



등록특허 10-2565284



(19) 대한민국특허청(KR)

(12) 등록특허공보(B1)

(45) 공고일자 2023년08월08일

(11) 등록번호 10-2565284

(24) 등록일자 2023년08월04일

(51) 국제특허분류(Int. Cl.)

C12Q 1/6897 (2018.01)

(52) CPC특허분류

C12Q 1/6897 (2018.05)

(21) 출원번호 10-2018-7000086

(22) 출원일자(국제) 2016년07월28일

심사청구일자 2021년06월15일

(85) 번역문제출일자 2018년01월02일

(65) 공개번호 10-2018-0033498

(43) 공개일자 2018년04월03일

(86) 국제출원번호 PCT/IB2016/001183

(87) 국제공개번호 WO 2017/021776

국제공개일자 2017년02월09일

(30) 우선권주장

62/199,681 2015년07월31일 미국(US)

(56) 선행기술조사문헌

CN101775410 A\*

W02008106551 A2\*

US5670367 A

W02014043535 A1

\*는 심사관에 의하여 인용된 문헌

(73) 특허권자

버베리안 노딕 에이/에스

덴마크 헬레루프 필립 헤이만 알레 3 (우: 2900)

(72) 발명자

델케이어, 알랭

미국 95124 캘리포니아 새너제이 렉스포드 애비뉴 2812

리, 쟁지

미국 94582 캘리포니아 사 라몬 윈드 테라스 드라이브 336

라운트리, 리안

미국 캘리포니아 새너제이 브라이어글렌 드라이브 4029

(74) 대리인

특허법인 남앤남

전체 청구항 수 : 총 15 항

심사관 : 이재영

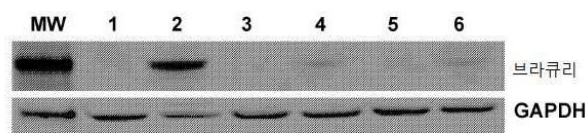
(54) 발명의 명칭 **폭스바이러스에서 발현을 증강하기 위한 프로모터**

## (57) 요약

본 발명은 이중성 유전자, 예를 들면, 브라큐리의 발현을 증강하는데 이용될 수 있는 하나 또는 그 이상의 프로모터 및/또는 발현 카세트에 관계한다. 특히, 하나 또는 그 이상의 프로모터 및/또는 발현 카세트는 바이러스 벡터, 예를 들면, 폭스바이러스의 일부로서 이중성 유전자의 발현을 증강한다.

## 대표도

인간 DCs에서 브라큐리의 발현



1. 감염되지 않은 DCs
2. MVA-브라큐리-TRICOM
3. FPV-WT
4. FPV-mBN343A
5. FPV-mBN344A
6. FPV-mBN345A

## 명세서

### 청구범위

#### 청구항 1

코딩 서열의 발현을 위한 발현 카세트로서,

(a) 서열 번호: 6 또는 서열 번호: 9를 갖고, 폭스바이러스 프로모터로서 활성적인 프로모터를 포함하는 핵산 서열; 및

(b) 상기 프로모터에 작동가능하게 연결된 이중성 코딩 서열로서, 코딩 서열은 브라큐리 항원을 인코딩하는, 이중성 코딩 서열을 포함하고,

상기 코딩 서열의 발현이 상기 프로모터에 의해 제어되는 것인 발현 카세트.

#### 청구항 2

청구항 1에 있어서, 코딩 서열은 L254V 돌연변이를 포함하는 브라큐리 항원을 인코딩하는 것인 발현 카세트.

#### 청구항 3

청구항 1에 있어서, 코딩 서열은 서열 번호: 11 및 서열 번호: 13으로 구성된 군에서 선택되는 브라큐리 항원을 인코딩하는 핵산 서열을 포함하는 것인 발현 카세트.

#### 청구항 4

청구항 1에 있어서, 코딩 서열은 서열 번호: 12, 서열 번호: 14 및 서열 번호: 15로 구성된 군에서 선택되는 것인 발현 카세트.

#### 청구항 5

청구항 1에 있어서, 코딩 서열은 다음과 같이 구성된 군에서 선택되는 브라큐리 항원을 인코딩하는 핵산을 포함하는 것인 발현 카세트: 서열 번호: 11, 13, 22, 또는 23.

#### 청구항 6

청구항 1에 있어서, 코딩 서열은 다음과 같이 구성된 군에서 선택되는 핵산을 포함하는 것인 발현 카세트: 서열 번호: 12, 14, 15, 20, 또는 21.

#### 청구항 7

서열 번호: 72 또는 서열 번호: 73을 갖는 핵산을 포함하는 발현 카세트.

#### 청구항 8

청구항 1 내지 청구항 7 중 어느 한 항의 발현 카세트를 포함하는 벡터.

#### 청구항 9

청구항 8에 있어서, 벡터는 오소폭스바이러스인 벡터.

#### 청구항 10

청구항 9에 있어서, 오소폭스바이러스는 MVA-BN인 변형된 양카라 바이러스 (MVA)인 벡터.

#### 청구항 11

청구항 1 내지 7 중 어느 한 항에 따른 발현 카세트를 포함하는, 백신, 약제, 또는 제약학적 조성물.

#### 청구항 12

청구항 1 내지 7 중 어느 한 항에 따른 발현 카세트를 포함하는 세포.

#### 청구항 13

코딩 서열을 표적 세포 내로 도입하기 위한 방법으로서, 청구항 1 내지 7 중 어느 한 항에 따른 발현 카세트를 표적 세포 내로 도입하는 것을 포함하는 방법.

#### 청구항 14

캡티드, 단백질, 또는 바이러스를 생산하기 위한 방법으로서, 청구항 8에 따른 벡터로 숙주 세포를 감염시키는 것을 포함하는 방법.

#### 청구항 15

코딩 서열을 표적 세포 내로 도입하는 방법에 사용하기 위한, 청구항 8에 따른 벡터.

#### 청구항 16

삭제

#### 청구항 17

삭제

#### 청구항 18

삭제

#### 청구항 19

삭제

#### 청구항 20

삭제

#### 청구항 21

삭제

#### 청구항 22

삭제

#### 청구항 23

삭제

#### 청구항 24

삭제

#### 청구항 25

삭제

#### 청구항 26

삭제

#### 청구항 27

삭제

청구항 28

삭제

청구항 29

삭제

청구항 30

삭제

청구항 31

삭제

청구항 32

삭제

청구항 33

삭제

청구항 34

삭제

청구항 35

삭제

청구항 36

삭제

청구항 37

삭제

청구항 38

삭제

청구항 39

삭제

청구항 40

삭제

청구항 41

삭제

청구항 42

삭제

청구항 43

삭제

청구항 44

삭제

청구항 45

삭제

청구항 46

삭제

청구항 47

삭제

청구항 48

삭제

청구항 49

삭제

**발명의 설명**

**기술 분야**

[0001] 발명의 분야

[0002] 본 발명은 폭스바이러스, 특히 조류폭스바이러스, 예를 들면, 계두바이러스 및 카나리아두바이러스에서 단백질 발현의 하나 또는 그 이상의 신규한 프로모터에 관계한다. 본 발명은 상기 프로모터를 포함하는 폴리뉴클레오티드, 특히 상기 프로모터 및 발현되는 핵산을 포함하는 폴리뉴클레오티드; 이를 포함하는 벡터; 상기 구조체를 포함하는 숙주 세포; 및 상기 중에서 한 가지를 포함하는 조성물에 더욱 관계한다. 추가적으로, 본 발명은 하나 또는 그 이상의 프로모터에 작동가능하게 연결된 하나 또는 그 이상의 신규한 폴리펩티드에 관계한다. 더욱 구체적으로, 하나 또는 그 이상의 신규한 폴리펩티드는 브라큐리 단백질에 관련된다.

**배경 기술**

[0003] 발명의 배경

[0004] 재조합 폭스바이러스는 감염성 생물체 및 더욱 최근에, 종양에 대한 백신으로서 이용되었다. Mastrangelo et al. J Clin Invest. 2000;105(8):1031-1034. 이들 폭스바이러스 군 중에서 2가지, 조류폭스바이러스 및 오소폭스바이러스는 종양과 싸우는데 효과적인 것으로 나타났고 잠재적 암 치료와 관련되었다. 동일 저자.

[0005] 한 가지 예시적인 조류폭스바이러스 중, 계두는 인간 투여를 위한 안전한 운반체인 것으로 나타났는데, 그 이유는 계두바이러스가 포유류 세포에 들어가고 단백질을 발현하지만, 불현적으로 복제하기 때문이다. Skinner et al. Expert Rev Vaccines. 2005 Feb;4(1):63-76. 추가적으로, 발현을 위한 운반체로서 계두바이러스의 이용은 암, 말라리아, 결핵, AIDS, 그리고 EBOLA에 대항하여 백신의 다양한 임상 시험에서 평가되고 있다. 동일 저자.

[0006] 재조합 폭스바이러스, 예를 들면, 계두는 여러 감염성 질환 및 종양 연관된 유전자, 예를 들면, p97, HER-2/neu, p53 및 ETA를 비롯한 넓은 범위의 삽입된 유전자를 발현하는데 이용되었다 (Paoletti, et al., 1993). 최근에 연구되는 한 가지 예시적인 종양 항원은 브라큐리 ("T"로서 또한 알려져 있음)이다.

[0007] 브라큐리는 생쥐에서, 또한 열성 치사성인 지배적인 짧은 꼬리 돌연변이체로서 확인되었다; 동형접합성 T/T 배아는 뒤쪽 중배엽 형성의 실패로 인해 임신 중기에서 죽는다 (Chesley, J. Exp. Zool., 70: 429-459, 1935). 뮐린 브라큐리 유전자 (Herrmann et al., Nature (Lond.), 343: 617-622, 1990)뿐만 아니라 다른 중, 예를 들면, 인간에서 동족체가 클로닝되었다. 생쥐 브라큐리의 인간 동족체의 발현은 임신 14-15 주에 인간 낙태아의 척삭 잔존물, 수질핵에서 RT-PCR에 의해 검출되었다 (Edwards et al., Genome Res., 6: 226-233, 1996).

[0008] 브라큐리는 중배엽 분화의 인식을 위한 귀중한 마커인 것으로 일반적으로 입증되었다 (Herrmann et al., Trends

Genet., 10: 280-286, 1994). 가령, 배아 그들 자체에서 발현과는 별개로, 브라큐리는 시험관내에서 중배엽 계통을 따라서 분화하는 일정한 뮤린 EC 및 ES 세포주의 분화 동안 활성화되는 것으로 보고되었다 (가령, Bain et al., Biochem. Biophys. Res. Commun., 223: 691-694, 1996을 참조한다). 인간에서, 브라큐리는 기형암종 (Gokhele et al., Cell Growth and Differentiation 11:157-62, 2000), 척삭종 (Vujovic et al., J. Pathol. 2: 157-65, 2006) 및 혈관모세포종 (Glasker et al., Cancer Res. 66: 4167-4172, 2006)에서 발현되는 것으로 밝혀졌다.

[0009] 더욱 최근에, 브라큐리를 발현하는 폭스바이러스는 종양에 대항하여 활성 면역치료제로서 효과적일 수 있는 것으로 설명되었다. (WO 2014/043535를 참조한다).

[0010] 활성 면역요법 및 백신을 비롯하여, 감염성 질환 및 암 치료를 향상시키기 위한 실제적인 충족되지 않은 의료 수요가 분명히 존재한다. 상기에 근거하여, 활성 면역요법에서 이용되는, 브라큐리와 같은 항원의 발현을 향상시키기 위한 요구가 당분야에서 존재한다. 본 발명은 이러한 필요를 충족시킨다.

## 발명의 내용

### 해결하려는 과제

[0011] 발명의 짧은 요약

[0012] 본 발명은 폭스바이러스의 일부로서 통합된 코딩 서열 및 항원의 발현에서 증강과 연관된 하나 또는 그 이상의 프로모터, 핵산, 발현 카세트, 재조합 펩티드 및 재조합 폭스바이러스를 제공한다.

[0013] 본 발명의 한 양상에서, 다음을 포함하는 발현 카세트가 제시된다:

[0014] a. 서열 번호: 1-10 및 77 중에서 한 가지에 최소한 70% 동일성을 갖는 핵 서열을 포함하는 프로모터; 및

[0015] b. 상기 프로모터에 작동가능하게 연결된 코딩 서열;

[0016] 여기서 상기 코딩 서열의 발현이 상기 프로모터에 의해 제어된다.

[0017] 본 발명의 추가의 양상에서, 발현 카세트는 그 안에 코딩 서열의 증강된 발현을 위한 벡터, 예를 들면, 바이러스 또는 플라스미드의 일부로서 통합될 수 있다. 바람직하게는, 발현 카세트는 폭스바이러스, 예를 들면, 하지만 제한 없이, 오소폭스바이러스 및 조류폭스바이러스의 일부이다. 더욱 바람직하게는, 발현 카세트는 조류폭스바이러스, 예를 들면, 계두바이러스이다. 다른 추가의 양상에서, 발현 카세트는 폭스바이러스의 유전체에서 자연적으로 발생하는 발현 카세트가 아니다.

[0018] 추가의 양상에서, a) 서열 번호: 1-10으로 구성된 군에서 선택되는 핵산 서열, b) 서열 번호: 1-10 및 77 중에서 한 가지의 하위서열, 또는 c) 서열 번호: 1-10 및 77 중에서 한 가지의 유도체 서열을 포함하거나 또는 이것으로 구성되는 하나 또는 그 이상의 프로모터가 제시되고, 여기서 상기 유도체는 하나 또는 그 이상의 치환, 결실 및/또는 삽입을 갖고, 그리고 여기서 상기 유도체 서열은 폭스바이러스 프로모터로서 활성적이고 및/또는 폭스바이러스 감염된 세포에서 활성적이다.

[0019] 다른 추가 양상에서, 본 발명은 다음과 같이 구성된 군에서 선택되는 하나 또는 그 이상의 신규한 브라큐리 단백질을 인코딩하는 하나 또는 그 이상의 핵산 서열을 포함한다: 서열 번호: 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71. 다른 추가 양상에서, 본 발명은 다음과 같이 구성된 군에서 선택되는 하나 또는 그 이상의 신규한 핵산 서열을 포함한다: 16-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69.

[0020] 다른 추가의 양상에서, 본 발명은 다음을 포함하는 발현 카세트를 포함한다:

[0021] a. 서열 번호: 1-10 및 77 중에서 한 가지에 최소한 70% 동일성을 갖는 핵산 서열을 포함하는 프로모터; 및

[0022] b. 상기 프로모터에 작동가능하게 연결된 코딩 서열, 여기서 상기 코딩 서열은 i) 서열 번호: 11, 13, 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71로 구성된 군에서 선택되는 브라큐리 단백질을 인코딩하는 핵산을 포함하거나, 또는 ii) 서열 번호: 12, 14-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69로 구성된 군에서 선택되는 핵산 중에서 한 가지에 최소한 70% 동일성을 갖는 핵산을 포함하거나 또는 이것으로 구성되고;

[0023] 여기서 상기 코딩 서열의 발현이 상기 프로모터에 의해 제어된다.

[0024] 본 발명의 추가 목적과 이점이 부분적으로, 아래의 설명에서 진술될 것이고, 그리고 부분적으로, 이러한 설명으로부터 명백할 것이고, 또는 본 발명의 실시예에 의해 학습될 수 있다. 본 발명의 목적과 이점은 특히 첨부된 청구항에서 지적된 요소와 조합에 의하여 실현되고 획득될 것이다.

[0025] 진술한 일반적인 설명 및 다음의 상세한 설명 둘 모두 단지 예시적이고 설명적이며 청구된 발명을 제한하지 않는 것으로 이해된다.

## 도면의 간단한 설명

[0026] 도면의 간단한 설명

도면 1은 비재조합 바이러스 FPV-WT, 또는 재조합 바이러스 FPV-mBN343A, FPV-mBN344A 또는 FPV-mBN345A로 감염된 인간 수지상 세포 (DCs)에서 브라큐리 단백질의 발현을 도해한다. 실시예 1에서 상술된 바와 같은 토끼 단일클론 항-브라큐리 항체를 이용하여 웨스턴 블롯 분석이 수행되었다.

도면 2는 비재조합 바이러스 MVA-WT 또는 FPV-WT, 또는 재조합 바이러스 MVA-브라큐리-TRICOM, FPV-mBN249B, FPV-mBN281A 클론 32, FPV-mBN281A 클론 35, FPV-mBN343A, FPV-mBN344A, FPV-mBN345A, FPV-mBN354A 또는 FPV-mBN355A로 감염된 인간 수지상 세포 (DCs)에서 브라큐리 단백질의 발현을 도해한다. 실시예 2에서 상술된 바와 같은 토끼 단일클론 항-브라큐리 항체를 이용하여 웨스턴 블롯 분석이 수행되었다.

도면 3은 실시예 2에서 상술된 바와 같이, 비재조합 바이러스 MVA-WT 또는 FPV-WT, 또는 재조합 바이러스 MVA-브라큐리-TRICOM, FPV-mBN249B, FPV-mBN281A 클론 32, FPV-mBN281A 클론 35, FPV-mBN343A, FPV-mBN344A, FPV-mBN345A, FPV-mBN354A 또는 FPV-mBN355A로 감염된 인간 수지상 세포 (DCs)의 웨스턴 블롯 분석으로부터 GAPDH와 비교하여 브라큐리 단백질의 상대적 발현을 묘사한다.

도면 4는 실시예 3에서 설명된 바와 같이, 각 단백질에 특정한 형광으로 표지화된 항체를 이용한 유세포분석법에 의해 사정된, 재조합 바이러스 MVA-브라큐리-TRICOM, FPV-mBN249B, FPV-mBN343A 또는 FPV-mBN345A로 감염된 CMMT 세포 (히말라야 원숭이 유방 종양 세포주)에서 브라큐리 및 TRICOM 단백질의 발현을 도해한다.

도면 5는 실시예 3에서 설명된 바와 같이, 유세포분석법에 의해 사정된, 재조합 바이러스 MVA-브라큐리-TRICOM, FPV-mBN249B, FPV-mBN343A 또는 FPV-mBN345A로 감염된 CMMT 세포 (히말라야 원숭이 유방 종양 세포주)에서 브라큐리 단백질의 중앙 발현 수준을 묘사한다.

도면 6은 실시예 4에서 설명된 바와 같이, 유세포분석법에 의해 사정된, 재조합 바이러스 FPV-mBN345B로 감염된 CMMT 세포 (히말라야 원숭이 유방 종양 세포주)에서 브라큐리 및 TRICOM 단백질의 발현을 묘사한다.

## 발명을 실시하기 위한 구체적인 내용

[0027] 발명의 상세한 설명

[0028] 본 발명은 서열 번호:1-10 및 77에서 명시된 바와 같은 프로모터 서열이 종양 항원 브라큐리의 발현을 증강한다는 놀라운 결정에 근거된다. 도면 1-6에서 도시되고 본원에서 더욱 상세하게 설명된 바와 같이, 브라큐리 항원의 발현은 본 발명의 프로모터를 이용할 때 증강된다. 발현은 이들 프로모터 및 브라큐리 항원이 재조합 폭스바이러스의 일부로서 이용될 때 더욱 증강된다.

[0029] 최소한 하나의 양상에서, 본 발명의 다양한 구체에는 공지된 우두 프로모터, 예를 들면, PrS 및 우두바이러스 40k (VV-40K)를 이용할 때 브라큐리 단백질의 불충분한 발현 수준의 결과로서 창출되었다. 브라큐리 발현 수준을 증강하려는 시도에서, 본 발명자들은 FPV에서 다양한 우두 프로모터 및 연관된 단백질 (가령, VV-40k, I3 등) 및 임의의 가능한 동족체를 분석하였다. 본 발명자들은 일부 FPV 상동성 서열이 이전에 발견되었다는 것을 인식하였다. 가령, Zantinge, J Gen Virol. 1996 Apr;77 ( Pt 4):603-14 및 NCBI 참고 서열: NP\_039051.1에서 유전자 FPV 088, Afonso, C.L et al., J. Virol. 74 (8), 3815-3831 (2000)을 참조한다.

[0030] 브라큐리 발현을 증강하려는 초기 시도에서, NCBI 참고 서열: NP\_039051.1에서 유전자 FPV 088의 가능한 프로모터 영역이 본 발명자들에 의해 브라큐리로 작제되고 시험되어 바람직하지 않은 결과 (가령, mBN344A에서 도면 1 및 2를 참조한다) 프로모터가 산출되었다. 본 발명자들은 유전자 FPV 088의 ORF로부터 뉴클레오티드의 부가적 차후 프로모터를 창출하였는데, 이것은 바람직하지 않은 결과를 유사하게 산출하였다. (가령, mBN354A에서 도면 1 및 2를 참조한다). 유전자 FPV 088의 ORF로부터 다른 추가 뉴클레오티드로 본 발명자들에 의해 추가 구조체가 창출되었는데, 이들은 본원에서 설명되고 예증된 바와 같이, 종양 항원 브라큐리의 발현을 증강한다.

- [0031] 따라서, 다양한 구체예에서, 본 발명은 발현 벡터, 예를 들면, 하지만 제한 없이, 플라스미드, 재조합 바이러스 기타 등등에서 코딩 서열 및/또는 핵산의 증강/발현을 위한 하나 또는 그 이상의 핵산 서열, 이들 핵산 서열을 포함하는 하나 또는 그 이상의 프로모터, 이들 핵산을 포함하는 하나 또는 그 이상의 발현 카세트, 하나 또는 그 이상의 펩티드 및/또는 펩티드 서열에 관계한다. 추가 구체예에서, 본 발명은 본원에서 설명된 핵산 서열, 프로모터, 발현 카세트, 및/또는 펩티드 및/또는 펩티드 서열 중에서 하나 또는 그 이상을 포함하는 하나 또는 그 이상의 재조합 폭스바이러스에 관계한다.
- [0032] 추가적으로, 본 발명의 다양한 구체예에서, 브라큐리 폴리펩티드를 인코딩하는 하나 또는 그 이상의 핵산 서열이 제시된다. 한 구체예에서, 브라큐리 폴리펩티드를 인코딩하는 하나 또는 그 이상의 핵산은 다음과 같이 구성된 군에서 선택된다: 서열 번호: 11, 13, 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71. 본 발명에 따른 브라큐리 폴리펩티드를 인코딩하는 뉴클레오티드 서열은 다음과 같이 구성된 군에서 선택된다: 서열 번호: 12, 14-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69.
- [0033] 정의
- [0034] 본원에서 이용된 바와 같이, 단수 형태 ("a", "an" 및 "the")는 문맥에서 달리 지시되지 않으면, 복수 지시대상을 포함하는 것으로 유의되어야 한다. 따라서, 예로서, "에피토프"에 대한 언급은 에피토프 중에서 하나 또는 그 이상을 포함하고, 그리고 "방법"에 대한 언급은 본원에서 설명된 방법에 대해 변형되거나 또는 이를 대체할 수 있는, 당업자에게 공지된 동등한 단계 및 방법에 대한 언급을 포함한다.
- [0035] 달리 지시되지 않으면, 일련의 요소를 선행하는 용어 "최소한"은 일련 내에 모든 요소를 지칭하는 것으로 이해된다. 당업자는 단지 일과적인 실험을 이용하여, 본원에서 설명된 발명의 특정한 구체예에 대한 많은 등가물을 인식하거나, 또는 확인할 수 있을 것이다. 이런 등가물은 본 발명에 의해 포괄되는 것으로 의도된다.
- [0036] 본 명세서 및 하기의 청구항 전반에서, 문맥에서 달리 요구되지 않으면, 단어 "포함한다" 및 변이, 예를 들면, "포함한다" 및 "포함하는"은 언급된 전체 또는 단계, 또는 전체 또는 단계의 군의 포함, 하지만 임의의 다른 전체 또는 단계, 또는 전체 또는 단계의 군의 배제하지 않음을 암시하는 것으로 이해될 것이다. 본원에서 이용될 때, 용어 "포함하는"은 용어 "내포하는" 또는 "포함하는" 또는 때때로, 본원에서 이용될 때 용어 "갖는"으로 치환될 수 있다. 비록 덜 바람직하긴 하지만, 임의의 전술한 용어 (포함하는, 내포하는, 포함하는, 갖는)은 본 발명의 양상 또는 구체예의 맥락에서 본원에서 이용될 때는 언제든지, 용어 "구성되는"으로 치환될 수 있다. 본원에서 이용될 때, "구성되는"은 청구항 요소에서 특정되지 않은 임의의 원소, 단계, 또는 성분을 배제한다. 본원에서 이용될 때, "본질적으로 구성되는"은 청구항의 기본 특징 및 신규한 특징에 실질적으로 영향을 주지 않는 물질 또는 단계를 배제하지 않는다.
- [0037] 본원에서 이용된 바와 같이, 복수의 언급된 요소 사이에 연결 용어 "및/또는"은 개별 옵션 및 합동된 옵션 둘 모두를 포괄하는 것으로 이해된다. 가령, 2개의 원소가 "및/또는"에 의해 연합되는 경우에, 첫 번째 옵션은 두 번째 요소 없이 첫 번째 요소의 적용가능성을 지칭한다. 두 번째 옵션은 첫 번째 요소 없이 두 번째 요소의 적용가능성을 지칭한다. 세 번째 옵션은 첫 번째와 두 번째 요소 공동의 적용가능성을 지칭한다. 이들 옵션 중에서 한 가지가 상기 의미 내에 들어가는 것으로 이해되고, 그리고 이런 이유로, 본원에서 이용된 바와 같은 용어 "및/또는"의 요건을 충족시킨다. 이들 옵션 중에서 하나 이상의 동시 적용가능성 또한, 상기 의미 내에 들어가는 것으로 이해되고, 그리고 이런 이유로, 용어 "및/또는"의 요건을 충족시킨다.
- [0038] 본원에서 이용된 바와 같이, 용어 "프로모터"는 예로서, 증진되는 유전자의 코딩 영역으로부터 RNA를 합성하는 것을 책임지는 단백질 전사 인자 및 중합효소에 의해 인식되고 결합되는 특정한 DNA 서열 요소를 내포하는, 발현되는 핵산의 서열의 상류에 위치한 핵산, 통상적으로 DNA의 조절 영역을 표시한다. 프로모터가 전형적으로, 문제되는 유전자에 바로 인접하기 때문에, 프로모터에서 위치는 특정 유전자에 대한 DNA의 전사가 시작되는 전사 시작 부위에 상대적으로 지정된다 (다시 말하면, 상류 위치는 -1로부터 역으로 계수되는 음수이다, 예를 들면, -100은 100개 염기쌍 상류 위치이다). 따라서, 프로모터 서열은 위치 -1까지 뉴클레오티드를 포함할 수 있다. 하지만, 위치 +1로부터 뉴클레오티드는 프로모터의 일부가 아니다, 다시 말하면, 이점에 관하여 주목해야 할 것은 번역 개시 코돈 (ATG 또는 AUG)이 프로모터의 일부가 아니라는 점이다. 따라서, 서열 번호: 1-10 및 77은 본 발명의 프로모터를 포함하는 폴리뉴클레오티드이다.
- [0039] 본원에서 이용된 바와 같이, 코딩 서열, 핵산, 단백질 및/또는 항원의 발현 수준에 대하여 이용될 때 용어 "증강하는" 또는 "증강된"은 당분야에서 공지된 프로모터 중에서 하나 또는 그 이상, 예를 들면 PrS 또는 VV-40k와

연관될 때 및/또는 이것의 일부로서 코딩 서열, 핵산, 단백질 및/또는 항원의 발현 수준에 비하여, 본 발명의 프로모터, 발현 카세트, 핵산, 단백질 및/또는 벡터 중에서 하나 또는 그 이상과 연관될 때 및/또는 이것의 일부로서 코딩 서열, 핵산, 단백질 및/또는 항원의 발현에서 증가를 지칭한다.

[0040] 본원에서 이용된 바와 같이, 서열 번호: 1-10 및 77에서 명시된 뉴클레오티드 서열과 "본질적으로 동일한 발현 특징"을 갖는 뉴클레오티드 서열은 생산된 재조합 단백질의 양에 의해 측정될 때, 서열 번호: 1-10 및 77의 프로모터 활성의 최소한 70%, 바람직하게는 최소한 80%, 이보다 더욱 바람직하게는 최소한 90%를 전시할 것이다. 문제되는 프로모터 서열이 서열 번호: 1-10 및 77 중에서 한 가지와 "본질적으로 동일한 발현 특징"을 갖는지의 여부는 본 출원의 실시예 1-4에서 진술된 방법을 이용하여 당업자에 의해 쉽게 결정될 수 있다. 본 발명에 따른 프로모터는 바람직하게는, 폭스바이러스 프로모터, 바람직하게는 조류폭스바이러스로서 활성적이거나, 또는 폭스바이러스 감염된 세포, 바람직하게는 조류폭스바이러스 감염된 세포에서 프로모터로서 활성적이다. 조류폭스바이러스는 바람직하게는 계두바이러스이다. "폭스바이러스 프로모터로서 활성적"은 프로모터가 폭스바이러스로 세포의 감염 후 상기 폭스바이러스 내에서 작동가능하게 연결된 유전자의 발현을 주도할 수 있다는 것을 의미한다. 세포는 바람직하게는, 폭스바이러스의 후기 및/또는 초기 및/또는 초기/후기 발현을 허용하는 세포이다. "폭스바이러스 감염된 세포에서 활성적인 프로모터"는 프로모터가 폭스바이러스 유전체의 일부가 아닌, 예를 들면, 플라스미드 또는 선형 폴리뉴클레오티드 또는 비-폭스바이러스 바이러스 유전체의 일부인 환경을 또한 포함한다; 이런 환경에서 본 발명에 따른 프로모터는 상기 프로모터를 포함하는 세포가 폭스바이러스 유전체를 또한 포함하면, 예를 들면, 상기 세포가 폭스바이러스로 감염되면 활성적이다. 이들 환경 하에, 바이러스 RNA 중합효소는 본 발명에 따른 프로모터를 인식하고, 그리고 상기 프로모터에 연결된 유전자/코딩 서열의 발현이 활성화된다.

[0041] 본원에서 이용된 바와 같이, 용어 "서열 번호: 1-10 및 77에서 명시된 핵산으로부터 유래된"은 서열 번호: 1-10 및 77의 뉴클레오티드 서열이 특정된 뉴클레오티드 변형, 예를 들면, 최소한 하나의 뉴클레오티드 부가, 결실, 치환 및/또는 전도를 달성하기 위한 기초로서 받아들여진다는 것을 의미한다. 용어 "유래된"은 예로서, 공지된 방법, 예를 들면, 오류 가능성 PCR에 의해 서열 번호: 1-10 및 77에 상응하는 물리적 서열을 실질적으로 변형하는 가능성을 포함한다. 용어 "유래된"은 서열 번호: 1-10 및 77의 서열에서 변형을 인실리코 수행하고, 그리고 이후, 이렇게 결정된 서열을 물리적 핵산으로서 합성하는 가능성을 부가적으로 포함한다. 가령, 용어 "유래된"은 예로서, 혼성화 안정성에 대하여 핵산 서열의 분석을 위해 임의의 공지된 컴퓨터 프로그램을 이용하는 가능성, 그리고 서열 번호: 1-10 및 77의 시작 서열을 변형함에 있어서 임의의 이차 핵산 구조의 가능성을 포괄한다. 바람직하게는, 10, 9, 8, 7, 6, 5, 4, 3, 2 또는 1개보다 많지 않은 뉴클레오티드가 서열 번호: 1-10 및 77의 핵산에 부가되고, 이들로부터 결실되고, 치환되고 및/또는 반전되었다. 게다가, 최소한 하나의 뉴클레오티드의 부가, 삽입 또는 결실은 핵산의 개시 코돈이 발현되도록 유발하지 않아야 한다.

