(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(51) 。Int. Cl. [6]
C12N 15/12

(45) 공고일자 2002년09월28일
(11) 등록번호 10-0353182
(24) 등록일자 2002년09월05일

(21) 출원번호 10-1999-0019060
(22) 출원일자 1999년05월26일
(62) 원출원 특허특1996-0703014
원출원일자 : 1996년06월07일

(65) 공개번호 특0000-0000000
(43) 공개일자 0000년00월00일

심사청구일자 1996년12월04일

(30) 우선권주장 08/164103 1993년12월07일 미국(US)
08/217780 1994년03월25일 미국(US)
08/333576 1994년11월02일 미국(US)

(73) 특허권자 제네틱스 인스티튜트, 엘엘씨
미국 매사추세츠 02140 캠브리지 캠브리지파크 드라이브 87
프레지던트 앤드 펠로우즈 오브 하바드 칼리지
미합중국, 메사추세츠 02138, 캠브리지, 퀸시 스트리트 17

(72) 발명자 톰슨제랄드에이치.
미국뉴욕11777포트제퍼슨베이뷰테라스201
멜튼더글라스에이.
미국메사추세트02173렉싱톤슬로컴로오드22
셀레스트안토니제이.
미국메사추세트01749휴드슨팩커드스트리트86
워즈니존엠.
미국메사추세트01749휴드슨올드볼튼로오드59
로오젠비키에이.
미국메사추세트02146브룩클린킬시트로오드#7127
울프만네일엠.
미국메사추세트02030도버롤링레인30

(74) 대리인 특허법인코리아나
조영원

심사관 : 임혜준

(54) 비엠피-12, 비엠피-13 및 이의 건-유도 조성물

요약

골 형태 발생 단백질 BMP-12 및 BMP-13 를 클론시킨다. 건/인대-류 조직 유도 활성을 갖는 상기 단백질의 조성물을 개시한다. 조성물은 건염 및 건 또는 인대 결함의 치료 및 관련 조직 치유에 유용하다.

대표도

도 1

색인어

BMP-12

명세서

도면의 간단한 설명

제 1 도는 인간 BMP-12 서열과 인간 MP52 서열의 비교이다.

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 출원 시리얼 번호 08/217,780 (1994, 3, 25 일 출원) 08/164,103 (1993, 12, 7 일 출원) 및 08/333,576 (1994, 11, 2 일 출원) 의 일부의 계속출원이다.

본 발명은 새로운 패밀리의 정제된 단백질, 및 상기 단백질을 함유하는 조성물에 관한 것이며, 이 조성물은 건(腱)/인대-류 조직 형성의 유도, 상처 치료 및 인대 및 기타 조직 치유에 유용하다. 또한 상기 단백질은 골 형태발생 단백질의 활성을 증가시키기위한 조성물로 사용할 수 있다.

골, 연골 및 골내의 건 및 기타 조직의 형성에 관여하는 분자 또는 분자들 및 기타 조직 추출물에 대한 연구로 골 형태발생 단백질 (BMPs) 이라 불리는 새로운 일련의 분자를 발견하게 되었다. BMP-1 에서 BMP-11 로 나타낸, 여러 단백질의 구조는 이미 설명되어있다. 골내의 상기 단백질의 존재와 더불어, 상기 단백질의 독특한 유도 활성은 이들이 골 치료과정의 중요한 조절물질이며, 골조직의 정상적 유지에 관여할 수 있다는 것을 시사한다. 기타 중추 조직을 형성시키는 역을 하는 기타 단백질을 확인할 필요가 있다. 본 발명은 건/인대-류 조직 유도 활성을 갖고, 건/인대-류 조직 형성의 유도 및 복구용 조성물로 유용한 단백질의 패밀리를 확인하는 것에 관한 것이다.

발명이 이루고자 하는 기술적 과제

한 구현예로, 본 발명은 본 발명자가 V1-1 로 명한 건/인대-류 유도 단백질을 코딩하는 DNA 분자를 포함한다. 이 새로운 단백질은 이제 BP-12 라 부르고 있다. 또한 본 발명은 BMP-12 관련 단백질을 코딩하는 DNA 분자를 포함한다.

BMP-12 관련 단백질은 예컨대, 감소된 스트린전시 (stringency); 유사정 정도의 지표가 되는 조건의 엄격성 조건하, PCR, BMP-12 특이 프라이머 (예, 후술된 프라이머 # 6 및 #7) 를 사용하여, 클론 및 확인된 DNA 서열로 코딩되는 건/인대-류 조직 유도 단백질로 정의된, BMP-12 및 V1-1 를 포함하여, BMP/TGF-β/Vg-1 패밀리의 단백질의 한 부분 집합이다. BMP-12 관련 단백질을 코딩하는 DNA 서열은 SEQ ID NO : 1 의 아미노산 #3 내지 #103 과 아미노산 수준에서 약 80 % 이상의 상동성을 공유하는 것이 바람직하다.

상기 DNA 분자는 바람직하게는 BMP-12 단백질을 코딩하는 DNA 서열, (그것의 서열은 SEQ ID NO : 1 내에 제공되어 있음), 또는 여기에 추가로 기재된 BMP-12 관련 단백질을 코딩하는 DNA 서열을 갖는 것이다. BMP-12 단백질 및 BMP-12 관련 단백질들은 실시예에 기재된 검정에서 건/인대-류 조직의 형성을 유도하는 능력을 특징으로한다.

본 발명의 DNA 분자는 바람직하게는 SEQ ID NO : 1 에 기재된 것과 같은 DNA 서열 ; 더욱 바람직하게는 SEQ ID NO : 1 의 뉴클레오티드 #496 내지 #882, #571 내지 #882 또는 #557 내지 #882 ; 또는 엄격한 혼성화 조건하에서 전술한 것에 하이브리다이즈하고 건/인대-류 조직을 형성하는 능력을 보이는 단백질을 코딩하는 DNA 서열을 포함한 것이다. 또한 본 발명의 DNA 분자는 SEQ ID NO : 25에 기재된 것과 같은 DNA 서열 ; 좀더 바람직하게는 SEQ ID NO : 25 의 뉴클레오티드 #604 또는 #658 내지 #964 를 포함한 것이다.

또한 본 발명의 DNA 분자는 SEQ ID NO : 2 또는 SEQ ID NO : 26 의 천연의 대립유전자 서열 및 등가의 (equivalent) 중첩성 코돈 서열, SEQ ID NO :2 또는 SEQ ID NO : 26 에 나타낸 아미노산 서열을 갖는 BMP-12 관련 단백질을 코딩하는 DNA 서열을 포함하는 DNA 분자를 포함한 것이기도 하다. 바람직하게는, 본 발명의 DNA 서열은 SEQ ID NO : 2 의 아미노산 #-25 내지 #104, #1 내지 #104 또는 #3 내지 #103; 또는 SEQ ID NO : 26 의 #1 내지 #120 또는 #19 내지 #120 을 코딩한 것이다. 상기의 DNA 서열은, 5' 에서 3' 방향에서, 프로펩티드를 코딩하는 뉴클레오티드, 및 SEQ ID NO : 2 의 아미노산 #-25 내지 #104#, 혹은 #1 내지 #104, #3 내지 #103 ; 또는 SEQ ID NO : 26 의 아미노산 #1 내지 #120 또는 # 19 내지 #120 를 코딩하는 뉴클레오티드를 포함할 수 있다. 전술한 구현예에 유용한 프로펩티드는 바람직하게는 천연 BMP-12 프로펩티드 및 TGF-B 슈퍼패밀리 (superfamily) 또는 BMP 패밀리의 다른 단백질 유래의 단백질 프로펩티드로 구성된 군으로부터 선택한다. 본 발명은 더 나아가 엄격한 혼성화 조건하에서 전술한 DNA 서열과 하이브리다이즈하고 건/인대-류 조직의 형성을 유도하는 능력을 보이는 BMP-12 관련 단백질을 코딩하는 DNA 서열을 포함하는 것이기도 하다.

기타 구현예로, 본 발명은 BMP-12 단백질, 또는 BMP-12 관련 단백질을 코딩하는 DNA 분자를 포함하는 숙주 세포 및 벡터를 포함한다. 상기 숙주 세포 및 벡터는 발현 조절 서열과 연계하여 기능하는 코딩 서열을 더 포함할 수 있다.

기타 구현예로, 본 발명은 정제된 BMP-12 관련 단백질을 제조하는 방법을 포함하고, 이 방법은 BMP-12 관련 단백질을 코딩하는 뉴클레오티드 서열을 포함하는 전술한 DNA 분자 또는 벡터로 형질 전환된 숙주 세포를 배양하는 단계 ; 및 (b) 배양 배지로부터 상기 BMP-12 관련 단백질을 회수 및 정제하는 단계로 구성된다. 바람직한 구현예로, 방법은 (a)SEQ ID NO : 1 에 나타낸 바와 같은 뉴클레오티드 #496, #571 또는 #577 내지 #879 또는 #882 의 뉴클레오티드 서열 ; 또는 SEQ ID NO : 25 의 #604 또는 #658 내지 #963 의 뉴클레오티드 서열을 포함하는 DNA 분자로 형질 전환된 세포를 배양하고; 및

(b) 상기 배양 배지로부터 SEQ ID NO : 2 에 나타낸아미노산 #-25, #1 또는 #3 내지 아미노산 #103 또는 #104 의 아미노산 서열 ; 또는 SEQ ID NO : 26 에 나타낸 아미노산 #1 또는 #19 내지 아미노산 #120 의 아미노산 서열을 포함하는 단백질을 상기 배양 배지로부터 회수 및 정제하는 것으로 구성된다. 또한 본 발명은 전술한 방법으로 제조된 정제 단백질을 포함한다.

본 발명은 더 나아가 건/인대-류 조직의 형성을 유도하는 능력을 특징으로하는 정제된 BMP-12 관련 단백질을 포함한 것이다. BMP-12 관련 폴리펩티드는 바람직하게는 SEQ ID NO : 2 에 나타낸 아미노산 서열을 포함한 것이다. 폴리펩티드는 좀더 바람직하게는 SEQ ID NO : 2 에서 나타낸 아미노산# -25, #1 또는 #3 내지 #103 또는 #104 ; 또는 아미노산 SEQ ID NO : 26 에 나타낸 #1 또는 #19 내지 #120 을 포함한 것이다. 바람직한 구현예로, 정제된 폴리펩티드는 각각이 SEQ ID NO : 2 의 아미노산 서열을 갖는, 2 개의 소단위체로 구성된 다이머 형태일 수 있다.

또 다른 구현예에서, 본 발명은 전술한 BMP-12 관련 단백질을 유효량으로 포함하는 조성물을 포함한다. 이 조성물에서, 단백질은 약학적으로 허용가능한 부형제와 혼합할 수 있다.

또한 본 발명은 건/인대-류 조직 치유 및 조직 복구방법, 건염 또는 기타 건 또는 인대 결함의 치료방법, 건/인대-류 조직의 형성유도가 필요한 환자에서 건/인대-류 조직 형성을 유도하는 방법을 포함한 것이기도 하며, 이는 전술한 조성물을 유효량으로 전술한 환자에게 투여하는 것을 특징으로한다.

기타 구현예는 BMP-12 관련 폴리펩티드를 코딩하는 DNA 서열에 정확한 해독구조[틀]로 연결되어 있는 TGF-β 수퍼패밀리의 단백질들중의 하나의 프로펩티드를 코딩하는 DNA 서열을 포함하는 키메릭 (chimeric) DNA 분자를 포함한다. 적합한 프로펩티드 하나는 BMP-2 의 프로펩티드이다. 또한 본 발명은 SEQ ID NO : 2 에 나타낸 아미노산 서열을 갖는 하나의 모노머, 및 TGF-β 서브패밀리의 기타 단백질의 아미노산 서열을 갖는 하나의 모노머로 구성된 헤테로다이머성 단백질 분자를 포함한 것이기도 하다.

최종적으로, 본 발명은 SEQ ID NO : 2 또는 SEQ ID NO :4 또는 SEQ ID NO : 26 에 나타낸 아미노산 서열을 갖으며, 건/인대-류 조직의 형성을 유도하는 능력을 보이는 단백질을 포함하는 유효량의 조성물을 건/인대-류 조직 형성의 유도가 필요한 환자에게 투여하는 것을 특징으로하는 환자에서의 건/인대-류 조직의 형성을 유도하는 방법을 포함한 것이다. 상기 아미노산 서열은 좀더 바람직하게는 하기 것들중의 하나 이다 : (a) SEQ ID NO : 2 의 아미노산 #-25, #1 또는 #3 내지 #103 또는 #104 ; (b) SEQ ID NO : 4 의 #1 또는 #19 내지 #119 또는 #120 ; (c) SEQ ID NO : 26 의 #1 또는 #19 내지 #119 또는 #120 ; (d) 건 및/또는 인대를 형성시키는 능력을 보이는 (a), (b) 또는 (c) 의 돌연변이체 및/또는 변형체. 전술한 방법의 기타 구현예로, 단백질은 SEQ ID NO : 1, SEQ ID NO : 3 또는 SEQ ID NO : 25 의 DNA 서열로 코딩되고, 좀더 바람직하게는 하기 것들중에서의 하나이다 : (a) SEQ ID NO : 1 의 뉴클레오티드 #496, #571 또는 #577 내지 #879 또는 #882 ; (b) SEQ ID NO : 3 의 뉴클레오티드 #845 또는 #899 내지 #1201 또는 #1204 ; (c) SEQ ID NO : 25 의 뉴클레오티드 #605 또는 #659 내지 #961 또는 #964 ; 및 (d) 엄격한 혼성화 조건하에서 (a) 또는 (b) 와 하이브리다이즈하고 건/인대-류 조직을 형성할 수 있는 능력을 보이는 단백질을 코딩하는 서열.

서열의 설명

SEQ ID NO : 1 은 인간 BMP-12 를 코딩하는 뉴클레오티드 서열이다.

SEQ ID NO : 2 는 성숙한 인간 BMP-12 폴리펩티드를 포함하는 아미노산 서열이다.

SEQ ID NO : 3 은 단백질 MP52 를 코딩하는 뉴클레오티드 서열이다.

SEQ ID NO : 4 는 성숙한 MP52 폴리펩티드를 포함하는 아미노산 서열이다.

SEQ ID NO : 5 는 인간 BMP-12 를 코딩하는 서열의 특별히 증폭된 부분의 뉴클레오티드 서열이다.

SEQ ID NO : 6 은 SEQ ID NO : 5 의 뉴클레오티드 서열에의해 코딩된 아미노산 서열이다.

SEQ ID NO : 7 은 인간 VL-1 을 코딩하는 서열의 특별히 증폭된 부분의 뉴클레오티드 서열이다.

SEQ ID NO : 8 은 SEQ ID NO : 7 의 뉴클레오티드 서열에의해 코딩되는 아미노산 서열이다.

SEQ ID NO : 9 는 E.coli. 에서의 BMP-12 의 발현을 위해 사용된, 플라스미드 pALV1-781 의 뉴클레오티드 서열이다.

SEQ ID NO : 10 은 뮤린 클론, mV1 의 한 단편의 뉴클레오티드 서열이다.

SEQ ID NO : 11 은 mV1 에 의해 코딩되는 뮤린 단백질의 한 단편의 아미노산 서열이다.

SEQ ID NO : 12 는 뮤린 클론, mV2 의 한 단편의 뉴클레오티드 서열이다.

SEQ ID NO : 13 은 mV2 에 의해 코딩되는 뮤린 단백질의 한 단편의 아미노산 서열이다.

SEQ ID NO : 14 는 뮤린 클론, mV9 의 한 단편의 뉴클레오티드 서열이다.

SEQ ID NO : 15 은 mV9 에 의해 코딩되는 뮤린 단백질의 한 단편의 아미노산 서열이다.

SEQ ID NO : 16 은 BMP/TGF-β/Vg-1 단백질 컨센서스 서열의 아미노산 서열이다. 처음의 Xaa 는 Gln 또는 Asn 을 나타낸다 ; 두번째의 Xaa 는 Val 또는 Ile 를 나타낸다.

SEQ ID NO : 17 은 올리고뉴클레오티드 #1 의 뉴클레오티드 서열이다.

SEQ ID NO : 18 은 BMP/TGF-β/Vg-1 단백질 컨센서스 서열의 아미노산 서열이다. Xaa 는 Val 또는 Leu 을 나타낸다.

SEQ ID NO : 19 는 올리고뉴클레오티드 #2 의 뉴클레오티드 서열이다.

SEQ ID NO : 20 은 올리고뉴클레오티드 #3 의 뉴클레오티드 서열이다.

SEQ ID NO : 21 은 올리고뉴클레오티드 #4 의 뉴클레오티드 서열이다.

SEQ ID NO : 22 는 올리고뉴클레오티드 #5 의 뉴클레오티드 서열이다.

SEQ ID NO : 23 은 올리고뉴클레오티드 #6 의 뉴클레오티드 서열이다.

SEQ ID NO : 24 는 올리고뉴클레오티드 #7 의 뉴클레오티드 서열이다.

SEQ ID NO : 25 는 인간 VL-1 (BMP-13) 을 코딩하는 서열의 뉴클레오티드 서열이다.

SEQ ID NO : 26 은 SEQ ID NO : 25 의 뉴클레오티드 서열에의해 코딩되는 아미노산 서열이다.

SEQ ID NO : 27 은 BMP-2 프로펩티드와 성숙한 BMP-12 의 코딩 서열의 융합단백질을 코딩하는 뉴클레오티드 서열이다.

SEQ ID NO : 28 은 SEQ ID NO : 27 의 뉴클레오티드 서열로 코딩되는 아미노산 서열이다.

SEQ ID NO : 29 는 뮤린 mV1 단백질을 코딩하는 뉴클레오티드 서열이다. X01 은 Val, Ala, Glu 또는 Gly 이고; X02 은 Ser, Pro, Thr 또는 Ala 이며; X03 는 Ser 또는 Arg 이고; X04 는 Leu, Pro, Gln 또는 Arg 이며; X05 은 Cys 또는 Trp 이고; X06 은 Val, Ala, Asp 또는 Gly 이며; X07 은 Val, Ala Glu 또는 Gly 이고; X08 은 Gln, Lys 또는 Glu 이다.

SEQ ID NO : 30 은 SEQ ID NO : 29 의 뉴클레오티드 서열에의해 코딩되는 아미노산 서열이다. X01 내지 X08 은 SEQ ID NO : 29 에서와 동일하다.

SEQ ID NO : 31 은 뮤린 mV2 단백질을 코딩하는 뉴클레오티드 서열이다. X01 은 Pro 또는 Thr 이고; X02 는 Val 이다.

SEQ ID NO : 32 는 SEQ ID NO : 31 의 뉴클레오티드 서열에 의해 코딩되는 아미노산 서열이다. X01 및 X02 는 SEQ ID NO : 31 에서와 동일하다.

SEQ ID NO : 33 은 인간 BMP-12 단백질을 코딩하는 뉴클레오티드 서열이다.

SEQ ID NO : 34 는 SEQ ID NO : 33 의 뉴클레오티드 서열에의해 코딩되는 아미노산 서열이다.

SEQ ID NO : 35 는 올리고뉴클레오티드 #8 의 뉴클레오티드 서열이다.

발명의 구성 및 작용

본 발명의 DNA 서열은 추후 기술되는 바와 같이, 건/인대-류 조직의 형성을 유도하는 단백질을 제조하는데 유용하다. 또한 본 발명의 DNA 서열은 유사한 활성을 갖는 BMP-12 관련 단백질을 코딩하는 DNA 서열을 추가로 단리 및 클로닝하는데 유용한 것이기도 하다. 상기 BMP-12 관련 단백질들은 다른종의 상동물들일 수 있으며, 또는, 동일한 종내의 관련된 단백질들일 수 있다.

또한, 본 발명의 또 하나의 면은 건/인대-류 조직을 유도하는 단백질의 발현을 코딩하는 DNA 서열이다. 이러한 서열은 유전 코드의 중첩성이 없다면, DNA 서열 SEQ ID NO : 1 또는 25 와 동일하고, SEQ ID NO 또는 26 의 단백질을 코딩하는 DNA 서열들인, SEQ ID NO : 1 또는 SEQ ID NO : 25 에 나타낸 5'→3' 방향의 뉴클레오티드들의 서열을 포함한다. 본 발명에는 엄격한 조건하에서 SEQ ID NO : 1 또는 25 의 DNA 서열과 하이브리다이즈하고 건 또는 인대의 형성을 유도하는 능력을 갖는 단백질을 코딩하는 DNA 서열도 또한 포함된다. 바람직한 DNA 서열은 문헌 (Maniatis et al, Molecular Cloning (A Laboratory Manual), Cold Spring Habor Laboratory (1982), p 387 내지 389) 에 기재된 것과 같은 엄격한 조건하에서 하이브리다이즈하는 것들을 포함한다. 최종적으로, 뉴클레오티드의 변화도 펩티드 서열이 변화되든 않든간에, 그러나 그 결과의 펩티드 서열은 여전히 건/인대-류 조직 유도 활성을 갖는 SEQ ID NO : 1 또는 25 의 서열의 대립유전자 또는 기타 형체들도 또한 본 발명에 포함된다.

인간 BMP-12 DNA 서열 (SEQ ID NO : 1) 및 아미노산 서열 (SEQ ID NO : 2) 을 서열 일람표에 나타내었다. 본 발명의 조성물 및 방법에 유용한 기타 단백질은 VL-1 이다. VL-1 은 BMP-12 의 서열을 이용하여 클론된 BMP-12 관련 단백질이다. 본 발명자들은 현재 VL-1 를 BMP-13 로 표시하였다. VL-1 의 부분 DNA 서열 (SEQ ID NO : 7) 및 코딩된 아미노산 서열 (SEQ ID NO : 8) ; 및 성숙한 VL-1 을 코딩하는 DNA 서열 (SEQ ID NO : 25) 및 코딩된 아미노산 서열 (SEQ ID NO : 26) 을 서열 일람표에 나타내었다. 비록 SEQ ID NO : 1 및 2 의 BMP-12 서열과 관련하여 추가로 기술되지만, 본 발명은 SEQ ID NO : 25 및 26 에 나타낸 VL-1 서열등과 같은, 기타 BMP-12 관련 서열들을 유사하게 부분변형 및 개선시킨 것도 또한 포함한다는 사실이 인식 될 것이다.

SEQ ID NO : 1 에 나타낸 BMP-12 의 서열은 완전한 성숙 서열 및 대략 190 아미노산의 프로펩티드를 포함한다. 성숙 인간 BMP-12 단백질의 코딩 서열은 뉴클레오티드 #496 또는 #571 에서 시작되는 것으로 보이며 SEQ ID NO : 1 의 뉴클레오티드 #882 를 거쳐 계속된다. TGF-β 단백질의 특징을 이루는 7 개의 시스테인 구조중의 제 1 시스테인은 뉴클레오티드 #577 에서 시작된다. 마지막 시스테인은 #879 에서 끝난다. 따라서, 활성 BMP-12 종을 코딩하는 DNA 서열들은 SEQ ID NO : 1 의 뉴클레오티드 #577 내지 #879 를 포함하리라 생각된다.

CHO 세포와 같은 포유류 세포에 의해서 발현되는, BMP-12 는 N-말단이 다양한 BMP-12 단백질의 활성 종의 이질집단분포로 존재하리라 생각된다. 모든 활성 종은 SEQ ID NO : 2 의 아미노산 #3에서 시스테인 잔기로 시작되는 아미노산 서열을 포함하며 적어도 아미노산 103 에서 적어도 시스테인 잔기를 지나거나 아미노산 104 이후의 종결 코돈까지 계속될 것으로 생각된다. 기타 활성 종은 N-말단 방향에 추가의 아미노산 서열을 포함한다. 여기에 추가로 기재한 바와 같이, 포유류 세포에의해 제조된 활성종의 N-말단은 펩티드 서열 Arg-X-X-Arg 를 코딩하는, 컨센서스 절

단 부위의 발생후 시작되리라 생각된다. 따라서, 활성 BMP-12 단백질을 코딩하는 DNA 서열들은 SEQ ID NO : 1 의 뉴클레오티드 #196, 199, 208, 217, 361, 388, 493, 496 또는 571 중의 어느 하나 내지 뉴클레오티드 #879 또는 882에서 시작하는 뉴클레오티드 서열을 포함하는 뉴클레오티드 서열을 갖을 것이로 예상된다.

인간 BMP-12 의 한 활성 종의 N-말단은 다음과 같음이 E.coli 에서의 발현에 의해 실험적으로 결정되었다 : [M]SRXSRKPLHVDF (여기서 X 는 시그널이 분명치 않은 아미노산 잔기를 나타내고, 이는 그 위치에서 시스테인 잔기와 일치한다. 따라서, 이러한 종의 BMP-12 의 N-말단은 SEQ ID NO : 1 의 아미노산 #1 에 있으며, 상기 종의 BMP-12 를 코딩하는 DNA 서열은 SEQ ID NO : 1 의 뉴클레오티드 #571 에서 시작할 수 있을 것으로 보인다. 상기 종의 인간 BMP-12 다이머의 겉보기 분자량은 Novex 16% 트리신 겔상 대략 20-22 kd 인 것으로 SDS-PAGE 에 의해 측정되었다. 인간 BMP-12 단백질은 투명, 무색의 0.1 % 트리플루오로아세트산 용액으로서 존재한다.

