
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,439	MLVBM_Q7SVK7_3mut
GGGGGS	17,440	BAEVM_P10272_3mutA
EAAAKEAAAK	17,441	AVIRE_P03360_3mut
GSSGGSEAAAK	17,442	PERV_Q4VFZ2_3mut
GGGEAAAK	17,443	WMSV_P03359_3mut
GSSGGGEAAAK	17,444	AVIRE_P03360_3mutA
GGG	17,445	XMRV6_A1Z651_3mut
GGGGSEAAAKGGGGS	17,446	BAEVM_P10272_3mut
GGGG	17,447	MLVMS_P03355_3mut
GGSGGS	17,448	MLVMS_P03355_3mutA_WS
GGSGGGGSS	17,449	MLVBM_Q7SVK7_3mutA_WS
GSSPAPGGS	17,450	PERV_Q4VFZ2_3mut
GSSPAPEAAAK	17,451	PERV_Q4VFZ2_3mutA_WS
EAAAKGGS	17,452	WMSV_P03359_3mut
GGSGGSGGSGGS	17,453	PERV_Q4VFZ2_3mut
GGGGSSEAAAK	17,454	KORV_Q9TTC1-Pro_3mut
PAPAPAPAPAPAP	17,455	MLVAV_P03356_3mut
EAAAKGSSGGG	17,456	MLVMS_P03355_PLV919
GGGGG	17,457	MLVBM_Q7SVK7_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,458	FFV_O93209_2mutA
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,459	KORV_Q9TTC1-Pro_3mut
GGSPAPGGG	17,460	MLVMS_P03355_3mutA_WS
GGGEAAAKGGS	17,461	MLVMS_P03355_3mut
GGGEAAAK	17,462	PERV_Q4VFZ2_3mut
PAPEAAAKGGG	17,463	MLVMS_P03355_3mut
GSSGSSGSSGSSGSSGSS	17,464	BAEVM_P10272_3mutA
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,465	GALV_P21414_3mutA
EAAAKGGSPAP	17,466	FFV_O93209-Pro
EAAAKEAAAK	17,467	MLVFF_P26809_3mut

[1645]

GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,468	PERV_Q4VFZ2_3mutA_WS
GGSGGSGGSGGS	17,469	MLVAV_P03356_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,470	SFV3L_P27401_2mutA
GSSGSSGSSGSSGSSGSS	17,471	BAEVM_P10272_3mut
GGGGS	17,472	MLVMS_P03355_PLV919
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAK A	17,473	SFV1_P23074
GGGGSGGGGS	17,474	KORV_Q9TTC1-Pro_3mutA
GGGGSGGGGS	17,475	MLVMS_P03355_3mut
GGSGSS	17,476	KORV_Q9TTC1_3mutA
GSSPAPGGG	17,477	PERV_Q4VFZ2_3mut
GSSGGSPAP	17,478	PERV_Q4VFZ2_3mutA_WS
PAPGGS	17,479	PERV_Q4VFZ2_3mutA_WS
GGSPAPEAAAK	17,480	FOAMV_P14350_2mutA
GGGPAPGGS	17,481	SFV3L_P27401_2mut
PAPGSSGGG	17,482	MLVCB_P08361_3mut
GSSGGGEAAAK	17,483	AVIRE_P03360_3mut
GSSGGG	17,484	XMRV6_A1Z651_3mut
GSSGSS	17,485	PERV_Q4VFZ2_3mut
GSSGGG	17,486	MLVAV_P03356_3mutA
PAPGGGGGS	17,487	PERV_Q4VFZ2_3mut
GSSEAAAK	17,488	MLVMS_P03355_3mut
PAPGGG	17,489	FLV_P10273_3mutA
GGGGSGGGGS	17,490	PERV_Q4VFZ2_3mut
GSSGGS	17,491	MLVMS_P03355_PLV919
GGGGSGGGGS	17,492	SFV3L_P27401_2mut
EAAAKGGSGSS	17,493	FLV_P10273_3mutA
GSSEAAAKGGS	17,494	MLVMS_P03355_3mutA_WS
PAPGSSEAAAK	17,495	SFV3L_P27401_2mutA
GGGGSGGGGS	17,496	SFV3L_P27401-Pro_2mutA
PAPGSSEAAAK	17,497	PERV_Q4VFZ2_3mut
PAPGSSEAAAK	17,498	PERV_Q4VFZ2
GGSPAPGGG	17,499	AVIRE_P03360_3mut
GGGGGS	17,500	PERV_Q4VFZ2_3mutA_WS
GGGGSSGGS	17,501	PERV_Q4VFZ2_3mut
PAPAPAPAP	17,502	AVIRE_P03360_3mutA
GGSGGS	17,503	WMSV_P03359_3mutA
GGGPAPGGS	17,504	PERV_Q4VFZ2_3mut
GGSGGSGGSGGSGGS	17,505	MLVMS_P03355_PLV919
GGSGGG	17,506	PERV_Q4VFZ2_3mut
EAAAKEAAAK	17,507	SFV3L_P27401_2mut
PAPGSS	17,508	XMRV6_A1Z651_3mut

[1646]

GSSEAAAK	17,509	MLVFF_P26809_3mut
GGSPAPGGG	17,510	MLVMS_P03355_3mut
EAAAKGGG	17,511	WMSV_P03359_3mutA
GSSEAAAKGGS	17,512	PERV_Q4VFZ2_3mutA_WS
GSSGGSPAP	17,513	FFV_O93209
GGGGGS	17,514	KORV_Q9TTC1-Pro_3mut
GSSGGG	17,515	MLVCB_P08361_3mut
GSSGSS	17,516	MLVCB_P08361_3mutA
GGSEAAAKPAP	17,517	BAEVM_P10272_3mut
EAAAKGGGGSS	17,518	MLVCB_P08361_3mut
EAAAKPAPGGS	17,519	KORV_Q9TTC1-Pro_3mutA
GSSGSSGSSGSSGSS	17,520	MLVAV_P03356_3mutA
GGGGSEAAAKGGGGS	17,521	PERV_Q4VFZ2_3mutA_WS
GGSGSS	17,522	KORV_Q9TTC1-Pro_3mut
GSS	17,523	SFV3L_P27401-Pro_2mutA
PAPAP	17,524	BAEVM_P10272_3mut
EAAAKPAP	17,525	BAEVM_P10272
EAAAKEAAAKEAAAKEAAAKEAAAK	17,526	KORV_Q9TTC1-Pro_3mut
GGGGGGG	17,527	PERV_Q4VFZ2_3mutA_WS
GGGGS	17,528	MLVMS_P03355_3mut
GSSGGG	17,529	FLV_P10273_3mutA
PAPAPAPAPAP	17,530	FLV_P10273_3mut
EAAAKEAAAKEAAAK	17,531	WMSV_P03359_3mutA
GSSGGS	17,532	MLVBM_Q7SVK7_3mutA_WS
EAAAKPAPGGG	17,533	MLVMS_P03355_3mut
GSSPAPGGS	17,534	WMSV_P03359_3mut
PAPGSSGGG	17,535	PERV_Q4VFZ2_3mutA_WS
GSSGGG	17,536	AVIRE_P03360_3mutA
PAPGGSGSS	17,537	MLVFF_P26809_3mut
PAPGSS	17,538	PERV_Q4VFZ2_3mut
GGGGGSGSS	17,539	WMSV_P03359_3mutA
EAAAKGGGGSS	17,540	MLVBM_Q7SVK7_3mutA_WS
GGGGGGG	17,541	BAEVM_P10272_3mut
PAPEAAAKGSS	17,542	MLVMS_P03355_3mut
GGSGGGEAAAK	17,543	MLVMS_P03355_PLV919
EAAAKGGGGGS	17,544	MLVCB_P08361_3mut
PAPGGS	17,545	KORV_Q9TTC1-Pro_3mut
GGGG	17,546	FLV_P10273_3mutA
EAAAKGGSGSS	17,547	MLVBM_Q7SVK7_3mutA_WS

[1647]

GGGGSSGGS	17,548	MLVMS_P03355_3mutA_WS
GGGGGGGG	17,549	WMSV_P03359_3mut
GGSGSSGGG	17,550	MLVMS_P03355_PLV919
GSSEAAAKGGS	17,551	KORV_Q9TTC1-Pro_3mutA
EAAAKPAPGSS	17,552	MLVCB_P08361_3mut
GGSPAPGSS	17,553	KORV_Q9TTC1_3mutA
PAPGSSGGG	17,554	BAEVM_P10272_3mut
EAAAKPAPGSS	17,555	WMSV_P03359_3mut
GGSPAPEAAAK	17,556	XMRV6_A1Z651_3mutA
GSSPAP	17,557	FLV_P10273_3mutA
GSS	17,558	BAEVM_P10272_3mutA
EAAAKPAPGGS	17,559	FLV_P10273_3mutA
GGSGSSPAP	17,560	FLV_P10273_3mutA
PAPGSSGGS	17,561	MLVMS_P03355_3mut
GSAGSAAGSGEF	17,562	PERV_Q4VFZ2_3mutA_WS
GSSGGSEAAAK	17,563	KORV_Q9TTC1_3mutA
GSSGGS	17,564	MLVMS_P03355_3mutA_WS
EAAAKGGGGSEAAAK	17,565	SFV3L_P27401_2mut
GSSGGS	17,566	PERV_Q4VFZ2_3mutA_WS
GGSPAPEAAAK	17,567	FLV_P10273_3mut
GGSEAAAKGSS	17,568	PERV_Q4VFZ2_3mutA_WS
GSSPAPEAAAK	17,569	PERV_Q4VFZ2_3mutA_WS
GGSGSSGGG	17,570	PERV_Q4VFZ2_3mut
GGGG	17,571	AVIRE_P03360_3mutA
GGSEAAAKPAP	17,572	WMSV_P03359_3mut
GSSGGSPAP	17,573	MLVAV_P03356_3mutA
GSSGGSEAAAK	17,574	MLVMS_P03355_3mut
PAPEAAAKGGS	17,575	KORV_Q9TTC1-Pro_3mut
GGSPAP	17,576	PERV_Q4VFZ2_3mutA_WS
GGSEAAAK	17,577	MLVAV_P03356_3mutA
EAAAKGGGGSEAAAK	17,578	KORV_Q9TTC1-Pro_3mut
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,579	MLVMS_P03355_PLV919
GSSEAAAK	17,580	KORV_Q9TTC1_3mutA
GGG	17,581	AVIRE_P03360
GGSEAAAKGSS	17,582	MLVBM_Q7SVK7_3mut
GGSEAAAKGSS	17,583	MLVMS_P03355_3mut
GGSPAPEAAAK	17,584	MLVCB_P08361_3mut
GGSGGGEAAAK	17,585	MLVCB_P08361_3mut
GGSEAAAKPAP	17,586	MLVMS_P03355_3mutA_WS
EAAAKGGSGSS	17,587	KORV_Q9TTC1-Pro_3mut
GGGEAAAKGGS	17,588	MLVCB_P08361_3mut

[1648]

EAAAKGGGGSEAAAK	17,589	FLV_P10273_3mutA
GGSPAP	17,590	MLVFF_P26809_3mut
GGGGSSGGS	17,591	XMRV6_A1Z651_3mutA
PAP	17,592	MLVCB_P08361_3mut
GGS	17,593	SFV3L_P27401-Pro_2mutA
GGGGSGGGGS	17,594	MLVMS_P03355_3mut
GGGEAAAKGGS	17,595	MLVAV_P03356_3mutA
GSSGSSGSSGSSGSSGSS	17,596	MLVMS_P03355_PLV919
PAPGSS	17,597	MLVCB_P08361_3mut
GGSGGSGGS	17,598	MLVMS_P03355_PLV919
PAPGGSGGG	17,599	FLV_P10273_3mutA
GGGGSGGGGSGGGGS	17,600	FLV_P10273_3mut
GGSGSSGGG	17,601	KORV_Q9TTC1-Pro_3mutA
GGSGGSGGS	17,602	GALV_P21414_3mutA
GGGEAAAKGGS	17,603	WMSV_P03359_3mut
SGSETPGTSESATPES	17,604	KORV_Q9TTC1_3mutA
EAAAKGGGGGS	17,605	KORV_Q9TTC1-Pro_3mut
EAAAKGSSPAP	17,606	BAEVM_P10272_3mut
GGGG	17,607	MLVCB_P08361_3mut
GGGGSGGGGSGGGGSGGGGSGGGGS	17,608	MLVBM_Q7SVK7_3mut
GSSGGSGGG	17,609	MLVMS_P03355_PLV919
GGSGSS	17,610	MLVFF_P26809_3mut
EAAAKGGS	17,611	AVIRE_P03360_3mutA
GSSEAAAKGGS	17,612	MLVBM_Q7SVK7_3mutA_WS
EAAAKPAPGGG	17,613	WMSV_P03359_3mut
PAPGSSGGG	17,614	MLVCB_P08361_3mutA
GGGGSSEAAAK	17,615	KORV_Q9TTC1-Pro_3mutA
GSSEAAAKPAP	17,616	BAEVM_P10272_3mutA
PAPGGGEAAAK	17,617	MLVBM_Q7SVK7_3mutA_WS
GGSGGGEAAAK	17,618	MLVCB_P08361_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,619	FFV_O93209
EAAAKGGGGGS	17,620	GALV_P21414_3mutA
GGSPAPGGG	17,621	MLVMS_P03355_3mut
GSSGSSGSS	17,622	FLV_P10273_3mutA
EAAAK	17,623	MLVBM_Q7SVK7_3mut
GGGGSSGGS	17,624	MLVMS_P03355_3mut
GGSGSSPAP	17,625	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAKEAAAK	17,626	BAEVM_P10272_3mut
GGGPAPGSS	17,627	MLVMS_P03355_3mut
GSSPAPGGS	17,628	PERV_Q4VFZ2_3mutA_WS

[1649]

[1650]

PAPAP	17,629	FLV_P10273_3mutA
PAPAPAPAP	17,630	PERV_Q4VFZ2_3mut
GGGGGSEAAAK	17,631	GALV_P21414_3mutA
GGGGGSGSS	17,632	BAEVM_P10272_3mutA
GGGEAAAKGSS	17,633	KORV_Q9TTC1_3mutA
GGGGGSPAP	17,634	AVIRE_P03360_3mut
GGGGGSEAAAK	17,635	SFV3L_P27401_2mutA
GGS	17,636	KORV_Q9TTC1_3mutA
GGGGGGG	17,637	PERV_Q4VFZ2_3mut
SGSETPGTSESATPES	17,638	SFV3L_P27401_2mutA
EAAAKGGSGGG	17,639	MLVMS_P03355_3mut
GGGGS	17,640	MLVFF_P26809_3mut
EAAAKGSSGGG	17,641	BAEVM_P10272_3mut
EAAAKPAPGGS	17,642	MLVF5_P26810_3mutA
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,643	SFV3L_P27401_2mutA
GGSPAPGGG	17,644	WMSV_P03359_3mutA
GSAGSAAGSGEF	17,645	MLVFF_P26809_3mut
GGGGSSGGS	17,646	MLVMS_P03355_3mutA_WS
GGGGGGG	17,647	MLVCB_P08361_3mut
GSSEAAAK	17,648	WMSV_P03359_3mut
PAPGSS	17,649	FLV_P10273_3mutA
GSSGGG	17,650	PERV_Q4VFZ2_3mutA_WS
PAPGGG	17,651	MLVFF_P26809_3mut
GGGGGSPAP	17,652	MLVMS_P03355_3mut
GGSEAAAK	17,653	XMRV6_A1Z651_3mut
GSSGGG	17,654	PERV_Q4VFZ2_3mut
GGSGGSGGSGGS	17,655	MLVMS_P03355_3mut
PAPAP	17,656	AVIRE_P03360_3mut
GGSEAAAK	17,657	PERV_Q4VFZ2_3mut
GGGGS	17,658	MLVMS_P03355_PLV919
GGGG	17,659	BAEVM_P10272_3mutA
EAAAKGGGGSS	17,660	MLVCB_P08361_3mutA
EAAAKEAAAKEAAAK	17,661	GALV_P21414_3mutA
PAPGGGEAAAK	17,662	KORV_Q9TTC1
EAAAKGGSPAP	17,663	MLVMS_P03355_3mut
GGSGSSEAAAK	17,664	MLVMS_P03355_3mut
GGSPAPEAAAK	17,665	FLV_P10273_3mutA
GGGGGGG	17,666	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,667	SFV1_P23074_2mutA
EAAAKGSSGGS	17,668	MLVMS_P03355_3mut
GSSEAAAKPAP	17,669	MLVFF_P26809_3mut

[1651]

GGGGSS	17,670	FLV_P10273_3mutA
EAAAKGGSGGG	17,671	AVIRE_P03360_3mutA
GGSGGS	17,672	PERV_Q4VFZ2_3mutA_WS
GGGGGSPAP	17,673	AVIRE_P03360_3mutA
EAAAKEAAAKEAAAK	17,674	XMRV6_A1Z651_3mut
PAPEAAAKGGS	17,675	FLV_P10273_3mutA
GSSGGSEAAAK	17,676	MLVCB_P08361_3mut
EAAAKGGSGGG	17,677	MLVMS_P03355
GGSGGGPAP	17,678	MLVMS_P03355_3mut
GGS	17,679	XMRV6_A1Z651_3mut
GGSEAAAKPAP	17,680	MLVFF_P26809_3mut
EAAAKGGG	17,681	MLVMS_P03355_PLV919
GSSGSSGSSGSS	17,682	WMSV_P03359_3mut
GGSGSSPAP	17,683	PERV_Q4VFZ2_3mut
GGGEAAAK	17,684	MLVMS_P03355_3mutA_WS
GSSPAPGGS	17,685	KORV_Q9TTC1-Pro_3mutA
GSSEAAAKGGG	17,686	SFV3L_P27401_2mut
EAAAKPAPGGS	17,687	MLVCB_P08361_3mut
GGSGGGEAAAK	17,688	PERV_Q4VFZ2
GGSGSS	17,689	MLVCB_P08361_3mut
GGSGGGEAAAK	17,690	MLVBM_Q7SVK7_3mutA_WS
GGSGGSGGSGGSGGSGGS	17,691	FLV_P10273_3mut
PAPEAAAKGSS	17,692	MLVMS_P03355_3mut
EAAAKGSSGGS	17,693	WMSV_P03359_3mutA
GGSGSSEAAAK	17,694	MLVCB_P08361_3mut
GGSGSSEAAAK	17,695	KORV_Q9TTC1_3mutA
GSSGGSGGG	17,696	MLVMS_P03355_PLV919
EAAAKGGSGGG	17,697	SFV3L_P27401-Pro_2mutA
GGSGGS	17,698	AVIRE_P03360_3mutA
GSAGSAAGSGEF	17,699	MLVMS_P03355_PLV919
GGSGSS	17,700	GALV_P21414_3mutA
GGGG	17,701	MLVFF_P26809_3mutA
GGGGSGGGGSGGGGSGGGGS	17,702	WMSV_P03359_3mut
SGSETPGTSESATPES	17,703	BAEVM_P10272_3mut
EAAAKEAAAKEAAAKEAAAK	17,704	FOAMV_P14350_2mutA
GGGEAAAKGGS	17,705	FLV_P10273_3mutA
GSSGGSEAAAK	17,706	MLVFF_P26809_3mut
EAAAKGGGGSS	17,707	MLVAV_P03356_3mut
PAPGGSEAAAK	17,708	KORV_Q9TTC1-Pro_3mut
EAAAK	17,709	XMRV6_A1Z651_3mut
GSSGSSGSSGSSGSSGSS	17,710	PERV_Q4VFZ2_3mut

[1652]

GGGG	17,711	MLVCB_P08361_3mutA
GSSGSS	17,712	WMSV_P03359_3mutA
GSSGGSPAP	17,713	AVIRE_P03360_3mut
GGSGGSGGS	17,714	MLVCB_P08361_3mut
EAAAKGGGPAP	17,715	FLV_P10273_3mutA
GGGGSGGGGS	17,716	MLVCB_P08361_3mut
GGSEAAAKGSS	17,717	PERV_Q4VFZ2_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,718	SFV3L_P27401_2mutA
GGSGSSEAAAK	17,719	PERV_Q4VFZ2_3mutA_WS
EAAAKEAAAKEAAAKEAAAK	17,720	SFV3L_P27401-Pro_2mutA
GSSEAAAKGGS	17,721	FLV_P10273_3mutA
GGSGSS	17,722	PERV_Q4VFZ2
GGSGSSEAAAK	17,723	SFV3L_P27401-Pro_2mutA
GSSGSSGSS	17,724	XMRV6_A1Z651_3mutA
EAAAKGSSPAP	17,725	KORV_Q9TTC1_3mutA
EAAAKPAP	17,726	FLV_P10273_3mutA
GGSGSSEAAAK	17,727	KORV_Q9TTC1-Pro_3mut
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,728	KORV_Q9TTC1_3mutA
GGGGSGGGGSGGGGS	17,729	KORV_Q9TTC1-Pro_3mutA
GGGGGGG	17,730	FLV_P10273_3mut
EAAAKGSS	17,731	WMSV_P03359_3mut
EAAAKGGGPAP	17,732	MLVCB_P08361_3mut
GSSGSS	17,733	MLVBM_Q7SVK7_3mutA_WS
EAAAKGGGGGS	17,734	MLVFF_P26809_3mut
GGSGGGEAAAK	17,735	FLV_P10273_3mutA
PAPGSS	17,736	MLVFF_P26809_3mutA
PAPGSS	17,737	BAEVM_P10272_3mutA
GGSPAPGSS	17,738	AVIRE_P03360_3mut
GGGGSSEAAAK	17,739	MLVMS_P03355_3mut
GSSGGGGGS	17,740	FFV_O93209-Pro
EAAAKGSSPAP	17,741	PERV_Q4VFZ2_3mut
GSSPAPGGS	17,742	PERV_Q4VFZ2_3mut
GGGGGG	17,743	BAEVM_P10272_3mut
EAAAKGGGGSS	17,744	PERV_Q4VFZ2_3mutA_WS
PAPGGSEAAAK	17,745	KORV_Q9TTC1_3mutA
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,746	MLVMS_P03355_3mutA_WS
GSSGSSGSSGSS	17,747	MLVMS_P03355_3mut
EAAAKGSSGGG	17,748	MLVMS_P03355_PLV919
GGSEAAAKPAP	17,749	AVIRE_P03360_3mutA
GSSGSSGSSGSSGSS	17,750	WMSV_P03359_3mutA
GGGEAAAKPAP	17,751	FLV_P10273_3mutA

PAPGSSGGG	17,752	KORV_Q9TTC1_3mutA
GSSGSS	17,753	MLVMS_P03355_3mutA_WS
PAPEAAAK	17,754	BAEVM_P10272_3mut
GGGPAPGSS	17,755	PERV_Q4VFZ2
GSSGGSPAP	17,756	MLVFF_P26809_3mut
GGGGSS	17,757	SFV3L_P27401_2mut
PAPEAAAKGSS	17,758	SFV3L_P27401_2mut
GGSGGGPAP	17,759	XMRV6_A1Z651_3mutA
PAPGGS	17,760	BAEVM_P10272_3mutA
EAAAKGGGGGS	17,761	AVIRE_P03360_3mut
GSSGGSPAP	17,762	KORV_Q9TTC1-Pro_3mutA
GSSGGGGGS	17,763	WMSV_P03359_3mut
GGGEAAAKGGS	17,764	AVIRE_P03360_3mut
GGGEAAAKGSS	17,765	BAEVM_P10272_3mut
PAPEAAAKGSS	17,766	MLVAV_P03356_3mutA
GSSGSSGSSGSSGSS	17,767	MLVCB_P08361_3mut
GGSPAPGSS	17,768	FLV_P10273_3mutA
EAAAKGSSPAP	17,769	BAEVM_P10272_3mutA
GGSGGSGGSGGSGGSGGS	17,770	PERV_Q4VFZ2
GGGGSSEAAAK	17,771	FLV_P10273_3mutA
GGGGSSPAP	17,772	FFV_O93209
GSSGGSPAP	17,773	MLVMS_P03355_3mut
GGGPAPGSS	17,774	MLVMS_P03355_PLV919
PAPGSSGGS	17,775	PERV_Q4VFZ2_3mut
GGGGGSPAP	17,776	MLVFF_P26809_3mut
SGSETPGTSESATPES	17,777	MLVMS_P03355_3mutA_WS
GSSGSSGSSGSSGSS	17,778	KORV_Q9TTC1_3mutA
GSSPAPGGG	17,779	WMSV_P03359_3mut
PAPAPAPAPAPAP	17,780	SFV3L_P27401_2mutA
GGGPAPGGS	17,781	MLVMS_P03355_3mut
PAPGGSEAAAK	17,782	WMSV_P03359_3mut
GGGGSSEAAAK	17,783	FFV_O93209-Pro
GGSPAPGGG	17,784	FLV_P10273_3mutA
GSSPAPEAAAK	17,785	AVIRE_P03360_3mut
GGGEAAAK	17,786	FLV_P10273_3mutA
PAPEAAAKGGG	17,787	MLVCB_P08361_3mut
GGSPAPGGG	17,788	MLVCB_P08361_3mut
GGSGGGGSS	17,789	BAEVM_P10272_3mutA
GSSPAPEAAAK	17,790	MLVCB_P08361_3mut
GGSPAPGGG	17,791	KORV_Q9TTC1-Pro_3mutA
PAPGGSGSS	17,792	KORV_Q9TTC1_3mutA

[1653]

[1654]

GSSPAP	17,793	KORV_Q9TTC1-Pro_3mutA
SGSETPGTSESATPES	17,794	MLVMS_P03355
GSSGSSGSS	17,795	MLVAV_P03356_3mutA
PAPGSSGGS	17,796	PERV_Q4VFZ2_3mutA_WS
PAPGGS	17,797	KORV_Q9TTC1-Pro_3mutA
PAPEAAAKGGG	17,798	SFV3L_P27401-Pro_2mutA
GGSGGSGGS	17,799	BAEVM_P10272_3mut
PAPGGS	17,800	MLVFF_P26809_3mut
GSSGGSPAP	17,801	MLVMS_P03355_PLV919
GSSGGGGGS	17,802	FLV_P10273_3mutA
GGGGGSPAP	17,803	KORV_Q9TTC1-Pro_3mut
EAAAKPAPGSS	17,804	SFV3L_P27401-Pro_2mutA
EAAAKGGSPAP	17,805	KORV_Q9TTC1-Pro
GGGPAPEAAAK	17,806	MLVMS_P03355_PLV919
GGSEAAAKGSS	17,807	MLVMS_P03355
PAPEAAAKGSS	17,808	KORV_Q9TTC1_3mutA
PAPEAAAKGGS	17,809	WMSV_P03359_3mutA
GSSGGG	17,810	PERV_Q4VFZ2_3mutA_WS
EAAAKGGGGSS	17,811	MLVMS_P03355_PLV919
EAAAKGGSPAP	17,812	AVIRE_P03360_3mutA
GGGGSSGGS	17,813	MLVMS_P03355_PLV919
PAPEAAAKGSS	17,814	PERV_Q4VFZ2_3mutA_WS
EAAAKGGGGGS	17,815	BAEVM_P10272_3mut
GSSGGGGGS	17,816	MLVMS_P03355_3mut
PAPAPAPAP	17,817	KORV_Q9TTC1_3mutA
GGSGGSGGSGGS	17,818	MLVAV_P03356_3mut
PAPAPAPAP	17,819	SFV3L_P27401_2mut
GSSEAAAKPAP	17,820	MLVMS_P03355_3mut
GGSGGGEAAAK	17,821	SFV3L_P27401_2mutA
GSSGGSGGG	17,822	MLVMS_P03355_3mutA_WS
GGGGGSPAP	17,823	MLVCB_P08361_3mutA
GGGEAAAKGSS	17,824	XMRV6_A1Z651_3mutA
GGGGSSPAP	17,825	BAEVM_P10272_3mut
GGSGGG	17,826	PERV_Q4VFZ2_3mut
GGGGSS	17,827	MLVBM_Q7SVK7_3mutA_WS
EAAAKGSSGGS	17,828	PERV_Q4VFZ2_3mutA_WS
GSSGGGGGS	17,829	PERV_Q4VFZ2
EAAAKGSSGGS	17,830	PERV_Q4VFZ2_3mut
EAAAKEAAAK	17,831	MLVAV_P03356_3mut
GSSGGGEAAAK	17,832	MLVAV_P03356_3mut
GSSPAPGGG	17,833	XMRV6_A1Z651_3mut

GGGGSGGGGSGGGGS	17,834	PERV_Q4VFZ2_3mut
EAAAKEAAAKEAAAKEAAAK	17,835	KORV_Q9TTC1_3mutA
EAAAKGGSGSS	17,836	MLVBM_Q7SVK7_3mut
PAPEAAAK	17,837	BLVJ_P03361
GSSGGG	17,838	FFV_O93209-Pro
GGSGGGEAAAK	17,839	KORV_Q9TTC1-Pro_3mutA
EAAAK	17,840	FLV_P10273_3mutA
GGGGSSPAP	17,841	MLVMS_P03355_3mut
GSS	17,842	SFV3L_P27401-Pro_2mut
PAPEAAAKGSS	17,843	BAEVM_P10272_3mut
GGGGGSPAP	17,844	PERV_Q4VFZ2_3mut
GSSGSSGSS	17,845	BAEVM_P10272_3mutA
GGGGSGGGGSGGGGSGGGGS	17,846	SFV1_P23074_2mut
GGGGSSEAAAK	17,847	SFV3L_P27401_2mutA
GGGGSGGGGSGGGGSGGGGS	17,848	FOAMV_P14350-Pro_2mut
PAPGSSEAAAK	17,849	MLVBM_Q7SVK7_3mutA_WS
GGGGGSGSS	17,850	MLVFF_P26809_3mutA
GGSEAAAKGGG	17,851	MLVBM_Q7SVK7_3mut
PAPGSSGGG	17,852	PERV_Q4VFZ2
GGS	17,853	PERV_Q4VFZ2_3mutA_WS
EAAAKGGSGSS	17,854	FLV_P10273_3mut
GGGEAAAK	17,855	WMSV_P03359_3mutA
GGSEAAAKPAP	17,856	MLVBM_Q7SVK7_3mut
SGSETPGTSESATPES	17,857	FOAMV_P14350-Pro_2mutA
EAAAKPAPGGS	17,858	AVIRE_P03360_3mut
EAAAKGGGGGS	17,859	KORV_Q9TTC1-Pro_3mutA
GGGGS	17,860	PERV_Q4VFZ2_3mut
GGSEAAAKGSS	17,861	MLVFF_P26809_3mutA
GGSEAAAKGGG	17,862	AVIRE_P03360
GGSGGSGGSGGSGGSGGS	17,863	SFV3L_P27401_2mut
GGSEAAAKGSS	17,864	SFV3L_P27401-Pro_2mutA
GGGEAAAKPAP	17,865	MLVCB_P08361_3mut
GGSEAAAK	17,866	MLVMS_P03355_PLV919
GGSPAPGSS	17,867	KORV_Q9TTC1-Pro_3mutA
GSSPAPEAAAK	17,868	WMSV_P03359_3mutA
GGSGSS	17,869	KORV_Q9TTC1-Pro_3mutA
PAPGGGGGS	17,870	AVIRE_P03360_3mut
PAPEAAAKGSS	17,871	FFV_O93209-Pro
GGSGGGEAAAK	17,872	WMSV_P03359_3mut
PAPGGG	17,873	MLVMS_P03355_3mut
EAAAKGGG	17,874	FLV_P10273_3mutA

[1655]

GSSGSSGSSGSS	17,875	MLVCB_P08361_3mut
EAAAKGGSGGG	17,876	FFV_O93209
GSSPAPGGS	17,877	PERV_Q4VFZ2_3mutA_WS
GSSPAPGGS	17,878	MLVCB_P08361_3mut
GGGPAP	17,879	WMSV_P03359_3mutA
GGGPAP	17,880	KORV_Q9TTC1_3mutA
GGSPAPGSS	17,881	KORV_Q9TTC1-Pro_3mut
PAPAP	17,882	MLVMS_P03355_3mut
GGGGGGG	17,883	MLVMS_P03355_3mut
GGGGG	17,884	KORV_Q9TTC1-Pro_3mut
GSAGSAAGSGEF	17,885	FOAMV_P14350_2mutA
PAPAP	17,886	KORV_Q9TTC1-Pro_3mutA
GGSEAAAKGGG	17,887	SFV3L_P27401-Pro_2mutA
PAPAP	17,888	WMSV_P03359_3mut
GGGGSGGGGSGGGGS	17,889	SFV3L_P27401_2mut
PAPGGS	17,890	KORV_Q9TTC1_3mutA
GGGEAAAKPAP	17,891	FLV_P10273_3mut
GGGGGS	17,892	MLVAV_P03356_3mutA
GSSEAAAKGGG	17,893	WMSV_P03359_3mut
EAAAKGGGGSS	17,894	GALV_P21414_3mutA
GSSGGS	17,895	MLVAV_P03356_3mutA
GSSGGG	17,896	MLVBM_Q7SVK7_3mut
PAPAPAP	17,897	SFV3L_P27401-Pro_2mutA
GGGG	17,898	KORV_Q9TTC1_3mutA
EAAAKPAPGGS	17,899	MLVFF_P26809_3mut
GGGGSGGGGS	17,900	XMRV6_A1Z651_3mut
EAAAKGGG	17,901	MLVCB_P08361_3mut
GGGGSSPAP	17,902	KORV_Q9TTC1_3mutA
GSSEAAAKGGG	17,903	KORV_Q9TTC1-Pro_3mutA
GGGGG	17,904	BLVJ_P03361_2mutB
GGGEAAAKGSS	17,905	FFV_O93209-Pro
GSSGSSGSS	17,906	BAEVM_P10272_3mut
GSSGGSPAP	17,907	PERV_Q4VFZ2_3mut
EAAAKGGS	17,908	KORV_Q9TTC1_3mut
GGSPAPEAAAK	17,909	AVIRE_P03360_3mut
GGSEAAAK	17,910	WMSV_P03359_3mut
GSSGGS	17,911	KORV_Q9TTC1-Pro_3mutA
GGGPAPEAAAK	17,912	KORV_Q9TTC1_3mutA
PAPGSS	17,913	WMSV_P03359_3mutA
GGSEAAAKGSS	17,914	FLV_P10273_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,915	SFV3L_P27401

[1656]

[1657]

GSSEAAAKGGG	17,916	SFV3L_P27401-Pro_2mutA
GGGGSEAAAKGGGGS	17,917	KORV_Q9TTC1-Pro_3mutA
GGSGGSGGS	17,918	WMSV_P03359_3mut
GGGGGSGSS	17,919	KORV_Q9TTC1-Pro
GGGGSGGGGSGGGGSGGGGS	17,920	MLVMS_P03355_3mut
EAAAKGGG	17,921	PERV_Q4VFZ2
GGSEAAAKGGG	17,922	KORV_Q9TTC1-Pro_3mut
GSSGGSGGG	17,923	PERV_Q4VFZ2_3mutA_WS
GGGGGS	17,924	PERV_Q4VFZ2_3mut
GSAGSAAGSGEF	17,925	PERV_Q4VFZ2
PAPEAAAKGSS	17,926	BAEVM_P10272_3mutA
GSSPAPGGG	17,927	MLVCB_P08361_3mut
GGGGSSPAP	17,928	KORV_Q9TTC1-Pro_3mutA
PAPGGSGGG	17,929	MLVFF_P26809_3mut
GSSPAP	17,930	KORV_Q9TTC1_3mutA
PAPGSS	17,931	SFV3L_P27401-Pro_2mut
GGSGGGGSS	17,932	MLVMS_P03355_PLV919
GSSGGS	17,933	WMSV_P03359_3mutA
EAAAKGGGGGS	17,934	PERV_Q4VFZ2
GGGGG	17,935	KORV_Q9TTC1_3mutA
EAAAKGSS	17,936	MLVMS_P03355_PLV919
EAAAKEAAAKEAAAKEAAAKEAAAK	17,937	FLV_P10273_3mut
EAAAKEAAAKEAAAKEAAAK	17,938	SFV3L_P27401-Pro_2mut
GSAGSAAGSGEF	17,939	SFV3L_P27401_2mutA
GGGPAPGGS	17,940	FLV_P10273_3mutA
GGSEAAAKGGG	17,941	MLVCB_P08361_3mut
PAPGGGEAAAK	17,942	BAEVM_P10272_3mut
EAAAKPAPGSS	17,943	FOAMV_P14350_2mut
GGSEAAAK	17,944	KORV_Q9TTC1_3mutA
GGSGSS	17,945	AVIRE_P03360
GGSPAPEAAAK	17,946	MLVMS_P03355_PLV919
GGGGS	17,947	XMRV6_A1Z651_3mut
GGSPAPGGG	17,948	XMRV6_A1Z651_3mut
EAAAKPAPGGS	17,949	PERV_Q4VFZ2
GSSPAP	17,950	BAEVM_P10272_3mut
GGSGSSGGG	17,951	FLV_P10273_3mutA
PAPGGG	17,952	PERV_Q4VFZ2_3mutA_WS
GSSGGSEAAAK	17,953	MLVBM_Q7SVK7_3mut
GGSEAAAK	17,954	MLVMS_P03355_3mut
GGGPAPGGS	17,955	MLVFF_P26809_3mut
GSAGSAAGSGEF	17,956	MLVBM_Q7SVK7_3mutA_WS