[0042] 본원에서 이용된 바와 같이, 용어 "발현된", "발현하다", "발현" 등은 관심되는 서열의 전사 단독뿐만 아니라 전사 및 번역 둘 모두를 표시한다. 따라서, DNA의 형태로 존재하는 핵산의 발현을 지칭함에 있어서, 이러한 발현으로부터 발생하는 산물은 RNA (발현되는 서열의 전사 단독으로부터 발생) 또는 폴리펩티드 서열 (발현되는 서열의 전사 및 번역 둘 모두로부터 발생)일 수 있다. 용어 "발현"은 따라서, RNA 및 폴리펩티드 산물 둘 모두 상기 발현으로부터 발생하고 동일한 공유된 환경 내에 함께 남아있는 가능성을 또한 포함한다. 가령, mRNA가 폴리펩티드 산물로의 번역 이후에 존속할 때의 사례가 여기에 해당한다.

[0043] 본원에서 이용된 바와 같이, 용어 "발현 카세트"는 클로닝 및/또는 형질전환에 전형적으로 이용되는 벡터 또는 재조합 바이러스의 부분으로서 규정된다. 발현 카세트는 전형적으로, a) 하나 또는 그 이상의 코딩 서열 (가령, 개방 해독틀 (ORF), 유전자, 단백질 및/또는 항원을 인코딩하는 핵산), 그리고 b) 하나 또는 그 이상의 코딩 서열의 발현을 제어하는 서열로 구성된다. 추가적으로, 발현 카세트는 진핵생물에서 폴리아데닐화 부위를 통상적으로 내포하는 3' 비번역 영역을 포함할 수 있다.

[0044] 용어 "재조합"은 자연에서 발생하지 않거나, 또는 자연에서 발견되지 않는 배열에서 다른 폴리뉴클레오티드에 연결된 반합성, 또는 합성 기원의 폴리뉴클레오티드 또는 폴리펩티드를 의미한다.

[0045] 본 발명의 한 가지 양상은 서열 번호: 1-10 및 77과 최소한 70%, 바람직하게는 75%, 80%, 85%, 90% 또는 95% 동일성을 갖고, 그리고 서열 번호: 1-10 및 77과 본질적으로 동일한 발현 특징을 갖는 핵산을 포함하는 프로모터를 포함한다.

[0046] 본원에서 설명된 핵산 서열에 대하여 "퍼센트 (%) 서열 상동성 또는 동일성"은 최대 퍼센트 서열 동일성이 달성 되도록 서열을 정렬하고 필요하면 갭을 도입한 후, 그리고 임의의 보존성 치환을 서열 동일성의 일부로서 고려

하지 않고, 참고 서열 (다시 말하면, 이것이 유래되는 핵산 서열)에서 뉴클레오티드와 동일한, 후보 서열 내에 뉴클레오티드의 백분율로서 규정된다. 퍼센트 뉴클레오티드 서열 동일성 또는 상동성을 결정하는 목적을 위한 정렬은 당분야의 기술 범위 안에 있는 다양한 방식으로, 예를 들면, 공개적으로 가용한 컴퓨터 소프트웨어, 예를 들면, BLAST, ALIGN 또는 Megalign (DNASTAR) 소프트웨어를 이용하여 달성될 수 있다. 당업자는 비교되는 서열의 전장에 걸쳐 최대 정렬을 달성하기 위해 필요한 임의의 알고리즘을 비롯하여, 정렬을 계속하기 위한 적절한 파라미터를 결정할 수 있다.

[0047] 가령, 핵산 서열에 대한 적절한 정렬은 Smith and Waterman, (1981), Advances in Applied Mathematics 2:482- 489의 국부 상동성 알고리즘에 의해 제공된다. 이러한 알고리즘은 Dayhoff, Atlas of Protein Sequences and Structure, M. O. Dayhoff ed., 5 suppl. 3:353-358, National Biomedical Research Foundation, Washington, D.C., USA에 의해 개발되고, 그리고 Gribskov (1986), Nucl. Acids Res. 14(6):6745-6763에 의해 정규화된 채점 매트릭스를 이용함으로써 아미노산 서열에 적용될 수 있다. 서열의 퍼센트 동일성을 결정하기 위한 이러한 알고리즘의 예시적인 실행은 Genetics Computer Group (Madison, Wis.)에 의해 "BestFit" 유틸리티 응용프로그램에서 제공된다. 이러한 방법에 대한 디폴트 파라미터는 Wisconsin 서열 분석 패키지 프로그램 매뉴얼, 버전 8 (1995)에서 설명된다 (Genetics Computer Group, Madison, Wis.로부터 가용함). 본 발명의 맥락에서 퍼센트 동일성을 확립하는 바람직한 방법은 University of Edinburgh에 의해 저작권이 소유되고, John F. Collins 및 Shane S. Sturrock에 의해 개발되고, 그리고 IntelliGenetics, Inc. (Mountain View, Calif)에 의해 배포된 프로그램의 MPSRCH 패키지를 이용하는 것이다. 패키지의 이러한 스위트로부터, Smith-Waterman 알고리즘이 이용될 수 있는데, 여기서 디폴트 파라미터가 채점 표에 이용된다 (가령, 12의 갭 열린 패널티, 1의 갭 연장 패널티, 그리고 6의 갭). 산출된 데이터로부터, "정합" 값은 "서열 동일성"을 반영한다. 서열 사이에 퍼센트 동일성 또는 유사성을 계산하기 위한 다른 적합한 프로그램은 당분야에서 전 반적으로 공지되어 있다, 예를 들면, 다른 정렬 프로그램은 디폴트 파라미터에서 이용된 BLAST이다. 가령, BLASTN 및 BLASTp가 다음의 디폴트 파라미터를 이용하여 이용될 수 있다: 유전자 코드 = 표준; 필터 = 없음; 가닥 = 양쪽; 컷오프 = 60; 기대 = 10; 매트릭스 = BLOSUM62; 설명 = 50 서열; 분류 = 높은 점수; 데이터베이스 = 비다중, GenBank+EMBL+DDBJ+PDB+ GenBank CDS 번역+Swiss 단백질+Spupdate+PIR. 이들 프로그램의 상세는 다음의 인터넷 주소: [http:// blast.ncbi.nlm.nih.gov](http://blast.ncbi.nlm.nih.gov)에서 발견될 수 있다.

#### [0048] 프로모터 및 핵산 서열

[0049] 한 구체예에서, 본 발명은 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 핵산 서열을 포함한다.

[0050] 다른 구체예에서, 본 발명은 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 핵산과 최소한 70% 동일성을 갖고, 그리고 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 핵산 서열과 본질적으로 동일한 발현 특징을 갖는 핵산 서열을 포함한다. 추가 구체예에서, 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 핵산과 최소한 75%, 80%, 85%, 90%, 또는 95% 동일성을 갖고, 그리고 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 핵산 서열과 본질적으로 동일한 발현 특징을 갖는 핵산 서열이 제시된다.

[0051] 다른 추가 구체예에서, 서열 번호: 1-10 및 77의 뉴클레오티드 서열과 비교하여 최소한 하나의 뉴클레오티드 부가, 결실, 치환 및/또는 전도를 포함하는, 서열 번호: 1-10 및 77에서 명시된 핵산으로부터 유래된 뉴클레오티드 서열을 갖고, 그리고 서열 번호: 1-10 및 77의 핵산과 본질적으로 동일한 발현 특징을 갖는 핵산이 제시된다.

[0052] 다른 추가 구체예에서, 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 핵산 서열에 혼성화할 수 있는 뉴클레오티드 서열을 갖고, 그리고 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 핵산 서열과 본질적으로 동일한 발현 특징을 갖는 핵산이 제시된다.

[0053] 본 발명의 한 양상에서, 코딩 및/또는 유전자 서열의 발현을 증강하기 위한 프로모터가 제시되고, 상기 프로모터는 다음과 같이 구성된 군에서 선택된다:

[0054] a) 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 핵산 서열 또는 이의 하위서열;

[0055] b) (a)의 뉴클레오티드 서열과 비교하여 최소한 하나의 뉴클레오티드 부가, 결실, 치환 및/또는 전도를 포함하는, (a)에서 명시된 핵산으로부터 유래된 뉴클레오티드 서열을 갖고, 그리고 (a)의 핵산과 본질적으로 동일한 발현 특징을 갖는 핵산;

[0056] c) (a)의 핵산과 최소한 70% 동일성을 갖고, 그리고 (a)의 핵산과 본질적으로 동일한 발현 특징을 갖는 핵산 서

열; 및

- [0057] d) (a), (b) 또는 (c)의 핵산 서열에 혼성화할 수 있고, 그리고 (a)의 핵산과 본질적으로 동일한 발현 특징을 갖는 핵산.
- [0058] e) 서열 번호:3, 4, 5, 6, 7, 8, 또는 9를 포함하는 핵산 서열, 여기서 상기 핵산 서열은 최대 245개 뉴클레오티드를 포함한다.
- [0059] 본 발명의 추가 양상은 (a), (b) 또는 (c)의 핵산 서열에 혼성화할 수 있고, 그리고 (a)의 핵산과 본질적으로 동일한 발현 특징을 갖는 핵산에 관계한다. 용어 "혼성화할 수 있는"은 (a), (b) 또는 (c)의 핵산의 임의의 부분이 (a)의 핵산과 본질적으로 동일한 발현 특징을 갖는 임의의 DNA 서열의 검출 및 단리를 허용하도록 다른 DNA 서열에 혼성화할 수 있는 조건에서 혼성화를 의미한다. 혼성화는 엄격한 조건에서 실행된다; 조건이 더욱 엄격할수록, 부분적으로 상보성 서열이 떨어지도록 강제될 가능성이 더욱 높다, 다시 말하면, 더욱 높은 엄격함은 혼성화의 확률을 낮춘다. 실제로, 용어 "엄격한 조건"은 높은 온도 (가령, 65℃) 및/또는 낮은 염 농도 (가령, 0.1xSSC)를 갖는 혼성화 조건을 의미한다. 온도를 증가시키고 및/또는 염 농도를 감소시키는 것은 엄격함을 증가시킨다. 엄격한 조건 하에 혼성화는 서열 사이에 최소한 70% 또는 바람직하게는 최소한 75, 80, 85, 90 또는 95% 동일성이 존재하는 경우에만 발생할 것이다. 혼성화는 Ausubel et al. (2002) Short Protocols in Molecular Biology Vol. 1 5th ed. Canada에서 명시된 바와 같이 실행될 수 있다.
- [0060] 바람직하게는, 본 발명의 프로모터 및/또는 핵산은 245 또는 242개 뉴클레오티드까지의 길이를 갖는다. 또한, 본 발명의 프로모터 및/또는 핵산은 232 또는 215개 뉴클레오티드까지의 길이를 갖는다. 더욱 바람직하게는, 본 발명의 프로모터는 212, 200개까지의 길이, 또는 195개 뉴클레오티드까지의 길이, 또는 183개까지의 길이, 또는 170 또는 167개 뉴클레오티드까지의 길이를 갖는다. 한 구체예에서, 상기에서 명시된 바와 같이 (b), (c) 또는 (d)의 핵산은 약 50-250개 뉴클레오티드, 약 60-220개 뉴클레오티드, 약 70-200개 뉴클레오티드, 약 80-190개 뉴클레오티드, 또는 약 90-180개 뉴클레오티드를 포함한다.
- [0061] 대안으로, 이들 프로모터의 유도체를 이용하는 것이 본 발명의 범위 안에 있는데, 이것은 서열 번호: 1-10 및 77 중에서 한 가지에서 규정된 바와 같은 서열의 하위서열일 수 있다. 용어 "서열 번호: 1-10 및 77 중에서 한 가지에 따른 서열의 하위서열"은 프로모터로서, 특히 오소포스바이러스, 예를 들면, FFPV에서, 또는 오소포스바이러스, 예를 들면, FFPV 감염된 세포에서 프로모터로서 여전히 활성적인 서열 번호: 1 내지 10 및 77 중에서 한 가지의 더욱 짧은 단편을 지칭한다.
- [0062] 다양한 구체예에서, 서열 번호:9의 예시적인 하위서열은 서열 번호:2-8에서 선택된다. 한 양상에서, 서열 번호: 4 또는 5는 서열 번호: 9의 선호되는 하위서열이다. 다른 양상에서, 서열 번호: 8의 하위서열은 서열 번호: 2-7을 포함하고, 서열 번호:7의 하위서열은 서열 번호: 2-6을 포함하고, 서열 번호:6의 하위서열은 서열 번호: 2-5를 포함하고, 서열 번호:5의 하위서열은 서열 번호: 2-4를 포함하고, 서열 번호:6의 하위서열은 서열 번호: 2-5를 포함하고, 서열 번호:5의 하위서열은 서열 번호: 2-4를 포함하고, 서열 번호: 4의 하위서열은 서열 번호: 2-3을 포함하고, 그리고 서열 번호: 3의 하위서열은 서열 번호: 2를 포함한다.
- [0063] 서열 번호: 1-10 및 77 중에서 한 가지 또는 이의 하위서열의 뉴클레오티드 서열을 포함하거나 또는 이것으로 구성되는 프로모터의 유도체는 또한, 서열 번호: 1-10 및 77의 서열 중에서 한 가지에 대하여 하나 또는 그 이상의 뉴클레오티드 치환, 결실 및/또는 삽입을 갖는 서열일 수 있다. 본 발명에 따른 유도체는 프로모터로서, 특히 오소포스바이러스 및/또는 조류포스바이러스 바이러스 감염된 세포에서 오소포스바이러스 및/또는 조류포스바이러스 프로모터로서, 더욱 바람직하게는 조류포스바이러스 감염된 세포에서 조류포스바이러스 프로모터로서 여전히 활성적이다. 본 발명에 따른 유도체에서 결실, 치환 및 삽입이 한 서열에서 합동될 수 있다.
- [0064] 바람직하게는, 유도체는 본원에서 설명된 바와 같은 서열 번호: 1-10 및 77의 서열 중에서 한 가지 또는 이의 하위서열과 비교할 때 최소한 40%, 더욱 바람직하게는 최소한 60%, 이보다 더욱 바람직하게는 최소한 80%, 가장 바람직하게는 최소한 90%의 상동성을 갖는다.
- [0065] **브라큐리 단백질**
- [0066] 본 발명의 여러 양상에서, 본 발명자들은 본원에서 설명된 다양한 프로모터를 합성한 결과로서, 본 발명자들이 하나 또는 그 이상의 재조합 브라큐리 단백질을 창출했다는 것을 확정하였다. 특히, 하나 또는 그 이상의 재조합 브라큐리 단백질은 본 발명의 프로모터 중에서 하나 또는 그 이상 및 브라큐리 단백질의 조합으로부터 발생하는 융합 단백질이다. 이들 재조합 브라큐리 단백질 중에서 하나 또는 그 이상이 백신, 약제 또는 다른 치료

조성물의 일부로서 유용할 수 있을 것으로 예기된다.

- [0067] 따라서, 다양한 양상에서, 본 발명은 하나 또는 그 이상의 합성 및 신규한 재조합 브라큐리 단백질 및/또는 재조합 브라큐리 단백질을 인코딩하는 합성 및 신규한 핵산을 포함한다.
- [0068] 한 구체예에서, 브라큐리 단백질을 인코딩하는 핵산과 최소한 70% 동일성을 갖는 하나 또는 그 이상의 핵산 서열이 제시되고, 여기서 브라큐리 단백질은 다음과 같이 구성된 군에서 선택된다: 서열 번호: 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71. 추가 구체예에서, 브라큐리 단백질을 인코딩하는 핵산과 최소한 75%, 80%, 85%, 90%, 또는 95% 동일성을 갖는 하나 또는 그 이상의 핵산이 제시되고, 여기서 브라큐리 단백질은 다음과 같이 구성된 군에서 선택된다: 서열 번호: 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71.
- [0069] 본 발명의 다양한 다른 구체예에서, 다음과 같이 구성된 군에서 선택되는 하나 또는 그 이상의 브라큐리 단백질이 제시된다: 서열 번호: 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71. 추가 구체예에서, 다음과 같이 구성된 군에서 선택되는 펩티드와 최소한 70% 동일성을 갖는 하나 또는 그 이상의 브라큐리 단백질이 제시된다: 서열 번호: 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71. 또 다른 추가의 구체예에서, 서열 번호: 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71로 구성된 군에서 선택되는 펩티드와 최소한 75%, 80%, 85%, 90%, 또는 95% 동일성을 갖는 펩티드가 제시된다.
- [0070] 다른 구체예에서, 브라큐리 단백질을 인코딩하는 하나 또는 그 이상의 핵산 서열이 제시되고, 상기 핵산 서열은 다음과 같이 구성된 군에서 선택된다: 서열 번호: 16-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69.
- [0071] 다른 구체예에서, 서열 번호: 16-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69로 구성된 군에서 선택되는 핵산과 최소한 70% 동일성을 갖고, 그리고 서열 번호: 16-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69로 구성된 군에서 선택되는 핵산 서열과 본질적으로 동일한 발현 특징을 갖는 하나 또는 그 이상의 핵산 서열이 제시된다. 추가 구체예에서, 서열 번호: 16-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69로 구성된 군에서 선택되는 핵산과 최소한 75%, 80%, 85%, 90%, 또는 95% 동일성을 갖고, 그리고 서열 번호: 16-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69로 구성된 군에서 선택되는 핵산 서열과 본질적으로 동일한 발현 특징을 갖는 핵산 서열이 제시된다.
- [0072] 바람직한 구체예에서, 서열 번호: 16-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69로 구성된 군에서 선택되는 하나 또는 그 이상의 핵산 서열이 제시된다.
- [0073] **발현 카세트**
- [0074] 추가 구체예에 따라, 본 발명은 이들 프로모터 중에서 하나 또는 그 이상 및/또는 브라큐리 재조합 단백질 및/또는 핵산을 포함하는 본 발명에 따른 발현 카세트를 제시한다.
- [0075] 한 구체예에서, 서열 번호: 1-10으로 구성된 군에서 선택되는 핵산과 최소한 70% 동일성을 갖는 프로모터, 그리고 서열 번호: 11, 13, 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71로 구성된 군에서 선택되는 브라큐리 펩티드를 인코딩하는 핵산을 포함하는 발현 카세트가 제시된다. 다른 구체예에서, 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 프로모터 및/또는 브라큐리 펩티드를 인코딩하는 핵산을 포함하는 발현 카세트가 제시되고, 여기서 상기 핵산은 다음과 같이 구성된 군에서 선택된다: 서열 번호: 12, 14-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69.
- [0076] 바람직한 구체예에서, 서열 번호: 72-76으로 구성된 군에서 선택되는 핵산 서열에 최소한 70% 상동한 핵산 서열을 포함하는 발현 카세트가 제시된다. 추가 구체예에서, 서열 번호: 72-76으로 구성된 군에서 선택되는 핵산과 최소한 75%, 80%, 85%, 90%, 또는 95% 동일성을 갖는 핵산 서열을 포함하는 발현 카세트가 제시된다.
- [0077] 더욱 바람직한 구체예에서, 서열 번호: 72를 포함하는 발현 카세트가 제시된다.
- [0078] 더욱 바람직한 구체예에서, 서열 번호: 73을 포함하는 발현 카세트가 제시된다.

- [0079] 더욱 바람직한 구체예에서, 서열 번호: 74를 포함하는 발현 카세트가 제시된다.
- [0080] **제조합 폭스바이러스**
- [0081] 추가 구체예에 따라, 본 발명에 따른 핵산, 프로모터, 제조합 단백질 및/또는 발현 카세트는 벡터의 일부일 수 있다. 용어 "벡터"는 당업자에게 공지된 임의의 벡터를 지칭한다. 벡터는 플라스미드 벡터, 예를 들면, pBR322 또는 pUC 시리즈의 벡터일 수 있다. 더욱 바람직하게는, 벡터는 제조합 바이러스이다. 본 발명의 맥락에서, 용어 "바이러스" 또는 "제조합 바이러스"는 바이러스 유전체를 포함하는 감염성 바이러스를 지칭한다. 이 경우에 있어서, 본 발명의 핵산, 프로모터, 제조합 단백질 및/또는 발현 카세트는 개별 제조합 바이러스의 바이러스 유전체의 일부이다. 제조합 바이러스 유전체는 포장되고, 그리고 획득된 제조합 바이러스는 세포 및 세포주의 감염, 특히 인간을 비롯한 살아있는 동물의 감염에 이용될 수 있다. 본 발명에 따라서 이용될 수 있는 전형적인 제조합 바이러스는 아데노바이러스 벡터, 레트로바이러스 벡터, 또는 아데노 연관된 바이러스 2 (AAV2)의 기초에서 벡터이다. 폭스바이러스 벡터가 가장 바람직하다.
- [0082] 여러 구체예에서, 본 발명에 따른 핵산, 프로모터, 폴리펩티드 및 발현 카세트는 바람직하게는, 폭스바이러스 프로모터로서 활성적이거나, 또는 폭스바이러스 감염된 세포에서 프로모터로서 활성적이다. 폭스바이러스는 바람직하게는 조류폭스바이러스 또는 오소폭스바이러스이다. 더욱 바람직하게는, 폭스바이러스는 조류폭스바이러스이다.
- [0083] 용어 "조류폭스바이러스"는 임의의 조류폭스바이러스, 예를 들면, 계두바이러스, 카나리아두바이러스, 운코폭스바이러스, 미나흐폭스바이러스, 구두바이러스, 프시타신폭스바이러스, 큐알리폭스바이러스, 피아콕폭스바이러스, 켄컨폭스바이러스, 스페로우폭스바이러스, 스탈링폭스바이러스 및 터키폭스바이러스를 지칭한다. 바람직한 조류폭스바이러스는 카나리아두바이러스 및 계두바이러스이다.
- [0084] 카나리아두바이러스의 실례는 군주 렌슬러 (Rentschler)이다. ALVAC로 명명된 플라크 정제된 카나리아 두창 군주 (U.S. 특허 번호 5,766,598)는 부다페스트 조약의 규정 하에, 수탁 번호 VR-2547로 American Type Culture Collection (ATCC)에 기탁되었다. 다른 카나리아 두창 군주는 Institute Merieux, Inc.로부터 가용한, LF2 CEP 524 24 10 75로서 지정된 상업적인 카나리아 두창 백신 군주이다.
- [0085] 계두바이러스의 실례는 군주 FP-1, FP-5, TROVAC (U.S. 특허 번호 5,766,598), 그리고 POXVAC-TC (U.S. 특허 7,410,644)이다. FP-1는 1일령 닭에서 백신으로서 이용되도록 변형된 Duvette 군주이다. 상기 군주는 O DCEP 25/CEP67/2309 October 1980으로서 지정된 상업적인 계두바이러스 백신 군주이고 Institute Merieux, Inc로부터 가용하다. FP-5는 American Scientific Laboratories (Division of Schering Corp.) Madison, Wis., 미국 수의학적 면허 번호 165, 일련 번호 30321로부터 가용한 닭 배아 기원의 상업적인 계두바이러스 백신 군주이다.
- [0086] 일정한 구체예에서, 제조합 FPV는 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 핵산 서열을 포함한다. 다양한 추가 구체예에서, 제조합 FPV는 서열 번호: 72-76에서 선택되는 핵산 서열을 포함하는 발현 카세트를 포함한다. 더욱 바람직한 구체예에서, 제조합 FPV는 서열 번호: 72, 73 및 74에서 선택되는 핵산 서열을 포함하는 발현 카세트를 포함한다.
- [0087] 다른 구체예에서, 제조합 FPV는 다음과 같이 구성된 군에서 선택되는 브라큐리 항원을 인코딩하는 핵산을 포함한다: 서열 번호: 11, 13, 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71. 추가 구체예에서, 제조합 FPV는 다음과 같이 구성된 군에서 선택되는 브라큐리 펩티드를 포함한다: 서열 번호: 11, 13, 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71. 추가 구체예에서, 제조합 FPV는 다음과 같이 구성된 군에서 선택되는 핵산을 포함한다: 서열 번호: 12, 14-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69. 더욱 바람직한 구체예에서, 제조합 FPV는 서열 번호: 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71로 구성된 군에서 선택되는 브라큐리 펩티드 및/또는 서열 번호: 16-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69로 구성된 군에서 선택되는 핵산을 포함한다.
- [0088] 본 발명의 다양한 다른 구체예에서, 제조합 폭스바이러스는 오소폭스바이러스, 예를 들면, 하지만 제한 없이, 우두바이러스, 변형된 우두바이러스 양카라 (MVA), 또는 MVA-BN이다.
- [0089] 우두바이러스 군주의 실례는 군주 Temple of Heaven, Copenhagen, Paris, Budapest, Dairen, Gam, MRIVP, Per, Tashkent, TBK, Tom, Bern, Patwadangar, BIEM, B-15, Lister, EM-63, New York City Board of Health,

Elstree, Ikeda 및 WR이다. 바람직한 우두바이러스 (VV) 균주는 Wyeth (DRYVAX) 균주 (U.S. 특허 7,410,644)이다. 다른 바람직한 VV 균주는 MVA (Sutter, G. et al. [1994], Vaccine 12: 1032-40)이다. 다른 바람직한 VV 균주는 MVA-BN이다.

- [0090] 본 발명의 실시에서 유용하고 부다페스트 조약의 규정에 따라 기탁된 MVA 바이러스 균주의 실례는 1994년 1월 27일자 수탁 번호 ECACC 94012707로 European Collection of Animal Cell Cultures (ECACC), Vaccine Research and Production Laboratory, Public Health Laboratory Service, Centre for Applied Microbiology and Research, Porton Down, Salisbury, Wiltshire SP4 0JG, United Kingdom에 기탁된 균주 MVA 572, 그리고 2000년 12월 7일자 ECACC 00120707 하에 기탁된 MVA 575이다. 2000년 8월 30일자에 번호 V00083008 하에 European Collection of Cell Cultures (ECACC)에 기탁된 MVA-BN, 그리고 이의 유도체가 추가 예시적인 균주이다.
- [0091] 비록 MVA-BN이 이의 더욱 높은 안전성 (더욱 적은 복제 적격성)으로 인해 선호되긴 하지만, 모든 MVA가 본 발명에 적합하다. 본 발명의 한 구체예에 따라, MVA 균주는 MVA-BN 및 이의 유도체이다. MVA-BN 및 이의 유도체의 정의는 본원에 참조로서 편입되는 PCT/EP01/13628에서 제공된다.
- [0092] 한 구체예에서, 본 발명은 암 요법을 위한 재조합 오소포스바이러스, 바람직하게는 우두바이러스 (VV), Wyeth 균주, ACAM 1000, ACAM 2000, MVA, 또는 MVA-BN을 포괄한다. 재조합 오소포스바이러스는 오소포스바이러스 내로 이중성 서열의 삽입에 의해 산출된다.
- [0093] 일정한 구체예에서, MVA는 2000년 8월 30일자에 번호 V00083008 하에 European Collection of Cell Cultures (ECACC)에 기탁되고, 그리고 국제 PCT 공보 WO2002042480 (예로서, U.S. 특허 번호 6,761,893 및 6,913,752를 또한 참조한다)에서 설명된 MVA-BN이고, 이들 문헌은 본원에 참조로서 편입된다.
- [0094] 일정한 구체예에서, 재조합 MVA는 MVA-BN의 유도체이다. 이런 "유도체"는 기탁된 균주 (ECACC 번호 V00083008)와 본질적으로 동일한 복제 특징을 전사하지만, 이의 유전체의 하나 또는 그 이상의 부분에서 차이를 전사하는 바이러스를 포함한다. 기탁된 바이러스와 동일한 "복제 특징"을 갖는 바이러스는 CEF 세포 및 세포주, HeLa, HaCat 및 143B에서 기탁된 균주와 유사한 증폭 비율로 복제하고; 그리고 생체내에서, 예를 들면, AGR129 유전자 도입 생쥐 모형에서 결정될 때 유사한 복제 특징을 보여주는 바이러스이다.
- [0095] 일정한 구체예에서, 재조합 오소포스바이러스는 서열 번호: 1-10 및 77로 구성된 군에서 선택되는 핵산을 포함한다. 다양한 추가 구체예에서, 재조합 오소포스바이러스는 서열 번호: 72-76에서 선택되는 핵산 서열을 포함하는 발현 카세트를 포함한다. 더욱 바람직한 구체예에서, 재조합 오소포스바이러스는 서열 번호: 72, 73 및 74에서 선택되는 핵산 서열을 포함하는 발현 카세트를 포함한다.
- [0096] 다른 구체예에서, 재조합 오소포스바이러스는 다음과 같이 구성된 군에서 선택되는 브라큐리 항원을 인코딩하는 핵산을 포함한다: 서열 번호: 11, 13, 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71. 추가 구체예에서, 재조합 오소포스바이러스는 다음과 같이 구성된 군에서 선택되는 핵산을 포함한다: 서열 번호: 12, 14-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69. 더욱 바람직한 구체예에서, 재조합 오소포스바이러스는 서열 번호: 18-19, 22-23, 26-27, 30-31, 34-35, 38-39, 42-43, 46-47, 50-51, 54-55, 58-59, 62-63, 66-67, 70 및 71로 구성된 군에서 선택되는 브라큐리 펩티드 및/또는 서열 번호: 16-17, 20-21, 24-25, 28-29, 32-33, 36-37, 40-41, 44-45, 48-49, 52-53, 56-57, 60-61, 64-65, 68 및 69로 구성된 군에서 선택되는 핵산을 포함한다.
- [0097] 본 발명에 따른 발현 카세트 또는 프로모터가 바이러스 유전체 내로, 특히 포스바이러스의 유전체 내로, 가장 바람직하게는 오소포스바이러스 및/또는 FPV의 유전체 내로 삽입될 수 있는 방법은 당업자에게 알려져 있다. 가령, 본 발명에 따른 발현 카세트 또는 프로모터 또는 이의 유도체는 상동성 재조합에 의해 포스바이러스의 유전체 내로 삽입될 수 있다. 이를 위해, 핵산은 허용적 세포주, 예를 들면, CEF 또는 BHK 세포 내로 형질감염되고, 여기서 상기 핵산은 본 발명에 따른 발현 카세트 또는 프로모터 또는 이의 유도체가 삽입되는 포스바이러스 유전체의 영역에 상동하는 뉴클레오티드 스트레치에 의해 측면에서 접하는 본 발명에 따른 발현 카세트 또는 프로모터 또는 이의 유도체를 포함한다. 세포는 포스바이러스에 의해 감염되고, 그리고 감염된 세포에서 상동성 재조합은 핵산 및 바이러스 유전체 사이에 일어난다. 대안으로, 세포를 포스바이러스로 먼저 감염시키고, 그리고 이후, 핵산을 감염된 세포 내로 형질감염시키는 것도 가능하다. 이번에도, 재조합이 세포 내에서 발생한다. 재조합 포스바이러스는 이후, 선행 기술에서 공지된 방법에 의해 선별된다. 재조합 포스바이러스의 작제는 이러한 특정 방법에 한정되지 않는다. 그 대신에, 당업자에게 공지된 임의의 적합한 방법이 이러한 목적으로 이용될 수

있다.