전술한 바와 같이, BMP-12 관련 단백질은 BMP-12 및 V1-1 를 포함하는, BMP/TGF-β/Vg-1 패밀리의 단백질의 한 부분 집합으로서, 엄격성이 감소된 조건하에서, 예를들어, PCR, BMP-12 특이 프라이머 (예를 들면, 아래에 기술되는 프라이머 # 6 및 #7) 를 사용하여, 클론 및 확인될 수 가 있는 DNA 서열로 코딩되는 건/인대-류 조직 유도 단백질로 정의될 수가 있다. 본 발명의 DNA 서열은 SEQ ID NO : 1 의 아미노산 #3 내지 #103 을 코딩하는 DNA 와 아미노산 수준에서 적어도 약 80 % 이상의 상동성을 공유하는 것이 바람직하다. 본 발명의 목적상, BMP-12 관련 단백질이라는 용어는 인간 MP52 단백질을 포함하지 아니한다. SEQ ID NO :1 및 3 의 서열 정보 및 제 1 도에 주어진 비교를 이용하여, BMP-12 관련 단백질을 코딩하는 유전자의 클로닝을 가능케해 줄 BMP-12 서열에 대한 프라이머을 설계하는 것은 당해 기술의 범위내에 있다.

본 발명의 BMP-12 관련 단백질의 한 예는, 현재 BMP-13 이라고 부르고 있는 VL-1 이다. 완전한 성숙 BMP-13 서열 및 BMP-13 의 프로펩티드의 적어도 일부의 서열이 SEQ ID NO : 25 에 주어져 있다. BMP-12 와 같이, CHO 세포와 같은 포유류 세포에 의해서 발현되는, BMP-13 은 N-말단이 다양한 활성종의 BMP-13 단백질의 이질 집단분포로 존재하리라 생각된다. 모든 활성종은 SEQ ID NO : 26 의 아미노산 #19 에서 시스테인 잔기로 시작되는 아미노산 서열을 포함하며 적어도 아미노산 119 에서 시스테인 잔기를 지나거나 아미노산 120 이후의 종결 코돈까지 계속될 것으로 예상되고 있다. 그 밖의 단른 활성종은 N-말단 방향에서 추가의 아미노산 서열을 포함한다. 여기에 추가로 기재한 바와 같이, 포유류 세포에 의해 제조된 활성종의 N-말단은 펩티드 서열 Arg-X-X-Arg 를 코딩하는, 컨센서스 절단 부위의 발생후 시작되리라 생각된다. 따라서, 활성 BMP-13 단백질을 코딩하는 DNA 서열들은 SEQ ID NO : 25 의 뉴클레오티드 #410, 458, 602, 605, 또는 659 중의 어느 하나 내지 뉴클레오티드 #961 또는 964 에서 시작되는 뉴클레오티드 서열을 포함하는 뉴클레오티드 서열을 가질 것이로 예상된다.

본 발명에 유용한 정제된 건/인대-류 조직 유도 단백질을 제조하기 위하여는, 적합한 코딩 서열를 포함하는 DNA 서열, 특히 SEQ ID NO : 1 의 뉴클레오티드 #496, #571 또는 #577 내지 #879 또는 #882 의 DNA 코딩 서열로 형질 전환시킨 숙주 세포를 배양하는 단계 ; 그리고 상기 아미노산 서열 또는 SEQ ID NO : 2 의 아미노산 #-25, #1 또는 #3 내지 #103 또는 #104 로 표시된 바의 실질적으로 상동인 서열을 포함하는 단백질을 배양 배지로부터 회수 및 정제하는 단계를 포함하는 방법이 사용된다. 또 다른 구현예에서, 사용된 방법은 적합한 코딩 서열을 포함하는 DNA 서열, 특히 SEQ ID NO : 25 의 뉴클레오티드 #605 또는 #659 내지 #961 또는 #964 의 DNA 코딩 서열로 형질 전환시킨 숙주 세포를 배양하는 단계; 그리고 상기 아미노산 서열 또는 SEQ ID NO : 26 의 아미노산 #1 또는 #19 내지 #119 또는 #120 으로 표시된 바의 실질적으로 상동인 서열을 포함하는 단백질을 배양 배지로부터 회수 및 정제하는 단계를 포함한 것이다.

인간 MP52 DNA 는 여기에 참조로한 WO93/16099 호에 기재되어 있다. 그러나, 이 기록문서는 건/인대-류 조직을 형성할 수 있는 상기 단백질의 능력, 또는 건/인대-류 조직의 유도를 위한 조성물에 사용되는 그 단백질ㅇ 용도를 밝히고 있지 않다. 인간 MP52 는 처음에는 인간 태아 조직의 RNA 를 이용하여 분리되었다. 인간 MP52 뉴클레오티드 서열 (SEQ ID NO : 3) 및 코딩된 아미노산 서열 (SEQ ID NO : 4) 는 여기의 서열표에 밝혀져 있다. MP52 단백질은 SEQ ID NO : 3 의 뉴클레오티드 #845 에서 시작하는 것으로 보이며 SEQ ID NO : 3 의 뉴클레오티드 #1204 를 지나 계속된다. TGF-β 단백질의 특징을 이루는 7 개의 시스테인 구조중의 첫번째 시스테인은 뉴클레오티드 #899 에서 시작된다. 마지막 시스테인은 #1201 에서 끝난다. 그밖의 MP52 단백질의 활성종은 SEQ ID NO : 3 의 뉴클레오티드 #845 의 N-말단 방향에서 추가의 뉴클레오티드를 더 갖을 수도 있다.

본 발명의 정제된 인간 MP52 단백질은 SEQ ID NO : 3 의 뉴클레오티드 #845 내지 #1204 의 DNA 코딩 서열을 포함하는 DNA 서열로 형질 전환된 숙주 세포를 배양하고 , SEQ ID NO : 4 의 아미노산 #1 내지 #120 으로 나타낸 아미노산 서열 또는 실질적으로 상동인 서열을 포함하는 단백질을 배양 배지로부터 회수 및 정제함으로써 제조될 수 있다. 또한 SEQ ID NO : 4 의 아미노산 #17 또는 #19 내지 #119 또는 #120 의 아미노산 서열도 활성을 유지하리라 생각된다. 따라서, 뉴클레오티드 #845, #893 또는 #899 내지 #1201 또는 #1204 의 DNA 서열은 활성 단백질을 코딩하리라 생각된다.

단백질을 포유류 숙주 세포에서 발현시키기 위하여는 그 숙주 세포를 성숙단백질의 코딩 서열에 적합한 해독구조[틀]로 연결되어있는 숙주 세포에 의한 단백질의 분비에 적당한 프로펩티드를 코딩하는 코딩 서열로 형질전환시킨다. 예로서, BMP-2 이외의 포유류 단백질의 전구체 부분을 코딩하는 DNA 를 성숙한 BMP-2 단백질을 코딩하는 DNA에 융합시키는, 여기에서 참조로한, 미합중국 특허 제 5,168,050 호의 개시를 참고한다. 따라서, 본 발명은 건/인대-류 조직을 유도하는 폴리펩티드를 코딩하는 DNA 서열에 정확한 해독구조[틀]로 연결되어 있는, TGF-β 슈퍼패밀리의 단백질들중의 하나의 프로펩티드를 코딩하는 DNA 서열을 포함하는 키메릭 DNA 분자를 포함한다. 용어 '키메릭' 은 프로펩티드가 코딩된 성숙 폴리펩티드외의 기타 폴리펩티드로부터 발생한 것임을 나타내는데 사용한다. 물론, 숙주 세포는 SEQ ID NO : 2, SEQ ID NO : 4, 또는 SEQ ID NO : 26 에 나타낸 성숙한 단백질을 코딩하는 코딩 서열에 정확한 해독틀로 연결된 천연 프로펩티드를 코딩하는 DNA 서열 코딩 서열로 형질 전환될 수 있다. 천연 프로펩티드의 완전한 서열은 SEQ ID NO : 1, SEQ ID NO : 3 또는 SEQ ID NO : 25 에 개시되어 있는 서열을 이용하여 당해 기술에서 공지된 방법으로 결정하여, 모든 클론을 확인 및 분리하는데에 적합한 탐식자를 설계할 수 있다.

본 발명은 또한 기타 단백질성 물질을 코딩하는 DNA 서열과 관련없이, 건/인대-류 조직 유도 단백질의 발현을 코딩하는 새로운 DNA 서열을 포함한 것이기도 하다. 상기 DNA 서열은 SEQ ID NO : 1 에서 5' → 3' 방향으로 묘사된 서열 및 엄격한 혼성화 조건 [예를 들면, 65℃ 에서의 0.1X SSC, 0.1% SDS ; T. Maniatis et al, Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory (1982), p 387 내지 389 참조] 하에서 이 서열에 하이브리다이즈하고 건/인대-류 조직 유도 활성을 갖는 단백질을 코딩하는 서열을 포함한다.

마찬가지로, SEQ ID NO : 1 또는 SEQ ID NO : 25 의 서열로 코딩되는 단백질, 또는 SEQ ID NO : 2 또는 SEQ ID NO : 26 의 아미노산 서열을 포함하나, 유전 암호의 중첩성 또는 대립유전자의 변형(아미노산이 변화되거나 또는 변화되지 안하는 결과를 가져올 수 있는 개체군에서의 자연발생의 염기 변화) 에 기인하여 코돈 서열이 상이한 단백질을 코딩하는 DNA 서열도 또한 여기에 기재된 건/인대-류 조직 유도 단백질을 코딩한다. 코딩된 폴리펩티드의 활성, 반감기 또는 생산을 향상시키기 위하여 점 돌연변이 또는 유도 변형 (삽입, 결실 및 치환을 포함) 으로 유발되는 SEQ ID NO : 1 또는 SEQ ID NO : 25 의 DNA 서열상의 변이도 또한 본 발명에 포함된다.

본 발명의 또 다른 면은 건/인대-류 조직을 유도하는 단백질을 제조하는 새로운 방법을 제공한다. 본 발명의 방법은 공지의 조절 서열의 조절하에서, 본 발명의 단백질을 코딩하는 DNA 서열로 형질 전환된 적합한 세포주의 배양을 포함한다. 형질 전환된 숙주 세포를 배양하고 그 배양 배지로부터 상기 단백질을 회수 및 정제한다. 정제된 단백질에는 부산물로 발생하는 기타 단백질 및 기타 오염물이 실질적으로 없다.

적합한 세포 또는 세포주는 차이니즈 햄스터 난소 세포(CHO) 와 같은 포유류 세포일 수 있다. 전술한 바와 같이, 포유류 세포에서의 단백질의 발현은 단백질의 분비를 확실히 하기위하여 적당한 프로펩티드를 필요로한다. 적당한 포유류 숙주 세포의 선택 및 형질전환, 배양, 증폭, 스크리닝, 산물 생산 및 정제 방법은 당해기술에 공지되어 있다. 예로서, Gething & Sambrook 의 Nature ,293: 620-625 (1981), 또는 Kaufman et al 의 Mol. Cell. Biol .,5(7) : 1750-1759 (1985) 또는 Howley et al 의 U.S. Patent 4,419,446 참조. 수반된 실시예에 기재되어 있는, 기타 적합한 포유류 세포주는 원숭이 COS-1 세포주이다. 또한 포유류 세포 CV-1 가 적합할 수 있다.

또한 박테리아 세포도 적합한 숙주일 수 있다. 예컨대, E.coli 의 다양한 균주 (예, HB101, MC1061) 는 생명공학분야에서 숙주 세포로 공지되어 있다. 또한 비. 서브틸리스 (B.subtilis), 슈도모나스(Pseudomonas), 기타 바실루스등의 다양한 균주를 상기의 방법에서 사용할 수 있다. 박테리아 세포에서의 단백질의 발현을 위하여, 프로펩티드를 코딩하는 DNA 는 불필요하다.

TGF-β 패밀리의 단백질들을 포함하여, 포유류 단백질의 박테리아에서의 발현은 비글리코실화 형태, 및 세포 봉입체로 알려진 불용성 펠릿 형의 단백질을 제조하는 것으로 알려져 있다. 상기 세포 봉입체의 가용화 기술, 카오트로픽 (chaotropic) 제를 사용한 단백질의 변성기술 및 가용형의 단백질생산을 가능케하는 충분히 정확한 단백질의 리폴딩 (refolding) 기술은 본 발명분야에서 기술되었다. 예로서, 여기에서 참조로한 EP 0433225 호의 개시를 참조한다.

또는, 목적 단백질이 융합 파트너와의 융합 단백질로 발현되는, 유전자 융합 단백질의 밟현등과 같은, 세포 봉입체 형성을 우회하는 방법이 고안되었다. 융합 단백질은 후에 절단되어서 목적 단백질을 내놓게된다. E.coli 의 상기 유전자 융합 발현 시스템의 한 예는 융합 파트너로 E.coli 티오레독신 유전자의 이용을 기초로한다. 여기에서 참조로한 문헌 (LaVallie et al., Bio/Technology , 11:187-193 (1993)) 를 참조한다.

또한 당업자에게 공지된 이스트 세포의 많은 균주는 본 발명의 폴리펩티드의 발현을 위한 숙주 세포로 이용할 수 있다. 또한, 원하는 경우, 곤충 세포는 본 발명의 방법에서 숙주 세포로 이용할 수 있다. 예를 들면, (Miller et al 의 Genetic Engineering , 8:277-298 (Plenum Press 1986)) 과 그 안에 인용된 참고문헌들을 보라.

본 발명의 다른 양상은 상기 건/인대-류 조직 유도 단백질의 발현 방법에서 유용한 벡터를 제공하는 것이다. 바람직하게는 벡터는 본 발명의 새로운 인자를 코딩하는 전술한 완전한 크기의 새로운 DNA 서열을 포함한다. 또한, 벡터는 단백질 서열의 발현을 가능케하는 적당한 발현 조절 서열을 포함한다. 또는, 전술한 바와같은 부분변형 서열을 연결시킨 벡터들은 본 발명의 구현예들 이기도 하다. 또한, SEQ ID NO : 1 또는 SEQ ID NO : 3 또는 SEQ ID NO : 25 의 서열은 프로펩티드 서열을 결실시키고 이를 BMP 단백질 또는 TGF-β 슈퍼패밀리의 단백질의 완전한 프로펩티드를 코딩하는 서열로 대체하는 조작으로 성숙한 단백질을 발현시킬 수 있다. 따라서, 본 발명은 SEQ ID NO : 2 또는 SEQ ID NO : 4 또는 SEQ ID NO : 26 의 아미노산 서열을 갖는 단백질을 코딩하는 DNA 서열에 정확한 해독틀로 연결된 TGF-β 슈퍼패밀리의 한 단백질의 프로펩티드를 코딩하는 키메릭 DNA 분자를 포함한다. 벡터는 세포주의 형질 전환방법에서 사용 가능하고, 선별된 숙주 세포에서 복제 및 발현을 지시할 수 있는 본 발명의 DNA 코딩 서열과 연계하여 기능하는 선별된 조절 서열을 포함할 수 있다. 이러한 벡터를 위한 조절 서열은 당업자에게 공지되어 있고, 숙주 세포에 따라 선택할 수있다. 이러한 선별은 통상적인 조작이며 본 발명의 일부가 되는 것은 아니다.

정상적으로 형성되지 못하는 환경하에서 건/인대-류 조직 또는 기타 조직 형성을 유도하는 본 발명의 단백질은 인간 및 기타 동물에서의 건 또는 인대 파열, 변형 및 기타 건 또는 인대의 결함의 치료에 적용되었다. 이러한 건/인대-류 조직 유도 단백질을 이용한 제제는 건 또는 인대의 기타 조직의 고정을 향상시키는 용도 및 건 또는 인대 조직의 결함을 치료하는 용도 뿐만아니라, 건 또는 인대 조직에 대한 손상을 예방하는 예방용도ㄹ 가질 수 있다. 본 발명의 조성물로 유도되는 건/인대-류 조직 신생 형성 (de novo formation) 은 선천성의, 유도된 외상, 또는 다른 원인의 기타 건 또는 인대 결함의 치료에 기여하며, 또한 건 또는 인대의 부착 또는 치유 목적의 성형 외과에 유용한 것이기도 하다. 또한 본 발명의 조성물은 건염, 수근관 증후군 및 기타 건 또는 인대 결함의 치료에 유용할 수 있다. 또한 본 발명의 조성물은 건 및/또는 인대 조직을 치료하거나 재생시키는 것이 바람직한 그 밖의 다른 징후들에도 유용할 수 있다. 이와 같은 징후들에는, 건염에서 발생하는 것과 같은 치근막 인대 손상의 재생 또는 복구,그리고 건-골 결합의 재생 또는 복구 등을 포함하며, 이것들에 한정되는 것은 아니다. 본 발명의 조성물은 건- 또는 인대-형성 세포를 유도하고, 건- 또는 인대-형성 세포의 성장을 자극하거나 건- 또는 인대-형성 세포의 근원세포터 (progenitor) 의 분화를 유도하는 환경을 제공할 수 있다.

BMP-12 관련 단백질은 배양 배지로부터 회수될 수 있으며 함께 나오는 기타 단백질성 물질 및 기타 존재하는 불순물로 부터 이들을 분리 정제할 수 있다. 본 발명의 단백질은 건/인대-류 조직의 형성을 유도할 수 있다. 또한 상기 단백질은 후술되는 쥐 전위 이식 검정에서 건/인대-류 조직 형성 활성을 나타내는 능력의 특징이 있다. 상기 단백질은 또한 인대와 같은 기타 유형의 조직의 형성을 유도하는 능력을 갖을 수 있다는 것이 예상된다.

또한 여기에 제공된 건/인대-류 조직 유도 단백질은 SEQ ID NO : 1 또는 SEQ ID NO : 25 의 것과 유사한 서열로 코딩되는 인자를 포함하나, 상기 서열에는 부분변형이 자연적으로 제공되거나 (예, 폴리펩티드상의 아미노산 변화로 결과될 수 있는 뉴클레오티드 서열상의 대립유전자 변이) 또는 의도적으로 조작되어 있다. 예컨대, 합성 폴리펩티드는 SEQ ID NO : 2 의 아미노산 잔기의 연속 서열을 전부 또는 일부 복제할 수 있다. 상기 서열은 SEQ ID NO : 2 의 건/인대-류 조직 성장 인자 폴리펩티드와 1차, 2차, 또는 3 차 구조 및 형태적 특징들을 공유하는 덕택으로, 그것과 더불어 공통적인 건/인대-류 또는 기타 조직 성장 인자의 생물학적 특성을 가질 수 있다. 따라서, 치료 조성물 및 치료 과정에서 천연의 건/인대-류 조직 유도 폴리펩티드의 생물학적 활성 대체물로 사용할 수 있다.

여기에 기재된 건/인대-류 조직 유도 단백질의 서열의 기타 특수한 돌연변이에는 글리코실화 부위의 부분변형이 포함된다. 이러한 부분변형에는 O-결합 또는 N-결합 글리코실화 부위가 포함될 수 있다. 일례로, 글리코실화의 부재 또는 오직 일부의 글리코실화는 아스파라진-결합 글리코실화 인식 부위에서의 아미노산 치환 또는 결실에서 발생한 것이다. 아스파라진-결합 글리코실화 인식 부위는 적합한 세포성 글리코실화 효소가 특이하게 인식하는 트리펩티드 서열들을 포함한다. 이들 트리펩티드 서열은 아스파라진-X-트레오닌, 아스파라진-X-세린 또는 아스파라진-X-시스테인일 수 있고, X 는 통상적으로 프롤린을 제외한 임의의 아미노산이다. 글리코실화 인식 부위의 제 1 또는 제 3 아미노산 위치의 하나 또는 양쪽에서의 다양한 아미노산 치환 또는 결실(및/또는 제 2 위치에서의 아미노산 결실)은 부분변형된 트리펩티드 서열의 비글리코실화를 초래한다. 또한 박테리아에 의한 단백질의 발현은 설령 글리코실화 부위가 부분변형되지 아니한 상태로 남아있더라도, 비글리코실화 단백질의 생산으로 결과될 것이다.

본 발명의 조성물은 SEQ ID NO : 1 또는 SEQ ID NO : 25 의 DNA 서열로 형질전환된 세포를 배양하고 이 배양 배지로부터 SEQ ID NO : 2 또는 SEQ ID NO : 26 의 아미노산 서열을 갖는 단백질을 회수 및 정제하여 제조할 수 있는 정제 BMP-12 관련 단백질을 포함한 것이다. 정제된 발현 단백질에는 함께 나오는 기타 단백질성 물질 및 기타 불순물이 실질적으로 없다. 회수된 정제 단백질은 건/인대-류 조직 형성 활성, 및 인대 재생과 같은 기타 조직 성장 활성을 보이리라 예상된다. 또한 본 발명의 단백질은 후술될 쥐 검정에서의 건/인대-류 조직 형성 활성을 나타내는 능력의 특징이 있다.

본 발명의 건/인대-류 조직 형성을 유도하는 조성물은 유효량의 건/인대-류 조직 유도 단백질을 함유할 수 있으며, 상기 단백질은 SEQ ID NO : 2 의 아미노산 서열, 바람직하게는 SEQ ID NO : 2 의 아미노산 #-25, #1 또는 #3 내지 #103 또는 #104 ; 또는 SEQ ID NO : 26 의 아미노산 #1 또는 #19 내지 #120 ; 뿐만아니라 건 및/또는 인대-류 조직을 형성할 수 있는 능력을 보이는, SEQ ID NO : 2 또는 SEQ ID NO : 26 의 돌연변이체 및/또는 변형체를 포함한다.

또한 본 발명의 조성물은 액티빈(activin) 등의, TGF-β 슈퍼패밀리의 추가된 단백질과 같은, 추가의 단백질을 더 포함할 수 있다. 본 발명의 또 다른 양상은 약학적으로 허용가능한 부형제 또는 담체내에, BMP-12 또는 VL-1과 같은, 치료상 유효한 양의 건/인대- 유도 단백질을 함유하는 약제 조성물을 제공한다. 이 조성물은 건/인대-류 조직 또는 기타 조직의 형성하는데 사용할 수 있다. 또한 상기 조성물은 건 및 인대의 복구, 상처 치유 그리고 피부복구 등의 기타 조직 복구에 사용할 수 있을 것으로 생각된다. 또한 본 발명의 단백질은 뉴론 생존을 증가시킬 수 있으며 따라서 이식 및 뉴론의 생존 감소가 보이는 증상의 치료에 유용하리라 생각된다. 또한 본 발명의 조성물은 예로서 미합중국 특허 제 5,108,922 ; 5,013,649 ; 5,116,738 ; 5,106,748 ; 5,187,076 ; 및 5,141,905 호에 개시된, BMP 단백질들인 BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6 및 BMP-7 ; PCT 공개공보 WO91/18098 에 개시된 BMP-8 ; PCT 공개공보 WO93/00432 에 개시된, BMP-9 ; 및 1993, 5, 12 일자로 출원된 시리얼 번호 08/061,695 및 08/061,464 의, 동사계류중인 특허출원에 개시되어 있는 BMP-10 또는 BMP-11 과 같은, 적어도 하나의 기타 치료학적 유용 제제를 포함할 수 있다. 전술한 문헌의 개시는 여기에서 참조로 하였다.

본 발명의 조성물은 BMP-12 또는 VL-1 (BMP-13) 과 같은 건/인대-유도 단백질외에도, MP52, 표피 성장 인자(EGF), 섬유아세포 성장 인자(FGF), 혈소판 유래 성장 인자(PDGF), 형질전환 성장 인자(TGF-α 및 TGFβ), 및 섬유아세포 성장인자-4 (FGF-4), 부갑상선 호르몬(PTH), 백혈병 억제 인자(LIF/HILDA/DIA), 인슐린류 성장 인자(IGF-I 및 IGF-II) 등을 포함하는 기타 치료상 유용한 제제를 포함할 수도 있다. 또한 상기 제제들의 일부를 본 발명의 조성물에 사용할 수 있다. 예로서, 함께 주입된 BMP-2 및 BMP-12 를 포함하는 조성물은 골과및 건/인대-류 조직 둘다를 발생시킨다. 이러한 조성물은 건, 인대 및 골이 해부학적으로 인접한 위치에서 동시에 형성되는 배 관절의 결함을 치료하고, 건이 골에 부착된 위치에서 조직을 재생시키는데 유용할 수 있다. 또한 본 발명의 조성물은 피부 치유 및 관련 조직 복구등의 상처 치유에 유용할 것으로 생각된다. 상처의 유형에는 화상, 절개 및 궤양등이 있으나 이에만 한정되는 것은 아니다 (예를 들면, 상처 치유 및 관련 조직 복구의 논의에 대한 PCT 공개공보 WO84/01106 를 보라).

본 발명의 단백질은 기타 관련 단백질 및 성장 인자와 협동적으로 또는 아마도 상승작용적으로 작용하리라 기대된다. 그러므로 본 발명의 치료 방법 및 조성물은 치료량의 본 발명의 1 종이상의 단백질과 치료량의 전술한 1 종이상의 BMP 단백질을 포함한 것이다. 이러한 조성물은 BMP 단백질의 별개의 분자 또는 상이한 BMP 부분으로 구성된 헤테로분자를 포함할 수 있다. 예로서, 본 발명의 방법 및 조성물은BMP-12 관련 단백질 소단위체 및 전술한 'BMP' 단백질중의 하나로부터의 소단위체를 포함하는 이황화물 결합 다이머를 포함할 수 있다. 따라서, 본 발명은 한 소단위체는 SEQ ID NO : 2 의 아미노산 #1 내지 아미노산 #104 의 아미노산 서열을 포함하고, 한 소단위체는 BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6, BMP-7, BMP-8, BMP-9, BMP-10 및 BMP-11 로 구성된 군으로부터 선택한 골 형태발생 단백질의 아미노산 서열을 포함하는 헤테로다이머인 정제된 BMP-12 관련 폴리펩티드를 함유하는 조성물을 포함한다. 추가의 구현예는 BMP-12, VL-1(BMP-13) 또는 MP52 와 같은 이황화물 결합 건/인대-류 조직 유도 부분의 헤테로다이머를 포함한다. 예로서 헤테로다이머는 SEQ ID NO : 2 의 #1 내지 #104 의 아미노산 서열을 포함하는 소단위체 하나를 함유할 수 있으며 다른 하나의 소단위체는 SEQ ID NO : 4 의 #1 내지 #120 또는 SEQ ID NO : 26 의 #1 내지 #120 의 아미노산 서열을 포함할 수 있다. 또한, 본 발명의 조성물은 결함, 상처 또는 문제의 조직의 치료에 유용한 기타 제제와 배합할 수 있다.

pH, 등장성 , 안정성등을 당연히 고려하는, 생리학적으로 허용가능한 단백질 조성물의 제조 및 제형은 당해 기술로 가능하다. 또한 치료 조성물은 현재 TGF-β 단백질의 종 특이성의 결핍으로인해 가축병 치료의 적용에 유용하다. 인간 이외에도 특히 가축 및 순종마가 본 발명의 조성물로 치료하기위한 바람직한 치료대상이다.

