



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 367 035**

51 Int. Cl.:

**C12N 15/12** (2006.01)      **C12N 15/70** (2006.01)  
**C12N 15/62** (2006.01)      **C12N 1/21** (2006.01)  
**C07K 14/51** (2006.01)      **A61K 38/17** (2006.01)  
**C12N 15/63** (2006.01)      **A61K 38/18** (2006.01)  
**A61P 19/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **06004492 .2**

96 Fecha de presentación : **06.12.1994**

97 Número de publicación de la solicitud: **1690874**

97 Fecha de publicación de la solicitud: **16.08.2006**

54

Título: **BMP-12, BMP-13 y composiciones de las mismas para inducir tendones.**

30

Prioridad: **07.12.1993 US 164103**  
**25.03.1994 US 217780**  
**02.11.1994 US 333576**

45

Fecha de publicación de la mención BOPI:  
**27.10.2011**

45

Fecha de la publicación del folleto de la patente:  
**27.10.2011**

73

Titular/es: **GENETICS INSTITUTE, L.L.C. et al.**  
**87 Cambridge Park Drive**  
**Cambridge, Massachusetts 02140, US**  
**President and Fellows of Harvard College**

72

Inventor/es: **Celeste, Anthony J.;**  
**Wozney, John M.;**  
**Rosen, Vicky A.;**  
**Wolfman, Neil M.;**  
**Thomson, Gerald H. y**  
**Melton, Douglas A.**

74

Agente: **Carpintero López, Mario**

ES 2 367 035 T3

**Aviso:** En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

**DESCRIPCIÓN**

BMP-12, BMP-13 y composiciones de las mismas para inducir tendones

**Campo de la invención**

5 La presente invención se refiere a una nueva familia de proteínas purificadas, y a composiciones que contienen dichas proteínas, siendo dichas composiciones útiles para la inducción de la formación de tejido similar a tendón/ligamento, para la curación de heridas y para la reparación de ligamentos y otros tejidos. Estas proteínas también pueden usarse en composiciones para aumentar la actividad de proteínas morfogenéticas óseas.

**Antecedentes de la invención**

10 La búsqueda de la molécula o moléculas responsables de la formación de hueso, cartílago, tendón y otros tejidos presentes en el hueso y otros extractos de tejidos ha llevado al descubrimiento de una nueva serie de moléculas denominadas las Proteínas Morfogenéticas Óseas (BMP). Previamente se han esclarecido las estructuras de varias proteínas, denominadas BMP-1 a BMP-11. Las actividades de inducción excepcionales de estas proteínas, junto con su presencia en el hueso, sugieren que son reguladores importantes de procesos de reparación ósea, y pueden estar implicadas en el mantenimiento normal del tejido óseo. Existe la necesidad de identificar proteínas adicionales  
15 que jueguen un papel en la formación de otros tejidos vitales. La presente invención se refiere a la identificación de una familia de proteínas que tienen actividad inductora de tejido similar a tendón/ligamento, y que son útiles en composiciones para la inducción de la formación y reparación de tejido similar a tendón/ligamento.

**Sumario de la invención**

20 En un primer aspecto, la presente invención se refiere a un polipéptido que tiene la capacidad de inducir la formación de un tejido similar a tendón o similar a ligamentos, en el que dicho polipéptido se codifica por una molécula de ADN que comprende una secuencia de nucleótidos seleccionada del grupo que consiste en

- (a) los nucleótidos nº 196, nº 199, nº 208, nº 217, nº 361, nº 388, nº 493, nº 496, nº 571 o nº 577 a nº 879 o nº 882 de la SEC ID Nº: 1;
- 25 (b) nucleótidos que codifican los aminoácidos nº -25, nº 1 o nº 3 a nº 103 o nº 104 de la SEC ID Nº: 2;
- (c) los nucleótidos nº 845, nº 893 o nº 899 a nº 1201 o nº 1204 de la SEC ID Nº: 3;
- (d) nucleótidos que codifican los aminoácidos nº 1 o nº 19 a nº 119 o nº 120 de la SEC ID Nº: 4;
- (e) los nucleótidos nº 410, nº 458, nº 602, nº 605 o nº 659 a nº 961 o nº 964 de la SEC ID Nº: 25;
- (f) nucleótidos que codifican los aminoácidos nº 1 o nº 19 a nº 119 o nº 120 de la SEC ID Nº: 26; o
- 30 (g) secuencias alélicas humanas naturales y secuencias de codones degenerados equivalentes de uno cualquiera de (a) a (f), o
- (h) secuencias que hibridan con cualquiera de (a) a (f) anteriores en condiciones de hibridación rigurosas.

Y donde dicho polipéptido es para uso en un procedimiento para inducir la formación de tejido similar a tendón/ligamento en un paciente que lo necesita.

35 En otro aspecto, la presente invención proporciona una composición farmacéutica que comprende un polipéptido que tiene la capacidad de inducir la formación de un tejido similar a tendón o similar a ligamento, en la que dicho polipéptido se codifica por una molécula de ADN que comprende una secuencia de nucleótidos seleccionada del grupo que consiste en

- 40 (a) los nucleótidos nº 196, nº 199, nº 208, nº 217, nº 361, nº 388, nº 493, nº 496, nº 571 o nº 577 a nº 879 o nº 882 de la SEC ID Nº: 1;
- (b) nucleótidos que codifican los aminoácidos nº -25, nº 1 o nº 3 a nº 103 o nº 104 de la SEC ID Nº: 2;
- (c) los nucleótidos nº 845, nº 893 o nº 899 a nº 1201 o nº 1204 de la SEC ID Nº: 3;
- (d) nucleótidos que codifican los aminoácidos nº 1 o nº 19 a nº 119 o nº 120 de la SEC ID Nº: 4;
- (e) los nucleótidos nº 410, nº 458, nº 602, nº 605 o nº 659 a nº 961 o nº 964 de la SEC ID Nº: 25;
- 45 (f) nucleótidos que codifican los aminoácidos nº 1 o nº 19 a nº 119 o nº 120 de la SEC ID Nº: 26; o
- (g) secuencias alélicas humanas naturales o secuencias de codones degenerados equivalentes de uno cualquiera de (a) a (f), o
- (h) secuencias que hibridan con cualquiera de (a) a (f) anteriores en condiciones de hibridación rigurosas

en la que la composición es para inducir la formación de tejido semejante a tendón/ligamento en un paciente que lo necesita.

50 De esta manera, la presente invención se refiere a moléculas de ADN que codifican una proteína inductora de tejido similar a tendón/ligamento que los inventores han denominado V1-1. Esta nueva proteína se denomina ahora BMP-12. La presente invención también incluye moléculas de ADN que codifican proteínas relacionadas con BMP-12.

Las proteínas relacionadas con BMP-12 son una subserie de la familia de proteínas BMP/TGF- $\beta$ /Vg-1, que incluyen

BMP-12 y VL-1, que se definen como proteínas inductoras de tejido similar a tendón/ligamento codificadas por secuencias de ADN que se clonan e identifican, por ejemplo, usando PCR, usando cebadores específicos de BMP-12, tales como cebadores nº 6 y nº 7 descritos más adelante, con condiciones de rigurosidad reducidas. Se prefiere que las secuencias de ADN que codifican proteínas relacionadas con BMP-12 compartan al menos aproximadamente un 80% de homología a nivel de los aminoácidos con los aminoácidos nº 3 a nº 103 de la SEC ID Nº: 1.

Las moléculas de ADN preferentemente tienen una secuencia de ADN que codifica la proteína BMP-12, proporcionándose dicha secuencia en la SEC ID Nº: 1, o una proteína relacionada con BMP-12 como se describe adicionalmente en el presente documento. Tanto la proteína BMP-12 como las proteínas relacionadas con BMP-12 se caracterizan por la capacidad de inducir la formación de un tejido similar a tendón/ligamento en el ensayo descrito en los ejemplos.

Las moléculas de ADN usadas por la invención preferentemente comprenden una secuencia de ADN, como se describe en la SECUENCIA ID Nº: 1, más preferentemente los nucleótidos nº 496 a nº 882, nº 571 a nº 882 o nº 577 a nº 882 de la SEC ID Nº: 1; o secuencias de ADN que hibridan con las anteriores en condiciones de hibridación rigurosas y codifican una proteína que presenta la capacidad de formar tejido similar a tendón/ligamento. El uso de moléculas de ADN por la invención también puede comprender una secuencia de ADN como se describe en la SEC ID Nº: 25; más preferentemente los nucleótidos nº 604 o nº 658 a nº 964 de la SEC ID Nº: 25.

Las moléculas de ADN usadas por la invención también incluyen moléculas de ADN que comprenden una secuencia de ADN que codifica una proteína relacionada con BMP-12 con la secuencia de aminoácidos mostrada en la SEC ID Nº: 2 o SEC ID Nº: 26, así como secuencias alélicas naturales y secuencias de codones degenerativos equivalentes de la SEC ID Nº: 2 o SEC ID Nº: 26. Preferentemente, la secuencia de ADN usada por la presente invención codifica los aminoácidos nº -25 a nº 104, nº 1 a nº 104 o nº 3 a nº 103 de la SEC ID Nº: 2; o los aminoácidos nº 1 a nº 120 o nº 19 a nº 120 de la SEC ID Nº: 26. La secuencia de ADN puede comprender, en una dirección 5' a 3', nucleótidos que codifican un propéptido y nucleótidos que codifican los aminoácidos nº -25 a nº 104, nº 1 a nº 104 o nº 3 a nº 103 de la SEC ID Nº: 2; o los aminoácidos nº 1 a nº 120 o nº 19 a nº 120 de la SEC ID Nº: 26. El propéptido útil en la realización anterior preferentemente se selecciona del grupo que consiste en el propéptido de BMP-12 nativo y un propéptido de proteína de un miembro diferente de la superfamilia de TGF-B o la familia de BMP. La invención además se refiere a secuencias de ADN que hibridan con las secuencias de ADN anteriores en condiciones de hibridación rigurosas y codifican una proteína relacionada con BMP-12 que presenta la capacidad de inducir la formación de tejido similar a tendón/ligamento.

La presente invención se refiere además a una proteína relacionada con BMP-12 purificada caracterizada por la capacidad de inducir la formación de tejido similar a tendón/ligamento. Los polipéptidos relacionados con BMP-12 preferentemente comprenden una secuencia de aminoácidos como la mostrada en la SEC ID Nº: 2. Más preferentemente, el polipéptido comprende los aminoácidos nº -25, nº 1 o nº 3 a nº 103 o nº 104 como se expone en la SEC ID Nº: 2; o los aminoácidos nº 1 o nº 19 a nº 120 como se expone en la SEC ID Nº: 26. En una realización preferida, el polipéptido purificado puede estar en forma de un dímero compuesto por dos subunidades, cada una con la secuencia de aminoácidos de la SEC ID Nº: 2.

En otra realización, la presente invención comprende composiciones que comprenden una cantidad eficaz de las proteínas relacionadas con BMP-12 descritas anteriormente. En las composiciones, la proteína puede mezclarse con un vehículo farmacéuticamente aceptable.

Otras realizaciones incluyen moléculas de ADN quiméricas que comprenden una secuencia de ADN que codifica un propéptido de un miembro de la superfamilia de proteínas TGP- $\beta$ , unida en el marco de lectura correcto a una secuencia de ADN que codifica un polipéptido relacionado con BMP-12. Un propéptido adecuado es el propéptido de BMP-2. La invención también incluye moléculas de proteína heterodiméricas que comprenden un monómero que tiene la secuencia de aminoácidos mostrada en la SEC ID Nº: 2, y un monómero que tiene la secuencia de aminoácidos de otra proteína de la subfamilia de TGF- $\beta$ .

### **Descripción de las secuencias**

La SEC ID Nº: 1 es la secuencia de nucleótidos que codifica la BMP-12 humana.

La SEC ID Nº: 2 es la secuencia de aminoácidos que comprende el polipéptido de BMP-12 humano maduro.

La SEC ID Nº: 3 es la secuencia de nucleótidos que codifica la proteína MP52.

La SEC ID Nº: 4 es la secuencia de aminoácidos que comprende el polipéptido de MP52 maduro.

La SEC ID Nº: 5 es la secuencia de nucleótidos de una parte amplificada específicamente de la secuencia que codifica BMP-12 humana.

- La SEC ID Nº: 6 es la secuencia de aminoácidos codificada por la secuencia de nucleótidos de la SEC ID Nº: 5.
- La SEC ID Nº: 7 es la secuencia de nucleótidos de una parte amplificada específicamente de la secuencia que codifica VL-1 humana.
- La SEC ID Nº: 8 es la secuencia de aminoácidos codificada por la secuencia de nucleótidos de la SEC ID Nº: 7.
- 5 La SEC ID Nº: 9 es la secuencia de nucleótidos del plásmido pALV1-781, usado para la expresión del BMP-12 en *E. coli*.
- La SEC ID Nº: 10 es la secuencia de nucleótidos de un fragmento del clon murino, mV1.
- La SEC ID Nº: 11 es la secuencia de aminoácidos de un fragmento de la proteína murina codificada por mV1.
- La SEC ID Nº: 12 es la secuencia de nucleótidos de un fragmento del clon murino, mV2.
- 10 La SEC ID Nº: 13 es la secuencia de aminoácidos de un fragmento de la proteína murina codificada por mV2.
- La SEC ID Nº: 14 es la secuencia de nucleótidos de un fragmento del clon murino, mV9.
- La SEC ID Nº: 15 es la secuencia de aminoácidos de un fragmento de la proteína murina codificada por mV9.
- La SEC ID Nº: 16 es la secuencia de aminoácidos de una secuencia consenso de la proteína BMP/TGF- $\beta$ /Vg-1. El primer Xaa representa Gln o Asn; el segundo Xaa representa Val o Ile.
- 15 La SEC ID Nº: 17 es la secuencia de nucleótidos del oligonucleótido nº 1.
- La SEC ID Nº: 18 es la secuencia de aminoácidos de una secuencia consenso de proteína BMP/TGF- $\beta$ /Vg-1. Xaa representa Val o Leu.
- La SEC ID Nº: 19 es la secuencia de nucleótidos del oligonucleótido nº 2.
- La SEC ID Nº: 20 es la secuencia de nucleótidos del oligonucleótido nº 3.
- 20 La SEC ID Nº: 21 es la secuencia de nucleótidos del oligonucleótido nº 4.
- La SEC ID Nº: 22 es la secuencia de nucleótidos del oligonucleótido nº 5.
- La SEC ID Nº: 23 es la secuencia de nucleótidos del oligonucleótido nº 6.
- La SEC ID Nº: 24 es la secuencia de nucleótidos del oligonucleótido nº 7.
- La SEC ID Nº: 25 es la secuencia de nucleótidos de la secuencia codificante de VL-1 humana (BMP-13).
- 25 La SEC ID Nº: 26 es la secuencia de aminoácidos codificada por la secuencia de nucleótidos de la SEC ID Nº: 25.
- La SEC ID Nº: 27 es la secuencia de nucleótidos que codifica una fusión del propéptido de BMP-2 y la secuencia codificante madura de BMP-12.
- La SEC ID Nº: 28 es la secuencia de aminoácidos codificada por la secuencia de nucleótidos de la SEC ID Nº: 27.
- La SEC ID Nº: 29 es la secuencia de nucleótidos que codifica la proteína mV1 murina. X01 es Val, Ala, Glu o Gly; X02 es Ser, Pro Thr o Ala; X03 es Ser o Arg; X04 es Leu, Pro, Gln o Arg; X05 es Cys o Trp; X06 es Val, Ala, Asp o Gly; X07 es Val, Ala, Glu o Gly; X08 es Gln, Lys o Glu.
- 30 La SEC ID Nº: 30 es la secuencia de aminoácidos codificada por la secuencia de nucleótidos de la SEC ID Nº: 29. X01 a X08 son los mismos que en la SEC ID Nº: 29.
- La SEC ID Nº: 31 es la secuencia de nucleótidos que codifica la proteína mV2 murina. X01 es Pro o Thr; X02 es Val.
- 35 La SEC ID Nº: 32 es la secuencia de aminoácidos codificada por la secuencia de nucleótidos de la SEC ID Nº: 31. X01 y X02 son los mismos que en la SEC ID Nº: 31.
- La SEC ID Nº: 33 es la secuencia de nucleótidos que codifica la proteína BMP-12 humana.
- La SEC ID Nº: 34 es la secuencia de aminoácidos codificada por la secuencia de nucleótidos de la SEC ID Nº: 33.
- La SEC ID Nº: 35 es la secuencia de nucleótidos del oligonucleótido nº 8.

**Breve descripción de las figuras**

La Figura 1 es una comparación de las secuencias de BMP-12 humana y MP52 un humana.

**Descripción detallada de la invención**

5 Las secuencias de ADN de la presente invención son útiles para producir proteínas que inducen la formación de tejido similar a tendón/ligamento, como se describe adicionalmente más adelante. Las secuencias de ADN de la presente invención son útiles además para aislar y clonar secuencias de ADN adicionales que codifican proteínas relacionadas con BMP-12 con actividad similar. Estas proteínas relacionadas con BMP-12 pueden ser homólogas de otras especies, o pueden ser proteínas relacionadas dentro de la misma especie.

10 Además, otro aspecto de la invención son secuencias de ADN que codifican la expresión de una proteína inductora de tejido similar a tendón/ligamento. Estas secuencias incluyen la secuencia de nucleótidos en la dirección 5' a 3' ilustrada en la SEC ID N°: 1 o en la SEC ID N°: 25, secuencias de ADN que, excepto por la degeneración del código genético, son idénticas a las secuencia de ADN de la SEC ID N°: 1 ó 25, y codifican la proteína de la SEC ID N°: 2 ó 26. También se incluyen en la presente invención secuencias de ADN que hibridan en condiciones rigurosas con la secuencia de ADN de la SEC ID N°: 1 ó 25 y codifican una proteína que tiene la capacidad de inducir la formación de un tendón o ligamento. Las secuencias de ADN preferidas incluyen las que hibridan en condiciones rigurosas como se describe en Maniatis y col, Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory (1982), páginas 387 a 389. Finalmente, también se incluyen en la presente invención variaciones alélicas u otras variaciones de las secuencias de la SEC ID N°: 1 ó 25, tanto si dichos cambios de nucleótidos dan como resultado cambios en la secuencia peptídica como si no, pero en las que la secuencia peptídica sigue teniendo actividad inductora de tejido similar a tendón/ligamento.

15 La secuencia de ADN del BMP-12 humana (SEC ID N°: 1) y la secuencia de aminoácidos (SEC ID N° 2) se exponen en el Listado de Secuencias. Otra proteína que es útil para las composiciones y procedimientos de la presente invención es VL-1. VL-1 es una proteína relacionada con BMP-12 que se clonó usando secuencias de BMP-12. Los inventores han denominado ahora a VL-1 con el nombre BMP-13. Una secuencia de ADN parcial de VL-1 (SEC ID N°: 7) y la secuencia de aminoácidos codificada (SEC ID N°: 8); así como una secuencia de ADN que codifica la VL-1 madura (SEC ID N°: 25) y la secuencia de aminoácidos codificada (SEC ID N°: 26) se exponen en el Listado de Secuencias. Aunque se realizan descripciones adicionales haciendo referencia a la secuencia de BMP-12 de las SEC ID N°: 1 y 2, se reconocerá que la invención incluye modificaciones y mejoras similares que pueden realizarse en otras secuencias relacionadas con BMP-12, tales como la secuencia de VL-1 mostrada en la SEC ID N°: 25 y 26.