[0098] 본 발명에 따른 발현 카세트 또는 프로모터는 바이러스 또는 바이러스 벡터의 임의의 적합한 부분 내로, 특히 바이러스 유전체 내로 도입될 수 있다. 오소폭스바이러스 및 조류폭스바이러스의 경우에, 삽입은 바이러스 유전체의 비필수적인 부분 내로 또는 바이러스 유전체의 유전자간 영역 내로 만들어질 수 있다. 용어 "유전자간 영역"은 바람직하게는, 코딩 서열을 포함하지 않는 2개의 인접한 유전자 사이에 위치된 바이러스 유전체의 부분을 지칭한다. 바이러스가 오소폭스바이러스 및 조류폭스바이러스이면, 삽입은 또한, 바이러스 유전체의 결실 부위 내로 만들어질 수 있다. 용어 "결실 부위"는 자연발생 오소폭스바이러스 또는 조류폭스바이러스의 유전체에 대하여 결실되는 바이러스 유전체의 부분을 지칭한다. 하지만, 삽입 부위는 오소폭스바이러스 및 조류폭스바이러스 유전체 내에 이들 바람직한 삽입 부위에 한정되지 않는데, 그 이유는 최소한 하나의 세포 배양 시스템, 예를 들면, 닭 배아 섬유모세포 (CEF 세포)에서 증폭되고 증식될 수 있는 재조합체를 획득하는 것이 가능하기만 하면, 발현 카세트가 바이러스 유전체 내에 어디든지 삽입될 수 있다는 것이 본 발명의 범위 안에 있기 때문이다.

[0099] 본 발명에 따른 프로모터는 벡터, 예를 들면, 오소폭스바이러스 및/또는 조류폭스바이러스의 유전체의 이미 일부인 유전자를 발현하는데 이용될 수 있다. 이런 유전자는 바이러스 유전체의 자연적으로 일부인 유전자 또는 벡터 내로 이미 삽입된 외래 유전자일 수 있다. 이들 사례에서, 본 발명에 따른 프로모터는 벡터 내에 유전자의 상류에 삽입되고, 이의 발현은 상기 프로모터에 의해 제어된다.

# [0100] **백신 및/또는 조성물**

[0101] 추가 구체예에 따라, 본 발명은 백신 또는 약제로서 본 발명에 따른 벡터에 관계한다. 더욱 일반적인 용어에서 본 발명은 본 발명에 따른 발현 카세트, DNA 또는 벡터를 포함하는 백신 또는 제약학적 조성물에 관계한다. 백신 또는 제약학적 조성물이 동물 또는 인체에 투여될 수 있는 방법은 당업자에게 알려져 있다. DNA 및 재조합 플라스미드 벡터의 경우에, DNA 및 벡터는 주사에 의해 간단히 투여될 수 있다. 백신 또는 조성물이 재조합 바이러스, 예를 들면, 오소폭스바이러스 또는 조류폭스바이러스, 특히 재조합 MVA 또는 재조합 FPV이면, 이것은 또한, 당업자의 지식에 따라, 예를 들면, 정맥내, 근육내, 비내, 피내 또는 피하 투여에 의해 동물 또는 인체에 투여될 수 있다. 투여되는 바이러스의 양에 관한 추가 상세는 아래에 제공된다.

[0102] 제약학적 조성물 또는 백신은 일반적으로, 본 발명에 따른 프로모터, 발현 카세트 또는 벡터에 더하여, 하나 또는 그 이상의 제약학적 허용되는 및/또는 승인된 담체, 첨가제, 항생제, 보존제, 어유번트, 희석제 및/또는 안정제를 포함할 수 있다. 이런 보조 물질은 물, 식염수, 글리세롤, 에탄올, 적십제 또는 유화제, pH 완충 물질, 또는 기타 유사한 것일 수 있다. 적합한 담체는 전형적으로 크고, 천천히 물질대사되는 분자, 예를 들면, 단백질, 다당류, 폴리유산, 폴리글리콜산, 중합성 아미노산, 아미노산 공중합체, 지질 응집체, 또는 기타 유사한 것이다.

[0103] 제약학적 조성물 또는 백신의 제조를 위해, 본 발명에 따른 DNA, 발현 카세트 또는 벡터, 특히 재조합 오소폭스바이러스 또는 조류폭스바이러스, 예를 들면, 재조합 MVA 또는 재조합 FPV는 생리학적으로 허용되는 형태로 전환된다. MVA 및 FPV의 경우에, 이것은 천연두에 대항하여 예방접종에 이용된 폭스바이러스 백신의 제조에서 경험에 근거하여 행해질 수 있다 (Stickl, H. et al. [1974] Dtsch. med. Wschr. 99, 2386-2392)에 의해 설명된 바와 같이). 가령, 정제된 바이러스는 약 10 mM Tris, 140 mM NaCl pH 7.4에서 조제된  $5 \times 10^8$  TCID<sub>50</sub>/ml의 역가로 -80℃에서 보관된다. 예방 주사의 제조를 위해, 예로서  $10^1$ - $10^9$  입자의 본 발명에 따른 재조합 바이러스가 앰플, 바람직하게는 유리 앰플 내에 2% 펩톤 및 1% 인간 알부민의 존재에서 인산염 완충된 식염수 (PBS)에서 동결건조된다. 대안으로, 예방 주사는 제제에서 바이러스의 단계별, 동결건조에 의해 생산될 수 있다. 이러한 제제는 추가 첨가제, 예를 들면, 만니톨, 텍스트란, 당, 글리신, 락토오스 또는 폴리비닐피롤리돈, 또는 생체내 투여에 적합한 다른 첨가제, 예를 들면, 항산화제 또는 불활성 가스, 안정제 또는 재조합 단백질 (가령, 인간 혈청 알부민)을 내포할 수 있다. 동결건조에 적합한 전형적인 바이러스 내포 제제는 10 mM Tris-완충액, 140 mM NaCl, 18.9 g/l 텍스트란 (MW 36000-40000), 45 g/l 수크로오스, 0.108 g/l L-글루탐산 모노 칼륨 염 이수화물 pH 7.4를 포함한다. 유리 앰플은 이후, 밀봉되고 4℃ 및 실온 사이에서 수개월 동안 보관될 수 있다. 하지만, 더 이상 필요하지 않으면, 앰플은 바람직하게는 -20℃ 미만의 온도에서 보관된다.

[0104] 예방접종 또는 요법을 위해, 동결건조물 또는 동결건조된 산물은 0.1 내지 0.5 ml의 수성 용액, 바람직하게는 물, 생리 식염수 또는 Tris 완충액에서 용해되고, 그리고 전신적으로 또는 국부적으로, 다시 말하면, 비경구, 근육내 또는 당업자에게 알려진 임의의 다른 투여 경로에 의해 투여될 수 있다. 투여 방식, 투여 용량 및 투여

횃수는 당업자에 의해 공지된 방식으로 최적화될 수 있다.

[0105] 하나 또는 그 이상의 구체예에서, 본 발명은 치료적 목적을 위해 코딩 서열을 표적 세포, 예를 들면, 인간 세포 내로 도입하기 위한 방법에 더욱 관계하고, 상기 방법은 본 발명에 따른 핵산, 프로모터, 재조합 단백질 및/또는 발현 카세트의 표적 세포 내로의 도입을 포함한다. 예시적인 인간 표적 세포는 항원 제시 세포 (APCs), 예를 들면, 수지상 세포, 대식세포 및 다른 비 APC, 예를 들면, 섬유모세포, 종양 세포 등을 포함할 수 있다.

[0106] 본 발명은 본 발명에 따른 재조합 바이러스로 숙주 세포의 감염, 그 이후에 적합한 조건 하에 감염된 숙주 세포의 배양, 그리고 그 이후에, 상기 숙주 세포에 의해 생산된 펩티드 및/또는 단백질 및/또는 바이러스의 단리 및/또는 농축을 포함하는 펩티드, 단백질 및/또는 바이러스를 생산하기 위한 방법에 더욱 관계한다. 본 발명에 따른 바이러스를 생산하는, 다시 말하면, 증폭하는 것이 의도되면, 세포는 바이러스가 복제할 수 있는 세포이어야 한다. 폭스바이러스, 특히 MVA의 경우에, 적합한 세포는 CEF (닭 배아 섬유모세포) 또는 BHK (아기 햄스터 신장) 세포이다. 조류폭스바이러스, 예를 들면, 계두바이러스의 경우에, 적합한 세포는 CEF 또는 CED (닭 배아 진피) 세포를 포함한다. 본 발명에 따른 재조합 바이러스에 의해 인코딩된 펩티드/단백질을 생산하는 것이 의도되면, 세포는 재조합 바이러스 벡터에 의해 감염될 수 있고, 그리고 바이러스 인코딩된 단백질/펩티드의 발현을 허용하는 임의의 세포일 수 있다.

[0107] 본 발명은 본 발명에 따른 발현 카세트, 핵산, 프로모터, 재조합 단백질 및/또는 발현 카세트 DNA로 세포의 형질감염, 그 이후에 폭스바이러스로 상기 세포의 감염을 포함하는 펩티드, 단백질 및/또는 바이러스를 생산하기 위한 방법에 더욱 관계한다. 감염된 숙주 세포는 적합한 조건 하에 배양된다. 추가 단계는 상기 숙주 세포에 의해 생산된 펩티드 및/또는 단백질 및/또는 바이러스의 단리 및/또는 농축을 포함한다. 이들 세포를 폭스바이러스로 감염시키는 단계는 세포의 형질감염의 단계 전후에 이루어질 수 있다.

[0108] 본 발명은 본 발명에 따른 핵산, 프로모터, 재조합 단백질 및/또는 발현 카세트를 포함하는 세포에 더욱 관계한다. 특히, 본 발명은 본 발명에 따른 재조합 바이러스로 감염된 세포에 관계한다.

#### [0109] 실시예

[0110] 다음 실시예는 본 발명을 더욱 예증한다. 제공된 실시예는 기술의 적용가능성을 제한하는 것으로 결코 해석될 수 없는 것으로 당업자에 의해 이해될 것이다.

#### [0111] 실시예 1: 재조합 계두바이러스로 감염된 인간 DCs에서 브라큐리의 발현

[0112] 브라큐리 단백질의 발현을 확인하기 위해, 인간 수지상 세포 (DC)는 2.5의 감염 다중도 (MOI)에서 양성 대조 바이러스, 브라큐리 및 TRICOM을 포함하는 재조합 MVA로 감염되었다. 인간 DCs는 또한, 본 발명에 따른 브라큐리 발현 카세트 및 TRICOM을 포함하는 음성 대조 비재조합 계두 (FPV-WT) 또는 재조합 계두바이러스 균주로 감염되었다. FPV-WT, FPV-mBN343A, FPV-mBN344A 및 FPV-mBN345A를 비롯한 각 FPV 균주 (표 1에서 더욱 상세하게 열거됨)가 20의 MOI에서 DCs를 감염시키는데 이용되었다. FPV브라큐리 발현은 토끼 단일클론 항-브라큐리 항체로 수행된 웨스턴 블롯 분석을 통해 검출되었다. 하우스키핑 단백질 글리세르알데히드 3-인산염 탈수소효소 (GAPDH) 또한, 부하 대조로서 웨스턴 블롯 분석을 통해 검출되었다.

[0113] 표 1. 브라큐리 발현 카세트 및 TRICOM을 포함하는 재조합 계두바이러스 균주

표 1

| 재조합 FPV 균주  | 발현 카세트    |
|-------------|-----------|
| FPV-mBN343A | 서열 번호: 74 |
| FPV-mBN344A | 서열 번호: 75 |
| FPV-mBN345A | 서열 번호: 72 |
| FPV-mBN354A | 서열 번호: 76 |
| FPV-mBN355A | 서열 번호: 73 |

[0115] 결과는 도면 1에서 도시된다. 브라큐리의 발현은 FPV-mBN343A 및 FPV-mBN345 재조합 FPVs로 검출되었지만, FPV-mBN344A 재조합 FPV를 이용하여 검출되지 않았다. 브라큐리 발현은 또한, MVA-브라큐리-TRICOM으로 검출되었다. 표본의 유사한 부하는 GAPDH 발현에 의해 증명되었다.

#### [0116] 실시예 2: 재조합 계두바이러스로 감염된 인간 DCs에서 브라큐리의 발현

- [0117] 브라큐리 단백질의 발현-수준은 또한, 상이한 프로모터로 브라큐리를 발현하는 추가 재조합 FPV 균주 사이에 비교되었다. 인간 수지상 세포 (DC)는 5의 MOI에서 브라큐리 및 TRICOM을 포함하는 양성 대조 재조합 MVA, 그리고 음성 대조 비재조합 균주 MVA-WT로 감염되었다. 인간 DCs는 또한, 본 발명에 따른 브라큐리 발현 카세트 및 TRICOM을 포함하는 재조합 FPVs (가령, FPV-mBN343A, FPV-mBN344A, FPV-mBN345A, FPV-mBN354A, FPV-mBN355A, 표 1을 참조한다)로 감염되었다. 인간 DCs는 우두바이러스 (VV)-40k 프로모터 또는 PrS 프로모터를 갖는 브라큐리 발현 카세트를 포함하는 재조합 계두바이러스 (FPV)로 부가적으로 감염되었다. 비재조합 FPV-WT 균주는 음성 대조로서 역할을 하였다. 이들 FPVs 모두 40의 MOI에서 이용되었다. 브라큐리 발현은 토끼 단일클론 항-브라큐리 항체로 수행된 웨스턴 블롯 분석을 통해 검출되었다. GAPDH는 또한, 부하 대조로서 검출되었다.
- [0118] 결과는 도면 2에서 도시된다. 브라큐리의 발현은 FPV-mBN343A 및 FPV-mBN345A 재조합 FPVs로 검출되었지만, FPV-mBN344A 및 FPV-mBN354A 재조합 FPV를 이용하여 검출되지 않았다. 더욱 구체적으로, 브라큐리를 주동하는 VV-40k 또는 PrS 프로모터를 갖는 재조합 FPVs (다시 말하면, FVP-mBN281A, FVP-mBN249B)의 경우에 브라큐리의 발현이 더욱 낮은 수준에서 검출되었다. 음성 대조 (감염되지 않은 DCs, MVA-WT, 또는 FPV-WT)에서는 브라큐리 발현이 검출되지 않았다. 표본의 유사한 부하는 GAPDH 발현에 의해 증명되었다.
- [0119] 실시예 2에서 도시된 웨스턴 블롯으로부터 브라큐리 발현 수준은 세포 사이에서 동등한 수준에서 발현될 것으로 예상되는 하우스키핑 유전자 GAPDH의 발현에 상대적으로 정규화되었다. 각 브라큐리 및 GAPDH 띠의 강도가 측정되었고, 그리고 동일한 표본 내에서 브라큐리 및 GAPDH 띠의 강도 사이에 비율이 계산되었다.
- [0120] 결과는 도면 3에서 도시된다. FPV 구조체 중에서, GAPDH에 비하여 브라큐리의 가장 높은 발현은 FPV-mBN355A로 감염된 DCs에서 검출되었다. 중간적 상대적 브라큐리 발현은 FPV-mBN343A 또는 FPV-mBN345A로 감염된 DCs에서 검출되었다. 가장 낮은 상대적 브라큐리 발현은 브라큐리를 주동하는 VV-40k 또는 PrS 프로모터를 갖는 재조합 FPVs (다시 말하면, FVP-mBN281A, FVP-mBN249B)로부터 검출되었다. 브라큐리의 가장 높은 상대적 발현은 양성 대조 바이러스 MVA-브라큐리-TRICOM으로 감염으로부터 관찰되었다; 음성 대조 표본에서는 브라큐리 발현이 검출되지 않았다. 이런 이유로, 재조합 FPV 균주 중에서, 브라큐리의 우수한 발현은 FPV-mBN355, FPV-mBN345, 또는 FPV-mBN344 프로모터로부터 브라큐리 발현을 주동하는 벡터에 의해 유도되었다.
- [0121] 실시예 3: 재조합 계두바이러스로 감염된 CMMT 세포에서 브라큐리 및 TRICOM의 발현
- [0122] 브라큐리 및 TRICOM 단백질의 발현 역시 각 단백질에 특정한 형광으로 표지화된 항체를 이용한 유세포분석법에 의해 사정되었다. CMMT 세포 (히말라야 원숭이 유방 종양 세포주)는 브라큐리 및 TRICOM을 포함하는 양성 대조 재조합 MVA로, 또는 본 발명에 따른 브라큐리 발현 카세트 및 TRICOM을 포함하는 재조합 FPVs (가령, FPV-mBN343A, FPV-mBN345A)로, 또는 브라큐리를 주동하는 PrS 프로모터를 갖는 재조합 FPV (FVP-mBN249B)로 감염되었다. 세포는 감염되지 않은 세포 및 감염된 세포의 혼합물이 분석되도록, 1 미만의 MOI에서 감염되었다. 감염되지 않은 CMMT 세포는 음성 대조로서 역할을 하였다.
- [0123] FACS 표본은 BD LSRII 또는 Fortessa에서 획득되었고 BD FACSDIVA 소프트웨어 (BD Bioscience, San Jose, CA) 또는 FlowJo (TreeStar Inc., Ashland, OR)를 이용하여 분석되었다.
- [0124] 결과는 도면 4에서 도시된다. 브라큐리 및 3가지 TRICOM 단백질 (CD80, CD54 및 CD58)에 대해 검출된 신호의 히스토그램은 플롯팅되었고, 그리고 양성 신호가 검출된 게이트가 뽑혔다 (흑색 라인). FPV 구조체 중에서, 브라큐리의 가장 높은 발현 (x 축을 따라서 신호에서 가장 큰 이동에 의해 도시됨)은 FPV-mBN343A 또는 FPV-mBN345A로 감염된 CMMT 세포에서 검출되었다. 가장 낮은 브라큐리 발현은 브라큐리를 주동하는 PrS 프로모터를 갖는 재조합 FPV (다시 말하면, FVP-mBN281A, FVP-mBN249B)로부터 검출되었다. 이들 FPV 구조체 사이에서 TRICOM 단백질의 유사한 발현 수준이 검출되었다. 브라큐리 및 TRICOM 단백질의 발현은 또한, 양성 대조 바이러스 MVA-브라큐리-TRICOM으로 감염으로부터 관찰되었다.
- [0125] 감염된 세포에서 브라큐리의 발현 수준을 정량하기 위해, 도면 4에서 게이팅된 브라큐리-양성 세포에 대한 중앙형광 강도 (MFI)가 계산되었다.
- [0126] 결과는 도면 5에서 도시된다. FPV 구조체 중에서, 가장 높은 브라큐리 MFI는 FPV-mBN343A 또는 FPV-mBN345A로 감염된 CMMT 세포에서 검출되었다. 가장 낮은 브라큐리 MFI는 브라큐리를 주동하는 PrS 프로모터를 갖는 재조합 FPV (FVP-mBN249B)로부터 검출되었다. 모든 시험된 구조체 중에서 가장 높은 브라큐리 MFI는 양성 대조 바이러스 MVA-브라큐리-TRICOM으로 감염된 세포에서 관찰되었다. 이런 이유로, 감염된 CMMT 세포에서 브라큐리의 중앙 발현 수준은 PrS 프로모터로부터 보다는 FPV-mBN345 또는 FPV-mBN344 프로모터로부터 브라큐리 발현을 주동하는 벡터로부터 더욱 높았다.

[0127] 실시예 4: FPV-mBN345B로부터 브라큐리의 발현

[0128] 재조합 FPV-mBN345A 균주를 초기에 산출하는데 이용된 약물 선별 카세트는 임상적 개발에 적합한 재조합 벡터를 산출하기 위해 제거되었다. 이것은 약물 선별 없이 상기 바이러스를 담 배아 섬유모세포 (CEF) 세포 상에서 계대배양하고, 개별 클론을 플라크 정제하고, 그리고 선별 카세트를 결여하는 클론을 PCR 및 DNA 염기서열결정에 의해 확인함으로써 달성되었다. 이것은 I3+15aa 프로모터에 의해 주동된 브라큐리 발현 카세트 및 TRICOM을 포함하지만, 약물 선별 카세트를 내포하지 않는 FPV-mBN345B의 산출을 유발하였다.

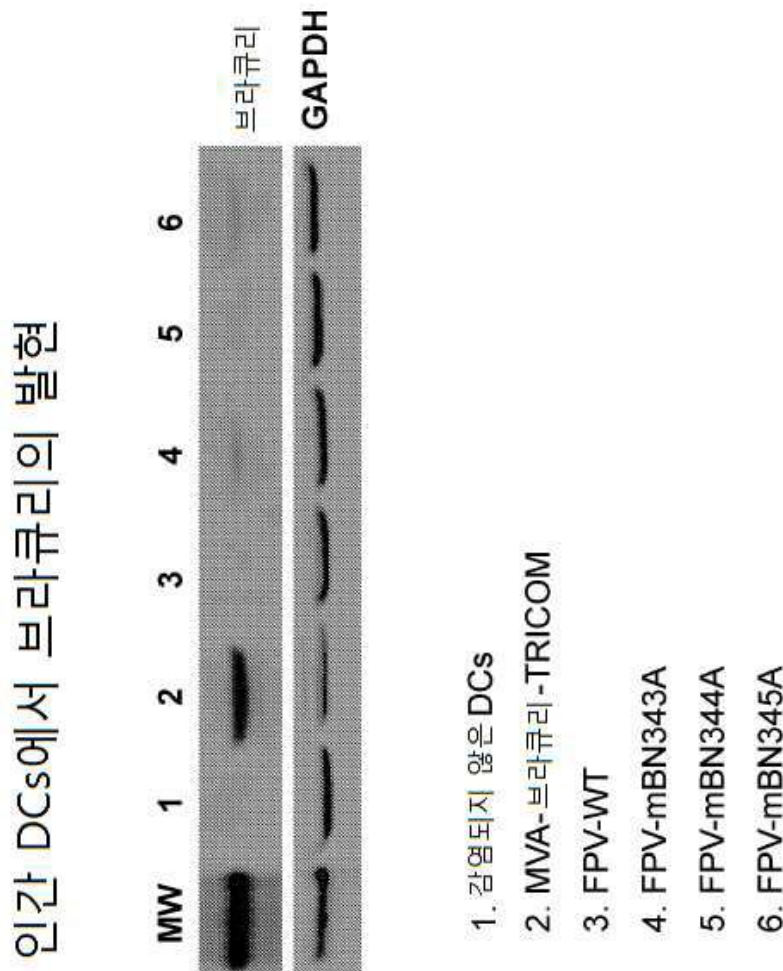
[0129] FPV-mBN345B로부터 브라큐리 및 TRICOM 단백질의 발현은 세포의 부분집합을 감염시키는 경우에 0.625, 또는 모든 세포를 감염시키는 경우에 40의 MOI에서 FPV-BN345B로 이들을 감염시킴으로써 CMMT 세포에서 확증되었다.

[0130] 결과는 도면 6에서 도시된다. 브라큐리 및 3가지 TRICOM 단백질 (CD80, CD54 및 CD58)에 대해 검출된 신호의 히스토그램은 0.625의 MOI에서 감염된 표본의 경우 적색선 및 40의 MOI에서 감염된 표본의 경우 청색선으로 플롯팅되었다. 감염되지 않은 세포는 음성 대조로서 역할을 하였다 (흑색 라인). 0.625의 MOI에서, 브라큐리 및 3가지 TRICOM 단백질의 발현은 감염되지 않은 세포로부터 더욱 높은 피크의 오른쪽으로 이동된 피크로서 검출되었다. 40의 MOI에서, 브라큐리 및 TRICOM의 발현은 모든 세포에서, 감염되지 않은 세포의 오른쪽으로 이동된 피크로서 검출되었다.

[0131] 본 발명의 다른 구체예는 본 명세서의 고려 및 본원에서 개시된 발명의 실시로부터 당업자에게 명백할 것이다. 본 명세서와 실시예는 단지 예시로서 고려되고, 본 발명의 진정한 범위와 사상은 아래 청구항에 의해 지시되는 것으로 의도된다.

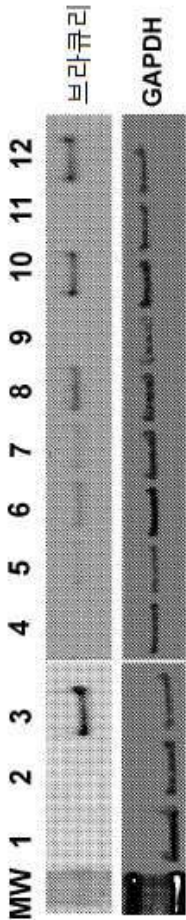
## 도면

### 도면1



도면2

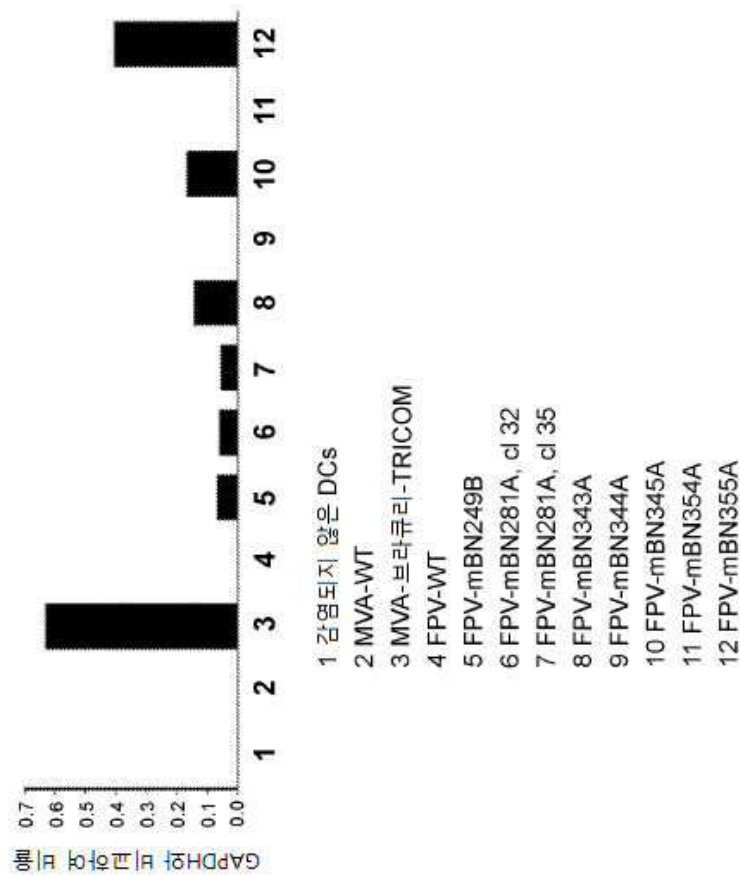
인간 DCs에서 브라큐리의 발현



- 1 감염되지 않은 DCs
- 2 MVA-WT
- 3 MVA-브라큐리-TRICOM
- 4 FPV-WT
- 5 FPV-mBN249B
- 6 FPV-mBN281A, cl 32
- 7 FPV-mBN281A, cl 35
- 8 FPV-mBN343A
- 9 FPV-mBN344A
- 10 FPV-mBN345A
- 11 FPV-mBN354A
- 12 FPV-mBN355A

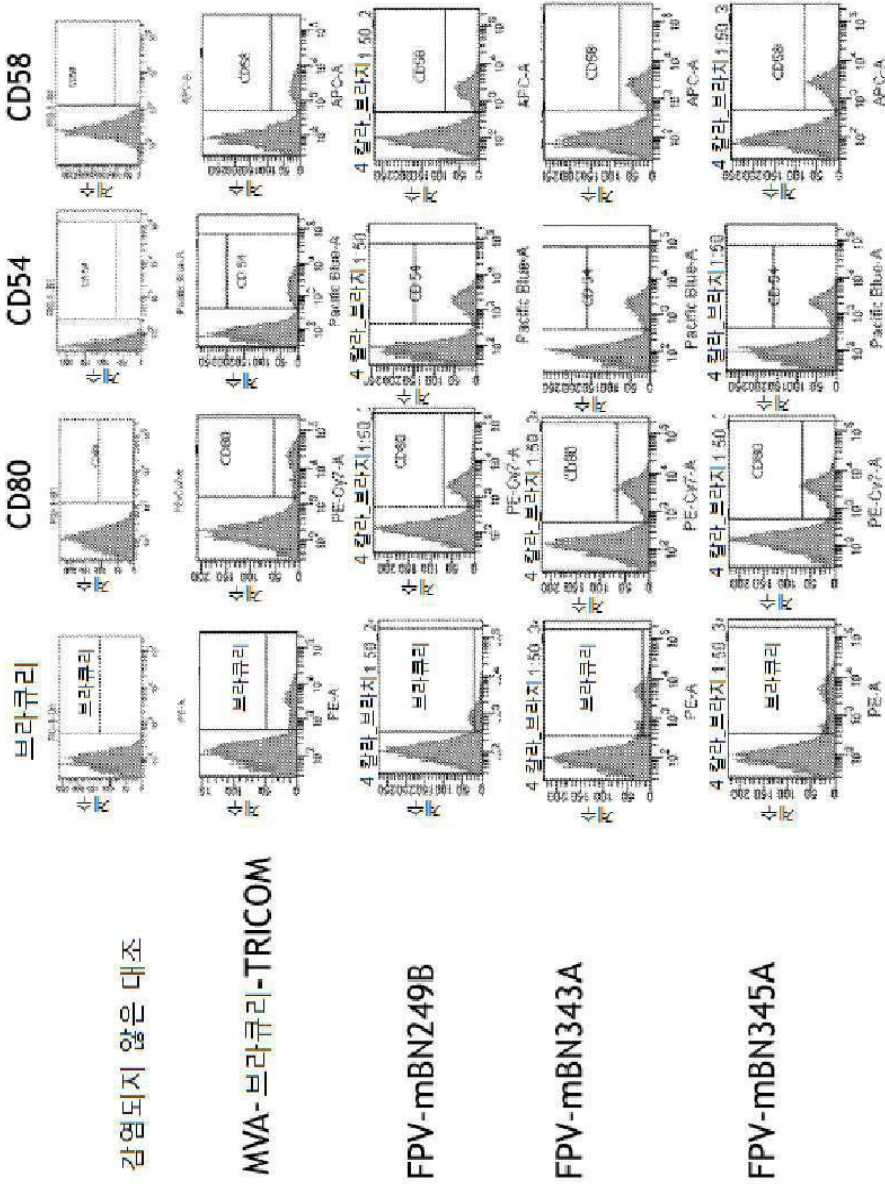
도면3

브라큐리의 상대적 발현



도면4

# CMMT 세포에서 브라큐리 및 TRICOM의 발현



감염되지 않은 대조

MVA-브라큐리-TRICOM

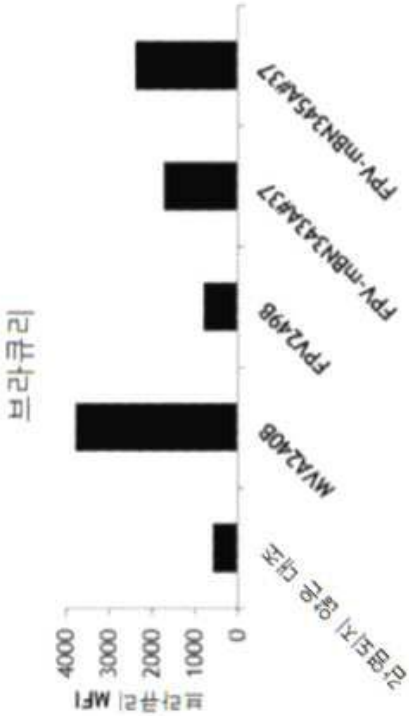
FPV-mBN249B

FPV-mBN343A

FPV-mBN345A

도면5

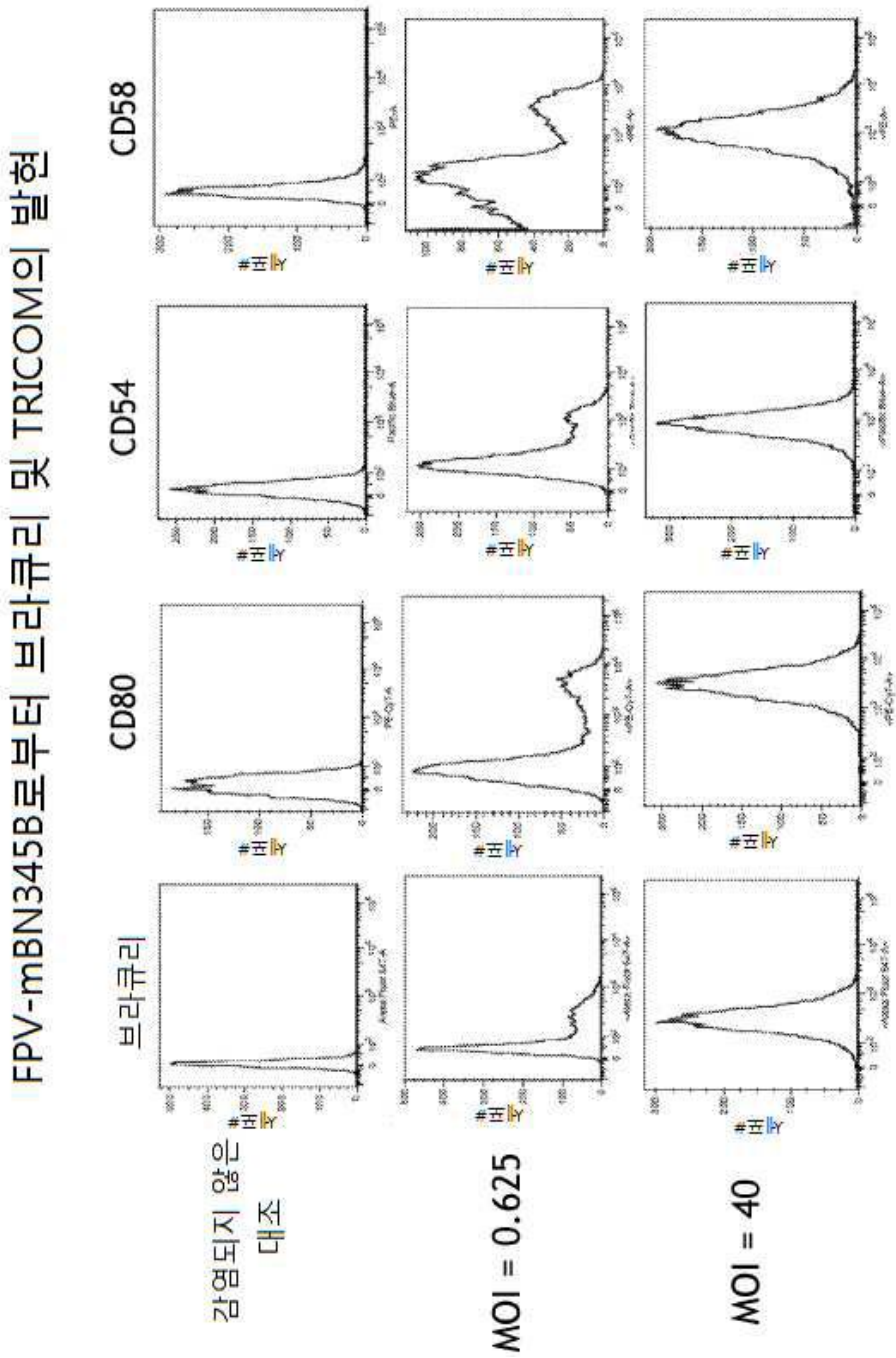
CMMT 세포에서 브라큐리 발현 수준



중앙 평균 강도 (MFI)

| 중앙 평균 강도 (MFI) |      |
|----------------|------|
| 감염되지 않은 대조     | 브라큐리 |
| MVA2408        | 584  |
| FPV2498        | 3783 |
| FPV-mBN343A#37 | 788  |
| FPV-mBN345A#37 | 1732 |
|                | 2391 |