EAAAKPAPGGS	17,957	SFVCP_Q87040
PAPGGG	17,958	PERV_Q4VFZ2_3mutA_WS
GSSPAPEAAAK	17,959	MLVBM_Q7SVK7
PAPEAAAK	17,960	MLVBM_Q7SVK7_3mut
PAPGGGGGS	17,961	AVIRE_P03360_3mutA
GGSEAAAKPAP	17,962	MLVBM_Q7SVK7_3mut
EAAAKGSS	17,963	WMSV_P03359_3mutA
GGGEAAAK	17,964	MLVFF_P26809_3mutA
EAAAKEAAAKEAAAK	17,965	MLVMS_P03355_3mut
PAPEAAAKGGG	17,966	BAEVM_P10272_3mut
PAPAPAP	17,967	MLVCB_P08361_3mut
EAAAKPAPGGS	17,968	BAEVM_P10272_3mut
GGGGSGGGGS	17,969	FLV_P10273_3mut
GGGGSEAAAKGGGGS	17,970	KORV_Q9TTC1_3mut
EAAAK	17,971	FLV_P10273_3mut
PAPAPAP	17,972	WMSV_P03359_3mut
GGGGSEAAAKGGGGS	17,973	FFV_O93209-Pro
GGSPAPEAAAK	17,974	MLVMS_P03355_3mut
GGSGSSGGG	17,975	XMRV6_A1Z651_3mut
GGSPAPGSS	17,976	PERV_Q4VFZ2_3mut
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,977	SFV3L_P27401-Pro_2mutA
EAAAKGGGPAP	17,978	BAEVM_P10272_3mutA
GSSGGSEAAAK	17,979	MLVMS_P03355_3mutA_WS
SGSETPGTSESATPES	17,980	PERV_Q4VFZ2_3mutA_WS
EAAAKEAAAKEAAAKEAAAKEAAAK	17,981	KORV_Q9TTC1-Pro_3mutA
GSSGSSGSS	17,982	KORV_Q9TTC1_3mutA
GSSPAPGGG	17,983	SFV3L_P27401-Pro_2mutA
GSSGGGEAAAK	17,984	KORV_Q9TTC1_3mutA
GGSGGGGSS	17,985	PERV_Q4VFZ2_3mutA_WS
GSSGGGEAAAK	17,986	MLVCB_P08361_3mut
GSSEAAAKGGG	17,987	MLVCB_P08361_3mut
GGSGGGGSS	17,988	KORV_Q9TTC1_3mutA
GGSGSSPAP	17,989	PERV_Q4VFZ2_3mutA_WS
GSSPAP	17,990	MLVMS_P03355_3mut
GGGGSSEAAAK	17,991	AVIRE_P03360
GGS	17,992	WMSV_P03359_3mut
EAAAKEAAAK	17,993	PERV_Q4VFZ2_3mut
PAPAPAPAP	17,994	MLVAV_P03356_3mut
GGSEAAAKGGG	17,995	KORV_Q9TTC1_3mutA
PAPGGG	17,996	MLVAV_P03356_3mut
EAAAKGSS	17,997	BAEVM_P10272_3mut

[1658]

[1659]

GGGGSGGGGS	17,998	WMSV_P03359_3mutA
GGSGGSGGS	17,999	SFV3L_P27401_2mut
EAAAK	18,000	MLVCB_P08361_3mut
GGGGSSGGS	18,001	WMSV_P03359_3mutA
GGGPAPEAAAK	18,002	MLVAV_P03356_3mutA
EAAAKEAAAKEAAAK	18,003	FFV_O93209
GSSEAAAKGGG	18,004	MLVBM_Q7SVK7_3mut
GGGPAPGGS	18,005	FLV_P10273_3mut
GGSEAAAKGGG	18,006	WMSV_P03359_3mut
EAAAKGGGGGS	18,007	XMRV6_A1Z651_3mutA
EAAAKGGSGGG	18,008	FLV_P10273_3mutA
GGSEAAAKGGG	18,009	SFV3L_P27401_2mutA
GGGGS	18,010	PERV_Q4VFZ2_3mutA_WS
GSSGGS	18,011	MLVMS_P03355_3mut
GSSGSS	18,012	MLVAV_P03356_3mutA
GGSPAPGGG	18,013	MLVBM_Q7SVK7_3mutA_WS
GSSGGGGGS	18,014	MLVF5_P26810_3mut
PAPAPAPAP	18,015	MLVCB_P08361_3mut
PAPAP	18,016	PERV_Q4VFZ2_3mutA_WS
PAPGSSGGS	18,017	KORV_Q9TTC1_3mut
PAPGSSGGG	18,018	PERV_Q4VFZ2_3mut
GGGEAAAK	18,019	MLVMS_P03355_PLV919
GGSGGSGGSGGSGGS	18,020	SFV3L_P27401-Pro_2mutA
GGSGGG	18,021	FLV_P10273_3mut
PAPEAAAKGGG	18,022	MLVFF_P26809_3mut
PAP	18,023	PERV_Q4VFZ2_3mutA_WS
PAPGGSGSS	18,024	FFV_O93209_2mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	18,025	FFV_O93209-Pro_2mut
GSSGSSGSSGSS	18,026	FFV_O93209-Pro
GSSGSSGSSGSSGSS	18,027	FLV_P10273_3mutA
GGGEAAAKPAP	18,028	PERV_Q4VFZ2
PAPGSSGGG	18,029	SFV3L_P27401_2mut
PAPGGSGSS	18,030	KORV_Q9TTC1-Pro_3mut
PAPAPAPAPAP	18,031	GALV_P21414_3mutA
GGSGGGEAAAK	18,032	PERV_Q4VFZ2_3mut
GSSPAP	18,033	MLVCB_P08361_3mut
EAAAKPAP	18,034	MLVF5_P26810_3mut
GGGGSGGGGSGGGGSGGGGS	18,035	MLVBM_Q7SVK7_3mut
GGSGGG	18,036	WMSV_P03359_3mut
GGSGGSGGS	18,037	KORV_Q9TTC1_3mut
GGGGGGGG	18,038	MLVFF_P26809_3mut

[1660]

GGGGSS	18,039	MLVAV_P03356_3mut
GSSGGGGGS	18,040	SFV3L_P27401_2mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	18,041	GALV_P21414_3mutA
GSSGSSGSS	18,042	PERV_Q4VFZ2_3mut
GSSPAPGGS	18,043	MLVFF_P26809_3mut
PAPAPAP	18,044	AVIRE_P03360_3mutA
EAAAKEAAAKEAAAKEAAAK	18,045	WMSV_P03359_3mutA
PAPAPAPAP	18,046	SFV3L_P27401_2mutA
GGGGSS	18,047	MLVAV_P03356_3mutA
GSSGSSGSSGSSGSS	18,048	SFV3L_P27401_2mutA
PAPGGS	18,049	WMSV_P03359_3mutA
GSSEAAAKGGG	18,050	PERV_Q4VFZ2
GSSGGSPAP	18,051	MLVMS_P03355_PLV919
GSSGSSGSSGSSGSSGSS	18,052	SFV3L_P27401_2mutA
GGSGSSGGG	18,053	MLVCB_P08361_3mut
GGGPAPGSS	18,054	SFV3L_P27401-Pro_2mutA
GSSEAAAKGGS	18,055	WMSV_P03359_3mut
GSSEAAAKGGG	18,056	MLVAV_P03356_3mut
GGSGGGPAP	18,057	FFV_O93209-Pro
GSSGSS	18,058	PERV_Q4VFZ2_3mut
PAPGGGGGS	18,059	GALV_P21414_3mutA
EAAAKPAPGGS	18,060	MLVAV_P03356_3mut
GSSGSS	18,061	MLVMS_P03355_3mut
EAAAKPAPGGS	18,062	FFV_O93209-Pro
GGGPAPEAAAK	18,063	MLVMS_P03355_3mutA_WS
GSSEAAAKGGG	18,064	MLVBM_Q7SVK7_3mut
GGGEAAAKGGS	18,065	BAEVM_P10272_3mut
GSSGSS	18,066	KORV_Q9TTC1-Pro_3mutA
EAAAKEAAAKEAAAK	18,067	SFV1_P23074
PAPGSSGGS	18,068	KORV_Q9TTC1-Pro_3mut
PAPAPAPAPAP	18,069	MLVMS_P03355
GSSEAAAK	18,070	SFV3L_P27401_2mut
PAP	18,071	PERV_Q4VFZ2_3mut
GGSEAAAKGGG	18,072	MLVBM_Q7SVK7_3mut
GGSGGGPAP	18,073	MLVBM_Q7SVK7_3mutA_WS
GSSGSS	18,074	MLVMS_P03355_3mut
GGSEAAAK	18,075	MLVMS_P03355
GSSEAAAKGGS	18,076	MLVMS_P03355_PLV919
PAPGGGGGS	18,077	MLVFF_P26809_3mut
GSSGGG	18,078	PERV_Q4VFZ2_3mut
GSSGGS	18,079	PERV_Q4VFZ2_3mutA_WS

PAPGGG	18,080	BAEVM_P10272_3mut
PAPGSSGGG	18,081	MLVBM_Q7SVK7_3mut
GGSEAAAK	18,082	SFV3L_P27401_2mut
GSSPAPEAAAK	18,083	SFV3L_P27401-Pro_2mut
GSSGGSPAP	18,084	BAEVM_P10272_3mut
GGSPAPGSS	18,085	PERV_Q4VFZ2_3mutA_WS
GGSGGSGGS	18,086	PERV_Q4VFZ2
GGSGGGPAP	18,087	FLV_P10273_3mut
GGGPAPEAAAK	18,088	SFV3L_P27401_2mutA
GGGGS	18,089	FLV_P10273_3mutA
GSSGGSGGG	18,090	XMRV6_A1Z651_3mut
EAAAKGGGGSS	18,091	PERV_Q4VFZ2
GGSGSSGGG	16,901	SFV3L_P27401-Pro_2mutA
GGSGGSGGS	16,902	MLVFF_P26809_3mut
GGGPAPEAAAK	16,903	FLV_P10273_3mut
GSSGGGEAAAK	16,904	MLVMS_P03355_3mut
GGG	16,905	SFV3L_P27401_2mut
GSAGSAAGSGEF	16,906	WMSV_P03359_3mut
GSSGGGPAP	16,907	MLVMS_P03355_PLV919
GGGGSS	16,908	KORV_Q9TTC1-Pro_3mut
GGGGSSEAAAK	16,909	KORV_Q9TTC1
PAPGGSGGG	16,910	SFV3L_P27401_2mut
GSSGSSGSSGSSGSS	16,911	FFV_O93209
GSSGGSPAP	16,912	MLVMS_P03355_3mut
GGSEAAAK	16,913	KORV_Q9TTC1-Pro_3mutA
GGGGSGGGGS	16,914	BAEVM_P10272_3mut
GSSEAAAKGGG	16,915	AVIRE_P03360_3mut
EAAAKPAPGGG	16,916	FLV_P10273_3mut
EAAAKGGSPAP	16,917	SFV3L_P27401-Pro_2mutA
GSSEAAAKPAP	16,918	MLVBM_Q7SVK7_3mut
GGGPAPGGS	16,919	MLVCB_P08361_3mut
GGG	16,920	SFV3L_P27401_2mutA
EAAAKGGGGSEAAAK	16,921	SFV3L_P27401_2mutA
GGSGSSGGG	16,922	MLVBM_Q7SVK7_3mut
GSAGSAAGSGEF	16,923	BAEVM_P10272_3mut
GGGEAAAK	16,924	FOAMV_P14350_2mutA
PAPEAAAKGGS	16,925	WMSV_P03359_3mut
PAPAPAPAPAPAP	16,926	MLVF5_P26810_3mutA
GGSGGGGSS	16,927	FLV_P10273_3mutA
PAPGSSGGS	16,928	BAEVM_P10272_3mut
PAPEAAAK	16,929	WMSV_P03359_3mutA

[1661]

GSSGSSGSSGSSGSSGSS	16,930	FFV_O93209-Pro_2mut
GGGGGSGSS	16,931	FFV_O93209-Pro
GGGGGGGG	16,932	SFV3L_P27401-Pro_2mutA
GGGGGG	16,933	FLV_P10273_3mut
GSSGGSGGG	16,934	MLVAV_P03356_3mutA
GGGGSS	16,935	SFV3L_P27401-Pro_2mutA
GGSGGGPAP	16,936	FOAMV_P14350_2mut
GSSGSS	16,937	AVIRE_P03360_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	16,938	SFV3L_P27401-Pro_2mutA
EAAAKEAAAK	16,939	BAEVM_P10272_3mut
GSSPAPEAAAK	16,940	GALV_P21414_3mutA
GGSEAAAKPAP	16,941	SFV3L_P27401_2mutA
GGSGGGEAAAK	16,942	SFV3L_P27401-Pro_2mutA
EAAAKGSSPAP	16,943	FOAMV_P14350_2mut
GGSGSSEAAAK	16,944	SFV3L_P27401_2mut
GGG	16,945	PERV_Q4VFZ2
GGGGGSGSS	16,946	FOAMV_P14350_2mut
GGSGGGEAAAK	16,947	KORV_Q9TTC1-Pro_3mut
GSSGGSGGG	16,948	AVIRE_P03360_3mutA
EAAAKPAPGGG	16,949	SFV3L_P27401_2mutA
PAPGGSGGG	16,950	KORV_Q9TTC1-Pro_3mut
PAPAPAP	16,951	WMSV_P03359_3mutA
GSSEAAAKPAP	16,952	SFV1_P23074
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	16,953	SRV2_P51517
GSSGGSGGG	16,954	PERV_Q4VFZ2_3mutA_WS
GSSGSSGSSGSSGSSGSS	16,955	FFV_O93209
GSSGGGPAP	16,956	WMSV_P03359_3mut
PAPAPAPAPAPAP	16,957	MLVBM_Q7SVK7_3mut
GGGGGSPAP	16,958	KORV_Q9TTC1-Pro_3mutA
PAPGSS	16,959	MLVBM_Q7SVK7_3mutA_WS
PAPEAAAKGGS	16,960	SFV3L_P27401-Pro_2mut
GGGGSSPAP	16,961	MLVMS_P03355_3mut
GGSEAAAK	16,962	FFV_O93209-Pro
EAAAKPAPGGS	16,963	AVIRE_P03360_3mutA
PAPGSS	16,964	WMSV_P03359_3mut
PAPGSSGGG	16,965	SFV3L_P27401-Pro_2mutA
EAAAKEAAAKEAAAK	16,966	SFV3L_P27401_2mut
GGS	16,967	MLVRD_P11227_3mut
GGGGS	16,968	KORV_Q9TTC1-Pro_3mut
GGSGGGGSS	16,969	KORV_Q9TTC1
GGSGGG	16,970	MLVMS_P03355_3mutA_WS

[1662]

GGGEAAAKPAP	16,971	BAEVM_P10272_3mut
EAAAKEAAAKEAAAKEAAAKEAAAK	16,972	FLV_P10273
PAPGGSGGG	16,973	KORV_Q9TTC1-Pro_3mutA
GSSGSSGSSGSSGSSGSS	16,974	HTL1L_P0C211
GGGEAAAKPAP	16,975	WMSV_P03359
GSSGGSPAP	16,976	FFV_O93209-Pro
PAPAPAPAPAP	16,977	SFV3L_P27401-Pro_2mutA
GSSGGSEAAAK	16,978	SFV3L_P27401_2mutA
GGSPAPGSS	16,979	SFV3L_P27401_2mut
GGSGGSGGS	16,980	KORV_Q9TTC1-Pro_3mut
PAPEAAAKGSS	16,981	KORV_Q9TTC1-Pro_3mut
EAAAKGGS	16,982	KORV_Q9TTC1_3mutA
EAAAKGGGGSEAAAK	16,983	SFV3L_P27401-Pro_2mut
GGGGSSPAP	16,984	FFV_O93209-Pro
EAAAK	16,985	SFV3L_P27401_2mut
EAAAKGGGGSS	16,986	BAEVM_P10272_3mut
GGGGGSEAAAK	16,987	MLVBM_Q7SVK7_3mut
GGGG	16,988	PERV_Q4VFZ2
GGGGGSEAAAK	16,989	FLV_P10273_3mut
EAAAKGGGPAP	16,990	KORV_Q9TTC1-Pro
GGGGSGGGGSGGGGSGGGGS	16,991	FFV_O93209_2mutA
GSSGGSGGG	16,992	PERV_Q4VFZ2_3mut
GGGGSGGGGSGGGGS	16,993	GALV_P21414_3mutA
GGSGGGEAAAK	16,994	AVIRE_P03360_3mutA
PAPEAAAKGGG	16,995	SFV3L_P27401_2mut
GGGGSGGGGS	16,996	AVIRE_P03360
GSSGGGEAAAK	16,997	SFV3L_P27401_2mutA
GGGGG	16,998	AVIRE_P03360_3mutA
GGSGSS	16,999	KORV_Q9TTC1_3mut
PAPAPAPAPAPAP	17,000	FOAMV_P14350_2mut
GGSEAAAKPAP	17,001	KORV_Q9TTC1-Pro_3mut
GGGGGG	17,002	PERV_Q4VFZ2_3mut
GSSGGGEAAAK	17,003	MLVBM_Q7SVK7
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,004	MLVAV_P03356
GGSPAPGSS	17,005	BAEVM_P10272_3mut
GGGGSSPAP	17,006	BAEVM_P10272
GGGGSEAAAKGGGGS	17,007	SFV3L_P27401_2mut
GGGGGGGG	17,008	GALV_P21414_3mutA
PAPAP	17,009	MLVAV_P03356_3mut
GGGEAAAK	17,010	PERV_Q4VFZ2_3mutA_WS
GSSPAPGGG	17,011	FFV_O93209_2mut

[1663]

[1664]

GGSGGSGGSGGSGGS	17,012	BAEVM_P10272
GGGGGS	17,013	MLVF5_P26810_3mutA
PAPGGGGSS	17,014	FLV_P10273_3mutA
GGGEAAAK	17,015	MLVBM_Q7SVK7_3mut
PAPEAAAKGGG	17,016	WMSV_P03359_3mut
GSSEAAAK	17,017	MLVBM_Q7SVK7_3mut
EAAAKEAAAK	17,018	AVIRE_P03360
EAAAKGGGGGS	17,019	MLVBM_Q7SVK7_3mut
GGGEAAAKGGS	17,020	SFV3L_P27401-Pro_2mutA
PAPAPAPAPAP	17,021	MLVF5_P26810_3mut
PAPGSSEAAAK	17,022	SFV3L_P27401-Pro_2mutA
EAAAKEAAAKEAAAK	17,023	BAEVM_P10272_3mutA
GGSPAPGSS	17,024	MLVMS_P03355
PAPGSSGGS	17,025	FLV_P10273_3mutA
EAAAKEAAAKEAAAKEAAAK	17,026	FOAMV_P14350-Pro_2mut
EAAAKGGG	17,027	KORV_Q9TTC1_3mutA
EAAAKGGSGGG	17,028	MLVBM_Q7SVK7_3mut
GGGGGS	17,029	KORV_Q9TTC1-Pro_3mutA
PAPGGSGGG	17,030	WMSV_P03359_3mut
GGGPAPGGS	17,031	KORV_Q9TTC1_3mutA
GSS	17,032	FFV_O93209
GGSGGSGGS	17,033	PERV_Q4VFZ2_3mut
GGGGS	17,034	GALV_P21414_3mutA
GGGG	17,035	MLVF5_P26810_3mut
GGSEAAAKPAP	17,036	FFV_O93209-Pro_2mut
PAPAPAPAP	17,037	FFV_O93209-Pro
PAP	17,038	MLVF5_P26810_3mut
EAAAKEAAAKEAAAK	17,039	FFV_O93209_2mut
EAAAKGSS	17,040	MLVCB_P08361_3mut
EAAAKGGG	17,041	MLVBM_Q7SVK7_3mut
PAPEAAAKGGG	17,042	FFV_O93209_2mut
GSSGGGEAAAK	17,043	SFV1_P23074-Pro_2mut
PAPGGGEAAAK	17,044	GALV_P21414_3mutA
GGGGSGGGGSGGGGSGGGGS	17,045	FOAMV_P14350-Pro_2mutA
GSSGGG	17,046	FOAMV_P14350_2mut
GGGGSGGGGSGGGGSGGGGS	17,047	SFV3L_P27401_2mutA
GGSGSS	17,048	AVIRE_P03360_3mut
GGSGSSEAAAK	17,049	MMTVB_P03365_WS
PAPAPAP	17,050	MLVAV_P03356_3mutA
GSSGGSPAP	17,051	SFV3L_P27401-Pro_2mut
GGSPAP	17,052	AVIRE_P03360

[1665]

GGSGGGPAP	17,053	FFV_O93209
GSSEAAAK	17,054	PERV_Q4VFZ2
GSSGGGPAP	17,055	PERV_Q4VFZ2_3mutA_WS
GGGGSSEAAAK	17,056	KORV_Q9TTC1_3mutA
GGSEAAAKPAP	17,057	SFVCP_Q87040
GGSGGGPAP	17,058	FOAMV_P14350_2mutA
GGGGSGGGGSGGGGSGGGGS	17,059	BLVJ_P03361_2mutB
GGGGSSPAP	17,060	SFV3L_P27401_2mutA
EAAAKGGS	17,061	MLVF5_P26810_3mut
GGSEAAAKGSS	17,062	MLVCB_P08361_3mut
GGGGSSEAAAK	17,063	SFV3L_P27401_2mut
EAAAKGGSGGG	17,064	FOAMV_P14350_2mut
GGSGGS	17,065	FLV_P10273_3mut
EAAAKGGG	17,066	FFV_O93209-Pro
GSSGSSGSSGSSGSS	17,067	SFV3L_P27401
GSSGGGPAP	17,068	PERV_Q4VFZ2_3mutA_WS
PAPGGSEAAAK	17,069	SFV3L_P27401-Pro_2mutA
GGSPAP	17,070	KORV_Q9TTC1
EAAAKPAPGSS	17,071	KORV_Q9TTC1_3mutA
SGSETPGTSESATPES	17,072	SFV1_P23074
GSSPAP	17,073	SFV3L_P27401-Pro_2mutA
GSSPAPGGG	17,074	SFV3L_P27401_2mut
GGGEAAAKGSS	17,075	SFV1_P23074_2mut
GGGPAPGGS	17,076	BAEVM_P10272_3mut
EAAAKGGG	17,077	KORV_Q9TTC1-Pro_3mutA
GSSGGG	17,078	SFV3L_P27401-Pro_2mut
GGSPAPEAAAK	17,079	BAEVM_P10272_3mut
EAAAKGSSPAP	17,080	FFV_O93209
EAAAKGGGGSEAAAK	17,081	SFV3L_P27401-Pro_2mutA
GSSGSSGSSGSSGSS	17,082	SFV1_P23074_2mut
EAAAKGGSPAP	17,083	FOAMV_P14350_2mut
GGSGGS	17,084	KORV_Q9TTC1-Pro_3mutA
EAAAKGSSGGS	17,085	GALV_P21414
GSSGGGPAP	17,086	MLVAV_P03356
PAPEAAAKGGS	17,087	FOAMV_P14350_2mut
EAAAKPAPGGG	17,088	AVIRE_P03360_3mut
GGSPAP	17,089	SFV3L_P27401_2mutA
GGGGSGGGGS	17,090	SFV3L_P27401_2mutA
GGGGSS	17,091	AVIRE_P03360_3mutA
GGSPAPGGG	17,092	SFV3L_P27401-Pro_2mutA
EAAAKPAPGSS	17,093	SFV3L_P27401

EAAAKPAP	17,094	FOAMV_P14350-Pro_2mut
PAPEAAAKGSS	17,095	PERV_Q4VFZ2_3mutA_WS
EAAAKGGSGSS	17,096	SFV3L_P27401_2mutA
GGGEAAAKGSS	17,097	GALV_P21414_3mutA
GGGGSEAAAKGGGGS	17,098	PERV_Q4VFZ2_3mut
PAPGGSGSS	17,099	FFV_O93209-Pro_2mutA
GGSEAAAKPAP	17,100	GALV_P21414_3mutA
GGSGGSGGSGGSGGS	17,101	FFV_O93209-Pro
GSSGGSEAAAK	17,102	SFV3L_P27401-Pro_2mut
GGS	17,103	GALV_P21414_3mutA
PAPGGSEAAAK	17,104	MLVMS_P03355
PAPEAAAKGGS	17,105	BAEVM_P10272_3mutA
GGSGSSPAP	17,106	SFV3L_P27401-Pro_2mutA
GSSPAP	17,107	WMSV_P03359_3mut
GGGEAAAK	17,108	MMTVB_P03365
GGGGSS	17,109	PERV_Q4VFZ2_3mut
GGSPAPGSS	17,110	SFV3L_P27401-Pro_2mut
PAPGGS	17,111	MLVBM_Q7SVK7_3mut
EAAAKGSSPAP	17,112	MLVBM_Q7SVK7_3mut
GGGGSSGGS	17,113	PERV_Q4VFZ2_3mut
PAPAPAPAPAPAP	17,114	SFV1_P23074
GGSEAAAKGGG	17,115	SFV3L_P27401-Pro_2mut
GGSGGS	17,116	SFV1_P23074_2mut
GSSGGGGGS	17,117	MLVF5_P26810_3mutA
EAAAKGGGPAP	17,118	SFV3L_P27401
EAAAKEAAAKEAAAKEAAAK	17,119	FOAMV_P14350-Pro_2mutA
GGGPAPGSS	17,120	SFV3L_P27401_2mutA
GGGGSGGGGSGGGGSGGGGS	17,121	SFV3L_P27401_2mut
EAAAKEAAAKEAAAKEAAAK	17,122	MMTVB_P03365_WS
PAPGSSGGS	17,123	KORV_Q9TTC1-Pro_3mutA
PAPGSSEAAAK	17,124	FOAMV_P14350-Pro_2mut
GSSPAPEAAAK	17,125	BAEVM_P10272_3mut
EAAAKGGGGSEAAAK	17,126	FFV_O93209-Pro
GGSPAP	17,127	PERV_Q4VFZ2
GGSGSSEAAAK	17,128	XMRV6_A1Z651_3mut
GGSEAAAKGGG	17,129	GALV_P21414_3mutA
PAPGGGGSS	17,130	AVIRE_P03360_3mutA
GGSGGSGGSGGS	17,131	PERV_Q4VFZ2
GGGGSSGGS	17,132	PERV_Q4VFZ2_3mutA_WS
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,133	BAEVM_P10272_3mutA
GGGPAP	17,134	MLVAV_P03356_3mut

[1666]

[1667]

GGGGSGGGGSGGGGSGGGGS	17,135	FFV_O93209_2mut
GSSEAAAK	17,136	FFV_O93209
GGSPAPEAAAK	17,137	FOAMV_P14350_2mut
GGGGGSEAAAK	17,138	FOAMV_P14350_2mut
GSSPAPGGS	17,139	MLVBM_Q7SVK7_3mut
GSS	17,140	SFVCP_Q87040_2mut
EAAAKPAP	17,141	FOAMV_P14350-Pro
EAAAKGGG	17,142	SFV3L_P27401_2mut
GGGEAAAK	17,143	AVIRE_P03360_3mutA
PAPGSSGGG	17,144	WMSV_P03359_3mut
EAAAKGGSPAP	17,145	SFV3L_P27401
GSSGGSGGG	17,146	SFV3L_P27401-Pro_2mutA
GSSGGGEAAAK	17,147	GALV_P21414_3mutA
GGGPAPGSS	17,148	MLVBM_Q7SVK7_3mutA_WS
PAPGGGEAAAK	17,149	FFV_O93209-Pro_2mut
GSSGSSGSSGSS	17,150	SFV1_P23074_2mut
GGSEAAAK	17,151	PERV_Q4VFZ2_3mutA_WS
GGGEAAAKPAP	17,152	SFV3L_P27401_2mut
EAAAKGGGPAP	17,153	SFV3L_P27401_2mut
GGGGSSPAP	17,154	FLV_P10273_3mut
EAAAKPAPGSS	17,155	FFV_O93209_2mut
GGGGSSPAP	17,156	SFV3L_P27401_2mut
GSSGSS	17,157	KORV_Q9TTC1_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGS	17,158	BLVJ_P03361_2mut
GGGGSSGGS	17,159	GALV_P21414_3mutA
EAAAKGGSGSS	17,160	FFV_O93209-Pro
EAAAKPAP	17,161	PERV_Q4VFZ2
GSSGGGEAAAK	17,162	MLVBM_Q7SVK7_3mut
PAPGGSGGG	17,163	BAEVM_P10272
EAAAKGGGPAP	17,164	MLVF5_P26810
GSSGSSGSS	17,165	MLVBM_Q7SVK7_3mut
GSSGGS	17,166	AVIRE_P03360_3mutA
GGSEAAAKGGG	17,167	FOAMV_P14350_2mut
EAAAKGGS	17,168	MLVF5_P26810_3mutA
GGSGSSGGG	17,169	WMSV_P03359_3mut
EAAAK	17,170	SFV1_P23074_2mut
GSSGGSPAP	17,171	SFV3L_P27401-Pro_2mutA
GGGGSSGGS	17,172	KORV_Q9TTC1_3mut
PAPGGSGGG	17,173	FFV_O93209-Pro_2mut
GGGPAPGGS	17,174	SFV3L_P27401_2mutA
GSSPAPEAAAK	17,175	FLV_P10273_3mut

GGSGSSPAP	17,176	SFV3L_P27401_2mut
GSSEAAAKGGS	17,177	SFV3L_P27401_2mut
PAPGGG	17,178	SFV3L_P27401_2mutA
SGSETPGTSESATPES	17,179	KORV_Q9TTC1-Pro_3mut
GGGGS	17,180	SFV1_P23074-Pro_2mutA
GSSGGGEAAAK	17,181	WMSV_P03359
EAAAKGGGGSEAAAK	17,182	MLVF5_P26810_3mutA
GSSEAAAKPAP	17,183	FFV_O93209
GGGGGG	17,184	SFV1_P23074_2mutA
EAAAKEAAAKEAAAK	17,185	MMTVB_P03365-Pro
EAAAKPAPGSS	17,186	MLVBM_Q7SVK7_3mut
GGSGSSEAAAK	17,187	SFV3L_P27401_2mutA
GGSEAAAK	17,188	MLVMS_P03355_3mut
GGSPAPEAAAK	17,189	SFV3L_P27401_2mut
GGGPAPGSS	17,190	SFV1_P23074
GGGGGSEAAAK	17,191	MLVBM_Q7SVK7_3mutA_WS
EAAAKPAPGSS	17,192	KORV_Q9TTC1-Pro
GSSGSSGSSGSS	17,193	SFV3L_P27401_2mut
EAAAKPAP	17,194	SFV3L_P27401_2mut
GGGEAAAK	17,195	PERV_Q4VFZ2_3mut
GGSGGS	17,196	SFV3L_P27401_2mutA
EAAAKGSSGGS	17,197	MMTVB_P03365
SGSETPGTSESATPES	17,198	SFV3L_P27401
EAAAKGSSGGG	17,199	PERV_Q4VFZ2
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,200	MMTVB_P03365
GGSGGGPAP	17,201	KORV_Q9TTC1_3mutA
PAPAPAPAP	17,202	SFV3L_P27401
GGGEAAAKGGS	17,203	SFV1_P23074_2mut
GSSGGSGGG	17,204	PERV_Q4VFZ2_3mut
PAPEAAAKGGS	17,205	FOAMV_P14350_2mutA
GGGEAAAKGSS	17,206	SFV3L_P27401_2mut
GGGGSGGGGSGGGGSGGGGS	17,207	MLVBM_Q7SVK7
PAPGSSGGG	17,208	FLV_P10273
GGSGSSGGG	17,209	FFV_O93209
EAAAKPAPGSS	17,210	MLVBM_Q7SVK7
GSSEAAAKGGG	17,211	SFV3L_P27401_2mutA
GGSGGSGGSGGSGGS	17,212	MLVF5_P26810
GGSEAAAKPAP	17,213	SFV3L_P27401-Pro_2mutA
EAAAKGGSPAP	17,214	SFV3L_P27401_2mutA
EAAAKGGGGGS	17,215	SFV3L_P27401_2mut
GSSPAPEAAAK	17,216	SFV3L_P27401_2mutA

[1668]

PAPAP	17,217	MLVBM_Q7SVK7_3mut
PAPGGSEAAAK	17,218	KORV_Q9TTC1-Pro
GGSGSS	17,219	MLVF5_P26810_3mutA
GGSEAAAKPAP	17,220	FFV_O93209_2mut
GSS	17,221	MLVMS_P03355
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,222	SFV3L_P27401-Pro
PAPGGGEAAAK	17,223	SFV3L_P27401_2mut
PAPGGGGGS	17,224	SFV3L_P27401-Pro_2mut
PAPGGSGSS	17,225	BAEVM_P10272_3mut
GSSGGGEAAAK	17,226	FFV_O93209
GGSEAAAKPAP	17,227	SFV1_P23074_2mut
GGGGG	17,228	FLV_P10273_3mut
GGGEAAAKGSS	17,229	SFV3L_P27401
GSSGSSGSSGSSGSS	17,230	SFV1_P23074-Pro
SGSETPGTSESATPES	17,231	AVIRE_P03360
PAPGSSGGG	17,232	MLVBM_Q7SVK7_3mut
GGGGSSPAP	17,233	HTL3P_Q4U0X6_2mut
GGGEAAAK	17,234	SFV1_P23074
GGSGGG	17,235	AVIRE_P03360
EAAAKGSSGGG	17,236	SFV3L_P27401_2mutA
GSSPAPEAAAK	17,237	FOAMV_P14350-Pro_2mutA
GGGPAPGSS	17,238	WMSV_P03359
EAAAKGSSGGG	17,239	MLVMS_P03355
GGGGGSEAAAK	17,240	MLVMS_P03355
EAAAKPAPGGS	17,241	SFV3L_P27401
EAAAKGSSPAP	17,242	SFV3L_P27401
GGGGGGG	17,243	FOAMV_P14350_2mutA
EAAAKEAAAKEAAAK	17,244	SFV3L_P27401
GSSPAPGGS	17,245	FFV_O93209_2mutA
GGGGSSEAAAK	17,246	SFV3L_P27401-Pro_2mutA
GGSEAAAKGSS	17,247	GALV_P21414_3mutA
GGSEAAAKGSS	17,248	BAEVM_P10272_3mutA
EAAAKPAPGGG	17,249	MLVCB_P08361
GSSGSSGSSGSSGSSGSS	17,250	SFV1_P23074-Pro
GGGGSEAAAKGGGGS	17,251	FOAMV_P14350_2mut
GSSPAPGGS	17,252	MLVMS_P03355_PLV919
GGGGSGGGGS	17,253	FFV_O93209-Pro
GSSGGSPAP	17,254	KORV_Q9TTC1_3mutA
GGSGGS	17,255	GALV_P21414_3mutA
PAPGSSEAAAK	17,256	WMSV_P03359
PAPGGGGSS	17,257	MMTVB_P03365-Pro

[1669]

GGGGSSGGS	17,258	PERV_Q4VFZ2_3mutA_WS
GGGGSGGGGS	17,259	FFV_O93209_2mut
GGGGSGGGGSGGGGSGGGGS	17,260	XMRV6_A1Z651
GGSGSSEAAAK	17,261	SFV1_P23074_2mut
GGSGGGGSS	17,262	GALV_P21414_3mutA
GGSEAAAKPAP	17,263	MLVBM_Q7SVK7
EAAAKGGSPAP	17,264	SFV1_P23074_2mutA
PAPAPAPAP	17,265	FFV_O93209
GSSGGSPAP	17,266	MMTVB_P03365-Pro
GGGGGSPAP	17,267	KORV_Q9TTC1_3mutA
EAAAKGGGPAP	17,268	PERV_Q4VFZ2
GSSGGSPAP	17,269	BAEVM_P10272
GGGGG	17,270	FFV_O93209
GGGGGS	17,271	FLV_P10273_3mutA
EAAAKEAAAKEAAAK	17,272	FOAMV_P14350
PAPGGG	17,273	MLVCB_P08361_3mut
GSSGGSEAAAK	17,274	FOAMV_P14350_2mutA
GGSPAPGGG	17,275	FLV_P10273_3mut
GSSGSSGSSGSSGSSGSS	17,276	SFV1_P23074-Pro_2mutA
GGSPAPEAAAK	17,277	SFV3L_P27401
PAPGGGGSS	17,278	HTL3P_Q4U0X6_2mutB
GGGGSSEAAAK	17,279	MMTVB_P03365_2mut_WS
PAPGGS	17,280	MLVRD_P11227_3mut
GGSGGSGGSGGSGGS	17,281	MMTVB_P03365
GSAGSAAGSGEF	17,282	AVIRE_P03360
GSSGGS	17,283	BAEVM_P10272_3mutA
GGSGGGGSS	17,284	MMTVB_P03365
GGSGGGGSS	17,285	WMSV_P03359
PAPEAAAKGSS	17,286	SFV1_P23074
GSSGSSGSSGSS	17,287	SFV1_P23074-Pro_2mutA
PAPAPAPAPAPAP	17,288	SFV3L_P27401
PAPGSSGGG	17,289	FLV_P10273_3mut
GGSGSSPAP	17,290	MLVMS_P03355
GGSGGGPAP	17,291	FOAMV_P14350
PAPGGGGGS	17,292	KORV_Q9TTC1_3mutA
EAAAKGSSPAP	17,293	GALV_P21414_3mutA
GGSGSSPAP	17,294	MLVBM_Q7SVK7_3mut
EAAAKGSS	17,295	SFV3L_P27401_2mut
GGGGGSEAAAK	17,296	WMSV_P03359
GGGGGGGG	17,297	SFV1_P23074-Pro
EAAAKEAAAK	17,298	MLVBM_Q7SVK7