20 La secuencia de BMP-12 mostrada en la SEC ID N°: 1 incluye la secuencia madura entera y aproximadamente 190 aminoácidos del propéptido. La secuencia codificante de la proteína BMP-12 humana madura parece empezar en el nucleótido n° 496 o n° 571 y continúa hasta el nucleótido n° 882 de la SEC ID N°: 1. La primera cisteína en la estructura de siete cisteínas característica de las proteínas TGF- $\beta$  empieza en el nucleótido n° 577. La última cisteína termina en el n° 879. De esta manera, es de esperar que las secuencias de ADN que codifican especies de BMP-12 activas comprendan los nucleótidos n° 577 a n° 879 de la SEC ID N°: 1.

25 Es de esperar que BMP-12, expresada por células de mamífero tales como células CHO, exista como una población heterogénea de especies activas de proteína BMP-12 con extremos N variables. Es de esperar que todas las especies activas contengan la secuencia de aminoácidos que empieza con el resto de cisteína en el aminoácido n° 3 de la SEC ID N°: 2 y continúa hasta al menos el resto de cisteína en el aminoácido 103 o hasta el codón de terminación después del aminoácido 104. Otras especies activas contienen otra secuencia de aminoácidos en la dirección N-terminal. Como se describe adicionalmente en el presente documento, es de esperar que el extremo N de las especies activas producidas por las células de mamífero comience después de la aparición de un sitio de escisión consenso, que codifica una secuencia peptídica Arg-X-X-Arg. De esta manera, es de esperar que las secuencias de ADN que codifican proteínas BMP-12 activas tengan una secuencia de nucleótidos que comprende la secuencia de nucleótidos que empieza en cualquiera de los nucleótidos n° 106, 199, 208, 217, 361, 388, 493, 496 ó 571 hasta el nucleótido n° 879 u 882 de la SEC ID N°: 1.

30 Se ha determinado experimentalmente por expresión en *E. coli* que el extremo N de una especie activa de BMP-12 humana es el siguiente: [M]SRXSRKPLHVDF, en el que X designa un resto de aminoácido sin señal evidente, que es coherente con un resto de cisteína en esa localización. De esta manera, parece ser que el extremo N de esta especie de BMP-12 está en el aminoácido n° 1 de la SEC ID N°: 1, y una secuencia de ADN que codifica dicha especie de BMP-12 empezaría en el nucleótido n° 571 de la SEC ID N°: 1. El peso molecular aparente de esta especie de dímero de BMP-12 humana, según se determinó por SDS-PAGE, es de aproximadamente 20-22 kd en un gel de tricina al 16% Novex. La proteína BMP-12 humana existe como una solución transparente incolora en ácido trifluoroacético al 0,1%.

35 Como se ha descrito anteriormente, las proteínas relacionadas con BMP-12 son una subserie de la familia de proteínas BMP/TGF- $\beta$ /Vg-1, incluyendo BMP-12 y VL-1, que pueden definirse como proteínas inductoras de tejido similar a tendón/ligamento codificadas por secuencias de ADN que pueden clonarse e identificarse, por ejemplo,

usando PCR, usando cebadores específicos de BMP-12, tales como los cebadores nº 6 y nº 7 descritos más adelante, con condiciones de rigurosidad reducida. Se prefiere que las secuencias de ADN de la presente invención compartan al menos aproximadamente un 80% de homología a nivel de aminoácidos con el ADN que codifica los aminoácidos nº 3 a nº 103 de la SEC ID Nº: 1. Para los fines de la presente invención, la expresión proteínas relacionadas con BMP-12 no incluye la proteína MP52 humana. Usando la información de secuencia de la SEC ID Nº: 1 y la SEC ID Nº: 3, y la comparación proporcionada en la Figura 1, está dentro de la experiencia en la técnica diseñar cebadores para la secuencia de BMP-12 que permitan la clonación de genes que codifican proteínas relacionadas con BMP-12.

Un ejemplo de las proteínas relacionadas con BMP-12 de la presente invención es VL-1, denominada actualmente BMP-13. La secuencia de la proteína BMP-13 madura completa y al menos una parte del propéptido de BMP-13 se proporcionan en la SEC ID Nº: 25. Al igual que BMP-12, es de esperar que BMP-13, expresada por células de mamífero tales como células CHO, exista como una población heterogénea de especies activas de proteína BMP-13 con extremos N variables. Es de esperar que todas las especies activas contengan la secuencia de aminoácidos que empieza con el resto de cisteína en el aminoácido nº 19 de la SEC ID Nº: 26 y continúa hasta al menos el resto de cisteína en el aminoácido 119 o hasta el codón de terminación después del aminoácido 120. Otras especies activas contienen una secuencia de aminoácidos adicional en la dirección N terminal. Como se describe adicionalmente en el presente documento, es de esperar que el extremo N de la especie activa producida por células de mamífero comience después de la aparición de un sitio de escisión consenso, que codifica una secuencia peptídica Arg-X-X-Arg. De esta manera, es de esperar que las secuencias de ADN que codifican proteínas BMP-13 activas tengan una secuencia de nucleótidos que comprende la secuencia de nucleótidos que empieza en cualquiera de los nucleótidos nº 410, 458, 602, 605 ó 659, hasta el nucleótido nº 961 ó 964 de la SEC ID Nº: 25.

Para producir las proteínas inductoras de tejido similar a tendón/ligamento purificadas útiles para la presente invención, se emplea un procedimiento que comprende cultivar una célula huésped transformada con una secuencia de ADN que comprende una secuencia codificante adecuada, particularmente la secuencia codificante de ADN desde el nucleótido nº 496, nº 571 o nº 577 a nº 879 o nº 882 de la SEC ID Nº: 1; y recuperar y purificar a partir del medio de cultivo una proteína que contenga la secuencia de aminoácidos o una secuencia sustancialmente homóloga representada por los aminoácidos nº -25, nº 1 o nº 3 a nº 103 o nº 104 de la SEC ID Nº: 2. En otra realización, el procedimiento empleado comprende cultivar una célula huésped transformada con una secuencia de ADN que comprende una secuencia codificante adecuada, particularmente la secuencia codificante de ADN desde el nucleótido nº 605 o nº 659 a nº 961 o nº 964 de la SEC ID Nº: 25; y recuperar y purificar del medio de cultivo una proteína que contiene la secuencia de aminoácidos o una secuencia sustancialmente homóloga a la representada por los aminoácidos nº 1 o nº 19 a nº 119 o nº 120 de la SEC ID Nº: 26

El ADN de MP52 humana se describe en el documento WO93/16099. Sin embargo, este documento no desvela la capacidad de la proteína de formar tejido similar a tendón/ligamento, o su uso en composiciones para la inducción de tejido similar a tendón/ligamento. La MP52 humana se aisló originalmente usando ARN de tejido embrionario humano. La secuencia de nucleótidos de MP52 humana (SEC ID Nº: 3) y las secuencias de aminoácidos codificadas (SEC ID Nº: 4) se exponen en el Listado de Secuencias del presente documento. La proteína MP52 parece empezar en el nucleótido nº 845 de la SEC ID Nº: 3 y continúa hasta el nucleótido nº 1204 de la SEC Nº: 3. La primera cisteína de la estructura de siete cisteínas característica de las proteínas TGF- $\beta$  empieza en el nucleótido nº 899. La última cisteína termina en el nº 1201. Otras especies activas de proteína MP52 pueden tener nucleótidos adicionales en la dirección N-terminal desde el nucleótido nº 845 de la SEC ID Nº: 3.

Pueden producirse proteínas MP52 humanas purificadas de la presente invención cultivando una célula huésped transformada con una secuencia de ADN que comprende la secuencia codificante de ADN de la SEC ID Nº: 3 desde el nucleótido nº 845 a nº 1204, y recuperando y purificando a partir del medio de cultivo una proteína que contiene la secuencia de aminoácidos o una secuencia sustancialmente homóloga a la representada por los aminoácidos nº 1 a nº 120 de la SEC ID Nº: 4. También es de esperar que la secuencia de aminoácidos desde los aminoácidos nº 17 o nº 19 a nº 119 o nº 120 de la SEC ID Nº: 4 retenga actividad. De esta manera, es de esperar que la secuencia de ADN de los nucleótidos nº 845, nº 893 o nº 899 a nº 1201 o nº 1204 codifique proteínas activas.

Para la expresión de la proteína en células huésped de mamífero, la célula huésped se transforma con una secuencia codificante que codifica un propéptido adecuado para la secreción de proteínas por la célula huésped, unida en un marco de lectura apropiado a la secuencia codificante para la proteína madura. Por ejemplo, véase la Patente de Estados Unidos 5.168.050, en la que un ADN que codifica una parte precursora de una proteína de mamífero distinta de BMP-2 se fusiona al ADN que codifica una proteína BMP-2 madura. De esta manera, la presente invención incluye moléculas de ADN químéricas que comprenden una secuencia de ADN que codifica un propéptido de un miembro de la superfamilia de proteínas TGF- $\beta$ , que está unida en el marco de lectura correcto a una secuencia de ADN que codifica un polipéptido de inducción de tejido similar a tendón/ligamento. El término "quimérico" se usa para hacer referencia a que el propéptido procede de un polipéptido diferente que el polipéptido maduro codificado. Por supuesto, la célula huésped puede transformarse con una secuencia de ADN codificante que codifica el propéptido nativo unido en el marco de lectura correcto a una secuencia codificante que codifica la proteína madura mostrada en la SEC ID Nº: 2, SEC ID Nº: 4 o SEC ID Nº: 26. La secuencia completa del propéptido

nativo puede determinarse por procedimientos conocidos en la técnica usando las secuencias desveladas en las SEC ID N°: 1, SEC ID N°: 3 o SEC ID N°: 25 para diseñar una sonda adecuada para identificar y aislar el clon entero.

La presente invención también incluye las nuevas secuencias de ADN, sin asociación con secuencias de ADN que codifican otros materiales proteicos, y que codifican la expresión de proteínas inductoras de tejido similar a tendón/ligamento. Estas secuencias de ADN incluyen las representadas en las SEC ID N°: 1 en una dirección 5' a 3' y las secuencias que hibridan con las mismas en condiciones de hibridación rigurosas [por ejemplo, SSC 0,1X, SDS al 19% a 65 °C; véase, T. Maniatis y col, Molecular Cloning (A Laboratory Manual), Cold Spring Harbor Laboratory (1982), páginas 387 a 389] y codifican una proteína que tiene actividad inductora de tejido similar a tendón/ligamento.

De forma similar, las secuencias de ADN que codifican proteínas codificadas por las secuencias de la SEC ID N°: 1 o SEC ID N°: 25, o proteínas que comprenden la secuencia de aminoácidos de la SEC ID N°: 2 o SEC ID N°: 26, pero que difieren en la secuencia de codones debido a la degeneración del código genético o variaciones alélicas (cambios de bases naturales en la población de especies que pueden dar como resultado o no un cambio de aminoácidos) también codifican las proteínas inductoras de tejido similar a tendón/ligamento descritas en el presente documento. También se incluyen en la invención variaciones en las secuencias de ADN de las SEC ID N°: 1 o SEC ID N°: 25 que se producen por mutaciones puntuales o por modificaciones inducidas (incluyendo inserción, delección y sustitución) para aumentar la actividad, semivida o producción de los polipéptidos codificados.

Otro aspecto de la presente invención proporciona un nuevo procedimiento para producir proteínas inductoras de tejido similar a tendón/ligamento. El procedimiento de la presente invención implica cultivar una línea celular adecuada, que se ha transformado con una secuencia de ADN que codifica una proteína de la invención, bajo el control de secuencias reguladoras conocidas. Las células huésped transformadas se cultivan y las proteínas se recuperan y purifican a partir del medio de cultivo. Las proteínas purificadas carecen sustancialmente de otras proteínas con las que se co-producen así como de otros contaminantes.

Las células o líneas celulares adecuadas pueden ser células de mamífero, tales como células de ovario de hámster chino (CHO). Como se ha descrito anteriormente, la expresión de proteínas en células de mamífero requiere un propéptido apropiado para asegurar la secreción de la proteína. En la técnica se conoce la selección de células huésped de mamífero adecuadas y procedimientos para la transformación, cultivo, amplificación, exploración, producción de producto y purificación. Véase, por ejemplo, Gething y Sambrook, Nature, 293: 620-625 (1981) o, como alternativa, Kaufman y col, Mol. Cell. Biol., 5(7): 1750-1759 (1985) o Howley y col, Patente de Estados Unidos 4.419.446. Otra línea celular de mamífero adecuada, que se describe en los ejemplos adjuntos, es la línea celular COS-1 de mono. También puede ser adecuada la célula CV-1 de mamífero.

También pueden ser huéspedes adecuados células bacterianas. Por ejemplo, las diversas cepas de *E. coli* (por ejemplo, HB101, MC1061) son bien conocidas como células huésped en el campo de la biotecnología. También pueden emplearse en este procedimiento diversas cepas de *B. subtilis*, *Pseudomonas*, otros bacilos y similares. Para la expresión de la proteína en células bacterianas, no es necesario ADN que codifique un propéptido.

Se sabe que la expresión en bacterias de proteínas de mamífero, incluyendo miembros de la familia de TGF- $\beta$ , produce las proteínas en una forma no glicosilada, y en forma de gránulos insolubles, conocidos como cuerpos de inclusión. Se han descrito técnicas en este campo para solubilizar estos cuerpos de inclusión, desnaturalizar la proteína usando un agente caotrópico y desplegar la proteína de una forma suficientemente correcta para permitir su producción en una forma soluble. Por ejemplo, véase el documento EP 0433225.

Como alternativa, se han ideado procedimientos que evitan la formación de cuerpos de inclusión, tales como la expresión de proteínas de fusión de genes, en los que la proteína deseada se expresa como una proteína de fusión con un compañero de fusión. La proteína de fusión posteriormente se somete a escisión para producir la proteína deseada. Un ejemplo de dicho sistema de expresión de fusión génica para *E. coli* se basa en el uso del gen de la tiorredoxina de *E. coli* como compañero de fusión, La Vallie y col., Bio/Technology, 11: 187-193 (1993).

Muchas cepas de células de levadura conocidas por los expertos en la materia también pueden estar disponibles como células huésped para la expresión de los polipéptidos de la presente invención. Además, cuando se desee, pueden utilizarse células de insecto como células huésped en el procedimiento de la presente invención. Véase, por ejemplo, Miller y col, Genetic Engineering, 8: 277-298 (Plenum Press 1986) y las referencias citadas en dicho documento.

Otro aspecto de la presente invención proporciona vectores para uso en el procedimiento de expresión de estas proteínas inductoras de tejido similar a tendón/ligamento. Preferentemente, los vectores contienen las secuencias de ADN nuevas completas descritas anteriormente que codifican los nuevos factores de la invención. Además, los vectores contienen secuencias de control de la expresión apropiadas que permiten la expresión de las secuencias de proteínas. Como alternativa, también son realizaciones de la presente invención vectores que incorporan secuencias modificadas como se han descrito anteriormente. Además, la secuencia de la SEC ID N°: 1 o la SEC ID N°: 3 o la SEC ID N°: 25 podría manipularse para expresar una proteína madura mediante la delección de secuencias

de propéptido y el reemplazo de las mismas con secuencias que codifican los propéptidos completos de proteínas BMP o miembros de la superfamilia de TGF- $\beta$ . De esta manera, la presente invención incluye moléculas de ADN quiméricas que codifican un propéptido de un miembro de la superfamilia de TGF- $\beta$  unido en el marco de lectura correcto a una secuencia de ADN que codifica una proteína que tiene la secuencia de aminoácidos de la SEC ID N<sup>o</sup>: 2 o la SEC ID N<sup>o</sup>: 4 o la SEC ID N<sup>o</sup>: 26. Los vectores pueden emplearse en el procedimiento de transformación de líneas celulares y contener secuencias reguladoras seleccionadas en la asociación operativa con las secuencias codificantes de ADN de la invención que son capaces de dirigir la replicación y expresión de las mismas en células huésped seleccionadas. Los expertos en la materia conocen secuencias reguladoras para dichos vectores y pueden seleccionarse dependiendo de las células huésped. Dicha selección es rutinaria y no forma parte de la presente invención.

Una proteína de la presente invención que induce tejido similar a tendón/ligamento u otra formación de tejido en circunstancias en las que normalmente no se forma dicho tejido, tiene aplicación en la curación de desgarros de tendones o ligamentos, deformidades y otros defectos de tendones o ligamentos en seres humanos y otros animales. Dicha preparación que emplea una proteína inductora de tejido similar a tendón/ligamento puede tener uso profiláctico en la prevención de lesiones en tejido de tendón o ligamento, así como uso en la mejora de la fijación del tendón o ligamento al hueso u otros tejidos, y en la reparación de defectos de tejido de tendón o ligamento. La formación de tejido similar a tendón/ligamento *de novo* inducida por una composición de la presente invención contribuye a la reparación de defectos congénitos, inducidos por traumatismo u otros defectos en tendón o ligamento de otro origen, y también es útil en la cirugía plástica cosmética para la unión o reparación de tendones o ligamentos. Las composiciones de la invención también pueden ser útiles en el tratamiento de tendinitis, síndrome del túnel del carpo y otros defectos de tendones o ligamentos. Las composiciones de la presente invención también pueden usarse en otras indicaciones en las que es deseable curar o regenerar tejido del tendón y/o ligamento. Estas indicaciones incluyen, sin limitación, la regeneración o reparación de lesiones en el ligamento periodontal, tales como las que se producen en las tendinitis, y la regeneración o reparación de la unión del tendón al hueso. Las composiciones de la presente invención pueden proporcionar un medio para unir células formadoras de tendón o ligamento, estimular el crecimiento de células formadoras de tendón o ligamento o inducir la diferenciación de progenitores de células formadoras de tendón o ligamento.

Las proteínas relacionadas con BMP-12 pueden recuperarse del medio de cultivo y purificarse por aislamiento de otros materiales proteicos con los que se co-producen y de otros contaminantes presentes. Las proteínas de la presente invención pueden inducir la formación de tejido similar a tendón/ligamento. Estas proteínas pueden caracterizarse adicionalmente por la capacidad de demostrar actividad de formación de tejido similar a tendón/ligamento en el ensayo de implante ectópico de rata descrito más adelante. Se contempla que estas proteínas también pueden tener capacidad de inducir la formación de otros tipos de tejido, tales como ligamentos.

Las proteínas inductoras de tejido similar a tendón/ligamento proporcionadas en el presente documento también incluyen factores codificados por secuencias similares a las de las SEC ID N<sup>o</sup>: 1 o la SEC ID N<sup>o</sup>: 25, pero en las que se proporcionan naturalmente modificaciones (por ejemplo, variaciones alélicas en la secuencia de nucleótidos que pueden dar como resultado cambios de aminoácidos en el polipéptido) o modificadas deliberadamente por ingeniería genética. Por ejemplo, los polipéptidos sintéticos pueden duplicar total o parcialmente secuencias continuas de los restos de aminoácidos de la SEC ID N<sup>o</sup>: 2. Estas secuencias, gracias a que comparten características conformacionales y de estructura primaria, secundaria o terciaria con polipéptidos de factor de crecimiento de tejido similar a tendón/ligamento de la SEC ID N<sup>o</sup>: 2, pueden poseer propiedades biológicas de factor de crecimiento de tejido similar a tendón/ligamento u otro tejido en común con los mismos. De esta manera, pueden emplearse como sustitutos biológicamente activos de polipéptidos inductores de tejido similar a tendón/ligamento naturales en las composiciones y procedimientos terapéuticos.