치료 방법은 조성물을 주입물 또는 장치로서 국소적으로, 전신적으로, 또는 국부적으로 투여하는 방법등이 있다. 투여시, 본 발명용 치료 조성물은, 물론, 피로겐이 없는, 생리학적으로 허용가능한 형태이다. 더우기, 조성물은 바람직하게는 조직 손상의 부위에 전달하기위하여 점성 형태로 캡슐화 또는 주입되는 것이 좋다. 국소 투여는 상처 치유 및 조직 복구에 적합할 수 있다. 전술한 바의 조성물내에 선택적으로 포함될 수도 있는 단백질들이외의 다른 치료학적으로 유용한 제제들은 본 발명의 방법의 조성물과 동시에 또는 순차적으로 투여될 수 있다.

또한 조성물은 적당한 매트릭스 및/또는 봉쇄제를 담체로서 포함할 수 있다. 예로서, 매트릭스는 조성물을 지지하거나 건/인대-류 조직 형성 및/또는 기타 조직 형성용 표면을 제공할 수 있다. 매트릭스는 단백질의 느린방출 및/또는 단백질을 내놓기 위한 적당한 환경을 제공할 수 있다. 봉쇄제는 주사 또는 기타 수단을 통한 투여를 용이하게 하는데 도움이 되는 물질이거나, 적용부위로부터 단백질의 이동을 늦출수 있다.

담체 물질의 선택은 생체적합성, 생분해성, 기계적 성질, 표면 외관 및 계면 특성에 기초한다. 조성물의 특정 사용은 적합한 제형을 규정할 것이다. 조성물의 잠재가능한 매트릭스는 생분해성이고 화학적으로 규정될 수 있다. 그리고 또 다른 매트릭스는 순수한 단백질 또는 세포외 매트릭스 성분으로 구성된다. 기타 가능한 매트릭스는 생분해성이아니며 화학적으로 규정되어 있다. 바람직한 매트릭스로는 헬리스탯 스폰지 (뉴저지, 플레인스보로, 인테그라 라이프사이언스)와 같은 콜라겐을 주성분으로 한 물질, 또는 주사제 형태의 콜라겐, 뿐만아니라 생분해성일 수 있고 알킬셀룰로오스성 물질 등을 포함할 수 있는 봉쇄제 등이 있다.

또다른 바람직한 종류의 담체는 폴리(락트산)의 중합체, 폴리(글리콜산)의 중합체 및 락트산 및 글리콜산의 공중합체를 포함하여, 다공성 입자 중합체 매트릭스이다. 이들 매트릭스는 또한 봉쇄제를 포함할 수도 있다. 적합한 중합체 매트릭스는 예를 들면, WO 93/00050 에 기재되어 있으며, 그것의 상세한 내용은 여기에 참고로 하고 있다.

바람직한 군의 봉쇄제는 메틸셀룰로오스, 에틸셀룰로오스, 히드록시에틸셀룰로오스, 히드록시프로필셀룰로오스, 히드록시프로필메틸셀룰로오스 및 카르복시메틸셀룰로오스를 포함하여, 알킬셀룰로오스 (히드록시알킬셀룰로오스등을 포함) 와 같은 셀룰로오스성 물질이며, 가장 바람직한 것은 카르복시메틸셀룰로오스(CMC) 의 양이온염이다. 기타 바람직한 봉쇄제로는 히알우론산, 나트륨 알기네이트, 폴리(에틸렌 글리콜), 폴리옥시에틸렌 옥시드, 카르복시비닐 중합체 및 폴리(비닐 알코올) 등이 있다. 여기에서 유용한 봉쇄제의 양은 총 중량을 기준으로 0.5-20 중량%, 바람직하게는 1-10 중량% 이고, 이는 중합체 매트릭스로부터 단백질의 탈착을 방지하고 조성물의 적당한 조작을 제공하기위하여 필요한 양이나, 근원 세포가 매트릭스에 침투하지 못하게 함으로써 단백질이 근원 세포의 활성을 보조하는 기회를 가질 수 있는 정도의 양은 아니다.

본 출원의 실제에 유용한 추가의 선택 성분의 예로는 만니톨, 수크로오스, 락토오스, 글루코오스 또는 글리신(동결 동안 단백질이 분해되는 것을 보호하기 위하여) 과같은 냉동 보호제, 메틸 및 프로필 파라벤 및 벤질 알코올 같은 항세균 방부제; EDTA, 시트레이트 및 BHT(부틸화 히드록시톨루엔)과 같은 산화방지제 ; 및 폴리(소르베이트) 및 폴리(옥시에틸렌) 과 같은 계면활성제등이다.

전술한 바와같이, 본 발명의 조성물은 건 또는인대 손상의 분야에서 건/인대-류 조직의 재생과 같은, 수많은 건 결함을 치료하기위한 방법에서 사용하여, 건 조직, 인대의 파열 및 각종 유형의 다른 조직 결함 또는 상처의 복구에 도움을 줄 수 있다. 본 발명에 따른 상기 방법은 건/인대-류 조직 또는 기타 조직의 복구가 필요한 환자에게, SEQ ID NO : 2, SEQ ID NO : 4 및/또는SEQ ID NO : 26 에 기재된 것과 같은 유효량의 건/인대-류 조직 유도 단백질을 함유하는 조성물을 투여하는것을 수반한다. 또한 상기 방법은 전술한 1 종이상의 BMP 단백질과 함께 건/인대-류 조직 유도 단백질을 투여하는 것을 수반한다.

또 다른 구현예로, 본방법은 단량체중의 하나가 BMP-12, VL-1(BMP-13) 또는 MP52와 같은, 건/인대-류 조직 유도 폴리펩티드이고 제 2 단량체는 TGF-β 슈퍼패밀리의 성장 인자들중의 하나인 헤테로다이머성 단백질의 투여를 수반할 수 있다. 또한, 상기 방법은EGF, FGF, TGF-α, TGF-β, 및 IGF 를 포함하는 기타 성장 인자와 함께 건/인대-류 조직 유도 단백질을 투여하는 것을 포함할 수 있다.

따라서, 본 발명의 또 하나의 면은 건/인대-류 조직의 복구, 건 또는 인대의 복구 그리고 건염 및 건 또는 인대의 결함에 관련된 기타 증상의 치료를 위한 치료 방법 및 조성물이기도 하다. 이와 같은 조성물은 약학적으로 허용가능한 부형제, 담체 또는 매트릭스와 배합한 상태로, BMP-12, BMP-12 관련 단백질, 또는 MP52 와 같은, 치료학적으로 유효한 양의 1 종이상의 건/인대-류 조직 유도 단백질을 포함한다.

처방 용량은 조성물의 작용을 변경시키는 각종 요인, 예컨대, 형성시키고자하는 건 또는 인대 조직의 양, 건 또는 인대 손상의 부위, 손상된 건 또는 인대의 증상, 상처의 크기, 손상된 조직의 유형, 환자의 연령, 성 및 식이법, 임의 감염의 심각성, 투여 시간 및 기타 임상 요인등을 고려하여 주치의에의해 결정될 것이다. 용량은 재구성에 사용하는 매트릭스의 유형과 조성물내의 추가 단백질의 유형에 따라 변동될 수 있다. IGF-I(인슐린류 성장 인자 I) 와 같은 기타 공지 성장 인자를 최종 조성물에 첨가하는 것도 또한 용량에 영향을 미칠 수 있다.

진행 과정은 건/인대-류 조직 형성, 또는 건 또는 인대의 성장 및/또는 복구의 주기적 분석으로 조사할 수 있다. 진행 과정은 당해 기술의 공지된 방법, 예컨대 X-레이, 관절경검사(athroscopy), 조직형태학적 결정 및 테트라시클린 표지 등의 방법으로 조사할 수 있다.

다음의 실시예들은 인간 건/인대-류 조직 유도 단백질의 회수 및 특성구명과 이를 이용하여 기타 건/인대-류 조직 유도 단백질을 회수, 상기 인간 단백질의 획득, 재조합 기술을 이용한 단백질의 발현에 있어서의 본발명의 실제, 그리고 생체내 모델에서 건/인대-류 조직을 형성할 수 있는 본 발명의 조성물의 능력의 증명을 구체적으로 설명한 것이다. 비록 실시예가 BMP-12 에 관련하여 본 발명을 설명하고 있다 해도, 당해 기술의 범위 내에서 약간의 부분변형을 가하면, MP52 및 VL-1 에대해서도 동일한 결과가 얻이질 수 있는 것으로 믿어진다.

실시예 1실시예 1

DNA 의 분리

BMP-12 및 BMP-12 관련 단백질을 코딩하는 DNA 서열은 당업자에게 공지된 다양한 방법으로 단리할 수 있다. 후술된 바와 같이, 올리고뉴클레오티드 프라이머는 기타 BMP 단백질, Vg-1 관련 단백질 및 TGF-β 슈퍼패밀리의 기타 단백질내에 존재하는 아미노산 서열을 기초로하여 설계할 수 있다. BMP 패밀리 단백질 및 TGF-β 슈퍼패밀리의 단백질들중 기타 단백질의 범위내에서 잘 보존되어 있는 아미노산 서열을 포함한 부위는 확인할 수 있으며 이들 잘 보존된 부위의 컨센서스 아미노산 서열은 개별적인 BMP/TGF-β/Vg-1 단백질의 대응하는 부위의 유사성을 기초로하여 구축할 수 있다. 이러한 컨센서스 아미노산 서열의 한 예는 하기에 나타낸다.

컨센서스 아미노산 서열 (1) :

Trp-Gln/Asn-Asp-Trp-Ile-Val/Ile-Ala (SEQ ID NO : 16)

(여기서 X/Y 는 둘중 어느 하나의 아미노산 잔기가 그 위치에 나타날 수 있음을 나타냄).

하기의 올리고뉴클레오티드는 앞서 확인된 컨센서스 아미노산 서열 (1) 에 기초하여 설계한다 :

#1 :CGGATCCTGGVANGAYTGGATHRTNGC (SEQ ID NO : 17)

상기 올리고뉴클레오티드 서열은 자동화 DNA 합성기로 합성한다. 앞서 확인된 올리고뉴클레오티드 프라이머에서의 표준 뉴클레오티드 기호는 하기와 같다 : A, 아데노신 ; C, 시토신 ; G, 구아닌 ; T, 티민 ; N, 아데노신 또는 시토신 또는 구아닌 또는 티민 ; R, 아데노신 또는 시토신 ; Y, 시토신 또는 티민 ; H, 아데노신 또는 시토신 또는 티민 ; V, 아데노신 또는 시토신 또는 구아닌 ; D, 아데노신 또는 구아닌 또는 티민.

올리고뉴클레오티드 #1 의 처음 7 개의 뉴클레오티드 (밑줄친) 는 BMP-12 단백질을 코딩하는 특정적으로 증폭된 DNA 서열의 조작을 촉진시키기위한 제한 엔도뉴클레아제 BamHI 의 인식 서열을 포함하고 따라서 상기의 컨센서스 아미노산 서열 (1) 에서 유도된 것은 아니다.

제 2 컨센서스 아미노산 서열은 후술된 바와 같이 BMP/TGF-β/Vg-1 단백질의 또 다른 잘 보존된 부위에서 유래된 것이다 :

His-Ala-Ile-Val/Leu-Gln-Thr (SEQ ID NO : 18)

하기의 올리고뉴클레오티드는 상기 확인된 컨센서스 아미노산 서열 (2) 를 기초로하여 설계한다 :

2 :TTTCTAGAARNGTYTGNACDATNGCRTG (SEQ ID NO : 19)

상기 올리고뉴클레오티드 서열은 자동 DNA 합성기로 합성한다. 상기한 바와 똑같은 뉴클레오티드 기호들이 사용된다. 올리고뉴클레오티드 #1 의 처음 7 개의 뉴클레오티드 (밑줄친) 는 BMP-12 단백질을 코딩하는 특이적으로 증폭된 DNA 서열의 조작을 촉진시키기위한 제한 엔도뉴클레아제 XbaI 의 인식 서열을 포함하고 따라서 전술한 컨센서스 아미노산 서열 (2) 에서 유래된 것은 아니다.

본 발명의 BMP-12 단백질 및 기타 BMP/TGF-β/Vg-1 관련 단백질은 전술한 컨센서스 아미노산 서열과 유사한 아미노산 서열을 포함할 수 있고 BMP-12 단백질 또는 기타 신규 관련 단백질내의 상기 서열의 위치는 그것들이 유래한 단백질내의 상대적인 위치에 대응하리라 생각된다. 또한 기타 BMP/TGF-β/Vg-1 단백질의 구조 및, 각각 컨센서스 아미노산 서열 (1) 및 (2) 로부터 유래하였던 올리고뉴클레오티드 서열 #1 및 #2 에서 도출된 상기 위치에 관한 정보는 BMP-12 단백질 또는 기타 BMP/TGF-β/Vg-1 관련 단백질의 대응하는 아미노산을 코딩하는 DNA 서열을 특정하게 증폭시키기는데에 이용할 수 있을 것으로 생각된다.

BMP/TGF-β/Vg-1 단백질의 유전자 구조의 지식을 기초로하면, 또한 인간 게놈 DNA 를 주형으로 사용하여 BMP-12 BMP/TGF-β/Vg-1 (BMP-12 관련 단백질) 을 코딩하는 서열을 확인케하는 특정의 증폭 반응을 행할 수 있으리라 생각된다. 이러한 인간 게놈 DNA 주형의 특정의 증폭 반응은 전술한 올리고뉴클레오티드 프라이머 #1 및 #2 의 사용으로 개시할 수 있다. 앞서 확인된 올리고뉴클레오티드 #1 및 #2 를 프라이머로 사용하여 인간 게놈 DNA 의 특정 뉴클레오티드 서열의 특정한 증폭을 가능케 할 수 있다. 증폭 반응은 하기와 같이 행한다 :

제네틱스 인스티튜트사의 켄 쟈콥에의해 제공된, 인간 게놈 DNA (출처 : 말초혈액 림프구) 를 25 게이지 니들을 통한 찌르기의 반복으로 전단하고, 100℃에서 5 분간 변성시킨다음 얼음에서 차게 한후에 각각의 데옥시뉴클레오티드 트리포스페이트 (dATP, dGTP, dCTP 및 dTTP) 200 μM, 10 mM Tris-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl2, 0.001 % 젤라틴, Taq DNA 폴리머라아제 1.25 단위, 100 pM 올리고뉴클레오티드 #1 및 100 pM 올리고뉴클레오티드 # 2 를 함유하는 반응 혼합물에 첨가한다. 상기 반응 혼합물을 94℃ 에서 2 분간 보온 반응시키고 이어서 하기 방법으로 열 순환을 시킨다 : 3 주기동안 94℃에서 1 분, 40℃에서 1 분, 72℃에서 1 분 ; 이어서 37 주기동안 94℃ 에서 1 분, 55℃에서 1 분, 72℃에서 1 분, 이어서 72℃에서 10 분간 보온 반응.

상기 반응으로 특별하게 증폭된 DNA를 에탄올 침강시키고, 제한 엔도뉴클레아제 BamHI 및 XbaI 로 소화시키며 아가로오즈 겔 전기영동시킨다. BMP-12 또는 기타 BMP/TGF-β/Vg-1 을 코딩하는 DNA 단편의 예상한크기에 해당하는 겔의부위를 절단하고 거기에 함유된 특정의 증폭 DNA 단편을 전기용리시키고 플라스미드 벡터 pGEM-3 의 폴리링커의 XbaI와 BamHI 부위 사이에 서브클론시킨다. 생성된 BMP-12 관련 서브클론중의 하나의 DNA 서열 분석은, 여기에 포함된 특정하게 증폭된 DNA 서열 산물은 본 발명의 BMP-12 단백질의 부분을 코딩하는 것을 보여준다.

이 BMP-12 의 특정의 증폭된 DNA 단편의 DNA 서열 (SEQ ID NO : 5) 및 유래된 아미노산 서열 (SEQ ID NO : 6) 은 서열표에 나타내었다.

SEQ ID NO : 5 의 뉴클레오티드 #1-#26 은 올리고뉴클레오티드 #1 의 일부분을 포함하고 뉴클레오티드 #103-#128 은 특정의 증폭 반응을 수행하기위하여 이용하는 올리고뉴클레오티드 #2 의 역상보의 일부분을 포함한다. 증폭 반응의 개시에서 올리고뉴클레오티드 #1 및 #2 의 기능으로 인하여, 이들은 BMP-12 단백질을 코딩하는 실제의 서열과 정확하게 일치하지 않을 수 있어 상응하는 아미노산 유도체 (SEQ ID NO : 6) 로 번역되지않는다.

다른 서브클론의 DNA 서열 분석은 거기에 포함된 특정의 증폭된 DNA 산물은 VL-1 로 칭한 본 발명의 또 하나의 BMP/TGF-β/Vg-1 (BMP-12관련) 단백질의 일부분을 코딩하는 것을 보여준다.

이 특정의 증폭된 DNA 단편의 DNA 서열 (SEQ ID NO : 7) 및 유래된 아미노산 서열 (SEQ ID NO : 8) 은 서열표에 나타내었다.

SEQ ID NO : 7 의 뉴클레오티드 #1-#26 은 올리고 뉴클레오티드 #1 의 일부를 포함하고 뉴클레오티드 #103-#128 은 특정의 증폭 반응을 수행하기위하여 이용하는 올리고뉴클레오티드 #2 의 역상보의 일부분을 포함한다. 증폭 반응의 개시에서 올리고뉴클레오티드 #1 및 #2 의 기능으로 인하여, 이들은 본 발명의 VL-1 단백질을 코딩하는 실제의 서열과 정확하게 일치하지 않을 수 있어 상응하는 아미노산 유도체 (SEQ ID NO : 8) 도 번역되지 않는다.

하기 올리고뉴클레오티드 탐식자는 전술한 특정의 증폭된 BMP-12 인간 DNA 서열 (SEQ ID NO : 5) 에 기초하여 설계하고 자동 DNA 합성기로 합성한다 :

3 : CCACTGCGAGGGCCTTTGCGACTTCCCTTTGCGTTCGCAC (SEQ ID NO : 20)

상기 올리고뉴클레오티드 탐식자는 32P 로 방사능 표지하고 벡터 λFIX (스트라타진 카타로그 #944201) 에 구축된 인간 게놈 라이브러리를 스크리닝하는데 사용한다. 인간 게놈 라이브러리의 500,000 재조합체를 플레이트당 대략 10,000 재조합체의 농도로 50 개의 플레이트상에 플레이트한다. 재조합 박테리오파아지 플라크의 복사 니트로셀룰로오스 레플리카스를 65℃ 에서 하룻밤 표준 혼성화 완충액 (SHB=5X SSC, 0.1% SDS, 5X Denhart's, 100μg/ml 연어 정자 DNA) 에서 올리고뉴클레오티드 탐식자 #3 와 혼성화시킨다. 다음날 혼성화 용액을 함유하는 방사능 표지 올리고뉴클레오티드를 제거하고 필터들을 65℃ 에서 0.2X SSC, 0.1% SDS 로 세척한다. 단일 양성 혼성화 재조합체를 확인하고 플라크를 정제한다. BMP-12 올리고뉴클레오티드 탐식자 #3 에 혼성화된 상기 플라크로부터 정제된 재조합 박테리오파아지 클론을 λHuG-48 로 나타낸다. 저장 박테리오파아지 플레이트를 만들고 λHuG-48 인간 게놈 클론으로부터 박테리오파아지 DNA를 단리한다. 박테리오파아지 λHuG-48 은 1993, 12, 7 일자로 기탁 번호 #75625 로 아메리칸 타입 컬처 콜렉션 'ATCC' (메릴랜드 록빌, 12301 파크론 드라이브) 에 기탁하였다. 상기 기탁물은 특허 절차 및 규제의 목적을 위한 미생물의 기탁의 국제적 인정의 부다페스트 조약의 요구조건을 충족시킨다. 상기 재조합체 λHuG-48 의 올리고뉴클레오티드 혼성화 부위는 3.2 kb BamHI 단편에 국한된다. 이 단편을 플라스미드 벡터 (pGEM-3) 에 서브클로닝하여 DNA 서열 분석을 행한다. 상기 플라스미드 서브클론은 PCR1-1#2 로 나타내고, 1993, 12, 7 일자로 기탁 번호 #69517 로 아메리칸 타입 컬처 콜렉션 'ATCC' (메릴랜드 록빌, 12301 파크론 드라이브) 에 기탁하였다. 상기 기탁물은 특허 절차 및 규제의 목적을 위한 미생물의 기탁의 국제적 인정의 부다페스트 조약의 요구조건을 충족시

킨다. 클론 λHuG-48 로부터 유도된 플라스미드 서브클론 PCR1-1#2 의 3.2 kb DNA 삽입체의 부분 DNA 서열 (SEQ ID NO : 1) 및 아미노산 서열 (SEQ ID NO : 2) 는 서열표에 나타내었다.

SEQ ID NO : 1 의 뉴클레오티드 #639-#714 은 SEQ ID NO : 5 에 나타낸 DNA 단편을 코딩하는 특정의 증폭된 BMP-12의 뉴클레오티드 #27-#102 에 해당하며 ; 따라서 인간 게놈 박테리오파아지 클론 λHuG-48 및 유도체 서브클론 PCR1-1#2 는 본 발명의 BMP-12 단백질의 적어도 일부를 코딩함이 분명함을 주목할 필요가 있다. 플라스미드 PCR1-1#2 의 3.2 kb BamHI 삽입체 부분의 뉴클레오티드 서열은 SEQ ID NO : 1 의 뉴클레오티드 #1-#882 로 정의되는 바와 같이, 적어도 882 bp 의 전사해독틀을 포함한다. 이 전사 해독틀은 본 발명의 인간 BMP-12 단백질의 적어도 294 아미노산을 코딩한다. 코딩된 294 아미노산 인간 BMP-12 단백질은 완전히 성숙한 인간 BMP-12 단백질 (SEQ ID NO : 2 의아미노산 #1-#104) 및 제 1 번역 산물의 프로펩티드 영역의 C-말단 부분 (SEQ ID NO : 2 의 아미노산 #-190 내지 #-1) 을 포함한다.

SEQ ID NO : 33 에 나타낸 플라스미드 PCR1-1#2 의 3.2 kb BamHI 삽입체의 추가 DNA 서열은 SEQ ID NO : 33 의 뉴클레오티드 #138 내지 #1301 로 정의되는 바와 같이, 1164 bp 전사해독틀의 존재를 증명한다 [주 : SEQ ID NO : 1 에 개시된 모든 서열은 SEQ ID NO : 33 내에 포함된다]. 이 서열은 게놈 클론에서 유도된 것이기 때문에 코딩 서열의 5' 쪽 범위와 개재된 서열 (인트론/비코딩 서열) 의 3' 리미트 사이의 경계를 결정하는것이 어렵다.

기타 BMP 단백질 및 TGF-β 패밀리내의 기타 단백질의 지식에 기초하여, 전구체 폴리펩티드는 Arg-X-X-Arg 의 제안된 컨센서스 단백질분해 가공 서열과 일치하는 다염기성 서열 Arg-Arg-Gly-Arg 에서 절단되리라 생각된다. BMP-12 전구체 폴리펩티드의 절단은 SEQ ID NO : 2 의 # 1 위치에서 아미노산 Ser 로 시작하는 104 아미노산 성숙 펩티드를 내놓을 것으로 예상된다. BMP-12 의 성숙한 형태로의 가공처리는 관련 단백질 TGF-β 의 가공처리와 유사한 방식으로 N-말단 부위의 다이머화 및 제거를 수반할 것으로 예상된다 [Gentry et al., Molec & Cell. Biol. , 8: 4162 (1988) ; Derynck et al. Nature , 316 : 701 (1985)].

그러므로 BMP-12 의 성숙한 활성종은 2 개의 폴리펩티드 소단위체로 이루어진 호모다이머를 포함하고, 각각의 소단위체가 대략 12,000 달톤의 예상 분자량을 갖는 SEQ ID NO : 2 의 아미노산 #1 내지 #104 를 포함하는 것으로 생각하여 볼 수 있다. 또한 활성종은 SEQ ID NO : 2 의 적어도 아미노산 #3 내지 #103 을 포함하여 첫번째 및 마지막의 보존성 시스테인 잔기를 포함하리라 생각된다. TGF-β/BMP 패밀리의 다른 단백질과 마찬가지로, BMP-12 단백질의 카르복실-말단 부분은 그 이상의 아미노-말단 부분보다 더 높은 서열 보존성을 나타낸다. 인간 BMP 단백질 및 TGF-β 패밀리내의 기타 단백질의 대응하는 부위에 대한 시스테인-풍부 C-말단 도메인 (아미노산 #3-#104) 에서의 인간 BMP-12 단백질의 아미노산 일치성의 백분율은 하기와 같다 : BMP-2, 55% ; BMP-3, 43% ; BMP-4, 53% ; BMP-5, 49% ; BMP-6, 49% ; BMP-7, 50% ; BMP-8, 57% ; BMP-9, 48% ; BMP-10, 57% ; 액티빈 WC(BMP-11), 38% ; Vg1, 46% ; GDF-1, 47% ; TGF-β1, 36% ; TGF-β2, 36% ; TGF-β3, 39% ; 인히빈 β(B), 36% ; 인히빈 β(A), 41%.