[1670]

GGGEAAAKGGS	17,299	MLVBM_Q7SVK7
EAAAKGGSPAP	17,300	SFV3L_P27401_2mut
GSSEAAAK	17,301	XMRV6_A1Z651
PAPGGGEAAAK	17,302	MMTVB_P03365_WS
GGSPAP	17,303	GALV_P21414_3mutA
GSSPAPGGG	17,304	MLVBM_Q7SVK7_3mutA_WS
GGSGSSPAP	17,305	SFV1_P23074_2mutA
GGS	17,306	HTL32_Q0R5R2_2mut
GGSGGGGSS	17,307	MMTVB_P03365-Pro
GGGGSGGGGSGGGGSGGGGS	17,308	SFVCP_Q87040_2mutA
EAAAKGGGPAP	17,309	FOAMV_P14350_2mut
GSSGGGEAAAK	17,310	MMTVB_P03365
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,311	MLVBM_Q7SVK7_3mutA_WS
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,312	MMTVB_P03365_WS
EAAAKEAAAK	17,313	FOAMV_P14350-Pro_2mut
GSSPAPEAAAK	17,314	FOAMV_P14350_2mutA
EAAAKPAPGGS	17,315	GALV_P21414_3mutA
GSSGGSPAP	17,316	KORV_Q9TTC1-Pro_3mut
GGGPAPEAAAK	17,317	MLVAV_P03356
GGGEAAAKPAP	17,318	SFV1_P23074-Pro_2mut
GGGGGSEAAAK	17,319	SFV3L_P27401_2mut
GGGPAPGSS	17,320	SFV3L_P27401_2mut
GGSEAAAKPAP	17,321	AVIRE_P03360
GSSGSSGSSGSSGSSGSS	17,322	SFV1_P23074-Pro_2mut
EAAAKGSSGGS	17,323	FOAMV_P14350_2mutA
GGGGGG	17,324	MLVBM_Q7SVK7_3mut
GSSPAPGGS	17,325	PERV_Q4VFZ2
GGSGSSPAP	17,326	GALV_P21414_3mutA
GGGPAPEAAAK	17,327	SFV3L_P27401
GGSGGGEAAAK	17,328	WMSV_P03359
GSAGSAAGSGEF	17,329	SFV1_P23074_2mut
GSSGGGEAAAK	17,330	MLVMS_P03355
GGG	17,331	MMTVB_P03365-Pro
PAPGSSGGS	17,332	FOAMV_P14350_2mut
GGGGSSPAP	17,333	FFV_O93209_2mut
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,334	MMTVB_P03365_WS
GGGGGGG	17,335	XMRV6_A1Z651
PAPAPAPAPAP	17,336	FOAMV_P14350
GGGGSGGGGSGGGGSGGGGS	17,337	MMTVB_P03365_2mut_WS

[1671]

GGSGGGPAP	17,338	SFV3L_P27401_2mut
GGGGGG	17,339	SFV1_P23074-Pro
EAAAKPAPGSS	17,340	SFV3L_P27401_2mut
GGGGSSGGS	17,341	HTL3P_Q4U0X6_2mut
PAPGSSEAAAK	17,342	MMTVB_P03365-Pro
GGGGSSPAP	17,343	FOAMV_P14350-Pro_2mut
PAPGSSGGS	17,344	MMTVB_P03365
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,345	SRV2_P51517
PAPAPAP	17,346	MMTVB_P03365_2mut_WS
PAPGGGGGS	17,347	MMTVB_P03365_2mutB
GGGGSS	17,348	SFV1_P23074-Pro_2mutA
EAAAKEAAAKEAAAKEAAAK	17,349	SFV3L_P27401-Pro
GGSGGSGGSGGSGGS	17,350	MMTVB_P03365-Pro
GGGGGGG	17,351	SFV3L_P27401_2mut
PAPGGGEAAAK	17,352	SFV3L_P27401
PAPGSS	17,353	FOAMV_P14350_2mutA
GGGGSGGGGS	17,354	SFVCP_Q87040_2mutA
GSSGGSGGG	17,355	XMRV6_A1Z651
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,356	MLVBM_Q7SVK7
GSSEAAAKGGG	17,357	FFV_O93209-Pro_2mut
GGSEAAAKPAP	17,358	SFV3L_P27401-Pro
GSSGGSGGG	17,359	SFV1_P23074_2mut
EAAAKGGGGSS	17,360	FOAMV_P14350_2mutA
GGGGGG	17,361	SFV3L_P27401_2mut
GGGGG	17,362	MLVBM_Q7SVK7_3mut
PAPEAAAKGGG	17,363	SFV3L_P27401
EAAAKGGSPAP	17,364	KORV_Q9TTC1_3mutA
GGGEAAAKPAP	17,365	SFV1_P23074_2mut
GSSGSSGSSGSSGSSGSS	17,366	KORV_Q9TTC1-Pro
EAAAKEAAAKEAAAKEAAAK	17,367	SFVCP_Q87040
PAPGSSEAAAK	17,368	MLVBM_Q7SVK7
GSSGSSGSS	17,369	FFV_O93209-Pro_2mut
GSSGGGPAP	17,370	SFV3L_P27401-Pro_2mut
GGGPAPEAAAK	17,371	WMSV_P03359_3mut
GGGEAAAK	17,372	MMTVB_P03365-Pro
GSSGSSGSSGSS	17,373	SFV3L_P27401-Pro_2mutA
PAPAPAPAPAP	17,374	FFV_O93209-Pro
GGSPAPEAAAK	17,375	FFV_O93209-Pro_2mut
GSSGSSGSSGSSGSSGSS	17,376	GALV_P21414
EAAAKEAAAKEAAAKEAAAKEAAAK	17,377	FOAMV_P14350
GGGPAPEAAAK	17,378	MMTVB_P03365-Pro

[1672]

PAPGGSGGG	17,379	MLVF5_P26810_3mutA
PAPGGSGGG	17,380	FLV_P10273_3mut
GGGEAAAKGGS	17,381	SFV3L_P27401
GSAGSAAGSGEF	17,382	MLVBM_Q7SVK7_3mut
GSSPAPGGG	17,383	MPMV_P07572_2mutB
GSSGSSGSSGSSGSSGSS	17,384	FOAMV_P14350
GGSGGGGSS	17,385	BLVJ_P03361_2mut
PAPEAAAKGSS	17,386	SFV1_P23074-Pro
GGG	17,387	FFV_O93209
EAAAKGGGGSS	17,388	SFV1_P23074_2mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,389	SRV2_P51517
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,390	MMTVB_P03365
GGGEAAAKGGS	17,391	MMTVB_P03365_WS
GSSGSS	17,392	SFV1_P23074
GSSGGGGGS	17,393	SFV3L_P27401
GGGGSSEAAAK	17,394	SFV1_P23074
EAAAKGSSGGS	17,395	HTL1A_P03362_2mutB
GSSEAAAKGGS	17,396	GALV_P21414_3mutA
EAAAKGSSPAP	17,397	SFV1_P23074
EAAAKPAPGSS	17,398	SFV3L_P27401_2mutA
PAPGSSGGG	17,399	SFV3L_P27401-Pro_2mut
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,400	SFV3L_P27401-Pro
EAAAKEAAAKEAAAKEAAAKEAAAK	17,401	MMTVB_P03365_WS
GGGGSSEAAAK	17,402	MLVF5_P26810_3mutA
EAAAKGGSPAP	17,403	GALV_P21414
PAPEAAAKGSS	17,404	MMTVB_P03365_WS
GSSGGGGGS	17,405	SFVCP_Q87040_2mut
GGGGSSPAP	17,406	SFV1_P23074
EAAAKGGGGSS	17,407	XMRV6_A1Z651
PAPAPAPAP	17,408	MMTVB_P03365
GGSEAAAKGSS	17,409	SFV3L_P27401_2mutA
GSSPAPGGG	17,410	MMTVB_P03365_WS
GGGGGG	17,411	SFV3L_P27401-Pro
GGSGGSGGS	17,412	FOAMV_P14350-Pro_2mut
PAPAPAPAPAPAP	17,413	WMSV_P03359
GSSPAP	17,414	MLVBM_Q7SVK7
GGGGGSGSS	17,415	MMTVB_P03365_2mut_WS
EAAAKGSSGGS	17,416	MMTVB_P03365_2mutB_WS
EAAAK	17,417	FFV_O93209_2mutA
PAPEAAAK	17,418	SFV1_P23074-Pro
EAAAKGGSGSS	17,419	SFV3L_P27401

[1673]

[1674]

GGSGGSGGS	17,420	FFV_O93209-Pro
GSSGGGEAAAK	17,421	MMTVB_P03365
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,422	MLVFF_P26809_3mutA
GGSGGSGGSGGSGGSGGS	17,423	HTL1L_P0C211_2mutB
GGGEAAAK	17,424	SFV3L_P27401-Pro_2mutA
GGGGGSGSS	17,425	MMTVB_P03365
GSSPAPGGS	17,426	FOAMV_P14350_2mutA
EAAAKGSS	17,427	MLVMS_P03355
GSSGGSGGG	17,428	FFV_O93209-Pro
GGSGGGGSS	17,429	MMTVB_P03365-Pro_2mut
GGSPAPGSS	17,430	FOAMV_P14350_2mut
GGSGGSGGSGGSGGSGGS	17,431	SFVCP_Q87040-Pro_2mut
GSSEAAAKGGG	17,432	FOAMV_P14350_2mutA
GGSGGSGGS	17,433	MMTVB_P03365-Pro
GSSGSSGSSGSSGSSGSS	17,434	MMTVB_P03365_2mut_WS
GSSGSSGSSGSSGSS	17,435	MMTVB_P03365-Pro
PAPEAAAK	17,436	WDSV_O92815
GSSGSSGSSGSSGSS	17,437	FFV_O93209-Pro_2mut
EAAAKGGGGSEAAAK	17,438	MMTVB_P03365-Pro
GGSPAPEAAAK	17,439	FOAMV_P14350
GSSGSS	17,440	PERV_Q4VFZ2
GGG	17,441	MMTVB_P03365-Pro
GGGGSGGGGSGGGGS	17,442	FFV_O93209_2mut
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,443	MMTVB_P03365-Pro
GGSGSSPAP	17,444	WMSV_P03359
GGGGGGGG	17,445	SFV3L_P27401_2mut
PAPGSSEAAAK	17,446	FOAMV_P14350-Pro_2mutA
GGGGSSPAP	17,447	FOAMV_P14350_2mut
GSSGGSPAP	17,448	MLVBM_Q7SVK7_3mut
GSSGGGGGS	17,449	GALV_P21414_3mutA
EAAAKEAAAKEAAAKEAAAKEAAAK	17,450	MMTVB_P03365
GSSGGGGGS	17,451	SFV1_P23074_2mut
GGGGSEAAAKGGGGS	17,452	SFV1_P23074
GGGEAAAKPAP	17,453	FFV_O93209
PAPGGGEAAAK	17,454	SFV1_P23074
GGSGGGEAAAK	17,455	PERV_Q4VFZ2_3mutA_WS
GSSGGG	17,456	MMTVB_P03365-Pro
EAAAKGSSGGS	17,457	FFV_O93209_2mut
GGGGG	17,458	SFV1_P23074_2mut
GGGPAP	17,459	SFV3L_P27401
GSSGGSEAAAK	17,460	FFV_O93209

SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,461	MMTVB_P03365-Pro
GSSGGGEAAAK	17,462	SFV1_P23074_2mutA
GSSGSSGSSGSSGSS	17,463	SFV3L_P27401_2mut
GGSEAAAKPAP	17,464	FLV_P10273
GGGGSGGGGS	17,465	FOAMV_P14350-Pro_2mutA
GSSEAAAKPAP	17,466	SFV3L_P27401
GGGGSEAAAKGGGGS	17,467	MMTVB_P03365-Pro
PAPGSSEAAAK	17,468	MLVF5_P26810_3mut
EAAAKGGSGGG	17,469	SFV3L_P27401
GGGPAPGGS	17,470	SFV3L_P27401
GSSEAAAKGGS	17,471	FOAMV_P14350_2mutA
EAAAKGGSGGG	17,472	HTL1L_P0C211
GSSGGSPAP	17,473	SFV3L_P27401_2mutA
PAPAP	17,474	FFV_O93209
PAPGGSGSS	17,475	MMTVB_P03365_WS
EAAAKGGGGGS	17,476	FOAMV_P14350_2mut
PAPEAAAKGGS	17,477	SFV3L_P27401_2mut
GSSEAAAKPAP	17,478	MMTVB_P03365-Pro
GGSGGS	17,479	PERV_Q4VFZ2_3mut
GSSEAAAKGGG	17,480	FFV_O93209-Pro_2mutA
EAAAK	17,481	HTL1L_P0C211
GSSPAP	17,482	MLVMS_P03355
EAAAKPAPGGG	17,483	FFV_O93209-Pro_2mut
GGGGSEAAAKGGGGS	17,484	SFV1_P23074-Pro_2mut
EAAAKGSSGGS	17,485	SFV3L_P27401
GSAGSAAGSGEF	17,486	FFV_O93209_2mutA
PAPEAAAKGGS	17,487	MMTVB_P03365_2mutB_WS
EAAAKEAAAKEAAAKEAAAKEAAAKEAAAK	17,488	MMTVB_P03365
GGS	17,489	MMTVB_P03365
GGSEAAAKPAP	17,490	SFV1_P23074
EAAAKGSSGGG	17,491	HTLV2_P03363_2mut
GGSEAAAKGGG	17,492	MMTVB_P03365_WS
GGSGGS	17,493	FFV_O93209-Pro
GSSEAAAKGGS	17,494	MMTVB_P03365-Pro
PAPAPAPAPAP	17,495	SFV1_P23074_2mutA
GGSEAAAKGGG	17,496	MMTVB_P03365_2mutB_WS
PAPAPAPAP	17,497	MMTVB_P03365_WS
GGGGSGGGGSGGGGSGGGGSGGGGS	17,498	HTL3P_Q4U0X6_2mut
PAPGGSEAAAK	17,499	SFV1_P23074-Pro_2mut
GGSGGGPAP	17,500	MMTVB_P03365
GSSGSSGSSGSSGSSGSS	17,501	MMTVB_P03365-Pro

[1675]

GGSEAAAKPAP	17,502	SFV1_P23074-Pro
GGGEAAAKGSS	17,503	SFV3L_P27401_2mutA
GGGPAPGGS	17,504	AVIRE_P03360
PAPGGG	17,505	MLVRD_P11227
GGSEAAAKGSS	17,506	SFV3L_P27401_2mut
GGGEAAAKGSS	17,507	FOAMV_P14350_2mut
GGGEAAAKGSS	17,508	SFV1_P23074-Pro
EAAAKEAAAKEAAAKEAAAK	17,509	MLVAV_P03356
EAAAKGGGPAP	17,510	JSRV_P31623_2mutB
EAAAKGGGGSS	17,511	FOAMV_P14350_2mut
EAAAKEAAAKEAAAKEAAAKEAAAK	17,512	SRV2_P51517
GSSGGGGGS	17,513	FFV_O93209
PAPAPAP	17,514	FOAMV_P14350_2mutA
GGSGGSGGSGGS	17,515	FOAMV_P14350
GGGEAAAK	17,516	MMTVB_P03365_WS
GGGGGS	17,517	SFV1_P23074_2mutA
GGSGGS	17,518	WMSV_P03359_3mut
EAAAKGGS	17,519	MMTVB_P03365-Pro
GGGGSS	17,520	BLVJ_P03361_2mut
PAPAP	17,521	MMTVB_P03365-Pro_2mut
PAPGGG	17,522	SMRVH_P03364
EAAAKGGGGSS	17,523	SFV3L_P27401
PAPAPAPAPAP	17,524	MMTVB_P03365
GGGPAP	17,525	MMTVB_P03365-Pro
GSSGGSGGG	17,526	MMTVB_P03365
EAAAKGGGPAP	17,527	FOAMV_P14350_2mutA
GSSGSSGSSGSS	17,528	SFV1_P23074
GGGGSGGGGS	17,529	SFV3L_P27401
GSSGGSGGG	17,530	MLVF5_P26810
GGGEAAAKPAP	17,531	MMTVB_P03365-Pro
PAPEAAAK	17,532	HTLV2_P03363_2mut
GSSGSSGSSGSS	17,533	FOAMV_P14350_2mut
GSSEAAAKPAP	17,534	MMTVB_P03365-Pro
PAPEAAAKGGG	17,535	HTL3P_Q4U0X6_2mut
GGSEAAAKGSS	17,536	MMTVB_P03365-Pro
EAAAKPAPGGS	17,537	MMTVB_P03365_2mut_WS
GSSGGSEAAAK	17,538	MLVF5_P26810_3mutA
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,539	MLVF5_P26810_3mut
EAAAKGGGGSS	17,540	MMTVB_P03365-Pro
GGGGGSGSS	17,541	HTL1A_P03362_2mutB
PAPAP	17,542	FFV_O93209-Pro_2mut

[1676]

GGGGGSPAP	17,543	HTL1C_P14078_2mut
GGGPAP	17,544	HTLV2_P03363_2mut
EAAAKGGGGSEAAAK	17,545	SFVCP_Q87040
GGSEAAAKGGG	17,546	FFV_O93209-Pro_2mutA
GSSPAPGGS	17,547	FOAMV_P14350-Pro_2mut
GGGGGGG	17,548	MMTVB_P03365-Pro
EAAAKGSS	17,549	SFV3L_P27401_2mutA
EAAAKGGGGSEAAAK	17,550	MMTVB_P03365-Pro
GGGGSEAAAKGGGGS	17,551	SFV1_P23074-Pro_2mutA
EAAAKGGGGSS	17,552	MMTVB_P03365
GGGEAAAKGGS	17,553	SFV1_P23074
PAPEAAAKGGG	17,554	MLVF5_P26810
GGGGSSGGS	17,555	MMTVB_P03365
GGSGSS	17,556	MMTVB_P03365
PAPAPAPAPAPAP	17,557	KORV_Q9TTC1
EAAAKGGG	17,558	SFV1_P23074-Pro_2mut
PAPAPAPAPAPAP	17,559	SRV2_P51517
GSSGSSGSSGSSGSS	17,560	FFV_O93209-Pro_2mutA
GGGGSS	17,561	FOAMV_P14350_2mut
PAPGGGEAAAK	17,562	MMTVB_P03365_WS
GGSGGGEAAAK	17,563	FFV_O93209-Pro_2mut
PAPAPAPAPAP	17,564	MMTVB_P03365_WS
GGGEAAAKGGS	17,565	MMTVB_P03365-Pro
GGGEAAAKGSS	17,566	MMTVB_P03365_2mutB
GSSPAPEAAAK	17,567	MMTVB_P03365_WS
EAAAKEAAAKEAAAKEAAAKEAAAK	17,568	SFV1_P23074-Pro_2mutA
PAPGGG	17,569	SFV3L_P27401
GSSEAAAKGGG	17,570	MMTVB_P03365_WS
GGGGSSEAAAK	17,571	FOAMV_P14350_2mut
PAPGSSGGS	17,572	SFV1_P23074-Pro_2mut
GSSGSSGSSGSSGSSGSS	17,573	SFV3L_P27401
EAAAKGSSGGG	17,574	MMTVB_P03365
PAPGGGGSS	17,575	WDSV_O92815_2mutA
GGSPAP	17,576	MMTVB_P03365-Pro
GGSGGSGGSGGSGGS	17,577	SFVCP_Q87040-Pro_2mut
PAPAPAPAP	17,578	MMTVB_P03365-Pro
GGGGG	17,579	HTL1A_P03362
GGSGGSGGSGGS	17,580	SFV1_P23074_2mutA
GSSGSSGSSGSSGSS	17,581	FOAMV_P14350-Pro_2mut
PAPGGSEAAAK	17,582	MMTVB_P03365_2mutB_WS
PAPAPAPAP	17,583	SFV1_P23074_2mut

[1677]

PAPGGGGSS	17,584	MMTVB_P03365
GGSGSS	17,585	SFV3L_P27401_2mut
EAAAKEAAAKEAAAKEAAAK	17,586	MMTVB_P03365_2mut
EAAAKGGSGGG	17,587	HTL3P_Q4U0X6_2mut
PAPGGGGSS	17,588	SFVCP_Q87040-Pro_2mutA
EAAAKGGGGGS	17,589	MLVAV_P03356
GGGGGS	17,590	FOAMV_P14350_2mut
GGGEAAAKGGS	17,591	FFV_O93209-Pro_2mutA
EAAAKPAPGGG	17,592	MMTVB_P03365_2mutB
GGSGGGPAP	17,593	FFV_O93209_2mut
GSSEAAAKPAP	17,594	MMTVB_P03365
PAPAPAPAPAPAP	17,595	SFV1_P23074_2mut
GGSPAPGGG	17,596	MMTVB_P03365-Pro
GGSGGGEAAAK	17,597	MMTVB_P03365
PAPAP	17,598	SFVCP_Q87040
GSSEAAAK	17,599	SFVCP_Q87040
GGGGSGGGGSGGGGS	17,600	MMTVB_P03365-Pro
GSSGSSGSS	17,601	SFV3L_P27401
EAAAKGGSGGG	17,602	MMTVB_P03365-Pro
GSSPAP	17,603	SFV1_P23074_2mut
GGGEAAAK	17,604	SFV1_P23074-Pro
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKEAAAKA	17,605	MMTVB_P03365-Pro
PAPGGS	17,606	HTL1C_P14078_2mut
PAPGSSGGS	17,607	SFV1_P23074_2mut
PAPEAAAK	17,608	MMTVB_P03365_WS
PAPAP	17,609	MMTVB_P03365-Pro
EAAAKGGS	17,610	HTL1A_P03362_2mut
GGGGSEAAAKGGGGS	17,611	HTL1C_P14078
EAAAKGSSGGS	17,612	FOAMV_P14350-Pro
PAPGGSGSS	17,613	MMTVB_P03365-Pro
PAPGGSEAAAK	17,614	SFV1_P23074_2mut
PAPGSSEAAAK	17,615	FFV_O93209-Pro_2mut
PAPGSSGGG	17,616	FOAMV_P14350-Pro_2mutA
GSSGGGEAAAK	17,617	AVIRE_P03360
GGGGGG	17,618	SMRVH_P03364_2mut
PAPEAAAKGGG	17,619	MMTVB_P03365-Pro
GGGEAAAKGGS	17,620	SFVCP_Q87040_2mutA
PAPAPAPAPAP	17,621	SRV2_P51517
GSSGSSGSSGSSGSSGSS	17,622	MMTVB_P03365
EAAAKGGGPAP	17,623	MLVAV_P03356
PAPAPAPAPAP	17,624	FOAMV_P14350-Pro_2mutA

[1678]

[1679]

PAPGGSEAAAK	17,625	FOAMV_P14350
GSSGGGPAP	17,626	HTL32_Q0R5R2_2mutB
GGGGGSPAP	17,627	HTL3P_Q4U0X6_2mutB
GSSGGSGGG	17,628	MMTVB_P03365-Pro
PAPAP	17,629	SFVCP_Q87040-Pro
GSSGGGPAP	17,630	MMTVB_P03365-Pro
GGSGSS	17,631	MMTVB_P03365-Pro_2mut
GGSPAPEAAAK	17,632	SFV1_P23074-Pro_2mut
EAAAKGGSGGG	17,633	SFV3L_P27401_2mut
GGGGSSEAAAK	17,634	MMTVB_P03365_WS
GGGGGSGSS	17,635	MMTVB_P03365_2mut
GGGGSSGGS	17,636	SFV1_P23074-Pro_2mutA
EAAAKGGGGSEAAAK	17,637	MMTVB_P03365_WS
PAPGGGEAAAK	17,638	SFV1_P23074-Pro
PAPEAAAKGGG	17,639	MMTVB_P03365
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,640	MMTVB_P03365
GSSGGSEAAAK	17,641	FOAMV_P14350-Pro_2mut
GGSPAP	17,642	MLVBM_Q7SVK7_3mut
GSSEAAAK	17,643	FOAMV_P14350
GSSEAAAK	17,644	MMTVB_P03365-Pro
EAAAKGSSGGS	17,645	HTL1A_P03362_2mut
GGGEAAAKPAP	17,646	FOAMV_P14350-Pro_2mut
EAAAKGGSPAP	17,647	FOAMV_P14350
GSSEAAAKPAP	17,648	MMTVB_P03365_WS
GSSGSSGSS	17,649	FOAMV_P14350_2mut
EAAAKEAAAKEAAAKEAAAK	17,650	MMTVB_P03365_WS
EAAAK	17,651	MMTVB_P03365
PAPGSS	17,652	BAEVM_P10272
PAPGGS	17,653	FFV_O93209-Pro_2mut
GGSGGS	17,654	SFV1_P23074-Pro_2mutA
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,655	HTLV2_P03363_2mut
GGSGGGEAAAK	17,656	MMTVB_P03365_WS
PAPGSSGGG	17,657	HTL1A_P03362
GGSGGS	17,658	SFV3L_P27401-Pro
GSSGSS	17,659	SFV1_P23074-Pro
PAPGGSEAAAK	17,660	MMTVB_P03365
GSAGSAAGSGEF	17,661	MMTVB_P03365-Pro
PAPGGG	17,662	FOAMV_P14350_2mut
EAAAKGGSGSS	17,663	MMTVB_P03365_WS
GSSGGGEAAAK	17,664	SFV3L_P27401-Pro
GGSGGGPAP	17,665	FOAMV_P14350-Pro_2mut

[1680]

PAPAPAPAPAPAP	17,666	WDSV_O92815
SGSETPGTSESATPES	17,667	SFVCP_Q87040-Pro_2mutA
GGSGGSGGS	17,668	SFV1_P23074
GGGGSS	17,669	SFVCP_Q87040_2mut
GGGGGSEAAAK	17,670	MMTVB_P03365
SGSETPGTSESATPES	17,671	MMTVB_P03365_WS
PAPAPAP	17,672	SFV3L_P27401
PAPEAAAKGSS	17,673	MMTVB_P03365_2mutB_WS
GSSGSSGSSGSSGSS	17,674	SRV2_P51517
GGGPAPGSS	17,675	HTL32_Q0R5R2_2mutB
GGSGGGGSS	17,676	MMTVB_P03365-Pro
SGSETPGTSESATPES	17,677	SRV2_P51517
EAAAKGSSGGS	17,678	MMTVB_P03365-Pro
GSSPAPEAAAK	17,679	MMTVB_P03365-Pro
GSSPAPEAAAK	17,680	SRV2_P51517
GGGGSSPAP	17,681	MMTVB_P03365-Pro
PAPGGGEAAAK	17,682	SFV1_P23074-Pro_2mutA
PAPEAAAKGGS	17,683	MMTVB_P03365
GSSGSSGSSGSSGSSGSS	17,684	FOAMV_P14350-Pro
GGSPAPGSS	17,685	SFV3L_P27401
GGGPAPGGS	17,686	SFV1_P23074-Pro_2mutA
GGGPAPGSS	17,687	MMTVB_P03365-Pro
EAAAKPAP	17,688	MLVBM_Q7SVK7
EAAAKEAAAKEAAAK	17,689	HTL1C_P14078
GSSGGSEAAAK	17,690	SRV2_P51517
PAPGGGGGS	17,691	SRV2_P51517
GGGEAAAK	17,692	FFV_O93209-Pro_2mut
EAAAKGGGPAP	17,693	HTL32_Q0R5R2
GGSGSSGGG	17,694	MMTVB_P03365
PAPEAAAKGSS	17,695	MMTVB_P03365-Pro
PAPGGGGGS	17,696	MMTVB_P03365-Pro
EAAAKGGGGGS	17,697	MMTVB_P03365_WS
GGGGGS	17,698	MMTVB_P03365-Pro
GGGGSGGGGSGGGGSGGGGSGGGGS	17,699	HTL1C_P14078
EAAAKGGSPAP	17,700	MMTVB_P03365
GGGGSSPAP	17,701	FFV_O93209-Pro_2mut
GGGGSSGGS	17,702	MMTVB_P03365-Pro
PAPGSSGGS	17,703	MMTVB_P03365-Pro
GGGGGS	17,704	SRV2_P51517
GGSGSSGGG	17,705	MMTVB_P03365
GSSGGSEAAAK	17,706	MMTVB_P03365-Pro

EAAAKEAAAKEAAAKEAAAK	17,707	GALV_P21414
GGSEAAAKGGG	17,708	MMTVB_P03365-Pro
SGGSSGGSSGSETPGTSESATPESSGGSSGGSS	17,709	MMTVB_P03365-Pro
GSSEAAAKGGS	17,710	MMTVB_P03365
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,711	HTL3P_Q4U0X6_2mutB
GGGEAAAK	17,712	MMTVB_P03365-Pro
PAPAPAPAP	17,713	MMTVB_P03365-Pro
PAPGSSGGG	17,714	MMTVB_P03365
GSSGSSGSSGSSGSS	17,715	GALV_P21414
GGSPAP	17,716	MMTVB_P03365_WS
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,717	MMTVB_P03365-Pro
PAPEAAAK	17,718	MMTVB_P03365-Pro
PAPGSSGGG	17,719	SFV1_P23074-Pro_2mutA
GGGGGSEAAAK	17,720	MMTVB_P03365_2mutB_WS
PAPAPAPAPAP	17,721	MMTVB_P03365-Pro
EAAAKGGSGSS	17,722	MMTVB_P03365-Pro
EAAAKEAAAKEAAAKEAAAK	17,723	MLVRD_P11227_3mut
PAPAPAPAP	17,724	FOAMV_P14350_2mutA
GGGPAPGSS	17,725	SFVCP_Q87040_2mut
PAPEAAAKGSS	17,726	SFVCP_Q87040_2mut
GGSPAPGGG	17,727	MMTVB_P03365-Pro
GGGGSGGGGSGGGGSGGGGS	17,728	MMTVB_P03365
EAAAKGGS	17,729	HTL3P_Q4U0X6_2mut
PAPGSSGGS	17,730	MMTVB_P03365_WS
GGGGSGGGGS	17,731	MMTVB_P03365
GGSGGS	17,732	FOAMV_P14350
EAAAKGGGGSEAAAK	17,733	SFVCP_Q87040-Pro_2mut
EAAAKEAAAKEAAAKEAAAK	17,734	MMTVB_P03365-Pro_2mutB
PAPGGGEAAAK	17,735	SFVCP_Q87040-Pro
GSSGSS	17,736	JSRV_P31623_2mutB
EAAAKGGGGGS	17,737	MMTVB_P03365_2mut_WS
GSSPAPEAAAK	17,738	MMTVB_P03365-Pro
GGGEAAAK	17,739	HTL1C_P14078
PAPEAAAKGSS	17,740	HTL32_Q0R5R2_2mutB
GGGGSSEAAAK	17,741	MMTVB_P03365-Pro
PAPGSSGGS	17,742	MMTVB_P03365-Pro
EAAAKGGGGGS	17,743	MMTVB_P03365
GGGGSGGGGSGGGGSGGGGS	17,744	MMTVB_P03365
EAAAKGGGGSS	17,745	HTL3P_Q4U0X6_2mut
GGGEAAAKGGS	17,746	SFVCP_Q87040-Pro
GGGGGSPAP	17,747	MMTVB_P03365-Pro_2mutB

[1681]

GGSGGGEAAAK	17,748	SFV3L_P27401-Pro
PAPGGGGGS	17,749	SFV3L_P27401-Pro
EAAAKGGGGSEAAAK	17,750	MMTVB_P03365
PAPEAAAKGSS	17,751	MMTVB_P03365-Pro
GGSEAAAKGGG	17,752	MMTVB_P03365-Pro
GGSGGSGGSGGSGGS	17,753	SMRVH_P03364_2mutB
GGSGGSGGSGGSGGS	17,754	HTL1L_P0C211_2mut
GGGGGG	17,755	WDSV_O92815
GGGGGSGSS	17,756	MMTVB_P03365-Pro
GGSEAAAKPAP	17,757	SFV3L_P27401-Pro_2mut
GGGPAPGSS	17,758	MMTVB_P03365_2mut_WS
GGGGGS	17,759	MMTVB_P03365_WS
GGSPAPEAAAK	17,760	MMTVB_P03365
PAPEAAAKGGS	17,761	HTL1A_P03362
EAAAKGGSGSS	17,762	MMTVB_P03365_2mut_WS
GGGPAPEAAAK	17,763	SFV3L_P27401-Pro_2mut
PAPGGGGSS	17,764	HTL32_Q0R5R2_2mut
GSSPAPGGG	17,765	HTL3P_Q4U0X6_2mut
GGGGSSGGS	17,766	BLVAU_P25059_2mut
EAAAKGGGGGS	17,767	HTL1L_P0C211
GGSEAAAKGSS	17,768	JSRV_P31623_2mutB
GSSGGG	17,769	JSRV_P31623
GGSGGSGGSGGS	17,770	MMTVB_P03365-Pro
EAAAKPAP	17,771	SFV1_P23074-Pro_2mutA
GGGGSSGGS	17,772	MMTVB_P03365_WS
GGSGGS	17,773	MMTVB_P03365_WS
EAAAKGGGGGS	17,774	MMTVB_P03365-Pro
GGGGSGGGGSGGGGSGGGGSGGGGSGGGGS	17,775	MMTVB_P03365
GGSGGSGGS	17,776	MMTVB_P03365
GGGGGSEAAAK	17,777	MLVBM_Q7SVK7
GGSGSSPAP	17,778	MMTVB_P03365_WS
EAAAKEAAAKEAAAK	17,779	JSRV_P31623
PAPEAAAKGGS	17,780	MMTVB_P03365-Pro
GGSGSSEAAAK	17,781	FOAMV_P14350
GGGGGSGSS	17,782	MMTVB_P03365-Pro_2mut
GGGPAPGGS	17,783	MMTVB_P03365
SGSETPGTSESATPES	17,784	SFVCP_Q87040_2mut
GSSPAPGGS	17,785	SFV1_P23074-Pro_2mutA
GSSGSSGSSGSSGSS	17,786	MMTVB_P03365
EAAAKGGGPAP	17,787	MMTVB_P03365
GSSGGG	17,788	MMTVB_P03365_2mut_WS

[1682]

GGGEAAAKPAP	17,789	MMTVB_P03365
PAPGGSGGG	17,790	MMTVB_P03365-Pro
GSSGGSGGG	17,791	WDSV_O92815_2mut
GGSGGG	17,792	HTL32_Q0R5R2_2mut
EAAAKGGSPAP	17,793	HTLV2_P03363_2mut
GGSPAPEAAAK	17,794	MMTVB_P03365-Pro
GSSGGSEAAAK	17,795	MMTVB_P03365_2mut
GSAGSAAGSGEF	17,796	MMTVB_P03365_WS
PAPGGSGSS	17,797	FFV_O93209
GGSEAAAKGGG	17,798	MMTVB_P03365
GGSPAPGSS	17,799	MMTVB_P03365-Pro
GSSGGSGGG	17,800	SFV3L_P27401
PAPEAAAKGGG	17,801	HTL1A_P03362_2mutB
GGGEAAAKPAP	17,802	MMTVB_P03365-Pro
GGSEAAAK	17,803	HTL32_Q0R5R2_2mutB
GGGEAAAKGSS	17,804	MPMV_P07572
GGGGGSEAAAK	17,805	MMTVB_P03365-Pro
PAPAPAPAPAP	17,806	SFVCP_Q87040-Pro_2mutA
PAPAPAPAPAP	17,807	HTL1L_P0C211_2mut
GGGGSSGGS	17,808	HTL3P_Q4U0X6
PAPGGSEAAAK	17,809	MMTVB_P03365_2mut_WS
PAPAPAPAPAP	17,810	HTL1A_P03362
EAAAKPAPGGG	17,811	MMTVB_P03365_2mut_WS
GGSEAAAK	17,812	MMTVB_P03365_2mut_WS
GGGEAAAKGSS	17,813	SFV1_P23074-Pro_2mutA
GGSPAPGSS	17,814	MMTVB_P03365-Pro
GGSEAAAKPAP	17,815	MLVBM_Q7SVK7
PAPEAAAKGGG	17,816	MMTVB_P03365_2mut_WS
GSSEAAAKPAP	17,817	MMTVB_P03365-Pro_2mutB
GGGGSEAAAKGGGGS	17,818	MMTVB_P03365-Pro_2mut
GSSEAAAKGGS	17,819	MMTVB_P03365-Pro_2mutB
GSSGSSGSSGSSGSS	17,820	SRV2_P51517_2mutB
GGGGGSPAP	17,821	HTL1L_P0C211_2mut
GGSEAAAK	17,822	MMTVB_P03365
GSSPAPEAAAK	17,823	SMRVH_P03364_2mutB
GGGPAPGGS	17,824	HTL1C_P14078_2mut
GGSPAPEAAAK	17,825	MMTVB_P03365_WS
GGSEAAAKPAP	17,826	HTL1A_P03362_2mut
PAPAPAPAP	17,827	HTLV2_P03363_2mut
GSSPAPGGG	17,828	MMTVB_P03365
GSSGSSGSSGSS	17,829	MMTVB_P03365-Pro

[1683]

GGSEAAAKGSS	17,830	MMTVB_P03365_WS
GGSGSSGGG	17,831	MMTVB_P03365_2mutB
GSSGSSGSSGSSGSSGSS	17,832	JSRV_P31623_2mutB
GGSEAAAKPAP	17,833	MMTVB_P03365-Pro
GSSGGSGGG	17,834	HTLV2_P03363_2mut
AEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKA	17,835	WDSV_O92815_2mut
GGSPAPEAAAK	17,836	MMTVB_P03365
GGGGSSEAAAK	17,837	MMTVB_P03365
GGSGGGEAAAK	17,838	SFV1_P23074-Pro_2mutA
GGGGSEAAAKGGGGS	17,839	WDSV_O92815_2mut
GGSGSSEAAAK	17,840	MMTVB_P03365_2mutB_WS
GGSEAAAKPAP	17,841	MMTVB_P03365_WS
GSSGGGEAAAK	17,842	SFVCP_Q87040-Pro
GSSGGS	17,843	SFVCP_Q87040-Pro_2mut
GGSEAAAKPAP	17,844	SFVCP_Q87040_2mut
GSSGGSEAAAK	17,845	SFVCP_Q87040_2mut
GSSPAPEAAAK	17,846	SRV2_P51517_2mutB
GGSGGSGGSGGSGGSGGS	17,847	BLVAU_P25059
GSSGSSGSSGSSGSS	17,848	HTL1C_P14078_2mut
EAAAKGGGGSS	17,849	MMTVB_P03365_2mutB
GGGEAAAKGSS	17,850	SFVCP_Q87040-Pro

[1684]

[1685] **실시예 6: 예시적인 드라이버 및 트랜스 주형의 생성**

[1686] 이 실시예는 실시예 1 내지 5에 기재된 3, 4, 또는 5개 구성요소 RNA-기반 유전자 변형 시스템에서 사용하기 위한 예시적인 드라이버 및 트랜스 주형을 설명한다.