Otras mutaciones específicas de las secuencias de proteínas inductoras de tejido similar a tendón/ligamento descritas en el presente documento implican modificaciones de sitios de glicosilación. Estas modificaciones pueden implicar sitios de glicosilación unidos a O o unidos a N. Por ejemplo, la ausencia de glicosilación o sólo una glicosilación parcial resulta de la sustitución o delección de aminoácidos en sitios de reconocimiento de glicosilación unidos a asparagina. Los sitios de reconocimiento de glicosilación unidos a asparagina comprenden secuencias tripeptídicas que se reconocen específicamente por enzimas de glicosilación celular apropiadas. Estas secuencias tripeptídicas pueden ser asparagina-X-treonina, asparagina-X-serina o asparagina-X-cisteína, donde X es normalmente cualquier aminoácido excepto prolina. Una diversidad de sustituciones o delecciones de aminoácidos en uno o los dos de la primera o tercera posiciones de aminoácidos de un sitio de reconocimiento de glicosilación (y/o delección de aminoácidos en la segunda posición) da como resultado una ausencia de glicosilación en la secuencia tripeptídica modificada. Además, la expresión bacteriana de proteínas también dará como resultado la producción de una proteína no glicosilada, aunque se dejen sin modificar los sitios de glicosilación.

Las composiciones de la presente invención comprenden una proteína relacionada con BMP-12 purificada que puede producirse por cultivo de una célula transformada con la secuencia de ADN de la SEC ID N<sup>o</sup>: 1 o la SEC ID N<sup>o</sup>: 25 y recuperación y purificación de la proteína que tiene la secuencia de aminoácidos de la SEC ID N<sup>o</sup>: 2 o la SEC ID N<sup>o</sup>: 26 a partir del medio de cultivo. La proteína expresada purificada carece sustancialmente de otros

materiales proteicos con los que se produce conjuntamente, así como de otros contaminantes. Se contempla que la proteína purificada recuperada presenta actividad de formación de tejido similar a tendón/ligamento y otra actividad de crecimiento de tejidos, tal como regeneración de ligamentos. Las proteínas de la invención pueden caracterizarse adicionalmente por la capacidad de demostrar actividad de formación de tejido similar a tendón/ligamento en el ensayo de rata descrito más adelante.

Las composiciones para inducir formación de tejido similar a tendón/ligamento de la presente invención pueden comprender una cantidad eficaz de una proteína inductora de tejido similar a tendón/ligamento, en las que dicha proteína comprende la secuencia de aminoácidos de la SEC ID N<sup>o</sup>:2, preferentemente los aminoácidos n<sup>o</sup> -25, n<sup>o</sup> 1 o n<sup>o</sup> 3 a n<sup>o</sup> 103 o n<sup>o</sup> 104 de la SEC ID N<sup>o</sup>: 2; o los aminoácidos n<sup>o</sup> 1 o n<sup>o</sup> 19 a n<sup>o</sup> 120 de la SEC ID N<sup>o</sup>: 26; así como mutantes y/o variantes de la SEC ID N<sup>o</sup>: 2 o la SEC ID N<sup>o</sup>: 26, que presenta la capacidad de formar tejido similar a tendón y/o ligamento.

Las composiciones de la presente invención pueden comprender además proteínas adicionales, tales como miembros adicionales de la superfamilia de proteínas TGF- $\beta$ , tales como activinas. Otro aspecto de la invención proporciona composiciones farmacéuticas que contienen una cantidad terapéuticamente eficaz de una proteína inductora de tendón/ligamento, tal como BMP-12 o VL-1, en un vehículo o excipiente farmacéuticamente aceptable. Estas composiciones pueden usarse para inducir la formación de tejido similar a tendón/ligamento u otro tejido. Se contempla que dichas composiciones también pueden usarse para la reparación de tendón y ligamento, curación de heridas y reparación de otros tejidos, tales como la reparación de la piel. También se contempla que las proteínas de la invención pueden aumentar la supervivencia neuronal y, por lo tanto, ser útiles en el trasplante y tratamiento de afecciones que presentan una reducción de la supervivencia neuronal. Las composiciones de la invención pueden incluir además al menos otro agente terapéuticamente útil, tal como las proteínas BMP BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6 y BMP-7, desveladas, por ejemplo, en las Patentes de Estados Unidos 5.108.922; 5.013.649; 5.116.738; 5.106.748; 5.187.076; y 5.141.905; BMP-8, desvelada en la publicación PCT WO91/18098; BMP-9, desvelada en la publicación PCT WO93/00432; y BMP-10 o BMP-11, desveladas en las solicitudes de patente en trámite junto con la presente con el número de serie 08/061.695 y 08/061.464, presentada el 12 de mayo de 1993.

Las composiciones de la invención pueden comprender, además de proteína inductora de tendón/ligamento tal como BMP-12 o VL-1 (BMP-13), otros agentes terapéuticamente útiles que incluyen MP52, factor de crecimiento epidérmico (EGF), factor de crecimiento de fibroblastos (FGF), factor de crecimiento derivado de plaquetas (PDGF), factores de crecimiento de transformación (TGF- $\alpha$  y TGF- $\beta$ ) y factor de crecimiento de fibroblastos-4 (FGF-4), hormona paratiroidea (PTH), factor inhibidor de leucemia (LIF/HILDA/DIA), factores de crecimiento similares a insulina (IGF-1 e IGF-II). En las composiciones de la presente invención también pueden usarse partes de estos agentes. Por ejemplo, una composición que comprende tanto BMP-2 como BMP-12 implantados conjuntamente da lugar a tejido similar tanto a hueso como a tendón/ligamento. Dicha composición puede ser útil para tratar defectos de la articulación embrionaria donde se forman simultáneamente tendón, ligamentos y hueso en localizaciones anatómicas contiguas, y puede ser útil para regenerar tejido en el sitio de la unión del tendón al hueso. Se contempla que las composiciones de la invención también pueden usarse en la curación de heridas, tal como en la curación de la piel y la reparación de tejidos relacionada. Los tipos de heridas incluyen, pero sin limitación, quemaduras, incisiones y úlceras. (Véase, por ejemplo, en la Publicación PCT WO84/0110, el análisis de la curación de heridas y la reparación de tejido relacionada).

Es de esperar que las proteínas de la invención puedan actuar conjuntamente con, o quizás de forma sinérgica, con otras proteínas y factores de crecimiento relacionados. Por lo tanto, otros procedimientos terapéuticos y composiciones de la invención comprenden una cantidad terapéutica de al menos una proteína de la invención con una cantidad terapéutica de al menos una de las proteínas BMP descritas anteriormente. Dichas composiciones pueden comprender moléculas separadas de las proteínas BMP o heteromoléculas compuestas por diferentes restos de BMP. Por ejemplo, un procedimiento y composición de la invención pueden comprender un dímero unido por enlaces disulfuro que comprende una subunidad de proteína relacionada con BMP-12 y una subunidad de una de las proteínas "BMP" descritas anteriormente. De esta manera, la presente invención incluye composiciones que comprenden un polipéptido relacionado con BMP-12 purificado es un heterodímero en el que una subunidad comprende la secuencia de aminoácidos desde el aminoácido n<sup>o</sup> 1 al aminoácido n<sup>o</sup> 104 de la SEC ID N<sup>o</sup>: 2, y otra subunidad comprende una secuencia de aminoácidos para una proteína morfogenética ósea seleccionada del grupo que consiste en BMP-1, BMP-2, BMP-3, BMP-4, BMP-5, BMP-6, BMP-7, BMP-8, 2. BMP-9, BMP-10 y BMP-11. Otra realización puede comprender un heterodímero de restos inductores de tejido similar a tendón/ligamento unidos por enlaces disulfuro tales como BMP-12, VL-1 (BMP-13) o MP52. Por ejemplo, el heterodímero puede comprender una subunidad que comprende una secuencia de aminoácidos desde el n<sup>o</sup> 1 al n<sup>o</sup> 104 de la SEC ID N<sup>o</sup>: 2 y la otra subunidad puede comprender una secuencia de aminoácidos desde el n<sup>o</sup> 1 al n<sup>o</sup> 120 de la SEC ID N<sup>o</sup>: 4 o desde el n<sup>o</sup> 1 al n<sup>o</sup> 120 de la SEC ID N<sup>o</sup>: 26. Además, pueden combinarse composiciones de la presente invención con otros agentes beneficiosos para el tratamiento del defecto, herida o tejido en cuestión.

La preparación y formulación de dichas composiciones de proteínas fisiológicamente aceptables, considerando debidamente el pH, la isotonicidad, la estabilidad y similares, está dentro de la experiencia de la técnica. Las composiciones terapéuticas también son valiosas actualmente para aplicaciones veterinarias debido a la ausencia

de especificidad de especie en proteínas TGF- $\beta$ . Particularmente, los animales domésticos y los caballos pura sangre además de los seres humanos son pacientes deseados para dicho tratamiento con las composiciones de la presente invención.

5 El procedimiento terapéutico incluye administrar la composición tópicamente, sistémicamente o localmente como un implante o dispositivo. Cuando se administra, la composición terapéutica para uso en la presente invención, por supuesto, está en una forma fisiológicamente aceptable sin pirógenos. Además, la composición puede encapsularse deseablemente o inyectarse en una forma viscosa para suministrarse en el sitio de lesión del tejido. La administración tópica puede ser adecuada para la curación de heridas y la reparación de tejidos. Ciertos agentes terapéuticamente útiles distintos de las proteínas que también pueden incluirse opcionalmente en la composición  
10 como se ha descrito anteriormente, pueden administrarse, como alternativa o adicionalmente, de forma simultánea o secuencial con la composición en los procedimientos de la invención.

15 Las composiciones también pueden incluir una matriz y/o agente secuestrante apropiado como excipiente. Por ejemplo, la matriz puede soportar la composición o proporcionar una superficie para la formación de tejido similar a tendón/ligamento y/u otra formación de tejido. La matriz puede proporcionar la liberación lenta de la proteína y/o el entorno apropiado para su presentación. El agente secuestrante puede ser una sustancia que ayuda a facilitar la administración por inyección u otro medio, o puede ralentizar la migración de la proteína desde el sitio de aplicación.

20 La elección de un material excipiente se basa en las propiedades de biocompatibilidad, biodegradabilidad, propiedades mecánicas, aspecto cosmético y propiedades de interfaz. La aplicación particular de las composiciones definirá la formulación apropiada. Las matrices potenciales para las composiciones pueden ser biodegradables y químicamente definidas. Otras matrices se componen de proteínas puras o componentes de matriz extracelular. Otras matrices potenciales son no biodegradables y definidas químicamente. Las matrices preferidas incluyen materiales basados en colágeno tales como esponja Helistat (Integra LifeSciences, Plainsboro, N.J.), o colágeno en una forma inyectable, así como agentes secuestrantes, que también pueden ser biodegradables y que pueden incluir materiales de alquilcelulosa.

25 Otra clase preferida de vehículos son matrices poliméricas porosas en forma de partículas, que incluyen polímeros de poli(ácido láctico), poli(ácido glicólico) y copolímeros de ácido láctico y ácido glicólico. Estas matrices también pueden incluir un agente secuestrante. Se describen matrices poliméricas adecuadas, por ejemplo, en el documento WO 93/00050.

30 Una familia preferida de agentes secuestrantes es la de materiales celulósicos tales como alquilcelulosas (incluyendo hidroxialquilcelulosas), incluyendo metilcelulosa, etilcelulosa, hidroxietilcelulosa, hidroxipropilcelulosa, hidroxipropilmetilcelulosa y carboximetilcelulosa, siendo las más preferidas sales catiónicas de carboximetilcelulosa (CMC). Otros agentes secuestrantes preferidos incluyen ácido hialurónico, alginato sódico, poli(etilenglicol), poli(óxido de oxietileno), polímero de carboxivinilo y poli(alcohol vinílico). La cantidad de agente secuestrante útil en el presente documento es del 0,5-20% en peso, preferentemente del 1-10% en peso con respecto al peso total de la  
35 formulación, que representa la cantidad necesaria para prevenir la desorción de la proteína de la matriz polimérica y permitir una manipulación apropiada de la composición, pero no tan grande como para impedir que las células progenitoras se infiltren en la matriz, proporcionando de esta manera a la proteína la oportunidad de ayudar a la actividad de las células progenitoras.

40 Otros componentes opcionales útiles en la práctica de la presente solicitud incluyen, por ejemplo, protectores criogénicos tales como manitol, sacarosa, lactosa, glucosa o glicina (para proteger a la proteína de la degradación durante la liofilización), conservantes antimicrobianos tales como metil y propil parabenos y alcohol bencílico; antioxidantes tales como EDTA, citrato y BHT (hidroxitolueno butilado); y tensioactivos tales como poli(sorbitos) y poli(oxietilenos), etc.

45 Como se ha descrito anteriormente, las composiciones de la invención pueden emplearse en procedimientos para tratar varios defectos de tendones, tales como para la regeneración de tejido similar a tendón/ligamento en áreas de lesiones de tendones o ligamentos, para ayudar a reparar desgarros de tejido de tendón, ligamentos y otros diversos tipos de defectos o heridas en tejidos. Estos procedimientos, de acuerdo con la invención, implican la administración a un paciente que necesita dicha reparación de tejido similar a tendón/ligamento u otro tejido, de una composición que comprende una cantidad eficaz de una proteína inductora de tejido similar a tendón/ligamento, tal como la descrita en la SEC ID N°: 2, SEC ID N°: 4 y/o SEC ID N°: 26. Estos procedimientos también pueden implicar la administración de una proteína inductora de tejido similar a tendón/ligamento junto con al menos una de las proteínas BMP descritas anteriormente.  
50

55 En otra realización, los procedimientos pueden implicar la administración de una proteína heterodimérica en la que uno de los monómeros es un polipéptido de inducción de tejido similar a tendón/ligamento, tal como BMP-12, VL-1 (BMP-13) o MP52, y el segundo monómero es un miembro de la superfamilia de factores de crecimiento TGF- $\beta$ . Además, estos procedimientos también pueden incluir la administración de una proteína inductora de tejido similar a tendón/ligamento con otros factores de crecimiento que incluyen EGF, FGF, TGF- $\alpha$ , TGF- $\beta$  e IGF.

De esta manera, otro aspecto de la invención es un procedimiento terapéutico y composición para reparar tejido similar a tendón/ligamento, para reparar tendones o ligamentos, así como para tratar tendinitis y otras afecciones relacionadas con defectos de tendones o ligamentos. Dichas composiciones comprenden una cantidad terapéuticamente eficaz de una o más proteínas inductoras de tejido similar a tendón/ligamento, tales como BMP-12, una proteína relacionada con BMP-12, o MP52, mezclada con un vehículo, excipiente o matriz farmacéuticamente aceptable.

El régimen de dosificación se determinará por el médico a cargo del caso considerando diversos factores que modifican la acción de la composición, por ejemplo, la cantidad de tejido de tendón o ligamento que se desea formar, el sitio de la lesión de tendón o ligamento, el estado del tendón o ligamento dañado, el tamaño de la herida, el tipo de tejido dañado, la edad, sexo y dieta del paciente, la gravedad de cualquier infección, el momento de administración y otros factores clínicos. La dosificación puede variar con el tipo de matriz usada en la reconstitución y los tipos de proteínas adicionales en la composición. También puede afectar a la dosificación la adición de otros factores de crecimiento conocidos, tales como IGF-I (factor de crecimiento similar a insulina I), a la composición final.

El progreso puede monitorizarse mediante la evaluación periódica de la formación de tejido similar a tendón/ligamento, o del crecimiento y/o reparación de tendón o ligamento. El progreso puede monitorizarse por procedimientos conocidos en la técnica, por ejemplo, rayos X, artroscopia, determinaciones histomorfométricas y marcaje con tetraciclina.

Los siguientes ejemplos ilustran la práctica de la presente invención en la recuperación y caracterización de proteínas inductoras de tejido similar a tendón/ligamento humano y su empleo para recuperar las otras proteínas inductoras de tejido similar a tendón/ligamento, obtener las proteínas humanas, expresar las proteínas mediante técnicas recombinantes y demostrar la capacidad de las composiciones de la presente invención para formar tejido similar a tendón/ligamento en un modelo *in vivo*. Aunque los ejemplos demuestran la invención con respecto a BMP-12, con modificaciones minoritarias dentro de la experiencia en la técnica, se cree que se pueden obtener los mismos resultados con MP52 y VL-1

## **Ejemplo 1**

### **Aislamiento de ADN**

Pueden aislarse secuencias de ADN que codifican BMP-12 y proteínas relacionadas con BMP-12 por diversas técnicas conocidas por los expertos en la materia. Como se describe más adelante, pueden diseñarse cebadores oligonucleotídicos basándose en secuencias de aminoácidos presentes en otras proteínas BMP, proteínas relacionadas con Vg-1 y otras proteínas de la superfamilia de TGF- $\beta$ . Pueden identificarse regiones que contienen secuencias de aminoácidos que están muy conservadas dentro de la familia de proteínas BMP y dentro de otros miembros de la superfamilia de TGF- $\beta$  y pueden construirse secuencias de aminoácidos consenso de estas regiones altamente conservadas basándose en la similitud de las regiones correspondientes de proteínas BMP/TGF- $\beta$ /Vg-1 individuales. A continuación se indica un ejemplo de dicha secuencia de aminoácidos consenso.

Secuencia de aminoácidos consenso (1):

Trp-Gln/Asn-Asp-Trp-Ile-Val/Ile-Ala (SEC ID N°: 16)

En la que X/Y indica que cualquier resto de aminoácido puede aparecer en esa posición.

El siguiente oligonucleótido se diseña basándose en la secuencia de aminoácidos consenso identificada anterior (1):

n° 1: CGGATCCTGGVANGAYTGGATHRTNGC (SEC ID N°: 17)

Esta secuencia oligonucleotídica se sintetiza en un sintetizador de ADN automático. Los símbolos de nucleótidos convencionales en el cebador oligonucleotídico identificado anterior son los siguientes: A, adenosina; C, citosina; G, guanina; T, timina; N, adenosina o citosina o guanina o timina; R, adenosina o citosina; Y, citosina o timina; H, adenosina o citosina o timina; V, adenosina o citosina o guanina; D, adenosina o guanina o timina.

Los siete primeros nucleótidos del oligonucleótido n° 1 (subrayado) contienen la secuencia de reconocimiento para la endonucleasa de restricción BamHI para facilitar la manipulación de una secuencia de ADN amplificada específicamente que codifica la proteína BMP-12 y, por lo tanto, no proceden de la secuencia de aminoácidos consenso (1) presentada anteriormente.

Una segunda secuencia de aminoácidos consenso procede de otra región altamente conservada de las proteínas BMP/TGF- $\beta$ /Vg-1 como se describe a continuación:

His-Ala-Ile-Val/Leu-Gln-Thr (SEC ID N°: 18)

El siguiente oligonucleótido se diseña basándose en la secuencia de aminoácidos consenso identificada anterior (2):

nº 2: TTTCTAGAARNGTYTGNACDATNGCRTG (SEC ID Nº: 19)

Esta secuencia oligonucleotídica se sintetiza en un sintetizador de ADN automático. Se usan los mismos símbolos de nucleótidos que se han descrito anteriormente.

5 Los siete primeros nucleótidos del oligonucleótido nº 1 (subrayado) contienen la secuencia de reconocimiento para la endonucleasa de restricción XbaI para facilitar la manipulación de una secuencia de ADN amplificada específicamente que codifica la proteína BMP-12 y, por lo tanto, no proceden de la secuencia de aminoácidos consenso (2) presentada anteriormente.