인간 BMP-12 DNA 서열 (SEQ ID NO : 1), 또는 이의 일부는 BMP-12 mRNA 를 합성하는 인간 세포주 또는 조직을 확인하기위한 탐식자로 사용할 수 있다. 간략히 설명하여, RNA 는 선별한 세포 또는 조직 원에서 추출하고 포름알데히드 아가로오즈 겔상에서 전기영동하여, 니트로셀룰로오스로 옮기거나, 포름알데히드와 반응시켜 직접 니트로셀룰로오스상에 스포트(spot)한다. 이어서 니트로셀룰로오스를 인간 BMP-12 의 코딩 서열 유래의 탐식자와 혼성화한다.

또는, 특정 증폭 반응을 수행하기위하여 이용하는 프라이머들 사이의 영역에 위치하는 BMP-12 코딩 서열의 일부를 특정하게 증폭시켜 줄 올리고뉴클레오티드 프라이머를 설계하는데에 인간 BMP-12 서열이 사용되기도 한다. 상기 인간 BMP-12 유래 프라이머로 mRNA, cDNA 또는 게놈 DNA 주형의 대응하는 BMP-12 코딩 서열을 특정하게 증폭할 수 있으리라 생각된다. 일단 전술한 방법중의 하나로 양성이 확인되면, mRNA 를 올리고 (dT) 셀룰로오즈 크로마토그

래피로 선별하고 cDNA 를 합성하여 당업자에게 공지된 λgt10 또는 기타 λ박테리오파아지 벡터, 예컨대 λZAP 에소 종래의 기술에의한 방법으로 클랩핑한다 (Toole et al., 상기참조). 또한 전술한 올리고뉴클레오티드 프라이머에의한 증폭 반응을 λ박테리오파아지 벡터에 클론된 기정의 인간 cDNA 또는 게놈 라이브러리에 대해 직접 행할 수 있다. 이런 경우에서는, 증폭된 BMP-12 코딩 DNA 의 단편을 탐식자로 이용하면, 인간 BMP-12 단백질의 부분을 코딩하는 특정의 증폭된 DNA산물을 내놓는 라이브러리를 직접 스크린할 수 있다.

인간 BMP-12 게놈 클론 λHuG-48 의 DNA 서열을 기초로하여 설계한 올리고뉴클레오티드 프라이머는 시판용의 기정의 인간 cDNA 라이브러리들(캘리포니아, 라호야, 스트라타진 또는 캘리포니아, 팔로 알토, 클론테크 레버러토리즈사) 로부터 인간 BMP-12 코딩 DNA 서열의 특정 증폭을 가능케하리라 예견된다 . 하기 올리고뉴클레오티드 프라이머는 SEQ ID NO : 1 에 나타낸 DNA 서열 의 뉴클레오티드 #571 내지 #590 에 기초하여 설계하고 자동 DNA 합성기로 합성한다 :

4 :TGCGGATCCAGCCGCTGCAGCCGCAAGCC (SEQ ID NO : 21)

프라이머 #4 의 처음의 9 개 뉴클레오티드 (밑줄친) 는 본 발명의 인간 BMP-12 단백질을 코딩하는 특정하게 증폭된 DNA 서열의 조작을 촉진시키기위하여 사용될 수 가 있는 제한 엔도뉴클레아제 BamHI 의 인식 서열을 포함하며 따라서 SEQ ID NO : 1 에 나타낸 DNA 서열에서 유래된 것은 아니다.

하기의 올리고뉴클레오티드 프라이머를 SEQ ID NO : 1 에 나타낸 DNA 서열의 뉴클레오티드 #866 내지 #885 에 기초하여 설계하고 자동 DNA 합성기로 합성한다 :

#5GACTCTAGACTACCTGCAGCCGCAGGCCT (SEQ ID NO : 22)

프라이머 #5 의 처음의 9 개 뉴클레오티드 (밑줄친) 는 본 발명의 인간 BMP-12 단백질을 코딩하는 특정의 증폭된 DNA 서열의 조작을 용이하게 하는데 사용될 수가 있는 제한 엔도뉴클레아제 XbaI 의 인식 서열을 포함하며 따라서 SEQ ID NO : 1 에 나타낸 DNA 서열에서 유래된 것은 아니다.

앞서 확인된 프라이머들에서 표준 뉴클레오티드 기호들은 하기와 같다 : A, 아데닌 ; C, 시토신 ; G, 구아닌 ; T, 티민.

앞서 확인된 프라이머 #4 및 #5 를 프라이머로 사용하여, 다음을 포함할 수 있는 기정의 cDNA 라이브러리로부터 특정 BMP-12 코딩 뉴클레오티드 서열을 증폭한다 : 사람 태아 뇌 cDNA/λZAPII (스트라타진 카타로그 #936206), 사람 간/λUNI-ZAP XR (스트라타진 카타로그 #937200), 사람 폐/λUNI-ZAP XR (스트라타진 카타로그 #937206) 및 사람 태아 비장/UNI-ZAP XR (스트라타진 카타로그 #937205).

상술한 바와 같은 인간 cDNA 삽입체를 포함하는 λ박테리오파아지 라이브러리 약 1×108 pfu (플라크 형성 단위) 를 95℃에서 5 분간 변성시킨후 각각의 데옥시뉴클레오티드 트리포스페이트 (dATP, dGTP, dCTP 및 dTTP) 200 μM, 10 mM 트리스-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl2, 0.001% 젤라틴, 1.25 단위 Taq DNA 폴리머라아제, 100 pM 올리고뉴클레오티드 프라이머 #4 및 100 pM 올리고뉴클레오티드 프라이머 #5 를 함유하는 반응 혼합물에 첨가한다. 이어서 반응 혼합물을 하기의 방법으로 열 순환시킨다 : 39 차례의 열 순환동안 94℃ 에서 1 분, 50℃ 에서 1 분, 72℃ 에서 1 분, 그 다음에는 72℃ 에서 10 분.

상기 반응으로 특정하게 증폭된 DNA 는 내부의 315bp 가 SEQ ID NO : 1 의 뉴클레오티드 #571 내지 #885 에 해당하고, 또한 프라이머 #4 및 #5 의 뉴클레오티드 #1-#9 로 정의된 제한 부위에 해당하는 BMP-12 의 특정단편의 각 말단의 9bp를 포함하는 것이기도 한 약 333 bp 의 BMP-12 코딩 산물을 제한 생성시키리라 기대된다. 그 결과로서

얻어지는 333 bp 의 DNA 산물을 제한 엔도뉴클레아제 BamHI 및 XbaI 로 소화시키고, 페놀 추출, 클로로포름 추출 및 에탄올 침강시킨다.

또는, 에탄올 침전을 위하여, 완충액 교환 및 BamHI/XbaI 제한 소화로부터 결과된 작은 DNA 단편의 제거는 10 mM 트리스-HCl pH 8.0, 1mM EDTA 에서의 소화 DNA 산물의 희석 및 후 속의 센트리콘TM 30 마이크로콘센트레이터로 원심분리하여 달성한다 (매사츄세트, 베버리, 더블유.알. 그레이스 앤 코오포레이트, 프로덕트 #4209). 그 결과로 얻어진 BamHI/XbaI 소화 증폭 DNA 산물을 폴리링커의 BamHI 및 XbaI 부위 사이의 플라스미드 벡터 (즉, pBluescript, pGEM-3 등)에 서브클로닝시킨다. 그 결과로 서 얻어지는 서브클론의 DNA 서열 분석은 BMP-12 코딩 삽입체의 완전함의 확인에 요구될 수 있을 것이다. 일단 양성의 cDNA 원이 상기 방법으로 확인되었으면, 333 bp BMP-12 특정 서열이 증폭되었던 해당하는 cDNA 라이브러리를 333 bp 삽입체 또는 기타 BMP-12 특이 탐식자로 직접 스크린하여 본 발명의 완전한 표준 크기의 BMP-12 단백질을 코딩하는 cDNA 클론을 확인 및 분리할 수 있다.

당업자에게 공지된 추가의 방법을 사용하면 인간 BMP-12 관련 단백질를 코딩하는 기타 완전한 표준 크기의 cDNA 들, 또는 사람 이외의 종 유래의, 특히 다른 포유동물 종의 본 발명의 BMP-12 관련 단백질을 코딩하는 완전한 표준 길이의 cDNA 클론들을 분리할 수 있다.

하기의 실시예들은 뮤린 게놈 DNA 라이브러리에서 BMP-12 관련 단백질의 상동체를 분리하기 데에 사용되는 인간 BMP-12 서열의 용도를 증명한 것이다.

본 발명의 인간 BMP-12 단백질을 코딩하는 DNA 서열은 상응하는 BMP-12 관련 단백질을 코딩하는 기타 종의 DNA 서열을 특정하게 증폭시키는데 이용될 수 있는 사람이외의 다른 종의 BMP-12 및 BMP-12 관련 서열과 상당한 상동성이 있는 것으로 예측된다. 구체적으로, 하기의 올리고뉴클레오티드는 인간 BMP-12 서열 (SEQ ID NO : 1) 을 기초로하여 설계하며 자동 DNA 합성기로 합성한다 :

#6 :GCGGATCCAAGGAGCTCGGCTGGGACGA (SEQ ID NO : 23)

#7 :GGAATTCCCCACCACCATGTCCTCGTAT (SEQ ID NO : 24)

올리고뉴클레오티드 프라이머 #6 과 #7 의 처음의 8 개 뉴클레오티드(밑줄친) 는, 각각, 제한 엔도뉴클레아제 BamHI과 EcoRI 으 인식 서열을 포함한다. 상기 서열은 사람이외의 다른 종의 BMP-12 및 BMP-12 관련 단백질을 코딩하는 특정의 증폭된 DNA 서열의 조작을 촉진시키는데 사용하며 따라서 SEQ ID NO : 1 에 나타낸 DNA 서열에서 유래된 것은 아니다. 올리고뉴클레오티드 프라이머 #6 은 SEQ ID NO : 1 의 뉴클레오티드 #607-#626 을 기초로하여 설계한다. 올리고뉴클레오티드 프라이머 #7 은 SEQ ID NO : 1 에 나타낸 DNA 서열의 뉴클레오티드 #846-#865 의 역상보 부분을 기초로하여 설계한다.

앞서 확인된 올리고뉴클레오티드 프라이머 #6 및 #7 은 사람이외의 다른 종에서 유래된 게놈 DNA 의 특정 BMP-12 관련 서열의 증폭을 가능케 하는 프라이머로서사용된다. 증폭 반응은 하기와 같이 행한다 :

뮤린 게놈 DNA (원 : 균주 Balb c) 을 25 게이지 니들을 통한 반복된 찌르기에 의해 전단시키고, 100℃ 에서 5 분간 변성시키고 이어서 얼음으로 냉각시킨후 각각 200μM 의 데옥시뉴클레오티드 트리포스페이트들 (dATP, dGTP, dCTP 및 dTTP), 10 mM 트리스-HCl pH 8.3, 50 mM KCl, 1.5 mM MgCl2, 0.001% 젤라틴, 1.25 단위 Taq DNA 폴리머라아제, 100 pM 올리고뉴클레오티드 프라이머 #6 및 100 pM 올리고뉴클레오티드 프라이머 #7 을 함유하는 반응 혼합물에 첨가한다. 그 다음에 상기 반응 혼합물을 하기 방법으로 열 순환을 시킨다 : 40 차례에 순환동안 95℃에서 1 분, 55℃에서 1 분, 72℃에서 1 분, 이어서 72℃에서 10 분간.

상기 반응으로 특정하게 증폭된 DNA를 에탄올 침강시키고, 제한 엔도뉴클레아제 BamHI 및 EcoRI 로 소화시키고 아가로오즈 겔 전기영동한다. 뮤린 BMP-12 또는 BMP-12 관련 단백질을 코딩하는 DNA 단편의 예측된 크기에 해당하는, 겔 부위를 잘라내어 여기에 함유된 특정의 증폭된 DNA 단편을 추출하고 (전기용리 또는 당업자에게 공지된 기타 방법으로) 폴리링커의 BamHI 및 EcoRI 부위 사이에서 pGEM-3 또는 pBluescript 등의 플라스미드 벡터에 서브클로닝한다. 그 결과로서 얻어진 mV1 로 칭한 서브클론중의 하나에 대한 DNA 서열 분석은, 여기에 포함된 특정의 증폭된 DNA 서열이 본발명의 BMP-12 또는 VL-1 서열의 뮤린 상동체인 것으로 보이는 단백질의 부분을 코딩하는 것임을 보여준다. 이 특정의 증폭된 뮤린 DNA 단편의 DNA 서열 (SEQ ID NO : 10) 및 유래된 아미노산 서열 (SEQ ID NO : 11) 은 서열표에 나타나 있다.

SEQ ID NO : 10 의 뉴클레오티드 #1-#26 은 올리고뉴클레오티드 #6 의 부분을 포함하고 뉴클레오티드 #246-#272 는 특정한 증폭 반응을 수행하기위하여 이용하는 올리고뉴클레오티드 #7 의 역상보 부분을 포함한다. SEQ ID NO : 10 의 뉴클레오티드 #27 은 코돈 3 염기쌍의 마지막 뉴클레오티드인 것으로 보이고, SEQ ID NO : 10 의 뉴클레오티드 #244-#245 는 코돈 3 염기쌍의 처음의 두 뉴클레오티드인 것으로 보인다. 그러므로, SEQ ID NO : 10 의 뉴클레오티드 #28 내지 #243 는 mV1 의 부분 코딩 서열과 일치한다. 증폭 반응의 개시에서의 올리고뉴클레오티드 #6 및 #7 의 기능으로 인하여, 이들은 본 발명의 인간 BMP-12 또는 VL-1단백질의 뮤린 상동체를 코딩하는 실제의 서열과 정확하게 일치하지 않을 수 있어 대응하는 아미노산 서열 유도 (SEQ ID NO : 11) 상 번역되지 않는다.

SEQ ID NO : 10 에 나타낸 특정의 증폭된 뮤린 BMP-12 또는 VL-1 DNA 서열에 기초하여 설계된 올리고뉴클레오티드 탐식자는 완전한 표준크기의 뮤린 BMP-12 또는 VL-1 을 코딩하는 클론 (cDNA 또는 게놈성) 을 확인하는데에 당업자에의해 이용될 수 있다.

그 결과로서 얻어지는 mV2 로 이름붙인 또 다른 서브클론의 DNA 서열 분석은 여기에 포함된 특정의 증폭된 DNA 서열은 본 발명의 뮤린 BMP-12 관련 서열의 일부를 코딩하는 것임을 보여준다.

상기 특정의 증폭된 뮤린 DNA 단편의 DNA 서열 (SEQ ID NO : 12) 및 유래된 아미노산 서열 (SEQ ID NO : 13) 은 서열 일람표에 나타내었다.

SEQ ID NO : 12 의 뉴클레오티드 #1-#26 은 올리고뉴클레오티드 #6 의 일부를 포함하고 뉴클레오티드 #246-#272 는 특정한 증폭 반응을 수행하기위하여 이용하는 올리고뉴클레오티드 #7 의 역 상보 부분을 포함한 것이다. SEQ ID NO : 12 의 뉴클레오티드 #27 은 코돈 3 염기쌍의 마지막 뉴클레오티드인 것으로 보이고, SEQ ID NO : 12 의 뉴클레오티드 #244-#245 는 코돈 3 염기쌍의 처음의 두 뉴클레오티드인 것으로 보인다. 그러므로, SEQ ID NO : 12 의 뉴클레오티드 #28 내지 #243 는 mV2 의 부분 코딩 서열과 일치한다. 증폭 반응의 개시에서 올리고뉴클레오티드 #6 및 #7 의 기능으로 인하여, 이들은 본 발명의 뮤린 BMP-12 관련 단백질을 코딩하는 실제의 서열과 정확하게 일치하지 않을 수 있어 대응하는 아미노산 서열 유도 (SEQ ID NO : 13) 에서 번역되지 않는다.

SEQ ID NO : 12 에 나타낸 특정하게 증폭된 뮤린 BMP-12 관련 DNA 서열에 기초하여 설계된 올리고뉴클레오티드 탐식자는 완전한 표준 크기의 뮤린 BMP-12 관련 단백질을 코딩하는 클론 (cDNA 또는 게놈성) 을 확인하는데에 당업자에의해 이용될 수 있다.

그 결과로서얻어지는 mV9 로 이름붙인 또 다른 서브클론의 DNA 서열 분석은, 여기에 포함된 특정의 증폭된 DNA 서열이 본 발명의 뮤린 BMP-12 관련 서열의 부분을 코딩하는 것임을 나타낸다. 상기 서열은 SEQ ID NO : 3 에서 나타낸 인간 MP52 DNA 서열의 뮤린 상동체인 것으로 보인다. 상기 특정의 증폭된 뮤린 DNA 단편의 DNA 서열 (SEQ ID NO : 14) 및 유래된 아미노산 서열 (SEQ ID NO : 15) 은 서열표에 나타나 있다.

SEQ ID NO : 14 의 뉴클레오티드 #1 - #26 은 올리고뉴클레오티드 #6 의 부분을 포함하고 뉴클레오티드 #246 - #272 은 특정한 증폭 반응을 수행하기위하여 이용하는 올리고뉴클레오티드 #7 의 역상보 부분을 포함한 것이다. SEQ ID NO : 14 의 뉴클레오티드 #27 은 코돈 3 염기쌍의 마지막 뉴클레오티드인 것으로 보이고, SEQ ID NO : 14 의 뉴클레오티드 #244 - #245 는 코돈 3 염기쌍의 처음의 두 뉴클레오티드인 것으로 보인다. 그러므로, SEQ ID NO : 14 의 뉴클레오티드 #28 내지 #243 는 mV9 의 부분 코딩 서열과 일치한다. 증폭 반응의 개시에서의 올리고뉴클레오티드 #6 및 #7 의 기능으로 인하여, 이들은 본 발명의 뮤린 BMP-12 관련 단백질을 코딩하는 실제의 서열과 정확하게 일치하지 않을 수 있어 대응하는 아미노산 서열 유도 (SEQ ID NO : 15) 상 번역되지 않는다.

SEQ ID NO : 14 에 나타낸 특정의 증폭된 뮤린 BMP-12 관련 DNA 서열에 기초하여 설계된 올리고뉴클레오티드 탐식자는 완전한 표준크기의 뮤린 BMP-12 관련 단백질을 코딩하는 클론 (cDNA 또는 게놈성) 을 확인하는데에 당업자에 의해 이용될 수 있다.

또는, 앞서 확인된 올리고뉴클레오티드 프라이머 #6 및 #7 은 275 bp DNA 탐식자의 특정 증폭을 가능케하는 프라이머로서 이용되기도 하는데, 그 내부의 259 bp 는 BMP-12 를 코딩하는 플라스미드 서브클론 PCR 1-1#2 유래의, SEQ ID NO : 1 의 뉴클레오티드 #607 내지 #865 에 해당한다. 상기 275bp DNA 탐식자를 32P 로 방사능 표지시켰고 벡터 λFIX II에 구축된 뮤린 게놈 라이브러리 (스트라타진 카타로그 #946306) 를 스크린하는데 사요하였다. 뮤린 게놈 라이브러리의 재조합체 백만개를 플레이트당 대략 20,000 재조합체의 농도로 50 개의 플레이트상에 플레이트한다. 재조합 박테리오파아지 플라크의 이중복사 니트로셀룰로오스를 60℃에서 하룻밤 표준 혼성화 완충액 (SHB=5X SSC, 0.1% SDS, 5X Denhart's, 100μg/ml 연어 정자 DNA) 의 특정의 증폭된 333bp 탐식자와, 엄격성이 감소된 조건하에서 혼성화시킨다. 다음날 방사능 표지 올리고뉴클레오티드를 함유하는 혼성화 용액을 여과기로 제거하고, 감소된 엄격성의 조건하에서, 60℃ 에서 2X SSC, 0.1% SDS 로 필터를 세척한다. 여러개의 양성의 혼성화 재조합체를 확인하고 플라크를 정제한다. 양성의 혼성화 뮤린 게놈 재조합체 클론의 단편을 표준 플라스미드 벡터 (즉, pGEM-3) 에 서브클로닝하여 DNA 서열분석을 행한다.

MVR3 으로 이름붙인 상기 서브클론중의 하나에 대한 DNA 서열 분석은, 그것이 상기한 PCR 산물 mV1 (SEQ ID NO : 1 에 나타낸 인간 BMP-12 서열의 뮤린 상동체) 에 상당하는 마우스 유전자의 부분을 코딩하는 것임을 나타낸다. 상기 서브클론의 부분 DNA 서열 및 상응하는 아미노산 번역은 각각 SEQ ID NO : 29 및 SEQ ID NO : 30 에 나타내었다.

MVR32 로 이름붙인 상기 서브클론들중의 다른 하나에 대한 DNA 서열 분석은 그것이 상기한 PCR 산물 mV2 (SEQ ID NO : 7 에 나타낸 인간 VL-1 서열의 뮤린 상동체) 에 상당하는 마우스 유전자의 부분을 코딩하는 것임을 나타낸다. 상기 서브클론의 부분 DNA 서열 및 상응하는 아미노산 번역은 각각 SEQ ID NO : 31 및 SEQ ID NO : 32 에 나타나 있다.

MVR23 으로 이름붙인 상기 서브클론중의 또 다른 하나에 대한 DNA 서열 분석은, 그것이 상기한 PCR 산물 mV9 (SEQ ID NO : 3 에 나타낸 MP-52 서열의 뮤린 상동체) 에 해당하는 마우스 유전자의 부분을 코딩하는 것을 나타낸다.

본 발명의 BMP-12 단백질을 코딩하는 인간 게놈 클론을 확인 및 분리하기위한 상기한 방법과 유사한 방법으로, SEQ ID NO : 7 에 나타낸 VL-1 코딩 서열에 해당하는 올리고뉴클레오티드 탐식자(들)을 설계하여 사용하면 VL-1 (BMP-13) 단백질을 코딩하는 인간 게놈 또는 cDNA 서열들을 확인할 수가 있다. 이들 올리고뉴클레오티드들은 VL-1 코딩 서열에 특이적인 부위까지 설계될 수 있을 것이며 그러므로 특정의 증폭된 VL-1 코딩 서열(SEQ ID NO : 7) 과 특정의 증폭된 BMP-12 코딩 서열 (SEQ ID NO : 5) 사이의 뉴클레오티드 서열 일치 정도가 가장 낮은 부위에서 유도된 것일 가능성이 있다.

또는, 앞서 확인된 올리고뉴클레오티드 프라이머 #4 및 #5 는 333 bp DNA 탐식자의 특정 증폭을 가능케 하는 프라이머로서 이용되기도 하며, 그 내부의 315 bp 는 BMP-12 를 코딩하는 플라스미드 서브클론 PCR 1-1#2 유래의, SEQ ID NO : 1 의 뉴클레오티드 #571 내지 #885 에 해당한다. 상기 333bp DNA 탐식자를 32P 로 방사능 표지시켰고 벡터 λDASH II 에 구축된 뮤린 게놈 라이브러리 (스트라타진 카타로그 #945203)을 스크린하는데 사용하였다. 뮤린 게놈 라이브러리의 재조합체 백만개를 플레이트당 대략 20,000 재조합체의 농도로 50 개의 플레이트상에 플레이트시킨다. 재조합 박테리오파아지 플라크의 이중복사 니트로셀룰로오스를 60℃ 에서 하룻밤 표준 혼성화 완충액 (SHB=5X SSC, 0.1% SDS, 5X Denhart's, 100μg/ml 연어 정자 DNA) 의 특정의 증폭된 333bp 탐식자와 엄격성이 감소된 조건하에서, 혼성화시킨다. 다음날 방사능 표지 올리고뉴클레오티드를 함유하는 혼성화 용액을 제거 하고, 감소된 엄격성의 하, 60℃ 에서 2X SSC, 0.1% SDS 로 필터를 세척한다. 여러개의 (대략 15) 양성의 혼성화 재조합체를 확인하고 플라크를 정제한다.

본 발명의 VL-1 단백질을 코딩하는 양성의 혼성화 재조합체를 본 스크리닝 절차에서 이용하는 BMP-12 코딩 플라스미드 PCR1-1#2 로부터 발생하는 333bp DNA 탐식자에 양성으로 혼성화하리라 예견되는 BMP-12 및 기타 BMP-12 관련 단백질을 코딩하는 재조합체와 구별하기 위하여, SEQ ID NO : 7 에 나타낸 VL-1 서열에 기초하여 다음의 올리고뉴클레오티드 탐식자를 설계하고 자동 DNA 합성기로 합성한다 :

#8 : TGTATGCGACTTCCCGC [SEQ ID NO : 35]

SEQ ID NO : 7 의 뉴클레오티드 #60 내지 #76 에 해당하는 올리고뉴클레오티드는 SEQ ID NO : 1 에 나타낸 BMP-12 코딩 서열의 대응하는 부분 (뉴클레오티드 #672 내지 #689) 과는 5 개의 뉴클레오티드가 다르다. VL-1 올리고뉴클레오티드 탐식자 #8 에 혼성화하는 재조합 박테리오파아지 클론들중의 하나를 λJLDc31 로 나타낸다. 이 재조합 박테리오파아지 클론을 플라크 정제하고, 박테리오파아지 저장 플레이트를 만들어 이 λJLDc31 인간 게놈 클론으로부터 박테리오파아지 DNA 를 분리한다. 박테리오파아지 λJLDc31 은 1994, 10, 20 일자로 수탁 번호 #75922 로 아메리칸 타입 컬처 콜렉션 'ATCC' (메릴랜드 록빌, 12301 파크론 드라이브) 에 기탁하였다. 상기 기탁물은 특허 절차 및 규제의 목적을 위한 미생물의 기탁의 국제적 인정의 부다페스트 조약의 요구조건을 충족시킨다. 상기 재조합체, λJLDc31의 올리고뉴클레오티드 혼성화 부위는 2.5 kb Eco RI 단편에 국한되어 있다. 상기 단편을 플라스미드 벡터 (pGEM-3) 에 서브클로닝하고 DNA 서열분석을 행한다. 상기 플라스미드 서브클론은 pGEMJLDc31/2.5 로 나타내고 1994, 10, 20 일자로 수탁 번호 #69710 로 아메리칸 타입 컬처 콜렉션 'ATCC' (메릴랜드 록빌, 12301 파크론 드라이브) 에 기탁하였다. 상기 기탁물은 특허 절차 및 규제의 목적을 위한 미생물의 기탁의 국제적 인정의 부다페스트 조약의 요구조건을 충족시킨다.