[1687] 1) 유전자 변형 폴리펩티드:

[1688] 다음을 함유하는 일련의 216개의 예시적인 유전자 변형 폴리펩티드를 생성하였다:

[1689] 1. 하나의 엔도뉴클레아제 도메인이 비활성화된 Cas-뉴클레아제(이 실시예에서는, Spy N863A Cas9),

[1690] 2. 2개의 역전사효소(RT)(PLV10993 또는 PLV10990/RNAIVT338에 의해 인코딩되는 유전자 변형 폴리펩티드의 RT) 중 하나, 및

[1691] 3. 1, 2, 또는 4개 RBP 반복부(이 실시예에서는, 표 31에 제공된 바와 같은 1, 2, 또는 4개의 MCPv2 도메인)를 함유하는 RNA 결합 단백질(RBP)(각각의 RBP 반복부는 표 10의 4개 링커 펩티드 중 하나에 의해 다른 RBP 반복부에 연결됨).

[1692] 이들 3개 도메인을, 예를 들어 표 10에서 확인되는 바와 같은, 펩티드 링커를 사용하여 연결하였다.

[1693] PLV10990/RNAIVT338에 의해 인코딩된 유전자 변형 폴리펩티드의 RT의 서열 및 PLV10993에 의해 인코딩된 유전자 변형 폴리펩티드의 RT의 서열은 하기에 제공되어 있다:

[1694] PLV10990/RNAIVT338 RT:

TAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPIT
LEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNP
YTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTL
FNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLC
QEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATR
PGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPV
AYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARIT
QYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFT
DGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAF
ATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNR
RADEVAREVAIRPLSTQATIS (서열번호 16,706)

[1695]

[1696] PLV10993 RT:

TLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYP
MSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVP
NPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKN
SPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKA
QICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAP
LYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGP
WRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDR
WLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPL
PDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKK
LNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHS
AEARGNRMADQAARKAAITETPDTSTLL (서열번호 16,707)

[1697]

[1698] 일련의 216개의 예시적인 유전자 변형 폴리펩티드는 실시예 1 및 2에 제시된 구성의 서브세트를 포함하였다:

[1699] · 구성 1: Cas9 - 링커 $(RBP)_n$ - 링커 - RT

[1700] · 구성 2: RT - 링커 - $(RBP)_n$ - 링커 - Cas9

[1701] · 구성 3: Cas9 - 링커 - RT - 링커 - $(RBP)_n$

[1702] 예시적인 유전자 변형 폴리펩티드의 서열은 하기에 제공되어 있다:

[1703] *표 S1. 구성 1(즉, Cas9 - 링커 (RBP)n - 링커 - RT)에 대한 예시적인 유전자 변형 서열*

식별자	전체 아미노산 서열	서열번호
...11941	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAK SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAK ALEAEAAAKEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRK FRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLT WTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQA ILQIPVPKTKRQVREFLGKAGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKR PVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHH CLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIY RERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARKVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,C
PL11942	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	

[1704]

[1705]

	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGSETPGTSESATP ESSGGSSGGSSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPW NTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTL FNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKAGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAA GWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQTQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDL TDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIK NAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARERVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,003
PL11943	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSTAPL EEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYR MVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLD HPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAE LAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLT REASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQTQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTD GSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPK RVAVMHCKGHQKDDAPTSTGNRRADEVARERVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,0
PL11944	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	

[1706]

	NPINASGVDAKAILSARLSKSRIRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	16,005
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG	
	LEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDL	
	REVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSL	
	LQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLY	
	AATRPGNDPLVWGEKEEEAFQSLKIALTQPPALALPSLDKPFQIFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKL	
	TFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLPETDDTLPHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRD	
	GKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVM	
	HCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11945	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	

	AEAAAKEAAAKEAAAKEAAAKALEAEAAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTAL PVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPE SQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEE VTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFV EETSGAAKGVLTQAIGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFK QTAALNPATLLPETDDTLPHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKS VNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGS EFESPKKKAKVE	16,006
PL11946	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWI5 SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY SGGSSGGSSGSETPGTSESATPESSGGSSGGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKR SLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEE GESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSL SNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQ ALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPET DDTLPHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATL HVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,0
PL11947	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	

[1707]

[1708]

	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	GSSGSSGSSGSSGSSGSSTAPLEEEYRLFLEAPIQNVTLLEQVKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPW	
	VHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEFGESGQLTWTRLPQGF	
	KNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAAQTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKT	
	KRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKR	
	LDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSL	
	TSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLT	
	AGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARELVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11948	MPAAKRVKLDGSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,008
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	GSSGGGEAAAKTAPLEEEYRLFLEAPIQNVTLLEQVKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPW	
	NTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEFGESGQLTWTRLPQGFKNSPTL	

[1709]

	FNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAA	
	GWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDL	
	TDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIK	
	NAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11949	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,009
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	GGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIF	
	ATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTI	
	KVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKA	
	TAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGT	
	SEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQG	
	FRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIP	
	GFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAA	
	AALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEAT	
	LFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAV	
	WLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11950	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,0
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	

[1710]

	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	GGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIF	
	ATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTI	
	KVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGSETPGTSESATPESSGGSSGGSSTAPLEEEYRLFLEAPI	
	QNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVN	
	KRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYV	
	DDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAAT	
	RPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFG	
	QDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKR	
	YAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCK	
	GHQKDDAPTSTGNRRADEVARREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11951	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,011
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG	

	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	GGGSGGGSGGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIF	
	ATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTI	
	KVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGGSSGGSSTAPLEEEYRLFLEAPIQNVTLLEQIWKREIPK	
	VWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTL	
	LSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSA	
	TRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEE	
	AFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSP	
	PDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWA	
	EPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRR	
	ADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11952	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,012
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYTV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNS8SQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNS8SQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	GGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIF	
	ATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTI	
	KVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGGEAAAKTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEIN	
	PPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLSLLPPD	
	RIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLM	
	TLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKL	
	ALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLT	
	NARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTS	

[1711]

	AQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAR EVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
11953	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSDGGSSGGGSSGSETPGTSESATPESSGGGSSGGSSMASNIFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKALEAEAAAKEAAAKE AAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWN TPLLPVRKSGTSEYRMVQDLREVNKRVVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLF NEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREF LGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11954	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	

[1712]

	EKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGGSSGSETPGTSESATPESSGGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGGSSGSETPGTSESATPESSGGGSSGGSSTAPL EEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYR MVQDLREVNKRVETIHPTVPNPYTLSLLPPDRIMYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLD HPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAE LAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLT REASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTD GSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPK RVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,015
PL11955	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGGSSGSETPGTSESATPESSGGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSGSSGSGSSGSSTAPLEEEYRLFLEAPIQNVT LLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVET IHPTVPNPYTLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIA ADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGND PLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITS SHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAA VVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKD DAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,0
PL11956	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	

[1713]

[1714]

	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQVPITLEAMR REIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVP NPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQA ACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWG EKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLES LLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDS VIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTST GNRRADEVARREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,017
PL11957	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFAN GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAA AKEAAAKALEAEAAAKEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQVPITLEAMR SLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEE	

	GESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSL	
	SNSRTQAILQWPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQ	
	ALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPET	
	DDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATL	
	HVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11958	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,018
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDSGGSSGGSSGSSETPGTSESATPESSGGGSSGGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR	
	QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFAN	
	GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGGSSETPG	
	TSESATPESSGGGSSGGGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVVRVRQYPITLEAKRSLRETIRKFRAAGILRP	
	VHSPWNTPLLPVRKSGTSEYRMVVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGF	
	KNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKT	
	KRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKR	
	LDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSL	
	TSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLT	
	AGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11959	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,0
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	

[1715]

[1716]

DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
QSITGLYETRIDLSQLGGDSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR	
QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGSSGSSGSSGSS	
GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGSSGSSGSSGSS	
GSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRK	
SGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDL	
QGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLF	
FIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQAPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAAL	
AAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLSTSTRPDLTDQPLAQAEA	
EATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALL	
TAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,020
DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
QSITGLYETRIDLSQLGGDSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR	
QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFAN	
GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGSSGGGEAAAKTAP	
LEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEY	
RMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRL	
DHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGF	
AELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQAPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAAL	
LTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLSTSTRPDLTDQPLAQAEATLFT	

11961	DGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWL PKRVAVMHCKGHQKDDAPTSTGNRRADEVARE VAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,021
PL11962	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSKGSSGGSSGGSSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSGSMASNFTQFVLVDNGGTGDVTVAPSNFAN GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGGSGGSGGSSG GGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ GLLKDGNPIPSAIAANSGIYGGSGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLN MELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPI QNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVN KRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYV DDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAAT RPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAAALLTREASKLTFG QDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKR YAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCK GHQKDDAPTSTGNRRADEVARE VAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,0

[1717]

[1718]

	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDSGGSGGSGGSEGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR	
	QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFAN	
	GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSG	
	GGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ	
	GLLKDGNPIPSAIAANSGIYGGSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLN	
	MELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPK	
	VWAEINPPGLASTQAPIHVQLLSTALPVRVRQVPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTL	
	LSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSA	
	TRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEE	
	AFQSLKIALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEIISSHNLESLLRSP	
	PDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWA	
	EPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRR	
	ADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL111963	MPAAKRVKLDGSGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,023
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDSGGSGGSGGSEGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR	
	QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFAN	
	GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSG	
	GGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ	
	GLLKDGNPIPSAIAANSGIYGGSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLN	

[1719]

	MELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQ APIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDL KDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRV SGKKAQILCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEEAFQSLKLALTQPPALA LPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQV LLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALT KALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQ ATISKRTADGSEFESPKKKAKVE	16,024
PL11964	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFAN GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSG GGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQG GLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLN MELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQL LSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCI PLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQ LCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEEAFQSLKLALTQPPALALPSLDKP FQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPR VRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWS KDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRT ADGSEFESPKKKAKVE	16,0
PL11965	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	

[1720]

	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGSGSGSGSGSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV	
	PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKATAPL	
	EEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRM	
	MVQDLREVNKRVETIHPTVPNPYTLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLD	
	HPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAE	
	LAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPPRCLRAIAAAALLT	
	REASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVREFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTD	
	GSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGAIYRERGWLTAGGKAIKNAPEILALLTAVWLPK	
	RVAVMHCKGHQKDDAPTSTGNRRADEVARERVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL111966	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,026
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGSGSGSGSGSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV	
	PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSGSGSETPGTSESATPESSGGSGSGGSSTAPLEEEYRLFLEAPIQNVT	
	LLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMMVQDLREVNKRVET	

[1721]

	IHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIA ADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGND PLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITS SHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAA VVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKD DAPTSTGNRRADEVARKEVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,027
PL11967	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQDSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCEIIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAE INPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPP DRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLL MTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSL KLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKW LTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIG TSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDDAPTSTGNRRADEVA REVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,0
PL11968	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	

[1722]

	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGSGSSGSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGGSGGSG

11970	TSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVPAAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLT AGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,030
	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGSGSGSGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGSGSETPGTSESATPESSGGSSG KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGSGSETPGTSESATPESSGGSSG GSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRK SGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDL QGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRL FIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAI AAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQA EATLFTDGSSYIRDGKRYAGAAVVTLDSVPAAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALL TAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,03
PL11971	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	

[1723]

[1724]

	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGSSGSSGSSGSSTAPLEEEYRLFLE APIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLRE VNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQ YVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYA ATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLT FGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRD GKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVM HCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,032
PL11972	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNRERIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGSSGSSGSSGSSTAPLEEEYRLFLEAPIQNV TLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVE TIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLI AADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGN DPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEI TSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGA AVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQK DDAPTSTGNRRADEVARFVAIRPLSTQATISKRTADGSEFESPKKKAKVE	

[1725]

n=11973	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,033
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV	
	PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY	
	KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAAN	
	SGIYGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELI	
	VKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPK	
	VWAEINPPGLASTQAPIHVQLLSTALPVRVRQVPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTL	
	LSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSA	
	TRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEE	
	AFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSP	
	PDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWA	
	EPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRR	
	ADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11974	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKV	16,0
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	

[1726]

	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGSSGGSSGGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSSGGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSSGGSGSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAAN SGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELI VKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGGSSGSETPGTSESATPESSGGSSGGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQA PIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLK DAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVS GKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALAL PSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVL VKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSGSSGETPGTSESATPESSGGSSGGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQA LLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAEIIALT KALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQ ATISKRTADGSEFESPKKKAKVE	16,035
PL11975	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGSSGGSSGGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSSGGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSSGGSGSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAAN SGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELI VKAMQGLLKDGNPIPSAIAANSGIYGSSGGSSGGSSGGSSGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVR QYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFA	

	FEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLG FKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSG AAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAAL NPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPGTSAQKAELIALTKALEWSKDKSVNIYT DSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESP KKKAKVE	16,036
PL11976	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGGSSGGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQDSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYV KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAAN SGIYGGSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELI VKAMQGLLKDGNPIPSAIAANSGIYGSSGGGGEAAAKTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITL EAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWA DAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHK GSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKG VLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATL LPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPGTSAQKAELIALTKALEWSKDKSVNIYTDSRYA FATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARFVAIRPLSTQATISKRTADGSEFESPKKKAK VE	16,0
PL11977	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	

[1727]

[1728]

	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKAARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKATAPLEEEYRLF LEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDL REVNKRVETIHPTVPNPYTLLSLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQA LQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLY AATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKL TFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRD GKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEMSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVM HCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,038
P11978	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKAARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSTAPLEEEYRLFLEAPIQNVTLLEQWK REIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVP NPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQA ACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWG	

	EKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPVWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLES LLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDS VIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTST GNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,039
J11979	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKAARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSGSSGSGSSGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAAEINPPGL ASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWY SVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAE LGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQ PPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARI TQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQK AELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAI RPLSTQATISKRTADGSEFESPKKKAKVE	
PL11980	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKAARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	16,0

[1729]

[1730]

	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKVTIKVEVPKGAW RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQA PIHVQLLSTALPVRVRIQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLK DAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVS GKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALAL PSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVL LLDPPRVREKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALT KALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQ ATISKRTADGSEFESPKKKAKVE	16,041
PL11981	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKVTIKVEVPKGAW RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAK EAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRIQYPITLEAKRSLRETIRKFRAAGILRPVHSPW NTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTL FNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAA GWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVREKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDL TDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIK NAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	

[1731]

PL11982	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGSSGGSETPGTSESATPESSGGSSGGGSSTAP LEEEYRLFLEAPIQNVTLLEQWKREIPKVVAENPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEY RMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRL DHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGF AELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAAL LTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLTSTRPDLTDQPLAQAEATLFT DGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWL PKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,042
PL11983	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	16,042

[1732]

	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGGSSGGSSGSSTAPLEEEYRLFLEAPIQNV TLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVE TIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLI AADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGN DPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAAALLTREASKLTFGQDIEI TSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRIFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGA AVVTLDSVNVAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQK DDAPTSTGNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,044
PL11984	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYTV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGGGEAAAKTAPLEEEYRLFLEAPIQNVTLLEQW KREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVP NPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQA ACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWG EKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAAALLTREASKLTFGQDIEITSSHNLES LLRSPPDKWLTNARITQYQVLLLDPPRVRIFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDS VIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTST GNRRADEVAREVAIRPLSTQATISKRTADGSEFESPKKKAKVE	16,0
PL11985	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	

[1733]

	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQG	
	RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV	
	RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSSGGGSGGGSMASNFTQFVLVDNGGT	
	GDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGG	
	SGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ	
	GLLKDGNPIPSAIAANSGIYAEAAAKEAAAAKEAAAAKEAAAAKEAAAAKEAAAAKEAAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEIN	
	PPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPD	
	RIMYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLM	
	TLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKL	
	ALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLT	
	NARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTS	
	AQKAEIIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAR	
	EVAIRPLSTQATISKRTADGSEFESPKKKAKVE	
PL11986	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,044
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	

	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW	
	RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV	
	RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSGGGSGGGSMASNFTQFVLVDNGGT	
	GDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGG	
	SGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ	
	GLLKDGNPIPSAIAANSGIYSGGSSGGSGGSSGETPGTSESATPESSGGSSGGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQL	
	LSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCI	
	PLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQ	
	LCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKP	
	FQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPR	
	VRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPKGTSAQKAELIALTKALEWS	
	KDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARE VAIRPLSTQATISKRT	
	ADGSEFESPKKKAKVE	
PL11987	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,047
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW	
	RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV	
	RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSGGGSGGGSMASNFTQFVLVDNGGT	
	GDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGG	
	SGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ	
	GLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITL	
	EAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWA	
	DAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHK	
	GSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKG	
	VLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATL	

[1734]

	LPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYA FATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARREVAIRPLSTQATISKRTADGSEFESPKKKAK VE	16,048
11988	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	

[1735]

[1736]

	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAK	
	ALEAEAAAKEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHI	
	QRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGI	
	SGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTP	
	EARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLS	
	KLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPEEGLQHNC	
	LLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTD	
	SRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAK	
	VE	
PL11990	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,050
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGSETPGTSESATP	
	ESSGGSSGGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQS	
	PWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTP	
	KTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLS	
	KKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPEEGLQHNC	

	LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEI	
	YRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
11991	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,051
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDILSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSGSGSSTLNIE	
	DEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN	
	DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQIWTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLA	
	DFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFC	
	RLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLR	
	MVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTD	
	QPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIK	
	NKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL11992	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,05:
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	

[1737]

[1738]

	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKTLNIEDEHRL	
	HETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPV	
	QDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQ	
	HPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIP	
	GFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAA	
	IAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPD	
	ADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEIL	
	ALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL11993	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,053
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	AEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKAT	
	STPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCL	
	RLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQ	
	ICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDL	
	TKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQ	
	ALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAE	
	LIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPD	
	TSTLLKRTADGSEFESPKKKAKVE	
PL11994	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,0
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	

[1739]

	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAK

[1740]

	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG
	GTGDVTVAPSNFANGIAEWISSNSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY
	GSSGSSGSSGSSGSSGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGI
	LVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWT
	RLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETV
	MGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWR
	RPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEE
	GLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFAT
	AHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE
PL11996	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG
	GTGDVTVAPSNFANGIAEWISSNSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY
	GSSGGGEAAAKTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQS
	PWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTP
	KTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLS
	KKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNNC
	LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEI
	YRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE
PL11997	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEI
	FSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
	DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNL

[1741]

	LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKF	
	IKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI	
	TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED	
	YFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR	
	KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENI	
	VIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSI	
	DNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE	
	NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNI	
	MNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS	
	PTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVN	
	FLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID	
	RKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAAKEAAAKEAAAKAMASNFTQFVLVD	
	NGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNP	
	IPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYL	
	NMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAE	
	WISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMAS	
	NFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAM	
	QGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEALEAEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLS	
	DFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK	
	RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQ	
	HPDLILQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLVKEGQRWLTEARKETVMGQPTPKTPRQLREFLGK	
	AGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKL	
	DPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEG	
	LQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYT	
	DSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGS	
	EFESPKKKAKVE	
PL11998	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,0
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	

[1742]

	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	GGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIF	
	ATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTI	
	KVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSGGSGGSSETPGTSESATPESSGGSSGGSSTLNIEDEHRLHETSKE	
	PDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLRE	
	VNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLIL	
	LQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEM	
	AAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTK	
	DAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT	
	WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKA	
	LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
P11999	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,059
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWIS	
	SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	GGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIF	
	ATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTI	

[1743]

	KVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSGSSGSSGSSGSSGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFP QAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNP YNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELD CQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNW GPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAP HAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKA GAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKG HSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,060
PL12000	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTI SNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSGSMASNFTQFVLVDNG GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY GGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIF ATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTI KVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSGSGGEAAAKTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAE TGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSG LPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQ KAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEA LVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVT TETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEA RGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,0
PL12001	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	

[1744]

	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	16,062
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR	
	QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKALEAEAAAAKEAAAKE	
	AAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSP	
	WNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFK	
	NSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSK	
	KLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCL	
	DILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIY	
	RRRGWLTSEGKEIKNKDEIIALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAAVE	
PL12002	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR	
	QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGSETPGTSESATPESSGGSSGGSSTLNIE	
	DEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN	

	DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLA	
	DFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFC	
	RLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLR	
	MVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTD	
	QPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIK	
	NKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL12003	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,063
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDSKGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVR	
	QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSGSGSGSSGSSTLNIEDEHRLHETSKEPDVS	
	LGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKR	
	VEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYV	
	DDLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLY	
	PLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKL	
	TMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDG	
	SSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKR	
	LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL12004	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,0
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	

[1745]

[1746]

	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	16,065
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDSGGSSGGGSSGSETPGTSESATPESSGGGSSGGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR	
	QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKTLNIEDEHRLHETSKEPDVSLGSTW	
	LSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP	
	TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLA	
	ATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKP	
	GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQP	
	QPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA	
	CPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL12005	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDSGGSSGGGSSGSETPGTSESATPESSGGGSSGGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR	
	QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFAN	
	GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAA	
	AKEAAAKALEAEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQE	
	ARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFE	
	WRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLL	
	KEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQ	
	GYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF	
	GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA	

12006	EGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADG SEFESPKKKAKVE	16,065
	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	16,065
PL12007	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	16,065

[1747]

[1748]

	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFAN GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGGSSGGSSGSS GSSTLNIEDEHRILHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLP VKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNE ALHRDLADFRIQHPDLILQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREF LGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAG WPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGT RPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEI EGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,068
PL112008	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFAN GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFAN EDEHRILHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGT NDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDL ADFRIQHPDLILQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGF CRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCL RMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLT DQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEI KNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	

PL12009	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,069
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR	
	QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFAN	
	GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGGSGG	
	GGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ	
	GLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLN	
	MELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKATLNIEDEHRLHETSKE	
	PDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLRE	
	VNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLIL	
	LQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEM	
	AAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTK	
	DAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHT	
	WYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKA	
	LFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL12010	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,0
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	

[1749]

[1750]

	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFAN GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGGSGGSG GGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ GLLKDGNPIPSAIAANSGIYGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLN MELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGGSSETPGTSESATPESSGGSSGGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFP QAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNP YNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELD CQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNW GPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAP HAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKA GAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALKALFLPKRLSIIHCPGHQKG HSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,071
PL12011	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFAN GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ GGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ GLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLN MELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGSSGGSSGGSSGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAV RQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWY	

	TVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGN LGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQA LLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDR WLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAK ALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMAD QAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,072
PL12012	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVR QSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFAN GIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSG GGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ GLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLN MELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGGEAAAKTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLII PLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLK DAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRA SAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAP ALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNA RMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAG TSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARK AAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,0
PL12013	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	

[1751]

[1752]

	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGSGSGSGSGSGSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV	
	PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKATLNIE	
	DEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN	
	DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLA	
	DFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFC	
	RLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLR	
	MVAAAAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTD	
	QPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIK	
	NKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL12014	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,074
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGSGSGSGSGSGSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV	
	PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGGSSGGGSSGGSETPGTSESATPESSGGGSSGGGSSTLNIEDEHRLHETSKEPDVS	
	LGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKR	
	VEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYV	
	DDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLY	

[1753]

	PLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKL TMGQPLVILAPHAVEALVKQPPDRWISNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDG SSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKR LSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,075
12015	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCEIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAW AETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLL SGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQ GTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPD QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAV EALVKQPPDRWISNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAA VTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSA EARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,0
PL12016	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	

[1754]

	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRIDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGM GLAVRQAPLIIPLKATSTPVSIKQYPMSQEAARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSH QWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQ EIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEV IWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRM MADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,077
PL12017	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKAARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRIDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGSSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKALEAEAAA KEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGI LVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWT RLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETV MGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWR RPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEE GLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFAT AHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	

[1755]

PL12018	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKAARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGSETPGTSESATPESSGGSSG GSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLP VKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNE ALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREF LGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAG WPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGT RPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTS EGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,078
PL12019	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKAARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	16,07

[1756]

	QSITGLYETRIDLSQLGGDGSSGGSSGSSGSSGSSGSSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSTLNIEDEHRLHET KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSTLNIEDEHRLHET SKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDL REVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDL ILLQYVDDLLIAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAE MAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVL TKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADH TWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLK ALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,080
PL12020	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSGSSTLNIEDEHRLHETSKEPDV KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSGSSTLNIEDEHRLHETSKEPDV SLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQY VDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAP LYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAG KLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTD GSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPK RLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,0
PL12021	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	

[1757]

	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELI	
	PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY	
	KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAAN	
	SGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELI	
	VKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQ	
	AWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPY	
	NLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDC	
	QQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWG	
	PDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPH	
	AVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAG	
	AAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGH	
	SAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL12022	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,08
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	

[1758]

	QSITGLYETRIDLSQLGGDGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSGGSGGGSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAAN SGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELI VKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSGGSGGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVR QAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQAL LTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRW LSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKA LPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQ AARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,083
PL12023	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSGGSGGGSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAAN SGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELI VKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPT SQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQ VKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLD	

	TDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALT QALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLL KRTADGSEFESPKKKAKVE	16,084
12024	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGGSSGGSSGGSSGGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEV PKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAY KVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSGGSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAAN SGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELI VKAMQGLLKDGNPIPSAIAANSGIYGSSGGGGEAAAKTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPM SQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFA FEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKVKYLGY LLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEK QGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQ FGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKM AEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTAD GSEFESPKKKAKVE	16,0
PL12025	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	

[1759]

[1760]

	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	16,086
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW	
	RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATLNIEDEHRL	
	HETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPV	
	QDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQ	
	HPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIP	
	GFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMG	
	IAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPD	
	ADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEIL	
	ALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL12026	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW	
	RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSTLNIEDEHRLHETSKEPDVSLGSTW	
	LSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHP	
	TVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLA	
	ATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKP	
	GTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMG	
	QPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDPDADHTWYTDGSSLL	

12027	QEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIH CPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,087
	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKAARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSGGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSGSGSGSGSGSGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGM GLAVRQAPLIIPLKATSTPVSIKQYPMSQEAARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSH QWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQ TLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQ EIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQ PPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEV IWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,088
PL12028	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKAARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	

[1761]

[1762]

	EKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW	
	RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGGEAAAKTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVR	
	QAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYT	
	VLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNL	
	GYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQQKAYQEIKQAL	
	LTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRW	
	LSNARMTHYQALLLDTDRVQFGPVVALNPATLLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKA	
	LPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQ	
	AARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,089
PL12029	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW	
	RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV	
	RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAK	
	EAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWILSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQS	
	PWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTP	
	KTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLS	
	KKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNNC	
	LDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEI	
	YRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITEIPDTSTLLKRTADGSEFESPKKKAKVE	16,0
PL12030	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKV	
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	

[1763]

	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW	
	RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV	
	RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGSGSSGSGSSGSGSSTLNIEDEHRLHETSKEPDV	
	EDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGT	
	KNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
	ADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGF	
	RMVAAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLT	
	DQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEI	
	KNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	16,091
PL12031	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW	
	RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV	
	RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV	
	EDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGT	
	KNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
	NDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDL	
	ADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGF	
	CRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCL	
	RMVAAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLT	
	DQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEI	
	KNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	

[1764]

	SLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK	
	RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQY	
	VDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAP	
	LYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAG	
	KLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTD	
	GSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPK	
	RLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL12032	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,092
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
	LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
	HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
	DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
	TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
	ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
	SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
	EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQDSSAQNRKYTIKVEVPKGAW	
	RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV	
	RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKTLNIEDEHRLHETSKEPDVSLGST	
	WLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDI	
	HPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLL	
	LAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTK	
	PGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTM	
	GQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSL	
	LQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSII	
	HCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE	
PL12033	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,0
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
	DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
	SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
	LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	

LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE		
HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV		
DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI		
TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK		
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD		
SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY		
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH		
QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW		
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV		
RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSGGGSMASNFTQFVLVDNGGT		
GDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGG		
SGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ		
GLLKDGNPIPSAIAANSGIYAEAAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAET		
GGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGL		
PPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTR		
ALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQK		
AYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEAL		
VKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTT		
ETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEAR		
GNRMADQAARKAAITETPDTSTLLKRTADGSEFESPKKKAKVE		
MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV		16,094
DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE		
NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS		
DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG		
SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL		
LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV		
LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE		
HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV		
DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI		
TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK		
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD		
SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY		
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH		
QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW		
RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV		
RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSGGGSMASNFTQFVLVDNGGT	PL12034	

[1765]

[1766]

	GDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGG SGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ GLLKDGNPIPSAIAANSGIYSGGSSGGSSGGSSETPGTSESATPESSGGSSGGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIP LKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKD AFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRAS AKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGT SAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKA AITETPDTSTLLKRTADGSEFESPKKKAKVE	16,095
PL12035	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAW RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSMASNFTQFVLVDNGGT GDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGG SGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ GLLKDGNPIPSAIAANSGIYGSSGGSSGGSSGGSSGGSSTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIPLKATSTPVSIKQYPM SQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFA FEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGY LLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEK QGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAVAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQ FGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKM AEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTAD GSEFESPKKKAKVE	

n	12036		16,096
		MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	
		DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
		NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLS	
		DILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNG	
		SIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSL	
		LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV	
		LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHE	
		HIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYV	
		DQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQI	
		TKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK	
		ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD	
		SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY	
		EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
		QSITGLYETRIDLSQLGGDGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ	
		RSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSV	
		RQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGGSSMASNFTQFVLVDNGGT	
		GDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGG	
		SGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQ	
		GLLKDGNPIPSAIAANSGIYGSSGGGGEAAAKTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARL	
		GIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRD	
		PEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEG	
		QRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYA	
		KGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPV	
		VALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGK	
		KLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLKRTADGSEFE	
		SPKKKAKVE	

[1767]

[1768] *표 S2. 구성 2(즉, RT - 링커 - (RBP)n - 링커 - Cas9)에 대한 예시적인 유전자 변형 서열*

식별자	전체 아미노산 서열	서열 번호
12037	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKALEAEAAAKEA AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS DCELIVKAMQGLLKDGNPIPSAIAANSGIYSAEAAAKEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVP SKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA YHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNL IALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPL ARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSR KLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMAREN QTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDN VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFY KVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLE AKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDK VLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKKAKVE	16,
P112038	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKALEAEAAAKEA AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS DCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGGSSGSETPGTSESATPESSGGGSSGGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA	

[1769]

	DLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG YIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI ECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKT ILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIRE QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,102
PL12039	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKALEAEAAAKEA AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS DCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSGSSGSGSSGSGSSDIKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTY DDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIK PILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRF NASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYP KLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12040	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	

[1770]

	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKALEAEAAAKEA	
	AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS	
	DCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLK	
	RTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG	
	DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN	
	LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD	
	GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQ	
	SFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH	
	DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDD	
	SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVEN	
	TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAER	
	GGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY	
	GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR	
	NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGEL	
	QKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY	
	FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12041	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,104
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEA	
	AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS	
	DCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVP	
	KGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKADKKYSI	
	GLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKK	
	HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK	
	SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY	
	DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY	
	PFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM	
	RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHL	

[1771]

	FDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL KDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANG EIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKS VKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYL DEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGS EFESPKKKAKVE	16,105
PL12042	MPAAKRVKLDGGTLNIEDEHRILHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEA AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS DCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVP KGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSGGSDKKYSIGLDIGTNSVGWAVIT DEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAY HEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVR QQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRI PYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYT GWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIV IEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKA RGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFR KDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILA DANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12043	MPAAKRVKLDGGTLNIEDEHRILHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	

[1772]

	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEA	
	AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS	
	DCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVP	
	KGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGGSGSGGSGSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGN	
	TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD	
	STDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTP	
	NFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK	
	NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA	
	WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE	
	DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDK	
	QSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQK	
	NSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKK	
	MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYH	
	HAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL	
	SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK	
	DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR	
	DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
P112044	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,107
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEA	
	AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS	
	DCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVP	
	KGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGGSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSI	
	KKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA	
	DLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF	
	DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG	
	YIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR	
	KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI	
	ECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKT	
	ILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER	
	MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNY	

[1773]

	WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIRE QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,108
PL12045	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKALEAEAAAKEA AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS DCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVP KGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEW ISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSMASNFTQFVLVDN GGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGI YAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF RGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFI KPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRF NASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYP KLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12046	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	

[1774]

[1775]

	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKEAAAKEA AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS DCELVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVP KGAWRSYLNMELTIPIFATNSDCELVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEW ISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDN GGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGI YSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQ DVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT NFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQLN EKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,110
PL12047	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKEAAAKEA AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS DCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVP KGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEW ISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDN GGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGI YGSSGSSGSSGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE MAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQ LFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI	

	LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD NGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE DIVLTLTFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY YVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF DSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,111
PL112048	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLAEAAAKEAAAKEAAAKALEAEAAAKEA AAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGI DCELIVKAMQGLLKDGNPIPSAIAANSGIYGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVP KGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEW ISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDN GGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGI YGSSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPI NASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVA QILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP	

[1776]

	EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDKRTADGSEFESPKKKAKVE	
12049	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGGSETPGTSESATPESSGGSSGGSS MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK DGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF DLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFI YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFI	16,112
PL12050	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGGSETPGTSESATPESSGGSSGGSS MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK DGNPIPSAIAANSGIYSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF RGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFI	16,113

[1777]

[1778]

	KPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRF NASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYP KLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,114
PL12051	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGSSEETPGTSESATPESSGGSSGGSS MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWSSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK DGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS DVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGD QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT NFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN EKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12052	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	

[1779]

	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGSETPGTSESATPESSGGSSGGSS	
	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	
	DGNPIPSAIAANSGIYGSSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI	
	CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLF	
	IQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL	
	AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL	
	LRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNL	
	PNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD	
	NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ	
	VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYY	
	LQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFI	
	KRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM	
	IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW	
	DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV	
	NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTK	
	EVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12053	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,116
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGSETPGTSESATPESSGGSSGGSS	
	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	
	DGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELT	
	IPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVI	
	TDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVA	
	YHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK	
	NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVR	
	QQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRI	
	PYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVD	
	LLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYT	
	GWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIV	
	IEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKA	

[1780]

	RGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFR KDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEK NPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILA DANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,117
PL12054	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEAERLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGSETPGTSESATPESSGGSSGGS MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK IPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGGSSETPGTSESATPESSGGSSGGSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLG NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTP NFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDK QSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQK NSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKK MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYH HAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12055	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEAERLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGSETPGTSESATPESSGGSSGGSS MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	

[1781]

	DGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELT	
	IPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSGSSGSGSSGSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALL	
	FDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALA	
	HMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKL	
	QLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQ	
	EEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPW	
	NFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIS	
	GVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDG	
	FANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIK	
	ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKL	
	ITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT	
	ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEV	
	QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELEN	
	GRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFT	
	LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKKAKVE	
PL12056	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,119
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGSETPGTSESATPESSGGSSGGS	
	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	
	DGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELT	
	IPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSGGSEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET	
	AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF	
	RGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT	
	YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFI	
	KPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV	
	DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRF	
	NASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF	
	MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL	
	KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF	
	DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYP	
	KLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS	
	KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM	