도면6



서열목록

SEQUENCE LISTING

- <110> BAVARIAN NORDIC A/S
- <120> PROMOTERS FOR ENHANCING EXPRESSION IN POXVIRUSES
- <130> BNIT0010PCT
- <140> 62/199681
- <141> 2015-07-31

<160> 77

<170> PatentIn version 3.5

<210> 1

<211> 167

<212> DNA

<213> Artificial Sequence

<220><223> Promoter Sequence with nucleotides 5' of Gene FPV 088

<400> 1

tatccgtaca ggtttgtttc tgaaattcac tttgtaagat acataattaa caaattcagg 60

gggaaaaatc ttacaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa 120

ataatcagaa cctaattttt ttatcaaaaa attaaaatat aaataaaa 167

<210> 2

<211> 170

<212> DNA

<213> Artificial Sequence

<220><223> Promoter Sequence with nucleotides 5' of Gene FPV 088 + atc  
sequence

<400> 2

tatccgtaca ggtttgtttc tgaaattcac tttgtaagat acataattaa caaattcagg 60

gggaaaaatc ttacaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa 120

ataatcagaa cctaattttt ttatcaaaaa attaaaatat aaataaaaatc 170

<210> 3

<211> 183

<212> DNA

<213> Artificial Sequence

<220><223> Promoter Sequence with nucleotides 5' of Gene FPV 088 plus  
addition of 16 nucleotides of FPV 088

<400> 3

tatccgtaca ggtttgtttc tgaaattcac tttgtaagat acataattaa caaattcagg 60

gggaaaaatc ttacaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa 120

ataatcagaa cctaattttt ttatcaaaaa attaaaatat aaataaaatg aaaaataact 180

tgt 183

<210> 4

<211> 194

<212> DNA

<213> Artificial Sequence

<220><223> Promoter Sequence with region from Gene FPV 088 plus addition of  
27 nucleotides

<400> 4

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| tatccgtaca ggtttgtttc tgaaattcac tttgtaagat acataattaa caaattcagg | 60  |
| gggaaaaatc ttacaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa  | 120 |
| ataatcagaa cctaattttt ttatcaaaaa attaaaatat aaataaaatg aaaaataact | 180 |
| tgtatgaaga aaaa                                                   | 194 |

<210> 5

<211> 200

<212> DNA

<213> Artificial Sequence

<220><223> Promoter Sequence with region from Gene FPV 088 plus addition of  
33 nucleotides

<400> 5

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| tatccgtaca ggtttgtttc tgaaattcac tttgtaagat acataattaa caaattcagg | 60  |
| gggaaaaatc ttacaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa  | 120 |
| ataatcagaa cctaattttt ttatcaaaaa attaaaatat aaataaaatg aaaaataact | 180 |
| tgtatgaaga aaaaatgaac                                             | 200 |

<210> 6

<211> 212

<212> DNA

<213> Artificial Sequence

<220><223> Promoter Sequence with region from Gene FPV 088 plus addition of  
45 nucleotides

<

400> 6

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| tatccgtaca ggtttgtttc tgaaattcac tttgtaagat acataattaa caaattcagg | 60  |
| gggaaaaatc ttacaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa  | 120 |
| ataatcagaa cctaattttt ttatcaaaaa attaaaatat aaataaaatg aaaaataact | 180 |
| tgtatgaaga aaaaatgaac atgagtaaga aa                               | 212 |

<210> 7

<211> 232

<212> DNA

<213> Artificial Sequence

<220><223> Promoter Sequence with region from Gene FPV 088 plus addition of  
65 nucleotides

<400> 7

tatccgtaca ggtttgtttc tgaaattcac ttgtaagat acataattaa caaattcagg 60

gggaaaaatc tttaaaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa 120

ataatcagaa cctaattttt ttatcaaaaa attaaaatat aaataaaatg aaaaataact 180

tgtatgaaga aaaaatgaac atgagtaaga aacaagtaaa aactcaaagt aa 232

<210> 8

<211> 242

<212> DNA

<213> Artificial Sequence

<220><223> Promoter Sequence with region from Gene FPV 088 plus addition of  
75 nucleotides

<400> 8

tatccgtaca ggtttgtttc tgaaattcac ttgtaagat acataattaa caaattcagg 60

gggaaaaatc tttaaaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa 120

ataatcagaa cctaattttt ttatcaaaaa attaaaatat aaataaaatg aaaaataact 180

tgtatgaaga aaaaatgaac atgagtaaga aacaagtaaa aactcaaagt aaatgtaata 240

at 242

<210> 9

<211> 245

<212> DNA

<213> Artificial Sequence

<220><223> Promoter Sequence with region from Gene FPV 088 plus addition of  
78 nucleotides

<400> 9

tatccgtaca ggtttgtttc tgaaattcac ttgtaagat acataattaa caaattcagg 60

gggaaaaatc tttaaaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa 120

ataatcagaa cctaattttt ttatcaaaaa attaaaaatat aaataaaatg aaaaataact 180  
tgtatgaaga aaaaatgaac atgagtaaga aacaagtaaa aactcaaagt aatgtaata 240  
atatc 245

<210> 10

<211> 119

<212> DNA

<213> Artificial Sequence

<220><223> Promoter Sequence with nucleotides 5' of the Gene FPV L2R

<400> 10

ttaatgtata gaactaattt ataataaaca tagtaaatat gggtaacttc ttaatagcca 60  
taattaaaat tgaaaaaaaa atatcattat aaaacgtaaa cgaacaaaaa acattaatt 119

<210> 11

<211> 435

<212> PRT

<213> Artificial Sequence

<220><223> Brachyury protein Isoform 1 from GenBank Accession No. 015178.1

<400> 11

Met Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr Arg

1 5 10 15

Val Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly Ser

20 25 30

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

35 40 45

Ser Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile Val

50 55 60

Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val Ser

65 70 75 80

Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val Ala

85 90 95

Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro Gly

100 105 110

Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
| 115                                                             | 120 | 125 |     |
| Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser |     |     |     |
| 130                                                             | 135 | 140 |     |
| Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met Leu |     |     |     |
| 145                                                             | 150 | 155 | 160 |
| Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val Gly |     |     |     |
| 165                                                             | 170 | 175 |     |
|                                                                 |     |     |     |
| Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln Phe |     |     |     |
| 180                                                             | 185 | 190 |     |
| Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile |     |     |     |
| 195                                                             | 200 | 205 |     |
| Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser |     |     |     |
| 210                                                             | 215 | 220 |     |
| Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly |     |     |     |
| 225                                                             | 230 | 235 | 240 |
|                                                                 |     |     |     |
| Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Leu Cys Pro |     |     |     |
| 245                                                             | 250 | 255 |     |
| Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser |     |     |     |
| 260                                                             | 265 | 270 |     |
| Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser |     |     |     |
| 275                                                             | 280 | 285 |     |
| Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser |     |     |     |
| 290                                                             | 295 | 300 |     |
|                                                                 |     |     |     |
| Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp |     |     |     |
| 305                                                             | 310 | 315 | 320 |
| Ser Ser Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser His |     |     |     |
| 325                                                             | 330 | 335 |     |
| Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser |     |     |     |
| 340                                                             | 345 | 350 |     |
| Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser |     |     |     |
| 355                                                             | 360 | 365 |     |

Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr  
 370 375 380  
 Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu  
 385 390 395 400  
 Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp  
 405 410 415  
 Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro  
 420 425 430

Pro Ser Met  
 435

<210> 12

<211> 1308

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence for Brachyury protein Isoform 1 GenBank Accession  
 No. 015178.1

<400> 12

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| atgagctccc ctggcaccga gagcgcggga aagagcctgc agtaccgagt ggaccacctg | 60  |
| ctgagcgccg tggagaatga gctgcaggcg ggagcgaga agggcgaccc cacagagcgc  | 120 |
| gaactgcgcg tgggcctgga ggagagcgag ctgtggctgc gcttcaagga gctcaccaat | 180 |
| gagatgatcg tgaccaagaa cggcaggagg atgtttccgg tgctgaaggt gaacgtgtct | 240 |
|                                                                   |     |
| ggcctggacc ccaacgcat gtactccttc ctgctggact tcgtggcggc ggacaaccac  | 300 |
| cgctggaagt acgtgaacgg ggaatgggtg ccggggggca agccggagcc gcaggcgccc | 360 |
| agctgcgtct acatccacc cgactcgccc aacttcgggg cccactggat gaaggtctcc  | 420 |
| gtctccttca gcaaagtcaa gtcaccaac aagctcaacg gagggggcca gatcatgctg  | 480 |
| aactccttgc ataagtatga gcctcgaatc cacatagtga gatttggggg tccacagcgc | 540 |
| atgatcacca gccactgctt ccctgagacc cagttcatag cggtgactgc ttatcagaac | 600 |
| gaggagatca cagctcttaa aattaagtac aatccatttg caaaagcttt ccttgatgca | 660 |
|                                                                   |     |
| aaggaaagaa gtgatcaca agagatgatg gaggaacccg gagacagcca gcaacctggg  | 720 |
| tactcccaat gggggtggct tcttctgga accagcacc tgtgtccacc tgcaaatcct   | 780 |
| catcctcagt ttggaggctg cctctccctc cctccacgc acagctgtga caggtaccca  | 840 |
| accctgagga gccaccggtc ctcaccctac ccagccct atgctcatcg gaacaattct   | 900 |

ccaacctatt ctgacaactc acctgcatgt ttatccatgc tgcaatccca tgacaattgg 960  
tccagccttg gaatgcctgc ccattcccagc atgctccccg tgagccacaa tgccagccca 1020  
cctaccagct ccagtcagta cccagcctg tggctctgtga gcaacggcgc cgtcacccccg 1080

ggctcccagg cagcagccgt gtccaacggg ctggggggccc agttcttccg gggtccccc 1140  
gcgcactaca caccctcac ccattccgtc tcggcgccct cttcctcggg atccccactg 1200  
tacgaagggg cgcccgccgc cacagacatc gtggacagcc agtacgacgc cgcagcccaa 1260  
ggccgcctca tagcctcatg gacacctgtg tcgccacctt ccatgtga 1308

<210> 13

<211> 435

<212> PRT

<213> Artificial Sequence

<220><223> Brachyury protein Isoform 1 (L254V)

<400> 13

Met Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr Arg  
1 5 10 15

Val Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly Ser  
20 25 30

Glu Lys Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly Leu Glu Glu  
35 40 45

Ser Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile Val  
50 55 60

Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val Ser  
65 70 75 80

Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val Ala  
85 90 95

Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro Gly  
100 105 110

Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp  
115 120 125

Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser  
130 135 140

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

385                                      390                                      395                                      400

Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp

405                                      410                                      415

Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro

420                                      425                                      430

Pro Ser Met

435

<210> 14

<211> 1308

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Brachyury protein Isoform 1 with L254V

<400> 14

atgagctccc ctggcaccga gagcgcgga aagagcctgc agtaccgagt ggaccacctg 60

ctgagcgccg tggagaatga gctgcaggcg ggcagcgaga agggcgaccc cacagagcgc 120

gaactgcgcg tgggcctgga ggagagcgag ctgtggctgc gcttcaagga gctcaccaat 180

gagatgatcg tgaccaagaa cggcaggagg atgtttccgg tgctgaaggt gaacgtgtct 240

ggcctggacc ccaacgcat gtactccttc ctgctggact tcgtggcggc ggacaaccac 300

cgctggaagt acgtgaacgg ggaatgggtg ccggggggca agccggagcc gcaggcgccc 360

agctgcgtct acatccacc cgactcgccc aacttcgggg cccactggat gaaggctccc 420

gtctccttca gcaaagtcaa gctcaccaac aagctcaacg gagggggcca gatcatgctg 480

aactccttgc ataagtatga gcctcgaatc cacatagtga gagttggggg tccacagcgc 540

atgatcacca gccactgctt ccctgagacc cagttcatag cggtgactgc ttatcagaac 600

gaggagatca cagctcttaa aattaagtac aatccatttg caaaggcttt ccttgatgca 660

aaggaaagaa gtgatcaca agagatgatg gaggaaccgg gagacagcca gcaacctggg 720

tactcccaat gggggtggct tcttctgga accagcaccg tttgtccacc tgcaaatcct 780

catectcagt ttggagtgct cctctccctc cctccacgc acagctgtga caggtacca 840

acctgagga gccaccggtc ctacacctac cccagccct atgctcatcg gaacaattct 900

ccaacctatt ctgacaactc acctgcatgt ttatccatgc tgcaatcca tgacaattgg 960

tccagccttg gaatgcctgc ccatccagc atgctccccg tgagccacaa tgccagccca 1020

cctaccagct ccagtcagta cccagcctg tggctctgtg gcaacggcgc cgtcaccccg 1080

ggctcccagg cagcagccgt gtccaacggg ctgggggccc agttcttcg gggtccccc 1140  
gcgcactaca caccctcac ccatccgtc tcggcgccct ctctctcggg atccccactg 1200  
tacgaagggg cggccgcggc cacagacatc gtggacagcc agtacgacgc cgcagcccaa 1260  
ggccgcctca tagcctcatg gacacctgtg tcgccacctt ccatgtga 1308

<210> 15

<211> 1305

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Brachyury protein Isoform 1 with L254V  
without 5' atg sequence

<400> 15

agctcccctg gcaccgagag cgcgggaaaag agcctgcagt accgagtgga ccacctgctg 60  
agcgccctgg agaattgagct gcaggcgggc agcgagaagg gcgacccac agagcgcgaa 120  
ctgcgcgtgg gcctggagga gagcgagctg tgctgcgt tcaaggagct caccaatgag 180  
atgatcgtga ccaagaacgg caggaggatg ttccgggtgc tgaagggtgaa cgtgtctggc 240  
ctggacccca acgccatgta ctcttctctg ctggacttcg tggcggcgga caaccaccgc 300

tggaagtacg tgaacgggga atgggtgccg gggggcaagc cggagccgca ggcgcccagc 360  
tgcgtctaca tccacccga ctgcaccaac ttcggggccc actggatgaa ggctcccgtc 420  
tccttcagca aagtcaagct caccaacaag ctcaacggag ggggccagat catgctgaac 480  
tccttgcata agtatgagcc tcgaatccac atagttagag ttgggggtcc acagcgcag 540  
atcaccagcc actgcttccc tgagaccag ttcatacgcg tgactgctta tcagaacgag 600  
gagatcacag ctcttaaaat taagtacaat ccatttgcaa aggccttcct tgatgcaaag 660  
gaaagaagtg atcacaaga gatgatggag gaaccggag acagccagca acctgggtac 720

tcccaatggg ggtggttct tctggaacc agcacgttt gtccactgc aaatctcat 780  
cctcagtttg gaggtgccct ctccctccc tccacgcaca gctgtgacag gtaccaacc 840  
ctgaggagcc accggtctc accctaccc agccctatg ctcatcgga caattctca 900  
acctattctg acaactcacc tgcatttta tccatgtgc aatccatga caattgtcc 960  
agccttgga tgctgccc tccagcatg ctccccgtga gccacaatgc cagcccacct 1020  
accagctcca gtcagtacc cagcctgtgg tctgtgagca acggcgccgt cccccgggc 1080  
tccaggcag cagccgtgtc caacgggctg ggggccagc tcttcgggg ctccccgcg 1140

cactacacac cctcacca tccgtctcg gcgccctct cctcgggatc cccactgtac 1200

gaagggcg cgcggccac agacatcgtg gacagccagt acgacgccg agcccaaggc 1260

cgccatcatag cctcatggac acctgtgtcg ccaccttcca tgtga 1305

<210> 16

<211> 1323

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding fusion protein comprising Brachyury  
(L254V) Protein plus addition of 15 nucleotides from Gene FPV 088

<400> 16

atgaaaaata acttgatgag ctcccttggc accgagagcg cgggaaagag cctgcagtac 60

cgagtggacc acctgctgag cgccgtggag aatgagctgc aggcgggcag cgagaagggc 120

gacccacag agcggaact gcgcgtgggc ctggaggaga gcgagctgtg gctgcgttc 180

aaggagctca ccaatgagat gatcgtgacc aagaacggca ggaggatgtt tccggtgctg 240

aaggtgaacg tgtctggcct ggaccccaac gccatgtact ctttctgct ggacttcgtg 300

gcggcggaca accaccgtg gaagtacgtg aacggggaat gggcgccggg gggcaagccg 360

gagccgcagg cgcccagctg cgtctacatc caccctgact cgcccaactt cggggccac 420

tggatgaagg ctcccgctc cttcagcaaa gtcaagctca ccaacaagct caacggaggg 480

ggccagatca tgctgaactc cttgcataag tatgagctc gaatccacat agtgagagtt 540

gggggtccac agcgcatgat caccagccac tgcttccctg agaccagtt catagcggtg 600

actgcttacc agaacgagga gatcacagct cttaaaatta agtacaatcc atttgcaaag 660

gctttccttg atgcaaagga aagaagtgat cacaagaga tgatggagga acccgagac 720

agccagcaac ctgggtactc ccaatggggg tggtttcttc ctggaaccag caccgttgt 780

ccacctgcaa atctcatcc tcagtttga ggtgcctct cctccctc caccgacagc 840

tgtgacaggt acccaacct gaggagccac cgttcctc cctacccag cccctatgt 900

catcggaaca attctcaac ctattctgac aactcacctg catgtttatc catgctgcaa 960

tcccatgaca attggtccag ccttgaatg cctgccatc ccagcatgct cccgtgagc 1020

cacaatgcca gccacctac cagctccagt cagtaccca gcctgtggtc tgtgagcaac 1080

ggcgccgtca cccgggctc ccaggcagca gccgtgtcca acgggctggg ggcccagttc 1140

ttccggggt ccccgcgca ctacacacc ctacccatc cgtctcggc gccctcttc 1200

tcgggatccc cactgtacga aggggcggcc gcggccacag acatcgtgga cagccagtac 1260

gacgccgcag cccaaggccg cctcatagcc tcattggacac ctgtgtgcc accttccatg 1320

tga 1323

<210> 17

<211> 1320

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding fusion protein comprising Brachyury  
(L254V) Protein plus addition of 15 nucleotides from Gene FPV 088  
without Brachyury protein atg start codon

<400> 17

atgaaaaata acttgagctc ccctggcacc gagagcgagg gaaagagcct gcagtaccga 60

gtggaccacc tgctgagcgc cgtggagaat gagctgcagg cgggcagcga gaaggcgac 120

cccacagagc gcgaactgcg cgtgggcctg gaggagagcg agctgtggct gcgcttcaag 180

gagctcacca atgagatgat cgtgaccaag aacggcagga ggatgtttcc ggtgctgaag 240

gtgaactgtg ctggcctgga ccccaacgcc atgtactcct tctgctgga ctctgtggcg 300

gcggacaacc accgctggaa gtactgaac ggggaatggg tgccgggggg caagccggag 360

ccgcaggcgc ccagctgcgt ctacatccac cccgactcgc ccaacttcgg ggcctactgg 420

atgaaggctc ccgtctcctt cagcaaagtc aagctcacca acaagctcaa cggagggggc 480

cagatcatgc tgaactcctt gcataagtat gagcctcgaa tccacatagt gagagtggg 540

ggtccacagc gcatgatcac cagccactgc ttcctgaga cccagtcat agcggtgact 600

gcttatcaga acgaggagat cacagctctt aaaattaagt acaatccatt tgcaaaggct 660

ttccttgatg caaaggaaag aagtgatcac aaagagatga tggaggaacc cggagacagc 720

cagcaacctg ggtactccca atgggggtgg ctctctctg gaaccagcac cgtttgtcca 780

cctgcaaate ctatctctca gtttgagggt gcctctccc tcccctccac gcacagctgt 840

gacaggtacc caacctgag gagccaccgg tctcaccct acccagccc ctatgctcat 900

cggacaatt ctccaacctt ttctgacaac tcacctgcat gtttatccat gctgcaatcc 960

catgacaatt ggtccagcct tggaatgcct gcccatccca gcatgctccc cgtgagccac 1020

aatgccagcc cacctaccag ctccagtcag taccacagcc tgtggtctgt gagcaacggc 1080

gccgtcacc cgggctccca ggcagcagcc gtgtccaacg ggctgggggc ccagtcttc 1140

cggggctccc ccgcgacta cacacccctc acccatccgg tctcggcgc ctcttctcgc 1200

ggatccccac tgtacgaagg ggccggcgcg gccacagaca tcgtggacag ccagtacgac 1260

gccgcagccc aaggccgcct catagcctca tggacacctg tgtcgccacc ttccatgtga 1320

<210> 18

<211> 440

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by

SEQ ID NO:16

<400> 18

Met Lys Asn Asn Leu Met Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys

1 5 10 15

Ser Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu

20 25 30

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

35 40 45

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

50 55 60

Asn Glu Met Ile Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu

65 70 75 80

Lys Val Asn Val Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu

85 90 95

Leu Asp Phe Val Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly

100 105 110

Glu Trp Val Pro Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val

115 120 125

Tyr Ile His Pro Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala

130 135 140

Pro Val Ser Phe Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly

145 150 155 160

Gly Gln Ile Met Leu Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His

165 170 175

Ile Val Arg Val Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe

180 185 190

Pro Glu Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile

195                      200                      205  
 Thr Ala Leu Lys Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp  
 210                      215                      220  
 Ala Lys Glu Arg Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp  
 225                      230                      235                      240  
 Ser Gln Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr  
 245                      250                      255  
  
 Ser Thr Val Cys Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala  
 260                      265                      270  
 Leu Ser Leu Pro Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg  
 275                      280                      285  
 Ser His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn  
 290                      295                      300  
 Ser Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln  
 305                      310                      315                      320  
  
 Ser His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met  
 325                      330                      335  
 Leu Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr  
 340                      345                      350  
 Pro Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln  
 355                      360                      365  
 Ala Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser  
 370                      375                      380  
  
 Pro Ala His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser  
 385                      390                      395                      400  
 Ser Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val  
 405                      410                      415  
 Asp Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp  
 420                      425                      430  
 Thr Pro Val Ser Pro Pro Ser Met  
 435                      440

<210> 19

<211> 439

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by  
SEQ ID NO:17 plus addition of 15 nucleotides from Gene FPV 088  
without Brachyury atg start codon.

<400> 19

Met Lys Asn Asn Leu Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser

1 5 10 15

Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu

20 25 30

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

35 40 45

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

50 55 60

Glu Met Ile Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys

65 70 75 80

Val Asn Val Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu

85 90 95

Asp Phe Val Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu

100 105 110

Trp Val Pro Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr

115 120 125

Ile His Pro Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro

130 135 140

Val Ser Phe Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly

145 150 155 160

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

165 170 175

Val Arg Val Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro

180 185 190

Glu Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr  
 195 200 205  
 Ala Leu Lys Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala  
 210 215 220  
 Lys Glu Arg Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser  
 225 230 235 240  
 Gln Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser  
 245 250 255  
 Thr Val Cys Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu  
 260 265 270  
 Ser Leu Pro Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser  
 275 280 285  
 His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser  
 290 295 300  
 Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser  
 305 310 315 320  
 His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met Leu  
 325 330 335  
 Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro  
 340 345 350  
 Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala  
 355 360 365  
 Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro  
 370 375 380  
 Ala His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser  
 385 390 395 400  
 Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp  
 405 410 415  
 Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr  
 420 425 430  
 Pro Val Ser Pro Pro Ser Met

435

<210> 20

<211> 1353

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding fusion protein comprising Brachyury  
(L254V) Protein plus addition of 45 nucleotides from Gene FPV 088

<400> 20

```

atgaaaaata acttgtatga agaaaaaatg aacatgagta agaaaatgag ctccctggc      60
accgagagcg cgggaaagag cctgcagtac cgagtggacc acctgctgag cgccgtggag      120
aatgagctgc aggcgggcag cgagaagggc gacccacag agcgcgaaact gcgcgtgggc      180

ctggaggaga gcgagctgtg gctgcgcttc aaggagctca ccaatgagat gatcgtgacc      240
aagaacggca ggaggatgtt tccggtgctg aaggtgaacg tgtctggcct ggacccaac      300
gccatgtact ctttctctgt ggacttcgtg gcggcggaca accaccgtg gaagtacgtg      360
aacggggaat ggggtgccggg gggcaagccg gagccgcagg cgcccagctg cgtctacatc      420
caccgcgact cgccaactt cggggccac tgatgaagg ctcccgcttc cttcagcaaa      480
gtcaagctca ccaacaagct caacggaggg ggccagatca tgtgaactc cttgcataag      540
tatgagcctc gaatccacat agtgagagtt ggggggtccac agcgcatgat caccagccac      600

tgcttcctg agaccagtt catagcgtg actgcttate agaacgagga gatcacagct      660
cttaaaatta agtacaatcc atttgcaaag gctttccttg atgcaaagga aagaagtgat      720
caciaagaga tgatggagga acccgagac agccagcaac ctgggtactc ccaatggggg      780
tggtttcttc ctggaaccag caccgtttgt ccacctgcaa atcctcatcc tcagtttgga      840
ggtgcctctt cctccctc cagcacagc tgtgacaggt acccaaccct gaggagccac      900
cggctctcac cctacccag cccctatgct catcggaaca attctcaac ctattctgac      960
aactcacctg catgtttatc catgctgcaa tcccatgaca attggtccag ccttggaatg      1020

ctgcccac cagcatgct ccccgtagc cacaatgcca gccacctac cagctccagt      1080
cagtaccca gcctgtggtc tgtgagcaac ggcccgctca ccccgggctc ccaggcagca      1140
gccgtgtcca acgggctggg ggcccagttc ttccggggct ccccgcgca ctacacacc      1200
ctcaccac cggctctggc gccctcttcc tcgggatccc cactgtacga aggggcggcc      1260
gcggccacag acatcgtgga cagccagtac gacgccgag ccaaggccg cctcatagcc      1320
tcatggacac ctgtgtcgcc accttccatg tga      1353

```

<210> 21

<211> 1350

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding fusion protein comprising Brachyury  
(L254V) Protein plus addition of 45 nucleotides from Gene FPV 088  
without Brachyury atg