클론 λJLDc31 로부터 유래한 플라스미드 서브클론 pGEMJLDc31/2.5 의 2.5 kb DNA 삽입체의 부분의 부분 DNA 서열 (SEQ ID NO : 25) 및 유래된 아미노산 서열 (SEQ ID NO : 26) 를 서열표에 나타나 있다.

SEQ ID NO : 25 에 나타낸 플라스미드 pGEMJLDc31/2.5 의 2.5 kb EcoRI 삽입체의 부분의 DNA 서열은 SEQ ID NO : 25 의 뉴클레오티드 #52 내지 #963 으로 정의된 바와 같은, 912 bp 전사해독틀을 함유한다. 상기 서열은 게놈 클론으로부터 유래하였기 때문에 코딩 서열의 5' 영역과 개재된 서열 (인트론/비코딩 서열)의 3' 말단사이의 경계를 결정하는것이 어렵다. 완전한 전사해독틀 (SEQ ID NO : 25 의 뉴클레오티드 #52 내지 #963) 은 304 아미노산 까지의 본 발명의 VL-1 단백질의 부분을 코딩한다.

기타 BMP 단백질 및 TGF-β 패밀리내의 기타 단백질의 지식에 기초하여, 전구체 폴리펩티드는 제안된 컨센서스 단백질 가수 분해 처리 서열 Arg-X-X-Arg 과 일치하는 다중염기 서열 Arg-Arg-Arg-Arg 에서 절단되리라 생각된다. VL-1 전구체 폴리펩티드의 절단은 SEQ ID NO : 26 의 # 1 위치에서 아미노산 Thr 로 시작하는 120 아미노산 성숙 펩티드를 발생시키는 것으로 예상되고 있다. 성숙한 형태로 VL-1 를 가공처리하는 것은 관련 단백질 TGF-β 의 가공처리 과정과 유사한 방법으로 다이머화 및 N-말단 부위의 제거를 수반하는 것으로 예상되고 있다 [Gentry et al.,Molec & Cell. Biol., 8:4162 (1988) ; Derynck et al. Nature , 316 : 701 (1985)].

그러므로 VL-1 의 성숙한 활성종은 각각의 소단위체가 대략 12,000 달톤의 예상 분자량을 가진 SEQ ID NO : 26 의 아미노산 #1 내지 #120 을 포함하는, 2 개의 폴리펩티드 소단위체의 호모다이머를 포함한다고 생각된다. 추가의 활성종은 SEQ ID NO : 26 의 적어도 아미노산 #19 내지 #119 또는 #120 을 포함하여 첫번째 및 마지막의 보존성 시스테인 잔기를 포함하리라 생각된다.

이와같은 방법을 사용하여, 성숙한 인간 VL-1(BMP-13) 을 코딩하는 클론이 얻어졌다. 상기 클론이 코딩하는 뉴클레오티드 서열 및 상응하는 아미노산 서열은 각각 서열 일람표내의 SEQ ID NO : 25 및 26 에 수록되어 있다.

실시예 2실시예 2

BMP-12 의 발현

인간 BMP-12 단백질을 제조하기 위하여, 이를 코딩하는 DNA 서열을 통상의 유전 공학 기술로 적당한 발현 벡터에 전이시켜 포유류 세포 또는 기타 바람직한 진핵 또는 원핵 세포 숙주에 도입한다.

박테리아 세포내에서 인간 BMP-12 단백질을 제조하기 위하여, 하기의 절차를 사용한다.

E.coli 에서의 BMP-12 의 발현

E.coli 내에서의 BMP-12 의 제조를 위하여, 다음의 중요한 특징을 갖는 발현 플라스미드 pALV1-781 을 구축하였다. 뉴클레오티드 1-2060 은 숙주 E.coli 균주에 항생물질 암피실린에 대한 저항성을 부여하는 β-락타마아제에 대한 유전자 및 colE1-유래의 복제 원점을 포함하는 서열을 함유하는 플라스미드 pUC-18 [Norrander 등, Gene 26: 101-106 (1983)] 로부터 발생하는 DNA 서열을 포함한다. 뉴클레오티드 2061-2221 은 3 개의 오퍼레이터 서열 0L1, 0L2 및 0L3 을 포함하여, 박테리오파아지λ [Sanger et al., J.Mol. Biol. 162 : 729-773 (1982)] 의 좌측의 주프로모터(pL) 의 DNA 서열을 포함한다 . 오퍼레이터는 λcI 레프레서 단백질의 결합부위이고, 그것의 세포내 농도는 pL 로부터 전사 개시량을 조절한다. 뉴클레오티드 2222-2723 은 Sanger 등의 J.Mol. Biol. 162: 729-773 (1982) 에 기재된 박테리오파아지 lambda 의 뉴클레오티드 35566 내지 35472 및 38137 내지 38361 유래의 서열상에 포함된 강력한 리보솜 결합 서열을 포함한다. 뉴클레오티드 2724-3041 은 모든 3' 비번역 서열이 제거된 성숙한 BMP-12 단백질을 코딩하는 DNA 서열을 포함한다. pALV1-781 발현 벡터에 도입된 BMP-12 DNA 서열은 코딩 능력의 변화없이 A+T 함량을 높여주기 위하여 5' 말단에서 부분변화시켰다. 이러한 부분변화로 E.coli 내의 BMP-12 mRNA 상에서 개시되는 번역의 효율이 증가하게 되었다. 뉴클레오티드 3042-3058 은 제한 엔도뉴클레아제 부위를 포함하는 '링커' DNA 서열을 제공한다. 뉴클레오티드 3059-3127 은 E.coli aspA 유전자의 전사 종결 서열에 기초로한 전사 종결 서열을 제공한다 [Takagi et al., Nucl. Acids Res. 13:2063-2074 (1985)]. 뉴클레오티드 3128-3532 는 pUC-18 로부터 유래한 DNA 서열이다.

플라스미드 pALV1-781 을 Dagert & Ehrlich, Gene 6:23 (1979)의 처리법에 따라 E.coli 숙주 균주 GI724(F, lacIq,lacpL8 , ampC::λcI+) 로 형질 전환시켰다. GI724 (ATCC 수탁번호 55151) 는 염색체의 ampC 자리에서 안정하게 통합된 야생형 λcI 레프레서의 복사물을 포함하며, 거기에서 살모넬라 티피뮤리움(Salmonella typhimurium) trp프로모터/오퍼레이터 서열의 전사 조절하에 놓여지게 되었다. GI724 에서, λCI 단백질은 최소 배지 등의 무(無)트립토판 배지 또는 상기한 IMC 등의 카사미노산으로 보충된 최소 배지에서 성장하는 동안에만 제조된다. GI724 의 배양물에 트립토판의 첨가는trp프로모터를 억압하여 λCI 의 합성을 중단시킬 것이며, pL 프로모터가 세포내에 존재한다면 pL 프로모터로부터 전사의 유도를 서서히 야기할 것이다.

1mM 의 MgSO4 를 함유하는 M9 배지 [Miller, 'Experiments in Molecular Genetics,' Cold Spring Harbor Laboratory, New York (1972)] 로 구성되어 있고, 0.5% w/v 글루코오스, 0.2% w/v 카사미노산 및 100μg/㎖ 암피실린으로 보충된 IMC 배지를 포함하는 1.5% 아가 플레이트상에서 형질전환체를 선별한다. pALV1-781 로 형질전환시킨 GI724 는 100μg/㎖ 암피실린을 포함하는 IMC 배지에서 37℃ 에서 A550 = 0.5 까지 성장시켰다. 이어서 트립토판을 최종농도 100μg/㎖ 까지 첨가하고 이 배양액을 4 시간더 배양시켰다. 상기 시간 동안에 BMP-12 단백질은 '봉입체' 분획물내에 축적된다.

단백질 단량체의 제조

동결세포 18 g 을 달아서 100 mM Tris, 10 mM EDTA, 1mM 페닐메틸술포닐 플루오라이드 [PMSF], pH 8.3 의 60 ml 에 재현탁시킨다. 마이크로플루라이저TM [모델 #MCF 100T] 에 3 차례 통과시켜 세포를 분해시켰다. 4℃ 에서 15,000 g 에서 20 분간 원심 분리하여 봉입체 펠릿을 수득하였다. 상청액을 따라 내고, 펠릿을 100 mM Tris, 0.1 M NaCl, 10 mM EDTA, 1mM PMSF, pH 8.3 의 100 ml 로 세척한다. 이 현탁액을 4℃, 15,000 g 에서 10 분간 원심 분리하였고, 상층액을 따라 냈다. 이어서 펠릿을 100 mM Tris, 10 mM EDTA, 1% 트리톤 X-100, 1mM PMSF, pH 8.3 의 100 ml 로 세척하였다. 이 현탁액을 다시 4℃, 15,000 g 에서 10 분간 원심 분리하였고, 상청액을 기울여 따라 냈다. 펠릿을 유리 조직 호모저나이저에서 1% DTT 를 함유하는, 20 mM Tris, 1 mM EDTA, 1mM PMSF, pH 8.3 의 50 ml 로 재현탁시켰다. 이어서 모노머 BMP-12 를 빙초산으로 pH 2.5 까지 산성화하여 용해시켰다. 4℃, 15,000 g 에서 20 분간 원심분리하여 가용성 분획물을 분리하였다.

상기 원심분리로부터의 상청액을 수집하여 20 ml 증량으로 세파크릴 S-100TM 크기 배제 컬럼 (83 cm × 2.6 cm ; 440 ml 층) 상에서 크로마토그래피를 하였다. 세파크릴 S-100TM 컬럼은 1.4 ml/분의 흐름 속도로 1% 아세트산의 유동상으로 흘러내리게 하였다. BMP-12 모노머에 해당하는 분획물의 존재를 280 nm 에서의 흡광도로 알아 내었고, 컴퓨터를 이용하여 흡광계수 18200M-1cm-1 및 분자량 (11667 달톤) 을 계산하였다. 상기 크기 배제 컬럼에 의해 수집된 물질을 리폴딩 반응의 출발 물질로 사용하였다.

상기의 방법에 대한 대안으로, -80℃ 에서 저장된 세포 1.0 g 을 측정한다. 용액 (3.4 ml 100 mM Tris, 10 mM EDTA, pH 8.5) 을 첨가한다. 용액을 세포가 잘 현탁될 때까지 와동시킨다. 40 μℓ 의 100 mM PMSF 의 이소프로판을 용액을 첨가한다. 프렌치 압력 셀내에서 1000 psi 로 세포를 분해시킨다. 봉입체를 에펜도르프 마이크로원심분리기로 4℃ 에서 20 분간 원심 분리하여 펠릿을 형성시킨다. 상청액을 기울여 따라 낸다. (총 4 개중의) 하나의 펠릿에, 250 mM DTT 를 포함하는, 탈기된 1.0 ml 8.0 M 구아니딘 히드로클로라이드, 0.5 M TRIS, 5 mM EDTA, pH 8.5 를 첨가한다. 펠릿을 용해시키고 이 액체에 아르곤을 30 초간 불어 넣는다. 이어서 이 용액을 37℃ 에서 1 시간동안 보온 반응시킨다. 불용성 물질을 23℃에서 에펜도르프 마이크로원심분리기로 2-3 분간 작은 덩어리로 뭉치게한다. 상청액 0.5-1.0 ml 을 수펠코(Supelco) 2 cm 가이드 카트리지(LC-304) 에 주입하고, 35 분에 걸쳐 0.1 % TFA 중의 아세토니트릴 농도기울기를 1-70 % 로하여 용출시킨다. 29 분 내지 31 분에 BMP-12가 용출된다. 분획물을 수집하고 아미노산 함량을 기초로 하는 이론적 빛 흡광계수를 이용하여, 단백질 농도를 0.1% TFA 와 대비하여 280 nm 에서의 흡광도로 결정한다.

상기의 방법에 대한 두번째 대안으로서, 상기한 바의 E.coli 형질 전환체로부터 수득한 동결 세포 펠릿을 30ml 의 TE 8.3(100:10) 완충액 (100 mM Tris-HCl pH8.3, 10 mM Na2EDTA, 1 mM PMSF) 에 녹인다. 세포를 마이크로플루이다이저TM [모델 #MCF 100T] 에 3 차례 통과시켜 세포를 분해시킨다. 초기의 세포 봉입체 물질 펠릿을 8 M 구아니딘-HCl, TE8.5 (100:10) 완충액 (100 mM Tris-HCl pH8.5, 100 mM DTT 에 함유된 상태의 10 mM Na2E

DTA) 에 용해시키고, 37℃ 에서 1 시간동안 보온배양시킨다. 이 물질을 실온에서 15 분간 12,000 xg 에서 원심분리한다.

CHAPS 시스템을 이용한 BMP-12 단백질의 리폴딩

충분한 부피의 BMP-12 수집물을 냉동건조시켜 10μg 의 단백질을 수득한다. 5 μℓ 의 유리 증류수를 첨가하여 잔류물을 재용해시키고, 이어서 100 μℓ 의 리폴딩용 혼합물 (50 mM Tris, 1.0 M NaCl, 2% 3-(3-클로라미도프로필)디메릴암모니아-1-프로판-술페이트 (CHAPS), 5 mM EDTA, 2 mM 글루타티온(환원) 1mM 글루타티온 (산화) ; 대략 8.5의 pH) 을 첨가한다. 이 용액을 부드럽게 혼합하고 23℃ 에서 1-4 일간 보관한다. 다이머 형성은 125 볼트로 노벡스 16% 트리신 겔상에서 일정분량을 2.5 시간동안 전개시킨 다음, 쿠마씨 블루 염색 및 탈색을 행하여 평가한다.

(A 완충액로 희석시킨) 1% B 완충액으로 평형화시킨 다음의 농도기울기 (A완충액 = 0.1% 트리플루오로아세트산, B완충액=95% 아세토니트릴, 0.1% 트리플루오로아세트산 [TFA]) 를 이용하는, 단백질이 용출되는 35 분간에 걸쳐 1 분당 1 ml 의 흐름속도로 흘러내린 C4 분석 RP-HPLC(역상 고성능 액체 크로마토그래피) 컬럼 (Vydac 214TP54) 을 이용하여 BMP-12 다이머를 정제하였다 :

1-5 분 20% B 완충액

5-10 분 20-30% B완충액

10-30 분 30-50% B완충액

30-35 분 50-100% B완충액

단백질을 280 nm 에서 흡광도로 모니터하였다. (29 분 내지 31 분에 용출하는) 피크 BMP-12 분획물들을 모았다. 순도는 SDS-PAGE 로 평가하였다. 농도를 280 nm 에서 측정하고, 상기한 바와 같이 컴퓨터를 이용하여 빛의 흡수 계수 및 분자량을 계산하였다.

포유류 세포에서의 BMP-12의 발현 :

생물학적으로 활성인 제조합 인간 BMP-12 의, 또 다른 생각하여 볼 수 있는 바람직한 발현 시스템은 안정하게 형질전환된 포유류 세포이다.

당업자라면 SEQ ID NO : 1 의 서열, 또는 BMP-12 단백질을 코딩하는 기타 DNA 서열 또는 기타 부분변화된, 서열들과 공지의 벡터들, 이를 테면 pCD [Okayama et al., Mol. Cell Biol. , 2:161-170 (1982)], pJL3, pJL4 [Gough et al., EMBO J., 4:645-653 (1985)] 및 pMT2 CXM 등을 사용하여 포유류 발현 벡터를 구축할 수 있다.

포유류 발현 벡터 pMT2 CXM 은 p91023(b) (Wong et al., Science 228 : 810-815, 1985) 의 유도체로서, 후자와 다른 점은 테트라사이클린 저항성 유전자 대신에 암피실린 저항성 유전자를 포함하고 추가로 cDNA 클론의 삽입을 위한 XholI 부위를 포함하고 있다는 것이다. pMT2 CXM 의 기능요소들은 기술되어 있고 (Kaufman, R.J., 1985, Proc. natl. Acad. Sci. USA 82:689-693) 아데노바이러스 VA 유전자들, 72 bp 엔핸서를 포함하는 SV 40 원점, 5' 스플라이스 부위 및 아데노바이러스 후기 mRNAs 상에 존재하는 대다수의 아데노바이러스 3 분할 리더 서열의 대부분, 3' 스플라이스 수용체 부위, DHFR 삽입체를 포함하는 아데노바이러스 주 후기 프로모터, SV 40 초기 폴리아데닐화 부위 (SV40), 및 E.coli 내에서의 증식에 필요한 pBR322 서열을 포함한다.

플라스미드 pMT2 CXM 는 pMT2-VWF 의 EcoRI 소화로 얻어진 것으로, ATCC 67122 의 수탁 번호로 (미국) 메사츄세트, 록빌, 아메리칸 유형 컬처 콜렉션 (ATCC) 에 기탁되어 있다. EcoRI 소화로 pMT2-VWF 내에 존재하는 cDNA 삽입체가 절단되어, 결찰 및 E.coli HB 101 또는 DH-5 를 암피실린 저항성을 갖도록 형질전환시키는 데에 사용할 수 있는 선형의 pMT2 가 생성된다. 플라스미드 pMT2 DNA 는 통상의 방법으로 제조할 수 있댜. 이어서 pMT2 CXM 은 루프아웃/인 돌연변이 생성법 (loopout/in mutagenesis) 를 이용하여 구축할 수 있다 [Morinaga, et al., Biotechonology 84 ; 636 (1984)]. 이는 SV 40 복제원점 및 pMT2 의 엔핸서 서열 부근의 Hind III 부위와 관계있는 염기 1075 내지 1145 를 제거한 것이다. 또한, 그것은 제한 엔도뉴클레아제 XhoI 의 인식 부위를 포함하는 서열을 삽입한 것이기도 하다. pMT23 이라 명명한, pMT2 CXM 의 유도체는 제한 엔도뉴클레아제 PstI, Eco RI, SalI, XhoI 의 인식 부위들을 포함한 것이다. 플라스미드 pMT2 CXM 및 pMT23 DNA 는 통상의 방법으로 제조할 수 있다.

또한 pMT21 유래의 pEMC2β1 도 본 발명의 실행에 적합할 수 있다. pMT21 은 pMT2-VWF 에서 유도된 pMT2 로부터 유도된 것이다. 상기한 바와 같이 EcoRI 소화로 pMT-VWF 내에 존재하는 cDNA 삽입체가 절단되어, 결찰 및 E.coli HB 101 또는 DH-5 를 암피실린 저항성을 갖도록 형질전환시키는 데에 사용할 수 있는 선형의 pMT2 가 생성된다. 플라스미드 pMT2 CXM 는 통상의 방법으로 제조할 수 있다.

pMT21 은 다음의 두가지 부분변형에 의해서 pMT2 로부터 유도된다. 먼저, cDNA 클로닝용 G/C 테일링(tailing) 의 연속된 19 개 G 잔기를 포함하는 DHFR cDNA 의 5' 비번역된 영역이 부위 76bp 를 탈락시킨다. 이 과정에서, XhoI 부위를 삽입하여 다음에 오는 DHFR 의 바로 윗쪽의 서열을 수득한다. 둘째, EcoRV 및 XbaI 에 의한 소화, DNA 폴리머라아제 I 의 클레나우 단편에 의한 처리 및 ClaI 링커 (CATCGATG) 와의 연결에 의해 유일한 ClaI 부위를 도입시킨다. 이것은 아데노바이러스 RNA (VAI) 부위의 250bp 단편을 탈락시킨 것이지만 VAI RNA 유전자의 발현 또는 기능을 방해하지 않는다. pMT21 은 EcoRI 및 XhoI 로 소화시키고, 벡터 pEMC2B1 를 유도하는데 이용한다.

EMCV 리더의 부분은 pMT2-ECAT1 [S.K. Jung, et ai, J.Virol 63:1651-1660(1989)] 으로부터 EcoRI 및 PstI 의 소화에 의해 2752 bp 단편으로 얻어진다. 이 단편을 TaqI 로 소화시켜 508 bp 의 Eco RI-TaqI 단편을 얻으며, 그것을 저용융 아가로오즈 겔상에서 전기영동시켜 정제한다. 68 bp 어답터 및 그것의 상보 가닥은 5' TaqI 돌출 말단과 뉴클레오티드 763 내지 827 의 EMC 바이러스 리더 서열과 합치하는 서열을 갖는 3' XhoI 돌출 말단으로 합성한다. 그것은 EMC 바이러스 리더내의 10 위치의 ATG 를 ATT 로 변경한 것이기고 하며 그 다음에 XhoI 부위가 온다. 세 방면의 pMT21 Eco RI-XhoI 단편, EMC 바이러스 EcoRI-TaqI 단편, 및 68 bp 올리고뉴클레오티드 어답터 TaqI-XhoI 어답터의 연결로 벡터 pEMC2β1 이 결과된다.

상기 벡터는 포유류 세포에서 원하는 cDNA 를 높은 수준으로 발현시키는데 적합한 관련이 있는, SV 40 복제원점 및 엔핸서, 아데노바이러스 주 후기 프로모터, 아데노바이러스 3 부할 리더 서열의 cDNA 복사물, 소(小) 혼성 개재 서열, SV40 폴리아데닐화 신호 및 아데노바이러스 VAI 유전자, DHFR 및 β-락타마아제 마커 및 EMC 서열을 함유한다.

벡터의 구축은 BMP-12 DNA 서열의 부분변형이 수반될 수 있다. 예로서, BMP-12 cDNA 는 코딩 영역의 5' 및 3' 말단상의 비코딩 뉴클레오티드를 제거함으로서 부분변형될 수 있다. 탈락된 비코딩 뉴클레오티드는 발현에 유익하다고 공지된 기타 서열로 대체되거나 되지 않을 수 있다. 상기 벡터는 BMP-12 단백질의 발현에 적합한 숙주 세포에 형질전환시킬 수 있다. 또한, SEQ ID NO : 1 의 서열 또는 BMP-12 단백질을 코딩하는 기타 서열은 SEQ ID NO : 1 의 뉴클레오티드 571 내지 882 의 성숙한 코딩 서열을 분리하여 다른 BMP 단백질의 완전한 프로펩티드를 코딩하는 5' 말단 서열에 부가시킴으로써 BMP-12 단백질이 발현될 수 있도록 조작할 수가 있다.

예를 들어, 당업자라면 BMP-2 프로펩티드를 코딩하는 DNA 서열이 성숙 BMP-12 펩티드를 코딩하는 DNA 서열에 적당한 해독틀로 연결되어 있는 DNA 벡터를 제조함으로써, BMP-2 의 프로펩티드가 성숙 BMP-12 펩티드에 기능가능한 방식으로 연결되어 있는 융합 단백질을 제조할 수 있다. 이러한 융합 단백질의 DNA 서열은 SEQ ID NO : 27 에 나타나 있다.

당업자라면 상기한 바와 같이, 박테리아 세포에 의한 세포내 또는 세포외 발현을 위한 박테리아 벡터를 만들기 위해 코딩 서열의 측면에 위치하여 있는 포유류 조절 서열을 제거하거나 코딩 서열의 측면 위치하여 있는 포유류 조절 서열을 박테리아 서열로 바꿈으로써 SEQ ID NO : 1 의 서열을 조작할 수가 있다. 또 하나의 예로서, 추가로 코딩 서열이 더 조작될 수 있다 (예를 들면, 기타 공지된 기술로 기타 공지된 링커에 연결시키거나 또는 비코딩 서열을 탈락시키거나 내부의 뉴클레오티드를 변화시킴으로써 부분변형시킴). 부분변형된 BMP-12 코딩 서열은 그 다음에 T. Taniguchi 등의Proc. Natl Acad. Sci. USA, 77:5230-5233(1980) 에 기재된 절차를 이용하여 공지된 박테리아 벡터에 삽입시킬 수 있다. 이어서 이 예의 박테리아 벡터는 박테리아 숙주 세포에 형질 전환시키고 그에 의해 BMP-12 단백질이 발현될 수 가 있다. 박테리아 세포내의 BMP-12 단백질의 세포외 발현을 얻기 위한 전략에 대해서는 유럽 특허 출원 EPA 제 177,343 호를 보라.

곤충 세포에서의 발현을 위한 곤충 벡터의 구축 [예를들면, 공고된 유럽 특허 출원 제 155,476 호에 기재된 처리법을 보라] 에도 유사한 조작을 행할 수 있다. 또한 이스트 벡터는 이스트 세포에 의한 본 발명의 인자의 세포내 또는 세포외 발현을 위한 이스트 조절 서열을 이용하여 구축할 수 있다 [예를들면 공고된 PCT 출원 WO86/00639 및 유럽 특허 출원 EPA 123,289 에 기재된 처리법을 보라].