12057	LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,120
PL12058	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGETPGTSESATPESSGGSSGGSS MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAA DGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELT IPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCS VRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFA NGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAA AKEAAAKALEAEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYT RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS DVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGD QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT NFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN EKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,

[1782]

[1783]

	DGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELT	
	IPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCS	
	VRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFA	
	NGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGSETP	
	GTSESATPESSGGSSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSN	
	EMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYN	
	QLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSD	
	AILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL	
	DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP	
	KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL	
	EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL	
	HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDM	
	YVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR	
	QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG	
	KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF	
	DSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH	
	YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH	
	QSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12059	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,122
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGSETPGTSESATPESSGGSSGGSS	
	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	
	DGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELT	
	IPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCS	
	VRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFA	
	NGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSS	
	GSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEES	
	FLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAK	
	AILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPL	
	SASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL	
	RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVK	
	YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL	

	KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQ TVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKY DENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKT EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKG KSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFV EQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD KRTADGSEFESPKKKAKVE	16,123
PL12060	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLSGGSSGGSSGSSETPGTSESATPESSGGSSGGSS MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGDSGGGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCS DGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELT IPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGDSGGGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCS VRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFA NGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKDK KYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEE DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSAR LSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMI KRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTY AHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVK VVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKR TADGSEFESPKKKAKVE	16,
PL12061	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	

[1784]

[1785]

	KNSPTLFNEALHRIDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGSSGSGSGSGSSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	AEAAAKEAAAKEAAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET	
	AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF	
	RGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT	
	YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFI	
	KPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV	
	DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRF	
	NASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ	
	MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKL	
	KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF	
	DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYP	
	KLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS	
	KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM	
	LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA	
	PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
P112062	MPAAKRVKLDGGTLNIEDEHRILHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,125
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRIDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGSSGSGSGSGSSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYS	
	GGSSGGSSGSETPGTSESATPESSGGSSGSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR	
	RKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSD	
	VDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ	
	YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK	
	LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTN	
	FDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDK	
	DFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ	
	KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKL	

	YLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDK AGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDV RKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKK DWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYT STKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,126
PL12063	MPAAKRVKLDGGTLNIEDEHRLIHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGSGSGSGSGSGSGSSMASNFTQFVLVDNG GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY GSGSGSGSGSGSGSGSGSGSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE MAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQ LFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAI LLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD NGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE DIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDM YVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF DSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12064	MPAAKRVKLDGGTLNIEDEHRLIHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGSGSGSGSGSGSGSSMASNFTQFVLVDNG GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	

[1786]

[1787]

	GSSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD	
	DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPI	
	NASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILR	
	VNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH	
	QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY	
	FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL	
	FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA	
	GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI	
	NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVA	
	QILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYF	
	FYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV	
	SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSP	
	EDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE	
	TRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12065	MPAAKRVKLDGGTLNIEDEHRLIETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,128
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKACMQGLLKDGNPIPSAIAANSGIY	
	GGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKACA	
	MQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVL	
	GNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK	
	LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGL	
	TPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQ	
	SKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSR	
	FAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL	
	KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR	
	DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKG	
	QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVV	
	KKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINN	
	YHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK	
	VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV	

[1788]

12067	KKDLIIKLPKYSLFELENGRKMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHFLTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAATETPDTSTLLGSSGSSGSSGSSGSSMASNFTQFVLVDNG GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKA GGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKA MQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRT ARRIYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLL AQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDG TEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQS FIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDD SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVEN TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAER GGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGEL QKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHFLTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,130
PL12068	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAATETPDTSTLLGSSGSSGSSGSSGSSMASNFTQFVLVDNG GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY GGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKA MQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRY TRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDN	16,1[illegible]

[1789]

	SDVDKLFQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGD	
	QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV	
	KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT	
	NFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK	
	DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE	
	DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN	
	EKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE	
	LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV	
	YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLI	
	ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL	
	ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR	
	KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12069	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,132
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNG	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	
	GGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKA	
	MQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIK	
	VEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQ	
	AYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAKALEAEA	
	AAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS	
	NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTY	
	NQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS	
	DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT	
	FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP	
	KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL	
	EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL	
	HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDM	
	YVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR	
	QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG	
	KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF	
	DSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH	

	YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
12070	MPAAKRVKLDGGTLNIEDEHRILHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSGGSGGSGGSSGGSSGGSSMASNFTQFVLVDNG GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY GGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKA MQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIK VEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQ AYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGGSETPGTSESATPESSGGSS GGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLE ESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELH AILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELT KVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMI EERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKK GILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRM NTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQK QLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQ LGGDKRTADGSEFESPKKKAKVE	
PL12071	MPAAKRVKLDGGTLNIEDEHRILHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSGGSGGSGGSSGGSSGGSSMASNFTQFVLVDNG GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY	16,

[1790]

[1791]

	GGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKA MQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIK VEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQ AYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGGSSGGSSGGSSGSDKKYSIGLDIGT NSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPI FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLEN LIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK GEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR KQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVM GRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSID NKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVI TLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPL IETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGI TIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQI SEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPK KKAKVE	16,135
PL12072	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGSSGSGSGSGSSSMASNFTQFVLVDNG GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIY GGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKA MQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIK VEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQ AYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGGSSGGSSGGSSGSDKKYSIGLDIGTNSVG WAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR	

[1792]

	RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKP ENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR SDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET GEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI LADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,136
PL12073	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGGGEAAAKMASNFTQFVLVDNGGTGDVT VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKE AAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKADKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKR TARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQ SFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYH DLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDD SLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVEN TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAER GGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGEL QKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL112074	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGGGEAAAKMASNFTQFVLVDNGGTGDVT	

[1793]

	VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCEUVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGG	
	SSGSETPGTSESATPESSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICY	
	LQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQ	
	LVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLA	
	AKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLL	
	RKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLP	
	NEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNE	
	ENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS	
	GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQ	
	NGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKR	
	QLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIA	
	KSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP	
	KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED	
	DATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12075	MPAAKRVKLDGGTLNIEDEHRLIHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,138
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGGGEAAAKMASNFTQFVLVDNGGTGDVT	
	VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCEUVKAMQGLLKDGNPIPSAIAANSGIYGSGSSSG	
	SSGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDS	
	FFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPIN	
	ASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRV	
	NTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQI	
	HLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFT	
	VYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFE	
	DREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG	
	SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN	
	RLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQI	
	LDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY	
	SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV	
	LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED	

	NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDKRTADGSEFESPKKKAKVE	
12076	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGGGEAAAKMASNFTQFVLVDNGGTGDVT VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELVKAMQGLLK AAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA DLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF	16,139
PL12077	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGGGEAAAKMASNFTQFVLVDNGGTGDVT VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELVKAMQGLLK MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELVKAMQGLLK DGNPIPSAIAANSGIYAEAAAAKEAAAAKEAAAAKALEAEAAAAKEAAAAKEAAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHS IKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKA DLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNF	16,1

[1794]

[1795]

	DLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAG	
	YIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR	
	KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKI	
	ECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKT	
	ILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER	
	MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNY	
	WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHD	
	AYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ	
	VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL	
	PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIRE	
	QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12078	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEAARLGIKPHIQRLLDQGILVPCQ	16,141
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGGGEAAAKMASNFTQFVLVDNGGTGDVT	
	VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	
	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	
	DGNPIPSAIAANSGIYSGGGSSGGGSSGGSETPGTSESATPESSGGGSSGGGSSDKKYSKGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGET	
	AEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKF	
	RGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT	
	YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFI	
	KPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV	
	DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRF	
	NASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF	
	MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL	
	KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKAARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF	
	DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYP	
	KLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFS	
	KESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM	
	LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA	
	PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12079	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEAARLGIKPHIQRLLDQGILVPCQ	16,
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	

	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGSGGEAAAKMASNFTQFVLVDNGGTGDVT VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGS MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK DGNPIPSAIAANSGIYGSSGGSSGGSSGGSSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI RRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNS DVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNL NFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN EKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,143
P112080	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGSGGEAAAKMASNFTQFVLVDNGGTGDVT VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGS MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK DGNPIPSAIAANSGIYGSSGGSSGGSSGGSSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLF IQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL LRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNL PNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD	

[1796]

	NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ	
	VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYY	
	LQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFI	
	KRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM	
	IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDW	
	DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYV	
	NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTK	
	EVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,144
PL12081	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGGGEAAAKMASNFTQFVLVDNGGTGDVT	
	VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGS	
	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	
	DGNPIPSAIAANSGIYGGSGGSGGGSGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKG	
	AWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTC	
	SVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKALEAEAAAAKEAA	
	AKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN	
	PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL	
	RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIP	
	HQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYE	
	YFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTL	
	TLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIAN	
	LAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQEL	
	DINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV	
	AQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKY	
	FFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVA	
	YSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS	
	PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLY	
	ETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12082	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	

[1797]

[1798]

	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGGGEAAAKMASNFTQFVLVDNGGTGDVT	
	VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGS	
	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	
	DGNPIPSAIAANSGIYGGGSGGGGSGGGGSGGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKG	
	AWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTC	
	SVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGSETPGTSESATPESSGGSSGGSSD	
	KKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVE	
	EDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSA	
	RLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS	
	MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR	
	QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV	
	TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT	
	YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTV	
	KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIV	
	PQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE	
	NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT	
	LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS	
	KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ	
	HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKR	
	TADGSEFESPKKKAKVE	
PL12083	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,1
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGGGEAAAKMASNFTQFVLVDNGGTGDVT	
	VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGS	
	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	
	DGNPIPSAIAANSGIYGGSGGSGGGGSGGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKG	
	AWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTC	
	SVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSDKKYSIGLDIGTNSVG	

[1799]

	WAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV	
	DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP	
	GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK	
	ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI	
	LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK	
	AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR	
	RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKP	
	ENIVIEMARENGQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR	
	SDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ	
	SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET	
	GEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF	
	FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI	
	LADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12084	MPAAKRVKLDGGTLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQ	16,147
	SPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGF	
	KNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK	
	TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKK	
	LDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDI	
	LAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR	
	RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLGSSGGGGEAAAKMASNFTQFVLVDNGGTGDVT	
	VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKDKKYSKGLDIGTNSVGWAVITDE	
	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLK	
	DGNPIPSAIAANSGIYGGGSGGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKG	
	AWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTC	
	SVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKDKKYSKGLDIGTNSVGWAVITDE	
	YKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHE	
	KYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGL	
	FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL	
	PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYY	
	VGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK	
	TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG	
	RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMA	
	RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKS	
	DNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ	
	FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK	
	GRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF	

	LEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLD	
	KVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
12109	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,148
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEEAA	
	AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMVELTIPIFATNSDCELIVKAMQG	
	LLKDGNPIPSAIAANSGIYSAEAAAAKEAAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTD	
	RHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDST	
	DKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNF	
	KSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNG	
	GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA	
	WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE	
	DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDK	
	QSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQK	
	NSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKK	
	MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYH	
	HAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEVVWDKGRDFATVRKVL	
	SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK	
	DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR	
	DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12110	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,14
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEEAA	
	AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMVELTIPIFATNSDCELIVKAMQG	
	LLKDGNPIPSAIAANSGIYSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS	
	GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHM	
	IKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK	
	DTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY	

[1800]

[1801]

	KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELG SQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIK KYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,150
PL12111	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEAAAKEAAAKEAAAKAALEAEAAAKEAAAKEAAAKEAA AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQIG LLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPD NSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN EKLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12112	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	

[1802]

	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEAAAKEAAAAKEAAAKEAAAKEAAAKEAAAKEAA AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRIKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQIG LLKDGNPIPSAIAANSGIYGGSSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKA QVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS TKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,152
P112113	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQVPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAA AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRIKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQIG LLKDGNPIPSAIAANSGIYGGSSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRIKYTIKVEVPKGAWRSYLNM ELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAA AVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVD EVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLK ALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKP ENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTR	

	SDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV	
	SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGET	
	GEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSS	
	FEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVI	
	LADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12114	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWVKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,153
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRIPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARKEVAIRPLSTQATISAEAAAKEAAAAKEAAAAKEAAAAKEAAAAKEAAAAKEAAAAKEAA	
	AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQG	
	LLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNM	
	ELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSGGSGSETPGTSESATPESSGGSGGSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFK	
	VLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLR	
	KKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL	
	GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF	
	DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARG	
	NSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV	
	KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN	
	GIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT	
	QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSE	
	EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVRE	
	INNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT	
	VRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY	
	KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAY	
	NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12115	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWVKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRIPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARKEVAIRPLSTQATISAEAAAKEAAAAKEAAAAKEAAAAKEAAAAKEAAAAKEAAAAKEAA	
	AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQG	

[1803]

	LLKDGNPIPSAIAANSGIYGGSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNM	
	ELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG	
	ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL	
	ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED	
	AKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG	
	ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI	
	TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS	
	VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLK	
	SDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE	
	EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLL	
	NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA	
	VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVK	
	KTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLF	
	ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENII	
	HLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12116	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILHPVHSP	16,155
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLIMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAA	
	AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQG	
	LLKDGNPIPSAIAANSGIYGGSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNM	
	ELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS	
	GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHM	
	IKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK	
	DTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY	
	KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE	
	EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE	
	DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN	
	RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELG	
	SQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ	
	RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIK	
	KYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG	
	GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK	

[1804]

	RMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN	
12117	LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,156
	MPAAKRVKLDGGTAPLEEEYRLFLEAPKQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARE VAIRPLSTQATISAEAAAKEAAAKEAAAKEALEAEAAAKEAAAKEAAAKEAAA	
	AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQG	
	LLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNM	
	ELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPS	
	TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPS	
	NFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAK	
	EAAAKEAAAKALEAEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR	
	RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP	
	DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI	
	GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEE	
	LLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER	
	MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI	
	IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFK	
	EDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQ	
	NEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS	
	ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK	
	VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL	
	IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE	
	LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID	
	RKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12118	MPAAKRVKLDGGTAPLEEEYRLFLEAPKQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARE VAIRPLSTQATISAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAA	
	AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQG	

[1805]

[1806]

	LLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNM	16,158
	ELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV	
	TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPS	
	NFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGGSSGS	
	ETPGTSESATPESSGGSSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF	
	SNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQT	
	YNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL	
	SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQR	
	TFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYN	
	PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE	
	LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDS	
	LHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD	
	MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE	
	TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQ	
	EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG	
	GFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA	
	SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT	
	LIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
P112119	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPVTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWILTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEAAAKEAAAKEAAAKLAEAEAAAKEAAAKEAA	
	AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQG	
	LLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNM	
	ELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV	
	TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPS	
	NFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGS	
	SGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFH	
	RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASG	
	VDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI	
	TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG	
	ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYN	
	ELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE	

	MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAI KKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD YDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNI MNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNE QIKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL SQLGGDKRTADGSEFESPKKKAKVE	16,159
PL12120	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQVPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLLHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAA AKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQG LLKDGNPIPSAIAANSGIYGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNM ELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPS NFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGA KDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFL VEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSA SMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTV KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIV PQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKR TADGSEFESPKKKAKVE	16,
PL12121	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQVPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	

[1807]

[1808]

	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSISNSRTQAILQIPVPKTKRQVRE	
	FLGKGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLFSLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGGSSGGGSSGSETPGTSESATPESSGGGSGGSSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMDGLLIKDSGNPIPSAIAANS	
	GIYAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN	
PL12122	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,161
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSISNSRTQAILQIPVPKTKRQVRE	
	FLGKGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLFSLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGGSSGGGSSGSETPGTSESATPESSGGGSGGSSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMDGLLIKDSGNPIPSAIAANS	
	GIYSGGSSGGGSSGSETPGTSESATPESSGGGSSGGSSDMKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR	
	YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPD	
	NSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG	
	DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL	
	VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM	
	TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK	
	DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE	
	DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN	

[1809]

	EKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLI ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,162
PL12123	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWVKREIPKVVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGGSSGGGSSGSETPGTSESATPESSGGGSSGGSSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS GIYGSSGSSGSSGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTY NQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDM YVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF DSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12124	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWVKREIPKVVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGGSSGGGSSGSETPGTSESATPESSGGGSSGGSSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS	

	GIYGSSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK	
	VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSD	
	ILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI	
	PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY	
	EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLT	
	LTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA	
	NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE	
	LDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH	
	VAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA	
	KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT	
	VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK	
	GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIT	
	GLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12125	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,164
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQICQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLIPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGGSSGGGSGGSETPGTSESATPESSGGGSSGGGSSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS	
	GIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIV	
	KAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFK	
	VLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLR	
	KKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSL	
	GLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF	
	DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARG	
	NSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV	
	KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN	
	GIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT	
	QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSE	
	EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVRE	
	INNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT	
	VRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY	

[1810]

[1811]

	KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAY	
	NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
12126	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,165
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARELVAIRPLSTQATISSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIV	
	GIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIV	
	KAMQGLLKDGNPIPSAIAAMSGIYSGGSGGSSGGSETPGTSESATPESSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG	
	ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL	
	ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED	
	AKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG	
	ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI	
	TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS	
	VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLK	
	SDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE	
	EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLL	
	NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA	
	VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVK	
	KTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLF	
	ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENII	
	HLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
P112127	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,1£
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARELVAIRPLSTQATISSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAAMS	
	GIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIV	
	KAMQGLLKDGNPIPSAIAAMSGIYSGGSGGSSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL	
	KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIE	

	GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGA SAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLT KAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESIL PKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASA GELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAA FKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,167
PL12128	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRIFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS GIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIV KAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARR RYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNP DNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEE LLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKI IKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFK EDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQ NEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTID RKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	

[1812]

PL12129	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,168
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSG	
	GIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIV	
	KAMQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKY	
	TIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNS	
	RSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGAEAAAKEAAAKEAAAKEAALE	
	AEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE	
	IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQ	
	TYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKN	
	LSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ	
	RTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEK	
	VLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENE	
	DILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG	
	DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGR	
	DMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV	
	ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSE	
	QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY	
	GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYL	
	ASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDA	
	TLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12130	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGGSSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS	
	GIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIV	
	KAMQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKY	

[1813]

[1814]

	TIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNS RSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSQSSETPGTSESATPESS GGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFF HRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINAS GVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNT EITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRL SDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILD SRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSN IMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV VAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRID LSQLGGDKRTADGSEEFESPKKKAKVE	16,170
PL12131	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WINTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKQALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLIRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARSVAIRPLSTQATSSGGSSGGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS GIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIV KAMQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKY TIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNS RSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGGSGGSGGSGGSGGSDKKYSIGL DIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFL KDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKP AFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDD KVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELV KVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKD	

[1815]

	DSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIRE VKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIR KRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVK ELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSE FESPKKKAKVE	16,171
PL12132	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGGSSGGGSSGGSETPGTSESATPESSGGGSSGGSSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKDKKYSIGLDIGTNS GIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIV KAMQGLLKDGNPIPSAIAANSGIYGSSGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKY TIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNS RSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKDKKYSIGLDIGTNS VGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFG NIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIA QLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNRE KIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRVTGWGRLLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMG RHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNK VLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITL KSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIE TNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITI MERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISE FSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKK AKVE	16,
PL12133	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	

[1816]

	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARevAIRPLSTQATISGSGSGSSGSGSSGSGSSMASNFTQFVLVDNGGTGDVTVAPSN	
	FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKE	
	AAAKEAAAKALEAEAAAKEAAAKEAAAKAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR	
	YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPD	
	NSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG	
	DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL	
	VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT	
	TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK	
	DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE	
	DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN	
	EKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR	
	LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV	
	YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLI	
	ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL	
	ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR	
	KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12134	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,173
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARevAIRPLSTQATISGSGSGSSGSGSSGSGSSMASNFTQFVLVDNGGTGDVTVAPSN	
	FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGSGSE	
	TPGTSESATPESSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS	
	NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTY	
	NQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS	
	DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT	
	FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP	
	KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL	
	EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL	
	HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDM	
	YVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR	
	QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG	

	KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF DSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,174
12135	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWVKREIPKVVWAEINPPGLASTQAPIHVQLLSTALPVRVRIQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSS GSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLV EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVD AKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITK APLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGEL HAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNEL TKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREM IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKK GILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRM NTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMN FFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQK QLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQ LGGDKRTADGSEFESPKKKAKVE	
PL12136	MPAAKRVKLDSGTAPLEEEYRLFLEAPIQNVTLLEQWVKREIPKVVWAEINPPGLASTQAPIHVQLLSTALPVRVRIQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARFVAIRPLSTQATISGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLV EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS	16,

[1817]

[1818]

	ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTV KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIV PQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE TADGSEFESPKKKAKVE	16,176
P112137	MPAAKRVKLDGGTAPLEEEYRLFLEAPRQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNF TQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP SAIAANSGIYAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAED AKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETI TPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLK SDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLL NAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVK KTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLF ELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENII HLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	

[1819]

[illegible]_12138	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLIRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP TQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP SAIAANSGIYSGGSSGGSSGGSSETPGTSESATPESSGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLF GDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGA SAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASL GTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLI HDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLT KAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESIL PKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASA GELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAA FKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,177
PL12139	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLIRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP TQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP SAIAANSGIYSGGSSGGSSGGSSGSSETPGTSESATPESSGGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRI CYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLF IQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDL	16,17

[1820]

	LRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNL PNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQ VSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYY LQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFI KRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG DPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA NFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTK EVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,179
PL12140	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP TQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP SAIAANSGIYGSSGGGGEAAAKDKKYSVGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIF SNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQT YNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNL SDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQR TFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVL PKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDS LHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQ EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYG GFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDAT LIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12141	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	

[1821]

	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN	
	FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP	
	TQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP	
	SAIAANSGIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYL	
	NMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSS	
	AQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKALEAEAAAAKEAAAA	
	KEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFH	
	RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASG	
	VDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEI	
	TKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLG	
	ELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYN	
	ELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRE	
	MIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAI	
	KKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSD	
	YDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR	
	MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNI	
	MNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV	
	AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNE	
	QKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL	
	SQLGGDKRTADGSEFESPKKKAKVE	16,1E
PL12142	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLIPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN	
	FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNF	
	TQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP	
	SAIAANSGIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYL	
	NMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSS	
	AQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSGGGSGGGSGETPGTSESATPESSGGGSSGGSSDKKYSIGL	
	DIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE	

[1822]

	RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR	
	RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE	
	HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFL	
	KDNIREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKP	
	AFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDD	
	KVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELV	
	KVMGRHKPENIVIEMAARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKD	
	DSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIRE	
	VKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIR	
	KRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVK	
	ELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD	
	EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSE	
	FESPKKKAKVE	
PL12143	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,182
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQICQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGAIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN	
	FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNF	
	TQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP	
	SAIAANSGIYGGGSGGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYL	
	NMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSS	
	AQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDE	
	YKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHE	
	KYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGL	
	FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL	
	PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYY	
	VGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK	
	TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWG	
	RLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMA	
	RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKS	
	DNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ	
	FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDK	
	GRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDF	

[1823]

	LEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLD	
	KVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
12144	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,183
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVEDIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSN	
	FANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSK	
	TQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIP	
	SAIAANSGIYGGGSGGGSGGSGGSGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYL	
	NMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSS	
	AQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSK	
	KFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIY	
	HLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLI	
	ALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYK	
	EIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA	
	RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKV	
	TVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRK	
	LINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQ	
	TTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVP	
	SEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV	
	REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF	
	ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK	
	GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLS	
	AYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12145	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVEDIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA	
	EWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKE	
	AAAKALEAEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRK	

	NRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD	
	KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA	
	DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN	
	REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD	
	KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDF	
	LDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKA	
	QVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLY	
	YLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAG	
	FIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK	
	MIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD	
	WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK	
	YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTS	
	TKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12146	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPTLEAKRSLRETIRKFRAAGILRPVHSP	16,185
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDIKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA	
	EWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGGSSGSETPGTSE	
	SATPESSGGGSSGGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK	
	VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE	
	NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSD	
	ILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI	
	PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY	
	EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLT	
	LTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA	
	NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE	
	LDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH	
	VAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA	
	KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT	
	VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK	
	GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIT	
	GLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	

[1824]

[1825]

PL12147	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKIALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPKGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA EWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGGSSGGSGSS DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLV EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRR QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTV KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIV PQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKR TADGSEFESPKKKAKVE	16,186
PL12148	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKIALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPKGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARELVAIRPLSTQATISGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA EWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGGSSGGSGSS LDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKH ERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKS RRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHL	16,187

[1826]

	FDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL KDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANG EIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKS VKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYL DEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGS EFESPKKKAKVE	16,188
PL12149	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQVPTLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVTWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA EWISSNSRSQAYKVTCSVRDSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS GIYAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMI KFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK DTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFAN RNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELG SQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIK KYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTN LGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12150	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQVPTLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	

[1827]

	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA	
	EWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS	
	GIYSGGSSGGSSGSETPGTSESATPESSGGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR	
	YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPD	
	NSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG	
	DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL	
	VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERM	
	TNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIK	
	DKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKE	
	DIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN	
	EKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE	
	LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVVTLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKV	
	YDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLI	
	ARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNEL	
	ALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR	
	KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	
PL12151	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,190
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEFVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRIAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWILTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARFVAIRPLSTQATISGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA	
	EWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS	
	GIYGSGSSGGSSGSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFS	
	NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTY	
	NQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLS	
	DAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT	
	FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP	
	KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL	
	EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSL	
	HEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDM	
	YVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETR	

[1828]

	QITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF DSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASH YEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIH QSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,191
PL12152	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLIRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARKVAIRPLSTQATISGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA EWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS GIYGSSGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEE NPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSD ILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSI PHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLY EYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLT LTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIA NLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE LDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH VAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATA KYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSIT GLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12153	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLIRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARKVAIRPLSTQATISGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA EWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLV	

[1829]

	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS	
	GIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIP	
	IFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYT	
	IKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKA	
	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLV	
	EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS	
	ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSAS	
	MIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRR	
	QEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYV	
	TEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT	
	YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTV	
	KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIV	
	PQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDE	
	NDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT	
	LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS	
	KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ	
	HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKR	
	TADGSEFESPKKKAKVE	
PL12154	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQVVKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP	16,193
	WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT	
	LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTDAILQIPVPKTKRQVRE	
	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKIALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG	
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARIITQYQVLLLDPPRVRIFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT	
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN	
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA	
	EWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLV	
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS	
	GIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIP	
	IFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYT	
	IKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYSGGSSGGSSGSETPGTSESATPESSGGSSGGSSDKKYSIGLDIGTNSV	
	GWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNI	
	VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQ	
	LPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTL	
	LKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI	
	EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQ	
	KKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQL	

[1830]

	KRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRH KPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVL TRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETN GETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIME RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS KRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKA KVE	16,194
PL12155	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA EWISSNSRSQAYKVTCSVRDSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSSMASNFTQFVLV DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS GIYGGGSGGGSGGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIP IFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYT IKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGSSGSSGSSGSSGSSGSSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKF KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHL RKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIAL SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEI FFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLAR GNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVT VKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLI NGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSE EVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVRE INNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFAT VRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY KEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAY NKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE	16,
PL12156	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSP WNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPT LFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVRE	

[1831]

	FLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAG
	WPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLT
	DQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKN
	APEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIA
	EWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLV
	DNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANS
	GIYGGGSGGGSGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIP
	IFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYT
	IKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGSSGGGGEAAAKDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGN
	TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
	STDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTP
	NFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSK
	NGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA
	WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKE
	DYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDK
	QSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQK
	NSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKK
	MKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYH
	HAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL
	SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK
	DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR
	DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDKRTADGSEFESPKKKAKVE

[1832] 표 S3. 구성 3(즉, Cas9 - 링커 - RT - 링커 - (RBP)n)에 대한 예시적인 유전자 변형 서열

식별자	전체 아미노산 서열	서열 번호
PL12085	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQNKREIPKVVWAEINPPGLASTGAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW SKDKSVNIYTDSRYAFATLHVHGAMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARERVAIRPLSTQATISAEA AAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNR KYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,200
PL12086	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL	16,

[1833]

[1834]

	FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA QLCQEFVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRVYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEA AAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNR KYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSN SRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,202
PL12087	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA	

[1835]

QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEA AAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNR KYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGGSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSN SRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSMASNFT QFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPS AIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS DCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,203
PL12088	MIPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLIRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGG SSGGSSGGSETPGTSESATPESSGGSSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRS YLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	

[1836]

PL12089	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKEAALEAEAAAKEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF FCPLAPESQLIFAFEWADAEEGESGQLTWTRLPDGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGG SSGGSSGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRS YLNMELTIPIFATNSDCELVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQS SAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,204
PL12090	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY	16,**

[1837]

	SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV	
	LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED	
	NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI	
	DLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAAEINPPGLASTQAPIH	
	VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF	
	FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA	
	QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK	
	PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAVAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP	
	RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW	
	SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISSGG	
	SSGGSSCGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSY	
	YLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSGGGSSMASNFTQFVLVDNGGTGDVT	
	SAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGGSGGGSGGGSGGGSSMASNFTQFVLVDNGGTGDVT	
	VAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGGSGGS	
	MASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKD	
	GNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	
PL12091	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,206
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN	
	PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL	
	RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH	
	QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY	
	FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL	
	FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA	
	GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI	
	NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ	
	ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY	
	SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV	
	LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED	
	NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI	
	DLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAAEINPPGLASTQAPIH	
	VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF	
	FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA	
	QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK	

[1838]

	PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP RVRFKQTAALNPATLLPETDDTLPHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSS GSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS DCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,207
PL12092	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAAKEAAAAKEAAAKALEAEAAAAKEAAAAKEAAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP RVRFKQTAALNPATLLPETDDTLPHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSS GSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS DCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE KGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,
PL12093	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH	

[1839]

	QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY	
	FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL	
	FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA	
	GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI	
	NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ	
	ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY	
	SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV	
	LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED	
	NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI	
	DLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIH	
	VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF	
	FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA	
	QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK	
	PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP	
	DLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVVAEINPPGLASTQAPIH	
	RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVTWAEPLPIGTSAQKAELIALTKALEW	
	VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF	
	SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSS	
	FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA	
	GSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNS	
	QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK	
	DCELVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVP	
	PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP	
	KGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWI	
	RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVTWAEPLPIGTSAQKAELIALTKALEW	
	SSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNG	
	SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSS	
	GTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYK	
	RTADGSEFESPKKKAKVE	
PL112094	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,2(
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN	
	PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL	
	RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH	
	QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY	
	FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL	
	FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA	
	GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI	
	NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ	
	ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY	
	SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV	

[1840]

	LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKEAEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQVKREIPKVVAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSS GGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVK AMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,210
PL12095	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQVKREIPKVVAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF FCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKA QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPP RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVARELVAIRPLSTQATISGSS GGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVK	

[1841]

	AMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRS YLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,211
12096	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIH VQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAF FCIPLAPESQLIFAFEWADAEFEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVVSGKKA QLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKLALTQPPALALPSLDK PFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNAARITQYQVLLLDPP RVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQKAELIALTKALEW SKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISGSS GGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVK AMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRS YLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRS QAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSMASNFTQFVLVDNGGTGDV TVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADG SEFESPKKKAKVE	16,
PL12097	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH	

QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY		
FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL		
FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA		
GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI		
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ		
ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY		
SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV		
LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED		
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI		
DLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQA		
PLIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL		
KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRA		
SAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA		
LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR		
MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS		
AQRAEIIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAI		
TETPDTSTLLAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV		
TCSVRQSSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE		
	PL12098	16,213
MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV		
DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN		
PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL		
RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH		
QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY		
FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL		
FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA		
GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI		
NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ		
ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY		
SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV		
LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED		
NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI		
DLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQA		
PLIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL		

[1842]

[1843]

	KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRA SAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS AQIRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAI TETPDTSTLLAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPS NFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFES PKKKAKVE	16,214
PL12099	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTAARRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQA PLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRA SAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS AQIRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAI TETPDTSTLLAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKAMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPS NFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGG GSGGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKA	

[1844]

12100	MQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSY LNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,215
	MPAAKRVKLDGSDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI	

[1845]

	NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ	
	ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY	
	SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV	
	LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED	
	NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI	
	DLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQA	
	PLIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL	
	KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRA	
	SAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA	
	LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR	
	MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS	
	AQRAEIIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAI	
	TETPDTSTLLSGGSSGGSGSSGETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYT	
	IKVEVPKGAWRSYLNMELTIPIFATNSDCEIVKAMQGLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRS	
	QAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCEIVKAMQGLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	
PL12102	MPAAKRVKLDGGDNKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	16,217
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN	
	PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL	
	RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH	
	QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY	
	FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL	
	FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA	
	GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI	
	NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ	
	ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY	
	SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV	
	LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED	
	NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI	
	DLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQA	
	PLIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL	
	KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRA	
	SAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA	
	LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR	

[1846]

	MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAI TETPDTSTLLSGGSSGGSSGGSETPGTSESATPESSGGSSGGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYT IKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRS QAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSGGSGGSMASNFTQFV LVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAA NSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCEL IVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,218
PL12103	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQA PLIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRA SAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAI TETPDTSTLLGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYL NMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,
PL12104	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN	

[1847]

	PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL	16,22
	RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH	
	QIHLGELHAILRRQEDFYPFLKDNREKEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY	
	FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL	
	FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA	
	GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI	
	NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ	
	ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY	
	SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV	
	LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED	
	NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI	
	DLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQA	
	PLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL	
	KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRA	
	SAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA	
	LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR	
	MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS	
	AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAI	
	TETPDTSTLLGSSGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYL	
	NMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSS	
	AQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	
P12105	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV	
	DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN	
	PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL	
	RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH	
	QIHLGELHAILRRQEDFYPFLKDNREKEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY	
	FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL	
	FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA	
	GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI	
	NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ	
	ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY	
	SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV	
	LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED	

[1848]

	NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQA PLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRA SAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAI TETPDTSTLLGSSGSSGSSGSSGSSGSSGSSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYL NMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSS AQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSGGSGGSGGSMASNFTQFVLVDNGGTGDVTV APSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSSM ASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDG NPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,221
PL12106	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNRERIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLA GSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDI NRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFY SNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSV LVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPED NEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDAEAAAKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKATLNIEDEHRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQA PLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRA SAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTS

[1849]

	AQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAI TETPDTSTLLGSSGGGEAAAKMASNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIP IFATNSDCELIVKAMQGLLKDGNPIPSAIAANSGIYKRTADGSEFESPKKKAKVE	16,222
J12107	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL	16,
PL12108	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKV DDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEEN PINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDIL RVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEY FTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL

[1850]

[1851] 2) 가이드 RNA:

[1852] 가이드 RNA(gRNA) 분자(5'-GCCGAAGCACTGCACGCCGTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC -3'; 서열번호 16,708)는 Cas9에 결합하고 표적 유전자좌에 상보적인 프로-스페이서 서열을 함유하여, 표적 게놈 유전자좌로 유전자 변형 복합체의 국재화를 초래한다.

[1853] 3) 트랜스-주형 RNA:

[1854] 다음 각각의 영역을 이용하여 일련의 384개 ttRNA를 구축하였다(288개 실험 구성 및 96개 음성 대조군):

[1855] 1. 닉이 생긴 DNA 가닥과 염기쌍을 형성하여 닉이 생긴 DNA의 프라이머 연장을 가능하게 하는 길이가 8, 13, 또는 17개 뉴클레오티드인 프라이머 결합 부위(PBS), 바로 이어서

[1856] 2. 표적 BFP/GFP 유전자좌와 상동성을 갖는 상동성 1 영역

[1857] 3. 예시적인 원하는 변형: BFP 표적을 GFP로 전환하는 C에서 T로의 치환(편집)뿐만 아니라, 표적 부위에서 다중 전환 절단을 방지하기 위한 PAM-무효화 돌연변이.

[1858] 4. 표적 BFP/GFP 유전자좌와 상동성을 갖는 12개 뉴클레오티드 또는 63개 뉴클레오티드 상동성 2 영역.

[1859] 5. 개재 RNA 링커 서열이 있는 1, 2, 또는 4개 MS2 서열(1x, 2x, 또는 4x MS2)을 함유하는 RBP 동원 부위(**RRS**).

[1860] 6. 8 nt 링커 서열을 갖는 ePEG를 포함하는 3' 말단-보호 RNA 2차 구조 또는 3' 말단-보호 RNA 이차 구조가 없음.

[1861] 7. 5' 말단-보호 RNA 2차 구조 또는 다음 5' 말단-보호 RNA 2차 구조 중 하나:

[1862] a. ePEG

[1863] b. 스페이서 서열이 결여된 Cas9 스캐폴드

[1864] c. 16개의 뉴클레오티드 스페이서 서열을 갖는 Cas9 스캐폴드

[1865] 8. RRS와 PBS 사이의 8 또는 16개 뉴클레오티드 링커

[1866] 이러한 구성요소를 하기 구성으로 조립하였다(3' 또는 5' 말단 보호 구조가 없는 것이 선택되는 경우, 그러한 구조는 주형 RNA에 포함되지 않음):

[1867] 3'-말단 블록 - $(RRS)_n$ - 링커 - PBS-주형 - 5' 말단 블록

[1868] 주형 서열은 하기 표 S4에 포함되어 있다.