10 Se contempla que la proteína BMP-12 de la invención y otras proteínas relacionadas con BMP/TGF-β/Vg-1 pueden contener secuencias de aminoácidos similares a las secuencias de aminoácidos consenso descritas anteriormente y que la localización de esta secuencia dentro de una proteína BMP-12 u otras proteínas relacionadas nuevas correspondería a las localizaciones relativas en las proteínas a partir de las cuales se obtuvieron. Además se contempla que esta información posicional derivada de la estructura de otras proteínas BMP/TGF-β/Vg-1 y las secuencias oligonucleotídicas nº 1 y nº 2 que se han obtenido a partir de las secuencias de aminoácidos consenso (1) y (2) respectivamente, podrían utilizarse para amplificar específicamente secuencias de ADN que codifican los aminoácidos correspondientes de una proteína BMP-12 u otras proteínas relacionadas con BMP/TGF-β/Vg-1.

15 Basándose en el conocimiento de las estructuras génicas de proteínas BMP/TGF-β/Vg-1, se contempla además que puede usarse ADN genómico humano como molde para realizar reacciones de amplificación específicas que darían como resultado la identificación de secuencias codificantes de BMP-12 HMB/TGF-β/Vg-1 (proteína relacionada con BMP-12). Dichas reacciones de amplificación específicas de un molde de ADN genómico humano podrían iniciarse con el uso de cebadores oligonucleotídicos nº 1 y nº 2 descritos anteriormente. Los oligonucleótidos nº 1 y nº 2 identificados anteriormente se utilizan como cebadores para permitir la amplificación específica de una secuencia de nucleótidos específica de ADN genómico humano. La reacción de amplificación se realiza como se indica continuación:

25 Se corta ADN genómico humano (fuente: linfocitos de sangre periférica), proporcionado por Ken Jacobs de Genetics Institute, mediante el pase repetido a través de una aguja de calibre 25, se desnaturaliza a 100 °C durante 5 minutos y después se enfría en hielo antes de añadirlo a una mezcla de reacción que contiene una concentración 200 μM de cada desoxinucleótido trifosfato (dATP, dGTP, dCTP y dTTP), Tris-HCl 10 mM, pH 8,3, KCl 50 mM, MgCl<sub>2</sub> 1,5 mM, gelatina al 0,001%, 1,25 unidades de ADN polimerasa Taq, una concentración 100 pM de oligonucleótido nº 1 y una concentración 100 pM de oligonucleótido nº 2. Esta mezcla de reacción se incuba a 94 °C durante dos minutos y después se somete a ciclos térmicos de la siguiente manera: 1 minuto a 94 °C, 1 minuto a 40 °C, 1 minuto a 72 °C durante tres ciclos; después 1 minuto a 94 °C, 1 minuto a 55 °C, 1 minuto a 72 °C durante treinta y siete ciclos, seguido de una incubación de 10 minutos a 72 °C.

35 El ADN que se amplifica específicamente por esta reacción se precipita con etanol, se digiere con las endonucleasas de restricción BamHI y XbaI y se somete a electroforesis en gel de agarosa. Se escinde una región del gel, que corresponde al tamaño previsto de la BMP-12 u otro fragmento de ADN que codifica BMP/TGF-β/Vg-1, y los fragmentos de ADN amplificados específicamente contenidos se electroeluyen y se subclonan en el vector plasmídico pGEM-3 entre los sitios XbaI y BamHI del poliengarce. El análisis de la secuencia de ADN de uno de los subclones relacionados con BMP-12 resultantes indica que la secuencia de ADN amplificada específicamente del producto contenido codifica una parte de la proteína BMP-12 de la invención.

40 La secuencia de ADN (SEC ID Nº: 5) y la secuencia de aminoácidos derivada (SEC ID Nº: 6) de este fragmento de ADN amplificado específicamente de BMP-12 se muestran en el Listado de SECUENCIAS.

45 Los nucleótidos nº 1-nº 26 de la SEC ID Nº: 5 comprenden una parte del oligonucleótido nº 1 y los nucleótidos nº 103 – nº 128 comprenden una parte del complemento inverso del oligonucleótido nº 2 utilizado para realizar la reacción de amplificación específica. Debido a la función de los oligonucleótidos nº 1 y nº 2 para iniciar la reacción de amplificación, es posible que no correspondan exactamente a la secuencia real que codifica una proteína BMP-12 y, por lo tanto, no se traducen en la derivación de aminoácidos correspondiente (SEC ID Nº: 6).

El análisis de la secuencia de ADN de otro subclón indica que el producto de ADN amplificado específicamente contenido en la misma codifica una parte de otra proteína BMP/TGF-β/Vg1 (relacionada con BMP-12) de la invención denominada VL-1.

50 La secuencia de ADN (SEC ID Nº: 7) y la secuencia de aminoácidos derivada (SEC ID Nº: 8) de este fragmento de ADN amplificado específicamente se muestran en el Listado de Secuencias.

55 Los nucleótidos nº 1 – nº 26 de la SEC ID Nº: 7 comprenden una parte del oligonucleótido nº 1 y los nucleótidos nº 103 – nº 128 comprenden una parte del complemento inverso del oligonucleótido nº 2 utilizado para realizar la reacción de amplificación específica. Debido a la función de los oligonucleótidos nº 1 y nº 2 en la iniciación de la reacción de amplificación, es posible que no correspondan exactamente a la secuencia real que codifica una

proteína VL-1 de la invención y, por lo tanto, no se traducen en la derivación de aminoácidos correspondiente (SEC ID N°: 8).

La siguiente sonda oligonucleotídica se diseña basándose en la secuencia de ADN humana de BMP-12 amplificada específicamente expuesta anteriormente (SEC ID N°: 5) y sintetizada en un sintetizador de ADN automático:

5 n° 3: CCACTGCGAGGGCCTTTGCGACTTCCCTTTGCGTTCGCAC (SEC ID N°: 20)

Esta sonda oligonucleotídica se marca radiactivamente con <sup>32</sup>P y se emplea para explorar una biblioteca genómica humana construida en el vector λFIX (Stratagene, n° de catálogo 944201). Se cultivan 500.000 recombinantes de la biblioteca genómica humana a una densidad de aproximadamente 10.000 recombinantes por placa en 50 placas. Réplicas en nitrocelulosa duplicadas de las placas de bacteriófagos recombinantes e hibridadas con la sonda oligonucleotídica n° 3 en tampón de hibridación convencional (SHB = SSC 5X, SDS al 0,1%, Denhardt 5X, 100 µg/ml de ADN de esperma de salmón) se dejan a 65 °C durante una noche. Al día siguiente, se retira la solución de hibridación que contiene oligonucleótido marcado radiactivamente y los filtros se lavan con SSC 0,2X, SDS al 0,1% a 65 °C. Se identifica un solo recombinante de hibridación positiva y se purifica en placas. Este clon de bacteriófago recombinante purificado en placa que hibrida con la sonda oligonucleotídica de BMP-12 n° 3 se denomina λHuG-48. Se prepara una solución madre en placa de bacteriófago y se aísla ADN del bacteriófago a partir del con genómico humano λHuG-48. El bacteriófago λHuG-48 se ha depositado en la Colección Americana de Cultivos Tipo, 12301 Parklawn Drive, Rockville, MD "ATCC" con el n° de acceso 75625 el 7 de diciembre de 1993. Este depósito cumple los requisitos del Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos a los Fines del Procedimiento en Materia de Patentes y Regulaciones expuestas en el mismo. La región de hibridación oligonucleotídica de este recombinante, λHuG-48, se localiza en un fragmento BamHI de 3,2 kb. Este fragmento se subclona en un vector plasmídico (pGEM-3) y se realiza el análisis de la secuencia de ADN. Este subclón plasmídico se denomina PCR1-1#2 y se ha depositado en la Colección Americana de Cultivos Tipo, 12301 Parklawn Drive, Rockville, MD "ATCC" con el n° de acceso 69517 el 7 de diciembre de 1993. Este depósito cumple los requisitos del Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos a los Fines del Procedimiento en Materia de Patentes y las Regulaciones expuestas en el mismo. La secuencia de ADN parcial (SEC ID N°: 1) y la secuencia de aminoácidos derivada (SEC ID N°: 2) del inserto de ADN de 3,2 kb del subclón plasmídico PCR1-1#2, derivada del clon λHuG-48, se muestran en la Lista de Secuencias.

Debe indicarse que los nucleótidos n° 639 – n° 714 de la SEC ID N°: 1 corresponden a los nucleótidos n° 27 – n° 102 del fragmento de ADN que codifica BMP-12 amplificado específicamente expuesto en la SEC ID N°: 5, confirmando de esta manera que el clon del bacteriófago genómico humano λHuG-48 y el subclón derivado PCR1-1#2 codifican al menos una parte de la proteína BMP-12 de la invención. La secuencia de nucleótidos de una parte del inserto BamHI de 3,2 kb del plásmido PCR1-1#2 contiene un marco de lectura abierto de al menos 882 pares de bases, como se define por los nucleótidos n° 1-n° 882 de la SEC ID N°: 1. Este marco de lectura de abierto codifica al menos 294 aminoácidos de la proteína BMP-12 humana de la invención. La proteína BMP-12 humana de 294 aminoácidos codificada incluye la proteína BMP-12 humana madura completa (aminoácidos n° 1-n° 104 de la SEC ID N°: 2) así como la parte C-terminal de la región del propéptido del producto de traducción primario (aminoácido n° -190 a n° -1 de la SEC ID N°: 2).

Otra secuencia de ADN del inserto BamHI de 3,2 kb del plásmido PCR1-1#2 expuesto en la SEC ID N°: 33 demuestra la presencia de un marco de lectura abierto de 1164 pb, como se define por los nucleótidos n° 138 a n° 1301 de la SEC ID N° 33. [OBSÉRVESE que toda la secuencia desvelada en la SEC ID N°: 1 está contenida dentro la SEC ID N°: 33]. Como esta secuencia procede de un clon genómico, es difícil determinar el límite entre el extremo 5' de la secuencia codificante y el límite 3' de la secuencia intermedia (secuencia de intrón/no codificante).

Basándose en el conocimiento de otras proteínas BMP y otras proteínas dentro de la familia de TGF-β, se prevé que el polipéptido precursor se escindiría en la secuencia mutibásica Arg-Arg-Gly-Arg de acuerdo con una secuencia de procesamiento proteolítico consenso propuesta de Arg-X-X-Arg. Es de esperar que la escisión del polipéptido precursor de BMP-12 genere un péptido maduro de 104 aminoácidos que empieza con el aminoácido Ser en la posición n° 1 de la SEC ID N°: 2. Se espera que el procesamiento de BMP-12 en la forma madura implique la dimerización y eliminación de la región N terminal de una manera análoga al procesamiento de la proteína relacionada TGF-β [Gentry y col., Molec & Cell. Biol., 8: 4162 (1988); Derynck y col. Nature, 316: 701 (1985)].

Por lo tanto, se contempla que la especie activa madura de BMP-12 comprende un homodímero de dos subunidades polipeptídicas, comprendiendo cada subunidad los aminoácidos n° 1 a n° 104 de la SEC ID N°: 2 con un peso molecular previsto de aproximadamente 12.000 daltons. Se contemplan otras especies activas que comprenden al menos los aminoácidos n° 3 a n° 103 de la SEC ID N°: 2, incluyendo de esta manera el primer y el último resto de cisteína conservado. Como ocurre con otros miembros de la familia de proteínas TGF-β/BMP, la parte carboxi-terminal de la proteína BMP-12 presenta una mayor conservación de secuencia que la parte más próxima al extremo amino. El porcentaje de identidad de aminoácidos de la proteína BMP-12 humana en el dominio C-terminal rico en cisteína (aminoácidos n° 3 – n° 104) con la región correspondiente de las proteínas BMP humanas y otras proteínas dentro de la familia de TGF-β es la siguiente: 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%; activina WC (BMP-11), 38%; Vgl, 46%; GDF-1, 47%; TGF- $\beta$ 1, 36%; TGF- $\beta$ 2, 36%; TGF- $\beta$ 3, 39%; inhibina  $\beta$ (B), 36%; inhibina  $\beta$ (A), 41%.

5 La secuencia de ADN de BMP-12 humana (SEC ID N<sup>o</sup>: 1), o una parte de la misma, puede usarse como sonda para identificar una línea celular humana o tejido que sintetiza ARNm de BMP-12. En resumen, se extrae ARN de una fuente celular o tisular seleccionada y se somete a electroforesis en un gel de agarosa-formaldehído y se transfiere a nitrocelulosa, o se hace reaccionar con formaldehído y se aplica puntualmente sobre nitrocelulosa directamente. Después, la nitrocelulosa se hibrida con una sonda derivada de la secuencia codificante de BMP-12 humana.

10 Como alternativa, la secuencia de BMP-12 humana se usa para diseñar cebadores oligonucleotídicos que amplificarán específicamente una parte de la secuencia codificante de BMP-12 localizada en la región entre los cebadores utilizados para realizar la reacción de amplificación específica. Se contempla que estos cebadores derivados de BMP-12 humana permitirían amplificar específicamente secuencias codificantes de BMP-12 correspondientes a partir de moldes de ARNm, ADNc o ADN genómico. Una vez que se ha identificado una fuente positiva por uno de los procedimientos descritos anteriormente, se selecciona ARNm por cromatografía de oligo (dT) celulosa y se sintetiza ADNc y se clona en  $\lambda$ gt10 u otros vectores del bacteriófago  $\lambda$  conocidos por los expertos en la materia, por ejemplo,  $\lambda$ ZAP por técnicas establecidas (Toole y col., mencionado anteriormente). También es posible realizar la reacción de amplificación dirigida a cebador oligonucleotídico, descrita anteriormente, directamente en una biblioteca de ADNc o genómica humana preestablecida que se ha clonado en un vector del bacteriófago  $\lambda$ . En estos casos, podría explorarse directamente una biblioteca que produce un producto de ADN amplificado específicamente que codifica una parte de la proteína BMP-12 humana, utilizando el fragmento de ADN que codifica BMP-12 amplificada como sonda.

20 Se predice que los cebadores oligonucleotídicos diseñados basándose en la secuencia de ADN del clon genómico de BMP-12 humana  $\lambda$ HuG-48 permiten la amplificación específica de secuencias de ADN que codifican BMP-12 humana a partir de bibliotecas de ADNc humano preestablecidas que están disponibles en el mercado (es decir, Stratagene, La Jolla, CA o Clontech Laboratories, Inc., Palo Alto, CA). El siguiente cebador oligonucleotídico se diseña basándose en los nucleótidos n<sup>o</sup> 571 a n<sup>o</sup> 590 de la secuencia de ADN expuesta en la SEC ID N<sup>o</sup>: 1 y se sintetiza en un sintetizador de ADN automático:

n<sup>o</sup> 4: TGCGGATCCAGCCGCTGCAGCCGCAAGCC (SEC ID N<sup>o</sup>: 21)

30 Los nueve primeros nucleótidos del cebador n<sup>o</sup> 4 (subrayados) comprenden la secuencia de reconocimiento para la endonucleasa de restricción BamHI que puede usarse para facilitar la manipulación de una secuencia de ADN amplificada específicamente que codifica la proteína BMP-12 humana de la invención y, por lo tanto, no derivan de la secuencia de ADN presentada en la SEC ID N<sup>o</sup>: 1. El siguiente cebador oligonucleotídico se diseña basándose en los nucleótidos n<sup>o</sup> 866 – n<sup>o</sup> 885 de la secuencia de ADN expuesta en la SEC ID N: 1 y se sintetiza en un sintetizador de ADN automático.

n<sup>o</sup> 5 GACTCTAGACTACCTGCAGCCGAGGCCT (SEC ID N<sup>o</sup>: 22)

35 Los nueve primeros nucleótidos del cebador n<sup>o</sup> 5 (subrayados) comprenden la secuencia de reconocimiento para la endonucleasa de restricción XbaI que puede usarse para facilitar la manipulación de una secuencia de ADN amplificada específicamente que codifica la proteína BMP-12 humana de la invención y, por lo tanto, no proceden de la secuencia de ADN presentada en la SEC ID N<sup>o</sup>: 1.

40 Los símbolos de nucleótidos convencionales en los cebadores identificados anteriormente son los siguientes: A, adenina; C, citosina; G, guanina; T, timina.

45 Los cebadores n<sup>o</sup> 4 y n<sup>o</sup> 5 identificados anteriormente se utilizan como cebadores para permitir la amplificación de una secuencia de nucleótidos que codifica BMP-12 específica a partir de bibliotecas de ADNc preestablecidas que pueden incluir las siguientes: ADNc de cerebro fetal humano/ $\lambda$ ZAPII (Stratagene, n<sup>o</sup> de catálogo 936206), hígado humano/ $\lambda$ UNI-ZAP XR (Stratagene, n<sup>o</sup> de catálogo 937200), pulmón humano/ $\lambda$ UNI-ZAP XR (Stratagene, n<sup>o</sup> de catálogo 937206) y bazo fetal humano/UNI-ZAP XR (Stratagene, n<sup>o</sup> de catálogo 937205).

50 Aproximadamente  $1 \times 10^8$  ufp (unidades formadoras de placas) de bibliotecas de bacteriófago  $\lambda$  que contienen insertos de ADNc humanos tales como los detallados anteriormente se desnaturalizan a 95 °C durante cinco minutos antes de la adición a una mezcla de reacción que contiene 200  $\mu$ M de cada desoxinucleótido trifosfato (dATP, dGTP, dCTP y dTTP), Tris-HCl 10 mM, pH 8,3, KCl 50 mM, MgCl<sub>2</sub> 1,5 mM, gelatina al 0,001%, 1,25 unidades de ADN polimerasa Taq, una concentración 100 pM de cebador oligonucleotídico n<sup>o</sup> 4 y una concentración 100 pM de cebador oligonucleotídico n<sup>o</sup> 5. La mezcla de reacción después se somete a ciclos términos de la siguiente manera: 1 minuto a 94 °C, 1 minuto a 50 °C, 1 minuto a 72 °C durante treinta y nueve ciclos, seguido de 10 minutos a 72 °C.

Sería de esperar que el ADN que se amplifica específicamente por esta reacción generara un producto codificante de BMP-12 de aproximadamente 333 pares de bases, cuyas 315 pb internas corresponden a los nucleótidos n<sup>o</sup> 571

a nº 885 de la SEC ID Nº: 1 y que también incluyen 9 pb en cada extremo del fragmento específico de BMP-12 que corresponden a los sitios de restricción definidos por los nucleótidos nº 1 – nº 9 de los cebadores nº 4 y nº 5. El producto de ADN de 333 pb resultante se digiere con las endonucleasas de restricción BamHI y XbaI, se extrae con fenol, se extrae con cloroformo y se precipita con etanol.

5 Como alternativa a la precipitación con etanol, se realiza intercambio de tampón y extracción de fragmentos pequeños de ADN resultante de la digestión de restricción con BamHI/XbaI por dilución del producto de ADN digerido en Tris-HCl 10 mM, pH 8,0, EDTA 1 mM seguido de centrifugación a través de un microconcentrador Centricon™ 30 (W.R. Grace & Co., Beverly, MA; nº de producto 4209). El producto de ADN amplificado digerido con BamHI/XbaI resultante se subclona en un vector plasmídico (es decir, pBluescript, pGEM-3 etc.) entre los sitios BamHI y XbaI de la región de poliengarce. Se necesitaría el análisis de la secuencia de ADN de los subclones resultantes para confirmar la integridad del inserto codificante de BMP-12. Una vez que se ha identificado una fuente de ADNc positiva de esta manera, podría explorarse directamente la biblioteca de ADNc correspondiente a partir de la cual se amplificó una secuencia específica de BMP-12 de 333 pb con el inserto de 333 pb u otras sondas específicas de BMP-12 para identificar y aislar clones de ADNc que codifiquen la proteína de BMP-12 de longitud completa de la invención.