<400> 21

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| atgaaaaata acttgatga agaaaaaatg aacatgagta agaaaagctc ccctggcacc  | 60  |
| gagagcgagg gaaagagcct gcagtaccga gtggaccacc tgctgagcgc cgtggagaat | 120 |
| gagctgcagg cgggcagcga gaagggcgac cccacagagc gcgaactgcg cgtgggcctg | 180 |
| gaggagagcg agctgtggct gcgcttcaag gagctcacca atgagatgat cgtgaccaag | 240 |
| aacggcagga ggatgtttcc ggtgctgaag gtgaacgtgt ctggcctgga cccaacgcc  | 300 |

|                                                                   |     |
|-------------------------------------------------------------------|-----|
| atgtactcct tctgctgga ctctgtggcg gcggacaacc accgctggaa gtacgtgaac  | 360 |
| ggggaatggg tgccgggggg caagccggag ccgcaggcgc ccagctgcgt ctacatccac | 420 |
| cccgactcgc ccaacttcgg ggcccactgg atgaaggctc ccgtctcctt cagcaaagtc | 480 |
| aagctcacca acaagctcaa cggagggggc cagatcatgc tgaactcctt gcataagtat | 540 |
| gagcctcgaa tccacatagt gagagtggg ggtccacagc gcatgatcac cagccactgc  | 600 |
| ttccctgaga cccagttcat agcggtgact gcttatcaga acgaggagat cacagctctt | 660 |
| aaaattaagt acaatccatt tgcaaaggct ttccttgatg caaaggaaag aagtgatcac | 720 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| aaagagatga tggaggaacc cggagacagc cagcaacctg ggtactccca atgggggtgg | 780  |
| cttcttctct gaaccagcac cgtttgtcca cctgcaaatc ctcatcctca gtttgagggt | 840  |
| gccctctccc tcccctccac gcacagctgt gacaggtagc caaccctgag gagccaccgg | 900  |
| tcctcacctt accccagccc ctatgctcat cggacaatt ctccaaccta ttctgacaac  | 960  |
| tcactgcat gtttatccat gctgcaatcc catgacaatt ggtccagcct tggaatgcct  | 1020 |
| gcccatacca gcatgtccc cgtgagccac aatgccagcc cacctaccag ctccagtcag  | 1080 |
| taccacagcc tgggtctgt gagcaacggc gccgtcacc cgggctccca ggcagcagcc   | 1140 |

|                                                                   |      |
|-------------------------------------------------------------------|------|
| gtgtccaacg ggctgggggc ccagttcttc cggggctccc ccgcgacta cacaccctc   | 1200 |
| accatccgg tctcggcgcc ctcttctctg ggatccccac tgtacgaagg ggcggccgcg  | 1260 |
| gccacagaca tcgtggacag ccagtacgac gccgcagccc aaggccgcct catagcctca | 1320 |
| tggacacctg tgctgccacc ttccatgtga                                  | 1350 |

<210> 22

<211> 450

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by  
SEQ ID NO:20

<400> 22

Met Lys Asn Asn Leu Tyr Glu Glu Lys Met Asn Met Ser Lys Lys Met

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

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

|     |     |     |
|-----|-----|-----|
| 130 | 135 | 140 |
| Pro | Asn | Phe |
| Gly | Ala | His |
| Trp | Met | Lys |
| Ala | Pro | Val |
| Ser | Phe | Ser |
| Lys |     |     |
| 145 | 150 | 155 |
| Val | Lys | Leu |
| Thr | Asn | Lys |
| Leu | Asn | Gly |
| Gly | Gly | Gln |
| Ile | Met | Leu |
| Asn |     |     |
| 165 | 170 | 175 |
| Ser | Leu | His |
| Lys | Tyr | Glu |
| Pro | Arg | Ile |
| His | Ile | Val |
| Arg | Val | Gly |
| Gly |     |     |
| 180 | 185 | 190 |
| Pro | Gln | Arg |
| Met | Ile | Thr |
| Ser | His | Cys |
| Phe | Pro | Glu |
| Thr | Gln | Phe |
| Ile |     |     |

195                      200                      205  
 Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile Lys  
 210                      215                      220  
 Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser Asp  
 225                      230                      235                      240  
 His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly Tyr  
 245                      250                      255  
 Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Val Cys Pro Pro  
  
 260                      265                      270  
 Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr  
 275                      280                      285  
 His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser Pro  
 290                      295                      300  
 Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser Asp  
 305                      310                      315                      320  
 Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp Ser  
  
 325                      330                      335  
 Ser Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser His Asn  
 340                      345                      350  
 Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val  
 355                      360                      365  
 Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser Asn  
 370                      375                      380  
 Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr Pro  
  
 385                      390                      395                      400  
 Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr  
 405                      410                      415  
 Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala  
 420                      425                      430  
 Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro Pro  
 435                      440                      445  
 Ser Met

450

<210> 23

<211> 449

<212

> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by

SEQ ID NO:21

<400> 23

Met Lys Asn Asn Leu Tyr Glu Glu Lys Met Asn Met Ser Lys Lys Ser

1 5 10 15

Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr Arg Val Asp

20 25 30

His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly Ser Glu Lys

35 40 45

Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly Leu Glu Glu Ser Glu

50 55 60

Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile Val Thr Lys

65 70 75 80

Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val Ser Gly Leu

85 90 95

Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val Ala Ala Asp

100 105 110

Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro Gly Gly Lys

115 120 125

Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp Ser Pro

130 135 140

Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser Lys Val

145 150 155 160

Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met Leu Asn Ser

165 170 175

Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val Gly Gly Pro

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
| 180                                                             | 185 | 190 |     |
| Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln Phe Ile Ala |     |     |     |
| 195                                                             | 200 | 205 |     |
| Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile Lys Tyr |     |     |     |
| 210                                                             | 215 | 220 |     |
| Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser Asp His |     |     |     |
| 225                                                             | 230 | 235 | 240 |
|                                                                 |     |     |     |
| Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly Tyr Ser |     |     |     |
| 245                                                             | 250 | 255 |     |
| Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Val Cys Pro Pro Ala |     |     |     |
| 260                                                             | 265 | 270 |     |
| Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr His |     |     |     |
| 275                                                             | 280 | 285 |     |
| Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser Pro Tyr |     |     |     |
| 290                                                             | 295 | 300 |     |
|                                                                 |     |     |     |
| Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser Asp Asn |     |     |     |
| 305                                                             | 310 | 315 | 320 |
| Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp Ser Ser |     |     |     |
| 325                                                             | 330 | 335 |     |
| Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser His Asn Ala |     |     |     |
| 340                                                             | 345 | 350 |     |
| Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val Ser |     |     |     |
| 355                                                             | 360 | 365 |     |
|                                                                 |     |     |     |
| Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser Asn Gly |     |     |     |
| 370                                                             | 375 | 380 |     |
| Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr Pro Leu |     |     |     |
| 385                                                             | 390 | 395 | 400 |
| Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr Glu |     |     |     |
| 405                                                             | 410 | 415 |     |
| Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala Ala |     |     |     |
| 420                                                             | 425 | 430 |     |

Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro Pro Ser

435

440

445

Met

<210> 24

<211> 1326

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding fusion protein comprising Brachyury  
(L254V) Protein plus addition of 18 nucleotides from Gene FPV 088

<400> 24

atgaacatga gtaagaaaat gagctccctt ggcaccgaga gcgcgggaaa gagcctgcag 60  
taccgagtgg accacctgct gagcgccgtg gagaatgagc tgcaggcggg cagcgagaag 120

ggcgacccca cagagcgaga actgcgcgtg ggcttgagg agagcgagct gtggctgcgc 180  
ttcaaggagc tcaccaatga gatgatcgtg accaagaacg gcaggaggat gtttccggtg 240  
ctgaaggatga acgtgtctgg cctggacccc aacgccatgt actccttctt gctggacttc 300  
gtggcggcgg acaaccaccg ctggaagtac gtgaacgggg aatgggtgcc ggggggcaag 360  
ccggagccgc aggcgcccag ctgcgtctac atccaccccg actcgcccaa cttcggggcc 420  
cactggatga aggtctccgt ctcttcagc aaagtcaagc tcaccaacaa gctcaacgga 480  
gggggccaga tcatgctgaa ctcttgcgt aagtatgagc ctgcaatcca catagtgaga 540

gttgggggtc cacagcgcat gatcaccagc cactgcttcc ctgagacca gtcatagcg 600  
gtgactgctt atcagaacga ggagatcaca gctcttaaaa ttaagtacaa tccatttgca 660  
aaggctttcc ttgatgcaaa ggaaagaagt gatcacaag agatgatgga ggaaccgga 720  
gacagccagc aacctgggta ctccaatgg gggtggcttc ttcttgaac cagcaccgtt 780  
tgtccacctg caaatctca tctcagttt ggaggtgccc tctccctccc ctccacgcac 840  
agctgtgaca ggiaccaaac cctgaggagc caccgttctt caccctaccc cagcccttat 900  
gctcatcgga acaattctcc aacctattct gacaactcac ctgcatgttt atccatgctg 960

caatcccatg acaattggc cagccttgga atgcctgccc atccagcat gctccccgtg 1020  
agccacaatg ccagcccacc taccagctcc agtcagtacc ccagcctgtg gtctgtgagc 1080  
aacggcgccg tcaccccgga ctccaggca gcagccgtgt ccaacgggct gggggcccag 1140  
ttcttccggg gctccccgc gcactacaca cccctaccc atccggtctc ggcgccctct 1200  
tcctcgggat cccactgta cgaaggggag gccgcggcca cagacatcgt ggacagccag 1260

tacgacgccg cagcccaagg ccgcctcata gcctcatgga cacctgtgtc gccacattcc 1320  
atgtga 1326

<210> 25  
<211> 1323  
<212> DNA  
<213> Artificial Sequence  
<220><223> Coding sequence encoding fusion protein comprising Brachyury  
(L254V) Protein plus addition of 18 nucleotides from Gene FPV 088  
without Brachyury atg

<400> 25

atgaacatga gtaagaaaag ctccccctggc accgagagcg cgggaaagag cctgcagtac 60  
cgagtggacc acctgctgag cgccgtggag aatgagctgc aggcgggcag cgagaagggc 120  
gacccacag agcgcgaact gcgcgtgggc ctggaggaga gcgagctgtg gctgcgcttc 180  
aaggagctca ccaatgagat gatcgtgacc aagaacggca ggaggatgtt tccggtgctg 240

aaggtgaacg tgcttggcct ggacccaac gccatgtact ccttctctgt ggacttcgtg 300  
gcggcggaca accaccgtg gaagtacgtg aacggggaat gggtgccggg gggcaagccg 360  
gagccgcagg cgcccagctg cgtctacatc caccctgact cgcccaactt cggggccac 420  
tggatgaagg ctcccgctc cttcagcaaa gtcaagctca ccaacaagct caacggaggg 480  
ggccagatca tctgaactc cttgcataag tatgagctc gaatccacat agtgagagtt 540  
gggggtccac agcgcagat caccagccac tgcttccctg agaccagtt catagcggtg 600  
actgcttacc agaacgagga gatcacagct cttaaaatta agtacaatcc atttgcaaag 660

gctttccttg atgcaaagga aagaagtgat cacaagaga tgatggagga acccgagac 720  
agccagcaac ctgggtactc ccaatggggg tggtttcttc ctggaaccag caccgtttgt 780  
ccacctgcaa atctcatcc tcagtttga ggtgccctct cctccctc cagcacagc 840  
tgtgacaggt acccaacct gaggagccac cgttctctac cctacccag cccctatgt 900  
catcggaaca attctcaac ctattctgac aactcactg catgtttatc catgctgcaa 960  
tcccatgaca attggtccag ccttggatg cctgccatc ccagcatgt cccgtgagc 1020  
cacaatgcca gccacctac cagctccagt cagtaccca gcctgtggc tgtgagcaac 1080

ggcgcctga ccccggtc ccagcagca gccgtgtcca acgggctggg ggcccagttc 1140  
ttccggggt ccccgcgca ctacacacc ctacccatc cgtctcggc gccctcttc 1200  
tcgggatccc cactgtacga agggcgggc gcggccacag acatcgtgga cagccagtac 1260

gacgccgcag cccaaggccg cctcatagcc tcatggacac ctgtgtcgcc accttccatg 1320  
 tga 1323  
 <210> 26  
 <211> 441  
 <212> PRT  
 <213> Artificial Sequence  
 <220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by

SEQ ID NO:24

<400> 26  
 Met Asn Met Ser Lys Lys Met Ser Ser Pro Gly Thr Glu Ser Ala Gly  
 1 5 10 15  
 Lys Ser Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn  
 20 25 30  
 Glu Leu Gln Ala Gly Ser Glu Lys Gly Asp Pro Thr Glu Arg Glu Leu  
 35 40 45  
 Arg Val Gly Leu Glu Glu Ser Glu Leu Trp Leu Arg Phe Lys Glu Leu  
 50 55 60  
  
 Thr Asn Glu Met Ile Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val  
 65 70 75 80  
 Leu Lys Val Asn Val Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe  
 85 90 95  
 Leu Leu Asp Phe Val Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn  
 100 105 110  
 Gly Glu Trp Val Pro Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys  
 115 120 125  
  
 Val Tyr Ile His Pro Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys  
 130 135 140  
 Ala Pro Val Ser Phe Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly  
 145 150 155 160  
 Gly Gly Gln Ile Met Leu Asn Ser Leu His Lys Tyr Glu Pro Arg Ile  
 165 170 175  
 His Ile Val Arg Val Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys

|                                                                 |     |     |
|-----------------------------------------------------------------|-----|-----|
| 180                                                             | 185 | 190 |
| Phe Pro Glu Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu |     |     |
| 195                                                             | 200 | 205 |
| Ile Thr Ala Leu Lys Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu |     |     |
| 210                                                             | 215 | 220 |
| Asp Ala Lys Glu Arg Ser Asp His Lys Glu Met Met Glu Glu Pro Gly |     |     |
| 225                                                             | 230 | 235 |
| Asp Ser Gln Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly |     |     |
| 245                                                             | 250 | 255 |
| Thr Ser Thr Val Cys Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly |     |     |
| 260                                                             | 265 | 270 |
| Ala Leu Ser Leu Pro Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu |     |     |
| 275                                                             | 280 | 285 |
| Arg Ser His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn |     |     |
| 290                                                             | 295 | 300 |
| Asn Ser Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu |     |     |
| 305                                                             | 310 | 315 |
| Gln Ser His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser |     |     |
| 325                                                             | 330 | 335 |
| Met Leu Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln |     |     |
| 340                                                             | 345 | 350 |
| Tyr Pro Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser |     |     |
| 355                                                             | 360 | 365 |
| Gln Ala Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly |     |     |
| 370                                                             | 375 | 380 |
| Ser Pro Ala His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser |     |     |
| 385                                                             | 390 | 395 |
| Ser Ser Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile |     |     |
| 405                                                             | 410 | 415 |
| Val Asp Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser |     |     |
| 420                                                             | 425 | 430 |

Trp Thr Pro Val Ser Pro Pro Ser Met

435

440

<210> 27

<211> 440

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by

SEQ ID NO:25

<400> 27

Met Asn Met Ser Lys Lys Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys

1 5 10 15

Ser Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu

20 25 30

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

35 40 45

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

50 55 60

Asn Glu Met Ile Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu

65 70 75 80

Lys Val Asn Val Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu

85 90 95

Leu Asp Phe Val Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly

100 105 110

Glu Trp Val Pro Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val

115 120 125

Tyr Ile His Pro Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala

130 135 140

Pro Val Ser Phe Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly

145 150 155 160

Gly Gln Ile Met Leu Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His

165 170 175

Ile Val Arg Val Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe  
 180 185 190  
 Pro Glu Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile  
 195 200 205  
 Thr Ala Leu Lys Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp  
 210 215 220  
 Ala Lys Glu Arg Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp  
 225 230 235 240  
  
 Ser Gln Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr  
 245 250 255  
 Ser Thr Val Cys Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala  
 260 265 270  
 Leu Ser Leu Pro Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg  
 275 280 285  
 Ser His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn  
 290 295 300  
  
 Ser Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln  
 305 310 315 320  
 Ser His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met  
 325 330 335  
 Leu Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr  
 340 345 350  
 Pro Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln  
 355 360 365  
  
 Ala Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser  
 370 375 380  
 Pro Ala His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser  
 385 390 395 400  
 Ser Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val  
 405 410 415  
 Asp Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp

420

425

430

Thr Pro Val Ser Pro Pro Ser Met

435

440

<210> 28

<211> 1320

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury

(L254V) Protein plus addition of 12 nucleotides from Gene FPV 088

<400> 28

```

atgagtaaga aaatgagctc ccctggcacc gagagcgcgg gaaagagcct gcagtaccga      60
gtggaccacc tgctgagcgc cgtggagaat gagctgcagg cgggcagcga gaagggcgac      120
cccacagagc gcgaactgcg cgtgggcctg gaggagagcg agctgtggct gcgcttcaag      180

gagctcacca atgagatgat cgtgaccaag aacggcagga ggatgtttcc ggtgctgaag      240
gtgaacgtgt ctggcctgga cccaacgcc atgtactcct tcctgctgga cttcgtggcg      300
gcggacaacc accgctggaa gtacgtgaac ggggaatggg tgccgggggg caagccggag      360
ccgcaggcgc ccagctgcgt ctacatccac ccgactcgc ccaacttcgg ggcccactgg      420
atgaaggctc ccgtctcctt cagcaaagtc aagctcacca acaagctcaa cggagggggc      480
cagatcatgc tgaactcctt gcataagtat gagcctcgaa tccacatagt gagagtggg      540
ggtccacagc gcatgatcac cagccactgc ttcctgaga ccagttcat agcggtgact      600

gcttatcaga acgaggagat cacagctctt aaaattaagt acaatccatt tgcaaaggct      660
ttccttgatg caaaggaaag aagtgatcac aaagagatga tggaggaacc cggagacagc      720
cagcaacctg ggtactccca atgggggtgg ctctctctg gaaccagcac cgtttgtcca      780
cctgcaaate ctatctctca gtttggaggt gccctctccc tcccctccac gcacagctgt      840
gacaggtacc caacctgag gageccacgg tctcaccct accccagccc ctatgtcat      900
cggaacaatt ctccaaccta ttctgacaac tcacctgcat gtttatccat gctgcaatcc      960
catgacaatt ggtccagcct tggaatgcct gcccatccca gcatgctccc cgtgagccac     1020

aatgccagcc cacctaccag ctccagtcag taccacagcc tgtggtctgt gagcaacggc     1080
gccgtcacc cgggctccca ggcagcagcc gtgtccaacg ggctgggggc ccagttcttc     1140
cggggtctcc ccgcgacta cacaccctc acccatcggg tctcggcgcc ctcttctctg     1200
ggatccccac tgiacgaagg ggcggccgcg gccacagaca tcgtggacag ccagtagcac     1260

```

gccgcagccc aaggccgcct catagcctca tggacacctg tgtcgccacc ttccatgtga 1320  
 <210> 29  
 <211> 1317  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Coding sequence encoding Fusion protein comprising Brachyury

(L254V) Protein plus addition of 12 nucleotides from Gene FPV 088  
 without Brachyury atg start codon

<400> 29  
 atgagtaaga aaagctcccc tggcaccgag agcgcgggaa agagcctgca gtaccgagtg 60  
 gaccacctgc tagcgccgt ggagaatgag ctgcaggcgg gcagcgagaa gggcgacccc 120  
 acagagcgcg aactgcgcgt gggcctggag gagagcgagc tgtggctgcg cttcaaggag 180  
 ctccaatg agatgatcgt gaccaagaac ggcaggagga tgtttccggt gctgaagggtg 240  
 aacgtgtctg gcctggaccc caacgcatg tactccttc tgcctggactt cgtggcgcg 300  
  
 gacaaccacc gctggaagta cgtgaacggg gaatgggtgc cggggggcaa gccggagccg 360  
 caggcgccca gctgcgtcta catccacccc gactcgcca acttcggggc cactggatg 420  
 aaggctcccg tctccttcag caaagtcaag ctaccaaca agctcaacgg agggggccag 480  
 atcatgctga actccttgca taagtatgag cctcgaatcc acatagttag agttgggggt 540  
 ccacagcgca tgatcaccag ccaactgttc cctgagaccc agttcatagc ggtgactgct 600  
 tatcagaacg aggagatcac agctcttaaa attaagtaca atccatttgc aaaggctttc 660  
 cttgatgcaa aggaaagaag tgatcacaaa gagatgatgg aggaacccgg agacagccag 720  
  
 caacctgggt actcccaatg ggggtggctt ctctctggaa ccagcaccgt ttgtccacct 780  
 gcaaatcctc atcctcagtt tggaggtgcc ctctccctcc cctccacgca cagctgtgac 840  
 aggtacccaa cctgaggag ccaccggtcc tcaccctacc ccagccccta tgctcatcgg 900  
 aacaattctc caacctatc tgacaactca cctgcatgtt tatccatgct gcaatcccat 960  
 gacaattggt ccagccttgg aatgcctgcc catcccagca tgctccccgt gagccacaat 1020  
 gccagccac ctaccagctc cagtcagtac ccagcctgt ggtctgtgag caacggcgcc 1080  
 gtcacccgg gctcccagc agcagccgtg tccaacgggc tgggggcca gttcttcgg 1140  
  
 ggctccccg cgactacac accctcacc catccgtct cggcgccctc tctctcgga 1200  
 tccccactgt acgaagggg gcccgggcc acagacatcg tggacagcca gtacgagcc 1260  
 gcagcccaag gccgcctcat agcctcatgg acacctgtgt cgccacctc catgtga 1317

<210> 30

<211> 439

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by  
SEQ ID NO:28

<400> 30

Met Ser Lys Lys Met Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser

1 5 10 15

Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu

20 25 30

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

35 40 45

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

50 55 60

Glu Met Ile Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys

65 70 75 80

Val Asn Val Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu

85 90 95

Asp Phe Val Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu

100 105 110

Trp Val Pro Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr

115 120 125

Ile His Pro Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro

130 135 140

Val Ser Phe Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly

145 150 155 160

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

165 170 175

Val Arg Val Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro

180 185 190

Glu Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
| 195                                                             | 200 | 205 |     |
| Ala Leu Lys Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala |     |     |     |
| 210                                                             | 215 | 220 |     |
| Lys Glu Arg Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser |     |     |     |
| 225                                                             | 230 | 235 | 240 |
| Gln Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser |     |     |     |
|                                                                 | 245 | 250 | 255 |
| Thr Val Cys Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu |     |     |     |
|                                                                 | 260 | 265 | 270 |
| Ser Leu Pro Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser |     |     |     |
|                                                                 | 275 | 280 | 285 |
| His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser |     |     |     |
|                                                                 | 290 | 295 | 300 |
| Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser |     |     |     |
| 305                                                             | 310 | 315 | 320 |
| His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met Leu |     |     |     |
|                                                                 | 325 | 330 | 335 |
| Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro |     |     |     |
|                                                                 | 340 | 345 | 350 |
| Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala |     |     |     |
|                                                                 | 355 | 360 | 365 |
| Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro |     |     |     |
|                                                                 | 370 | 375 | 380 |
| Ala His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser |     |     |     |
| 385                                                             | 390 | 395 | 400 |
| Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp |     |     |     |
|                                                                 | 405 | 410 | 415 |
| Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr |     |     |     |
|                                                                 | 420 | 425 | 430 |
| Pro Val Ser Pro Pro Ser Met                                     |     |     |     |
|                                                                 | 435 |     |     |

<210> 31

<211> 438

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by  
SEQ ID NO:29

<400> 31

Met Ser Lys Lys Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu

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

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

|                                                                 |     |     |
|-----------------------------------------------------------------|-----|-----|
| 130                                                             | 135 | 140 |
| Ser Phe Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln |     |     |
| 145                                                             | 150 | 155 |
| Ile Met Leu Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile Val |     |     |
|                                                                 | 165 | 170 |
| Arg Val Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu |     |     |
|                                                                 | 180 | 185 |
|                                                                 |     | 190 |

Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala

195 200 205

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

210 215 220

Glu Arg Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln

225 230 235 240

Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr

245 250 255

Val Cys Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser

260 265 270

Leu Pro Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His

275 280 285

Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro

290 295 300

Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His

305 310 315 320

Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met Leu Pro

325 330 335

Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser

340 345 350

Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala

355 360 365

Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala

370 375 380

His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly

385 390 395 400

Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser

405 410 415

Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro

420 425 430

Val Ser Pro Pro Ser Met

435

<210> 32

<211> 1386

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury

(L254V) Protein plus addition of 75 nucleotides from Gene FPV 088

and an atc

<400> 32

```

atgaaaaata acttgtatga agaaaaaatg aacatgagta agaaacaagt aaaaactcaa      60
agtaaagtga ataatatcat gagctcccct ggccaccgaga gcgcgggaaa gaggctgcag      120
taccgagtgg accacctgct gagcgccgtg gagaatgagc tgcaggcggg cagcgagaag      180
ggcgacccca cagagcgca actgcgcgtg ggcttgagg agagcgagct gtggctgcgc      240
ttcaaggagc tcaccaatga gatgatcgtg accaagaacg gcaggaggat gtttcggtg      300
ctgaaggatga acgtgtctgg cctggacccc aacgcatgt actccttct gctggacttc      360

gtggcggcgg acaaccaccg ctggaagtac gtgaacgggg aatgggtgcc ggggggcaag      420
ccggagccgc aggcgcccat ctgcgtctac atccaccccg actcgcccaa cttcggggcc      480
cactggatga aggtcccgt ctccttcagc aaagtcaagc tcaccaacaa gctcaacgga      540
gggggcccaga tcatgctgaa ctcttgcac aagtatgagc ctggaatcca catagtgaga      600
gttgggggtc cacagcgcat gatcaccagc cactgcttcc ctgagacca gtcatagcg      660
gtgactgctt atcagaacga ggagatcaca gctcttaaaa ttaagtacaa tccatttgca      720
aaggctttcc ttgatgcaaa ggaaagaagt gatcacaag agatgatgga ggaacccgga      780

gacagccagc aacctgggta ctccaatgg ggttggttc ttcctggaac cagcacggtt      840
tgtccacctg caaatctca tctcagttt ggaggtgccc tctccctccc ctccacgcac      900
agctgtgaca ggtacceaac cctgaggagc caccggtcct caccctacce cagccctat      960
gtcatcgga acaattctcc aacctattct gacaactcac ctgcatgttt atccatgctg     1020
caatcccatg acaattggc cagccttga atgctgccc atccagcat gctccccgtg     1080
agccacaatg ccagcccacc taccagctcc agtcagtacc ccagcctgtg gtctgtgagc     1140
aacggcgccg tcacccggg ctccaggca gcagcgtgt ccaacgggct gggggcccag     1200

ttttccggg gtcceccgc gcactacaca cccctaccc atccggtctc ggccccctct     1260
tctcgggat cccactgta cgaaggggcg gccgcggcca cagacatcgt ggacagccag     1320

```

tacgacgccg cagcccaagg ccgcctcata gcctcatgga cacctgtgtc gccacattcc 1380  
atgtga 1386  
<210> 33  
<211> 1383  
<212> DNA  
<213> Artificial Sequence  
<220><223> Coding sequence encoding Fusion protein comprising Brachyury  
(L254V) Protein plus addition of 75 nucleotides from Gene FPV 088  
  
and an atc, without Brachyury atg start codon  
<400> 33  
atgaaaaata acttgtatga agaaaaaatg aacatgagta agaaacaagt aaaaactcaa 60  
agtaaatgta ataatatcag ctcccctggc accgagagcg cgggaaagag cctgcagtac 120  
cgagtggacc acctgctgag cgccgtggag aatgagctgc aggcgggcag cgagaagggc 180  
gacccacag agcgcgaact gcgcgtgggc ctggaggaga gcgagctgtg gctgcgttc 240  
aaggagctca ccaatgagat gatcgtgacc aagaacggca ggaggatgtt tccggtgctg 300  
aaggtgaacg tgtctggcct ggacccaac gccatgtact ccttcctgct ggacttcgtg 360  
  
gcggcggaca accaccgtg gaagtacgtg aacggggaat gggtgccggg gggcaagccg 420  
gagccgcagg cgcccagctg cgtctacatc caccaccgact cgccaactt cggggccac 480  
tggatgaagg ctcccgtctc cttcagcaaa gtcaagctca ccaacaagct caacggaggg 540  
ggccagatca tgctgaactc cttgcataag tatgagcctc gaatccacat agtgagagtt 600  
gggggtccac agcgcagatg caccagccac tgcttccttg agaccagtt catagcggtg 660  
actgcttacc agaacgagga gatcacagct cttaaaatta agtacaatcc atttgcaaag 720  
gctttccttg atgcaaagga aagaagtgat cacaagaga tgatggagga acccgagac 780  
  
agccagcaac ctgggtactc ccaatggggg tggtctcttc ctggaaccag caccgttgt 840  
ccacctgcaa atcctcatcc tcagtttgga ggtgccctct cctccctc caccacagc 900  
tgtgacaggt acccaacct gaggagccac cgttcctcac cctacccag ccctatgct 960  
catcggaaca attctcaac ctattctgac aactcacctg catgtttatc catgctgcaa 1020  
tcccatgaca attggtccag ctttggatg cctgcccac ccagcatgct ccccgtagc 1080  
cacaatgcca gccacctac cagctccagt cagtaccca gcctgtggtc tgtgagcaac 1140  
ggcgccgtca ccccggtc ccaggcagca gccgtgtcca acgggctggg ggcccagttc 1200  
  
ttccggggct ccccgcgca ctacacacc ctacccatc cgtctcggc gcctcttcc 1260

tcgggatccc cactgtacga aggggcggcc gcggccacag acatcgtgga cagccagtac 1320  
gacgccgcag cccaaggccg cctcatagcc tcatggacac ctgtgtcgcc accttccatg 1380  
tga 1383

<210> 34

<211> 461

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by  
SEQ ID NO: 32

<400> 34

Met Lys Asn Asn Leu Tyr Glu Glu Lys Met Asn Met Ser Lys Lys Gln

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

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

130 135 140  
Ala Pro Ser Cys Val Tyr Ile His Pro Asp Ser Pro Asn Phe Gly Ala  
145 150 155 160  
His Trp Met Lys Ala Pro Val Ser Phe Ser Lys Val Lys Leu Thr Asn  
165 170 175

Lys Leu Asn Gly Gly Gly Gln Ile Met Leu Asn Ser Leu His Lys Tyr  
 180 185 190  
 Glu Pro Arg Ile His Ile Val Arg Val Gly Gly Pro Gln Arg Met Ile  
 195 200 205  
 Thr Ser His Cys Phe Pro Glu Thr Gln Phe Ile Ala Val Thr Ala Tyr  
 210 215 220  
 Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile Lys Tyr Asn Pro Phe Ala  
 225 230 235 240  
 Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser Asp His Lys Glu Met Met  
 245 250 255  
 Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly Tyr Ser Gln Trp Gly Trp  
 260 265 270  
 Leu Leu Pro Gly Thr Ser Thr Val Cys Pro Pro Ala Asn Pro His Pro  
 275 280 285  
 Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr His Ser Cys Asp Arg  
 290 295 300  
 Tyr Pro Thr Leu Arg Ser His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr  
 305 310 315 320  
 Ala His Arg Asn Asn Ser Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys  
 325 330 335  
 Leu Ser Met Leu Gln Ser His Asp Asn Trp Ser Ser Leu Gly Met Pro  
 340 345 350  
 Ala His Pro Ser Met Leu Pro Val Ser His Asn Ala Ser Pro Pro Thr  
 355 360 365  
 Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val Ser Asn Gly Ala Val  
 370 375 380  
 Thr Pro Gly Ser Gln Ala Ala Ala Val Ser Asn Gly Leu Gly Ala Gln  
 385 390 395 400  
 Phe Phe Arg Gly Ser Pro Ala His Tyr Thr Pro Leu Thr His Pro Val  
 405 410 415  
 Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala

420 425 430  
Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg  
435 440 445  
Leu Ile Ala Ser Trp Thr Pro Val Ser Pro Pro Ser Met

450 455 460  
<210> 35  
<211> 460  
<212> PRT  
<213> Artificial Sequence  
<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by  
SEQ ID NO: 33

<400> 35  
Met Lys Asn Asn Leu Tyr Glu Glu Lys Met Asn Met Ser Lys Lys Gln  
1 5 10 15  
Val Lys Thr Gln Ser Lys Cys Asn Asn Ile Ser Ser Pro Gly Thr Glu  
20 25 30  
Ser Ala Gly Lys Ser Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala

35 40 45  
Val Glu Asn Glu Leu Gln Ala Gly Ser Glu Lys Gly Asp Pro Thr Glu  
50 55 60  
Arg Glu Leu Arg Val Gly Leu Glu Glu Ser Glu Leu Trp Leu Arg Phe  
65 70 75 80  
Lys Glu Leu Thr Asn Glu Met Ile Val Thr Lys Asn Gly Arg Arg Met  
85 90 95  
Phe Pro Val Leu Lys Val Asn Val Ser Gly Leu Asp Pro Asn Ala Met