[1869] *표 S4. 예시적인 주형 서열*

식별자	주형 서열	서열 번호
13280	GACCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGATCACCCATGTgc	16,301
PL13281	GACCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATA ACCCTCAAAGTCGGGGGGGC	16,302
PL13282	GACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,303
PL13283	GACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGG TCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,304
PL13284	GACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccctctctctctctc tctcgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,305
PL13285	GACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccctctctctctctc tctcgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAAC CCTCAAAGTCGGGGGGGC	16,306
PL13286	GACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,307
PL13287	GACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGGTCAGGAGCCCCCCCCCTGAACC CAGGATAACCCTCAAAGTCGGGGGGC	16,308
PL13288	GACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,309
PL13289	GACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccTCT CTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,310
PL13290	GACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccctctc tctctctctctctcgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,311
PL13291	GACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccctctc tctctctctctctcgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGGTCAGGAGCCCCCCCCCTGAACCC AGGATAACCCTCAAAGTCGGGGGGC	16,312
PL13292	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATGTgc	16,31
PL13293	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGGTCAGGAGCCCCCCCCCTGAACCCAG GATAACCCTCAAAGTCGGGGGGC	16,31
PL13294	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,31
PL13295	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccTCTCTC TCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,31
PL13296	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccctctctct ctctctctctcgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,31

[1870]

[1871]

PL13297	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctct ctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGG ATAACCCTCAAAGTCGGGGGGGC	16,318
PL13298	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,319
PL13299	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCT GAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,320
PL13300	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctt ccc	16,321
PL13301	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctt cccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,322
PL13302	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctt ccctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,323
PL13303	GACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctt ccctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTG AACCCAGGATAACCCTCAAAGTCGGGGGGC	16,324
PL13304	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcACATGAGGATCACCCATGTgc	16,325
PL13305	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGAAC CCAGGATAACCCTCAAAGTCGGGGGGC	16,326
PL13306	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,327
PL13307	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTC TCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,328
PL13308	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccct ctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,329
PL13309	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccct ctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCC AGGATAACCCTCAAAGTCGGGGGGC	16,330
PL13310	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,33
PL13311	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCC CCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,33
PL13312	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGg agtctttccc	16,33
PL13313	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGg agtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,33
PL13314	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGg agtctttcccctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,33

[1872]

PL13315	GACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGg agtcttccctctctctctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCC CCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,336
PL13316	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgc	16,337
PL13317	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAG GATAACCCTCAAAGTCGGGGGGGC	16,338
PL13318	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,339
PL13319	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCT CGGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,340
PL13320	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctct ctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,341
PL13321	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctct ctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc ATAACCCTCAAAGTCGGGGGGGC	16,342
PL13322	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgc	16,343
PL13323	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCTCGGGTCAGGAGCCCCCCCCT GAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,344
PL13324	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttc cc	16,345
PL13325	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttc ccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,346
PL13326	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttc cctctctctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,347
PL13327	GACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttc cctctctctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGA ACCCAGGATAACCCTCAAAGTCGGGGGGC	16,348
PL13328	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGA TCACCCATGTgc	16,34
PL13329	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGA TCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,35
PL13330	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGA GGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,35
PL13331	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGA GGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGG GGC	16,35

[1873]

n_13332	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGA GGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTG GGgagtcttccc	16,353
_13333	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGA GGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTG GGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,354
PL13334	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcAC ATGAGGATCACCCATGTgc	16,355
PL13335	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcAC ATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,356
PL13336	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAggga gcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,357
PL13337	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAggga gcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAA GTCGGGGGGC	16,358
PL13338	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAggga gcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCAC AGTCACTGGGgagtctttccc	16,359
PL13339	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAggga gcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCAC AGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,360
PL13340	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATG AGGATCACCCATGTgc	16,361
PL13341	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATG AGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,362
PL13342	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcA CATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,363
PL13343	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcA CATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGT CGGGGGGC	16,364
PL13344	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcA CATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGT CACTGGGgagtctttccc	16,365

[1874]

PL13345	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAggaggcA CATGAGGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGT CACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,366
13346	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACAT AgcACATGAGGATCACCCATGTgc	16,367
PL13347	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACAT AgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,368
PL13348	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACAT AgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,369
PL13349	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACAT AgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCC TCAAAGTCGGGGGGGC	16,370
PL13350	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACAT AgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactc CCACAGTCACTGGGgagtctttccc	16,371
PL13351	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACAT AgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactc CCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,372
PL13352	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcA CATGAGGATCACCCATGTgc	16,373
PL13353	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcA CATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,374
PL13354	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAggg agcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,375
PL13355	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAggg agcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAA AGTCGGGGGGGC	16,376
PL13356	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAggg agcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCA CAGTCACTGGGgagtctttccc	16,37
PL13357	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAggg agcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCA CAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,37
PL13358	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACT AACATAgcACATGAGGATCACCCATGTgc	16,37

[1875]

m13359	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACT AACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,380
13360	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACT AACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,381
13361	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACT AACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGAT AACCCTCAAAGTCGGGGGGGC	16,382
PL13362	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACT AACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctctgggagcACATGAGGATCACCCATGT gcgactcCCACAGTCACTGGGgagtctttccc	16,383
PL13363	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACT AACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctctgggagcACATGAGGATCACCCATGT gcgactcCCACAGTCACTGGGgagtctttccc TCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,384
PL13364	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATG AGGATCACCCATGTgc	16,385
PL13365	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATG AGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,386
PL13366	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcAC ATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,387
PL13367	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcAC ATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc TCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTC GGGGGGGC	16,388
PL13368	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcAC ATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctctgggagcACATGAGGATCACCCATGTC ACTGGGgagtctttccc	16,389
PL13369	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcAC ATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctctgggagcACATGAGGATCACCCATGTC ACTGGGgagtctttccc TCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,390
PL13370	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACAT AgcACATGAGGATCACCCATGTgc	16,391
PL13371	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACAT AgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,392
PL13372	GAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACAT AgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,393

[1876]

PL13373	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACAT AgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCC TCAAAGTCGGGGGGGC	16,394
13374	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACAT AgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactc CCACAGTCACTGGGgagtctttccc	16,395
PL13375	GAAGTTCATCTGTACCACCGGGCAAACTGCCCGTGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACAT AgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactc CCACAGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,396
PL13376	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGATCACCCATGTgc	16,397
PL13377	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGATCACCCATGTgcTCT CTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,398
PL13378	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTg cgactcCCACAGTCACTGGGgagtctttccc	16,399
PL13379	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTg cgactcCCACAGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,400
PL13380	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTg cgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,401
PL13381	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTg gactcCCACAGTCACTGGGgagtctttccctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctCTCT CTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,402
PL13382	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAACTAACTAACATAgcACATGAGGATCACCC ATGTgc	16,403
PL13383	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAACTAACTAACATAgcACATGAGGATCACCC ATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,404
PL13384	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAACTAACTAACATAgggagcACATGAGGATCA CCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,40
PL13385	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAACTAACTAACATAgggagcACATGAGGATCA CCCATGTgcgactcCCACAGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,40
PL13386	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAACTAACTAACATAgggagcACATGAGGATCA CCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtc ttccc	16,40
PL13387	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGACATACAACTAACTAACATAgggagcACATGAGGATCA CCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtc ttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,40

[1877]

[illegible]13388	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATG Tgc	16,409
_13389	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATG TgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,410
_13390	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAggggagcACATGAGGATCACCC ATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,411
PL13391	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCC ATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,412
PL13392	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCC ATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcc c	16,413
PL13393	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCC ATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcc cTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,414
PL13394	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgcACATGAGGAT CACCCATGTgc	16,415
PL13395	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgcACATGAGGAT CACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,416
PL13396	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAG GATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,417
PL13397	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAG GATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGG GC	16,418
PL13398	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAG GATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccctctctctctctctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGG Ggagtcttccc	16,419
PL13399	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAG GATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccctctctctctctctctctctctcgaggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGG Ggagtcttccc TCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,42
PL13400	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcACATGAGGATCACC CATGTgc	16,42
PL13401	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcACATGAGGATCACC CATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,42
PL13402	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATC ACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,42

[1878]

PL13403	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATC ACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,424
13404	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATC ACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagt cttccc	16,425
PL13405	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATC ACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagt cttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,426
PL13406	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGA GGATCACCCATGTgc	16,427
PL13407	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGA GGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,428
PL13408	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACA TGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,429
PL13409	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACA TGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCG GGGGGC	16,430
PL13410	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACA TGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCA CTGGGgagtctttccc	16,431
PL13411	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACA TGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCA CTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,432
PL13412	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGT gc	16,433
PL13413	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGT gcTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,43
PL13414	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCC ATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,43
PL13415	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCC ATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,43
PL13416	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCC ATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,43

[1879]

PL13417	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCC ATGTgcgactccCCACAGTCACTGGGgagtcttcccctctctctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcc cTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,438
_13418	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgcACATGAGGATC ACCCATGTgc	16,439
PL13419	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgcACATGAGGATC ACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,440
PL13420	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAG GATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttccc	16,441
PL13421	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAG GATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGG GC	16,442
PL13422	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAG GATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccctctctctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGG Ggagtcttccc	16,443
PL13423	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAG GATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccctctctctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGG Ggagtcttccc TCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,444
PL13424	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAAgcACATGAGGATCACCCATGTgc	16,445
PL13425	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAA AGTCGGGGGGC	16,446
PL13426	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttccc	16,447
PL13427	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAG CCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,44P
PL13428	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccctctctctctctctctcgggagc ACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttccc	16,44!
PL13429	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccctctctctctctctctcgggagc ACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAG TCGGGGGGC	16,45I

[1880]

PL13430	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC	16,451
13431	TGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgc CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATA ACCCTCAAAGTCGGGGGGGC	16,452
PL13432	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttccc	16,453
PL13433	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGG TCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,454
PL13434	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttcccctctctctctctc tctcgggagCACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttccc	16,455
PL13435	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttcccctctctctctc tctcgggagCACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAAC CCTCAAAGTCGGGGGGGC	16,456
PL13436	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATGTgc	16,457
PL13437	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCC TCAAAGTCGGGGGGGC	16,458
PL13438	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttccc	16,459
PL13439	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCA GGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,460
PL13440	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttcccctctctctctctc gggagcACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttccc	16,46
PL13441	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttcccctctctctctctc gggagcACATGAGGATCACCCATGTgcactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCT CAAAGTCGGGGGGGC	16,46

[1881]

m_13442	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,463
_13443	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAG GATAACCCTCAAAGTCGGGGGGGC	16,464
PL13444	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,465
PL13445	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCT CGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,466
PL13446	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCT ctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,467
PL13447	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCT ctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc ATAACCCTCAAAGTCGGGGGGGC	16,468
PL13448	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcACATGAGGATCACCCATGTgc	16,469
PL13450	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,471
PL13451	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGG GTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,472
PL13452	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccctctctctctctc tctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,473
PL13453	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccctctctctctc tctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAA CCCTCAAAGTCGGGGGGGC	16,47
PL13454	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,47

n13455	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGAAC CCAGGATAACCCTCAAAGTCGGGGGGC	16,476
13456	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,477
PL13457	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTC TCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,478
PL13458	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctct ctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCC AGGGATAACCCTCAAAGTCGGGGGGGC	16,479
PL13459	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctct ctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCC AGGGATAACCCTCAAAGTCGGGGGGGC	16,480
PL13460	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgc	16,481
PL13461	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCC TCAAAGTCGGGGGGGC	16,482
PL13462	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,483
PL13463	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCA GGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,484
PL13464	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccctctctctctctctct

[1883]

PL13467	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAG GATAACCCTCAAAGTCGGGGGGC	16,488
13468	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,489
PL13469	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCT CGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,490
PL13470	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCT ctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,491
PL13471	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAGAAGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCC TGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCT ctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGG ATAACCCTCAAAGTCGGGGGGC	16,492
PL13472	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGATCACCCATGTgc	16,493
PL13473	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACC CTCAAAGTCGGGGGGC	16,494
PL13474	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,495
PL13475	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTC AGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,496
PL13476	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCTCT gggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,497
PL13477	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCTCT gggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCC TCAAAGTCGGGGGGC	16,49
PL13478	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,49

[1884]

n 13479	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCA GGATAACCCTCAAAGTCGGGGGGGC	16,500
_13480	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,501
PL13481	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccTCTCT CTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGGATAACCCTCAAAGTCGGGGGGGC	16,502
PL13482	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccctctct ctctctctctctcgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,503
PL13483	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccctctct ctctctctctctcgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAG GATAACCCTCAAAGTCGGGGGGGC	16,504
PL13484	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATGTgc	16,505
PL13485	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGGA TAACCCTCAAAGTCGGGGGGGC	16,506
PL13486	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,507
PL13487	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccTCTCTCTC GGGTCAGGAGCCCCCCCCCTGAACCCAGGGATAACCCTCAAAGTCGGGGGGGC	16,508
PL13488	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccctctctctc tctctctgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,509
PL13489	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccctctctctc tctctctgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATA ACCCTCAAAGTCGGGGGGGC	16,511
PL13490	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,51

[1885]

PL13491	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGA ACCCAGGATAACCCTCAAAGTCGGGGGGC	16,512
13492	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,513
PL13493	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc TCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,514
PL13494	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc tctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,515
PL13495	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc tctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccCTCTCTCTCGGGTCAGGAGCCCCCCCCTGAAC CCAGGATAACCCTCAAAGTCGGGGGGC	16,516
PL13496	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcACATGAGGATCACCCATGTgc	16,517
PL13497	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCC AGGATAACCCTCAAAGTCGGGGGGC	16,518
PL13498	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,519
PL13499	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTC TCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,520
PL13500	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccctctc tctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,521
PL13501	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccctctc tctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAG GATAACCCTCAAAGTCGGGGGGC	16,52
PL13502	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,52

[1886]

n_13503	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCC CTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,524
_13504	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCACAGTCACTGGGgagt cttccc	16,525
PL13505	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCACAGTCACTGGGgagt cttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,526
PL13506	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCACAGTCACTGGGgagt cttcccttctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttccc	16,527
PL13507	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCACAGTCACTGGGgagt cttcccttctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCT GAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,528
PL13508	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgc	16,529
PL13509	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGA TAACCCTCAAAGTCGGGGGGC	16,530
PL13510	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttccc	16,531
PL13511	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccTCTCTCTCG GGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,532
PL13512	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccttctctctctc tctctctgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttccc	16,533
PL13513	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccttctctctc tctctctgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATA ACCCTCAAAGTCGGGGGGC	16,534
PL13514	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,535

[1887]

PL13515	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGA ACCCAGGATAACCCTCAAAGTCGGGGGGC	16,536
13516	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCACAGTCACTGGGgagtctttccc	16,537
PL13517	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCACAGTCACTGGGgagtctttccc TCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,538
PL13519	CAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA CCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactccCACAGTCACTGGGgagtctttccct ctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAAC CCAGGATAACCCTCAAAGTCGGGGGGC	16,540
PL13520	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGATC ACCCATGTgc	16,541
PL13521	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGATC ACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,542
PL13522	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGG ATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccc	16,543
PL13523	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGG ATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGG C	16,544
PL13524	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGG ATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccctctctctctctctctctctctctgggagcACATGAGGATCACCCATGTCACTGGG gagtctttccc	16,545
PL13525	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgggagcACATGAGG ATCACCCATGTgcgactccCCACAGTCACTGGGgagtctttccctctctctctctctctctctctctgggagcACATGAGGATCACCCATGTCACTGGG gagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,546
PL13526	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACAT GAGGATCACCCATGTgc	16,54

[1888]

PL13527	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACAT GAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,548
13528	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagc ACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,549
PL13529	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagc ACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAG TCGGGGGGC	16,550
PL13530	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgggagc ACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccctctctctctctctctctctctgcgagcACATGAGGATCACCCATGTgcgactcCCACAG TCACTGGGgagtcttccc	16,551
PL13532	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAG GATCACCCATGTgc	16,553
PL13533	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAG GATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,554
PL13534	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACAT GAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,555
PL13535	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACAT GAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGG GGGGC	16,556
PL13536	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACAT GAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctctgcgagcACATGAGGATCACCCATGTgcgactcCCACAGTCAC TGGGgagtcttccc	16,55
PL13537	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgggagcACAT GAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctctgcgagcACATGAGGATCACCCATGTgcgactcCCACAGTCAC TGGGgagtcttccCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,55i

[1889]

PL13538	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAg cACATGAGGATCACCCATGTgc	16,559
13539	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAg cACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,560
PL13540	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAg ggagCACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,561
PL13541	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAg ggagCACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTC AAAGTCGGGGGGGC	16,562
PL13542	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAg ggagCACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTC ACAGTCACTGGGgagtcttccc	16,563
PL13543	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAACTAACATAg ggagCACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTC ACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,564
PL13544	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcACA TGAGGATCACCCATGT	16,565
PL13545	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgcACA TGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,566
PL13546	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggag cACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,56
PL13547	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggag cACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAA GTCGGGGGGGC	16,56i
PL13548	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggag	16,56i

[1890]

	cACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctcgaggagcACATGAGGATCACCCATGTggactcCCACA GTCACTGGGgagtctttccc	
13549	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAAgggag cACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctcgaggagcACATGAGGATCACCCATGTggactcCCACA GTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,570
PL13550	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAAC ATAgcACATGAGGATCACCCATGTgc	16,571
PL13551	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAAC ATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,572
PL13552	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAAC ATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,573
PL13553	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAAC ATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACC CCTCAAAGTCGGGGGGGC	16,574
PL13554	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAAC ATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACC ctcCCACAGTCACTGGGgagtctttccc	16,575
PL13555	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACACCCTGACGTACGGCGTGCAGTGCTTCGGCACATACAACTAAC ATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC ctcCCACAGTCACTGGGgagtctttccc	16,576
PL13556	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAG GATCACCCATGTgc	16,57
PL13557	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAG GATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,571
PL13558	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACAT GAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,57

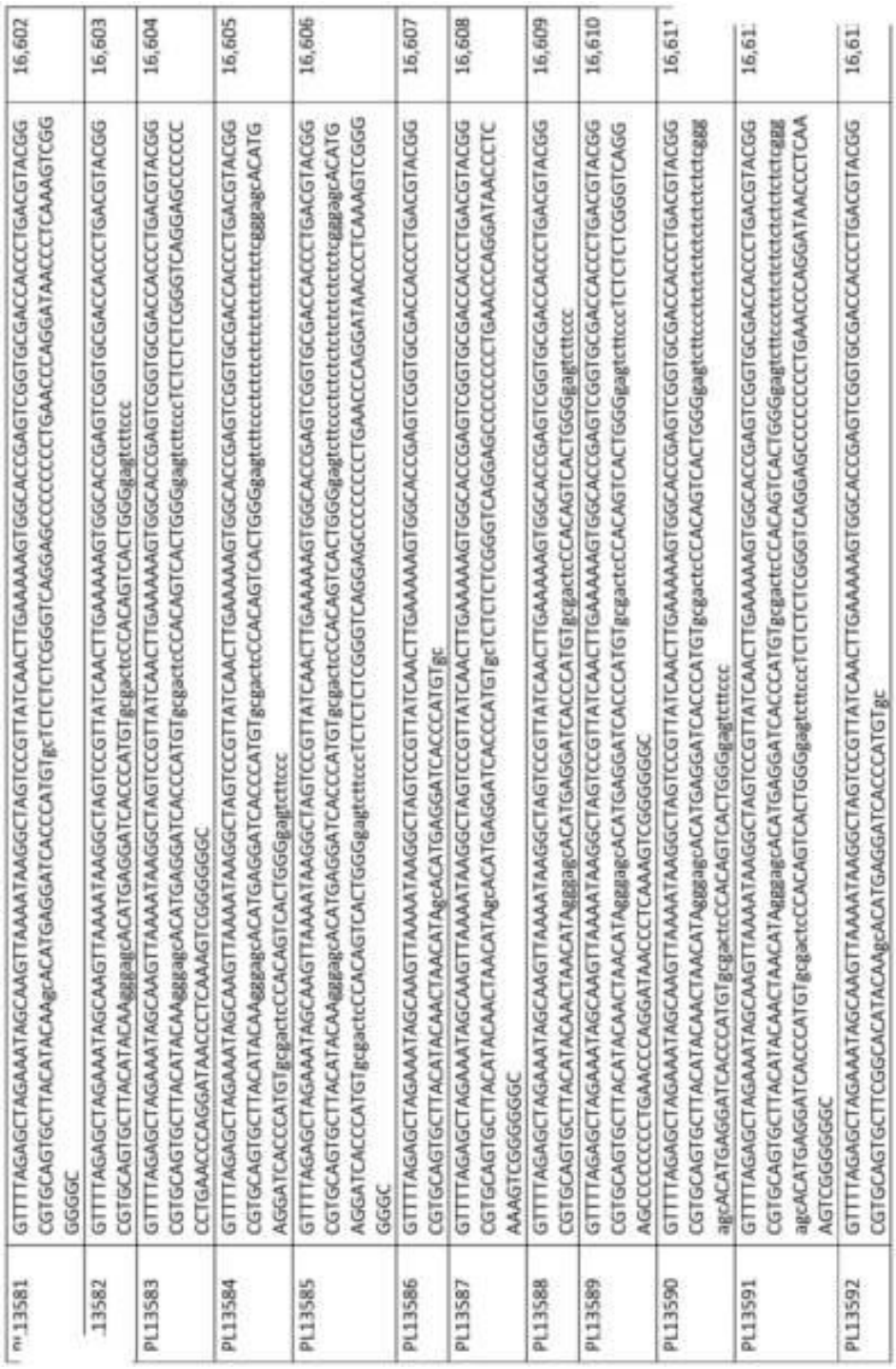
[1891]

m_13559	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAggaigcACAT GAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGG GGGGC	16,580
_13560	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAggaigcACAT GAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGG TGGGgagtctttccc	16,581
PL13561	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAggaigcACAT GAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGG TGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,582
PL13562	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAg cACATGAGGATCACCCATGTgc	16,583
PL13563	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAg cACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,584
PL13564	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAg ggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTC	16,585
PL13565	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAg ggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTC AAAGTCGGGGGGGC	16,586
PL13566	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAg ggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTC ACAGTCACTGGGgagtctttccc	16,587
PL13567	AGGGCGATGCCACCTAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGA AGTTCATCTGTACCACCGGCAAACTGCCCCGTGCCTTGGCCCCACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAg ggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTC ACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,588
PL13568	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAAgcACATGAGGATCACCCATGTgc	16,589

[1892]

PL13569	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,590
_13570	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,591
PL13571	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGA ACCCAGGATAACCCTCAAAGTCGGGGGGC	16,592
PL13572	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC CACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,593
PL13573	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC CACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,594
PL13574	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,595
PL13575	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGT CGGGGGGC	16,596
PL13576	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,597
PL13577	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCC CCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,598
PL13578	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCC ATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,599
PL13579	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCC ATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTC GGGGGGC	16,600
PL13580	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATGTgc	16,601

[1893]

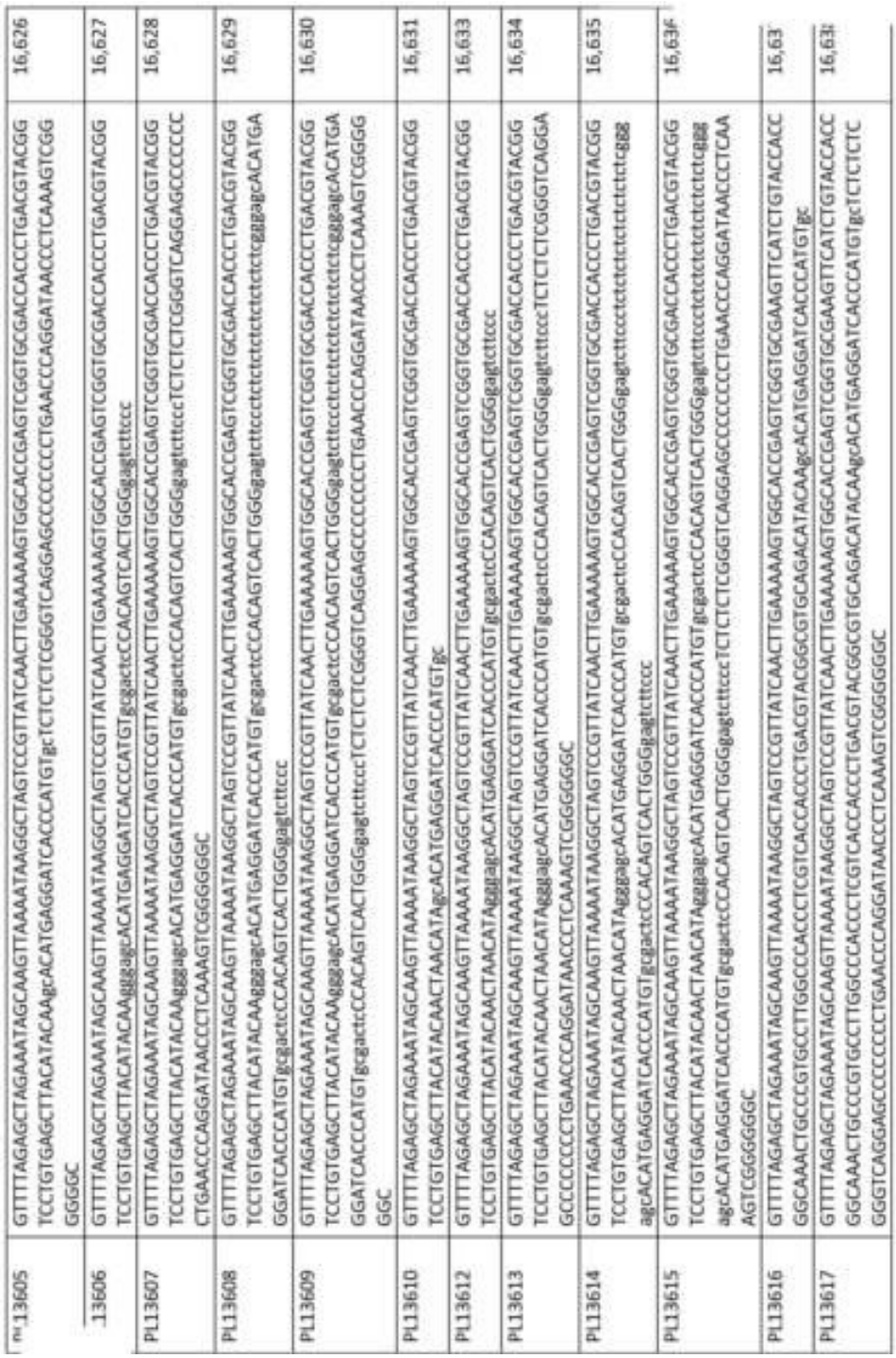


PL13581	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGG GGGGGC	16,602
_13582	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,603
PL13583	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCC CCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,604
PL13584	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctgggagcACATG AGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,605
PL13585	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctgggagcACATG AGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGG GGGC	16,606
PL13586	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,607
PL13587	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTC AAAGTCGGGGGGC	16,608
PL13588	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,609
PL13589	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGG AGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,610
PL13590	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccctctctctctctctctctgggg agcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,611
PL13591	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccctctctctctctctctctgggg agcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAA AGTCGGGGGGC	16,612
PL13592	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAAgcACATGAGGATCACCCATGTgc	16,613

[1894]

[illegible]13593	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAG TCGGGGGGC	16,614
13594	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,615
P113595	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCC CCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,616
P113596	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcctctctctctctctctcgggagcA CATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,617
P113597	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcctctctctctctctctcgggagcA CATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGT CGGGGGGC	16,618
P113598	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,619
P113599	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGAGGATCACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAAC CCTCAAAGTCGGGGGGC	16,620
P113600	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,621
P113601	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGGT CAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,622
P113602	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcctctctctctctctctc tcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcc	16,623
P113603	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG CGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcctctctctctctctctc tcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCC TCAAAGTCGGGGGGC	16,624
P113604	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGACCACCCTGACGTACGG TCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgc	16,625

[1895]



[1896]

[illegible]13618	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAggagcACATGAGGATCACCCATGTgcgactc CCACAGTCACTGGGgagtctttccc	16,639
_13619	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAggagcACATGAGGATCACCCATGTgcgactc CCACAGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,640
PL13620	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAggagcACATGAGGATCACCCATGTgcgactc CCACAGTCACTGGGgagtctttccctctctctctctctctctcggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccc	16,641
PL13621	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAggagcACATGAGGATCACCCATGTgcgactc CCACAGTCACTGGGgagtctttccctctctctctctctctctcggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctCTCTCTCG GGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,642
PL13622	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgc	16,643
PL13623	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgc TCTCTCTCGGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,644
PL13624	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgc GTgcgactcCCACAGTCACTGGGgagtctttccc	16,645
PL13625	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgc GTgcgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,646
PL13626	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgc GTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctcggagcACATGAGGATCACCCATGTcgactcCCACAGTCACTGGGgagtctttccc	16,647
PL13627	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAACTAACATAgcACATGAGGATCACCCATGTgc GTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctcggagcACATGAGGATCACCCATGTcgactcCCACAGTCACTGGGgagtctttccctCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,648
PL13628	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGACATACAAgcACATGAGGATCACCCATGTgc	16,649
PL13629	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAgcACATGAGGATCACCCATGTgcTCT CTCTCGGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,650

[1897]

[illegible]_13630	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAggagcACATGAGGATCACCCATGTg cgactcCCACAGTCACTGGGgagtctttccc	16,651
_13631	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAggagcACATGAGGATCACCCATGTg cgactcCCACAGTCACTGGGgagtctttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,652
PL13632	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGTACGGCGTGCAGTGCTTACATACAAggagcACATGAGGATCACCCATGTg cgactcCCACAGTCACTGGGgagtctttcccTCTCTCTC

[1898]

m_13641	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAAgcACATGAGGATCACCCATGT gcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,662
_13642	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCA TGTgcgactcCCACAGTCACTGGGgagtcttccc	16,663
PL13643	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCA TGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,664
PL13644	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCA TGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,665
PL13645	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAAgggagcACATGAGGATCACCCA TGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc CTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,666
PL13646	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGAGGATC ACCCATGTgc	16,667
PL13647	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgcACATGAGGATC ACCCATGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGGC	16,668
PL13648	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGG ATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,669
PL13649	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGG ATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGG C	16,670
PL13650	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGG ATCACCCATGTgcgactcCCACAGTCACTGGGgagtctttccctctctctctctctctctctctgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGG gagtcttccc	16,67
PL13651	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCCGTGCCTTGGCCCACCCTCGTCACCACCCTGACGGTACGGCGTGCAGTGCTTCGGCACATACAACTAACATAgggagcACATGAGG	16,67

[1899]

	ATCACCCATGTgcgactcCCACAGTCACTGGGgagtcttccctctctctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGG gagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	
_13652	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgc	16,673
_13653	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgcACATGAGGATCACCCATGTgcTCT CTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,674
PL13654	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgc gactcCCACAGTCACTGGGgagtcttccc	16,675
PL13655	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgc gactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,676
PL13656	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgc gactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,677
PL13657	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAAgggagcACATGAGGATCACCCATGTgc gactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC CTCGGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,678
PL13658	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgcACATGAGGATCACCCA TGTgc	16,679
PL13659	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgcACATGAGGATCACCCA TGTgcTCTCTCTCGGGTCAGGAGCCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,680
PL13660	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAGGATCA CCCATGTgcgactcCCACAGTCACTGGGgagtcttccc	16,681
PL13661	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAGGATCA CCCATGTgcgactcCCACAGTCACTGGGgagtcttcccTCTCTCTCGGGTCAGGAGCCCCCCCCTGAACCCAGGATAACCCTCAAAGTCGGGGGGC	16,682
PL13662	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGAAGTTCATCTGTACCACC GGCAAACTGCCCGTGCCTTGGCCCACACCCTCGTCACCACCCTGACGTACGGTCCTGTGAGCTTACATACAACTAACATAgggagcACATGAGGATCA CCCATGTgcgactcCCACAGTCACTGGGgagtcttcccttctctctctctctctctctctcgggagcACATGAGGATCACCCATGTgcgactcCCACAGTCACTGGGgagtc ttccc	16,683

[1900]

[1901] **실시예 7: 시스 주형 RNA를 사용한 예시적인 유전자 변형 폴리펩티드의 재작성 활성의 평가**

[1902] 이 실시예는 벤치마크 시스 주형 RNA를 사용하여 실시예 6에서 생성되고 예시적인 유전자 변형 폴리펩티드 PLV4921에 대해 정규화된 예시적인 유전자 변형 시스템의 재작성 활성을 입증한다.

[1903] 트랜스 주형을 사용한 실시예 6의 유전자 변형 폴리펩티드의 재작성 활성을 평가하기 전에, 가이드 RNA와 PBS 및 주형 서열을 포함하는 3'-연장부를 둘 다 함유하는 '시스' 주형 RNA를 사용하여 214개의 폴리펩티드를 평가하였으며, 여기서 주형은 상동성 1, 편집물 및 상동성 2로 구성되어 있다. 각각의 예시적인 유전자 변형 시스템(예시적인 유전자 변형 폴리펩티드 및 '시스' 주형 RNA의 조합)을 시험하여 BFP-GFP SNP 리포터를 발현하는 게놈 랜딩 패드를 보유하는 포유동물 U2OS 세포를 사용하여 게놈-편집 능력을 결정하였다. U2OS 세포는 BFP-발현 세포이고 시스 주형 RNA에 의해 프로그래밍된 단일 염기쌍 치환을 통해 GFP-발현 세포로 전환될 수 있다. U2OS 세포를 예시적인 개별 유전자 변형 폴리펩티드를 인코딩하는 플라스미드 및 시스 주형 RNA를 인코딩하는 제2 플라스미드의 형태로 각각의 예시적인 유전자 발현 시스템을 이용하여 전기천공하였다.

[1904] 유전자 변형 시스템의 전기천공 4일 후 유세포 분석법으로 세포를 분석하였다. GFP 세포의 백분율을 재작성 활성에 대한 척도로 사용하였고 예시적인 유전자 변형 폴리펩티드 및 시스 주형 RNA를 함유하는 벤치마크 유전자 변형 시스템의 재작성 활성과 비교하였다(예를 들어, 도 7a에 나타낸 작업 흐름 참조).

[1905] 벤치마크 유전자 변형 폴리펩티드 및 시스 주형 RNA의 서열은 여기에 제공되어 있다:

[1906] PLV4921:

MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF
DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHP
IFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV
DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALS
LGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVN
TEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFY
KFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNR
EKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN
LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK
EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDR
EMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF
MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPE
NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRD
MYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQ
LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKL
IREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY
KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG
RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY
SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEI
IEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDR
KRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSS
GGSSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVS
IKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVED
IHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRL
PQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGY
RASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPG
FAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGV
LTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALV
KQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPD
LTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKM
AEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCP
GHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSEFEKRTADGSEF
ESPKKKAKVE (서열번호 16,709)

[1907]

[1908] PLV4114

5'-
AGAAGTCGTGCTGCTTCATGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCG
TTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCACCCTGACCTACGGCGTGCAGTGCTTCG
[1909] GCCGCTACCCCGATCACATGAAGCAGCACGAC -3' (서열번호 16,710)

[1910] 도 7b는 선택된 폴리펩티드 구성(도 6a 참조)을 갖고 1, 2, 또는 4개의 MCP 반복부를 포함하는 예시적인 RBP 도메인을 포함하는 선택된 예시적인 유전자 변형 폴리펩티드에 대한 재작성 활성(PLV4921에 대해 정규화됨)의 그래프를 나타낸다. 예시적인 유전자 변형 폴리펩티드는 제1 예시적인 RT 도메인(왼쪽 음영 막대) 또는 제2 예시적인 RT 도메인(오른쪽 밝은 음영 막대)을 포함하였다.