Pueden usarse otros procedimientos conocidos por los expertos en la materia para aislar otros ADNc de longitud completa que codifiquen proteínas relacionadas con BMP-12 humanas, o clones de ADNc de longitud completa que codifiquen proteínas relacionadas con BMP-12 de la invención de especies distintas de seres humanos, particularmente otras especies de mamífero.

20 Los siguientes ejemplos demuestran el uso de la secuencia de BMP-12 humana para aislar homólogos de proteínas relacionadas con BMP-12 en una biblioteca de ADN genómico murino.

Se predice que la secuencia de ADN que codifica la proteína BMP-12 humana de la invención es significativamente homóloga a BMP-12 y secuencias relacionadas con BMP-12 de especies distintas de seres humanos, de tal forma que podría utilizarse para amplificar específicamente secuencias de ADN de esas otras especies que codifiquen las proteínas relacionadas con BMP-12 correspondientes. Específicamente, los siguientes oligonucleótidos se diseñan basándose en la secuencia de BMP-12 humana (SEC ID Nº: 1) y se sintetizan en un sintetizador de ADN automático:

nº 6: GCGGATCCAAGGAGCTCGGCTGGGACGA (SEC ID Nº: 23)

nº 7: GGAATTC~~CCC~~ACCACCATGTCCTCGTAT (SEC ID Nº: 24)

30 Los ocho primeros nucleótidos de los cebadores oligonucleotídicos nº 6 y nº 7 (subrayados) comprenden la secuencia de reconocimiento para las endonucleasas de restricción BamHI y EcoRI, respectivamente. Estas secuencias se utilizan para facilitar la manipulación de una secuencia de ADN amplificada específicamente que codifica BMP-12 o una proteína relacionada con BMP-12 de una especie distinta del ser humano y, por lo tanto, no derivada de la secuencia de ADN presentada en la SEC ID Nº: 1. El cebador oligonucleotídico nº 6 se diseña basándose en los nucleótidos nº 607-nº 626 de la SEC ID Nº: 1. El cebador oligonucleotídico nº 7 se diseña basándose en el complemento inverso de los nucleótidos nº 846-nº 865 de la secuencia de ADN expuesta en la SEC ID Nº: 1.

40 Los cebadores oligonucleotídicos nº 6 y nº 7 identificados anteriormente se utilizan como cebadores para permitir la amplificación de secuencias relacionadas con BMP-12 específicas a partir de ADN genómico derivado de especies distintas de seres humanos. La reacción de amplificación se realiza como se indica a continuación:

Se corta ADN genómico (fuente: cepa Balb c) por pase repetido a través de una aguja de calibre 25, se desnatura a 100 °C durante cinco minutos y después se enfría en hielo antes de la adición a una mezcla de reacción que contiene una concentración 200 µM de cada desoxinucleótido trifosfato (dATP, dGTP, dCTP y dTTP), Tris-HCl 10 mM, pH 8,3, KCl 50 mM, MgCl<sub>2</sub> 1,5 mM, gelatina al 0,001%, 1,25 unidades de ADN polimerasa Taq, una concentración 100 pM de cebador oligonucleotídico nº 6 y una concentración 100 pM de cebador oligonucleotídico nº 7. Después, la mezcla de reacción se somete a ciclos térmicos de la siguiente manera: 1 minuto a 95 °C, 1 minuto a 55 °C, 1 minuto a 72 °C durante cuarenta ciclos, seguido de 10 minutos a 72 °C.

50 El ADN que se amplifica específicamente por esta reacción se precipita con etanol, se digiere con las endonucleasas de restricción BamHI y EcoRI y se somete a electroforesis en gel de agarosa. Una región del gel, que corresponde al tamaño previsto del fragmento de ADN que codifica BMP-12 murina o una proteína relacionada con BMP-12, se escinde y los fragmentos de ADN amplificados específicamente contenidos en el gel se extraen (por electroelución o por otros procedimientos conocidos por los expertos en la materia) y se subclona en un vector plasmídico, tal como pGEM-3 o pBluescript entre los sitios BamHI y EcoRI del poliengarce. El análisis de la secuencia de ADN de uno de los subclones resultantes, denominado mV1, indica que la secuencia de ADN amplificada específicamente contenida en su interior codifica una parte de una proteína que parece ser el homólogo murino de la secuencia de BMP-12 o

VL-1 de la invención. La secuencia de ADN (SEC ID N°: 10) y la secuencia de aminoácidos derivada (SEC ID N°: 11) de este fragmento de ADN murino amplificado específicamente se muestran en el listado de secuencias.

5 Los nucleótidos n° 1-n° 26 de la SEC ID N°: 10 comprenden una parte de oligonucleótido. El nucleótido n° 6 y los nucleótidos n° 246-n° 272 comprenden una parte del complemento inverso del oligonucleótido n° 7 utilizado para realizar la reacción de amplificación específica. El nucleótido n° 27 de la SEC ID N°: 10 parece ser el último nucleótido de un triplete de codón, y los nucleótidos n° 244-n° 245 de la SEC ID N°: 10 parecen ser los dos primeros nucleótidos de un triplete de codón. Por lo tanto, los nucleótidos n° 28 a n° 243 de la SEC ID N°: 10 corresponden a una secuencia codificante parcial de mV1. Debido a la función de los oligonucleótidos n° 6 y n° 7 en la iniciación de la reacción de amplificación, pueden no corresponder exactamente a la secuencia real que codifica el homólogo murino a la proteína BMP-12 humana o VIL-1 de la invención y, por lo tanto, pueden no traducirse en la derivación de secuencia de aminoácidos correspondiente (SEC ID N°: 11).

Los expertos en la materia pueden utilizar sondas oligonucleotídicas diseñadas basándose en la secuencia de ADN de BMP-12 o VL-1 murina amplificada específicamente expuesta en la SEC ID N°: 10 para identificar clones codificantes de BMP-12 o VL-1 murinas de longitud completa (ADNc o genómico).

15 El análisis de la secuencia de ADN de otro de los subclones resultantes, denominado mV2, indica que la secuencia de ADN amplificada específicamente contenida en su interior codifica una parte de una secuencia relacionada con \*BMP-12 murina de la invención. La secuencia de ADN (SEC ID N°: 12) y la secuencia de aminoácidos derivada (SEC ID N°: 13) de este fragmento de ADN murino amplificado específicamente se muestran en el listado de secuencias.

20 Los nucleótidos n° 1-n° 26 de la SEC ID N°: 12 comprenden una parte del oligonucleótido n° 6 y los nucleótidos n° 246-n° 272 comprenden una parte del complemento inverso del 31 oligonucleótido n° 7 utilizado para realizar la reacción de amplificación específica. El nucleótido 27 de la SEC ID N°: 12 parece ser el último nucleótido de un triplete de codón, y los nucleótidos n° 244-n° 245 de la SEC ID N°: 12 parecen ser los dos primeros nucleótidos de un triplete de codón. Por lo tanto, los nucleótidos n° 28 a n° 243 de la SEC ID N°: 12 corresponden a una secuencia codificante parcial de mV2. Debido a la función de los oligonucleótidos n° 6 y n° 7 en la iniciación de la reacción de amplificación, pueden no corresponder exactamente a la secuencia real que codifica la proteína relacionada con BMP-12 murina de la invención y, por lo tanto, no se traducen en la derivación de secuencia de aminoácidos correspondiente (SEC ID N°: 13).

30 Las sondas oligonucleotídicas diseñadas basándose en la secuencia de ADN relacionada con BMP-12 murina amplificada específicamente expuesta en la SEC ID N°: 12 pueden utilizarse por los expertos en la materia para identificar clones que codifican proteínas relacionadas con BMP-12 murina de longitud completa (ADNc o genómico).

35 El análisis de la secuencia de ADN de otro de los subclones resultantes denominado mV9 indica que la secuencia de ADN amplificada específicamente contenida en su interior codifica una parte de una secuencia relacionada con BMP-12 murina de la invención. Esta secuencia parece ser el homólogo murino de la secuencia de ADN de BMP52 descrita en la SEC ID N°: 3. La secuencia de ADN (SEC ID N°: 14) y la secuencia de aminoácidos derivada (SEC ID N°: 15) de este fragmento de ADN murino amplificado específicamente se muestran en el listado de secuencias.

40 Los nucleótidos n° 1-n° 26 de la SEC ID N°: 14 comprenden una parte del oligonucleótido n° 6 y los nucleótidos n° 246-n° 272 comprenden una parte del complemento inverso del oligonucleótido n° 7 utilizado para realizar la reacción de amplificación específica. El nucleótido n° 27 de la SEC ID N°: 14 parece ser el último nucleótido de un triplete de codón, y los nucleótidos n° 244-n° 245 de la SEC ID N°: 14 parecen ser los dos primeros nucleótidos de un triplete de codón. Por lo tanto, los nucleótidos n° 28 a n° 243 de la SEC ID N°: 14 corresponden a una secuencia codificante parcial de mV9. Debido a la función de los oligonucleótidos n° 6 y n° 7 en la iniciación de la reacción de amplificación, pueden no corresponder exactamente a la secuencia real que codifica la proteína relacionada con BMP-12 murina de la invención y, por lo tanto, no traducirse en la derivación de secuencia de aminoácidos correspondiente (SEC ID N°: 15).

Los expertos en la materia pueden utilizar sondas oligonucleotídicas diseñadas basándose en la secuencia de ADN relacionada con BMP-12 murina amplificada específicamente expuesta en la SEC ID N°: 14 para identificar clones que codifican proteína relacionada con BMP-12 murina de longitud completa (ADNc o genómico).

50 Como alternativa, se utilizan cebadores oligonucleotídicos n° 6 y n° 7 identificados anteriormente como cebadores para permitir la amplificación específica de una sonda de ADN de 275 pares de bases, de la que 259 pb internos corresponden a los nucleótidos n° 607 a n° 865 de la SEC ID N°: 1 del subclón del plásmido codificante de BMP-12 PCR1-1#2. Esta sonda de ADN de 275 pb se marcó radiactivamente con <sup>32</sup>P y se empleó para explorar una biblioteca genómica murina construida en el vector X FI XX (Stratagene, n° de catálogo 946306). 1 millón de recombinantes de la biblioteca genómica murina se cultivan a una densidad de aproximadamente 20.000 recombinantes por placa en 50 placas. Se hibridan réplicas en nitrocelulosa duplicadas de las placas de bacteriófago recombinante en condiciones de rigurosidad reducida, con sonda de 333 pb amplificada específicamente en tampón de hibridación convencional (SHB = SSC 5X, SDS al 0,1%, Denhardt 5X, 100 µg/ml de ADN de espermatozoos de salmón)

a 60 °C durante una noche. Al día siguiente se retira la solución de hibridación que contiene oligonucleótido marcado radiactivamente y los filtros se lavan, en condiciones de rigurosidad reducida, con SSC 2X, SDS al 0,1% a 60 °C. Se identifican múltiples recombinantes de hibridación positiva y se purifican en placa. Los fragmentos de los clones recombinantes genómicos murinos de hibridación positiva se subclonan en vectores plasmídicos convencionales (es decir, pGEM-3) y se someten a análisis de la secuencia de ADN.

El análisis de la secuencia de ADN de uno de estos subclones denominado MvR3 indica que codifica una parte del gen de ratón correspondiente al producto de PCR mV1 (homólogo murino de la secuencia de BMP-12 humana expuesta en la SEC ID N°: 1) descrita anteriormente. La secuencia de ADN parcial de este subclón y la traducción en aminoácidos correspondiente se exponen en la SEC ID N°: 29 y SEC ID N°: 30, respectivamente.

El análisis de la secuencia de ADN de otro de estos subclones denominado MVR32 indica que codifica una parte del gen de ratón correspondiente al producto de PCR mV2 (homólogo murino de la secuencia de VL-1 humana expuesta en la SEC ID N°: 7) descrito anteriormente. La secuencia de ADN parcial de este subclón y la traducción de aminoácidos correspondiente se exponen en las SEC ID N°: 31 y SEC ID N°: 32 respectivamente.

El análisis de la secuencia de ADN de otro de estos subclones denominado MVR23 indica que codifica una parte del gen de ratón correspondiente al producto de PCR mV9 (homólogo murino de la secuencia de MP-52 expuesta en la SEC ID N°: 3) descrita anteriormente.

De una manera similar a lo que se ha descrito anteriormente para identificar y aislar clones genómicos humanos que codifican la proteína BMP-12 de la invención, pueden diseñarse sondas oligonucleotídicas correspondientes a la secuencia codificante de VL-1 expuesta en la SEC ID N°: 7 y utilizarse para identificar secuencias genómicas o de ADNc humanas que codifican la proteína VL-1 (BMP-13). Estos oligonucleótidos se diseñarían para regiones específicas para secuencias codificantes de VL-1 y, por lo tanto, probablemente procederían de regiones del menor grado de identidad de secuencia de nucleótidos entre la secuencia que codifica VL-1 amplificada específicamente (SEC ID N°: 7) y la secuencia codificante de BMP-12 amplificada específicamente (SEC ID N°: 5).

Como alternativa, los cebadores oligonucleotídicos n° 4 y n° 5 identificados anteriormente se utilizan como cebadores para permitir la amplificación específica de una sonda de ADN de 333 pares de bases, de la que 315 pb internos corresponden a los nucleótidos n° 571 a n° 885 de la SEC ID N°: 1, a partir del subclón de plásmido codificante de BMP-12 PCRI-1#2. Esta sonda de ADN de 333 pb se marcó radiactivamente con <sup>32</sup>P y se empleó para explorar una biblioteca genómica humana construida en el vector XDASH II (Stratagene, n° de catálogo 945203). 1 millón de recombinantes de la biblioteca genómica humana se cultivan a una densidad de aproximadamente 20.000 recombinantes por placa en 50 placas. Se hibridan réplicas de nitrocelulosa duplicadas de las placas de bacteriófago recombinante en condiciones de rigurosidad reducida, con la sonda de 333 pb amplificada específicamente en tampón de hibridación convencional (SHB = SSC 5X, SDS al 0,1%, Denhardt 5X, 100 µg/ml de ADN de esperma de salmón) a 60 °C durante una noche. Al día siguiente, se retira la solución de hibridación que contiene oligonucleótido marcada radiactivamente y los filtros se lavan, en condiciones de rigurosidad reducida, con SSC 2X, SDS al 0,1% a 60 °C. Se identifican múltiples (aproximadamente 15) recombinantes de hibridación positiva y se purifican en placa.

Para distinguir recombinantes de hibridación positiva que codifican la proteína VL-1 de la invención de BMP-12 y otros recombinantes que codifican proteínas relacionadas con BMP-12 que, según podría predecirse, hibridarían positivamente con la sonda de ADN de 333 pb generada a partir del plásmido codificante de BMP-12 PCRI-1#2 utilizado este procedimiento de exploración, se diseña la siguiente sonda oligonucleotídica basándose en la secuencia de VL-1 expuesta en la SEC ID N°: 7, y se sintetiza en un sintetizador de ADN automático:

n° 8: TGTATGCGACTTCCCGC [SECUENCIA ID N°: 35]

Un oligonucleótido correspondiente a los nucleótidos n° 60 a n° 76 de la SEC ID N°: 7 que contiene 5 diferencias de nucleótidos con la región correspondiente de la secuencia codificante de BMP-12 expuesta en la SEC ID N°: 1 (nucleótidos n° 672 a n° 689). Uno de los clones de bacteriófago recombinante que hibrida con la sonda oligonucleotídica de VL-1 n° 8 se denomina λJLDc31. Este clon de bacteriófago recombinante se purifica en placa, se realiza una reserva de placa de bacteriófago y se aísla ADN del bacteriófago a partir del clon genómico humano λJLDc31. El bacteriófago λJLDc31 se ha depositado en la Colección Americana de Cultivos Tipo, 12301 Parklawn Drive, Rockville, MD "ATCC" con el número de acceso n° 75922 el 20 de octubre de 1994. Este depósito satisface los requisitos del Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos a los Fines del Procedimiento en Materia de Patentes y Regulaciones expuestas en el mismo. La región de hibridación de oligonucleótidos de este recombinante, λJLDc31, está localizada en un fragmento Eco RI de 2,5 kb. Este fragmento se subclona en un vector plasmídico (pGEM-3) y se realiza el análisis de la secuencia de ADN. Este subclón de plásmido se denomina pGEMJLDc31/2.5 y se ha depositado en la Colección Americana de Cultivos Tipo, 12301 Parklawn Drive, Rockville, MD "ATCC" con el número de acceso n° 69710 el 20 de octubre de 1994. Este depósito satisface los requisitos del Tratado de Budapest sobre el Reconocimiento Internacional del Depósito de Microorganismos a los Fines del Procedimiento en Materia de Patentes y Regulaciones expuestas en el mismo.

La secuencia de ADN parcial (SEC ID N°: 25) y la secuencia de aminoácidos derivada (SEC ID N°: 26) de una parte del inserto de ADN de 2,5 kb del subclón plasmídico pGEMJLDc31/2.5, derivado del clon  $\lambda$ JLDc31, se muestran en el Listado de Secuencias.

5 La secuencia de ADN de una parte del inserto EcoRI de 2,5 kb del plásmido pGEMJLDc31/2.5 se expone en la SEC ID N°: 25.

10 Contiene un marco de lectura abierto de 912 pb, como se define por los nucleótidos n° 52 a n° 963 de la SEC ID N°: 25. Como esta secuencia deriva de un clon genómico, es difícil determinar el límite entre el extremo 5' de la secuencia codificante y el límite 3' de la secuencia intermedia (secuencia de intrón/no codificante). La fase de lectura abierta entera (nucleótidos n° 52 a n° 963 de la SEC ID N°: 25) codifica una parte de la proteína VL-1 de la invención de hasta 304 aminoácidos.

15 Basándose en el conocimiento de otras proteínas BMP y otras proteínas dentro de la familia de TGF- $\beta$ , se predice que el polipéptido precursor se escindiría en la secuencia multibásica Arg-Arg-Arg-Arg de acuerdo con una secuencia de procesamiento proteolítico consenso Arg-X-X-Arg. Es de esperar que la escisión del polipéptido precursor de VL-1 genere un péptido maduro de 120 aminoácidos que empieza con el aminoácido Thr en la posición n° 1 de la SEC ID N°: 26. Es de esperar que el procesamiento de VL-1 en la forma madura implique la dimerización y eliminación de la región N-terminal de una manera análoga al procesamiento de la proteína relacionada con TGF- $\beta$  [Gentry y col., Molec & Cell. Biol., 8:4162 (1988); Derynck y col. Nature, 316:701 (1985)].

20 Por lo tanto, se contempla que la especie activa madura de VL-1 comprende un homodímero de dos subunidades polipeptídicas, comprendiendo cada subunidad los aminoácidos n° 1 a n° 120 de la SEC ID N°: 26 con un peso molecular previsto de aproximadamente 12.000 daltons. Se contemplan otras especies activas que comprenden al menos los aminoácidos n° 19 a n° 119 o n° 120 de la SEC ID N°: 26, incluyendo de esta manera el primer y el último restos de cisteína conservados.