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

Trp Met Lys Ala Pro Val Ser Phe Ser Lys Val Lys Leu Thr Asn Lys

165 170 175

Leu Asn Gly Gly Gly Gln Ile Met Leu Asn Ser Leu His Lys Tyr Glu

180 185 190

Pro Arg Ile His Ile Val Arg Val Gly Gly Pro Gln Arg Met Ile Thr

195 200 205

Ser His Cys Phe Pro Glu Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln

210 215 220

Asn Glu Glu Ile Thr Ala Leu Lys Ile Lys Tyr Asn Pro Phe Ala Lys

225 230 235 240

Ala Phe Leu Asp Ala Lys Glu Arg Ser Asp His Lys Glu Met Met Glu

245 250 255

Glu Pro Gly Asp Ser Gln Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu

260 265 270

Leu Pro Gly Thr Ser Thr Val Cys Pro Pro Ala Asn Pro His Pro Gln

275 280 285

Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr His Ser Cys Asp Arg Tyr

290 295 300

Pro Thr Leu Arg Ser His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala

305 310 315 320

His Arg Asn Asn Ser Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu

325 330 335

Ser Met Leu Gln Ser His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala

340 345 350

His Pro Ser Met Leu Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser

355 360 365

Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr

370 375 380

Pro Gly Ser Gln Ala Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe

385 390 395 400

Phe Arg Gly Ser Pro Ala His Tyr Thr Pro Leu Thr His Pro Val Ser

405 410 415  
Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala

420 425 430  
Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu

435 440 445  
Ile Ala Ser Trp Thr Pro Val Ser Pro Pro Ser Met

450 455 460

<210> 36

<211> 1359

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury  
(L254V) Protein plus addition of 48 nucleotides from Gene FPV 088

and an atc

<400> 36

atgaacatga gtaagaaaca agtaaaaact caaagtaaat gtaataatat catgagctcc 60

cctggcaccg agagcgcggg aaagagcctg cagtaccgag tggaccacct gctgagcgcc 120

gtggagaatg agctgcaggc gggcagcgag aaggcgacc ccacagagcg cgaactgcgc 180

gtgggcctgg aggagagcga gctgtggctg cgcttcaagg agctcaccaa tgagatgatc 240

gtgaccaaga acggcaggag gatgtttccg gtgctgaagg tgaacgtgtc tggcctggac 300

cccaacgcca tgtactcctt cctgtggac ttctgtggcg cggacaacca ccgctggaag 360

tacgtgaacg gggaatgggt gccggggggc aagccggagc cgcaggcgcc cagctgcgtc 420

tacatccacc ccgactcgcc caacttcggg gccactgga tgaaggctcc cgtctccttc 480

agcaaagtca agctcaccaa caagctcaac ggagggggcc agatcatgct gaactccttg 540

cataagtatg agcctcgaat ccacatagtg agagttgggg gtcacagcg catgatcacc 600

agccactgct tcctgagac ccagttcata gcggtgactg cttatcagaa cgaggagatc 660

acagctctta aaattaagta caatccattt gcaaagcctt tccttgatgc aaaggaaaga 720

agtgatcaca aagagatgat ggaggaaccc ggagacagcc agcaacctgg gtactcccaa 780

tgggggtggc ttcttcctgg aaccagcacc gtttgtccac ctgcaaatcc tcctcctcag 840

tttggaggtg cctctccct cccctccacg cacagctgtg acaggtaccc aacctgagg 900

agccaccggt cctcacccta cccagcccc tatgctcacc ggaacaattc tccaacctat 960

tctgacaact cacctgcatg tttatccatg ctgcaatccc atgacaattg gtccagcctt 1020

ggaatgcctg cccatcccag catgtctccc gtgagccaca atgccagccc acctaccagc 1080

tccagtcagt accccagcct gtggtctgtg agcaacggcg ccgtcacccc gggtctccag 1140

gcagcagccg tgtccaacgg gctgggggcc cagttcttcc ggggtctccc cgcgcactac 1200

acaccctca cccatccggt ctgcgcgcc tcttctctcg gatccccact gtacgaaggg 1260

gcggccgagg ccacagacat cgtggacagc cagtacgacg ccgcagccca aggcgcctc 1320

atagcctcat ggacacctgt gtcgccacct tccatgtga 1359

<210> 37

<211> 1356

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury  
(L254V) Protein plus addition of 48 nucleotides from Gene FPV 088  
and an atc, without Brachyury atg start codon

<400> 37

atgaacatga gtaagaaaca agtaaaaact caaagtaa atgtaataat cagctcccct 60

ggcacccgaga gcgcgggaaa gagcctgcag taccgagtgg accacctgct gagcgccgtg 120

gagaatgagc tgcaggcggg cagcgagaag ggcgacccca cagagcgcga actgcgcgtg 180

ggcctggagg agagcgagct gtggctgcgc ttcaaggagc tcaccaatga gatgatcgtg 240

accaagaacg gcaggaggat gtttccggtg ctgaaggatga acgtgtctgg cctggacccc 300

aacgccatgt actccttctt cctggacttc gtggcgccgg acaaccaccg ctggaagtac 360

gtgaacgggg aatgggtgcc ggggggcaag ccggagccgc aggcgcccag ctgcgtctac 420

atccaccccg actcgcccaa cttcggggcc cactggatga aggtccccgt ctcttcagc 480

aaagtcaagc tcaccaacaa gctcaacgga gggggccaga tcattgctgaa ctcttgcatt 540

aagtatgagc ctgcaatcca catagtgaga gttgggggtc cacagcgcat gatcaccagc 600

cactgcttcc ctgagaccca gttcatagcg gtgactgctt atcagaacga ggagatcaca 660

gctcttaaaa ttaagtacaa tccatttgca aaggttttcc ttgatgcaaa ggaaagaagt 720

gatcacaag agatgatgga ggaacccgga gacagccagc aacctgggta ctcccaatgg 780

gggtggcttc ttcttggaac cagcaccgtt tgtccacctg caaatctca tctcagttt 840

ggaggtgccc tctccctccc ctccacgcac agctgtgaca ggtacccaac cctgaggagc 900

caccggtcct caccctaccc cagcccctat gctcatcgga acaattctcc aacctattct 960

gacaactcac ctgcatgttt atccatgctg caatcccatg acaattggtc cagccttgga 1020  
atgcctgccc atcccagcat gctccccgtg agccacaatg ccagcccacc taccagctcc 1080  
agtcagtacc ccagcctgtg gtctgtgagc aacggcgccg tcaccccggg ctcccaggca 1140  
gcagccgtgt ccaacgggct gggggcccag ttcttcggg gctccccgc gcactacaca 1200  
ccctcacc atccggtctc ggcgcctct tcctcgggat cccactgta cgaaggggcg 1260  
gccgcggcca cagacatcgt ggacagccag tacgacgccg cagcccaagg ccgcctcata 1320

gcctcatgga cactgtgtc gccacattcc atgtga 1356

<210> 38

<211> 452

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by  
SEQ ID NO:36

<400> 38

Met Asn Met Ser Lys Lys Gln Val Lys Thr Gln Ser Lys Cys Asn Asn

1 5 10 15

Ile Met Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr

20 25 30

Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly

35 40 45

Ser Glu Lys Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly Leu Glu

50 55 60

Glu Ser Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile

65 70 75 80

Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val

85 90 95

Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val

100 105 110

Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro

115 120 125

Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro

130 135 140

Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe  
145 150 155 160

Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met  
165 170 175

Leu Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val  
180 185 190

Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln  
195 200 205

Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys  
210 215 220

Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg  
225 230 235 240

Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro  
245 250 255

Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Val Cys  
260 265 270

Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro  
275 280 285

Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser  
290 295 300

Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr  
305 310 315 320

Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn  
325 330 335

Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser  
340 345 350

His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp  
355 360 365

Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val  
370 375 380

Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr

385                      390                      395                      400  
 Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro  
                                  405                      410                      415

Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr  
                                  420                      425                      430  
 Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser  
                                  435                      440                      445

Pro Pro Ser Met  
                                  450

<210> 39

<211> 451

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by  
 SEQ ID NO:37

<400> 39

Met Asn Met Ser Lys Lys Gln Val Lys Thr Gln Ser Lys Cys Asn Asn

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

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

Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp

130 135 140

Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser

145 150 155 160

Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met Leu

165 170 175

Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val Gly

180 185 190

Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln Phe

195 200 205

Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile

210 215 220

Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser

225 230 235 240

Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly

245 250 255

Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Val Cys Pro

260 265 270

Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser

275 280 285

Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser

290 295 300

Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser

305 310 315 320

Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp

325 330 335

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

340 345 350

Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser

355 360 365

Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser

370 375 380  
 Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr

385 390 395 400  
 Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu

405 410 415  
 Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp

420 425 430  
 Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro

435 440 445  
 Pro Ser Met

450  
 <210> 40  
 <211> 1353

<212> DNA  
 <213> Artificial Sequence  
 <220><223> Coding sequence encoding Fusion protein comprising Brachyury  
 (L254V) Protein plus addition of 42 nucleotides from Gene FPV 088  
 and an atc

<400> 40

atgagtaaga aacaagtaaa aactcaaagt aaatgtaata atatcatgag ctccctggc 60  
 accgagagcg cgaggaaagag cctgcagtac cgagtggacc acctgctgag cgccgtggag 120  
 aatgagctgc aggcgggcag cgagaagggc gacccacag agcggaact gcgcgtgggc 180  
 ctggaggaga gcgagctgtg gctgcgcttc aaggagctca ccaatgagat gatcgtgacc 240

aagaacggca ggaggatgtt tccggtgctg aaggtgaacg tgtctggcct ggacccaac 300  
 gccatgtact ctttctgct ggacttcgtg gcggcggaca accaccgtg gaagtacgtg 360  
 aacggggaat gggtgccggg gggcaagccg gagccgcagg cgcccagctg cgtctacatc 420  
 caccgcgact cgccaactt cggggccac tggatgaagg ctccgtctc cttcagcaaa 480  
 gtcaagctca ccaacaagct caacggaggg ggccagatca tgctgaactc cttgcataag 540  
 tatgagctc gaatccacat agtgagagtt ggggtccac agcgcatgat caccagccac 600  
 tgcttcctg agaccagtt catagcgtg actgcttate agaacgagga gatcacagct 660

cttaaaatta agtacaatcc atttgcaaag gctttccttg atgcaaagga aagaagtgat 720

cacaaagaga tgatggagga acccgagac agccagcaac ctgggtactc ccaatggggg 780  
 tggctttctt ctggaaccag caccgtttgt ccacctgcaa atcctcatcc tcagtttga 840  
 ggtgccctct cctccccctc cagcacagc tgtgacaggt acccaaccct gaggagccac 900  
 cggctctcac cctaccccag cccctatgct catcggaaca attctccaac ctattctgac 960  
 aactcacctg catgtttatc catgctgcaa tcccatgaca attggtccag ccttggaatg 1020  
 cctgcccac cagcatgct ccccgtagc cacaatgcca gcccacctac cagctccagt 1080

cagtaccca gcctgtggtc tgtgagcaac ggccgctca ccccggtc ccaggcagca 1140  
 gccgtgtcca acgggctggg ggcccagttc ttccggggt ccccgcgca ctacacacc 1200  
 ctacccatc cgtctcggc gccctcttc tcgggatccc cactgtacga agggcgggc 1260  
 gcggccacag acatcgtgga cagccagtac gacgccgag cccaaggccg cctcatagcc 1320  
 tcatggacac ctgtgtcgcc acctccatg tga 1353

<210> 41

<211> 1350

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury

(L254V) Protein plus addition of 42 nucleotides from Gene FPV 088  
 and an atc, without Brachyury atg start codon

<400> 41

atgagtaaga aacaagtaaa aactcaaagt aaatgtaata atatcagctc ccctggcacc 60  
 gagagcgcgg gaaagagcct gcagtaccga gtggaccacc tgctgagcgc cgtggagaat 120  
 gagctgcagg cgggcagcga gaaggcgac cccacagagc gcgaactgcg cgtgggcctg 180  
 gaggagagcg agctgtggct gcgcttcaag gagctacca atgagatgat cgtgaccaag 240  
 aacggcagga ggatgtttcc ggtgctgaag gtgaacgtgt ctggcctgga cccaacgcc 300

atgtactcct tctgtctgga ctctgtggcg gcggacaacc accgtggaa gtacgtgaac 360  
 ggggaatggg tgccgggggg caagccggag ccgcaggcgc ccagctgcgt ctacatccac 420  
 cccgactcgc ccaacttcgg ggcccactgg atgaaggctc ccgtctcctt cagcaaagtc 480  
 aagctacca acaagctcaa cggagggggc cagatcatgc tgaactcctt gcataagtat 540  
 gagcctcgaa tccacatagt gagagttggg ggtccacagc gcatgatcac cagccactgc 600  
 ttccctgaga cccagttcat agcgtgact gcttatcaga acgaggagat cacagctctt 660  
 aaaattaagt acaatccatt tgcaaaggct ttccttgatg caaaggaaag aagtgatcac 720

aaagagatga tggaggaacc cggagacagc cagcaacctg ggtactccca atgggggtgg 780  
cttcttctcg gaaccagcac cgtttgtcca cctgcaaate ctcacctca gtttgaggt 840  
gccctctccc tcccctccac gcacagctgt gacaggtacc caaccctgag gagccaccgg 900  
tcctcacctt accccagccc ctatgctcat cggacaatt ctccaaccta ttctgacaac 960  
tcacctgcat gtttatccat gctgcaatcc catgacaatt ggtccagcct tggaatgcct 1020  
gcccaccca gcatgctccc cgtgagccac aatgccagcc cacctaccag ctccagtcag 1080  
taccacagcc tgtggtctgt gagcaacggc gccgtcacc cgggctccca ggcagcagcc 1140

gtgtccaacg ggctgggggc ccagttcttc cggggctccc ccgcgacta cacaccctc 1200  
acctatccgg tctcggcgcc ctcttctctg ggatcccccac tgtacgaagg ggcggccgcg 1260  
gccacagaca tcgtggacag ccagtacgac gccgcagccc aaggccgcct catagcctca 1320  
tggacacctg tgtcgccacc ttccatgtga 1350

<210> 42

<211> 450

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by

SEQ ID NO:40

<400> 42

Met Ser Lys Lys Gln Val Lys Thr Gln Ser Lys Cys Asn Asn Ile Met

1 5 10 15  
Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr Arg Val

20 25 30  
Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly Ser Glu

35 40 45  
Lys Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly Leu Glu Glu Ser

50 55 60  
Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile Val Thr

65 70 75 80  
Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val Ser Gly

85 90 95  
Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val Ala Ala

100 105 110

Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro Gly Gly  
 115 120 125  
 Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp Ser  
 130 135 140  
 Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser Lys  
 145 150 155 160  
 Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met Leu Asn  
 165 170 175  
 Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val Gly Gly  
 180 185 190  
 Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln Phe Ile  
 195 200 205  
 Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile Lys  
 210 215 220  
 Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser Asp  
 225 230 235 240  
 His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly Tyr  
 245 250 255  
 Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Val Cys Pro Pro  
 260 265 270  
 Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr  
 275 280 285  
 His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser Pro  
 290 295 300  
 Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser Asp  
 305 310 315 320  
 Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp Ser  
 325 330 335  
 Ser Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser His Asn  
 340 345 350  
 Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val

355                      360                      365  
 Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser Asn  
 370                      375                      380  
 Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr Pro  
  
 385                      390                      395                      400  
 Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr  
 405                      410                      415  
 Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala  
 420                      425                      430  
 Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro Pro  
 435                      440                      445  
 Ser Met  
 450  
 <210> 43  
 <211> 449  
 <212>  
 > PRT  
 <213> Artificial Sequence  
 <220><223> Fusion protein comprising Brachyury (L254V) Protein encoded by  
 SEQ ID NO:41  
 <400> 43  
 Met Ser Lys Lys Gln Val Lys Thr Gln Ser Lys Cys Asn Asn Ile Ser  
 1                      5                      10                      15  
 Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr Arg Val Asp  
 20                      25                      30  
 His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly Ser Glu Lys  
 35                      40                      45  
  
 Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly Leu Glu Glu Ser Glu  
 50                      55                      60  
 Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile Val Thr Lys  
 65                      70                      75                      80  
 Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val Ser Gly Leu  
 85                      90                      95

Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val Ala Ala Asp  
 100 105 110

Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro Gly Gly Lys  
 115 120 125

Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp Ser Pro  
 130 135 140

Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser Lys Val  
 145 150 155 160

Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met Leu Asn Ser  
 165 170 175

Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val Gly Gly Pro  
 180 185 190

Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln Phe Ile Ala  
 195 200 205

Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile Lys Tyr  
 210 215 220

Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser Asp His  
 225 230 235 240

Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly Tyr Ser  
 245 250 255

Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Val Cys Pro Pro Ala  
 260 265 270

Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr His  
 275 280 285

Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser Pro Tyr  
 290 295 300

Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser Asp Asn  
 305 310 315 320

Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp Ser Ser  
 325 330 335

Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser His Asn Ala

340 345 350  
 Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val Ser  
 355 360 365

Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser Asn Gly  
 370 375 380  
 Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr Pro Leu  
 385 390 395 400  
 Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr Glu  
 405 410 415  
 Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala Ala  
 420 425 430

Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro Pro Ser  
 435 440 445

Met

<210> 44

<211> 1323

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding fusion protein comprising Brachyury  
 protein Isoform 1 plus addition of 15 nucleotides from Gene FPV  
 088

<400> 44

|                                                                    |     |
|--------------------------------------------------------------------|-----|
| atgaaaaata acttgatgag ctccctggc accgagagcg cgggaaagag cctgcagtac   | 60  |
| cgagtggacc acctgctgag cgccgtggag aatgagctgc aggcgggcag cgagaagggc  | 120 |
|                                                                    |     |
| gacccacag agcgcgaact gcgcgtgggc ctggaggaga gcgagctgtg gctgcgttc    | 180 |
| aaggagctca ccaatgatgat gatcgtgacc aagaacggca ggaggatgtt tccggtgctg | 240 |
| aaggtgaacg tgictggcct ggacccaac gccatgtact ccttcctgct ggacttcgtg   | 300 |
| gcggcggaca accaccgctg gaagtacgtg aacggggaat gggtgccggg gggcaagccg  | 360 |
| gagccgcagg cgcccagctg cgtctacatc caccgcgact cgcccaactt cggggccac   | 420 |
| tggatgaagg ctcccgctc cttcagcaaa gtcaagctca ccaacaagct caacggaggg   | 480 |

ggccagatca tgctgaactc cttgcataag tatgagcctc gaatccacat agtgagagtt 540

gggggtccac agcgcgatgat caccagccac tgcttccctg agaccagtt catagcggtg 600

actgcttate agaacgagga gatcacagct cttaaaatta agtacaatcc atttgcaaaa 660

gctttccttg atgcaaagga aagaagtgat cacaagaga tgatggagga acccgagac 720

agccagaac ctgggtactc ccaatggggg tggttcttc ctggaaccag caccgtttgt 780

ccacctgcaa atcctcatcc tcagtttga ggtgccctct cctccctc cacgcacagc 840

tgtgacaggt acccaacct gaggagccac cgttcctcac cctaccag ccctatgct 900

catcggaaca attctcaac ctattctgac aactcacctg catgtttatc catgctgcaa 960

tcccatgaca attggtccag ccttgaatg cctgccatc ccagcatgct cccgtgagc 1020

cacaatgcca gccacctac cagctccagt cagtaccca gcctgtggtc tgtgagcaac 1080

ggcgccgtca cccgggctc ccaggcagca gccgtgtcca acgggctggg ggcccagttc 1140

ttccggggt ccccgcgca ctacacacc ctacccatc cgtctcggc gccctcttc 1200

tcgggatccc cactgtacga agggcgggc gcggccacag acatcgtgga cagccagtac 1260

gacgcgcag cccaaggccg cctcatagcc tcattggacac ctgtgtgcc accttccatg 1320

tga 1323

<210> 45

<211> 1320

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding fusion protein comprising Brachyury protein Isoform 1 plus addition of 15 nucleotides from Gene FPV 088 without Brachyury protein atg start codon

<400> 45

atgaaaaata acttgagctc cctggcacc gagagcgg gaaagagcct gcagtaccga 60

gtggaccacc tgctgagcgc cgtggagaat gagctgcagg cgggcagcga gaaggcgac 120

cccacagagc gcgaactgcg cgtgggcctg gaggagagcg agctgtggct gcgcttcaag 180

gagctacca atgagatgat cgtgaccaag aacggcagga ggatgtttcc ggtgctgaag 240

gtgaacgtgt ctggcctgga cccaacgcc atgtactcct tcctgctgga ctctgtggcg 300

gcggacaacc accgctgga gtacgtgaac gggaatggg tgccgggggg caagccggag 360

ccgcaggcgc ccagctgcgt ctacatccac ccgactcgc ccaacttcgg ggcccactgg 420

atgaaggctc ccgtctcctt cagcaaagtc aagctacca acaagctcaa cggagggggc 480

cagatcatgc tgaactcctt gcataagtat gagcctcgaa tccacatagt gagagtggg 540  
 ggtccacagc gcatgatcac cagccactgc ttcctgaga ccagttcat agcggtgact 600

gcttatcaga acgaggagat cacagctctt aaaattaagt acaatccatt tgcaaaagct 660  
 ttccttgatg caaaggaaag aagtgatcac aaagagatga tggaggaacc cggagacagc 720  
 cagcaacctg ggtactccca atgggggtgg ctcttctctg gaaccagcac cgtttgtcca 780  
 cctgcaaate ctcatctca gtttggaggt gccctctccc tcccctccac gcacagctgt 840  
 gagaggtacc caacctgag gagccaccgg tctcacct accccagccc ctatgctcat 900  
 cggacaatt ctccaaccta ttctgacaac tcacctgcat gtttatecat gctgcaatcc 960  
 catgacaatt ggtccagcct tggaatgcct gcccatccca gcatgctccc cgtgagccac 1020

aatgccagcc cacctaccag ctccagtcag taccacagcc tgtggtctgt gagcaacggc 1080  
 gccgtcacc cgggctccca ggcagcagcc gtgtccaacg ggctgggggc ccagttcttc 1140  
 cggggctccc ccgcgacta cacacccctc acccatccgg tctcggcgcc ctcttctctg 1200  
 ggatccccac tgtacgaagg ggcggcccg gccacagaca tcgtggacag ccagtacgac 1260  
 gccgcagccc aaggcgcct catagcctca tggacacctg tgtcgccacc ttccatgtga 1320

<210> 46

<211> 440

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by

SEQ ID NO:44

<400> 46

Met Lys Asn Asn Leu Met Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys

1 5 10 15

Ser Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu

20 25 30

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

35 40 45

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

50 55 60

Asn Glu Met Ile Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu

65 70 75 80

Lys Val Asn Val Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu  
 85 90 95  
 Leu Asp Phe Val Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly  
 100 105 110  
 Glu Trp Val Pro Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val  
 115 120 125  
  
 Tyr Ile His Pro Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala  
 130 135 140  
 Pro Val Ser Phe Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly  
 145 150 155 160  
 Gly Gln Ile Met Leu Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His  
 165 170 175  
 Ile Val Arg Val Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe  
 180 185 190  
  
 Pro Glu Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile  
 195 200 205  
 Thr Ala Leu Lys Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp  
 210 215 220  
 Ala Lys Glu Arg Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp  
 225 230 235 240  
 Ser Gln Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr  
 245 250 255  
  
 Ser Thr Leu Cys Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala  
 260 265 270  
 Leu Ser Leu Pro Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg  
 275 280 285  
 Ser His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn  
 290 295 300  
 Ser Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln  
 305 310 315 320  
  
 Ser His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met

325 330 335  
 Leu Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr  
 340 345 350  
 Pro Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln  
 355 360 365  
 Ala Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser  
 370 375 380

Pro Ala His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser  
 385 390 395 400  
 Ser Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val  
 405 410 415  
 Asp Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp  
 420 425 430  
 Thr Pro Val Ser Pro Pro Ser Met  
 435 440

<210> 47

<211> 439

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by  
 SEQ ID NO:45

<400> 47

Met Lys Asn Asn Leu Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser  
 1 5 10 15  
 Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu  
 20 25 30  
 Gln Ala Gly Ser Glu Lys Gly Asp Pro Thr Glu Arg Glu Leu Arg Val  
 35 40 45  
 Gly Leu Glu Glu Ser Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn  
 50 55 60  
 Glu Met Ile Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys  
 65 70 75 80

Val Asn Val Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu  
85 90 95  
Asp Phe Val Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu  
100 105 110  
Trp Val Pro Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr  
115 120 125  
Ile His Pro Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro  
130 135 140  
Val Ser Phe Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly  
145 150 155 160  
Gln Ile Met Leu Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile  
165 170 175  
Val Arg Val Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro  
180 185 190  
Glu Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr  
195 200 205  
Ala Leu Lys Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala  
210 215 220  
Lys Glu Arg Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser  
225 230 235 240  
Gln Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser  
245 250 255  
Thr Leu Cys Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu  
260 265 270  
Ser Leu Pro Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser  
275 280 285  
His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser  
290 295 300  
Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser  
305 310 315 320  
His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met Leu

325                      330                      335  
 Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro  
 340                      345                      350  
 Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala  
 355                      360                      365  
 Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro  
  
 370                      375                      380  
 Ala His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser  
 385                      390                      395                      400  
 Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp  
 405                      410                      415  
 Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr  
 420                      425                      430  
 Pro Val Ser Pro Pro Ser Met  
 435

<210> 48

<211> 1353

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding fusion protein comprising Brachyury  
 protein Isoform 1 plus addition of 45 nucleotides from Gene FPV  
 088

<400> 48

|                                                                     |     |
|---------------------------------------------------------------------|-----|
| atgaaaaata acttgtatga agaaaaaatg aacatgagta agaaaatgag ctcccctggc   | 60  |
| accgagagcg cgggaaagag cctgcagtac cgagtggacc acctgctgag cgccgtggag   | 120 |
| aatgagctgc aggcgggcag cgagaagggc gacccacag agcgcgaact gcgcgtgggc    | 180 |
| ctggaggaga gcgagctgtg gctgcgcttc aaggagctca ccaatgagat gatcgtgacc   | 240 |
|                                                                     |     |
| aagaacggca ggaggatgtt tccggtgctg aaggtgaacg tgtctggcct ggaccecaac   | 300 |
| gccatgtact ccttcctgct ggacttcgtg gcggcggaca accaccgctg gaagtacgtg   | 360 |
| aacggggaat gggtgccggg gggcaagccg gagccgcagg cgcccagctg cgtctacatc   | 420 |
| caccccgact cgccaactt cggggcccccac tggatgaagg ctcccgctctc cttcagcaaa | 480 |
| gtcaagctca ccaacaagct caacggaggg ggccagatca tgctgaactc cttgcataag   | 540 |

tatgagcctc gaatccacat agtgagagtt ggggggtccac agcgcatgat caccagccac 600  
 tgcttccctg agaccagtt catagcgggtg actgcttatac agaacgagga gatcacagct 660

cttaaaatta agtacaatcc atttgcaaag gctttccttg atgcaaagga aagaagtgat 720  
 cacaagaga tgatggagga acccgagac agccagcaac ctgggtactc ccaatggggg 780  
 tggtctcttc ctggaaccag caccgtttgt ccacctgcaa atcctcatcc tcagtttga 840  
 ggtgccctct cctccctc cagcacagc tgtgacaggt acccaaccct gaggagccac 900  
 cggctctcac cctacccag cccctatgct catcggaaca attctcaac ctattctgac 960  
 aactcacctg catgtttata catgctgcaa tcccatgaca attggtccag ccttggaatg 1020  
 cctgcccatac ccagcatgct ccccgtagc cacaatgcca gccacctac cagctccagt 1080

cagtaccca gcctgtggtc tgtgagcaac ggcgccgtca ccccggtc ccaggcagca 1140  
 gccgtgtcca acgggctggg ggcccagttc ttccggggct ccccgcgca ctacacacc 1200  
 ctcaccatac cggctctggc gccctcttc tcgggatccc cactgtacga aggggcggcc 1260  
 gcggccacag acatcgtgga cagccagtac gacgccgag cccaaggccg cctcatagcc 1320  
 tcatggacac ctgtgtcgcc accttccatg tga 1353

<210> 49  
 <211> 1350  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Coding sequence encoding fusion protein comprising Brachyury

protein Isoform 1 plus addition of 45 nucleotides from Gene FPV  
 088 without Brachyury atg

<400> 49

atgaaaaata acttgatatg agaaaaaatg aacatgagta agaaaagctc cctggcacc 60  
 gagagcgcg gaaagagcct gcagtaccga gtggaccacc tgctgagcgc cgtggagaat 120  
 gagctgcagg cgggcagcga gaaggcgac cccacagagc gcgaactgcg cgtgggcctg 180  
 gaggagagcg agctgtggct gcgcttcaag gagctacca atgagatgat cgtgaccaag 240  
 aacggcagga ggatgtttcc ggtgctgaag gtgaacgtgt ctggcctgga cccaacgcc 300  
 atgtactcct tctgtctgga ctctgtggcg gcggacaacc accgtggaa gtactgaac 360

ggggaatggg tgccgggggg caagccggag ccgcagcgcc ccagctgcgt ctacatccac 420  
 cccgactcgc ccaacttcgg ggccactgg atgaaggctc ccgtctcctt cagcaaagtc 480  
 aagctacca acaagctcaa cggagggggc cagatcatgc tgaactcctt gcataagtat 540

gagcctcgaa tccacatagt gagagttggg ggtccacagc gcatgacac cagccactgc 600  
 ttccctgaga cccagttcat agcggtagct gcttatcaga acgaggagat cacagctctt 660  
 aaaattaagt acaatccatt tgcaaaagct ttccttgatg caaaggaaag aagtgatcac 720  
 aaagagatga tggaggaacc cggagacagc cagcaacctg ggtactccca atgggggtgg 780

cttcttctg gaaccagcac cgtttgcca cctgcaaate ctcatectca gtttgaggt 840  
 gccctctccc tcccctccac gcacagctgt gacaggtacc caaccctgag gagccaccgg 900  
 tcctcacctt accccagccc ctatgctcat cggacaatt ctccaaccta ttctgacaac 960  
 tcacgtgcat gtttatccat gctgcaatcc catgacaatt ggtccagcct tggaatgcct 1020  
 gcccatccca gcatgctccc cgtgagccac aatgccagcc cacctaccag ctccagtcag 1080  
 taccacagcc tgtggtctgt gagcaacggc gccgtcacc cgggctccca ggcagcagcc 1140  
 gtgtccaacg ggctgggggc ccagttcttc cggggctccc ccgcgcacta cacaccctc 1200

acccatccgg tctcggcgcc ctcttcctcg ggatccccac tgtacgaagg ggcggccgcg 1260  
 gccacagaca tcgtggacag ccagtacgac gccgcagccc aaggccgcct catagcctca 1320  
 tggacacctg tgtegccacc ttccatgtga 1350

<210> 50

<211> 450

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by  
 SEQ ID NO:48

<400> 50

Met Lys Asn Asn Leu Tyr Glu Glu Lys Met Asn Met Ser Lys Lys Met

1 5 10 15

Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr Arg Val

20 25 30

Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly Ser Glu

35 40 45

Lys Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly Leu Glu Glu Ser

50 55 60

Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile Val Thr

65 70 75 80

Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val Ser Gly  
 85 90 95  
 Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val Ala Ala  
 100 105 110  
 Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro Gly Gly  
 115 120 125  
 Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp Ser  
 130 135 140  
  
 Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser Lys  
 145 150 155 160  
 Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met Leu Asn  
 165 170 175  
 Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val Gly Gly  
 180 185 190  
 Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln Phe Ile  
 195 200 205  
  
 Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile Lys  
 210 215 220  
 Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser Asp  
 225 230 235 240  
 His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly Tyr  
 245 250 255  
 Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Leu Cys Pro Pro  
 260 265 270  
  
 Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr  
 275 280 285  
 His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser Pro  
 290 295 300  
 Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser Asp  
 305 310 315 320  
 Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp Ser

325 330 335

Ser Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser His Asn

340 345 350

Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val

355 360 365

Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser Asn

370 375 380

Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr Pro

385 390 395 400

Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr

405 410 415

Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala

420 425 430

Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro Pro

435 440 445

Ser Met

450

<210> 51

<211> 449

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by

SEQ ID NO:49

<400> 51

Met Lys Asn Asn Leu Tyr Glu Glu Lys Met Asn Met Ser Lys Lys Ser

1 5 10 15

Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr Arg Val Asp

20 25 30

His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly Ser Glu Lys

35 40 45

Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly Leu Glu Glu Ser Glu

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
| 50                                                              | 55  | 60  |     |
| Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile Val Thr Lys |     |     |     |
| 65                                                              | 70  | 75  | 80  |
| Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val Ser Gly Leu |     |     |     |
|                                                                 | 85  | 90  | 95  |
| Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val Ala Ala Asp |     |     |     |
|                                                                 | 100 | 105 | 110 |
| Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro Gly Gly Lys |     |     |     |
|                                                                 | 115 | 120 | 125 |
|                                                                 |     |     |     |
| Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp Ser Pro |     |     |     |
|                                                                 | 130 | 135 | 140 |
| Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser Lys Val |     |     |     |
| 145                                                             | 150 | 155 | 160 |
| Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met Leu Asn Ser |     |     |     |
|                                                                 | 165 | 170 | 175 |
| Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val Gly Gly Pro |     |     |     |
|                                                                 | 180 | 185 | 190 |
|                                                                 |     |     |     |
| Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln Phe Ile Ala |     |     |     |
|                                                                 | 195 | 200 | 205 |
| Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile Lys Tyr |     |     |     |
|                                                                 | 210 | 215 | 220 |
| Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser Asp His |     |     |     |
| 225                                                             | 230 | 235 | 240 |
| Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly Tyr Ser |     |     |     |
|                                                                 | 245 | 250 | 255 |
|                                                                 |     |     |     |
| Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Leu Cys Pro Pro Ala |     |     |     |
|                                                                 | 260 | 265 | 270 |
| Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr His |     |     |     |
|                                                                 | 275 | 280 | 285 |
| Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser Pro Tyr |     |     |     |
|                                                                 | 290 | 295 | 300 |

Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser Asp Asn  
305 310 315 320

Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp Ser Ser  
325 330 335

Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser His Asn Ala  
340 345 350

Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val Ser  
355 360 365

Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser Asn Gly  
370 375 380

Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr Pro Leu  
385 390 395 400

Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr Glu  
405 410 415

Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala Ala  
420 425 430

Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro Pro Ser  
435 440 445

Met

<210> 52

<211> 1326

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding fusion protein comprising Brachyury  
protein Isoform 1 plus addition of 18 nucleotides from Gene FPV  
088

<400> 52

atgaacatga gtaagaaaat gagctcccct ggcaccgaga ggcgaggaaa gaggctgcag 60  
taccgagtgg accacctgct gagcgccgtg gagaatgagc tgcaggcggg cagcgagaag 120  
ggcgacccca cagagcgca actgcgcgtg ggcctggagg agagcgagct gtggctgcgc 180

ttcaaggagc tcaccaatga gatgatcgtg accaagaacg gcaggaggat gtttccggtg 240

ctgaaggatga acgtgtctgg cctggacccc aacgcatgt actccttctt gctggacttc 300

gtggcggcgg acaaccaccg ctggaagtac gtgaacgggg aatgggtgcc ggggggcaag 360

ccggagccgc aggcgcccag ctgcgtctac atccacccc actcgcccaa cttcggggcc 420

cactggatga aggtccctt ctccttcagc aaagtcaagc tcaccaacaa gctcaacgga 480

gggggcccaga tcatgtgaa ctccttgcat aagtatgagc ctggaatcca catagtgaga 540

gttgggggtc cacagcgcat gatcaccagc cactgcttcc ctgagacca gtcatagcg 600

gtgactgctt atcagaacga ggagatcaca gctcttaaaa ttaagtacaa tccatttgca 660

aaggctttcc ttgatgcaaa ggaaagaagt gatcacaagc agatgatgga ggaaccgga 720

gacagccagc aacctgggta ctccaatgg gggtggcttc ttcttggaa cagcaccgtt 780

tgtccacctg caaatctca tcttcagttt ggaggtgccc tctccctccc ctccacgac 840

agctgtgaca ggiaccaac cctgaggagc caccggtcct caccctacc cagccctat 900

gctcatcgga acaattctcc aacctattct gacaactcac ctgcatgttt atccatgctg 960

caatcccatg acaattggc cagccttggga atgcctgccc atccagcat gctccccgtg 1020

agccacaatg ccagcccacc taccagctcc agtcagtacc ccagcctgtg gtctgtgagc 1080

aacggcgccg tcaccccggg ctcccaggca gcagccgtgt ccaacgggct gggggcccag 1140

ttcttccggg gctccccgc gcactacaca cccctaccc atccggtctc ggcgccctct 1200

tcctcgggat cccactgta cgaaggggcg gccgcggcca cagacatcgt ggacagccag 1260

tacgagccg cagcccaagg ccgcctcata gcctcatgga cactgtgtc gccaccttc 1320

atgtga 1326

<210> 53

<211> 1323

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding fusion protein comprising Brachyury

protein Isoform 1 plus addition of 18 nucleotides from Gene FPV

088 without brachyury atg

<400> 53

atgaacatga gtaagaaaag ctcccctggc accgagagcg cgggaaagag cctgcagtac 60

cgagtggacc acctgctgag gcgcgtggag aatgagctgc aggcgggcag cgagaagggc 120

gacccacag agcggaact gcgcgtgggc ctggaggaga gcgagctgtg gctgcgcttc 180

aaggagctca ccaatgagat gatcgtgacc aagaacggca ggaggatgtt tccggtgctg 240  
aaggtgaacg tgtctggcct ggaccccaac gccatgtact ccttcctgct ggacttcgtg 300  
gcggcggaca accaccgctg gaagtacgtg aacggggaat gggtgccggg gggcaagccg 360

gagccgcagg cgcccagctg cgtctacatc caccocgact cgcccaactt cggggccac 420  
tggatgaagg ctcccgctc cttcagcaaa gtcaagctca ccaacaagct caacggaggg 480  
ggccagatca tgttgaactc cttgcataag tatgagcctc gaatccacat agtgagagtt 540  
gggggtccac agcgcattgat caccagccac tgcttccttg agaccagtt catagcggtg 600  
actgcttacc agaacgagga gatcacagct cttaaaatta agtacaatcc atttgcaaaa 660  
gctttccttg atgcaaagga aagaagtgat cacaagaga tgatggagga acccgagac 720  
agccagcaac ctgggtactc ccaatggggg tggtttcttc ctggaaccag caccgtttgt 780

ccacctgcaa atctcatcc tcagtttga ggtgccctct cctccctc cagcacagc 840  
tgtgacaggt acccaacct gaggagccac cgttcctcac cctacccag ccctatgct 900  
catcggaaca attctcaac ctattctgac aactcacctg catgtttacc catgctgcaa 960  
tcccatgaca attggtccag ccttggaatg cctgccacc cagcatgct ccccgtagc 1020  
cacaatgcca gccacctac cagctccagt cagtaccca gccgtggtc tgtgagcaac 1080  
ggcgccgtca ccccggtc ccaggcagca gccgtgtcca acgggctggg ggcccagttc 1140  
ttccggggt ccccgcgca ctacacacc ctacccacc cggctcggc gccctcttc 1200

tcgggatccc cactgtacga agggcgggc gcggccacag acatcgtga cagccagtac 1260  
gacgcgcag cccaagccg cctcatagcc tcatggacac ctgtgtgcc accttccatg 1320  
tga 1323

<210> 54

<211> 441

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by  
SEQ ID NO:52

<400> 54

Met Asn Met Ser Lys Lys Met Ser Ser Pro Gly Thr Glu Ser Ala Gly

1 5 10 15

Lys Ser Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn

20 25 30

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

275                      280                      285  
 Arg Ser His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn  
 290                      295                      300  
 Asn Ser Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu  
 305                      310                      315                      320  
 Gln Ser His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser  
                          325                      330                      335  
  
 Met Leu Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln  
                          340                      345                      350  
 Tyr Pro Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser  
                          355                      360                      365  
 Gln Ala Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly  
                          370                      375                      380  
 Ser Pro Ala His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser  
 385                      390                      395                      400

Ser Ser Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile  
                          405                      410                      415  
 Val Asp Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser  
                          420                      425                      430  
 Trp Thr Pro Val Ser Pro Pro Ser Met  
                          435                      440

<210> 55

<211> 440

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by  
 SEQ ID NO:53

<400

> 55

Met Asn Met Ser Lys Lys Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys  
 1                      5                      10                      15  
 Ser Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu  
                          20                      25                      30

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

275 280 285  
 Ser His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn  
 290 295 300  
 Ser Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln  
 305 310 315 320

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

Pro Ala His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser  
 385 390 395 400  
 Ser Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val  
 405 410 415  
 Asp Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp  
 420 425 430  
 Thr Pro Val Ser Pro Pro Ser Met  
 435 440

<210> 56

<211> 1320

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury  
 protein Isoform 1 plus addition of 12 nucleotides from Gene FPV  
 088

<400> 56

atgagtaaga aaatgagctc ccctggcacc gagagcgagg gaaagagcct gcagtaccga 60  
 gtggaccacc tgctgagcgc cgtggagaat gagctgcagg cgggcagcga gaaggcgac 120  
 cccacagagc gcgaactgcg cgtgggcctg gaggagagcg agctgtggct gcgcttcaag 180

gagctcacca atgagatgat cgtgaccaag aacggcagga ggatgtttcc ggtgctgaag 240  
 gtgaacgtgt ctggcctgga cccaacgcc atgtactcct tcctgctgga cttcgtggcg 300

gcggacaacc accgctggaa gtacgtgaac ggggaatggg tgccgggggg caagccggag 360  
 ccgcaggcgc ccagctgcgt ctacatccac cccgactcgc ccaacttcgg ggcccactgg 420  
 atgaaggctc ccgtctcctt cagcaaagtc aagctcacca acaagctcaa cggagggggc 480  
 cagatcatgc tgaactcctt gcataagtat gagcctcgaa tccacatagt gagagtggg 540  
 ggtccacagc gcatgatcac cagccactgc ttccttgaga ccagttcat agcggtgact 600  
 gcttatcaga acgaggagat cacagctctt aaaattaagt acaatccatt tgcaaaagct 660  
 ttccttgatg caaaggaaag aagtgatcac aaagagatga tggaggaacc cggagacagc 720

cagcaacctg ggtactccca atgggggtgg ctctctcctg gaaccagcac cgtttgtcca 780  
 cctgcaaate ctcatcctca gtttggaggt gccctctccc tcccctccac gcacagctgt 840  
 gacaggtacc caacctgag gagccaccgg tcctcacct accccagccc ctatgctcat 900  
 cggaacaatt ctccaaccta ttctgacaac tcacctgcat gtttatccat gctgcaatcc 960  
 catgacaatt ggtccagcct tggaatgcct gcccatccca gcatgctccc cgtgagccac 1020  
 aatgccagcc cacctaccag ctccagtcag taccagcc tgtggtctgt gagcaacggc 1080  
 gccgtcacc cgggctccca ggcagcagcc gtgtccaacg ggctgggggc ccagttcttc 1140

cggggtctcc ccgcgacta cacaccctc acccatccgg tctcggcgc ctcttcctcg 1200  
 ggatccccac tgtacgaagg ggcggccgc gccacagaca tcgtggacag ccagtacgac 1260  
 gccgcagccc aaggccgct catagcctca tggacacctg tgtcgccacc ttccatgtga 1320

<210> 57  
 <211> 1317  
 <212> DNA  
 <213> Artificial Sequence  
 <220><223> Coding sequence encoding Fusion protein comprising Brachyury  
 protein Isoform 1 plus addition of 12 nucleotides from Gene FPV  
 088 without Brachyury atg start codon  
 <400> 57

atgagtaaga aaagctcccc tggcaccgag agcgcgggaa agagcctgca gtaccgagtg 60

gaccacctgc tgagcgccgt ggagaatgag ctgcaggcgg gcagcgagaa gggcgacccc 120  
 acagagcgcg aactgcgcgt gggcctggag gagagcgagc tgtggctgcg cttcaaggag 180  
 ctaccaatg agatgatcgt gaccaagaac ggcaggagga tgtttccggt gctgaaggtg 240

aacgtgtctg gcctggaccc caacgccatg tactccttcc tgctggactt cgtggcggcg 300  
gacaaccacc gctggaagta cgtgaacggg gaatgggtgc cggggggcaa gccggagccg 360  
caggcgccca gctgcgtcta catccacccc gactcgccca acttcggggc cactgggatg 420  
aaggctcccg tctccttcag caaagtcaag ctaccaaca agctcaacgg agggggccag 480

atcatgctga actccttgca taagtatgag cctcgaatcc acatagttag agttgggggt 540  
ccacagcgca tgatcaccag cactgcttc cctgagaccc agttcatagc ggtgactgct 600  
tatcagaacg aggagatcac agctcttaaa attaagtaca atccatttgc aaaagctttc 660  
cttgatgcaa aggaagaag tgatcacaaa gagatgatgg aggaacccgg agacagccag 720  
caacctgggt actcccaatg ggggtggctt cttcctggaa ccagcaccgt ttgtccacct 780  
gcaaatcctc atcctcagtt tggaggtgcc ctctccctcc cctccacgca cagctgtgac 840  
aggtacccaa cctgaggag ccaccggtcc tcaccctacc ccagccccta tgctcatcgg 900

aacaattctc caacctattc tgacaactca cctgcatggt tatccatgct gcaatcccat 960  
gacaattggt ccagccttgg aatgcctgcc catcccagca tgctccccgt gagccacaat 1020  
gccagccca ctagcagctc cagtcatgac ccagcctgt ggtctgtgag caacggcgcc 1080  
gtcaccccg gctcccagc agcagccgtg tccaacgggc tggggggcca gttcttcggg 1140  
ggctccccg cgcactacac accctcacc catccggtct cggcgccctc ttctcggga 1200  
tccccactgt acgaaggggc ggccgcgcc acagacatcg tggacagcca gtacgagcc 1260  
gcagcccaag gccgcctcat agcctcatgg acacctgtgt cgccacctc catgtga 1317

<210> 58

<211> 439

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by  
SEQ ID NO: 56

<400> 58

Met Ser Lys Lys Met Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser

1 5 10 15

Leu Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu

20 25 30

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

35

40

45

Gly Leu Glu Glu Ser Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn  
 50 55 60  
 Glu Met Ile Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys  
 65 70 75 80  
 Val Asn Val Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu  
 85 90 95  
 Asp Phe Val Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu  
 100 105 110  
 Trp Val Pro Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr  
 115 120 125  
 Ile His Pro Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro  
 130 135 140  
 Val Ser Phe Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly  
 145 150 155 160  
 Gln Ile Met Leu Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile  
 165 170 175  
 Val Arg Val Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro  
 180 185 190  
 Glu Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr  
 195 200 205  
 Ala Leu Lys Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala  
 210 215 220  
 Lys Glu Arg Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser  
 225 230 235 240  
 Gln Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser  
 245 250 255  
 Thr Leu Cys Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu  
 260 265 270  
 Ser Leu Pro Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser  
 275 280 285  
 His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser

290 295 300  
 Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser  
 305 310 315 320  
 His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met Leu  
 325 330 335  
 Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro  
 340 345 350  
 Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala

355 360 365  
 Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro  
 370 375 380  
 Ala His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser  
 385 390 395 400  
 Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp  
 405 410 415  
 Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr

420 425 430  
 Pro Val Ser Pro Pro Ser Met

435

<210> 59

<211> 438

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by  
 SEQ ID NO:57

<400> 59

Met Ser Lys Lys Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu  
 1 5 10 15  
 Gln Tyr Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln  
 20 25 30

Ala Gly Ser Glu Lys Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly  
 35 40 45

Leu Glu Glu Ser Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu  
 50 55 60  
 Met Ile Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val  
 65 70 75 80  
 Asn Val Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp  
 85 90 95  
  
 Phe Val Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp  
 100 105 110  
 Val Pro Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile  
 115 120 125  
 His Pro Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro Val  
 130 135 140  
 Ser Phe Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln  
 145 150 155 160  
  
 Ile Met Leu Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile Val  
 165 170 175  
 Arg Val Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu  
 180 185 190  
 Thr Gln Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala  
 195 200 205  
 Leu Lys Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys  
 210 215 220  
  
 Glu Arg Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln  
 225 230 235 240  
 Gln Pro Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr  
 245 250 255  
 Leu Cys Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser  
 260 265 270  
 Leu Pro Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His  
 275 280 285  
  
 Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro

290                      295                      300  
 Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His  
 305                      310                      315                      320  
 Asp Asn Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met Leu Pro  
                          325                      330                      335  
 Val Ser His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser  
                          340                      345                      350  
  
 Leu Trp Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala  
                          355                      360                      365  
 Ala Val Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala  
                          370                      375                      380  
 His Tyr Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly  
 385                      390                      395                      400  
 Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser  
                          405                      410                      415  
  
 Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro  
                          420                      425                      430  
 Val Ser Pro Pro Ser Met  
                          435

<210> 60

<211> 1386

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury  
 protein Isoform 1 plus addition of 75 nucleotides from Gene FPV  
 088 and an atc

<400> 60

atgaaaaata acttgtatga agaaaaaatg aacatgagta agaaacaagt aaaaactcaa      60  
  
 agtaaatgta ataatatcat gagctcccct ggcaccgaga gcgcgggaaa gaggctgcag      120  
 taccgagtgg accacctgct gagcgccgtg gagaatgagc tgcaggcgagg cagcgagaag      180  
 ggcgacccca cagagcgcca actgcgcgtg ggcctggagg agagcgagct gtggctgcgc      240  
 ttcaaggagc tcaccaatga gatgatcgtg accaagaacg gcaggaggat gtttcgggtg      300

ctgaagggtga acgtgtctgg cctggacccc aacgccatgt actccttcct gctggacttc 360  
gtggcgccgg acaaccaccg ctggaagtac gtgaacgggg aatgggtgcc ggggggcaag 420  
ccggagccgc aggcgcccag ctgcgtctac atccaccccg actcgcccaa cttcggggcc 480

cactggatga aggtctccgt ctcttcagc aaagtcaagc tcaccaacaa gctcaacgga 540  
gggggccaga tcatgctgaa ctcttgcat aagtatgagc ctggaatcca catagtgaga 600  
gttgggggtc cacagcgcat gatcaccagc cactgcttcc ctgagacca gttcatagcg 660  
gtgactgctt atcagaacga ggagatcaca gctcttaaaa ttaagtacaa tccatttgca 720  
aaagctttcc ttgatgcaaa ggaaagaagt gatcacaaag agatgatgga ggaacccgga 780  
gacagccagc aacctgggta ctccaatgg ggggtggctt ttcttggaac cagcacctgt 840  
tgtccacctg caaatcctca tcttcagttt ggaggtgccc tctccctccc ctccacgcac 900

agctgtgaca ggtacceaac cctgaggagc caccgttctt caccctaccc cagcccttat 960  
gctcatcgga acaattctcc aacctattct gacaactcac ctgcatgttt atccatgctg 1020  
caatcccatg acaattggtc cagccttgga atgctgccc atcccagcat gctccccgtg 1080  
agccacaatg ccagcccacc taccagctcc agtcagtacc ccagcctgtg gtctgtgagc 1140  
aacggcgccg tcaccccggt ctcccaggca gcagccgtgt ccaacgggct gggggcccag 1200  
ttttccggg gtcccccg gcactacaca cccctaccc atccggtctc ggcgccctct 1260  
tctcgggat cccactgta cgaagggcg gccgcggcca cagacatcgt ggacagccag 1320

tacgagccg cagcccaagg ccgctcata gcctcatgga caccgtgtc gccaccttc 1380  
atgtga 1386

<210> 61  
<211> 1383  
<212> DNA  
<213> Artificial Sequence  
<220><223> Coding sequence encoding Fusion protein comprising Brachyury  
protein Isoform 1 plus addition of 75 nucleotides from Gene FPV  
088 and an atc, without Brachyury atg start codon  
<400> 61  
atgaaaaata acitgtatga agaaaaatg aacatgagta agaaacaagt aaaaactcaa 60  
agtaaatgta ataatatcag ctcccctggc accgagagcg cgggaaagag cctgcagtac 120  
cgagtggacc acctgctgag gcgcgtggag aatgagctgc aggcgggcag cgagaagggc 180  
gacccacag agcggaact gcgcgtgggc ctggaggaga gcgagctgtg gctgcgcttc 240

aaggagctca ccaatgagat gatcgtgacc aagaacggca ggaggatgtt tccggtgctg 300  
aaggtgaacg tgtctggcct ggaccccaac gccatgtact ccttcctgct ggacttcgtg 360  
gcggcggaca accaccgctg gaagtacgtg aacggggaat ggggtgccggg gggcaagccg 420  
gagccgcagg cgcccagctg cgtctacatc caccgccact cgcccaactt cggggcccac 480

tggatgaagg ctcccgtctc cttcagcaaa gtcaagctca ccaacaagct caacggaggg 540  
ggccagatca tgttgaactc cttgcataag tatgagcctc gaatccacat agtgagagtt 600  
gggggtccac agcgcatgat caccagccac tgcttccttg agaccagtt catagcggtg 660  
actgcttacc agaacgagga gatcacagct cttaaaatta agtacaatcc atttgcaaaa 720  
gctttccttg atgcaaagga aagaagtgat cacaagaga tgatggagga acccgagac 780  
agccagcaac ctgggtactc ccaatggggg tggtctcttc ctggaaccag caccgtttgt 840  
ccacctgcaa atcctcatcc tcagtttga ggtgccctct cctccctc cagcacagc 900

tgtgacaggt acccaacct gaggagccac cggtcctcac cctacccag cccctatgct 960  
catcggaaca attctcaac ctattctgac aactcacctg catgtttatc catgctgcaa 1020  
tcccatgaca attggtccag ccttggaatg cctgcccac ccagcatgct ccccgtagc 1080  
cacaatgcca gccacctac cagctccagt cagtaccca gccgtggtc tgtgagcaac 1140  
ggcgccgtca ccccggtc ccaggcagca gccgtgtcca acgggctggg ggcccagttc 1200  
ttccggggt ccccgcgca ctacacacc ctacccac cggctcggc gccctcttc 1260  
tcgggatccc cactgtacga aggggcggc gcggccacag acatcgtgga cagccagtac 1320

gacgcgcag cccaaggccg cctcatagcc tcattggacac ctgtgtgcc accttccatg 1380  
tga 1383

<210> 62

<211> 461

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by  
SEQ ID NO:60

<400> 62

Met Lys Asn Asn Leu Tyr Glu Glu Lys Met Asn Met Ser Lys Lys Gln

1 5 10 15

Val Lys Thr Gln Ser Lys Cys Asn Asn Ile Met Ser Ser Pro Gly Thr

20

25

30

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

275 280 285  
 Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr His Ser Cys Asp Arg  
 290 295 300  
 Tyr Pro Thr Leu Arg Ser His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr  
 305 310 315 320  
 Ala His Arg Asn Asn Ser Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys  
 325 330 335  
 Leu Ser Met Leu Gln Ser His Asp Asn Trp Ser Ser Leu Gly Met Pro

340 345 350  
 Ala His Pro Ser Met Leu Pro Val Ser His Asn Ala Ser Pro Pro Thr  
 355 360 365  
 Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val Ser Asn Gly Ala Val  
 370 375 380  
 Thr Pro Gly Ser Gln Ala Ala Ala Val Ser Asn Gly Leu Gly Ala Gln  
 385 390 395 400  
 Phe Phe Arg Gly Ser Pro Ala His Tyr Thr Pro Leu Thr His Pro Val

405 410 415  
 Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala  
 420 425 430  
 Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg  
 435 440 445  
 Leu Ile Ala Ser Trp Thr Pro Val Ser Pro Pro Ser Met  
 450 455 460

<210> 63

<211> 460

<212> PRT

<213> Artificial Sequence

<220

><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by  
 SEQ ID NO:61

<400> 63

Met Lys Asn Asn Leu Tyr Glu Glu Lys Met Asn Met Ser Lys Lys Gln  
 1 5 10 15

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

260 265 270  
Leu Pro Gly Thr Ser Thr Leu Cys Pro Pro Ala Asn Pro His Pro Gln

275 280 285  
Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr His Ser Cys Asp Arg Tyr

290 295 300  
Pro Thr Leu Arg Ser His Arg Ser Ser Pro Tyr Pro Ser Pro Tyr Ala

305 310 315 320  
His Arg Asn Asn Ser Pro Thr Tyr Ser Asp Asn Ser Pro Ala Cys Leu

325 330 335  
Ser Met Leu Gln Ser His Asp Asn Trp Ser Ser Leu Gly Met Pro Ala  
340 345 350

His Pro Ser Met Leu Pro Val Ser His Asn Ala Ser Pro Pro Thr Ser  
355 360 365

Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val Ser Asn Gly Ala Val Thr

370 375 380  
Pro Gly Ser Gln Ala Ala Ala Val Ser Asn Gly Leu Gly Ala Gln Phe

385 390 395 400  
Phe Arg Gly Ser Pro Ala His Tyr Thr Pro Leu Thr His Pro Val Ser

405 410 415  
Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr Glu Gly Ala Ala Ala Ala

420 425 430  
Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala Ala Ala Gln Gly Arg Leu

435 440 445  
Ile Ala Ser Trp Thr Pro Val Ser Pro Pro Ser Met

450 455 460

<210> 64

<211> 1359

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury  
protein Isoform 1 plus addition of 48 nucleotides from Gene FPV  
088 and an atc

<400> 64

|                                                                    |      |
|--------------------------------------------------------------------|------|
| atgaacatga gtaagaaaca agtaaaaact caaagtaaata gtaataatat catgagctcc | 60   |
| cctggcaccg agagcgcgagg aaagagcctg cagtaccgag tggaccacct gctgagcgcc | 120  |
|                                                                    |      |
| gtggagaatg agctgcaggc gggcagcgag aaggcgacc ccacagagcg cgaactgcgc   | 180  |
| gtgggccttg aggagagcga gctgtggctg cgcttcaagg agtcaccaa tgagatgac    | 240  |
| gtgaccaaga acggcaggag gatgtttccg gtgctgaagg tgaacgtgtc tggcctggac  | 300  |
| cccaacgcca tgiactcctt cctgctggac ttcgtggcgg cggacaacca ccgctggaag  | 360  |
| tacgtgaacg gggaatgggt gccggggggc aagccggagc cgcaggcgcc cagctgcgtc  | 420  |
| tacatccacc ccgactcgcc caacttcggg gccactgga tgaaggctcc cgtctccttc   | 480  |
| agcaaagtca agtcaccaa caagctcaac ggagggggcc agatcatgct gaactccttg   | 540  |
|                                                                    |      |
| cataagtatg agcctcgaat ccacatagtg agagttgggg gtccacagcg catgatcacc  | 600  |
| agccactgct tccttgagac ccagttcata gcggtgactg cttatcagaa cgaggagatc  | 660  |
| acagctctta aaattaagta caatccattt gcaaaagctt tccttgatgc aaaggaaaga  | 720  |
| agtatcaca aagagatgat ggaggaaccc ggagacagcc agcaacctgg gtactcccaa   | 780  |
| tgggggtggc ttcttccttg aaccagcacc gttgtccac ctgcaaatcc tcacctcag    | 840  |
| tttggagggt cctctccct cccctccacg cacagctgtg acaggtaacc aaccctgagg   | 900  |
| agccaccggt cctcaccta cccagcccc tatgtcacc ggaacaattc tccaacctat     | 960  |
|                                                                    |      |
| tctgacaact cacctgcatg tttatccatg ctgcaatccc atgacaattg gtccagcctt  | 1020 |
| ggaatgcctg cccatccag catgtcccc gtgagccaca atgccagccc acctaccagc    | 1080 |
| tccagtcagt accccagcct gtggctctgt agcaacggcg ccgtacccc gggtcccag    | 1140 |
| gcagcagccg tgtccaacgg gctgggggcc cagttcttcc ggggtcccc cgcgcactac   | 1200 |
| acacccctca cccatccgt ctcggcgccc tcttctctgg gatccccact gtacgaaggg   | 1260 |
| gcggcccgcg ccacagacat cgtggacagc cagtacgacg ccgcagcca aggccgcctc   | 1320 |
| atagcctcat ggacacctgt gtcgccacct tccatgtga                         | 1359 |

<210> 65

<211> 1356

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury protein Isoform 1 plus addition of 48 nucleotides from Gene FPV 088 and an atc, without Brachyury atg start codon

<400> 65

|                                                                   |      |
|-------------------------------------------------------------------|------|
| atgaacatga gtaagaaaca agtaaaaact caaagtaaata gtaataatat cagctccct | 60   |
| ggcaccgaga gcgcgggaaa gagcctgcag taccgagtgg accacctgct gagcgccgtg | 120  |
| gagaatgagc tgcaggcggg cagcgagaag ggcgaccca cagagcgca actgcgcgtg   | 180  |
|                                                                   |      |
| ggcctggagg agagcgagct gtggctgcgc ttcaaggagc tcaccaatga gatgatcgtg | 240  |
| accaagaacg gcaggaggat gtttcgggtg ctgaaggtga acgtgtctgg cctggacccc | 300  |
| aacgccatgt actccttct gctggacttc gtggcgccgg acaaccaccg ctggaagtac  | 360  |
| gtgaacgggg aatgggtgcc ggggggcaag ccggagccgc aggcgcccag ctgcgtctac | 420  |
| atccaccccg actcgccaa cttcggggcc cactggatga aggtccctg ctccttcagc   | 480  |
| aaagtcaagc tcaccaacaa gctcaacgga gggggccaga tcatgctgaa ctccttgcac | 540  |
| aagtatgagc ctgaatcca catagtgaga gttgggggtc cacagcgcat gatcaccagc  | 600  |
|                                                                   |      |
| cactgcttcc ctgagaccca gttcatagcg gtgactgctt atcagaacga ggagatcaca | 660  |
| gctcttaaaa ttaagtacaa tccatttgca aaagctttcc ttgatgcaa ggaaagaagt  | 720  |
| gatcacaag agatgatgga ggaacccgga gacagccagc aacctgggtg ctcccaatgg  | 780  |
| gggtggcttc ttcttggaac cagcacgtt tgtccactg caaatctca tctcagttt     | 840  |
| ggaggtgccc tctccctccc ctccacgac agctgtgaca ggtacccaac cctgaggagc  | 900  |
| caccggtcct caccctacce cageccctat gtcacgga acaattctcc aacctattct   | 960  |
| gacaactcac ctgcatgttt atccatgctg caatcccatg acaattggtc cagccttgga | 1020 |
|                                                                   |      |
| atgcctgccc atcccagcat gtcctccgtg agccacaatg ccagcccacc taccagctcc | 1080 |
| agtcagtacc ccagcctgtg gtctgtgagc aacggcgccg tcacccggg ctcccaggca  | 1140 |
| gcagcctgt ccaacgggtg gggggcccag ttcttcggg gtcctccgc gactacaca     | 1200 |
| ccccacccc atccggtctc ggcgcctct tctcgggat cccactgta cgaagggcg      | 1260 |
| gccgcggcca cagacatcgt ggacagccag tacgagccg cagcccaagg ccgcctcata  | 1320 |
| gcctcatgga cactgtgtc gccaccttc atgtga                             | 1356 |