[1911] 도 7b에 나타낸 예시적인 유전자 변형 폴리펩티드:

[1912]

제1 예시적인 RT를 함유하는 드라이버 (인코딩하는 플라스미드의 ID)	제2 예시적인 RT를 함유하는 드라이버 (인코딩하는 플라스미드의 ID)
PL11941	PL11989
PL11944	PL11992
PL11981	PL12018
PL11945	PL11993
PL11949	PL12034
PL11973	PL11998
PL12109	PL12037
PL12113	PL12053
PL12115	
PL12117	PL12045
PL12120	PL12047
PL12085	PL12097
PL12088	
PL12086	PL12101
PL12095	PL12104
PL12087	PL12108
PL12096	PL12099

[1913] 결과는 시험한 모든 구성, MCP 반복부의 수, 및 RT 도메인 둘 다를 포함하는 많은 예시적인 유전자 변형 폴리펩티드가 U2OS 세포에서 BFP에서 GFP로의 재작성 활성을 나타내고 이들의 재작성 활성이 시스 주형 RNA와 조합될 때 벤치마크 유전자 변형 폴리펩티드 PLV4921의 재작성 활성과 비슷하다는 것을 나타낸다. 추가로, 실험은 시스 주형 RNA와 조합될 때 활성이 벤치마크 유전자 변형 폴리펩티드의 75% 초과인 36개의 예시적인 유전자 변형 폴리펩티드, 및 이들 중 활성이 벤치마크의 90% 초과인 12개의 예시적인 유전자 변형 폴리펩티드를 확인하였으며; 예시적인 유전자 변형 폴리펩티드의 서열은 하기 표 X0에 제공되어 있다. 시스 주형 RNA와의 활성은 예시적인 유전자 변형 폴리펩티드가 활성이고 잠재적으로 트랜스 주형 RNA를 사용하여 재작성할 수 있음을 입증한다.

[1914] PLV4921에 비해 >75% 재작성 활성을 갖는 예시적인 유전자 변형 폴리펩티드가 표 X0에 열거되어 있다.

[1915] *표 X0: PLV4921에 비해 >75% 재작성 활성을 갖는 예시적인 유전자 변형 폴리펩티드*

드라이버(인코딩하는 플라스미드의 ID)	재작성 활성(PLV4921에 대해 정규화됨)
PL12086	1.057
PL11941	1.039
PL11957	1.035
PL11981	1.026
PL12089	0.956
PL11949	0.941
PL12095	0.921
PL12092	0.917
PL12018	0.889
PL11945	0.885
PL11993	0.885
PL11973	0.865
PL11994	0.849
PL12088	0.847
PL11996	0.842
PL12053	0.838
PL11995	0.827
PL11982	0.820
PL12034	0.820
PL11958	0.813
PL11947	0.811
PL12085	0.801
PL12101	0.801
PL12029	0.800
PL11984	0.789
PL11983	0.773
PL11950	0.771
PL12006	0.771
PL11960	0.762
PL12020	0.758
PL11944	0.758
PL12104	0.757
PL12113	0.755
PL11969	0.753
PL12017	0.753
PL11985	0.746

[1916]

[1917] **실시예 9: 예시적인 유전자 변형 폴리펩티드와 예시적인 트랜스 주형 풀의 조합 특성화**

[1918] 이 실시예는 실시예 6에서 생성되고 실시예 7에서 시스 주형 RNA와 조합될 때 재작성이 가능한 것으로 결정된 예시적인 유전자 변형 폴리펩티드 중 대략 18개의 재작성 활성 및 트랜스 주형 선호도를 특성화한다. 하기 열거된 18개의 예시적인 유전자 변형 폴리펩티드(실시예 6에서 제공된 아미노산 서열) 각각은 서열번호 16,711의 아미노산 서열을 포함하는 RT 도메인 및 서열번호 16,712의 아미노산 서열을 포함하는 Cas9 엔도뉴클레아제 도메인을 포함한다.

TAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEA
KRSLRETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHPTVPNPYTLLSLLP
PDRIWYSVLDLKDAFFCIPLAPESQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGF
RLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGKKAQLCQEEVTYLGFKIHKGSRSL
SNSRTQAILQIPVPKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAFQSLKL
ALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALL
TREASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFKQTAALNPATLLPETDD
TLPIHHCLDTLDSLTSTRPDLTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSA
QKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERGWLTAGGKAIKNAPEILALLTAVWLPK
RVAVMHCKGHQKDDAPTSTGNRRADEVAREVAIRPLSTQATIS (서열번호 16,711)

[1919]

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR
YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKK
LVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI
LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQ
IGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIF
FDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHA
ILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF
IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTV
KQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREM
IEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDS
LTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQ
KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI
VPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSEL
DKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNY
HHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEIT
LANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIA
RKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVK
KDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTI
DRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (서열번호 16,712)

[1920]

[1921] 예시적인 유전자 변형 폴리펩티드는 다음과 같다:

구성 세부정보	드라이버(인코딩하는 플라스미드의 ID)
구성1_1MCP A	PL11941
구성1_1MCP B	PL11944
구성1_2MCP A	PL11981
구성1_2MCP B	PL11945
구성1_4MCP A	PL11949
구성1_4MCP B	PL11973
구성2_1MCP A	PL12109
구성2_2MCP A	PL12113
구성2_2MCP B	PL12115
구성2_4MCP A	PL12117
구성2_4MCP B	PL12120
구성3_1MCP A	PL12085
구성3_1MCP B	PL12088
구성3_2MCP A	PL12086
구성3_2MCP B	PL12095
구성3_4MCP A	PL12087
구성3_4MCP B	PL12096
구성2_1MCP A	PL12037

[1922]

[1923] 예시적인 유전자 변형 폴리펩티드는 실시예 7에서 선택된 각각의 구성으로부터 6개를 포함하고 1, 2 또는 4개의 RBP 반복부 및 3개의 단백질 도메인 각각을 연결하는 2개의 상이한 링커 조합을 갖는 예시를 포함하였다.

[1924] 풀의 예시적인 트랜스 주형 RNA는 8 또는 16개 뉴클레오티드의 링커; 1, 2 또는 4개의 RRS 반복부(MS2 서열); 8 또는 13개 뉴클레오티드의 PBS(또는 음성 대조군의 경우 스크램블드 무작위 PBS); 및 16개의 뉴클레오티드 스페이서가 있는 5' 말단-보호 Cas9 스캐폴드 또는 5' 말단 보호 특징부가 없는 것을 갖는 예시를 포함하였다.

풀 번호	플라스미드(등록 ID)
1xMS2_PBS8	PL13280
1xMS2_PBS8	PL13472
1xMS2_PBS8	PL13286
1xMS2_PBS8	PL13478
2xMS2_PBS8	PL13282
2xMS2_PBS8	PL13474
2xMS2_PBS8	PL13288
2xMS2_PBS8	PL13480
4xMS2_PBS8	PL13284
4xMS2_PBS8	PL13476
4xMS2_PBS8	PL13290
4xMS2_PBS8	PL13578
1xMS2_PBS13	PL13292
1xMS2_PBS13	PL13484
1xMS2_PBS13	PL13298
1xMS2_PBS13	PL13490
2xMS2_PBS13	PL13294
2xMS2_PBS13	PL13582
2xMS2_PBS13	PL13300
2xMS2_PBS13	PL13492
4xMS2_PBS13	PL13296
4xMS2_PBS13	PL13488
4xMS2_PBS13	PL13302
4xMS2_PBS13	PL13494
1xMS2_scrPBS	PL13316
1xMS2_scrPBS	PL13604
1xMS2_scrPBS	PL13322
1xMS2_scrPBS	PL13610
2xMS2_scrPBS	PL13318
2xMS2_scrPBS	PL13606
2xMS2_scrPBS	PL13324
2xMS2_scrPBS	PL13612
4xMS2_scrPBS	PL13320
4xMS2_scrPBS	PL13512
4xMS2_scrPBS	PL13326

[1925]

[1926] 이 실시예에 기재된 예시적인 유전자 변형 시스템의 재작성 활성은 예시적인 유전자 변형 폴리펩티드(PLV4921) 및 동일한 편집물을 인코딩하는 시스 주형 RNA(PLV4106)를 함유하는 벤치마크 유전자 변형 시스템에 대해 정규화되었다. PLV4921의 서열은 실시예 7에서 제공되었으며, 이 실시예에서 정규화에 사용된 시스-주형의 서열은 하기에 제공되어 있다:

[1927] PLV4106

[1928] GCCGAAGCACTGCACGCCGTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC
TTGAAAAAGTGGCACCGAGTCGGTGCACCCTGACGTACGGCGTGCAGTGCTT (서열번호 16,713)

[1929] 유전자 변형 시스템의 제3 구성요소를 구성하는 gRNA의 서열은 하기와 같다:

[1930] PLV4165

[1931] GCCGAAGCACTGCACGCCGTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC TTGAAAAAGTGGCACCGAGTCGGTGC (서열번호 16,714)

[1932] 각각의 예시적인 트랜스-유전자 변형 시스템(예시적인 유전자 변형 폴리펩티드, 예시적인 트랜스 주형 RNA, 및 예시적인 gRNA의 조합)을 시험하여 BFP-GFP SNP 리포터를 발현하는 게놈 랜딩 패드를 보유하는 포유동물 U2OS 세포를 사용하여 게놈-편집 능력을 결정하였다. U2OS 세포는 BFP-발현 세포이고 트랜스 주형 RNA에 의해 프로그래밍된 단일 염기쌍 치환을 통해 GFP-발현 세포로 전환될 수 있다. U2OS 세포를 예시적인 개별 유전자 변형 폴리펩티드를 인코딩하는 플라스미드, 예시적인 gRNA를 인코딩하는 제2 플라스미드, 및 예시적인 트랜스 주형 RNA의 서브세트를 인코딩하는 플라스미드의 풀의 형태로 각각의 예시적인 트랜스-유전자 발현 시스템을 이용하여 전기천공하였다. 주어진 풀의 트랜스 주형 RNA는 유사한 프라이머 결합 부위(PBS) 길이(8 또는 13개의 뉴클레오티드)와 동일한 수의 MS2 서열(1, 2 또는 4개)을 함유하는 RBP 동원 부위(RRS)를 갖는 것을 특징으로 하였다. 음성 대조군 풀은 스크램블드 서열 PBS를 함유하는 트랜스 주형 RNA를 함유하였다.

[1933] 유전자 변형 시스템의 전기천공 4일 후 유세포 분석법으로 세포를 분석하였다. GFP 세포의 백분율을 재작성 활성에 대한 척도로 사용하였고 예시적인 유전자 변형 폴리펩티드(PLV4921) 및 동일한 편집물을 인코딩하는 시스 주형 RNA를 함유하는 벤치마크 유전자 변형 시스템의 재작성 활성과 비교하였다.

[1934] 도 8은 BFP/GFP SNP 리포터 분석에서 예시적인 트랜스-유전자 변형 시스템(시스 주형 RNA를 이용하여 PLV4921에 대해 정규화됨)의 재작성 활성에 대한 히트 맵을 나타낸다. 결과는 트랜스 주형 RNA의 여러 풀이 시스 주형 RNA와 비슷하거나 더 나은 재작성 활성을 제공한다는 것을 나타낸다. 이러한 결과로 상이한 특성의 트랜스 주형에 걸쳐 조사된 18개의 예시적인 유전자 변형 폴리펩티드의 재작성 활성이 추가로 확인되며, 또한 예시적인 유전자 변형 폴리펩티드를 트랜스 주형과 쌍을 형성함으로써 재작성 활성이 개선될 수 있음을 입증한다. 결과는 또한 평균 13개 뉴클레오티드 길이의 PBS를 함유하는 트랜스 주형 RNA가 8개 뉴클레오티드 PBS를 함유하는 트랜스 주형 RNA보다 평가된 유전자 변형 폴리펩티드에 걸쳐 더 나은 재작성 활성을 제공한다는 것을 시사하였다. 결과는 또한 2 또는 4개의 RBP 반복부(이 실시예에서는 2 또는 4개의 MCP 도메인)를 함유하는 유전자 변형 폴리펩티드가 일반적으로 1개의 RBP(1개의 MCP 도메인)를 함유하는 유전자 변형 폴리펩티드보다 더 높은 재작성 활성을 갖는다는 것을 시사하였다. 결과는 또한 1 또는 4개의 RRS 반복부(이 실시예에서는 1 또는 4개의 MS2 서열)을 함유하는 트랜스 주형 RNA가 일반적으로 2개의 RRS 반복부(2개의 MS2 서열)를 함유하는 트랜스 주형 RNA보다 평가된 유전자 변형 폴리펩티드에 걸쳐 더 나은 재작성 활성을 제공한다는 것을 시사하였다. 데이터는 유전자 변형 폴리펩티드의 시험된 각각의 구성이 재작성 활성을 가지고 있음을 반복한다.

[1935] **실시예 10: 예시적인 유전자 변형 폴리펩티드와 예시적인 개별 트랜스 주형의 조합 특성화**

[1936] 이 실시예는 실시예 6에서 생성되고 실시예 7 및 9에서 재작성 및 트랜스 동원이 가능한 것으로 결정된 예시적인 유전자 변형 폴리펩티드 중 6개의 재작성 활성 및 트랜스 주형 선호도를 특성화한다. 하기 열거된 6개의 예시적인 유전자 변형 폴리펩티드(실시예 6에서 제공된 아미노산 서열) 각각은 서열번호 16,711의 아미노산 서열을 포함하는 RT 도메인 및 서열번호 16,712의 아미노산 서열을 포함하는 Cas9 엔도뉴클레아제 도메인을 포함한다. 실시예 6에서 생성된 384개 중 29개의 예시적인 트랜스 주형 RNA를 유세포 분석법(본 명세서에 기재된 바와 같이 BFP에서 GFP로의 전환을 평가함)을 사용하여 시험하였다.

[1937] 예시적인 유전자 변형 폴리펩티드는 다음과 같다: PL11945, PL12095, PL12096, PL12109, PL12113, 및 PL12115.

[1938] 예시적인 유전자 변형 폴리펩티드는 실시예 7 및 9에서 선택된 각각의 구성으로부터 적어도 하나의 예시를 포함하고 1, 2 또는 4개의 RBP 및 3개의 단백질 도메인 각각을 연결하는 2개의 상이한 링커 조합을 갖는 예시를 포함하였다.

[1939] 풀의 예시적인 트랜스 주형 RNA는 8 또는 16개 뉴클레오티드의 링커; 1 또는 4개의 RRS 반복부(MS2 서열); 13개 뉴클레오티드의 PBS(또는 음성 대조군의 경우 스크램블드 무작위 PBS); 및 16개의 뉴클레오티드 스페이서가 있는 5' 말단-보호 Cas9 스캐폴드, 스페이서가 없는 5' 말단-보호 Cas9 스캐폴드, 대조군 5' 말단-보호ePEG, 또는 5' 말단 보호 특징부가 없는 것을 갖는 예시를 포함하였다.

[1940] 이 실시예에 사용된 예시적인 트랜스 주형

PL13292
PL13293
PL13296
PL13297
PL13298
PL13299
PL13302
PL13303
PL13388
PL13389
PL13392
PL13393
PL13394
PL13395
PL13398
PL13399
PL13484
PL13488
PL13489
PL13490
PL13491
PL13494
PL13580
PL13581
PL13584
PL13585
PL13586
PL13590
PL13591

[1941]

[1942] PLV4921 및 시스 주형 RNA의 서열은 실시예 7에서 제공되었으며, 예시적인 유전자 변형 폴리펩티드 및 트랜스 주형 RNA의 서열은 실시예 6에서 제공되었다.

[1943] 각각의 예시적인 트랜스-유전자 변형 시스템(예시적인 유전자 변형 폴리펩티드, 예시적인 트랜스 주형 RNA 및 예시적인 gRNA의 조합)을 시험하여 BFP-GFP SNP 리포터를 발현하는 게놈 랜딩 패드를 보유하는 포유동물 U2OS 세포를 사용하여 게놈-편집 능력을 결정하였다. U2OS 세포는 BFP-발현 세포이고 트랜스 주형 RNA에 의해 프로그래밍된 단일 염기쌍 치환을 통해 GFP-발현 세포로 전환될 수 있다. U2OS 세포를 예시적인 개별 유전자 변형 폴리펩티드를 인코딩하는 플라스미드, 예시적인 gRNA를 인코딩하는 제2 플라스미드, 및 예시적인 개별 트랜스 주형 RNA를 인코딩하는 제3 플라스미드의 형태로 각각의 예시적인 트랜스-유전자 발현 시스템을 이용하여 전기천공하였다.

[1944] 유전자 변형 시스템의 전기천공 4일 후 유세포 분석법으로 세포를 분석하였다. GFP 세포의 백분율을 재작성 활성에 대한 척도로 사용하였고 동일한 유전자 변형 폴리펩티드 및 시스 주형 RNA를 함유하는 유전자 변형 시스템의 재작성 활성에 대해 정규화하였다.

[1945] 도 9는 BFP/GFP SNP 리포터 분석에서 예시적인 트랜스-유전자 변형 시스템의 재작성 활성에 대한 히트 맵을 나

타낸다. 결과는 다수의 트랜스 주형 RNA가 시스 주형 RNA와 함께 동일한 유전자 변형 폴리펩티드를 함유하는 유전자 변형 시스템과 비슷하거나 더 나은 재작성 활성을 제공한다는 것을 나타낸다. 이러한 결과는 또한 Cas9 스캐폴드와 16개 뉴클레오티드 스페이서를 포함하는 5' 말단 블록을 함유하는 트랜스 주형 RNA가 스페이서가 없는 5' 말단 블록 Cas9 스캐폴드를 함유하는 트랜스 주형 RNA보다 대부분의 평가된 유전자 변형 폴리펩티드에 걸쳐 더 나은 재작성 활성을 제공한다는 것을 나타낸다. 결과는 16 nt 스페이서가 있는 5' 말단 보호 Cas9 스캐폴드가 대조군 5' 말단 보호 ePEG, 또는 5' 말단 보호 특징부가 없는 것보다 더 나은 재작성 활성을 제공한다는 것을 추가로 나타낸다.

[1946] **실시예 11: 재작성 활성에 대한 트랜스 주형 RNA 구성요소의 기여도 평가**

[1947] 이 실시예는 트랜스 주형 RNA 및 예시적인 RBP-함유 유전자 변형 폴리펩티드를 포함하는 유전자 변형 시스템의 재작성 활성에 대한 RRS-RBP 상호작용의 중요성을 입증한다.

[1948] 예시적인 유전자 변형 시스템은 RBP-함유 유전자 변형 폴리펩티드 E1 또는 RBP G1이 결여된 예시적인 유전자 변형 폴리펩티드(각각은 시스 주형을 사용하여 이전에 입증된 재작성 활성을 가짐, 데이터는 나타내지 않음) 및 Cas9 스캐폴드를 포함하고 스페이서를 포함하지 않으며 4개의 RRS(이 실시예에서는 4개의 MS2 서열) 또는 1개의 RRS(1개의 MS2 서열)을 포함하는 다양한 트랜스 주형 RNA를 포함하였다. 이 실시예에서 시험된 예시적인 유전자 변형 폴리펩티드에 대한 서열은 하기 표 X1에 제공되어 있다.

[1949] *표 X1. 예시적인 유전자 변형 폴리펩티드*

[1950]

유전자 변형 폴리펩티드	플라스미드 (등록 ID)	아미노산 서열	서열번호
E1*	PL12113	MPAAKRVKLDGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPK VWAEINPPGLASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLR ETIRKFRAAGILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREV NKRVETIHPTVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPE SQLIFAFEWADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRD LQGFRLDHPSVSLLQYVDDLLIAADTQAACLSATRDLLMTLAE LGYRVSGKKAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPV PKTKRQVREFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLV WGEKEEEAFQSLKLALTQPPALALPSLDKPFQLFVEETSGAAK GVLTQALGPWKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLT REASKLTFGQDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLL DPPRVRFKQTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPD LTDQPLAQAEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEP LPIGTSAQKAELIALTKALEWSKDKSVNIYTDSRYAFATLHVH GMIYRERGWLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCK GHQKDDAPTSTGNRRADEVAREVAIRPLSTQATISAEAAAKEA AAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAAKAMA SNFTQFVLVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKV TCSVRQSSAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSD CELIVKAMQGLLKDGNPIPSAIAANSGIYGGSGGSMASNFTQFV LVDNGGTGDVTVAPSNFANGIAEWISSNSRSQAYKVTCSVRQS SAQNRKYTIKVEVPKGAWRSYLNMELTIPIFATNSDCELIVKA MQGLLKDGNPIPSAIAANSGIYAEAAAKEAAAKEAAAKEAAA KALEAEAAAKEAAAKEAAAKEAAAKADKKYSIGLDIGTNSVG WAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEE SFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQT	16,801

[1951]

		YNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKN GLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNL LAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGS IPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPL ARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVE DRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDR EMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIR DKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPE NIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPV ENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMKNYWRQ LLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQF YKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYK VYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIR KRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQ TGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYK EVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSK YVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISE FSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLG APAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQ LGGDKRTADGSEFESPKKKAKVE	
G2	PLV10990	MPAAKRVKLDGGDKKYSIGLDIGTNSVGWAVITDEYKVPSKK FKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTR RKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPI FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHM IKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASG VDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLT PNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFL AAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLL KALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAIL RRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTR KSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHD LLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAH LFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLK SDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYL YYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNK VLTRSDKARGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKF DNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMN TKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHH	16,802

		AHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETG EIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPK RNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSK KLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPK YSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADAN LDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTT IDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGAEAA AKEAAAKEAAAKEAAAKALEAEAAAKEAAAKEAAAKEAAA KAGGTAPLEEEYRLFLEAPIQNVTLLEQWKREIPKVWAEINPPG LASTQAPIHVQLLSTALPVRVRQYPITLEAKRSLRETIRKFRAA GILRPVHSPWNTPLLPVRKSGTSEYRMVQDLREVNKRVETIHP TVPNPYTLLSLLPPDRIWYSVLDLKDAFFCIPLAPESQLIFAFEW ADAEEGESGQLTWTRLPQGFKNSPTLFNEALNRDLQGFRLDHP SVSLLQYVDDLLIAADTQAACLSATRDLLMTLAELGYRVSGK KAQLCQEEVTYLGFKIHKGSRSLSNSRTQAILQIPVPKTKRQVR EFLGKIGYCRLFIPGFAELAQPLYAATRPGNDPLVWGEKEEEAF QSLKLALTQPPALALPSLDKPFQLFVEETSGAAKGVLTQALGP WKRPVAYLSKRLDPVAAGWPRCLRAIAAAALLTREASKLTFG QDIEITSSHNLESLLRSPPDKWLTNARITQYQVLLLDPPRVRFK QTAALNPATLLPETDDTLPIHHCLDTLDSLTSTRPDLTDQPLAQ AEATLFTDGSSYIRDGKRYAGAAVVTLDSVIWAEPLPIGTSAQ KAELIALTKALEWSKDKSVNIYTDSRYAFATLHVHGMIYRERG WLTAGGKAIKNAPEILALLTAVWLPKRVAVMHCKGHQKDDA PTSTGNRRADEVAREVAIRPLSTQATISAGKRTADGSEFEKRTA DGSEFESPKKKAKVE	
N863A Cas9	PLV1425	MGKPIPNPLLGLDSTAPKKKRKVGIHGVPAADKKYSIGLDIGT NSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFH RLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQ LVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPG EKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDD LDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLS ASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMR KPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTL TLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRK LINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQ KAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVM GRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQI LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDDSIDNKVLTRSDKARGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVE TRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFR KDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFV	16,803

[1952]

		YGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITL ANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIV KKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNE LALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLD EIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLF TLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSRADPKKKRKVEFSR	

[1953]

[1954] * 서열은 드라이버 표에 포함된다.

[1955] 예시적인 시스 주형 RNA는 PLV4106이다(앞서 실시예 9에 기재된 서열). 예시적인 트랜스 주형 RNA는 하기 표 X2에 열거되어 있다.

[1956] *표 X2. 예시적인 트랜스 주형 RNA*

트랜스 주형 RNA	플라스미드 (등록 ID)	플라스미드 명칭(표 S4 참조)
4x MS2 스페이서 없음	PL13584	305 스캐폴드 _RT12_P13_L8_4xMS2_
1x MS2 스페이서 없음	PL13580	301 스캐폴드 _RT12_P13_L8_1xMS2_

[1957]

[1958] 음성 대조군 폴리펩티드는 N863A Cas9(RT 또는 RBP가 결여됨)를 포함하였다. 예시적인 유전자 변형 시스템을 시험하여 BFP-GFP SNP 리포터를 발현하는 게놈 랜딩 패드를 보유하는 포유동물 U2OS 세포를 사용하여 게놈-편집 능력을 결정하였다. U2OS 세포는 BFP-발현 세포이고 트랜스 주형 RNA에 의해 프로그래밍된 단일 염기쌍 치환을 통해 GFP-발현 세포로 전환될 수 있다. U2OS 세포를 예시적인 트랜스 주형 RNA를 인코딩하는 다양한 용량(37.5 ng, 75 ng, 또는 150 ng)의 플라스미드, 예시적인 유전자 변형 폴리펩티드를 인코딩하는 300 ng의 플라스미드, 및 예시적인 gRNA(즉, 실시예 9에 기재된 바와 같은 gRNA PLV4165)를 인코딩하는 200 ng의 제2 플라스미드를 이용하여 전기천공하였다.

[1959] 유전자 변형 시스템의 전기천공 4일 후 유세포 분석법으로 세포를 분석하였다. GFP-발현 세포의 백분율을 재작성 활성에 대한 척도로 사용하였다.

[1960] 도 10은 예시적인 유전자 변형 시스템에 대한 재작성 활성의 그래프를 나타낸다. 결과는 음성 대조군 N863A Cas9가 최소한의 재작성 활성을 나타냈음을 보인다. RBP가 없는 유전자 변형 폴리펩티드 G2는 4x 또는 1x RRS 및 Cas9 스캐폴드를 포함하고 스페이서가 결여된 5' 말단 블록을 포함하는 트랜스 주형 RNA와 조합될 때 재작성 활성을 나타내지 않았다. 결과는 예시적인 RBP-함유 유전자 변형 폴리펩티드 E1이 4x RRS 및 스페이서가 결여된 5' 말단 블록 Cas9 스캐폴드를 포함하는 트랜스 주형 RNA와 조합될 때 재작성 활성을 나타내고, 트랜스 주형 RNA의 용량이 증가함에 따라 재작성 활성이 증가한다는 것을 나타낸다. E1은 1x RRS 및 스페이서가 결여된 5' 말단 블록 Cas9 스캐폴드를 포함하는 트랜스 주형 RNA와 조합될 때 낮은 수준의 재작성 활성을 나타내지만, 트랜스 주형 RNA의 용량이 증가함에 따라 재작성 활성이 증가한다는 것을 나타낸다. 결과는 RRS-RBP 상호작용이 5' 말단 블록 Cas9 스캐폴드를 포함하고 스페이서가 결여된 트랜스 주형 RNA를 사용할 때 트랜스 주형 RNA:유전자 변형 폴리펩티드:표적 DNA 복합체의 조립에 중요할 수 있음을 나타낸다. 이론에 구속되고자 함이 없이, 1x RRS와 상호작용하는 E1의 RBP에 의해 제공될 수 있는 더 약한 RBP:RRS 상호작용조차도 트랜스 주형 RNA의 용량이 증가하여 관찰 가능한 재작성 활성을 제공함에 따라 충분한 상호작용을 제공하는 것으로 보인다.

[1961] 종합하면, 이러한 결과는 트랜스 주형 RNA와 예시적인 유전자 변형 폴리펩티드 사이의 RRS:RBP 상호작용이 트랜스 주형 RNA가 Cas9 스캐폴드를 포함하고 스페이서가 결여된 5' 말단 블록을 함유할 경우 상기 구성요소를 포함하는 유전자 변형 시스템의 동원 및 재작성 활성에 중요함을 나타낸다.

[1962] **실시예 12: 재작성 활성에 대한 5' 말단 블록 스페이서의 기여도 평가**

[1963] 이 실시예는 트랜스 주형 RNA 및 예시적인 RBP-함유 유전자 변형 폴리펩티드를 포함하는 유전자 변형 시스템의 재작성 활성에 대한 5' 말단 블록의 중요성 및 역할을 입증한다. Cas9 스캐폴드 및 스페이서가 있는 5' 말단 블록의 존재는 5' 말단 블록이 Cas9 스캐폴드를 함유하고 스페이서가 없는 유사한 시스템에 비해 실시예 10의 유전자 변형 시스템의 재작성 활성을 증가시켰다. 5' 말단 블록에서 스페이서의 역할을 추가로 조사하기 위해, 예시적인 RBP-함유 유전자 변형 폴리펩티드 E1과, Cas9 스캐폴드를 포함하고 스페이서가 없거나(스페이서 없음) 16개 뉴클레오티드 스페이서를 포함하는(스페이서) 5' 말단 블록, 및 4개의 RRS(이 실시예에서는 4개의 MS2 서열) 또는 1개의 RRS(1개의 MS2 서열)을 포함하는 다양한 트랜스 주형 RNA를 포함하는 예시적인 유전자 변형 시스템을 조립하였다. 이론에 구속되고자 함이 없이, 16개 뉴클레오티드 스페이서는 E1의 S. 피오게네스 Cas9 도메인에 의한 제한된 닉 생성 활성만을 가능하게 하지만, E1:스페이서가 표적 DNA의 스페이서의 상보적 서열에 결합할 수 있게 하기에 충분하다고 생각된다. Cas9 스캐폴드(유사 매듭(pseudoknot))를 포함하지 않는 5' 말단 블록을 포함하는 대조군 트랜스 주형 RNA도 또한 포함하였다. 16개의 뉴클레오티드 스페이서가 (전체 스페이서 길이가 20 nt인) 닉을 표적으로 하는 부위를 표적화하는 gRNA를 일부 샘플(+2번째 닉)에 포함하였다. N863A Cas9(RT 또는 RBP가 결여됨)를 대조군 폴리펩티드로 사용하였다. 이 실시예에 사용된 예시적인 sgRNA는 PLV4165(실시예 9에 기재된 서열)이었다. 이 실시예에 사용된 예시적인 유전자 변형 폴리펩티드는 실시예 11에 기재된 것들이었다. 이 실시예에 사용된 예시적인 시스 주형 RNA는 PLV4106(실시예 9에 기재된 서열)이었다. 이 실시예에 사용된 예시적인 트랜스 주형 RNA는 하기 표 X3에 열거되어 있다.

[1964] *표 X3. 예시적인 트랜스 주형 RNA*

트랜스 주형 RNA	플라스미드 (등록 ID)	플라스미드 명칭(표 S4 참조)
4xMS2 유사매듭	PLV13392	113 유사매듭 _RT12_P13_L8_4xMS2_
4xMS2 스페이서 없음	PLV13584	305 스캐폴드_RT12_P13_L8_4xMS2_
4xMS2 + 스페이서	PLV13488	209 g16A_RT12_P13_L8_4xMS2_
1xMS2 유사매듭	PLV13388	109 유사매듭 _RT12_P13_L8_1xMS2_
1xMS2 스페이서 없음	PLV13580	301 스캐폴드_RT12_P13_L8_1xMS2_
1xMS2 + 스페이서	PLV13484	205 g16A_RT12_P13_L8_1xMS2_

[1965]

[1966] PL13488 및 PL13484의 5' 스페이서와 동일한 표적에 결합/닉을 생성하는 예시적인 제2 닉 gRNA는 다음 핵산 서열을 갖는 PLV4163이었다:

[1967] CCGGCAAGCTGCCCGTGCCCGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAAC TTGAAAAAGTGGCACCGAGTCGGTGC (서열번호 16,715).

[1968] 예시적인 유전자 변형 시스템을 시험하여 BFP-GFP SNP 리포터를 발현하는 게놈 랜딩 패드를 보유하는 포유동물 U2OS 세포를 사용하여 게놈-편집 능력을 결정하였다. U2OS 세포는 BFP-발현 세포이고 트랜스 주형 RNA에 의해 프로그래밍된 단일 염기쌍 치환을 통해 GFP-발현 세포로 전환될 수 있다. U2OS 세포를 예시적인 개별 유전자 변형 폴리펩티드를 인코딩하는 플라스미드, 예시적인 gRNA를 인코딩하는 제2 플라스미드, 예시적인 개별 트랜스 주형 RNA를 인코딩하는 제3 플라스미드, 및 (해당 샘플에 필요한 경우) 제2 닉 gRNA를 인코딩하는 제4 플라스미드의 형태로 각각의 예시적인 트랜스-유전자 발현 시스템을 이용하여 전기천공하였다.

[1969] 유전자 변형 시스템의 전기천공 4일 후 유세포 분석법으로 세포를 분석하였다. GFP-발현 세포의 백분율을 재작성 활성에 대한 척도로 사용하였다.

[1970] 도 11a 내지 11b는 예시적인 유전자 변형 시스템에 대한 재작성 활성의 그래프를 나타낸다. 결과는 음성 대조군 N863A Cas9가 4x RRS 및 Cas9 스캐폴드와 스페이서가 있는 5' 말단 블록을 함유하는 트랜스 주형 RNA를 이용하여 최소한의 재작성 활성을 나타냈음을 보인다. 양성 대조군 시스 주형 RNA는 유전자 변형 폴리펩티드와 조합될 때 상당한 재작성 활성을 나타내었다(도 11a 내지 11b).

[1971] 결과는 트랜스 주형 5' 말단 블록의 스페이서(줄무늬 막대)와 동일한 표적 부위 DNA에 닉을 지시하는 gRNA의 추가가 재작성 활성을 증가시키지 않았음을 나타낸다(열린 막대; 유사매듭, 스페이서 없음, 및 + 스페이서(제2 닉이 있거나 없음)). gRNA를 이용한 재작성 증가가 없다는 것은 이러한 예시적인 유전자 변형 시스템의 경우 5' 말단 블록 스페이서에 의해 표적화된 부위에서의 닉 생성이 재작성 활성에 중요하지 않음을 시사한다.

[1972] 대조적으로, 결과는 트랜스 주형 RNA의 5' 말단 블록에서 스페이서의 존재가 재작성 활성을 상당히 증가시켰음을 나타낸다(도 11a 내지 11b). 스페이서의 존재는 4x RRS 트랜스 주형 RNA 시스템뿐만 아니라, 1x RRS 트랜스 주형 RNA 시스템 둘 다에서 재작성 활성을 증가시켰으며, 이는 재작성 활성에 대한 RRS:RBP 상호작용과 5' 말단 블록 스페이서:추가 유전자 변형 폴리펩티드 상호작용의 기여가 부가적임을 시사한다. 이러한 결과는 5' 말단 블록 스페이서가 다른 약한 RRS:RBP 상호작용을 구제할 수 있음을 추가로 시사한다(도 11a의 스페이서 없음과 + 스페이서 대 도 1b의 스페이서 없음과 +스페이서를 비교). Cas9 스캐폴드 및 스페이서가 있는 5' 말단 블록 및 4x RRS를 포함하는 트랜스 주형 RNA를 포함하는 시스템을 사용한 재작성 활성은 닉 생성을 제공하는 gRNA의 추가에 의해 영향을 받지 않았다(제2 닉이 있거나 없는 +스페이서를 비교). 이러한 결과는 5' 말단 블록 스페이서의 주요 역할이 이러한 예시적인 유전자 변형 시스템에서 표적 유전자좌에 대한 트랜스 주형 RNA의 Cas9-매개 결합을 제공하는 것이며, 스페이서에 의해 표적화된 부위에서 닉 생성이 이들 예시적인 시스템에 대한 재작성 활성에 중요하지 않음을 시사한다.

[1973] 종합하면, 이러한 결과는 BFP에서 GFP로의 재작성 분석에서 이들 예시적인 시스템의 5' 말단 블록 스페이서의 역할이 트랜스 주형을 표적 부위로 동원하는 제2 모드를 위한 것임을 시사한다.

[1974] 이론에 구속되고자 함이 없이, 실시예 및 개시내용은 부분적으로 2개 이상의 상호작용에 의해 트랜스 주형 RNA를 유전자 변형 폴리펩티드:sgRNA:표적 게놈 DNA 복합체에 고정시키는 것이 높은 재작성 활성을 달성하는 데, 예를 들어 단일 또는 여러 뉴클레오티드 길이의 편집물을 달성하는 데 중요하다는 발견에 관한 것이다. 이 실시예는, 예를 들어 1) RRS:RBP 상호작용 및 2) (추가 유전자 변형 폴리펩티드를 통해 매개되는) 표적 DNA 상호작

용에 대한 5' 말단 블록 Cas9 스캐폴드 및 스페이서는 트랜스 주형 RNA를 유전자 변형 폴리펩티드:sgRNA:표적 게놈 DNA 복합체에 함께 고정시켜 재작성을 가능하게 하는 예시적인 상호작용을 나타냄을 시사한다. RRS:RBP 상호작용이 5' 말단 블록 스페이서의 부재 하에 중요하지만, 둘 다의 존재는 높은 재작성 활성을 제공하고 5' 말단 블록 스페이서의 존재는 더 약한 RRS:RBP 상호작용과 조합하여 재작성 활성을 구제한다.