25 Usando dicho procedimiento, se obtuvo un clon que codificaba la VL-1 humana madura (BMP-13). La secuencia de nucleótidos y la secuencia de aminoácidos correspondientes codificadas por este clon se presentan en el Listado de Secuencias en las SEC ID N°: 25 y 26, respectivamente.

## **Ejemplo 2**

### **Expresión de BMP-12**

30 Para producir proteínas BMP-12 humanas, el ADN que las codifica se transfiere a un vector de expresión apropiado y se introduce en células de mamífero u otros huéspedes eucariotas o procariontes preferidos por técnicas de ingeniería genética convencionales.

Para producir a producir la proteína BMP-12 humana en células bacterianas, se emplea el siguiente procedimiento.

### **Expresión de BMP-12 en *E. coli***

35 Se construyó un plásmido de expresión pALV1-781, para la producción de BMP-12 en *E. coli*, que contiene las siguientes características principales. Los nucleótidos 1-2060 contienen secuencias de ADN procedentes del plásmido pUC-18 [Norrander y col., Gene 26: 101-106 (1983)] incluyendo secuencias que contienen el gen para la  $\beta$ -lactamasa que confiere resistencia al antibiótico ampicilina en cepas de *E. coli* huésped, y un origen de replicación derivado de *colE1*. Los nucleótidos 2061-2221 contienen secuencias de ADN para el promotor a la izquierda principal (pL) del bacteriófago  $\lambda$  [Sanger y col., J. Mol. Biol. 162: 729-773 (1982)], incluyendo, tres secuencias operadoras O<sub>L1</sub>, O<sub>L2</sub> y O<sub>L3</sub>. Los operadores son los sitios de unión para la proteína represora  $\lambda$ cI, cuyos niveles intracelulares controlan la cantidad de iniciación de la transcripción a partir de pL. Los nucleótidos 2222-2723 contienen una secuencia de unión a ribosomas fuerte incluida en una secuencia procedente de los nucleótidos 35566 a 35472 y 38137 a 38361 del bacteriófago lambda como se describe en Sanger y col., J. Mol. Biol. 162: 729-773 (1982). Los nucleótidos 2724-3041 contienen una secuencia de ADN que codifica la proteína BMP-12 madura con toda la secuencia 3' no traducida retirada. Las secuencias de ADN de BMP-12 introducidas en el vector de expresión 40 pALV1-781 se modificaron en el extremo 5' para elevar el contenido de A+T sin alterar la capacidad codificante. Estos cambios se realizaron para aumentar la eficacia de la traducción iniciada en el ARNm del BMP-12 en *E. coli*. Los nucleótidos 3042-3058 proporcionan una secuencia de ADN de "enganche" que contiene sitios de endonucleasas de restricción. Los nucleótidos 3059-3127 proporcionan una secuencia de terminación de la transcripción basada en la del gen de asp A de *E. coli* [Takagi y col., Nucl. Acids Res. 13:2063-2074 (1985)]. Los nucleótidos 3128-3532 son 45 secuencias de ADN derivadas de pUC-18.

50 El plásmido pALV1-781 se utilizó para transformar la cepa huésped *E. coli* G1724 (F,  $\text{lacI}^{\Delta}$ ,  $\text{lacp}^{\text{L8}}$ , ampC::  $\lambda$ cI+) por el procedimiento de Dagert y Ehrlich, Gene 6:23 (1979). G1724 (ATCC, n° de acceso 55151) contiene una copia del gen represor de  $\lambda$ cI de tipo silvestre integrado de forma estable en el cromosoma en el locus ampC, donde se ha colocado bajo control transcripcional de las secuencias de promotor/operador de trp de *Salmonella typhimurium*. En

GI724, la proteína  $\lambda$ CI se obtiene únicamente durante el crecimiento en medio sin triptófano, tal como medio mínimo o un medio mínimo suplementado con casaminoácidos tales como IMC, descrito anteriormente. La adición de triptófano a un cultivo de GI724 reprimirá el promotor de *trp* e inactivará la síntesis de  $\lambda$ CI, causando gradualmente la inducción de la transcripción a partir de promotores de pL si están presentes en la célula.

- 5 Se seleccionaron transformantes en placas de agar al 1,5% p/v que contenían medio IMC, que está compuesto de medio M9 [Miller, "Experiments in Molecular Genetics", Cold Spring Harbor Laboratory, New York (1972)] que contiene  $\text{MgSO}_4$  1 mM y suplementado con glucosa al 0,5% p/v, casaminoácidos al 0,2% p/v y 100  $\mu\text{g/ml}$  de ampicilina. Se cultivó GI724 transformado con pALV1-781 a 37 °C a una  $A_{550}$  de 0,5 en medio IMC que contenía 100  $\mu\text{g/ml}$  de ampicilina. Después se añadió triptófano a una concentración final de 100  $\mu\text{g/ml}$  y el cultivo se incubó durante 4 horas más. Durante este tiempo, la proteína BMP-12 se acumula dentro de la fracción de "cuerpos de inclusión".

### **Preparación de monómero de proteína**

- 15 Se pesaron 18 g de células congeladas y se resuspendieron en 60 ml de Tris 100 mM, EDTA 10 mM, fluoruro de fenilmetilsulfonilo [PMSF], 1 mM, pH 8,3. Las células se lisaron por 3 pases a través de un Microfluidizer™ [modelo nº MCF 100 T]. El sedimento de cuerpos de inclusión se obtuvo por centrifugación a 15.000 g a 4 °C durante 20 minutos. El sobrenadante se decantó y el sedimento se lavó con 100 ml de Tris 100 mM, NaCl 1,0 M, EDTA 10 mM, PMSF 1 mM, pH 8,3. La suspensión se centrifugó de nuevo a 15.000 g a 4 °C durante 10 minutos y el sobrenadante se decantó. Después, el sedimento se lavó con 100 ml de Tris 100 mM, EDTA 10 mM, Triton X-100 al 1%, PMSF 1 mM, pH 8,3. La suspensión se centrifugó de nuevo a 15.000 g a 4 °C durante 10 minutos y el sobrenadante se decantó. El sedimento se resuspendió con 50 ml de Tris 20 mM, EDTA 1 mM, PMSF 1 mM, pH 8,3, que contenía DTT al 1% en un homogeneizador de tejidos de vidrio. Después, la BMP-12 monomérica se solubilizó por acidificación a pH 2,5 con ácido acético glacial. La fracción soluble se aisló por centrifugación a 15.000 g durante 20 minutos a 4 °C.

- 25 El sobrenadante de esta centrifugación se recogió y se cromatografió en una columna de exclusión molecular Sephacryl S-100™ (83 cm x 2,6 cm; lecho  $\approx$ 440 ml) en incrementos de 20 ml. La columna de Sephacryl S-100™ se procesó con una fase móvil de ácido acético al 1% a un caudal de 1,4 ml/min. Las fracciones correspondientes al monómero de BMP-12 se detectaron por absorbancia a 280 nm y usando un ordenador se calculó un coeficiente de extinción de  $18200 \text{ M}^{-1}\text{m}^{-1}$  y un peso molecular (11667 daltons). El material reunido en la columna de exclusión molecular se usó como material de partida para reacciones de repliegamiento.

- 30 Como alternativa a lo anterior, se miden 1,0 g de células almacenadas a -80 °C. Se añade solución (3,4 ml de TRIS 100 mM, EDTA 10 mM, pH 8,5). La solución se somete a agitación vorticial hasta que las células se suspenden bien. Se añaden 40  $\mu\text{l}$  de PMSF 100 mM en isopropanol. Las células se lisan a 6,89 MPa (1000 psi) en una celda de presión French. Los cuerpos de inclusión se centrifugan a 4 °C durante 20 minutos en una microcentrífuga Eppendorf para formar sedimentos. Los sobrenadantes se decantan. A un sedimento (de 4 en total) se le añaden 1,0 ml de clorhidrato de guanidina 8,0 M desgasificado, TRIS 0,5 M, EDTA 5 mM, pH 8,5, que contiene DTT 250 mM. El sedimento se disuelve y se insufla argón sobre el líquido durante 30 segundos. A continuación, la solución se incuba a 37 °C durante una hora. El material insoluble se sedimenta durante 2-3 minutos en una microcentrífuga Eppendorf a 23 °C. Se inyectan 0,5-1,0 ml de sobrenadante en un cartucho auxiliar Supelco 2 cm (LC-304) y se eluye con un gradiente de acetonitrilo en TFA al 0,1% del 1-70% durante 35 minutos. La BMP-12 eluye entre 29 y 31 minutos. Las fracciones se reúnen y la concentración de proteína se determina por absorbancia a 280 nanómetros frente a TFA al 0,1%, usando el coeficiente de extinción teórico basado en el contenido de aminoácidos.

- 45 Como segundo procedimiento alternativo al anterior, se descongelan sedimentos de células congelados obtenidos a partir de los transformantes de *E. coli* como se ha descrito anteriormente en 30 ml de tampón TE8.3 (100:10) (Tris-HCl 100 mM pH 8,3,  $\text{Na}_2\text{EDTA}$  10 mM, PMSF 1 mM). Las células se lisan por tres pases a través de un Microfluidizer™ [modelo nº MCF 100 T]. El sedimento de material de cuerpos de inclusión inicial se disuelve en HCl-guanidina 8 M, tampón TE8.5 (100:10) (Tris-HCl 100 mM, pH 8,5,  $\text{Na}_2\text{EDTA}$  10 mM que contenía DTT 100 mM) y se incuba a 37 °C durante 1 hora. Este material se centrifuga a 12.000 x g durante 15 minutos a temperatura ambiente.

### **Replegamiento de la proteína BMP-12 usando el sistema CHAPS**

- 50 Se liofiliza un volumen suficiente de reserva de BMP-12 para dar 10  $\mu\text{g}$  de proteína. Se añaden 5  $\mu\text{l}$  de agua destilada en alambique de vidrio para redissolver el residuo, y después se añaden 100  $\mu\text{l}$  de mezcla de replegamiento – (Tris 50 mM, NaCl 1,0 M, 1-propano-sulfonato de 3-(3-cloramidopropil)dimetilamonio (CHAPS) al 2%, EDTA 5 mM, glutatión 2 mM (reducido), glutatión 1 mM (oxidado); a pH de aproximadamente 8,5). La solución se mezcla suavemente y se almacena a 23 °C durante 1-4 días. La formación de dímeros se evalúa por procesamiento de una alícuota en un gel de tricina Novex al 16% a 125 voltios durante 2,5 horas, seguido de tinción con Azul de Coomassie y destinción.

El dímero de BMP-12 se purificó usando una columna C4 de RP-HPLC analítica (cromatografía líquida de alta

resolución de fase inversa) (Vydac 214TP54) que se equilibró hasta Tampón B al 1% (diluido en tampón A) y se procesó durante 35 minutos, tiempo durante el cual la proteína eluye, usando el siguiente gradiente (tampón A = ácido trifluoroacético al 0,1%, tampón B = acetonitrilo a 95%, ácido trifluoroacético [TFA] al 0,1% con un caudal de 1 ml/min:

- 5 1-5 minutos de tampón B al 20%
- 5-10 minutos de tampón B al 20-30%
- 10-30 minutos de tampón B al 30-50%
- 30-35 minutos de tampón B al 50-100%

- 10 La proteína se monitorizó por absorbancia a 280 nm. Se reunieron fracciones pico de BMP-12 (que eluían entre 29 y 31 minutos). La pureza se evaluó por SDS-PAGE. La concentración se determinó por absorbancia a 280 nm, y usando el coeficiente de extinción calculado por ordenador y peso molecular indicados anteriormente.

### **Expresión de BMP-12 en células de mamífero**

Otro sistema de expresión preferido contemplado para BMP-12 humana recombinante biológicamente activa es células de mamífero transformadas de forma estable.

- 15 Un experto en la materia puede construir vectores de expresión de mamífero empleando la secuencia de la SEC ID N° 1 u otras secuencias de ADN que codifican proteínas BMP-12 u otras secuencias modificadas y vectores conocidos, tales como pCD [Okayama y col., Mol. Cell Biol., 2: 161-170 (1982)], pJL3, pJL4 [Gough y col., EMBO J., 4: 645-653 (1985)] y pMT2 CXM.

- 20 El vector de expresión de mamífero pMT2 CXM es un derivado de p91023(b) (Wong y col., Science 228: 810-815, 1985) que difiere de este último en que contiene el gen de resistencia a ampicilina en lugar del gen de resistencia a tetraciclina y contiene además un sitio XhoI para la inserción de clones de ADNc. Se han descrito los elementos funcionales de pMT2 CXM (Kaufman, R.J., 1985, Proc. Natl. Acad. Sci. USA 82: 689-693) e incluyen los genes VA de adenovirus, el origen de replicación de SV40 que incluye el potenciador de 72 pb, el promotor tardío principal de adenovirus que incluye un sitio de corte y empalme 5' y la mayor parte de la secuencia líder tripartita de adenovirus
- 25 presente en ARNm tardíos de adenovirus, un sitio aceptor de corte y empalme 3', un inserto de DHFR, el sitio de poliadenilación temprano de SV40 (SV40) y secuencias de pBR322 necesarias para la propagación en *E. coli*.

- 30 El plásmido pMT2 CXM se obtiene por digestión con EcoRI de pMT2-VWF, que se ha depositado con la Colección Americana de Cultivos Tipo (ATCC), Rockville, MD (USA) con el número de acceso ATCC 67122. La digestión con EcoRI escinde el inserto de ADNc presente en pMT2-VWF, produciendo pMT2 en forma lineal que puede ligarse y usarse para transformar *E. coli* HB 101 o DH-5 en bacterias con resistencia a ampicilina. El ADN del plásmido pMT2 puede prepararse por procedimientos convencionales. Después se construye pMT2 CXM usando mutagénesis loopout/in [Morinaga, y col., Biotechnology 84: 636 (1984). Esto elimina las bases 1075 a 1145 con respecto al sitio Hind III cerca del origen de replicación de SV40 y las secuencias potenciadoras de pMT2. Además, inserta una secuencia que contiene el sitio de reconocimiento para la endonucleasa de restricción Xho I. Un derivado de
- 35 pMT2CXM, denominado pMT23, contiene sitios de reconocimiento para las endonucleasas de restricción PstI, Eco RI, Sall y XhoI. El ADN del plásmido pMT2 CXM y pMT23 puede prepararse por procedimientos convencionales.

- 40 pEMC2β1 derivado de pMT21 también puede ser adecuado en la puesta en práctica de la invención. pMT21 procede de pMT2 que procede de pMT2-VWF. Como se ha descrito anteriormente, la digestión con EcoRI escinde el inserto de ADNc presente en pMT-VWF, produciendo pMT2 en forma lineal que puede ligarse y usarse para transformar *E. coli* HB 101 o DH-5 en bacterias con resistencia a ampicilina. El ADN del plásmido pMT2 puede prepararse por procedimientos convencionales.

- 45 pMT21 procede de pMT2 mediante las dos siguientes modificaciones. En primer lugar, se delecionan 76 pb de la región 5' no traducida del ADNc de DHFR incluyendo un tramo de 19 restos de G de la cola G/C para la clonación del ADNc. En este procedimiento, se inserta un sitio XhoI para obtener la siguiente secuencia inmediatamente cadena arriba de DHFR. En segundo lugar, se introduce un único sitio ClaI por digestión con EcoRV y XbaI, tratamiento con fragmento de Klenow de ADN polimerasa I y ligamiento a un engarce ClaI (CATCGATG). Esto deleciona un segmento de 250 pb de la región de ARN asociada a adenovirus (VAI) pero no interfiere con la expresión o función del gen de ARN VAI. pMT21 se digiere con EcoRI y XhoI y se usa para obtener el vector pEMC2B1.

- 50 Se obtiene una parte del líder de EMCV a partir de pMT2-ECAT1 [S.K. Jung, y col., J. Virol 63:1651-1660 (1989)] por digestión con Eco RI y PstI, dando como resultado un fragmento de 2752 pb. Este fragmento se digiere con TaqI produciendo un fragmento Eco RI-TaqI de 508 pb que se purifica por electroforesis en un gel de agarosa de bajo punto de fusión. Se sintetiza un adaptador de 68 pb y su cadena complementaria con un extremo saliente 5' TaqI y un extremo saliente 3' XhoI que tiene una secuencia que corresponde a la secuencia líder del virus EMC desde el
- 55 nucleótido 763 al 827. También cambia ATG en la posición 10 dentro del líder del virus EMC a ATT y se continúa por un sitio XhoI. Un ligamiento de tres vías del fragmento Eco RI-XhoI de pMT21; el fragmento EcoRI-TaqI del virus

EMC y el adaptador oligonucleotídico de 68 pb TaqI-XhoI da como resultado el vector pEMC2β1.

Este vector contiene el origen de replicación y potenciador de SV40, el promotor tardío principal de adenovirus, una copia de ADNc de la mayor parte de la secuencia líder tripartita de adenovirus, una pequeña secuencia intermedia híbrida, una señal de poliadenilación de SV40 y el gen del adenovirus VA I, marcadores de DHFR y β-lactamasa y una secuencia de EMC, en relaciones apropiadas para dirigir el alto nivel expresión del ADNc deseado en células de mamífero.

La construcción de vectores puede implicar la modificación de las secuencias de ADN de BMP-12. Por ejemplo, el ADNc de BMP-12 puede modificarse retirando los nucleótidos no codificantes de los extremos 5' y 3' de la región codificante. Los nucleótidos no codificantes delecionados pueden reemplazarse o no por otras secuencias consideradas beneficiosas para la expresión. Estos vectores se utilizan para transformar células huésped apropiadas para la expresión de proteínas BMP-12. Además, la secuencia de la SEC ID N°: 1 u otras secuencias que codifican proteínas BMP-12 pueden manipularse para expresar la proteína BMP-12 por aislamiento de la secuencia codificante madura de los nucleótidos 571 a 882 de la SEC ID N°: 1 y adición a las secuencias del extremo 5' que codifican los propéptidos completos de otras proteínas BMP.

Por ejemplo, un experto en la materia puede fabricar una proteína de fusión en la que el propéptido de BMP-2 está unido de forma operativa al péptido de BMP-12 madura mediante la preparación de un vector de ADN en el que la secuencia de ADN que codifica el propéptido de BMP-2 está unida en el marco de lectura apropiado a la secuencia de ADN que codifica el péptido de BMP-12 madura. La secuencia de ADN de dicha proteína de fusión se muestra en la SECUENCIA ID N°: 27.

Un experto en la materia puede manipular las secuencias de la SEC ID N°: 1 eliminando o reemplazando las secuencias reguladoras de mamífero que flanquean la secuencia codificante con secuencias bacterianas para crear vectores bacterianos para la expresión intracelular o extracelular por células bacterianas, como se ha descrito anteriormente. Como otro ejemplo, las secuencias codificantes podrían manipularse adicionalmente (por ejemplo, ligarse a otros engarces conocidos o modificarse por deleción de secuencias no codificantes de las mismas o alteración de nucleótidos por otras técnicas conocidas). La secuencia codificante de BMP-12 modificada después podría insertarse en un vector bacteriano conocido usando procedimientos tales como los descritos en T. Taniguchi y col., Proc. Natl Acad. Sci. USA, 77: 5230-5233 (1980). Este vector bacteriano ejemplar después se podría utilizar para transformar células huésped bacterianas y expresar de esta manera una proteína BMP-12. Como estrategia para producir la expresión extracelular de proteínas BMP-12 en células bacterianas, véase, por ejemplo, la solicitud de patente europea EPA 177.343.