<210> 66

<211> 452

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by

SEQ ID NO:64

<400> 66

Met Asn Met Ser Lys Lys Gln Val Lys Thr Gln Ser Lys Cys Asn Asn  
1 5 10 15  
Ile Met Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr  
20 25 30  
Arg Val Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly  
35 40 45  
Ser Glu Lys Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly Leu Glu  
50 55 60  
Glu Ser Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile  
65 70 75 80  
Val Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val  
85 90 95  
Ser Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val  
100 105 110  
Ala Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro  
115 120 125  
Gly Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro  
130 135 140  
Asp Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe  
145 150 155 160  
Ser Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met  
165 170 175  
Leu Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val  
180 185 190  
Gly Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln  
195 200 205  
Phe Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys  
210 215 220  
Ile Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg  
225 230 235 240  
Ser Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro

245 250 255  
 Gly Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Leu Cys  
 260 265 270  
 Pro Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro  
 275 280 285  
 Ser Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser  
 290 295 300  
 Ser Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr

305 310 315 320  
 Ser Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn  
 325 330 335  
 Trp Ser Ser Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser  
 340 345 350  
 His Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp  
 355 360 365  
 Ser Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val

370 375 380  
 Ser Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr  
 385 390 395 400  
 Thr Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro  
 405 410 415  
 Leu Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr  
 420 425 430  
 Asp Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser

435 440 445  
 Pro Pro Ser Met

450

<210> 67

<211> 451

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by

SEQ ID NO:65

<400> 67

Met Asn Met Ser Lys Lys Gln Val Lys Thr Gln Ser Lys Cys Asn Asn

1 5 10 15

Ile Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr Arg

20 25 30

Val Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly Ser

35 40 45

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

50 55 60

Ser Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile Val

65 70 75 80

Thr Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val Ser

85 90 95

Gly Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val Ala

100 105 110

Ala Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro Gly

115 120 125

Gly Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp

130 135 140

Ser Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser

145 150 155 160

Lys Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met Leu

165 170 175

Asn Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val Gly

180 185 190

Gly Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln Phe

195 200 205

Ile Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile

210 215 220

Lys Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser

225                      230                      235                      240  
 Asp His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly  
                                  245                      250                      255  
 Tyr Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Leu Cys Pro  
                                  260                      265                      270  
 Pro Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser  
                                  275                      280                      285  
  
 Thr His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser  
                                  290                      295                      300  
 Pro Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser  
 305                      310                      315                      320  
 Asp Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp  
                                  325                      330                      335  
 Ser Ser Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser His  
                                  340                      345                      350  
  
 Asn Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser  
                                  355                      360                      365  
 Val Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser  
                                  370                      375                      380  
 Asn Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr  
 385                      390                      395                      400  
 Pro Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu  
                                  405                      410                      415  
  
 Tyr Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp  
                                  420                      425                      430  
 Ala Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro  
                                  435                      440                      445  
 Pro Ser Met  
                                  450  
 <210> 68  
 <211> 1353  
 <212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury protein Isoform 1 plus addition of 42 nucleotides from Gene FPV 088 and an atc

<400> 68

|                                                                   |      |
|-------------------------------------------------------------------|------|
| atgagtaaga aacaagtaaa aactcaaagt aaatgtaata atatcatgag ctccctggc  | 60   |
| accgagagcg cgggaaagag cctgcagtac cgagtggacc acctgctgag cgccgtggag | 120  |
| aatgagctgc aggcgggcag cgagaagggc gacccacag agcgcgaact gcgcgtgggc  | 180  |
| ctggaggaga gcgagctgtg gctgcgcttc aaggagctca ccaatgagat gatcgtgacc | 240  |
| aagaacggca ggaggatgtt tccggtgctg aaggtgaacg tgtctggcct ggacccaac  | 300  |
| gccatgtact ctttctgct ggacttctg gcggcggaca accaccgtg gaagtacgtg    | 360  |
| aacggggaat gggtgccggg gggcaagccg gagccgcagg cgcccagctg cgtctacatc | 420  |
|                                                                   |      |
| caccccgact cgccaactt cggggccac tggatgaagg ctccgtctc cttcagcaaa    | 480  |
| gtcaagctca ccaacaagct caacggaggg ggccagatca tgtgaactc cttgcataag  | 540  |
| tatgagcctc gaatccacat agtgagagtt gggggtccac agcgcagat caccagccac  | 600  |
| tgtttcctg agaccagtt catagcgtg actgcttacc agaacgagga gatcacagct    | 660  |
| cttaaaatta agtacaatcc atttgcaaaa gctttccttg atgcaaagga aagaagtgat | 720  |
| cacaaagaga tgatggagga acccgagac agccagcaac ctgggtactc ccaatggggg  | 780  |
| tggtttcttc ctggaaccag caccgtttgt ccacctgcaa atctcatcc tcagtttga   | 840  |
|                                                                   |      |
| ggtgacctct cctccctc cagcacagc tgtgacaggt acccaaccct gaggagccac    | 900  |
| cggctctcac cctaccag cccctatgct catcggaaca attctcaac ctattctgac    | 960  |
| aactcacctg catgtttacc catgctgcaa tcccatgaca attggtccag ccttggaatg | 1020 |
| cctgcccacc ccagcatgct cccgtgagc cacaatgcca gccacctac cagctccagt   | 1080 |
| cagtaccca gcctgtggtc tgtgagcaac ggcgccgtca cccgggctc ccaggcagca   | 1140 |
| gccgtgtcca acgggctggg ggcccagttc ttccggggct ccccgcgca ctacacacc   | 1200 |
| ctcaccatc cggctctggc gccctcttc tcgggatccc cactgtacga agggcgggcc   | 1260 |
|                                                                   |      |
| gcggccacag acatcgtgga cagccagtac gacgccgag cccaaggccg cctcatagcc  | 1320 |
| tcattggacac ctgtgtcgcc accttccatg tga                             | 1353 |

<210> 69

<211> 1350

<212> DNA

<213> Artificial Sequence

<220><223> Coding sequence encoding Fusion protein comprising Brachyury protein Isoform 1 plus addition of 42 nucleotides from Gene FPV 088 and an atc, without Brachyury atg start codon

<400> 69

```

atgagtaaga aacaagtaaa aactcaaagt aaatgtaata atatcagctc ccctggcacc      60

gagagcgcg gaaagagcct gcagtaccga gtggaccacc tgctgagcgc cgtggagaat      120
gagctgcagg cgggcagcga gaaggcgac cccacagagc gcgaactgcg cgtgggcctg      180
gaggagagcg agctgtggct gcgcttcaag gagctcacca atgagatgat cgtgaccaag      240
aacggcagga ggatgtttcc ggtgtgaag gtgaacgtgt ctggcctgga cccaacgcc      300
atgtactcct tctgtctgga cttcgtggcg gcggacaacc accgctggaa gtacgtgaac      360
ggggaatggg tgccgggggg caagccggag ccgcaggcgc ccagctgcgt ctacatccac      420
cccgactcgc ccaacttcgg ggcccactgg atgaaggctc ccgtctcctt cagcaaagtc      480

aagtcacca acaagctcaa cggagggggc cagatcatgc tgaactcctt gcataagtat      540
gagcctcgaa tccacatagt gagagttggg ggtccacagc gcatgacac cagccactgc      600
ttccctgaga ccagttcat agcggtgact gcttatcaga acgaggagat cacagctctt      660
aaaattaagt acaatccatt tgcaaaagct ttccttgatg caaaggaaaag aagtgatcac      720
aaagagatga tggaggaacc cggagacagc cagcaacctg ggtactccca atgggggtgg      780
cttcttctg gaaccagcac cgtttgtcca cctgcaaate ctcactctca gtttgagggt      840
gccctctccc tcccctccac gcacagctgt gacaggtacc caaccctgag gagccaccgg      900

tcctcacct accccagccc ctatgctcat cggaacaatt ctccaaccta ttctgacaac      960
tcacctgcat gtttatccat gctgcaatcc catgacaatt ggtccagcct tggaatgcct      1020
gcccaccca gcatgctccc cgtgagccac aatgccagcc cacctaccag ctccagtcag      1080
taccacgccc tgtggtctgt gagcaacggc gccgtcacc cgggctccca ggcagcagcc      1140
gtgtccaacg ggtggggggc ccagttcttc cggggctccc ccgcgacta cacaccctc      1200
acccatccgg tctcggcgcc ctcttctctg ggatccccac tgtacgaagg ggcggccgcg      1260
gccacagaca tcgtggacag ccagtacgac gccgcagccc aaggccgcct catagcctca      1320

tggacacctg tgcgccacc ttccatgtga      1350

```

<210> 70

<211> 450

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by  
SEQ ID NO:68

<400> 70

Met Ser Lys Lys Gln Val Lys Thr Gln Ser Lys Cys Asn Asn Ile Met

1 5 10 15

Ser Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr Arg Val

20 25 30

Asp His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly Ser Glu

35 40 45

Lys Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly Leu Glu Glu Ser

50 55 60

Glu Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile Val Thr

65 70 75 80

Lys Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val Ser Gly

85 90 95

Leu Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val Ala Ala

100 105 110

Asp Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro Gly Gly

115 120 125

Lys Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp Ser

130 135 140

Pro Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser Lys

145 150 155 160

Val Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met Leu Asn

165 170 175

Ser Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val Gly Gly

180 185 190

Pro Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln Phe Ile

195 200 205

Ala Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile Lys

|                                                                 |     |     |     |
|-----------------------------------------------------------------|-----|-----|-----|
| 210                                                             | 215 | 220 |     |
| Tyr Asn Pro Phe Ala Lys Ala Phe Leu Asp Ala Lys Glu Arg Ser Asp |     |     |     |
| 225                                                             | 230 | 235 | 240 |
| His Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly Tyr |     |     |     |
|                                                                 | 245 | 250 | 255 |
| Ser Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Leu Cys Pro Pro |     |     |     |
|                                                                 | 260 | 265 | 270 |
| Ala Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr |     |     |     |
|                                                                 | 275 | 280 | 285 |
| His Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser Pro |     |     |     |
|                                                                 | 290 | 295 | 300 |
| Tyr Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser Asp |     |     |     |
| 305                                                             | 310 | 315 | 320 |
| Asn Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp Ser |     |     |     |
|                                                                 | 325 | 330 | 335 |
| Ser Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser His Asn |     |     |     |
|                                                                 | 340 | 345 | 350 |
| Ala Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val |     |     |     |
|                                                                 | 355 | 360 | 365 |
| Ser Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser Asn |     |     |     |
|                                                                 | 370 | 375 | 380 |
| Gly Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr Pro |     |     |     |
| 385                                                             | 390 | 395 | 400 |
| Leu Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr |     |     |     |
|                                                                 | 405 | 410 | 415 |
| Glu Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala |     |     |     |
|                                                                 | 420 | 425 | 430 |
| Ala Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro Pro |     |     |     |
|                                                                 | 435 | 440 | 445 |
| Ser Met                                                         |     |     |     |
| 450                                                             |     |     |     |

<210> 71

<211> 449

<212> PRT

<213> Artificial Sequence

<220><223> Fusion protein comprising Brachyury protein Isoform 1 encoded by  
SEQ ID NO:69

<400> 71

Met Ser Lys Lys Gln Val Lys Thr Gln Ser Lys Cys Asn Asn Ile Ser

1 5 10 15  
Ser Pro Gly Thr Glu Ser Ala Gly Lys Ser Leu Gln Tyr Arg Val Asp

20 25 30  
His Leu Leu Ser Ala Val Glu Asn Glu Leu Gln Ala Gly Ser Glu Lys

35 40 45  
Gly Asp Pro Thr Glu Arg Glu Leu Arg Val Gly Leu Glu Glu Ser Glu

50 55 60  
Leu Trp Leu Arg Phe Lys Glu Leu Thr Asn Glu Met Ile Val Thr Lys

65 70 75 80  
Asn Gly Arg Arg Met Phe Pro Val Leu Lys Val Asn Val Ser Gly Leu

85 90 95  
Asp Pro Asn Ala Met Tyr Ser Phe Leu Leu Asp Phe Val Ala Ala Asp

100 105 110  
Asn His Arg Trp Lys Tyr Val Asn Gly Glu Trp Val Pro Gly Gly Lys

115 120 125  
Pro Glu Pro Gln Ala Pro Ser Cys Val Tyr Ile His Pro Asp Ser Pro

130 135 140  
Asn Phe Gly Ala His Trp Met Lys Ala Pro Val Ser Phe Ser Lys Val

145 150 155 160  
Lys Leu Thr Asn Lys Leu Asn Gly Gly Gly Gln Ile Met Leu Asn Ser

165 170 175  
Leu His Lys Tyr Glu Pro Arg Ile His Ile Val Arg Val Gly Gly Pro

180 185 190

Gln Arg Met Ile Thr Ser His Cys Phe Pro Glu Thr Gln Phe Ile Ala

195 200 205

Val Thr Ala Tyr Gln Asn Glu Glu Ile Thr Ala Leu Lys Ile Lys Tyr

210 215 220

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

225 230 235 240

Lys Glu Met Met Glu Glu Pro Gly Asp Ser Gln Gln Pro Gly Tyr Ser

245 250 255

Gln Trp Gly Trp Leu Leu Pro Gly Thr Ser Thr Leu Cys Pro Pro Ala

260 265 270

Asn Pro His Pro Gln Phe Gly Gly Ala Leu Ser Leu Pro Ser Thr His

275 280 285

Ser Cys Asp Arg Tyr Pro Thr Leu Arg Ser His Arg Ser Ser Pro Tyr

290 295 300

Pro Ser Pro Tyr Ala His Arg Asn Asn Ser Pro Thr Tyr Ser Asp Asn

305 310 315 320

Ser Pro Ala Cys Leu Ser Met Leu Gln Ser His Asp Asn Trp Ser Ser

325 330 335

Leu Gly Met Pro Ala His Pro Ser Met Leu Pro Val Ser His Asn Ala

340 345 350

Ser Pro Pro Thr Ser Ser Ser Gln Tyr Pro Ser Leu Trp Ser Val Ser

355 360 365

Asn Gly Ala Val Thr Pro Gly Ser Gln Ala Ala Ala Val Ser Asn Gly

370 375 380

Leu Gly Ala Gln Phe Phe Arg Gly Ser Pro Ala His Tyr Thr Pro Leu

385 390 395 400

Thr His Pro Val Ser Ala Pro Ser Ser Ser Gly Ser Pro Leu Tyr Glu

405 410 415

Gly Ala Ala Ala Ala Thr Asp Ile Val Asp Ser Gln Tyr Asp Ala Ala

420 425 430

Ala Gln Gly Arg Leu Ile Ala Ser Trp Thr Pro Val Ser Pro Pro Ser

|                                                                           |      |     |
|---------------------------------------------------------------------------|------|-----|
| 435                                                                       | 440  | 445 |
| Met                                                                       |      |     |
| <210> 72                                                                  |      |     |
| <211> 1516                                                                |      |     |
| <212> DNA                                                                 |      |     |
| <213> Artificial Sequence                                                 |      |     |
| <220><223> Expression Cassette found in mBN345 - Promoter + cDNA encoding |      |     |
| Fusion protein comprising Brachyury (L254V) Protein                       |      |     |
| <400> 72                                                                  |      |     |
| atccgtacag gtttgtttct gaaattcact ttgtaagata cataattaac aaattcaggg         | 60   |     |
| ggaaaaatct ttacaaaatt agtatagaag ctatagatat atcaaaaggt agacaacaaa         | 120  |     |
| taatcagaac ctaatTTTTT tatcaaaaaa ttaaaatata aataaaatga aaaataactt         | 180  |     |
| gtatgaagaa aaaaatgaaca tgagtaagaa aagctccctt ggccaccgaga gcgcgggaaa       | 240  |     |
| gagcctgcag taccgagtgg accacctgct gagcgccgtg gagaatgagc tgcaggcggg         | 300  |     |
| cagcgagaag ggcgacccca cagagcgca actgcgcgtg ggccctggagg agagcgagct         | 360  |     |
| gtggctgcgc ttcaaggagc tcaccaatga gatgatcgtg accaagaacg gcaggaggat         | 420  |     |
| gtttccgggt ctgaagggtg acgtgtcttg cctggacccc aacgccatgt actccttctt         | 480  |     |
| gctggacttc gtggcggcgg acaaccaccg ctggaagtac gtgaacgggg aatgggtgcc         | 540  |     |
| ggggggcaag ccggagccgc aggcgcccag ctgcgtctac atccacccc actcgcccaa          | 600  |     |
| cttcggggcc cactggatga aggcctccct ctccttcagc aaagtcaagc tcaccaacaa         | 660  |     |
| gctcaacgga gggggccaga tcatgctgaa ctccttgcct aagtatgagc ctgcaatcca         | 720  |     |
| catagtgaga gttgggggtc cacagcgcat gatcaccagc cactgcttcc ctgagaccca         | 780  |     |
| gttcatacg gtgactgctt atcagaacga ggagatcaca gctcttaaaa ttaagtacaa          | 840  |     |
| tccatttgca aaggctttcc ttgatgcaaa ggaaagaagt gatcacaag agatgatgga          | 900  |     |
| ggaacccgga gacagccagc aacctgggtt ctcccaatgg gggtggcttc ttcctggaac         | 960  |     |
| cagcaccgtt tgtccacctg caaatcctca tctcagttt ggaggtgccc tctcctccc           | 1020 |     |
| ctccacgcac agctgtgaca ggtaccaac cctgaggagc caccggtcct caccctaccc          | 1080 |     |
| cagcccctat gctcatcgga acaattctcc aacctattct gacaactcac ctgcatgttt         | 1140 |     |
| atccatgctg caatcccatg acaattggtc cagccttgga atgcctgccc atcccagcat         | 1200 |     |
| gctccccgtg agccacaatg ccagcccacc taccagctcc agtcagtacc ccagcctgtg         | 1260 |     |

gtctgtgagc aacggcgccg tcaccccggt ctcccaggca gcagccgtgt ccaacgggct 1320  
 gggggcccag ttcttcggg gctccccgc gcactacaca cccctcacc atccgtctc 1380  
 ggcgccctct tcttcgggat cccactgta cgaagggcg gccgcggcca cagacatcgt 1440  
 ggacagccag tacgacgccc cagcccaagg ccgcctcata gcctcatgga cacctgtgtc 1500  
 gccaccttcc atgtga 1516

<210> 73

<211> 1552

<212> DNA

<213> Artificial Sequence

<220><223> Expression Cassette found in mBN355 - Promoter + cDNA encoding  
 Fusion protein comprising Brachyury (L254V) Protein

<400> 73

atccgtacag gtttgtttct gaaattcact ttgtaagata cataattaac aaattcaggg 60  
 ggaaaaatct ttacaaaatt agtatagaag ctatagatat atcaaaaggt agacaacaaa 120  
 taatcagaac ctaatTTTTT tatcaaaaaa ttaaaatata aataaaatga aaaataactt 180  
 glatgaagaa aaaatgaaca tgagtaagaa acaagtaaaa actcaaagta aatgtaataa 240

tatcatgagc tcccctggca ccgagagcgc gggaaagagc ctgcagtacc gagtggacca 300  
 cctgtctgagc gccgtggaga atgagctgca ggcgggcagc gagaagggcg accccacaga 360  
 gcgcgaactg cgcgtgggcc tggaggagag cgagctgtgg ctgcgcttca aggagctcac 420  
 caatgagatg atcgtgacca agaacggcag gaggatgttt ccggtgctga aggtgaacgt 480  
 gtctggcctg gaccccaacg ccatgtactc ctctctctg gacttcgtgg cggcggacaa 540  
 ccaccgttg aagtacgtga acggggaatg ggtgccgggg ggcaagccgg agccgcaggc 600  
 gccagctgc gtctacatcc accccgactc gcccaacttc ggggcccact ggatgaaggc 660

tcccgtctcc ttacgaaaag tcaagctcac caacaagctc aacggagggg gccagatcat 720  
 gctgaactcc ttgcataagt atgagcctcg aatccacata gtgagagtgt ggggtccaca 780  
 gcgcatgac accagccact gcttcctga gaccagttc atagcgtga ctgcttatca 840  
 gaacgaggag atcacagctc ttaaaattaa gtacaatcca tttgcaaagg ctttccttga 900  
 tgcaaaggaa agaagtgatc acaaagagat gatggaggaa cccggagaca gccagcaacc 960  
 tgggtactcc caatgggggt ggcttcttcc tggaaccagc accgtttgtc cacctgcaaa 1020  
 tcctcatcct cagtttggag gtgccctctc cctccctcc acgcacagct gtgacaggta 1080

cccaaccctg aggagccacc ggtctctacc ctacccagc ccctatgctc atcggaacaa 1140

ttctccaacc tattctgaca actcacctgc atgtttatcc atgctgcaat cccatgacaa 1200  
 ttggtccagc ctiggaatgc ctgcccaccc cagcatgctc cccgtgagcc acaatgccag 1260  
 cccacctacc agctccagtc agtaccaccc cctgtggtct gtgagcaacg gcgccgtcac 1320  
 cccgggctcc caggcagcag ccgtgtccaa cgggctgggg gccagttct tccggggctc 1380  
 ccccgcgcac tacacacccc tcacccatcc ggtctcggcg cctcttcct cgggatcccc 1440  
 actgtacgaa ggggcggccg cgccacaga catcgtggac agccagtacg acgccgcagc 1500

ccaaggccgc ctcatagcct catggacacc tgtgtcgcca cttccatgt ga 1552

<210> 74

<211> 1464

<212> DNA

<213> Artificial Sequence

<220><223> Expression Cassette found in mBN343 - Promoter + cDNA encoding  
 Fusion protein comprising Brachyury (L254V) Protein

<400> 74

taatgtatag aactaattta taataaacat agtaaataatg ggtaacttct taatagccat 60  
 aattaaaatt gaaaaaaaaa tatcattata aaacgtaac gaacaaaaaa cattaattga 120  
 tccccccct gtctccccc tccagattga gcaatcatga gctcccctgg caccgagagc 180

gcgggaaaga gcctgcagta ccgagtggac cacctgctga gcgccgtgga gaatgagctg 240  
 caggcgggca gcgagaaggc cgaccccaca gagcgcgaac tgcgcgtggg cctggaggag 300  
 agcagactgt ggctgcgttt caaggagctc accaatgaga tgatcgtgac caagaacggc 360  
 aggaggatgt ttccggtgct gaaggtgaac gtgtctggcc tggaccccaa cgccatgtac 420  
 tccttcctgc tggacttcgt ggccgaggac aaccaccgct ggaagtacgt gaacggggaa 480  
 tgggtgccgg ggggcaagcc ggagccgcag gcgccagct gcgtctacat ccaccccgac 540  
 tcgccaact tcggggccca ctggatgaag gctcccgtct ccttcagcaa agtcaagctc 600

accaacaagc tcaacggagg gggccagatc atgctgaact ccttgcataa gtatgagcct 660  
 cgaatccaca tagtgagagt tgggggtcca cagcgcatga tcaccagcca ctgcttcctt 720  
 gagaccagc tcatagcgtt gactgcttat cagaacgagg agatcacagc tcttaaaatt 780  
 aagtacaatc catttgcaa ggctttcctt gatgcaaagg aaagaagtga tcacaaagag 840  
 atgatggagg aacccggaga cagccagcaa cctgggtact cccaatgggg gtggcttctt 900  
 cctggaacca gcaccgtttg tccacctgca aatcctcacc ctcagtttgg aggtgcctc 960  
 tccctccctt ccacgcacag ctgtgacagg tacccaaccc tgaggagcca cggctcctca 1020

ccctacccca gccctatgc tcatcggaac aattctccaa cctattctga caactcacct 1080  
gcatgtttat ccatgctgca atcccatgac aattggtcca gccttggaat gcctgcccac 1140  
cccagcatgc tccccgtgag ccacaatgcc agcccaccta ccagctccag tcagtacccc 1200  
agcctgtggt ctgtgagcaa cggcgccgtc accccgggct cccaggcagc agccgtgtcc 1260  
aacgggctgg gggcccagtt cttccggggc tccccgcgc actacacacc cctcacccat 1320  
ccggtctcgg cgcctcttc ctcgggatcc cactgtacg aaggggcggc cgcggccaca 1380  
gacatcgtgg acagccagta cgacgccga gcccaaggcc gcctcatagc ctcattggaca 1440

cctgtgtcgc caccttccat gtga 1464

<210> 75

<211> 1478

<212> DNA

<213> Artificial Sequence

<220><223> Expression Cassette found in mBN344 - Promoter + cDNA encoding

Fusion protein comprising Brachyury (L254V) Protein

<400> 75

tatccgtaca ggttgtttc tgaaattcac ttgtgaagat acataattaa caaattcagg 60  
gggaaaaatc tttaaaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa 120  
ataatcagaa cctaattttt ttatcaaaaa attaaaatat aaataaaatc atgagctccc 180

ctggcaccga gagcgcgga aagagcctgc agtaccgagt ggaccacctg ctgagcgccg 240  
tggagaatga gctgcaggcg ggcagcgaga agggcgaccc cacagagcgc gaactgcgcg 300  
tgggcctgga ggagagcgag ctgtggctgc gctcaagga gctcaccaat gagatgatcg 360  
tgaccaagaa cggcaggagg atgtttccgg tgctgaaggt gaacgtgtct ggcctggacc 420  
ccaacgcat gtactccttc ctgctggact tcgtggcggc ggacaaccac cgctggaagt 480  
acgtgaacgg ggaatgggtg cccgggggca agccggagcc gcaggcgccc agctgcgtct 540  
acatccacc cgactcgccc aacttcgggg cccactggat gaaggctccc gtctccttca 600

gcaaagtcaa gctcaccaac aagctcaac gagggggcca gatcatgctg aactccttgc 660  
ataagatga gcctgaatc cacatagtga gatttggggg tccacagcgc atgatcacca 720  
gccactgctt cctgagacc cagttcatag cgttgactgc ttatcagaac gaggagatca 780  
cagctcttaa aattaagtac aatccatttg caaaggcttt ccttgatgca aaggaaagaa 840  
tgatcacaa agatgatgag gaggaaccg gagacagcca gcaacctggg tactcccaat 900  
gggggtggct tcttcttga accagaccg ttgtccacc tgcaaatcct catcctcagt 960

ttggagggtgc cctctccctc ccctccacgc acagctgtga caggtaccca accctgagga 1020

gccaccggtc ctcaccctac ccagcccct atgtcatcg gaacaattct ccaacctatt 1080

ctgacaactc acctgcatgt ttatccatgc tgcaatccca tgacaattgg tccagccttg 1140

gaatgcctgc ccatcccagc atgtccccg tgagccacaa tgccagccca cctaccagct 1200

ccagtcagta cccagcctg tggctctgtga gcaacggcgc cgtcaccccg ggctcccagg 1260

cagcagccgt gtccaacggg ctgggggccc agttcttccg gggtccccc gcgcactaca 1320

caccctcac ccatccggtc tcggcgccct ctctctcggg atccccactg tacgaagggg 1380

cggccggcgc cacagacatc gtggacagcc agtacgacgc cgcagcccaa ggccgcctca 1440

tagcctcatg gacacctgtg tcgccacctt ccatgtga 1478

<210> 76

<211> 1486

<212> DNA

<213> Artificial Sequence

<220><223> Expression Cassette found in mBN354 - Promoter + cDNA encoding  
Fusion protein comprising Brachyury (L254V) Protein

<400> 76

atccgtacag gtttgtttct gaaattcact ttgtaagata cataattaac aaattcaggg 60

ggaaaaatct ttacaaaatt agtatagaag ctatagatat atcaaaaggt agacaacaaa 120

taatcagaac ctaatTTTTT tatcaaaaaa ttaaaatata aataaaatga aaaataactt 180

gagctccctt ggcaccgaga gcgcgggaaa gagcctgcag taccgagtgg accacctgct 240

gagcgccgtg gagaatgagc tgcaggcggg cagcgagaag ggcgaccca cagagcgcca 300

actgcgcgtg ggcctggagg agagcgagct gtggctgcgc ttcaaggagc tcaccaatga 360

gatgatcgtg accaagaacg gcaggaggat gtttccggtg ctgaaggtga acgtgtctgg 420

cctggacccc aacgccatgt actcttctct gctggacttc gtggcggcgg acaaccaccg 480

ctggaagtac gtgaacgggg aatgggtgcc ggggggcaag ccggagccgc aggcgcccag 540

ctgcgtctac atccaccccg actcgcccaa cttcggggcc cactggatga aggtccctt 600

ctcttcagc aaagtcaagc tcaccaacaa gctcaacgga gggggccaga tcatgtctgaa 660

ctccttgcat aagtatgagc ctcgaaatcca catagtgaga gttgggggtc cacagcgcat 720

gatcaccagc cactgcttcc ctgagacca gttcatagcg gtgactgctt atcagaacga 780

ggagatcaca gctcttaaaa ttaagtacaa tcattttgca aagcttttcc ttgatgcaaa 840

ggaaagaagt gatcacaag agatgatgga ggaacccgga gacagccagc aacctgggta 900

|                                                                       |      |
|-----------------------------------------------------------------------|------|
| ctcccaatgg ggggtggcttc ttcttgggaac cagcaccggtt tgtccacctg caaatcctca  | 960  |
| tcctcagttt ggaggtgccc tctccctccc ctccacgcac agctgtgaca ggtaccaac      | 1020 |
|                                                                       |      |
| cctgaggagc caccggtcct caccctaccc cagccctat gctcatcgga acaattctec      | 1080 |
| aacctattct gacaactcac ctgcatgttt atccatgctg caatcccatg acaattggtc     | 1140 |
| cagccttgga atgcctgccc atcccagcat gctccccgtg agccacaatg ccagcccacc     | 1200 |
| taccagctcc agtcagtacc ccagcctgtg gtctgtgagc aacggcgccg tcaccccggg     | 1260 |
| ctcccaggca gcagccgtgt ccaacgggct gggggcccag ttcttccggg gctccccgc      | 1320 |
| gcactacaca cccctaccc atccggtctc ggcgccctct tctcgggat cccactgta        | 1380 |
| cgaaggggcg gccgcgcca cagacatcgt ggacagccag tacgacgccc cagccaagg       | 1440 |
|                                                                       |      |
| ccgcctcata gcctcatgga cacctgtgtc gccaccttc atgtga                     | 1486 |
| <210> 77                                                              |      |
| <211> 182                                                             |      |
| <212> DNA                                                             |      |
| <213> Artificial Sequence                                             |      |
| <220><223> Promoter Sequence with nucleotides 5' of Gene FPV 088 + 15 |      |
| nucleotides                                                           |      |
| <400> 77                                                              |      |
| tatccgtaca ggtttgtttc tgaaattcac ttgtgaagat acataattaa caaattcagg     | 60   |
| gggaaaaatc ttacaaaat tagtatagaa gctatagata tatcaaaagg tagacaacaa      | 120  |
| ataatcagaa cctaattttt ttatcaaaaa attaaaatat aaataaaatg aaaaataact     | 180  |
| tg                                                                    | 182  |