포유류 세포에서 본 발명의 BMP-12 단백질을 높은 수준으로 생성시키기위한 방법은 이종 BMP-12 유전자의 다중 복사물을 함유하는 세포의 구축을 수반할 수 있다. 이종 유전자는 증폭가능한 마커, 예를들면 디히드로폴레이트 환원효소(DHFR) 유전자에 연결되며 이 경우 증가된 유전자 복사무을 함유하는 세포는 Kaufman & Sharp, [J. Mol. Biol., 159:601-629 (1982)] 의 처리법에 따라 메토트렉세이트(MTX) 의 농도를 증가시킬때의 증식으로 선택할 수가 있다. 이러한 접근방법은 여러 다른 세포유형들에 사용할 수 있다.

예를들면, 발현을 가능케하는 기타 플라스미드 서열과 연계하여 기능하는 본 발명의 BMP-12 의 DNA 서열을 함유하는 플라스미드와 DHFR 발현 플라스미드 pAdA26SV(A)3 [Kaufman & Sharp, J. Mol. Biol., 159:601-629 (1982)] 을 칼슘 포스페이트 공침전법 및 트랜스팩션, 일렉트로폴이션 또는 프로토플라스트 융합법 등을 포함한 다양한 방법으로, DHFR-결핍 CHO 세포, DUKX-BII 에 동시-도입시킬 수 있다. DHFR 발현 형질 전환체는 투석시킨 태아 소혈청이 든 알파 배지에서의 성장으로 선별하고, 이어서 Kaufman & Sharp, J. Mol. Biol., 159:601-629 (1982) 에 기재된 것과 같이 MTX 의 농도를 증가시킬 때 (예를 들어, 0.02, 0.2, 1.0 및 5 uM MTX 의 순차적 단계) 의 성장에 의한 증폭으로 선별한다. 형질 전환체를 클론하고, 생물학적으로 활성인 BMP-12 발현을 하기 실시예 5 에 기재된 로젠법-변형 삼파트-레디 쥐 검정법으로 모니터한다. BMP-12 발현은 MTX 저항성의 수준이 증가함에 따라 증가하여야 한다. BMP-12 폴리펩티드는 [35S] 메티오닌 또는 시스테인에 의한 펄스 라벨링 및 폴리아크릴아미드 겔 전기영동법 등의 당해 분야에 공지된 표준 기법들을 사용하여 그 특성을 구명한다. 기타 관련 BMP-12 단백질의 제조에도 유사한 처리법을 따를 수 있다.

실시예 3실시예 3

BMP-2 프로펩티드/BMP-12 성숙 펩티드 융합체의 제조

BMP-2 프로펩티드/BMP-12 성숙 펩티드 융합체를 코딩하는 벡터의 구축을 위하여, 다음의 클로닝 절차를 이용하여 두개의 서열을 함께 융합시켰다.

먼저, SEQ ID NO : 27 의 뉴클레오티드 1 내지 843 을 포함하는, 인간 BMP-2 단백질의 프로펩티드를 포함하는 DNA 제한 효소 단편을 pBMP2△EMC 로부터 절단한다. pBMP2△EMC 는 상기 벡터에서 BMP 의 비번역 5' 및 3' 서열이 탈락된 인간 BMP-2 단백질의 완전한 코딩 서열을 포함하는 lambda U20S-39 (ATCC#40345) 에서 유도된 플라스미드이다. 사용된 5' 제한 효소는 BgI II 였고 그것은 벡터 중의 뉴클레오티드 979 에서 pBMP2△EMC 를 절단한다. 사용된 3' 제한 효소는 Mae II 였으며 그것은 BMP-2 프로펩티드내의 카르복시 말단에 못미치는, 뉴클레오티드 1925 에서 pBMP2△EMC 를 절단한다. 그 결과로 얻어지는 954 bp 산물을 그후 젤 분리하여 유전자를 정제하였다. 두번째로, 인간 BMP-12 성숙 펩티드 DNA 서열의 5' 부분을 포함하는 DNA 제한 효소 단편을 pPCR1-1#2 V1-1 (ATCC#69517) 로부터 절단한다. 사용된 5' 제한효소는 Eae I 이었으며 그것은 인간 BMP-12 성숙 펩티드 서열의 N-말단의 바로 3' 쪽의 pPCR1-1#2 V1-1 을 절단한다. 그 결과로 얻어지는 259 bp 산물을 겔 분리하여 유전자를 정제시킨다. 세 번째로, 2 개의 DNA 올리고들을 설계 합성하였으며, 어닐링시 인간 BMP-12 프로펩티드의 가장자리 3' 말단 및 BMP-12 성숙 펩티드의 5' 말단의 융합 서열을 포함하는 작은 DNA 단편이 형성될 수가 있었다. 이 DNA 단편은 인간 BMP-2 프로펩티드를 포함하는 3' 제한 효소 단편과 어닐링하는 5' Mae II 상보성 접착 말단을 갖는다. 어닐링된 올리고 DNA 단편은 인간 BMP-12 의 성숙 펩티드를 포함하는 제한효소 단편의 5' 에 어닐링하는 3'Eae I 상보성 접착 말단을 갖는다. 코딩 가닥 올리고는 B2/12 로 명명하며 13 bp 의 길이를 갖는 것이다. 다음, BMP-12 성숙 펩티드 단편의 3' 말단의 123 bp 를 코딩하는 DNA 단편을 하기와 같이 수득하였다. 먼저, 뉴클레오티드 1 내지 846 을 포함하는, 인간 BMP-2 단백질의 프로펩티드를 포함하는 DNA 단편을 pBMP2△EMC 로부터 PCR 로 증폭시킨다. 5'프라이머 (올리고 655a) 는 폴리링커의 바로 5' 에 어닐링한다. 3' 프라이머(BMPpro3) 는 BMP-2 프로펩티드 3' 말단에 어닐링하고 사일런트 서열 돌연변이에 의해서 Bgl II 제한 효소 부위를 도입시킨다. 생성된 PCR 산물을 Sal I (폴리링커에서 절단함) 및 Bgl II 로 절단한다. (아미노산 서열 REKR 로 끝나는) 850bp 제한효소단편을 겔 분리하여 유전자를 정제하였다. BMP-12 성숙 펩티드는 사일런트 서열 돌연 변이에 의한, Bgl II 제한 효소 부위를 코딩하는 5' 말단에 어닐링하는 5' 프라이머(올리고 5-1)를 이용하여 PCR 증폭시켰다. 3' 프라이머 (V1-1 3) 는 BMP-12 성숙 펩티드 3' 말단에 어닐링하며 종결 코돈 이후의 Xba I 제한 효소 부위를 도입시킨다. 그 결과로서 얻어지는 PCR 산물을 Bgl II 및 Xba I 로 절단하였다. 321 bp 제한 효소 단편을 겔 분히하여 유전자를 정제하였다.

2 개의 제한 단편들은 앞서의 SalI 및 XbaI 절단 벡터에 셋방면 연결시켰다. 그 결과로서 얻어진 구축물을 PCR 유발 실수를 조사하기 위하여 시퀀싱하였고 사일런트 C → T 돌연변이가 프로펩티드내의 bp 185 에서 관찰되었다. 상기 플라스미드를 pREKRSRC 로 나타내었다. 이어서 pREKRSRC 를 BglII 및 NgoMI 로 절단하였고, 그로부터 BMP-12 성숙 서열의 마지막 123 bp 를 포함하는 벡터 단편이 분리되었다. 3 개의 제한 단편들 및 어닐링된 올리고링커를 네 군데 연결시켜 3' 말단의 성숙 절단 부위가 BMP-12 성숙 펩티드의 (TAL) 5' 말단에 융합된 BMP-2 프로펩티드를 갖는 pREKR-TAL 을 얻었다. 그 결과로서 얻어진 결찰 벡터의 코딩 서열은 SEQ ID NO : 27 에 나타나 있다.

실시예 4실시예 4

발현된 BMP-12의 생물학적 활성

상기 실시예 2 에서 수득한 발현 BMP-12 단백질의 생물학적 활성을 측정하기 위하여, 단백질을 세포 배양물로부터 회수하고 부산물로 생성된 기타 단백질성 물질 및 기타 불순물로부터 BMP-12 단백질을 분리 정제한다. 정제된 단백질은 하기 실시예 5 에 기재된 쥐 검정법에 따라 검정할 수 있다.

정제는 당업자에게 공지된 표준 기법을 이용하여 행한다.

단백질 분석은 쿠마씨 블루 또는 실버 [Oakley, 등 Anal.Biochem. 105:361 (1980)] 로 염색하는 SDS-PAGE 아크릴아미드 [Laemmli, Nature 227:680 (1970)] 및 면역블롯 [Towbin, et al. Proc. natl. Acad. Sci. USA 76:4350 (1979)] 등의 표준 기법을 사용하여 행한다.

실시예 5실시예 5

로젠법-변형 삼파스-레디 검정법

Sampath & Reddi, [Proc. Natl. Acad. Sci. USA ,80: 6591-6595(1983)] 에 기재된 쥐 전위 이식 검정의 부분변형법을 이용하여 BMP-12 단백질의 활성을 평가한다. 이 부분변형 검정법을 여기서 로젠법 변형 삼파스-레디 검정법이라 부른다. 이 검정법은 BMP 의 골 및 연골-유도 활성을 평가하는데 널리 사용되와 왔다. 삼파스-레디 처리법의 에탄올 침강 단계는 분석하과자 하는 분획물의 물에 대한 투석 (조성물이 용액인 경우) 또는 투석여과(diafiltering) (조성물이 현탁액인 경우) 로 바꾼 것이다. 어어서 용액 또는 현탁액을 0.1% TFA 로 평형호시킨다. 그 결과로서 얻어진 용액을 20 mg 의 쥐 매트릭스에 첨가한다. 단백질로 처리되지 않은 가짜 쥐 매트릭스 샘플을 대조군으로 한다. 이 물질을 동결 건조시키고 그 결과로서 얻어진 분말을 # 5 젤라틴 캡슐에 봉입한다. 이 캡슐을 생후 21-49 일된 수컷 롱 에반스 쥐의 복부 흡곽부에 피하 이식시킨다. 이식물을 10 일후 떼어낸다. 각각의 이식물의 절단면을 고정시키고, 가공처리하여 조직 분석을 한다. 1㎛ 글리콜메타크릴레이트 절단면을 폰 코사 및 산 푸신(fuschin) 으로 염색시키고 각각의 이식물에 존재하는 유도된 건/인대-류 조직 형성의 양을 기록한다.

BMP-12 를 이식물당 1, 5, 25 및 50 ㎍ 의 용량으로 쥐에 10일간 이식하였다. 5 ㎍ 의 용량에서의 BMP-2 는 양성 대조군으로 포함되었다. 시험된 BMP-12 의 모든 용량에 대해서, 10 일 경과후 골 또는 연골 형성이 어느 이식물에서도 관찰되지 않았다. 대신, 이식물에는 태아 건과 유사한 조직이 가득차 있었으며, 이는 동일 평면에 편중된 단단하게 함께 채워진, 섬유아세포의 빡빡한 다발의 존재로 쉽게 알 수 있다. [건/인대-류 조직에 관한 기술이 있는 곳은, 예를 들면 (Ham & Cormack, Histology JB Lippincott Co. (1979), pp. 367-369 등이며, 그 상세한 내용은 여기에 참고로 하였다] 상기 발견사실들은 건/인대-류 조직이 BMP-12 를 함유하는 모든 이식물 내에 존재하는 제 2 의 일련의 검정에서 재현되었다. 대조적이게도, BMP-2 이식물은, 예상한 바와 같이, 연골 및 골 형성을 보였지만 건/인대-류 조직은 전혀 함유하고 있지 않았다.

본 발명의 BMP-12 단백질 및 관련 단백질은 상기 검정으로 활성이 평가될 수 있다.

실시예 6실시예 6

본 발명 분양의 기술의 범위내에서 약간의 부분변형을 가한 상기의 실시예에 따른 방법을 이용하여, 인간 MP 52 단백질 및 BMP-13 단백질의 상동체를 발현시켜 건/인대-류 조직 유도 활성에 대한 검정을 하였다. 모든 단백질들은 인간 BMP-12 의 상기한 결과와 유사한, 비교할 만한 결과를 나타낸다.

전술한 설명으로 본발명의 현재 바람직한 구현예를 상술하였다. 본 발명의 실제에 있어 많은 부분가감 및 변형은 상기 설명들을 고려하면 당업자에게 이루어질 수 있는 것으로 생각된다. 이러한 부분가감 및 변형은 여기에 첨부된 특허청구의 범위내에 포함하는 것으로 믿어진다. 여기에서 논의된 모든 참고문헌들의 상세한 내용은 여기에서 참조로 하였다.

서열표

(1) 일반 정보 :

(i) 출원인: 제네틱스 인스티튜트, 인코포레이티드

프레지던트 앤드 펠로우스 오브 하버드 칼리지

(ii) 발명의 명칭 : 건-유도 조성물

(iii) 서열 번호 : 35

(iv) 통신 주소 :

(A) 수취인 : 제네틱스 인스티튜트, 인코포레이트

(B) 거리 : 87 캐임브리지파크 드라이브

(C) 시 : 캐임브리지

(D) 주 : 메사츄세츠

(E) 나라 : 미국

(F) 우편번호 : 02140

(v) 컴퓨터 판독형 :

(A) 매체형 : 플로피 디스크

(B) 컴퓨터 : IBM PC 호환식

(C) 작동 시스템 : PC-DOS/MS-DOS

(D) 소프트웨어 : 패턴트인 릴리스 #1.0, 버전 #1.25

(vi) 현재의 출원자료 :

(A) 출원 번호 :

(B) 출원일 : 동봉

(C) 분류:

(vii) 선(先) 출원자료 :

(A) 출원번호 : US 08/164,103

(B) 출원일 : 1993. 12. 7

(C) 출원번호 : US 08/217,780

(D) 출원일 : 1994. 3. 25

(E) 출원 번호 : US 08/333,576

(F) 출원일 : 1994. 11. 2

(viii) 소송대리인/대리인 정보 :

(A) 이름 : 라자르, 스티븐 알.

(B) 등록 번호 : 32,618

(C) 참조/사건 일람 번호 : 5202D-PCT

(ix)전기통신정보 :

(A) 전화 : 617 498-8260

(B) 팩시밀리 : 617 876-5851

(2) SEQ ID NO : 1 에 대한 정보 :

(i) 서열 특징 :

(A) 길이 : 926 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vi) 최초의 원천 :

(A) 유기체 : 호모 사피엔스

(vii) 직접얻을 수 있는 원천 :

(B) 클론 : V1-1

(ix) 특징 :

(A) 이름/키 : 매트_펩티드

(B) 위치 : 571..882

(ix) 특징 :

(A) 이름/키 : CDS

(B) 위치 : 1..882

(xi) 서열의 기술 : SEQ ID NO : 1 :

```
GCG CGT AAT ACG ACT CAC TAT AGG GCG AAT TGG GTA CGG GGC CCA GGC          48
Ala Arg Asn Thr Thr His Tyr Arg Ala Asn Trp Val Arg Gly Pro Gly
-190                 -185                 -180                 -175

AGC TGG ACT TCT CCG CCG TTG CTG CTG CTG TCC ACG TGC CCG GGC GCC          96
Ser Trp Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys Pro Gly Ala
                -170                 -165                 -160

GCC CGA GCG CCA CGC CTG CTG TAC TCG CGG GCA GCT GAG CCC CTA GTC         144
Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val
            -155                 -150                 -145

GGT CAG CGC TGG GAG GCG TTC GAC GTG GCG GAC GCC ATG AGG CGC CAC         192
Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His
        -140                 -135                 -130

CGT CGT GAA CCG CGC CCC CCC CGC GCG TTC TGC CTC TTG CTG CGC GCA         240
Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Leu Arg Ala
    -125                 -120                 -115

GTG GCA GGC CCG GTG CCG AGC CCG TTG GCA CTG CGG CGA CTG GGC TTC         288
Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe
-110                 -105                 -100                 -95

GGC TGG CCG GGC GGA GGG GGC TCT GCG GCA GAG GAG CGC GCG GTG CTA         336
Gly Trp Pro Gly Gly Gly Gly Ser Ala Ala Glu Glu Arg Ala Val Leu
                -90                  -85                  -80

GTC GTC TCC TCC CGC ACG CAG AGG AAA GAG AGC TTA TTC CGG GAG ATC         384
Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile
            -75                  -70                  -65

CGC GCC CAG GCC CGC GCG CTC GGG GCC GCT CTG GCC TCA GAG CCG CTG         432
Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser Glu Pro Leu
        -60                  -55                  -50

CCC GAC CCA GGA ACC GGC ACC GCG TCG CCA AGG GCA GTC ATT GGC GGC         480
Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val Ile Gly Gly
    -45                  -40                  -35

CGC AGA CGG AGG AGG ACG GCG TTG GCC GGG ACG CGG ACA GCG CAG GGC         528
Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr Ala Gln Gly
-30                  -25                  -20                  -15

AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC AGG GGC CGG AGC CGC         576
Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg
                -10                   -5                    1

TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG GAG CTC GGC TGG GAC         624
Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp
        5                   10                  15

GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG TAC CAC TGC GAG GGC         672
Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly
    20                  25                  30

CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG CCC ACC AAC CAT GCC         720
Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
35                  40                  45                  50
```

```
ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA CCA GAC GCG GCG CCG GCC         768
Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala
                55                  60                  65

TCC TGC TGT GTG CCA GCG CGC CTC AGC CCC ATC AGC ATC CTC TAC ATC         816
Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile
            70                  75                  80

GAC GCC GCC AAC AAC GTT GTC TAC AAG CAA TAC GAG GAC ATG GTG GTG         864
Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
        85                  90                  95

GAG GCC TGC GGC TGC AGG TAGCGCGCGG GCCGGGGAGG GGGCAGCCAC                912
Glu Ala Cys Gly Cys Arg
    100

GCGGCCGAGG ATCC                                                        926
```

(2) SEQ ID NO : 2 에 관한 정보 :

(i) 서열의 특징 :

(A) 길이 : 294 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 2 :

```
Ala Arg Asn Thr Thr His Tyr Arg Ala Asn Trp Val Arg Gly Pro Gly
-190                 -185                 -180                 -175

Ser Trp Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys Pro Gly Ala
            -170                 -165                 -160

Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val
        -155                 -150                 -145

Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His
    -140                 -135                 -130

Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Leu Arg Ala
-125                 -120                 -115

Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe
-110                 -105                 -100                 -95

Gly Trp Pro Gly Gly Gly Gly Ser Ala Ala Glu Glu Arg Ala Val Leu
            -90                  -85                  -80

Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile
        -75                  -70                  -65

Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser Glu Pro Leu
    -60                  -55                  -50

Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val Ile Gly Gly
-45                  -40                  -35

Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr Ala Gln Gly
-30                  -25                  -20                  -15

Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg
            -10                  -5                   1
```

```
Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp
 5                   10                  15

Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly
     20                  25                  30

Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala
 35                  40                  45                  50

Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala
                55                  60                  65

Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile
            70                  75                  80

Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val
        85                  90                  95

Glu Ala Cys Gly Cys Arg
    100
```

(2) SEQ ID NO : 3 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 1207 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vi) 최초의 원천 :

(A) 유기체 : 호모 사피엔스

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : MP52

(ix) 특징:

(A) 이름/키 : CDS

(B) 위치 : 845..1204

(xi) 서열의 기술 : SEQ ID NO : 3 :

```
ACCGGGCGGC CCTGAACCCA AGCCAGGACA CCCTCCCCAA ACAAGGCAGG CTACAGCCCG      60
GACTGTGACC CCAAAAGGAC AGCTTCCCGG AGGCAAGGCA CCCCCAAAAG CAGGATCTGT     120
CCCCAGCTCC TTCCTGCTGA AGAAGGCCAG GGAGCCCGGG CCCCCACGAG AGCCCAAGGA     180
GCCGTTTCGC CCACCCCCCA TCACACCCCA CGAGTACATG CTCTCGCTGT ACAGGACGCT     240
GTCCGATGCT GACAGAAAGG GAGGCAACAG CAGCGTGAAG TTGGAGGCTG GCCTGGCCAA     300
CACCATCACC AGCTTTATTG ACAAAGGGCA AGATGACCGA GGTCCCGTGG TCAGGAAGCA     360
GAGGTACGTG TTTGACATTA GTGCCCTGGA GAAGGATGGG CTGCTGGGGG CCGAGCTCCG     420
GATCTTGCGG AAGAAGCCCT CGGACACGGC CAAGCCAGCG GCCCCCGGAG GCGGGCGGGC     480
TGCCCAGCTG AAGCTGTCCA GCTGCCCCAG CGGCCGGCAG CCGGCCTCCT TGCTGGATGT     540
GCGCTCCGTG CCAGGCCTGG ACGGATCTGG CTGGGAGGTG TTCGACATCT GGAAGCTCTT     600
CCGAAACTTT AAGAACTCGG CCCAGCTGTG CCTGGAGCTG GAGGCCTGGG AACGGGGCAG     660
GGCCGTGGAC CTCCGTGGCC TGGGCTTCGA CCGCGCCGCC CGGCAGGTCC ACGAGAAGGC     720
```

```
CCTGTTCCTG GTGTTTGGCC GCACCAAGAA ACGGGACCTG TTCTTTAATG AGATTAAGGC      780

CCGCTCTGGC CAGGACGATA AGACCGTGTA TGAGTACCTG TTCAGCCAGC GGCGAAAACG      840

GCGG GCC CCA CTG GCC ACT CGC CAG GGC AAG CGA CCC AGC AAG AAC CTT      889
     Ala Pro Leu Ala Thr Arg Gln Gly Lys Arg Pro Ser Lys Asn Leu
       1               5                  10                  15

AAG GCT CGC TGC AGT CGG AAG GCA CTG CAT GTC AAC TTC AAG GAC ATG      937
Lys Ala Arg Cys Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met
                   20                  25                  30

GGC TGG GAC GAC TGG ATC ATC GCA CCC CTT GAG TAC GAG GCT TTC CAC      985
Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His
               35                  40                  45

TGC GAG GGG CTG TGC GAG TTC CCA TTG CGC TCC CAC CTG GAG CCC ACG     1033
Cys Glu Gly Leu Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr
           50                  55                  60

AAT CAT GCA GTC ATC CAG ACC CTG ATG AAC TCC ATG GAC CCC GAG TCC     1081
Asn His Ala Val Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser
       65                  70                  75

ACA CCA CCC ACC TGC TGT GTG CCC ACG CGG CTG AGT CCC ATC AGC ATC     1129
Thr Pro Pro Thr Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile
   80                  85                  90                  95

CTC TTC ATT GAC TCT GCC AAC AAC GTG GTG TAT AAG CAG TAT GAG GAC     1177
Leu Phe Ile Asp Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp
               100                 105                 110

ATG GTC GTG GAG TCG TGT GGC TGC AGG TAG                             1207
Met Val Val Glu Ser Cys Gly Cys Arg
           115                 120
```

(2) SEQ ID NO : 4 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 120 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 4 :

```
Ala Pro Leu Ala Thr Arg Gln Gly Lys Arg Pro Ser Lys Asn Leu Lys
  1               5                  10                  15

Ala Arg Cys Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met Gly
             20                  25                  30

Trp Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys
         35                  40                  45

Glu Gly Leu Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn
     50                  55                  60

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

Pro Pro Thr Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu
                 85                  90                  95
```

```
Phe Ile Asp Ser Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met
            100                 105                 110

Val Val Glu Ser Cys Gly Cys Arg
        115                 120
```

(2) SEQ ID NO : 5 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 128 bp

(B) 유형 : 핵산

(C) 가닥 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vi) 최초의 원천 :

(A) 유기체 : 호모 사피엔스

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : V1-1 단편

(ix) 특징

(A) 이름/키 : CDS

(B) 위치 : 28..102

(xi) 서열의 기술 : SEQ ID NO : 5 :

```
GGATCCTGGA AGGATTGGAT CATTGCG CCG CTG GAC TAC GAG GCG TAC CAC          51
                              Pro Leu Asp Tyr Glu Ala Tyr His
                                1               5

TGC GAG GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG CCC ACC        99
Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr
     10                  15                  20

AAC CACGCTATAG TCCAAACCTT TCTAGA                                     128
Asn
 25
```

(2) SEQ ID NO : 6 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 25 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 6 :

```
Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Leu Cys Asp Phe Pro
 1               5                  10                  15

Leu Arg Ser His Leu Glu Pro Thr Asn
            20                  25
```

(2) SEQ ID NO : 7 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 128 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vi) 최초의 원천 :

(A) 유기체 : 호모 사피엔스

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : VL-1

(ix) 특징:

(A) 이름/키 : CDS

(B) 위치 : 28..102

(xi) 서열의 기술 : SEQ ID NO : 7 :

```
GGATCCTGGG ATGACTGGAT TATGGCG CCG CTG GAC TAC GAG GCG TAC CAC          51
                              Pro Leu Asp Tyr Glu Ala Tyr His
                                1               5

TGC GAG GGT GTA TGC GAC TTC CCG CTG CGC TCG CAC CTG GAG CCC ACC        99
Cys Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr
     10                  15                  20

AAC CACGCCATGC TACAAACGCT TCTAGA                                     128
Asn
 25
```

(2) SEQ ID NO : 8 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 25 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 8 :

```
Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro
 1               5                  10                  15

Leu Arg Ser His Leu Glu Pro Thr Asn
            20                  25
```

(2) SEQ ID NO : 9 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 3585 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : pALV1-781

(xi) 서열의 기술 : SEQ ID NO : 9 :

```
CTAACTACCC AACTCAAAAA AAAAAAAAAA AAAAACCCCC TCTAACCCCC ATTGACGAAA     60

GGGCCTCGTG ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC    120

GTCAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT ATTTGTTTAT TTTTCTAAAT    180
```

ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG 240
AAAAAGGAAG AGTATGAGTA TTCAACATTT CCGTGTCGCC CTTATTCCCT TTTTTGCGGC 300
ATTTTGCCTT CCTGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA 360
TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA 420
GAGTTTTCGC CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG 480
CGCGGTATTA TCCCGTATTG ACGCCGGGCA AGAGCAACTC GGTCGCCGCA TACACTATTC 540
TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC 600
AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCGG CCAACTTACT 660
TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA 720
TGTAACTCGC CTTGATCGTT GGGAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG 780
TGACACCACG ATGCCTGTAG CAATGGCAAC AACGTTGCGC AAACTATTAA CTGGCGAACT 840
ACTTACTCTA GCTTCCCGGC AACAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG 900
ACCACTTCTG CGCTCGGCCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG 960
TGAGCGTGGG TCTCGCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT 1020
CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC 1080
TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAACTGTCA GACCAAGTTT ACTCATATAT 1140
ACTTTAGATT GATTTAAAAC TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT 1200
TGATAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTTCG TTCCACTGAG CGTCAGACCC 1260
CGTAGAAAAG ATCAAAGGAT CTTCTTGAGA TCCTTTTTTT CTGCGCGTAA TCTGCTGCTT 1320
GCAAACAAAA AAACCACCGC TACCAGCGGT GGTTTGTTTG CCGGATCAAG AGCTACCAAC 1380
TCTTTTTCCG AAGGTAACTG GCTTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT 1440
GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT 1500
GCTAATCCTG TTACCAGTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA 1560
CTCAAGACGA TAGTTACCGG ATAAGGCGCA GCGGTCGGGC TGAACGGGGG GTTCGTGCAC 1620
ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCATTG 1680
AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GGCGGACAGG TATCCGGTAA GCGGCAGGGT 1740
CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGGAAAC GCCTGGTATC TTTATAGTCC 1800
TGTCGGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG 1860
GAGCCTATGG AAAAACGCCA GCAACGCGGC CTTTTTACGG TTCCTGGCCT TTTGCTGGCC 1920
TTTTGCTCAC ATGTTCTTTC CTGCGTTATC CCCTGATTCT GTGGATAACC GTATTACCGC 1980
CTTTGAGTGA GCTGATACCG CTCGCCGCAG CCGAACGACC GAGCGCAGCG AGTCAGTGAG 2040
CGAGGAAGCG GAAGAGCGCC CAATACGCAA ACCGCCTCTC CCCGCGCGTT GGCCGATTCA 2100
TTAATGCAGA ATTGATCTCT CACCTACCAA ACAATGCCCC CCTGCAAAAA ATAAATTCAT 2160
ATAAAAAACA TACAGATAAC CATCTGCGGT GATAAATTAT CTCTGGCGGT GTTGACATAA 2220

```
ATACCACTGG CGGTGATACT GAGCACATCA GCAGGACGCA CTGACCACCA TGAAGGTGAC    2280
GCTCTTAAAA ATTAAGCCCT GAAGAAGGGC AGCATTCAAA GCAGAAGGCT TTGGGGTGTG    2340
TGATACGAAA CGAAGCATTG GCCGTAAGTG CGATTCCGGA TTAGCTGCCA ATGTGCCAAT    2400
CGCGGGGGGT TTTCGTTCAG GACTACAACT GCCACACACC ACCAAAGCTA ACTGACAGGA    2460
GAATCCAGAT GGATGCACAA ACACGCCGCC GCGAACGTCG CGCAGAGAAA CAGGCTCAAT    2520
GGAAAGCAGC AAATCCCCTG TTGGTTGGGG TAAGCGCAAA ACCAGTTCCG AAAGATTTTT    2580
TTAACTATAA ACGCTGATGG AAGCGTTTAT GCGGAAGAGG TAAAGCCCTT CCCGAGTAAC    2640
AAAAAAACAA CAGCATAAAT AACCCCGCTC TTACACATTC CAGCCCTGAA AAAGGGCATC    2700
AAATTAAACC ACACCTATGG TGTATGCATT TATTTGCATA CATTCAATCA ATTGTTATCT    2760
AAGGAAATAC TTACATATGT CTCGTTGTTC TCGTAAACCA CTGCATGTAG ATTTTAAAGA    2820
GCTCGGCTGG GACGACTGGA TCATCGCGCC GCTGGACTAC GAGGCGTACC ACTGCGAGGG    2880
CCTTTGCGAC TTCCCTTTGC GTTCGCACCT CGAGCCCACC AACCATGCCA TCATTCAGAC    2940
GCTGCTCAAC TCCATGGCAC CAGACGCGGC GCCGGCCTCC TGCTGTGTGC CAGCGCGCCT    3000
CAGCCCCATC AGCATCCTCT ACATCGACGC CGCCAACAAC GTTGTCTACA AGCAATACGA    3060
GGACATGGTG GTGGAGGCCT GCGGCTGCAG GTAGTCTAGA GTCGACCTGC AGTAATCGTA    3120
CAGGGTAGTA CAAATAAAAA AGGCACGTCA GATGACGTGC CTTTTTTCTT GTGAGCAGTA    3180
AGCTTGGCAC TGGCCGTCGT TTTACAACGT CGTGACTGGG AAAACCCTGG CGTTACCCAA    3240
CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC    3300
ACCGATCGCC CTTCCCAACA GTTGCGCAGC CTGAATGGCG AATGGCGCCT GATGCGGTAT    3360
TTTCTCCTTA CGCATCTGTG CGGTATTTCA CACCGCATAT ATGGTGCACT CTCAGTACAA    3420
TCTGCTCTGA TGCCGCATAG TTAAGCCAGC CCCGACACCC GCCAACACCC GCTGACGCGC    3480
CCTGACGGGC TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA    3540
GCTGCATGTG TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGA                    3585
```

(2) SEQ ID NO : 10 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 272 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vi) 최초의 원천 :

(A) 유기체 : 마우스

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : mV1

(ix) 특징:

(A) 이름/키 : CDS

(B) 위치 : 28..243

(xi) 서열의 기술 : SEQ ID NO : 10 :

```
GGATCCAAGG AGCTCGGCTG GGACGAC TGG ATC ATC GCG CCA TTA GAC TAC          51
                              Trp Ile Ile Ala Pro Leu Asp Tyr
                                1               5

GAG GCA TAC CAC TGC GAG GGC GTT TGC GAC TTT CCT CTG CGC TCG CAC        99
Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His
     10                  15                  20

CTG GAG CCT ACC AAC CAC GCC ATC ATT CAG ACG CTG CTC AAC TCC ATG       147
Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met
 25                  30                  35                  40

GCG CCC GAC GCT GCG CCA GCC TCC TGC TGC GTG CCC GCA AGG CTC AGT       195
Ala Pro Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser
                 45                  50                  55

CCC ATC AGC ATT CTC TAC ATC GAT GCC GCC AAC AAC GTG GTC TAC AAG       243
Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys
            60                  65                  70

CAATACGAGG ACATGGTGGT GGGGAATTC                                       272
```

(2) SEQ ID NO : 11 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 72 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 11 :

```
Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val
  1               5                  10                  15

Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile
            20                  25                  30

Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala Ser
        35                  40                  45

Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp
    50                  55                  60

Ala Ala Asn Asn Val Val Tyr Lys
65                  70
```

(2) SEQ ID NO : 12 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 272 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vi) 최초의 원천 :

(A) 유기체 : 마우스

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : mV2

(ix) 특징:

(A) 이름/키 : CDS

(B) 위치 : 28..243

(xi) 서열의 기술 : SEQ ID NO : 12 :

```
GGATCCAAGG AGCTCGGCTG GGACGAC TGG ATT ATC GCG CCC CTA GAG TAC          51
                                  Trp Ile Ile Ala Pro Leu Glu Tyr
                                    1               5

GAG GCC TAT CAC TGC GAG GGC GTG TGC GAC TTT CCG CTG CGC TCG CAC        99
Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His
     10                  15                  20

CTT GAG CCC ACT AAC CAT GCC ATC ATT CAG ACG CTG ATG AAC TCC ATG       147
Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu Met Asn Ser Met
 25                  30                  35                  40

GAC CCG GGC TCC ACC CCG CCT AGC TGC TGC GTT CCC ACC AAA CTG ACT       195
Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val Pro Thr Lys Leu Thr
                 45                  50                  55

CCC ATT AGC ATC CTG TAC ATC GAC GCG GGC AAT AAT GTA GTC TAC AAG       243
Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn Asn Val Val Tyr Lys
             60                  65                  70

CAATACGAGG ACATGGTGGT GGGGAATTC                                       272
```

(2) SEQ ID NO : 13 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 72 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 13 :

```
Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly Val
 1               5                  10                  15

Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile
            20                  25                  30

Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro Ser
        35                  40                  45

Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp
    50                  55                  60

Ala Gly Asn Asn Val Val Tyr Lys
65                  70
```

(2) SEQ ID NO : 14 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 272 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vi) 최초의 원천 :

(A) 유기체 : 마우스

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : mV9

(ix) 특징

(A) 이름/키 : CDS

(B) 위치 : 28..243

(xi) 서열의 기술 : SEQ ID NO : 14 :

```
Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys Glu Gly Leu
 1               5                  10                  15

Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Val
            20                  25                  30

Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr Pro Pro Thr
        35                  40                  45

Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp
    50                  55                  60

Ser Ala Asn Asn Val Val Tyr Lys
65                  70
```

(2) SEQ ID NO : 15 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 72 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 15 :

```
Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe His Cys Glu Gly Leu
 1               5                  10                  15

Cys Glu Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Val
            20                  25                  30

Ile Gln Thr Leu Met Asn Ser Met Asp Pro Glu Ser Thr Pro Pro Thr
        35                  40                  45

Cys Cys Val Pro Thr Arg Leu Ser Pro Ile Ser Ile Leu Phe Ile Asp
    50                  55                  60

Ser Ala Asn Asn Val Val Tyr Lys
65                  70
```

(2) SEQ ID NO : 16 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 7 아미노산

(B) 유형 : 아미노산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : 펩티드

(vi) 최초의 원천 :

(A) 유기체 : BMP/TGF-β 컨센서스 서열

(xi) 서열의 기술 : SEQ ID NO : 16 :

Trp Xaa Asp Trp Ile Xaa Ala
1 5

(2) SEQ ID NO : 17 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 27 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : 올리고뉴클레오티드 #1

(xi) 서열의 기술 : SEQ ID NO : 17 :

CGGATCCTGG VANGAYTGGA THRTNGC

(2) SEQ ID NO : 18 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 6 아미노산

(B) 유형 : 아미노산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : 펩티드

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : BMP/TGF - β 컨센서스 서열

(xi) 서열의 기술 : SEQ ID NO : 18 :

His Ala Ile Xaa Gln Thr
1 5

(2) SEQ ID NO : 19 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 28 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : 올리고뉴클레오티드 #2

(xi) 서열의 기술 : SEQ ID NO : 19 :

TTTCTAGAAR NGTYTGNACD ATNGCRTG 28

(2) SEQ ID NO : 20 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 40 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : 올리고뉴클레오티드 #3

(xi) 서열의 기술 : SEQ ID NO : 20 :

CCACTGCGAG GGCCTTTGCG ACTTCCCTTT GCGTTCGCAC 40

(2) SEQ ID NO : 21 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 29 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(A) 라이브러리 : 올리고뉴클레오티드 #4

(xi) 서열의 기술 : SEQ ID NO : 21 :

TGCGGATCCA GCCGCTGCAG CCGCAAGCC 29

(2) SEQ ID NO : 22 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 29 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : 올리고뉴클레오티드 #5

(xi) 서열의 기술 : SEQ ID NO : 22 :

GACTCTAGAC TACCTGCAGC CGCAGGCCT 29

(2) SEQ ID NO : 23 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 28 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(A) 라이브러리 : 올리고뉴클레오티드 #6

(xi) 서열의 기술 : SEQ ID NO : 23 :

GCGGATCCAA GGAGCTCGGC TGGGACGA 28

(2) SEQ ID NO : 24 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 28 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : 올리고뉴클레오티드 #7

(xi) 서열의 기술 : SEQ ID NO : 24 :

GGAATTCCCC ACCACCATGT CCTCGTAT 28

(2) SEQ ID NO : 25 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 1171 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : 인간 VL-1 단백질

(ix) 특징:

(A) 이름/키 : CDS

(B) 위치 : 2..964

(vii) 특징 :

(A) 이름/키 : 매트_펩티드

(B) 위치 : 605..964

(xi) 서열의 기술 : SEQ ID NO : 25 :

```
G AAT TCG GAT CTC TCG CAC ACT CCT CTC CGG AGA CAG AAG TAT TTG          46
  Asn Ser Asp Leu Ser His Thr Pro Leu Arg Arg Gln Lys Tyr Leu
  -201-200              -195              -190

TTT GAT GTG TCC ATG CTC TCA GAC AAA GAA GAG CTG GTG GGC GCG GAG        94
Phe Asp Val Ser Met Leu Ser Asp Lys Glu Glu Leu Val Gly Ala Glu
    -185              -180              -175
```

```
CTG CGG CTC TTT CGC CAG GCG CCC TCA GCG CCC TGG GGG CCA CCA GCC        142
Leu Arg Leu Phe Arg Gln Ala Pro Ser Ala Pro Trp Gly Pro Pro Ala
-170                 -165                 -160                 -155

GGG CCG CTC CAC GTG CAG CTC TTC CCT TGC CTT TCG CCC CTA CTG CTG        190
Gly Pro Leu His Val Gln Leu Phe Pro Cys Leu Ser Pro Leu Leu Leu
                -150                 -145                 -140

GAC GCG CGG ACC CTG GAC CCG CAG GGG GCG CCG CCG GCC GGC TGG GAA        238
Asp Ala Arg Thr Leu Asp Pro Gln Gly Ala Pro Pro Ala Gly Trp Glu
            -135                 -130                 -125

GTC TTC GAC GTG TGG CAG GGC CTG CGC CAC CAG CCC TGG AAG CAG CTG        286
Val Phe Asp Val Trp Gln Gly Leu Arg His Gln Pro Trp Lys Gln Leu
        -120                 -115                 -110

TGC TTG GAG CTG CGG GCC GCA TGG GGC GAG CTG GAC GCC GGG GAG GCC        334
Cys Leu Glu Leu Arg Ala Ala Trp Gly Glu Leu Asp Ala Gly Glu Ala
    -105                 -100                 -95

GAG GCG CGC GCG CGG GGA CCC CAG CAA CCG CCG CCC CCG GAC CTG CGG        382
Glu Ala Arg Ala Arg Gly Pro Gln Gln Pro Pro Pro Pro Asp Leu Arg
-90                  -85                  -80                  -75

AGT CTG GGC TTC GGC CGG AGG GTG CGG CCT CCC CAG GAG CGG GCC CTG        430
Ser Leu Gly Phe Gly Arg Arg Val Arg Pro Pro Gln Glu Arg Ala Leu
                -70                  -65                  -60

CTG GTG GTA TTC ACC AGA TCC CAG CGC AAG AAC CTG TTC GCA GAG ATG        478
Leu Val Val Phe Thr Arg Ser Gln Arg Lys Asn Leu Phe Ala Glu Met
            -55                  -50                  -45

CGC GAG CAG CTG GGC TCG GCC GAG GCT GCG GGC CCG GGC GCG GGC GCC        526
Arg Glu Gln Leu Gly Ser Ala Glu Ala Ala Gly Pro Gly Ala Gly Ala
        -40                  -35                  -30

GAG GGG TCG TGG CCG CCG CCG TCG GGC GCC CCG GAT GCC AGG CCT TGG        574
Glu Gly Ser Trp Pro Pro Pro Ser Gly Ala Pro Asp Ala Arg Pro Trp
    -25                  -20                  -15

CTG CCC TCG CCC GGC CGC CGG CGG CGG CGC ACG GCC TTC GCC AGT CGC        622
Leu Pro Ser Pro Gly Arg Arg Arg Arg Arg Thr Ala Phe Ala Ser Arg
-10                  -5                   1                    5

CAT GGC AAG CGG CAC GGC AAG AAG TCC AGG CTA CGC TGC AGC AAG AAG        670
His Gly Lys Arg His Gly Lys Lys Ser Arg Leu Arg Cys Ser Lys Lys
            10                   15                   20

CCC CTG CAC GTG AAC TTC AAG GAG CTG GGC TGG GAC GAC TGG ATT ATC        718
Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile
        25                   30                   35

GCG CCC CTG GAG TAC GAG GCC TAT CAC TGC GAG GGT GTA TGC GAC TTC        766
Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe
    40                   45                   50

CCG CTG CGC TCG CAC CTG GAG CCC ACC AAC CAC GCC ATC ATC CAG ACG        814
Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr
55                   60                   65                   70

CTG ATG AAC TCC ATG GAC CCC GGC TCC ACC CCG CCC AGC TGC TGC GTG        862
Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val
                75                   80                   85

CCC ACC AAA TTG ACT CCC ATC AGC ATT CTA TAC ATC GAC GCG GGC AAT        910
Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn
            90                   95                   100

AAT GTG GTC TAC AAG CAG TAC GAG GAC ATG GTG GTG GAG TCG TGC GGC        958
Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys Gly
        105                  110                  115

TGC AGG TAGCGGTGCC TTTCCCGCCG CCTTGGCCCG GAACCAAGGT GGGCCAAGGT        1014
Cys Arg
    120

CCGCCTTGCA GGGGAGGCCT GGCTGCAGAG AGGCGGAGGA GGAAGCTGGC GCTGGGGGAG    1074

GCTGAGGGTG AGGGAACAGC CTGGATGTGA GAGCCGGTGG GAGAGAAGGG AGCGCACCTT    1134

CCCAGTAACT TCTACCTGCC AGCCCAGAGG GAAATAT                              1171
```

(2) SEQ ID NO : 26 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 321 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(vi) 최초의 원천 :

(xi) 서열의 기술 : SEQ ID NO : 26 :

```
Asn Ser Asp Leu Ser His Thr Pro Leu Arg Arg Gln Lys Tyr Leu Phe
-201 -200                -195                -190

Asp Val Ser Met Leu Ser Asp Lys Glu Glu Leu Val Gly Ala Glu Leu
-185                 -180                 -175                 -170

Arg Leu Phe Arg Gln Ala Pro Ser Ala Pro Trp Gly Pro Pro Ala Gly
                 -165                 -160                 -155

Pro Leu His Val Gln Leu Phe Pro Cys Leu Ser Pro Leu Leu Leu Asp
             -150                 -145                 -140

Ala Arg Thr Leu Asp Pro Gln Gly Ala Pro Pro Ala Gly Trp Glu Val
         -135                 -130                 -125

Phe Asp Val Trp Gln Gly Leu Arg His Gln Pro Trp Lys Gln Leu Cys
     -120                 -115                 -110

Leu Glu Leu Arg Ala Ala Trp Gly Glu Leu Asp Ala Gly Glu Ala Glu
-105                 -100                 -95                  -90

Ala Arg Ala Arg Gly Pro Gln Gln Pro Pro Pro Pro Asp Leu Arg Ser
                 -85                  -80                  -75

Leu Gly Phe Gly Arg Arg Val Arg Pro Pro Gln Glu Arg Ala Leu Leu
             -70                  -65                  -60

Val Val Phe Thr Arg Ser Gln Arg Lys Asn Leu Phe Ala Glu Met Arg
         -55                  -50                  -45

Glu Gln Leu Gly Ser Ala Glu Ala Ala Gly Pro Gly Ala Gly Ala Glu
     -40                  -35                  -30

Gly Ser Trp Pro Pro Pro Ser Gly Ala Pro Asp Ala Arg Pro Trp Leu
-25                  -20                  -15                  -10

Pro Ser Pro Gly Arg Arg Arg Arg Arg Thr Ala Phe Ala Ser Arg His
                 -5                   1                    5

Gly Lys Arg His Gly Lys Lys Ser Arg Leu Arg Cys Ser Lys Lys Pro
         10                   15                   20

Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala
     25                   30                   35

Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro
 40                   45                   50                   55

Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu
                 60                   65                   70

Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val Pro
             75                   80                   85

Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn Asn
         90                   95                   100

Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys Gly Cys
     105                  110                  115

Arg
120
```

(2) SEQ ID NO : 27 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 1233 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : BMP2 프로펩티드/BMP-12 성숙 펩티드를 코딩하는 DNA

(ix) 특징:

(A) 이름/키 : CDS

(B) 위치 : 1..1233

(ix) 특징:

(A) 이름/키 : 매트_펩티드

(B) 위치 : 847..1233

(xi) 서열의 기술 : SEQ ID NO : 27 :

```
ATG GTG GCC GGG ACC CGC TGT CTT CTA GCG TTG CTG CTT CCC CAG GTC        48
Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val
-282    -280                -275                -270

CTC CTG GGC GGC GCG GCT GGC CTC GTT CCG GAG CTG GGC CGC AGG AAG        96
Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys
    -265                -260                -255

TTC GCG GCG GCG TCG TCG GGC CGC CCC TCA TCC CAG CCC TCT GAC GAG       144
Phe Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu
-250                -245                -240                -235

GTC CTG AGC GAG TTC GAG TTG CGG CTG CTC AGC ATG TTC GGC CTG AAA       192
Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys
                -230                -225                -220

CAG AGA CCC ACC CCC AGC AGG GAC GCC GTG GTG CCC CCC TAC ATG CTA       240
Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu
            -215                -210                -205
```

```
GAC CTG TAT CGC AGG CAC TCA GGT CAG CCG GGC TCA CCC GCC CCA GAC      288
Asp Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp
-200                 -195                 -190

CAC CGG TTG GAG AGG GCA GCC AGC CGA GCC AAC ACT GTG CGC AGC TTC      336
His Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe
-185                 -180                 -175

CAC CAT GAA GAA TCT TTG GAA GAA CTA CCA GAA ACG AGT GGG AAA ACA      384
His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr
-170                 -165                 -160                 -155

ACC CGG AGA TTC TTC TTT AAT TTA AGT TCT ATC CCC ACG GAG GAG TTT      432
Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe
                -150                 -145                 -140

ATC ACC TCA GCA GAG CTT CAG GTT TTC CGA GAA CAG ATG CAA GAT GCT      480
Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala
            -135                 -130                 -125

TTA GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT AAT ATT TAT GAA ATC      528
Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile
        -120                 -115                 -110

ATA AAA CCT GCA ACA GCC AAC TCG AAA TTC CCC GTG ACC AGA CTT TTG      576
Ile Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu
    -105                 -100                 -95

GAC ACC AGG TTG GTG AAT CAG AAT GCA AGC AGG TGG GAA AGT TTT GAT      624
Asp Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp
-90                  -85                  -80                  -75

GTC ACC CCC GCT GTG ATG CGG TGG ACT GCA CAG GGA CAC GCC AAC CAT      672
Val Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His
                -70                  -65                  -60

GGA TTC GTG GTG GAA GTG GCC CAC TTG GAG GAG AAA CAA GGT GTC TCC      720
Gly Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser
            -55                  -50                  -45

AAG AGA CAT GTT AGG ATA AGC AGG TCT TTG CAC CAA GAT GAA CAC AGC      768
Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser
        -40                  -35                  -30

TGG TCA CAG ATA AGG CCA TTG CTA GTA ACT TTT GGC CAT GAT GGA AAA      816
Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys
    -25                  -20                  -15

GGG CAT CCT CTC CAC AAA AGA GAA AAA CGT ACG GCG TTG GCC GGG ACG      864
Gly His Pro Leu His Lys Arg Glu Lys Arg Thr Ala Leu Ala Gly Thr
-10                  -5                   1                    5

CGG ACA GCG CAG GGC AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC      912
Arg Thr Ala Gln Gly Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg
                10                   15                   20

AGG GGC CGG AGC CGC TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG      960
Arg Gly Arg Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys
            25                   30                   35

GAG CTC GGC TGG GAC GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG     1008
Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala
        40                   45                   50

TAC CAC TGC GAG GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG     1056
Tyr His Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu
55                   60                   65                   70

CCC ACC AAC CAT GCC ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA CCA     1104
Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro
                75                   80                   85

GAC GCG GCG CCG GCC TCC TGC TGT GTG CCA GCG CGC CTC AGC CCC ATC     1152
Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile
            90                   95                   100

AGC ATC CTC TAC ATC GAC GCC GCC AAC AAC GTT GTC TAC AAG CAA TAC     1200
Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr
        105                  110                  115

GAG GAC ATG GTG GTG GAG GCC TGC GGC TGC AGG                        1233
Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg
    120                  125
```

(2) SEQ ID NO : 28 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 411 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 28 :

```
Val Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His
                -70                 -65                 -60

Gly Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser
            -55                 -50                 -45

Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser
        -40                 -35                 -30

Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys
    -25                 -20                 -15

Gly His Pro Leu His Lys Arg Glu Lys Arg Thr Ala Leu Ala Gly Thr
-10                 -5                  1                   5

Arg Thr Ala Gln Gly Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg
            10                  15                  20

Arg Gly Arg Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys
        25                  30                  35

Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala
    40                  45                  50

Tyr His Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu
55                  60                  65                  70

Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro
                75                  80                  85

Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile
            90                  95                  100

Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr
        105                 110                 115

Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg
    120                 125
```

(2) SEQ ID NO : 29 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 1203 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : 뮤린 MV1

(ix) 특징:

(A) 이름/키 : CDS

(B) 위치 : 2..721

(xi) 서열의 기술 : SEQ ID NO : 29 :

```
A AAG TTC TGC CTG GTG CTG GNG NCG GTG ACG GCC TCG GAG AGC AGN          46
  Lys Phe Cys Leu Val Leu X01 X02 Val Thr Ala Ser Glu Ser X03
  1               5                  10                  15

CNG CTG GCC CTG AGA CGA CTG GGC TTC GGC TGN CCG GGC GGT GGC GAC        94
X04 Leu Ala Leu Arg Arg Leu Gly Phe Gly X05 Pro Gly Gly Gly Asp
              20                  25                  30

GGC GGC GGC ACT GCG GNC GAG GAG CGC GCG CTG TTG GTG ATC TCC TCC       142
Gly Gly Gly Thr Ala X06 Glu Glu Arg Ala Leu Leu Val Ile Ser Ser
          35                  40                  45

CGT ACG CAA AGG AAA GAG AGT CTG TTC CGG GAG ATC CGA GCC CAG GCC       190
Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile Arg Ala Gln Ala
      50                  55                  60

CGT GCT CTC CGG GCC GCT GCA GAG CCG CCA CCG GAT CCA GGA CCA GGC       238
Arg Ala Leu Arg Ala Ala Ala Glu Pro Pro Pro Asp Pro Gly Pro Gly
  65                  70                  75

GCT GGG TCA CGC AAA GCC AAC CTG GGC GGT CGC AGG CGG CAG CGG ACT       286
Ala Gly Ser Arg Lys Ala Asn Leu Gly Gly Arg Arg Arg Gln Arg Thr
80                  85                  90                  95

GCG CTG GCT GGG ACT CGG GGA GNG NAG GGA AGC GGT GGT GGC GGC GGT       334
Ala Leu Ala Gly Thr Arg Gly X07 X08 Gly Ser Gly Gly Gly Gly Gly
              100                 105                 110

GGC GGT GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GGC GCA GGC       382
Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala Gly
          115                 120                 125

AGG GGC CAC GGG CGC AGA GGC CGG AGC CGC TGC GGT CGC AAG TCA CTG       430
Arg Gly His Gly Arg Arg Gly Arg Ser Arg Cys Gly Arg Lys Ser Leu
      130                 135                 140

CAC GTG GAC TTT AAG GAG CTG GGC TGG GAC GAC TGG ATC ATC GCG CCA       478
His Val Asp Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala Pro
  145                 150                 155

TTA GAC TAC GAG GCA TAC CAC TGC GAG GGC GTT TGC GAC TTT CCT CTG       526
Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro Leu
160                 165                 170                 175

CGC TCG CAC CTG GAG CCT ACC AAC CAC GCC ATC ATT CAG ACG CTG CTC       574
Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu
              180                 185                 190

AAC TCC ATG GCG CCC GAC GCT GCG CCA GCC TCC TGC TGC GTG CCC GCA       622
Asn Ser Met Ala Pro Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala
          195                 200                 205

AGG CTC AGT CCC ATC AGC ATT CTC TAC ATC GAT GCC GCC AAC AAC GTG       670
Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val
      210                 215                 220

GTC TAC AAG CAG TAC GAA GAC ATG GTG GTG GAG GCC TGC GGC TGC AGG       718
Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg
  225                 230                 235

TAGCATGCGG TCTGGGGAGG GTCTGGCCGC CCAGGACCCT AGCTCAAGAG                771
240

CAGGTGTCAT CAGGCCCGAG GGACGGCGGA CTATGGCCTC TGCCAGCACA GAGGAGAGCA     831

CACAGTTAAC ACTCACATTT ACACACTCCT TCACTCACGC ACATGTTTAC CGTGGACGGC     891

AGGCGCTAAA AGCCTTGCTT ATTTGCTACC ATTGATACAA ACCTCTGTCC TTTTCGGGAG     951

AGGGAAGGGC ATCTGTGTTT ATGTTGCAGT AATTGGCACT AAATCCAAGT AGAAATGGGT    1011

TAGCATTGGA TTCTCCTTTT AGTTGGAGGC GGTGTGGCTG GATTCCTGAC GTTGGATATG    1071

GAGTGCACTG CAGGGCTGGG ATACCCAGAT TCTCTGGAGT GGGCATTGGG AACCTTCAAA    1131
```

AGTAAGGAGC CACTGGGGCT TGGGAGGGAG CACCCGGTTC CTAAACAAGT CTGATGTGTA 1191

CTGCTCAGTT TG 1203

(2) SEQ ID NO : 30 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 240 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 30 :

Lys Phe Cys Leu Val Leu X01 X02 Val Thr Ala Ser Glu Ser X03 X04
1 5 10 15

Leu Ala Leu Arg Arg Leu Gly Phe Gly X05 Pro Gly Gly Gly Asp Gly
20 25 30

Gly Gly Thr Ala X06 Glu Glu Arg Ala Leu Leu Val Ile Ser Ser Arg
35 40 45

Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile Arg Ala Gln Ala Arg
50 55 60

Ala Leu Arg Ala Ala Ala Glu Pro Pro Pro Asp Pro Gly Pro Gly Ala
65 70 75 80

Gly Ser Arg Lys Ala Asn Leu Gly Gly Arg Arg Arg Gln Arg Thr Ala
85 90 95

Leu Ala Gly Thr Arg Gly X07 X08 Gly Ser Gly Gly Gly Gly Gly Gly
100 105 110

Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Gly Ala Gly
115 120 125

Arg Gly His Gly Arg Arg Gly Arg Ser Arg Cys Gly Arg Lys Ser Leu
130 135 140

His Val Asp Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala Pro
145 150 155 160

Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro Leu
165 170 175

Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu
180 185 190

Asn Ser Met Ala Pro Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala
195 200 205

Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val
210 215 220

Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg
225 230 235 240

(2) SEQ ID NO : 31 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 1046 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(iii) 가정(HYPOTHETICAL) : 없음

(iv) 안티센스 : 없음

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : 뮤린 MV2

(ix) 특징

(A) 이름/키 : CDS

(B) 위치 : 2..790

(xi) 서열의 기술 : SEQ ID NO : 31 :

```
A AGA AAA CAA GCT TGC ATT CCT GCA GGT CCG ACT CTA AGA GGA TCC          46
  Arg Lys Gln Ala Cys Ile Pro Ala Gly Pro Thr Leu Arg Gly Ser
    1               5                  10                  15

TCA GGG ACC CAA CCC AGG CCG GCT GGG AAG TCT TTC GAC GTG TGG CAG        94
Ser Gly Thr Gln Pro Arg Pro Ala Gly Lys Ser Phe Asp Val Trp Gln
                20                  25                  30

GGC CTG CGC CCT CAG CCT TGG AAG CAG CTG TGC CTG GAG TTG CGG GCA       142
Gly Leu Arg Pro Gln Pro Trp Lys Gln Leu Cys Leu Glu Leu Arg Ala
            35                  40                  45

GCC TGG GGT GAG CTG GAC RCC GGG GAT ACG GGG GCG CGC GCG AGG GGT       190
Ala Trp Gly Glu Leu Asp X01 Gly Asp Thr Gly Ala Arg Ala Arg Gly
        50                  55                  60

CCC CAG CAG CCA CCG CCT CTG GAC CTG CGG AGT CTG GGC TTC GGT CGG       238
Pro Gln Gln Pro Pro Pro Leu Asp Leu Arg Ser Leu Gly Phe Gly Arg
    65                  70                  75

AGG GTG AGA CCG CCC CAG GAG CGC GCC CTG CTT GTA GTG TTC ACC AGA       286
Arg Val Arg Pro Pro Gln Glu Arg Ala Leu Leu Val Val Phe Thr Arg
80                  85                  90                  95

TCG CAG CGC AAG AAC CTG TTC ACT GAG ATG CAT GAG CAG CTG GGC TCT       334
Ser Gln Arg Lys Asn Leu Phe Thr Glu Met His Glu Gln Leu Gly Ser
                100                 105                 110

GCA GAG GCT GCG GGA GCC GAG GGG TCA TGT CCA GCG CCG TCG GGC TCC       382
Ala Glu Ala Ala Gly Ala Glu Gly Ser Cys Pro Ala Pro Ser Gly Ser
            115                 120                 125

CCA GAC ACC GGG TCT TGG CTG CCC TCG CCC GGC CGC CGG CGG CGA CGC       430
Pro Asp Thr Gly Ser Trp Leu Pro Ser Pro Gly Arg Arg Arg Arg Arg
        130                 135                 140

ACC GCC TTC GCC AGC CGT CAC GGC AAG CGA CAT GGC AAG AAG TCC AGG       478
Thr Ala Phe Ala Ser Arg His Gly Lys Arg His Gly Lys Lys Ser Arg
    145                 150                 155

CTG CGC TGC AGC AGA AAG CCT CTG CAC GTG AAT TTT AAG GAG TTA GGC       526
Leu Arg Cys Ser Arg Lys Pro Leu His Val Asn Phe Lys Glu Leu Gly
160                 165                 170                 175

TGG GAC GAC TGG ATT ATC GCG CCC CTA GAG TAC GAG GCC TAT CAC TGC       574
Trp Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys
                180                 185                 190
```

```
GAG GGC GTG TGC GAC TTT CCG CTG CGC TCG CAC CTT GAG CCC ACT AAC        622
Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn
            195                 200                 205

CAT GCC ATC ATT CAG ACG CTG ATG AAC TCC ATG GAC CCG GGC TCC ACC        670
His Ala Ile Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr
        210                 215                 220

CCG CCT AGC TGC TGC GTT CCC ACC AAA CTG ACT CCC ATT AGC ATC CTG        718
Pro Pro Ser Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu
    225                 230                 235

TAC ATC GAC GCG GGC AAT AAT GTN GTC TAC AAG CAG TAT GAG GAC ATG        766
Tyr Ile Asp Ala Gly Asn Asn X02 Val Tyr Lys Gln Tyr Glu Asp Met
240                 245                 250                 255

GTG GTG GAG TCC TGC GGC TGT AGG TAGCGGTGCT GTCCCGCCAC CTGGGCCAGG       820
Val Val Glu Ser Cys Gly Cys Arg
                260

GACCATGGAG GGAGGCCTGA CTGCCGAGAA AGGAGCAGGA GCTGGCCTTG GAAGAGGCCA      880

CAGGTGGGGG ACAGCCTGAA AGTAGGAGCA CAGTAAGAAG CAGCCCAGCC TTCCCAGAAC      940

CTTCCAATCC CCCAACCCAG AAGCAGCTAA GGGGTTTCAC AACTTTTGGC CTTGCCAGCC     1000

TGGAAAGACT AGACAAGAGG GATTCTTCTC TTTTTATTAT GGCTTG                    1046
```

(2) SEQ ID NO : 32 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 263 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 32 :

```
Arg Lys Gln Ala Cys Ile Pro Ala Gly Pro Thr Leu Arg Gly Ser Ser
  1               5                  10                  15

Gly Thr Gln Pro Arg Pro Ala Gly Lys Ser Phe Asp Val Trp Gln Gly
            20                  25                  30

Leu Arg Pro Gln Pro Trp Lys Gln Leu Cys Leu Glu Leu Arg Ala Ala
        35                  40                  45

Trp Gly Glu Leu Asp X01 Gly Asp Thr Gly Ala Arg Ala Arg Gly Pro
    50                  55                  60

Gln Gln Pro Pro Pro Leu Asp Leu Arg Ser Leu Gly Phe Gly Arg Arg
65                  70                  75                  80

Val Arg Pro Pro Gln Glu Arg Ala Leu Leu Val Val Phe Thr Arg Ser
                85                  90                  95

Gln Arg Lys Asn Leu Phe Thr Glu Met His Glu Gln Leu Gly Ser Ala
            100                 105                 110

Glu Ala Ala Gly Ala Glu Gly Ser Cys Pro Ala Pro Ser Gly Ser Pro
        115                 120                 125

Asp Thr Gly Ser Trp Leu Pro Ser Pro Gly Arg Arg Arg Arg Arg Thr
    130                 135                 140
```

```
Ala Phe Ala Ser Arg His Gly Lys Arg His Gly Lys Lys Ser Arg Leu
145                 150                 155                 160

Arg Cys Ser Arg Lys Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp
                165                 170                 175

Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu
            180                 185                 190

Gly Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His
        195                 200                 205

Ala Ile Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro
    210                 215                 220

Pro Ser Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr
225                 230                 235                 240

Ile Asp Ala Gly Asn Asn X02 Val Tyr Lys Gln Tyr Glu Asp Met Val
                245                 250                 255

Val Glu Ser Cys Gly Cys Arg
            260
```

(2) SEQ ID NO : 33 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 1345 bp

(B) 유형 : 핵산

(C) 가닥상태 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(iii) 가설 : 없음

(iv) 안티센스 : 없음

(vii) 직접 얻을 수 있는 원천 :

(B) 클론 : 인간 V1-1

(ix) 특징:

(A) 이름/키 : CDS

(B) 위치 : 138..1301

(ix) 특징:

(A) 이름/키 : 매트_펩티드

(B) 위치 : 990..1301

(xi) 서열의 기술 : SEQ ID NO : 33 :

```
AACTATAGCA CCTGCAGTCC CTGGTCTTGG GTGTAGGGGT GCGCTCCTGG TCCCGCGGCT      60

CAGGGATATG CAGTGACCAA TGGGTTGTTG GCCTGATGGG ACTTTTGGCT TGCTAAACCA     120

AAGCTCGGTT CGGATAG CCC GGG CGA AGA CGT CCG CTG CTC TGG GCC AGG          170
                 Pro Gly Arg Arg Arg Pro Leu Leu Trp Ala Arg
                 -284            -280                -275

CTG GCA GCG TTC AGG CTG GGG CAG AGA CGC GGA GTC GGG CGC TGG CTC        218
Leu Ala Ala Phe Arg Leu Gly Gln Arg Arg Gly Val Gly Arg Trp Leu
          -270                -265                -260

CAA CAG GCC TGG CTC CCA CAT CGA AGA CAG CTG GGC CAT TTG CTG TTA        266
Gln Gln Ala Trp Leu Pro His Arg Arg Gln Leu Gly His Leu Leu Leu
     -255                -250                -245

GGA GGC CCC GCG CTG ACA GTG TGC AGG ATT TGC TCT TAC ACA GCT CTT        314
Gly Gly Pro Ala Leu Thr Val Cys Arg Ile Cys Ser Tyr Thr Ala Leu
-240                -235                -230

TCT CTC TGT CCC TGC CGG TCC CCC GCA GAC GAA TCG GCA GCC GAA ACA        362
Ser Leu Cys Pro Cys Arg Ser Pro Ala Asp Glu Ser Ala Ala Glu Thr
-225                -220                -215                -210

GGC CAG AGC TTC CTG TTC GAC GTG TCC AGC CTT AAC GAC GCA GAC GAG        410
Gly Gln Ser Phe Leu Phe Asp Val Ser Ser Leu Asn Asp Ala Asp Glu
               -205                -200                -195

GTG GTG GGT GCC GAG CTG CGC GTG CTG CGC CGG GGA TCT CCA GAG TCG        458
Val Val Gly Ala Glu Leu Arg Val Leu Arg Arg Gly Ser Pro Glu Ser
          -190                -185                -180

GGC CCA GGC AGC TGG ACT TCT CCG CCG TTG CTG CTG CTG TCC ACG TGC        506
Gly Pro Gly Ser Trp Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys
     -175                -170                -165

CCG GGC GCC GCC CGA GCG CCA CGC CTG CTG TAC TCG CGG GCA GCT GAG        554
Pro Gly Ala Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu
-160                -155                -150

CCC CTA GTC GGT CAG CGC TGG GAG GCG TTC GAC GTG GCG GAC GCC ATG        602
Pro Leu Val Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met
-145                -140                -135                -130

AGG CGC CAC CGT CGT GAA CCG CGC CCC CCC CGC GCG TTC TGC CTC TTG        650
Arg Arg His Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu
               -125                -120                -115

CTG CGC GCA GTG GCA GGC CCG GTG CCG AGC CCG TTG GCA CTG CGG CGA        698
Leu Arg Ala Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg
          -110                -105                -100

CTG GGC TTC GGC TGG CCG GGC GGA GGG GGC TCT GCG GCA GAG GAG CGC        746
Leu Gly Phe Gly Trp Pro Gly Gly Gly Gly Ser Ala Ala Glu Glu Arg
     -95                 -90                 -85

GCG GTG CTA GTC GTC TCC TCC CGC ACG CAG AGG AAA GAG AGC TTA TTC        794
Ala Val Leu Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe
-80                 -75                 -70

CGG GAG ATC CGC GCC CAG GCC CGC GCG CTC GGG GCC GCT CTG GCC TCA        842
Arg Glu Ile Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser
-65                 -60                 -55                 -50

GAG CCG CTG CCC GAC CCA GGA ACC GGC ACC GCG TCG CCA AGG GCA GTC        890
Glu Pro Leu Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val
               -45                 -40                 -35

ATT GGC GGC CGC AGA CGG AGG AGG ACG GCG TTG GCC GGG ACG CGG ACA        938
Ile Gly Gly Arg Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr
          -30                 -25                 -20

GCG CAG GGC AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC AGG GGC        986
Ala Gln Gly Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly
     -15                 -10                 -5

CGG AGC CGC TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG GAG CTC       1034
Arg Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu
  1                 5                  10                  15
```

```
GGC TGG GAC GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG TAC CAC      1082
Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His
                20                  25                  30

TGC GAG GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG CCC ACC      1130
Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr
            35                  40                  45

AAC CAT GCC ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA CCA GAC GCG      1178
Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala
        50                  55                  60

GCG CCG GCC TCC TGC TGT GTG CCA GCG CGC CTC AGC CCC ATC AGC ATC      1226
Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile
    65                  70                  75

CTC TAC ATC GAC GCC GCC AAC AAC GTT GTC TAC AAG CAA TAC GAG GAC      1274
Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp
80                  85                  90                  95

ATG GTG GTG GAG GCC TGC GGC TGC AGG TAGCGCGCGG GCCGGGGAGG            1321
Met Val Val Glu Ala Cys Gly Cys Arg
                100

GGGCAGCCAC GCGGCCGAGG ATCC                                           1345
```

(2) SEQ ID NO : 34 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 388 아미노산

(B) 유형 : 아미노산

(D) 위상 : 선형

(ii) 분자 유형 : 단백질

(xi) 서열의 기술 : SEQ ID NO : 34 :

```
Pro Gly Arg Arg Arg Pro Leu Leu Trp Ala Arg Leu Ala Ala Phe Arg
-284            -280                -275                -270

Leu Gly Gln Arg Arg Gly Val Gly Arg Trp Leu Gln Gln Ala Trp Leu
            -265                -260                -255

Pro His Arg Arg Gln Leu Gly His Leu Leu Leu Gly Gly Pro Ala Leu
        -250                -245                -240

Thr Val Cys Arg Ile Cys Ser Tyr Thr Ala Leu Ser Leu Cys Pro Cys
    -235                -230                -225

Arg Ser Pro Ala Asp Glu Ser Ala Ala Glu Thr Gly Gln Ser Phe Leu
-220                -215                -210                -205

Phe Asp Val Ser Ser Leu Asn Asp Ala Asp Glu Val Val Gly Ala Glu
                -200                -195                -190

Leu Arg Val Leu Arg Arg Gly Ser Pro Glu Ser Gly Pro Gly Ser Trp
            -185                -180                -175

Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys Pro Gly Ala Ala Arg
        -170                -165                -160

Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu Pro Leu Val Gly Gln
    -155                -150                -145

Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met Arg Arg His Arg Arg
-140                -135                -130                -125
```

Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu Leu Arg Ala Val Ala
-120 -115 -110

Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg Leu Gly Phe Gly Trp
-105 -100 -95

Pro Gly Gly Gly Gly Ser Ala Ala Glu Glu Arg Ala Val Leu Val Val
-90 -85 -80

Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile Arg Ala
-75 -70 -65

Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser Glu Pro Leu Pro Asp
-60 -55 -50 -45

Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val Ile Gly Gly Arg Arg
-40 -35 -30

Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr Ala Gln Gly Ser Gly
-25 -20 -15

Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg Cys Ser
-10 -5 1

Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp Asp Trp
5 10 15 20

Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Leu Cys
25 30 35

Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile
40 45 50

Gln Thr Leu Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala Ser Cys
55 60 65

Cys Val Pro Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp Ala
70 75 80

Ala Asn Asn Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ala
85 90 95 100

Cys Gly Cys Arg

(2) SEQ ID NO : 35 에 관한 정보 :

(i) 서열 특징 :

(A) 길이 : 17 bp

(B) 유형 : 핵산

(C) 가닥 : 단일

(D) 위상 : 선형

(ii) 분자 유형 : DNA (게놈)

(iv) 가설 : 없음

(vi) 안티 - 센스 : 없음

(vi) 최초의 원천 :

(C) 개체 (INDIVIDUAL)단리물 : 프라이머 번호 8

(xi) 서열의 기술 : SEQ ID NO : 35 :

TGTATGCGAC TTCCCGC

17

발명의 효과

본 발명의 조성물은 건염 및 건 또는 인대 결함의 치료 및 관련 조직 치유에 유용하다.

(57) 청구의 범위

청구항 1.

SEQ ID NO : 2 또는 26 아미노산 서열을 코딩하는 DNA 서열에 정확한 해독 구조[틀] 로 연결되어 있는 TGF-β 슈퍼패밀리의 단백질들 중의 하나의 프로펩티드를 코딩하는 DNA 서열을 포함하는 키메릭 DNA 분자.

청구항 2.

삭제

청구항 3.

하기군에서 선택된 폴리펩티드의 아미노산 서열을 갖는 모노머 하나와 BMP-1 내지 BMP-13 또는 MP52 단백질로 구성된 TGF-β 슈퍼패밀리에서 선택된 다른 단백질의 아미노산 서열을 갖는 모노머 하나를 포함하는 헤테로다이머성(heterodimeric) 단백질 분자 :

(a) SEQ ID NO : 2 에 나타낸 아미노산 #-25 내지 아미노산 #104 ;

(b) SEQ ID NO : 2 에 나타낸 아미노산 #1 내지 아미노산 #104 ;

(c) SEQ ID NO : 2 에 나타낸 아미노산 #3 내지 아미노산 #104 ;

(d) SEQ ID NO : 26 에 나타낸 아미노산 #1 내지 아미노산 #120 ; 및

(e) SEQ ID NO : 26 에 나타낸 아미노산 #19 내지 아미노산 #120.

청구항 4.

제 3 항에 있어서, 모노머 하나는 BMP-12 의 아미노산 서열을 포함하고, 다른 모노머는 BMP-2, BMP-3, BMP-4, BMP-5, BMP-6, BMP-7, BMP-8, BMP-9, BMP-10 및 BMP-11 로 구성된 군에서 선택된 단백질의 아미노산 서열을 포함하는 헤테로다이머성 단백질 분자.

청구항 5.

각각의 모노머는 BMP-12, BMP-13 및 MP-52 로부터 선택된 폴리펩티드를 코딩하는 개별적으로 선택된 아미노산 서열을 포함하는 헤테로다이머성 단백질 분자.

도면

도면 1

인간 V1-1 대 인간 MP-52 의 비교

```
V1-1 Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu
  1  AGC CGC TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG GAG CTC
        ||| ||| ||  ||  ||| |  ||  ||  ||  || ||| ||| ||  |
MP52 GCT CGC TGC AGT CGG AAG GCA CTG CAT GTC AAC TTC AAG GAC ATG
  1  Ala Arg Cys Ser Arg Lys Ala Leu His Val Asn Phe Lys Asp Met

 16  Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala Tyr
 46  GGC TGG GAC GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG TAC
     ||| ||| ||| ||| ||| ||| ||| ||  ||  ||  ||  ||| ||| ||  | |
 46  GGC TGG GAC GAC TGG ATC ATC GCA CCC CTT GAG TAC GAG GCT TTC
 16  Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Glu Tyr Glu Ala Phe

 31  His Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu
 91  CAC TGC GAG GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG
     ||| ||| ||  ||  ||| ||  ||| ||  ||| ||  ||  ||| ||  ||| |||
 91  CAC TGC GAG GGG CTG TGC GAG TTC CCA TTG CGC TCC CAC CTG GAG
 31  His Cys Glu Gly Leu Cys Glu Phe Pro Leu Arg Ser His Leu Glu

 46  Pro Thr Asn His Ala Ile Ile Gln Thr Leu Leu Asn Ser Met Ala
121  CCC ACC AAC CAT GCC ATC ATT CAG ACG CTG CTC AAC TCC ATG GCA
     ||| ||  ||  ||| ||   ||  ||  ||| ||  |||  |  ||| ||| ||| |
121  CCC ACG AAT CAT GCA GTC ATC CAG ACC CTG ATG AAC TCC ATG GAC
 46  Pro Thr Asn His Ala Val Ile Gln Thr Leu Met Asn Ser Met Asp

 61  Pro Asp Ala Ala Pro Ala Ser Cys Cys Val Pro Ala Arg Leu Ser
181  CCA GAC GCG GCG CCG GCC TCC TGC TGT GTG CCA GCG CGC CTC AGC
     ||  ||   |   |  ||   ||  ||| ||| ||| ||   ||  ||  ||  ||
181  CCC GAG TCC ACA CCA CCC ACC TGC TGT GTG CCC ACG CGG CTG AGT
 61  Pro Glu Ser Thr Pro Pro Thr Cys Cys Val Pro Thr Arg Leu Ser

 76  Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn Val Val Tyr
226  CCC ATC AGC ATC CTC TAC ATC GAC GCC GCC AAC AAC GTT GTC TAC
     ||| ||| ||| ||| ||| | | ||  |||  |  ||| ||| ||| ||  ||  ||
226  CCC ATC AGC ATC CTC TTC ATT GAC TCT GCC AAC AAC GTG GTG TAT
 76  Pro Ile Ser Ile Leu Phe Ile Asp Ser Ala Asn Asn Val Val Tyr

 91  Lys Gln Tyr Glu Asp Met Val Val Glu Ala Cys Gly Cys Arg
271  AAG CAA TAC GAG GAC ATG GTG GTG GAG GCC TGC GGC TGC AGG
     ||| ||  ||  ||| ||| ||| ||  ||| |||  |  ||  ||| ||| |||
271  AAG CAG TAT GAG GAC ATG GTC GTG GAG TCG TGT GGC TGC AGG
 91  Lys Gln Tyr Glu Asp Met Val Val Glu Ser Cys Gly Cys Arg
```

뉴클레오티드 수준에서의 상동성 : 249/312 = 79.8%
아미노산 수준에서의 상동성 : 84/104 = 80.8%