Pueden realizarse manipulaciones similares para la construcción de un vector de insecto [Véanse, por ejemplo, los procedimientos descritos en la publicación de la solicitud de patente europea 155.476] para la expresión en células de insecto. También podría construirse un vector de levadura empleando secuencias reguladoras de levadura para la expresión intracelular o extracelular de los factores de la presente invención por células de levadura [Véanse, por ejemplo, los procedimientos descritos en la publicación de la solicitud PTC WO 86/00639 y en la solicitud de patente europea EPA 123.289).

Un procedimiento para producir altos niveles de una proteína BMP-12 de la invención en células de mamífero puede implicar la construcción de células que contienen múltiples copias del gen de BMP-12 heterólogo. El gen heterólogo está unido a un marcador amplificable, por ejemplo, el gen de la dihidrofolato reductasa (DHFR) para el que pueden seleccionarse células que contienen más copias de genes para la propagación en concentraciones crecientes de metotrexato (MTX) de acuerdo con los procedimientos de Kaufman y Sharp, J. Mol. Biol., 159:601-629 (1982). Esta estrategia puede emplearse con varios tipos celulares diferentes.

Por ejemplo, puede co-introducirse un plásmido que contiene una secuencia de ADN para una BMP-12 de la invención en asociación operativa con otras secuencias de plásmido que permiten la expresión de la misma y el plásmido de expresión de DHFR pAdA26SV(A)3 [Kaufman y Sharp, Mol. Cell. Biol., 2: 1304 (1982)] en células CHO deficientes de DHFR, DUKX-BII, por diversos procedimientos incluyendo co-precipitación con fosfato cálcico y transfección, electroporación o fusión de protoplastos. Se seleccionan transformantes que expresan DHFR con respecto al crecimiento en medio alfa con suero bovino fetal dializado y posteriormente se seleccionan para la amplificación por crecimiento en concentraciones crecientes de MTX (por ejemplo, etapas secuenciales en MTX 0,02, 0,2, 1,0 y 5 μM) como se describe en Kaufman y col., Mol Cell Biol., 5: 1750 (1983). Los transformantes se clonan y se monitoriza la expresión de BMP-12 biológicamente activa por el ensayo en rata de Sampath-Reddi modificado por Rosen descrito más adelante en el Ejemplo 5. La expresión de BMP-12 aumentaría al aumentar los niveles de resistencia a MTX. Los polipéptidos de BMP-12 se caracterizan usando técnicas convencionales conocidas en este campo tales como el marcaje pulsátil con [35S] metionina o cisteína y electroforesis en gel de poliacrilamida. Pueden seguirse procedimientos similares para producir otras proteínas relacionadas con BMP-12.

**Ejemplo 3****Preparación de fusión de propéptido de BMP-2/péptido maduro de BMP-12**

Para construir un vector que codificara la fusión de propéptido de BMP-2/péptido maduro de BMP-12, se usó el siguiente procedimiento de clonación para fusionar las dos secuencias entre sí.

5 En primer lugar, se cortó un fragmento de enzimas de restricción de ADN que comprendía el propéptido de la proteína BMP-12 humana, que comprendía los nucleótidos 1 a 843 de la SEC ID N°: 27, a partir de pBMP2nEMC. pBMP2nEMC es un plásmido derivado de lambda U20S-39 (ATCC n° 40345) que comprende la secuencia codificante entera de la proteína BMP-2 humana con las secuencias 5' y 3' no traducidas de BMP-2 delecionadas del vector. La enzima de restricción 5' usada fue Bgl II y corta a pBMP2nEMC en el vector en el nucleótido 979. La enzima de restricción 3' usada era Mae II y corta pBMP2n EMC en el propéptido de BMP-2 en el nucleótido 1925, cerca del extremo carboxi. El producto de 954 pares de bases resultante después se aisló en gel y se sometió a GeneClean. En segundo lugar, se cortó un fragmento de ADN de enzimas de restricción que comprendía la parte 5' de la secuencia de ADN del péptido maduro de BMP-12 humana, a partir de pPCR1-1#2 V1-1 (ATCC n° 69517). La enzima de restricción 5' usada fue Eae I y corta pPCR1-#12 V1-1 en posición 3' con respecto al extremo N de la secuencia del péptido maduro de BMP-12 humana. El producto de 259 pares de bases resultante se aisló en gel y se sometió a GeneClean. En tercer lugar, se diseñaron dos oligos de ADN y se sintetizaron, de forma que cuando se hibridaran formarían un fragmento de ADN pequeño que comprendiera la secuencia de fusión del extremo 3' del propéptido de BMP-2 humana y el extremo 5' del péptido maduro de BMP-12. El fragmento de ADN tiene un extremo adherente complementario 5' Mae II que hibrida con el fragmento de enzimas de restricción 3' que comprende el propéptido de BMP-12 humana. El fragmento de ADN oligo hibridado tiene un extremo adherente complementario Eae I 3' que hibrida con el extremo 5' del fragmento de enzimas de restricción que comprende el péptido maduro de BMP-12 humana. El oligo de la cadena codificante se denomina B2/12 y tiene 13 pares de bases de longitud. A continuación, se obtuvo un fragmento de ADN que codifica los 123 pares de bases en el extremo 3' del fragmento del péptido maduro de BMP-12 como se indica a continuación. En primer lugar, se amplificó por PCR un fragmento de ADN que comprendía el propéptido de la proteína BMP-2 humana, que comprende los nucleótidos 1 a 846, a partir de pBMP2ΔEMC. El cebador 5' (oligo 655a) hibrida en la posición 5' del poliengarce. El cebador 3' (BMPpro3) hibrida con el propéptido de BMP-2 en el extremo 3' e introduce un sitio de enzima de restricción Bgl II por mutaciones de secuencia silenciosas. El producto de PCR resultante se cortó con Sal I, que escinde en el poliengarce, y Bgl II. El fragmento de enzimas de restricción de 850 pares de bases (que terminaba en la secuencia de aminoácidos REKR) se aisló en gel y se sometió a GeneClean. El péptido maduro de BMP-12 se amplificó por PCR usando un cebador 5' (oligo 5-1) que codificaba el sitio de enzimas de restricción Bgl II por mutaciones de secuencia silenciosas, e hibridación con el extremo 5' de un posible producto de escisión maduro, empezando con la secuencia de aminoácidos SCRS. El cebador 3' (V1-1 3) hibrida con el extremo 3' del péptido maduro de BMP-12 e introduce un sitio de enzima de restricción Xba I después del codón de terminación. El producto de PCR resultante se cortó con Bgl II y Xba I. El fragmento de enzimas de restricción de 321 pares bases se aisló en gel y se sometió a GeneClean.

Los dos fragmentos de restricción se ligaron por tres vías en un vector cortado previamente con Sall y XbaI. La construcción resultante se secuenció para comprobar los errores inducidos por la PCR y se observó una mutación silenciosa de C a T en el par de bases 185 en el propéptido. Este plásmido se denominó pREKRSRC. Después, pREKRSRC se cortó con BgIII y Ngo MI, y el fragmento de vector que incluía los últimos 123 pares de bases de la secuencia madura de BMP12 se aisló de esta manera. Los tres fragmentos de restricción y el oligoengarce hibridado se ligaron por cuatro vías para producir pREKR-TAL con el propéptido de BMP-2 con el sitio de escisión maduro en el extremo 3' fusionado al extremo 5' (TAL) del péptido maduro de BMP-12. La secuencia codificante del vector ligado resultante se muestra en la SEC ID N°: 27.

**Ejemplo 4****Actividad biológica de BMP-12 expresada**

Para medir la actividad biológica de las proteínas BMP-12 expresadas obtenidas en el Ejemplo 2 anterior, las proteínas se recuperan del cultivo celular y se purifican aislando las proteínas BMP-12 de otros materiales proteicos con los que se co-producen, así como de otros contaminantes. La proteína purificada puede ensayarse de acuerdo con el ensayo de rata descrito más adelante en el Ejemplo 5.

La purificación se realiza usando técnicas convencionales conocidas por los expertos en la materia.

El análisis de proteínas se realiza usando técnicas convencionales tales como SDS-PAGE acrilamida [Laemmli, Nature 227: 680 (1970)] teñida con Azul de Coomassie o plata [Oakley, y col. Anal: Biochem. 105: 361 (1980)] y por inmunotransferencia [Towbin, y col. Proc. Natl. Acad. Sci. USA 76: 4350 (1979)]

55

**Ejemplo 5****Ensayo de SAMPATH-REDDI modificado por ROSEN**

Se usa una versión modificada del ensayo de implante ectópico de rata descrito en Sampath y Reddi, Proc. Natl. Acad. Sci. USA, 80: 6591-6595 (1983) para evaluar la actividad de las proteínas BMP-12. Este ensayo modificado se denomina en el presente documento el ensayo de Sampath-Reddi modificado por Rosen. El ensayo se ha usado ampliamente para evaluar la actividad de inducción de cartílago y hueso de las BMP. La etapa de precipitación con etanol del procedimiento de Sampath-Reddi se reemplaza por diálisis (si la composición es una solución) o diafiltración (si la composición es una suspensión) de la fracción a ensayar frente a agua. La solución o suspensión después se equilibra a TFA al 0,1%. La solución resultante se añade a 20 mg de matriz de rata. Como control se utiliza una muestra de matriz de rata simulada no tratada con la proteína. Este material se congela y se liofiliza y el polvo resultante se encierra en cápsulas de gelatina nº 5. Las cápsulas se implantan por vía subcutánea en el área torácica abdominal de ratas Long Evans macho de 21-49 días de edad. Los implantes se retiran después de 10 días. Una sección de cada implante se fija y se procesa para el análisis histológico. Se tiñen secciones de metacrilato de glicol de 1 µm con Von Kossa y fucsina ácida para evaluar la cantidad de formación de tejido similar a tendón/ligamento inducido presente en cada implante.

BMP-12 se implantó en las ratas en dosis de 1, 5, 25 y 50 µm por implante durante 10 días. BMP-2 a una dosis de 5 µg se incluyó como control positivo. Para todas las dosis de BMP-12 ensayadas, no se observó formación de hueso o cartílago en los implantes después de diez días. En su lugar, los implantes se rellenaron con tejido parecido a tendón embrionario, que se reconoce fácilmente por la presencia de haces densos de fibroblastos orientados en el mismo plano y apretados entre sí. [Se describe tejido similar a tendón/ligamento, por ejemplo, en Ham y Cormack, Histology (JB Lippincott Co. (1979), págs. 367-369. Estos descubrimientos se reprodujeron en una segunda serie de ensayos en los que estaban presentes tejidos similares a tendón/ligamento en todos los implantes que contenían BMP-12. Por el contrario, los implantes de BMP-2, como era de esperar, mostraron formación de cartílago y hueso pero no contenían tejido similar a tendón/ligamento.

Las proteínas BMP-12 y proteínas relacionadas de la presente invención pueden ensayarse con respecto a la actividad en este ensayo.

**Ejemplo 6**

Usando procedimientos de acuerdo con los ejemplos anteriores, con modificaciones minoritarias dentro de la experiencia de la técnica, se expresaron proteína MP52 humana y el homólogo murino de la proteína BMP-13 y se ensayaron con respecto a la actividad inductora de tejido similar a tendón/ligamento. Todas las proteínas mostraron resultados comparables, similares a los descritos anteriormente para BMP-12 humana.

Las descripciones anteriores detallan realizaciones preferidas actualmente de la presente invención. Es de esperar que a los expertos en la materia se les ocurran numerosas modificaciones y variaciones en la práctica de la presente invención tras la consideración de estas descripciones. Se considera que estas modificaciones y variaciones se incluyen dentro de las reivindicaciones adjuntas.

## LISTADO DE SECUENCIAS

## (1) INFORMACIÓN GENERAL

(i) SOLICITANTE: GENETICS INSTITUTE, INC. PRESIDENT AND FELLOWS OF HARVARD COLLEGE

(ii) TÍTULO DE LA INVENCIÓN: COMPOSICIONES PARA INDUCIR TENDONES

(iii) NÚMERO DE SECUENCIAS: 35

(iv) DIRECCIÓN POSTAL:

(A) DESTINATARIO: GENETICS INSTITUTE, INC.

(B) CALLE: 87 CambridgePark Drive

(C) CIUDAD: Cambridge

(D) ESTADO: Massachusetts

(D) PAÍS: EEUU

(E) CÓDIGO POSTAL: 02140

(v) FORMA LEGIBLE POR ORDENADOR:

- (A) TIPO DE SOPORTE: Disquete
- (B) ORDENADOR: Compatible con PC IBM
- (C) SISTEMA OPERATIVO: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release # 1.0, Versión # 1.25

5 (vi) DATOS DE LA SOLICITUD ACTUAL:

- (A) NÚMERO DE SOLICITUD:
- (B) FECHA DE PRESENTACIÓN: Junto con la presente
- (C) CLASIFICACIÓN:

(vii) DATOS DE LA SOLICITUD ANTERIOR:

- 10 (A) NÚMERO DE SOLICITUD: US 08/164.103
- (B) FECHA DE PRESENTACIÓN: 07-DIC-1993
- (C) NÚMERO DE SOLICITUD: US 08/217.780
- (D) FECHA DE PRESENTACIÓN: 25-MAR-1994
- (E) NÚMERO DE SOLICITUD: US 08/333.576
- 15 (F) FECHA DE PRESENTACIÓN: 02-NOV-1994

(viii) INFORMACIÓN DEL MANDATARIO/AGENTE:

- (A) NOMBRE: Lazar, Steven R.
- (B) NÚMERO DE REGISTRO: 32.618
- (C) NÚMERO DE REFERENCIA/EXPEDIENTE: 5202D-PCT

20 (ix) INFORMACIÓN DE TELECOMUNICACIONES:

- (A) TELÉFONO: 617 498-8260
- (B) TELEFAX: 617 876-5851

(2) INFORMACIÓN PARA LA SEC ID Nº: 1

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 25 (A) LONGITUD: 926 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

30 (vi) FUENTE ORIGINAL:

- (A) ORGANISMO: Homo sapiens

(vii) FUENTE INMEDIATA:

- (B) CLON: v1-1

(ix) CARACTERÍSTICA:

- 35 (A) NOMBRE/CLAVE: mat\_peptide
- (B) LOCALIZACIÓN: 571..882

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 1..882

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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 CCG 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 CCG 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 CCG 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

GCGGCCGAGG ATCC                                                    926
    
```

(2) INFORMACIÓN PARA LA SEC ID Nº: 2

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 294 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 2

```

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

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

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

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

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

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

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

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

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

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

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

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



ACCGGGCGGC CCTGAACCCA AGCCAGGACA CCCTCCCCAA ACAAGGCAGG CTACAGCCCG 60  
 GACTGTGACC CCAAAGGAC AGCTTCCCGG AGGCAAGGCA CCCCAAAAG CAGGATCTGT 120  
 CCCCAGCTCC TTCCTGCTGA AGAAGGCCAG GGAGCCCGGG CCCCCACGAG AGCCCAAGGA 180  
 GCCGTTTCGC CCACCCCCA TCACACCCA CGAGTACATG CTCTCGCTGT ACAGGACGCT 240  
 GTCCGATGCT GACAGAAAGG GAGGCAACAG CAGCGTGAAG TTGGAGGCTG GCCTGGCCAA 300  
 CACCATCACC AGCTTTATG ACAAGGGCA AGATGACCGA GSTCCCGTGG TCAGGAAGCA 360  
 GAGGTACGTG TTTGACATTA GTGCCCTGGA GAAGGATGGG CTGCTGGGGG CCGAGCTCCG 420  
 GATCTTGC GG AAGAAGCCCT CGGACACGGC CAAGCCAGCG GCCCCCGGAG GCGGGCGGGC 480  
 TGCCCAGCTG AAGCTGTCCA GCTGCCCCAG CGGCCGGCAG CCGGCCTCCT TGCTGGATGT 540  
 GCGCTCCGTG CCAGGCCTGG ACGGATCTGG CTGGGAGGTG TTCGACATCT GGAAGCTCTT 600  
 CCGAACTTT AAGAATCGG CCCAGCTGTG CCTGGAGCTG GAGGCCTGGG AACGGGGCAG 660  
 GCGCGTGGAC CTCCTGTGCC TGGGCTTGA CCGCGCCGCC CCGCAGGTCC ACGAGAAGGC 720

CCTGTTCTTG GTGTTGGCC GCACCAAGAA ACGGGACCTG TTCTTTAATG AGATTAAGGC 780

CCGCTCTGGC CAGGACGATA AGACCGTGTG 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) INFORMACIÓN PARA LA SEC ID Nº: 4

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 120 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

5

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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) INFORMACIÓN PARA LA SEC ID Nº: 5

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 5 (A) LONGITUD: 128 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

10 (vi) FUENTE ORIGINAL:

- (A) ORGANISMO: Homo sapiens

(vii) FUENTE INMEDIATA:

- (B) CLON: fragmentos V1-1

(ix) CARACTERÍSTICA:

- 15 (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 28..102

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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) INFORMACIÓN PARA LA SEC ID Nº: 6

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 25 aminoácidos

(B) TIPO: aminoácido

5 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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) INFORMACIÓN PARA LA SEC ID Nº: 7

10 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 128 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

15 (ii) TIPO DE MOLÉCULA: ADN (genómico)

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: Homo sapiens

(vii) FUENTE INMEDIATA:

(B) CLON: VL-1

20 (ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 28..102

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 7

```

GGATCCTGGG ATGACTGGAT TATGSCG 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
    
```

25 (2) INFORMACIÓN PARA LA SEC ID Nº: 8

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 25 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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) INFORMACIÓN PARA LA SEC ID Nº: 9

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 3585 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: sencilla
- (D) TOPOLOGÍA: lineal

10 (ii) TIPO DE MOLÉCULA: ADN (genómico)

(vii) FUENTE INMEDIATA:

- (B) CLON: pALV1-781

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 9

```

CTAACTACCC AACTCAAAAA AAAAAAAAAA AAAAACCCCC TCTAACCCCC ATTGACGAAA      60
GGGCCTCGTG ATACGCCTAT TTTTATAGGT TAATGTCATG ATAATAATGG TTTCTTAGAC      120
GTCAGGTGGC ACTTTTCGGG GAAATGTGCG CGGAACCCCT AATTGTTTAT TTTTCTAAAT      180
    
```