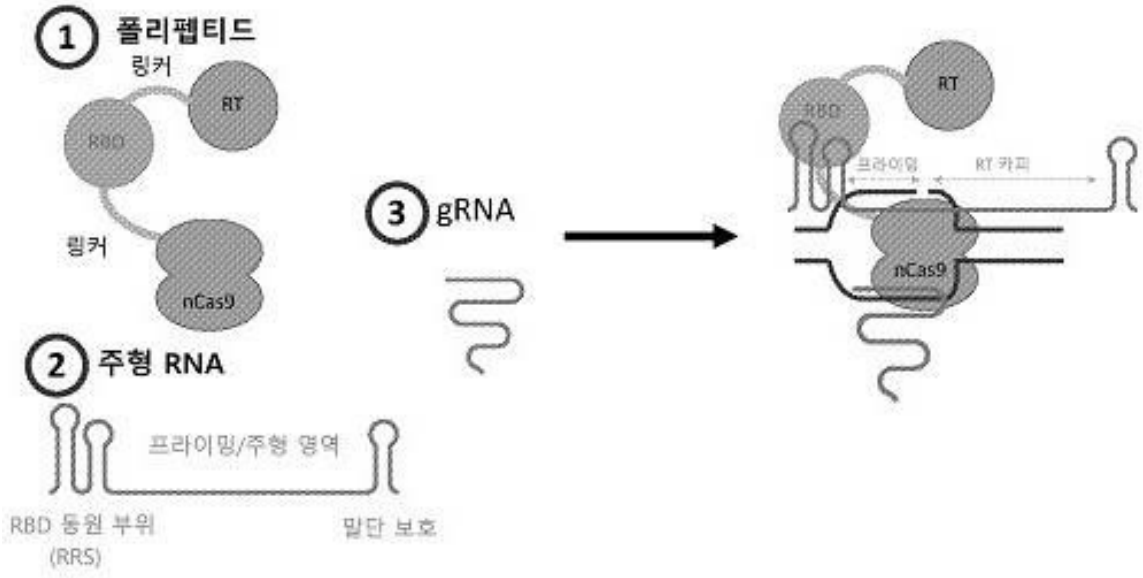
[1975] "약", "적어도", "미만", 및 "초과"와 같은 본 출원에서 일부 파라미터를 기재하는 모든 수치 범위에 대하여, 기재는 또한 언급된 값에 의해 제한되는 임의의 범위를 필수적으로 포함하는 것으로 이해되어야 한다. 따라서, 예를 들어 "적어도 1, 2, 3, 4, 또는 5"라는 기재는 특히 1 내지 2, 1 내지 3, 1 내지 4, 1 내지 5, 2 내지 3, 2 내지 4, 2 내지 5, 3 내지 4, 3 내지 5, 및 4 내지 5 등의 범위를 또한 기재한다.

[1976] 본 명세서에 인용된 모든 특허, 출원, 또는 기타 참조문헌, 예컨대 비-특허 문헌 및 참조 서열 정보에 있어서, 이들은 모든 목적을 위해 그리고 언급된 명제를 위해 그 전체가 참고로 포함되는 것으로 이해되어야 한다. 참조로 포함된 문서와 본 출원 사이에 임의의 상충이 존재하는 경우, 본 출원이 우선할 것이다. 예를 들어 게놈 유전자좌, 게놈 서열, 기능적 주석, 대립유전자 변이체, 및 참조 mRNA(예를 들어, 엑손 경계 또는 반응 요소를 포함함) 및 단백질 서열(예컨대, 보존된 도메인 구조)뿐만 아니라, 화학적 참조(예를 들어, PubChem 화합물, PubChem 물질, 또는 PubChem 생물검정 항목(구조 및 분석과 같은 그 안의 주석을 포함함) 등)를 포함하여, 유전자ID 또는 수탁 번호(통상적으로 NCBI 수탁 번호를 참조함)와 같은, 본 출원에 개시된 참조 유전자 서열과 연관된 모든 정보는 그 전체가 본 명세서에 참고로 포함된다.

[1977] 본 출원에 사용된 제목은 단지 편의를 위한 것이며 본 출원의 해석에 영향을 미치지 않는다.

도면

도면1



도면2a

가능한 구성

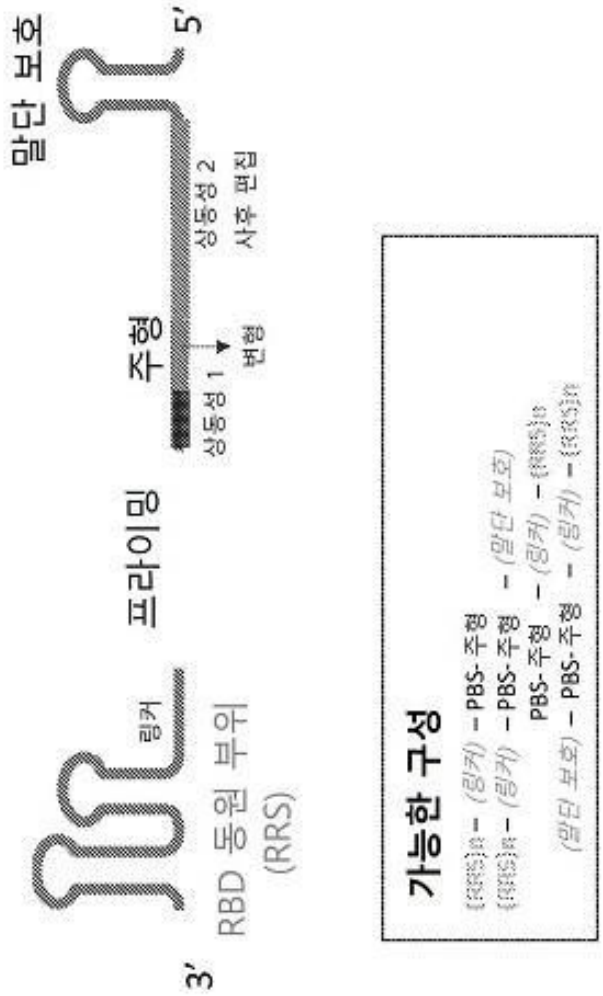
- Cas9 – (링커) - RT – (링커) – {RBD}n
- RT – (링커) - Cas9 – (링커) – {RBD}n
- {RBD}n – (링커) – Cas9 (링커) – RT
- {RBD}n – (링커) – RT (링커) – Cas9
- Cas9 – (링커) – {RBD}n – (링커) – RT
- RT – (링커) – {RBD}n – (링커) – Cas9

도면2b

가능한 구성

- 이량체화 도메인은 각각의 도메인에 대해 탠덤 반복부(n=1 내지 30)로 여러 번 존재할 수 있음

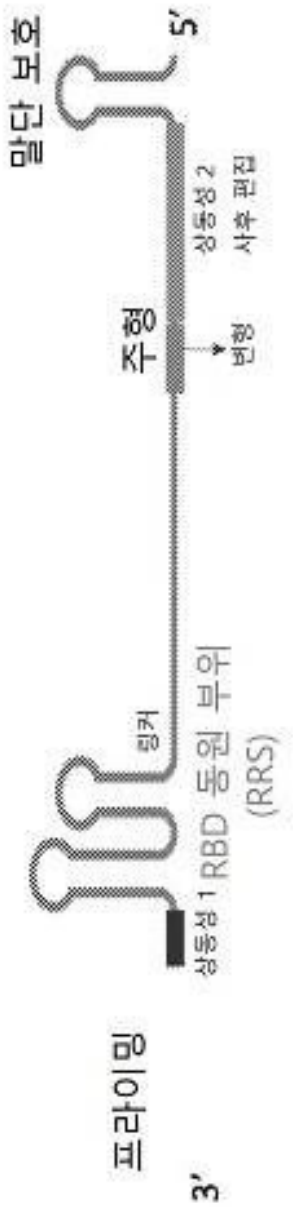
도면3a


말단 보호
5'
주형
상동성 2
사후 편집
상동성 1
변형
프라이밍
링커
RBD 동원 부위
(RRS)
3'
가능한 구성
(RRS) – (링커) – PBS- 주형
(RRS) – (링커) – PBS- 주형 – (말단 보호)
PBS- 주형 – (링커) – (RRS)
(말단 보호) – PBS- 주형 – (링커) – (RRS)

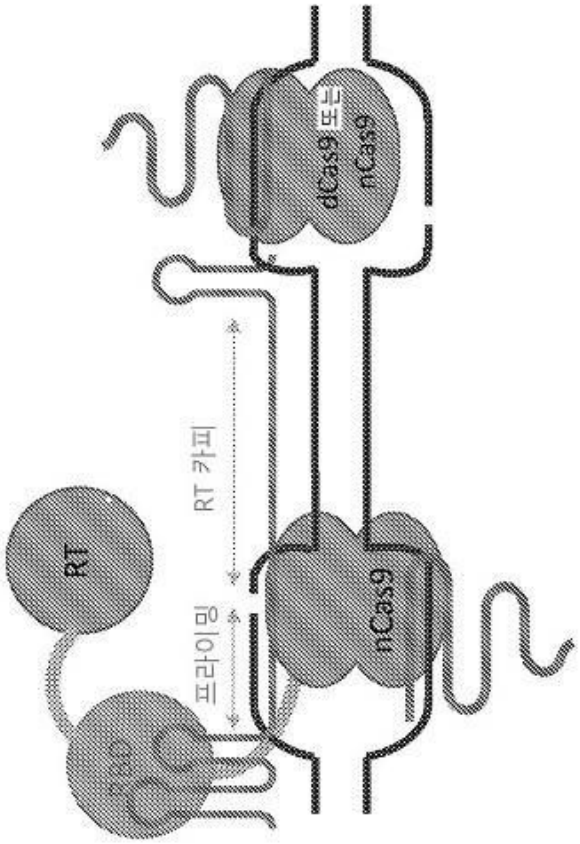
도면3b

- **RBD 동원 부위(RRS)**
  - 다양한 기원(서열 표 참조)
  - n회 존재할 수 있음: 1 내지 10회
- **링커:** 0 내지 20 nt -> 서열 표 참조
- **프라이밍:** 8 내지 17 nt -> 서열 표 참조
- **주형:**
  - 상동성 영역 사전 편집 1: 0 내지 20 nt
  - 상동성 영역 사후 편집: 5 내지 500 nt
  - 변형: 긴 및 짧은 삽입, 결실, 점 돌연변이
- **말단 블록**(서열 표 참조)

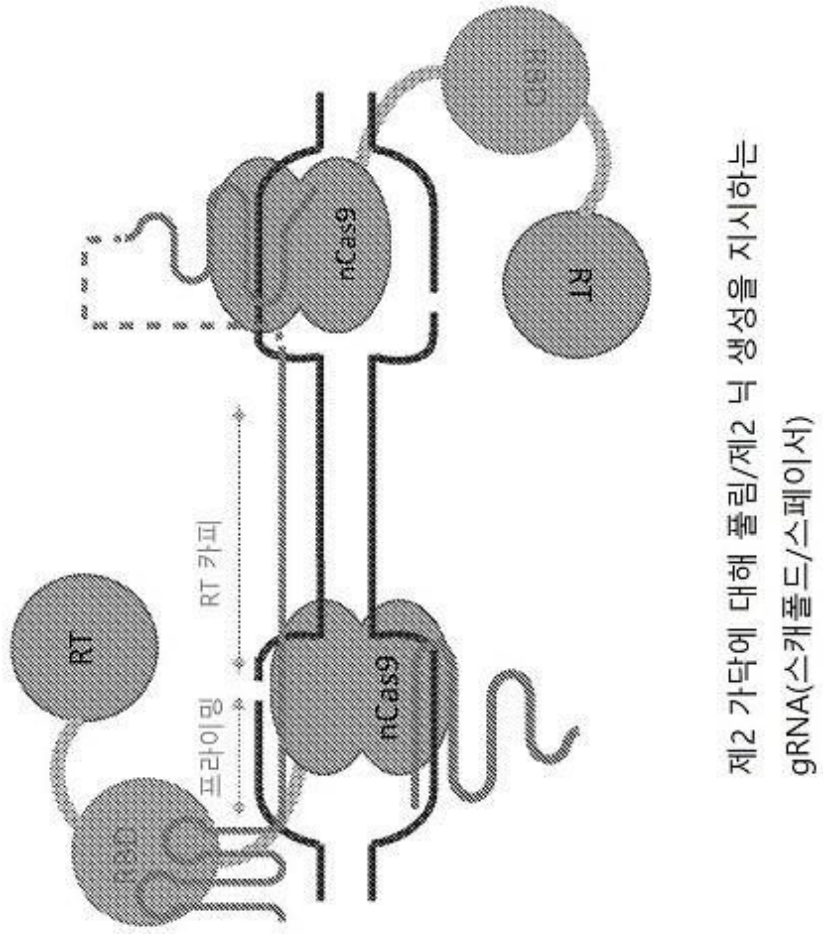
도면3c



도면4a

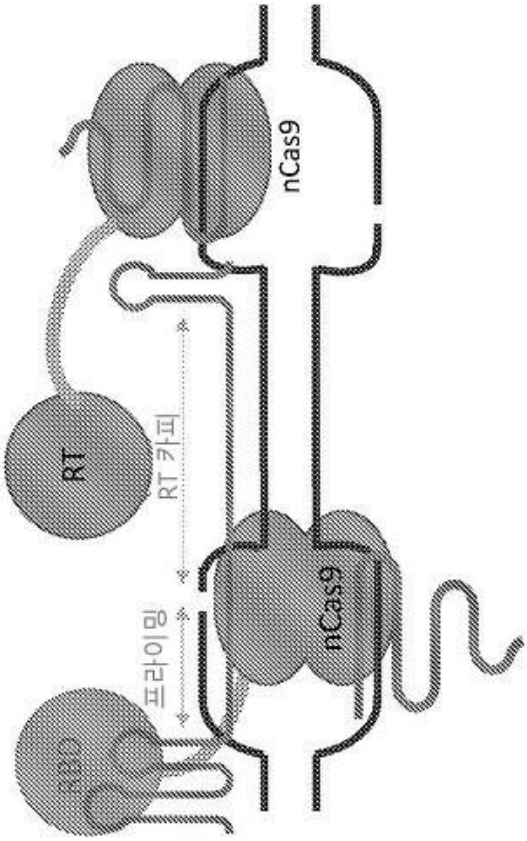


도면4b

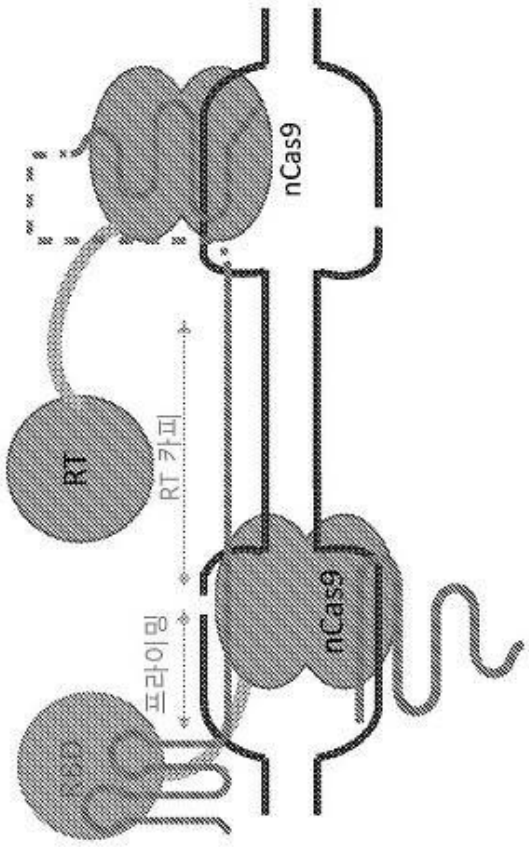


제2 가닥에 대해 풀림/제2 닉 생성을 지시하는
gRNA(스캐폴드/스페이서)

도면5a



도면5b



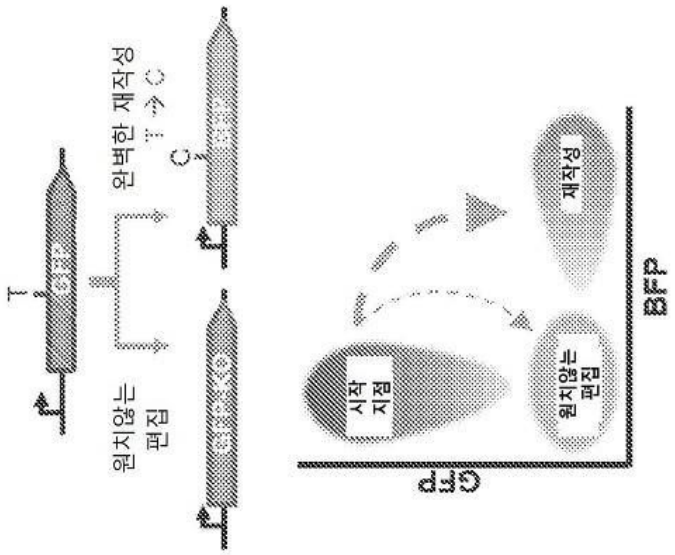
도면6a

216개 드라이버
구성 1: 96개 드라이버
Cas9
(RBP)n
4
n = 1,2,4
RT
4
2 RTs
구성 2: 96개 드라이버
RT
2 RTs
4
(RBP)n
n = 1,2,4
4
Cas9
4
구성 3: 24개 드라이버
Cas9
RT
2 RTs
4
(RBP)n
n = 1,2,4
드라이버 변수:
• 구성(3)
• RT 도메인(2)
• RBP 반복부의 수(3)
• 링커 조합(16)
(구성 3에 대한 4개 링커 옵션)

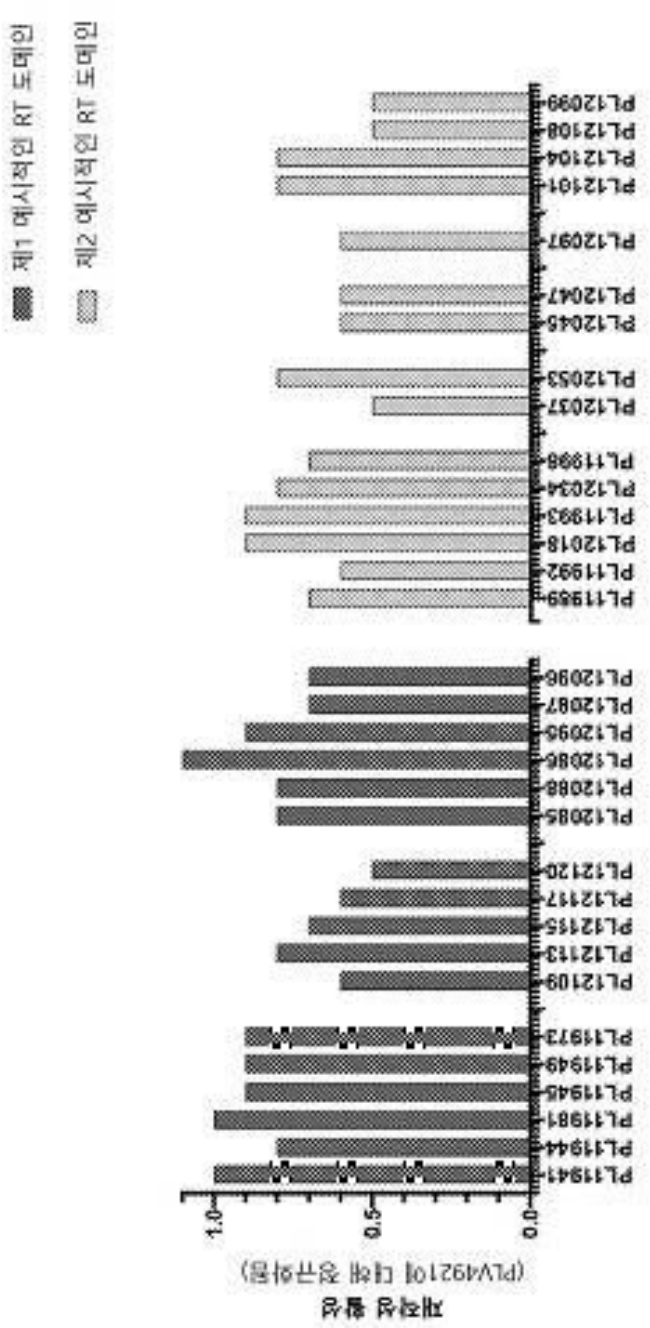
도면6b

384개 주형(288개 활성 구성 + 96개 음성 대조군)
3' 말단 블록
5' 말단 블록
링커
프라이밍
RT 주형
상동성 1
상동성 2
5'
3'
BAP 효소
보 바이오틴 변형부
라입
주형 변수:
• 3' 말단 블록(2)
• 말단 블록 없음
• 대조군 ePEG 말단-블록
• RNS 반복부의 수(3)
• 1x MS2
• 2x MS2
• 4x MS2
• 링커 길이(2)
• 8 nt
• 16 nt
• 프라이밍(PBS) 길이(3+1)
• 8 nt
• 13 nt
• 17 nt
• (비-상보적 PBS)
• 상동성 2 길이(2)
• 12 nt
• 규칙적인 간격으로 4개 SNP가 있는 63 nt
• 5' 말단 블록(4)
• 말단 블록 없음
• 대조군 ePEG 말단-블록
• Cas9 스캐폴드(스페이서 없음)
• 16 nt 스페이서가 있는 Cas9 스캐폴드

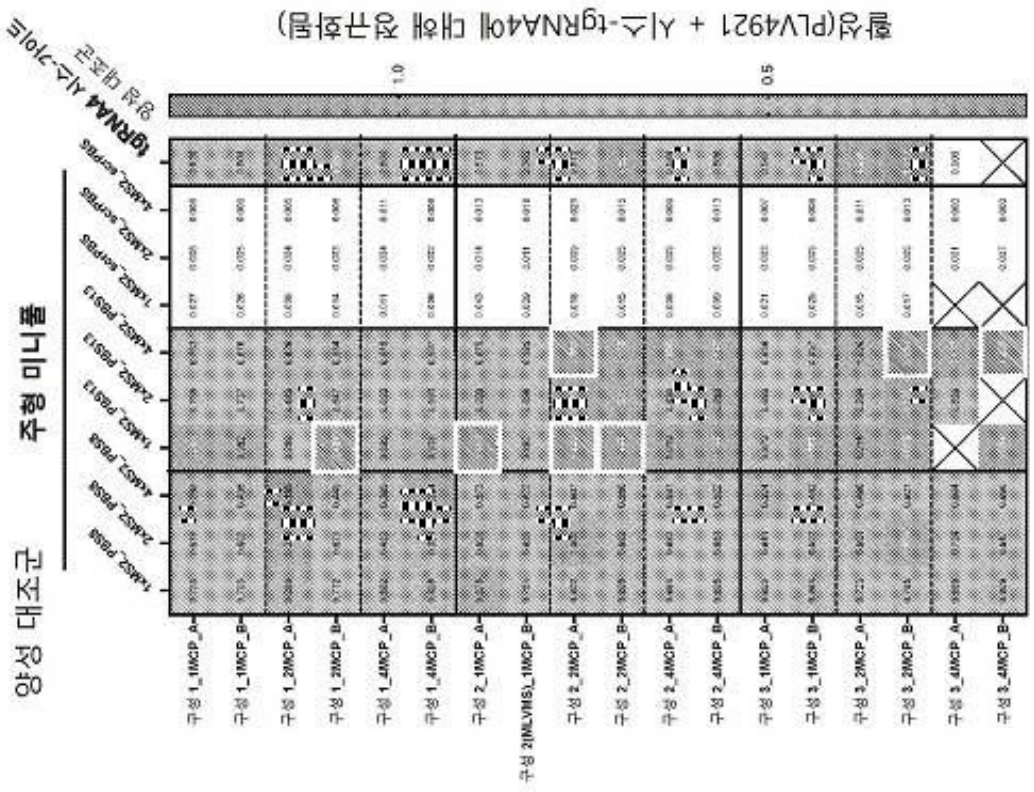
도면7a



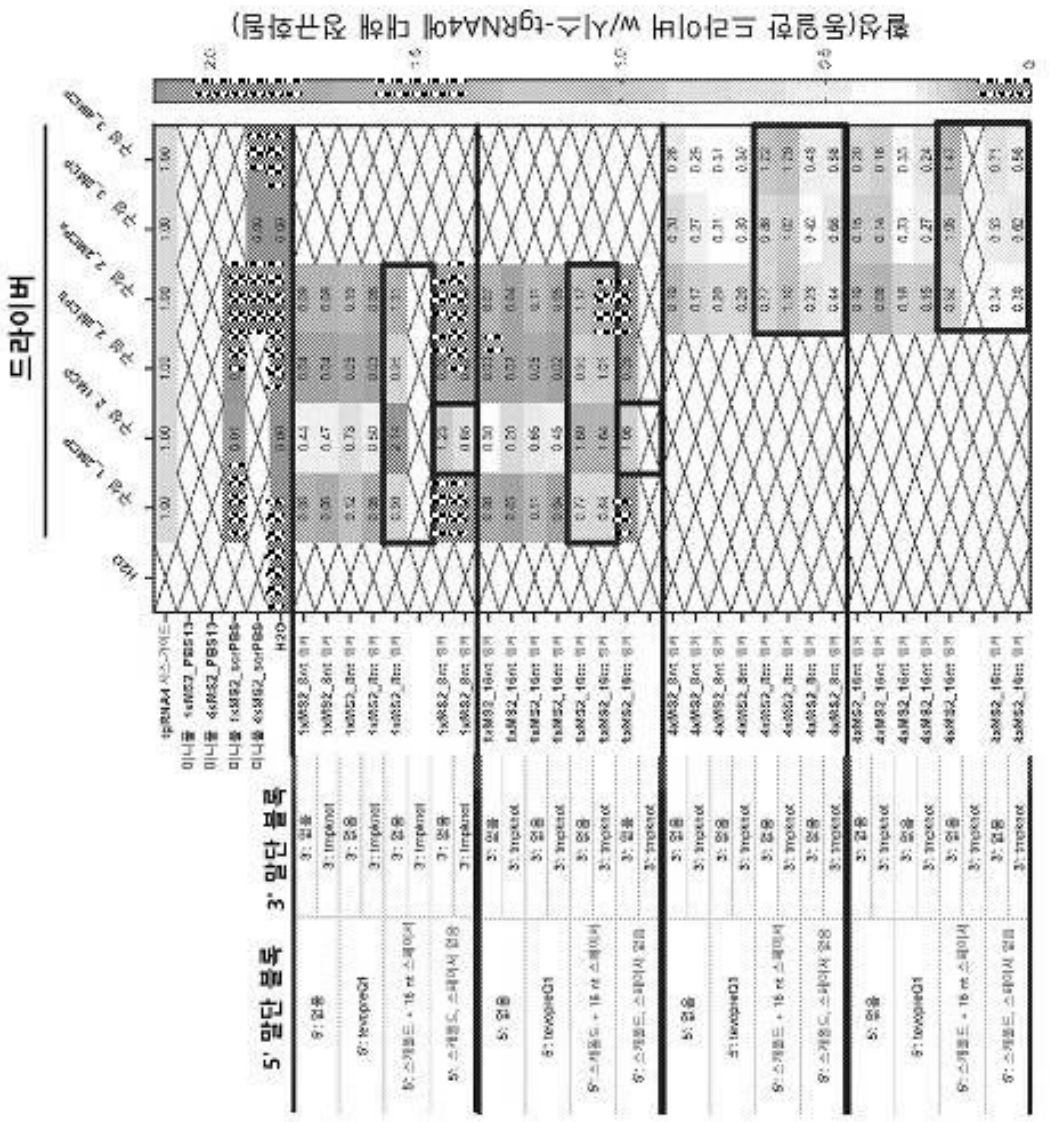
도면7b


제1 예시적인 RT 도메인
제2 예시적인 RT 도메인
PL12099
PL12108
PL12104
PL12101
PL12097
PL12047
PL12045
PL12053
PL12037
PL11998
PL12034
PL11993
PL12018
PL11992
PL11989
PL12096
PL12087
PL12093
PL12086
PL12088
PL12085
PL12120
PL12117
PL12115
PL12113
PL12109
PL11973
PL11949
PL11945
PL11981
PL11944
PL11941
1.0
0.5
0.0
(PLV492101 대해 정규화됨)
재조성 활성

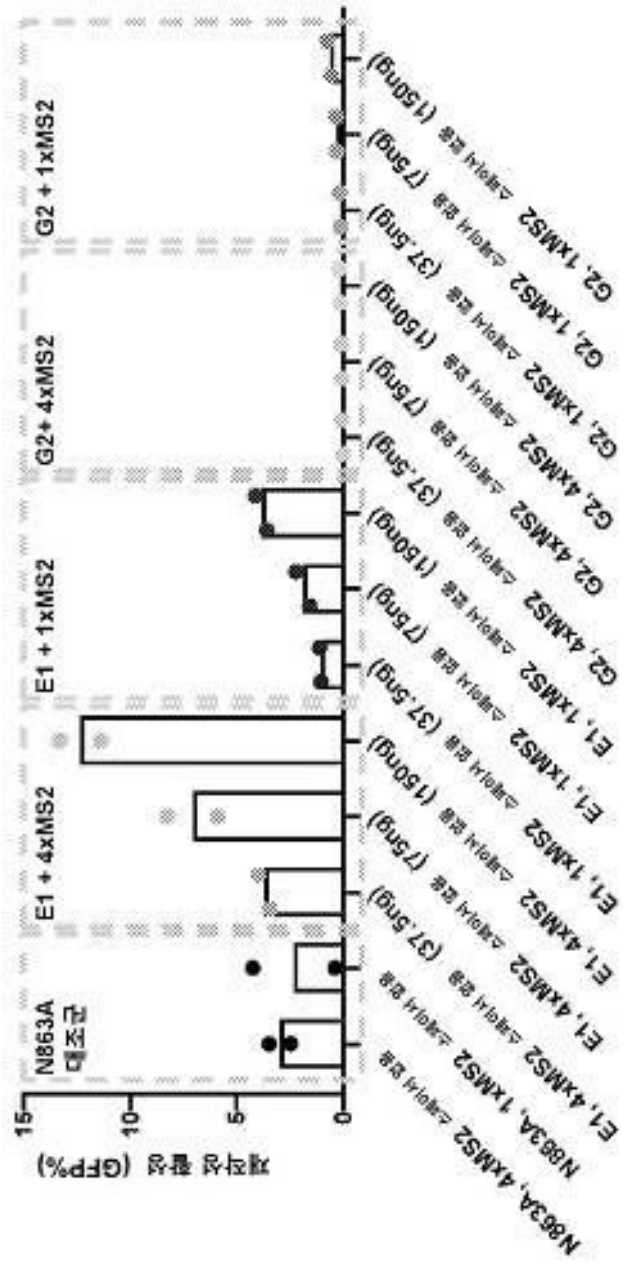
도면8


활성(PLV4921 + 시스-tgRNA44에 대해 정규화됨)
1.0
0.5
양성 대조군
주형 미니홀
tgRNA44 시스-기이드
양성 대조군
1xMS2_PBS8
2xMS2_PBS8
4xMS2_PBS8
1xMS2_PBS13
2xMS2_PBS13
4xMS2_PBS13
1xMS2_scrPBS
2xMS2_scrPBS
4xMS2_scrPBS
구성 1_1MCP_A
구성 1_1MCP_B
구성 1_2MCP_A
구성 1_2MCP_B
구성 1_4MCP_A
구성 1_4MCP_B
구성 2_1MCP_A
구성 2(ML/MS)_1MCP_B
구성 2_2MCP_A
구성 2_2MCP_B
구성 2_4MCP_A
구성 2_4MCP_B
구성 3_1MCP_A
구성 3_1MCP_B
구성 3_2MCP_A
구성 3_2MCP_B
구성 3_4MCP_A
구성 3_4MCP_B

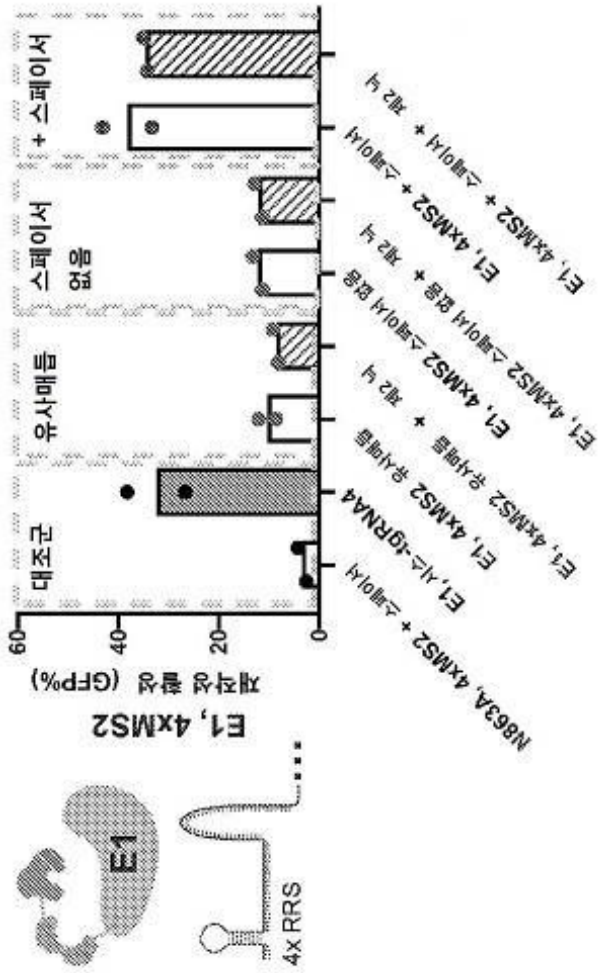
도면9



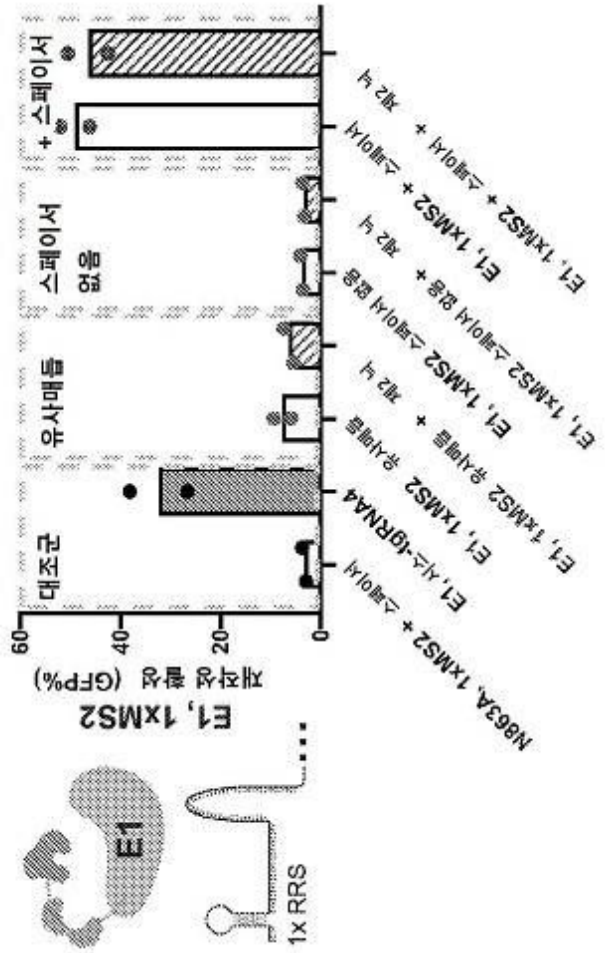
도면10


제작성 활성 (GFP%)
15
10
5
0
N863A
대조군
E1 + 4xMS2
E1 + 1xMS2
G2+ 4xMS2
G2 + 1xMS2
N863A, 4xMS2 스페이서 함유 (150ng)
N863A, 1xMS2 스페이서 함유 (75ng)
E1, 4xMS2 스페이서 함유 (37.5ng)
E1, 4xMS2 스페이서 함유 (150ng)
E1, 1xMS2 스페이서 함유 (75ng)
E1, 1xMS2 스페이서 함유 (37.5ng)
E1, 1xMS2 스페이서 함유 (150ng)
G2, 4xMS2 스페이서 함유 (75ng)
G2, 4xMS2 스페이서 함유 (37.5ng)
G2, 4xMS2 스페이서 함유 (150ng)
G2, 1xMS2 스페이서 함유 (75ng)
G2, 1xMS2 스페이서 함유 (37.5ng)
G2, 1xMS2 스페이서 함유 (150ng)

도면11a


E1, 4xMS2
재작성 활성 (GFP%)
0
20
40
60
대조군
유사펩틀
스페이서 없음
스페이서 + 스페이서
N863A, 4xMS2 + 스페이서
E1, 시스-gRNA4
E1, 4xMS2 유사펩틀
E1, 4xMS2 + 유사펩틀
E1, 4xMS2 스페이서 없음
E1, 4xMS2 + 스페이서 없음
E1, 4xMS2 스페이서
E1, 4xMS2 + 스페이서
E1
4x RRS

도면11b


E1, 1xMS2
재작성 활성 (GFP%)
60
40
20
0
대조군
유사매듭
스페이서
없음
스페이서
+ 스페이서
E1, 1xMS2 + 스페이서
E1, 1xMS2 없음 + 재2 닉
E1, 1xMS2 스페이서 없음
E1, 1xMS2 없음 + 재2 닉
E1, 1xMS2 유사매듭
E1, 1xMS2 + 재2 닉
E1, 시스-tgRNA4
E1, 1xMS2 + 스페이서
N863A, 1xMS2 + 스페이서
E1
1x RRS

서 열 목 록 (첨부)

아이콘을 클릭하시면 서열목록 파일이 열립니다.

본 공보 PDF는 첨부파일을 가지고 있습니다. Acrobat Reader PDF뷰어를 제공하지 않는 브라우저(크롬, 파이어폭스, 사파리 등)의 경우 첨부파일 열기가 제한되어 있으므로 Acrobat Reader PDF뷰어 설치 후 공보 PDF를 다운로드 받아 해당 뷰어에서 조회해주시기 바랍니다.