ACATTCAAAT ATGTATCCGC TCATGAGACA ATAACCCTGA TAAATGCTTC AATAATATTG	240
AAAAAGGAAG AGTATGAGTA TTCAACATTT CCGTGTCCGC CTTATTCCTT TTTTTCGGC	300
ATTTTGCCTT CCTGTTTTTG CTCACCCAGA AACGCTGGTG AAAGTAAAAG ATGCTGAAGA	360
TCAGTTGGGT GCACGAGTGG GTTACATCGA ACTGGATCTC AACAGCGGTA AGATCCTTGA	420
GAGTTTTTCG CCCGAAGAAC GTTTTCCAAT GATGAGCACT TTTAAAGTTC TGCTATGTGG	480
CGCGGTATTA TCCCGTATG ACGCCGGGCA AGAGCAACTC GGTCCGCCA TACACTATTC	540
TCAGAATGAC TTGGTTGAGT ACTCACCAGT CACAGAAAAG CATCTTACGG ATGGCATGAC	600
AGTAAGAGAA TTATGCAGTG CTGCCATAAC CATGAGTGAT AACACTGCCG CCAACTTACT	660
TCTGACAACG ATCGGAGGAC CGAAGGAGCT AACCGCTTTT TTGCACAACA TGGGGGATCA	720
TGTAACTCGC CTTGATCGTT GGGAACCGGA GCTGAATGAA GCCATACCAA ACGACGAGCG	780
TGACACCACG ATGCCTGTAG CAATGGCAAC AACGTTGCGC AACTATTAA CTGGCGAACT	840
ACTTACTCTA GCTTCCCGGC AACAAATTAAT AGACTGGATG GAGGCGGATA AAGTTGCAGG	900
ACCACTTCTG CGCTCGGCC TTCCGGCTGG CTGGTTTATT GCTGATAAAT CTGGAGCCGG	960
TGAGCGTGGG TCTCCGGTA TCATTGCAGC ACTGGGGCCA GATGGTAAGC CCTCCCGTAT	1020
CGTAGTTATC TACACGACGG GGAGTCAGGC AACTATGGAT GAACGAAATA GACAGATCGC	1080
TGAGATAGGT GCCTCACTGA TTAAGCATTG GTAACGTCA GACCAAGTTT ACTCATATAT	1140
ACTTTAGATT GATTTAAAAC TTCATTTTTA ATTTAAAAGG ATCTAGGTGA AGATCCTTTT	1200
TGATAATCTC ATGACCAAAA TCCCTTAACG TGAGTTTTCG TTCCACTGAG CGTCAGACCC	1260
CGTAGAAAAG ATCAAAGGAT CTCTTGAGA TCCTTTTTT CTGCGCGTAA TCTGCTGCTT	1320
GCAAACAAA AAACCACCGC TACCAGCGGT GGTGTGTTG CCGGATCAAG AGCTACCAAC	1380
TCTTTTTCCG AAGGTAAC TGCTCAGCAG AGCGCAGATA CCAAATACTG TCCTTCTAGT	1440
GTAGCCGTAG TTAGGCCACC ACTTCAAGAA CTCTGTAGCA CCGCCTACAT ACCTCGCTCT	1500
GCTAATCCTG TTACCAGTGG CTGCTGCCAG TGGCGATAAG TCGTGTCTTA CCGGGTTGGA	1560
CTCAAGACGA TAGTTACCGG ATAAGGCGCA GCGGTCCGGC TGAACGGGGG GTTCGTGCAC	1620
ACAGCCCAGC TTGGAGCGAA CGACCTACAC CGAACTGAGA TACCTACAGC GTGAGCATTG	1680
AGAAAGCGCC ACGCTTCCCG AAGGGAGAAA GCGGCACAGG TATCCGGTAA GCGGCAGGGT	1740
CGGAACAGGA GAGCGCACGA GGGAGCTTCC AGGGGAAAC GCCTGGTATC TTTATAGTCC	1800
TGTCGGGTTT CGCCACCTCT GACTTGAGCG TCGATTTTTG TGATGCTCGT CAGGGGGGCG	1860
GAGCCTATGG AAAACGCCA GCAACGCGGC CTTTTTACGG TTCTGGCCT TTTGCTGGCC	1920
TTTTGCTCAC ATGTTCTTTC CTGCGTTATC CCCTGATTCT GTGGATAACC GTATPACCGC	1980
CTTTGAGTGA GCTGATACCG CTCGCCGAG CCGAACGACC GAGCGCAGCG AGTCAGTGAG	2040
CGAGGAAGCG GAAGAGCGCC CAATACGCAA ACCGCCTCTC CCGCGCGTT GSCCGATTCA	2100
TTAATGCAGA ATTGATCTCT CACCTACCAA ACAATGCCCC CCTGCAAAA ATAAATTCAT	2160
ATAAAAAACA TACAGATAAC CATCTGCGGT GATAAATTAT CTCTGGCGGT GTTGACATAA	2220

ATACCACTGG CGGTGATACT GAGCACATCA GCAGGACGCA CTGACCACCA TGAAGGTGAC	2280
GCTCTTAAAA ATTAAGCCCT GAAGAAGGGC AGCATTCAA GCAGAAGGCT TTGGGGTGTG	2340
TGATACGAAA CGAAGCATTG GCCGTAAGTG CGATCCGGA TTAGCTGCCA ATGTGCCAAT	2400
CGCGGGGGGT TTTCGTTTCA GACTACAACCT GCCACACACC ACCAAGCTA ACTGACAGGA	2460
GAATCCAGAT GGATGCACAA ACACGCCGCC GCGAACGTCG CGCAGAGAAA CAGGCTCAAT	2520
GGAAAGCAGC AAATCCCCTG TTGGTTGGGG TAAGCGCAA ACCAGTCCG AAAGATTTT	2580
TTAACTATAA ACGCTGATGG AAGCGTTTAT GCGGAAGAGG TAAAGCCCTT CCCGAGTAAC	2640
AAAAAACA CAGCATAAAT AACCCCGCTC TTACACATTC CAGCCCTGAA AAAGGGCATC	2700
AAATTAACC ACACCTATGG TGTATGCATT TATTTGCATA CATTCAATCA ATTGTTATCT	2760
AAGGAAATAC TTACATATGT CTCGTTGTTT TCGTAAACCA CTGCATGTAG ATTTTAAAGA	2820
GCTCGGCTGG GACGACTGGA TCATCGCGCC GCTGGACTAC GAGGCGTACC ACTGCGAGGG	2880
CCTTTGCGAC TTCCCTTTGC GTTCGCACCT CGAGCCCAAC AACCATGCCA TCATTAGAC	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 CTTTTTCTT GTGAGCAGTA	3180
AGCITGGCAC TGGCCGTCGT TTACAACGT CGTGACTGGG AAAACCCTGG CGTTACCCAA	3240
CTTAATCGCC TTGCAGCACA TCCCCCTTTC GCCAGCTGGC GTAATAGCGA AGAGGCCCGC	3300
ACCGATCGCC CTTCCCAACA GTTGCAGCAGC CTGAATGGCG AATGGCGCCT GATGCGGTAT	3360
TTTCTCCTTA CGCATCTGTG CGGTATTTCA CACCGCATAT ATGGTGCCT CTCAGTACAA	3420
TCTGCTCTGA TGCCGCATAG TTAAGCCAGC CCCGACACCC GCCAACACCC GCTGACGCGC	3480
CCTGACGGGC TTGTCTGCTC CCGGCATCCG CTTACAGACA AGCTGTGACC GTCTCCGGGA	3540
GCTGCATGTG TCAGAGGTTT TCACCGTCAT CACCGAAACG CGCGA	3585

(2) INFORMACIÓN PARA LA SEC ID Nº: 10

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 272 pares de bases

5 (B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(vi) FUENTE ORIGINAL:

10 (A) ORGANISMO: ratón

(vii) FUENTE INMEDIATA:

(B) CLON: mV1

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 28..243

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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) INFORMACIÓN PARA LA SEC ID Nº: 11

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 72 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

10 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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) INFORMACIÓN PARA LA SEC ID Nº: 12

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 272 pares de bases

15 (B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(vi) FUENTE ORIGINAL:

(A) ORGANISMO: ratón

(vii) FUENTE INMEDIATA:

(B) CLON: mV2

(ix) CARACTERÍSTICA:

5 (A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 28..243

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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) INFORMACIÓN PARA LA SEC ID Nº: 13

10 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 72 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

15 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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) INFORMACIÓN PARA LA SEC ID Nº: 14

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 272 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

5 (vi) FUENTE ORIGINAL:

(A) ORGANISMO: ratón

(vii) FUENTE INMEDIATA:

(B) CLON: mV9

(ix) CARACTERÍSTICA:

10 (A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 28..243

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 14

GGATCCAAGG AGCTCGGCTG GGACGAC TGG ATC ATC GCA CCT CTT GAG TAT	51
Trp Ile Ile Ala Pro Leu Glu Tyr	
1 5	
GAG GCC TTC CAC TGC GAA GGA CTG TGT GAG TTC CCC TTG CGC TCC CAC	99
Glu Ala Phe His Cys Glu Gly Leu Cys Glu Phe Pro Leu Arg Ser His	
10 15 20	
TTG GAG CCC ACA AAC CAC GCA GTC ATT CAG ACC CTA ATG AAC TCT ATG	147
Leu Glu Pro Thr Asn His Ala Val Ile Gln Thr Leu Met Asn Ser Met	
25 30 35 40	
GAC CCT GAA TCC ACA CCA CCC ACT TGT TGT GTG CCT ACA CGG CTG AGT	195
Asp Pro Glu Ser Thr Pro Pro Thr Cys Cys Val Pro Thr Arg Leu Ser	
45 50 55	
CCT ATT AGC ATC CTC TTC ATC GAC TCT GCC AAC AAC GTG GTG TAT AAA	243
Pro Ile Ser Ile Leu Phe Ile Asp Ser Ala Asn Asn Val Val Tyr Lys	
60 65 70	
CAATACGAGG ACATGGTGGT GGGGAATTC	272

(2) INFORMACIÓN PARA LA SEC ID Nº: 15

15 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 72 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

20 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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) INFORMACIÓN PARA LA SEC ID Nº: 16

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 7 aminoácidos

5 (B) TIPO: aminoácido

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: péptido

(vi) FUENTE ORIGINAL:

10 (A) ORGANISMO: secuencia consenso de BMP/TGF-beta

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 16

Trp Xaa Asp Trp Ile Xaa Ala  
 1 5

(2) INFORMACIÓN PARA LA SEC ID Nº: 17

(i) CARACTERÍSTICAS DE LA SECUENCIA:

15 (A) LONGITUD: 27 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

20 (vii) FUENTE INMEDIATA:

(B) CLON: oligonucleótido nº 1

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 17

CGGATCCTGG VANGAYTGGA THRTNGC 27

(2) INFORMACIÓN PARA LA SEC ID Nº: 18

25 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 6 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: péptido

(vii) FUENTE INMEDIATA:

(B) CLON: secuencia consenso de BMP/TGF-beta

5 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID N°: 18

His Ala Ile Xaa Gln Thr  
1 5

(2) INFORMACIÓN PARA LA SEC ID N°: 19

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 28 pares de bases

10 (B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(vii) FUENTE INMEDIATA:

15 (B) CLON: oligonucleótido n° 2

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID N°: 19

TTTCTAGAAR NGTYTGNACD ATNGCRTG 28

(2) INFORMACIÓN PARA LA SEC ID N°: 20

(i) CARACTERÍSTICAS DE LA SECUENCIA:

20 (A) LONGITUD: 40 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

25 (vii) FUENTE INMEDIATA:

(B) CLON: oligonucleótido n° 3

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID N°: 20

CCACTGCGAG GGCCTTTGCG ACTTCCCTTT GCGTTCGCAC 40

(2) INFORMACIÓN PARA LA SEC ID N°: 21

30 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 29 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

35 (ii) TIPO DE MOLÉCULA: ADN (genómico)

(vii) FUENTE INMEDIATA:

(A) BIBLIOTECA: oligonucleótido nº 4

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 21

TGCGGATCCA GCCGCTGCAG CCGCAAGCC 29

5 (2) INFORMACIÓN PARA LA SEC ID Nº: 22

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 29 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

10 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(vii) FUENTE INMEDIATA:

(B) CLON: oligonucleótido nº 5

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 22

15 GACTCTAGAC TACCTGCAGC CGCAGGCCT 29

(2) INFORMACIÓN PARA LA SEC ID Nº: 23

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 28 pares de bases

(B) TIPO: ácido nucleico

20 (C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(vii) FUENTE INMEDIATA:

(A) BIBLIOTECA: oligonucleótido nº 6

25 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 23

GCGGATCCAA GGAGCTCGGC TGGGACGA 28

(2) INFORMACIÓN PARA LA SEC ID Nº: 24

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 28 pares de bases

30 (B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(vii) FUENTE INMEDIATA:

35 (B) CLON: oligonucleótido nº 7

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 24

GGAATTCCCC ACCACCATGT CCTCGTAT 28

(2) INFORMACIÓN PARA LA SEC ID Nº: 25

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 5 (A) LONGITUD: 1171 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

10 (vii) FUENTE INMEDIATA:

(B) CLON: proteína VL-1 humana

(ix) CARACTERÍSTICA:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 2..964

15 (ix) CARACTERÍSTICA:

- (A) NOMBRE/CLAVE: mat\_peptide
- (B) LOCALIZACIÓN: 605..964

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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 Leu Arg Leu Phe Arg Gln Ala Pro Ser Ala Pro Trp Gly Pro Pro Ala -170 -165 -160 -155	142
GGG CCG CTC CAC GTG CAG CTC TTC CCT TGC CTT TCG CCC CTA CTG CTG Gly Pro Leu His Val Gln Leu Phe Pro Cys Leu Ser Pro Leu Leu Leu -150 -145 -140	190
GAC GCG CGG ACC CTG GAC CCG CAG GGG GCG CCG CCG GCC GGC TGG GAA Asp Ala Arg Thr Leu Asp Pro Gln Gly Ala Pro Pro Ala Gly Trp Glu -135 -130 -125	238
GTC TTC GAC GTG TGG CAG GGC CTG CCG CAC CAG CCC TGG AAG CAG CTG Val Phe Asp Val Trp Gln Gly Leu Arg His Gln Pro Trp Lys Gln Leu -120 -115 -110	286
TGC TTG GAG CTG CCG GCC GCA TGG GGC GAG CTG GAC GCC GGG GAG GCC Cys Leu Glu Leu Arg Ala Trp Gly Glu Leu Asp Ala Gly Glu Ala -105 -100 -95	334
GAG GCG CGC GCG CGG GGA CCC CAG CAA CCG CCG CCC CCG GAC CTG CCG Glu Ala Arg Ala Arg Gly Pro Gln Gln Pro Pro Pro Asp Leu Arg -90 -85 -80 -75	382
AGT CTG GGC TTC GGC CGS AGG GTG CCG CCT CCC CAG GAG CCG GCC CTG Ser Leu Gly Phe Gly Arg Arg Val Arg Pro Pro Gln Glu Arg Ala Leu -70 -65 -60	430
CTG GTG GTA TTC ACC AGA TCC CAG CGC AAG AAC CTG TTC GCA GAG ATG Leu Val Val Phe Thr Arg Ser Gln Arg Lys Asn Leu Phe Ala Glu Met -55 -50 -45	478
CGC GAG CAG CTG GGC TCG GCC GAG GCT GCG GGC CCG GGC GCG GCC GCC Arg Glu Gln Leu Gly Ser Ala Glu Ala Ala Gly Pro Gly Ala Gly Ala -40 -35 -30	526
GAG GGG TCG TGG CCG CCG CCG TCG GGC GCC CCG GAT GCC AGG CCT TGG Glu Gly Ser Trp Pro Pro Pro Ser Gly Ala Pro Asp Ala Arg Pro Trp -25 -20 -15	574
CTG CCC TCG CCC GGC CGC CGG CCG CCG CGC ACG GCC TTC GCC AGT CGC Leu Pro Ser Pro Gly Arg Arg Arg Arg Arg Thr Ala Phe Ala Ser Arg -10 -5 1 5	622
CAT GGC AAG CGG CAC GGC AAG AAG TCC AGG CTA CGC TGC AGC AAG AAG His Gly Lys Arg His Gly Lys Lys Ser Arg Leu Arg Cys Ser Lys Lys 10 15 20	670
CCC CTG CAC GTG AAC TTC AAG GAG CTG GGC TGG GAC GAC TGG ATT ATC Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile 25 30 35	718
GCG CCC CTG GAG TAC GAG GCC TAT CAC TGC GAG GGT GTA TGC GAC TTC Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe 40 45 50	766
CCG CTG CGC TCG CAC CTG GAG CCC ACC AAC CAC GCC ATC ATC CAG ACG Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr 55 60 65 70	814
CTG ATG AAC TCC ATG GAC CCC GGC TCC ACC CCG CCC AGC TGC TGC GTG Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val 75 80 85	862
CCC ACC AAA TTG ACT CCC ATC AGC ATT CTA TAC ATC GAC GCG GCC AAT Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn 90 95 100	910

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
CCCACTAAT TCTACCTGCC AGCCAGAGG GAAATAT	1171

(2) INFORMACIÓN PARA LA SEC ID N°: 26

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 321 aminoácidos

5

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID N°: 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 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) INFORMACIÓN PARA LA SEC ID Nº: 27

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1233 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

5 (ii) TIPO DE MOLÉCULA: ADN (genómico)

(vii) FUENTE INMEDIATA:

(B) CLON: ADN que codifica propéptido de BMP2/péptido maduro de BMP-12

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: CDS

10 (B) LOCALIZACIÓN: 1..1233

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: mat\_peptide

(B) LOCALIZACIÓN: 847..1233

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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	

15

GAC CTG TAT CGC AGG CAC TCA GGT CAG CCG GGC TCA CCC GCC CCA GAC Asp Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp -200 -195 -190	288
CAC CGG TTG GAG AGG GCA GCC AGC CGA GCC AAC ACT GTG CGC AGC TTC His Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe -185 -180 -175	336
CAC CAT GAA GAA TCT TTG GAA GAA CTA CCA GAA ACG AGT GGG AAA ACA His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr -170 -165 -160 -155	384
ACC CGG AGA TTC TTC TTT AAT TTA AGT TCT ATC CCC ACG GAG GAG TTT Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe -150 -145 -140	432
ATC ACC TCA GCA GAG CTT CAG GTT TTC CGA GAA CAG ATG CAA GAT GCT Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala -135 -130 -125	480
TTA GGA AAC AAT AGC AGT TTC CAT CAC CGA ATT AAT ATT TAT GAA ATC Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile -120 -115 -110	528
ATA AAA CCT GCA ACA GCC AAC TCG AAA TTC CCC GTG ACC AGA CTT TTG Ile Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu -105 -100 -95	576
GAC ACC AGG TTG GTG AAT CAG AAT GCA AGC AGG TGG GAA AGT TTT GAT Asp Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp -90 -85 -80 -75	624
GTC ACC CCC GCT GTG ATG CCG TGG ACT GCA CAG GGA CAC GCC AAC CAT Val Thr Pro Ala Val Met Arg Trp Thr Ala Gln Gly His Ala Asn His -70 -65 -60	672
GGA TTC GTG GTG GAA GTG GCC CAC TTG GAG GAG AAA CAA GGT GTC TCC Gly Phe Val Val Glu Val Ala His Leu Glu Glu Lys Gln Gly Val Ser -55 -50 -45	720
AAG AGA CAT GTT AGG ATA AGC AGG TCT TTG CAC CAA GAT GAA CAC AGC Lys Arg His Val Arg Ile Ser Arg Ser Leu His Gln Asp Glu His Ser -40 -35 -30	768
TGG TCA CAG ATA AGG CCA TTG CTA GTA ACT TTT GGC CAT GAT GGA AAA Trp Ser Gln Ile Arg Pro Leu Leu Val Thr Phe Gly His Asp Gly Lys -25 -20 -15	816
GGG CAT CCT CTC CAC AAA AGA GAA AAA CGT ACG GCG TTG GCC GGG ACG Gly His Pro Leu His Lys Arg Glu Lys Arg Thr Ala Leu Ala Gly Thr -10 -5 1 5	864
CGG ACA GCG CAG GGC AGC GGC GGG GGC GCG GGC CGG GGC CAC GGG CGC Arg Thr Ala Gln Gly Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg 10 15 20	912
AGG GGC CGG AGC CGC TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG Arg Gly Arg Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys 25 30 35	960
GAG CTC GGC TGG GAC GAC TGG ATC ATC GCG CCG CTG GAC TAC GAG GCG Glu Leu Gly Trp Asp Asp Trp Ile Ile Ala Pro Leu Asp Tyr Glu Ala 40 45 50	1008
TAC CAC TGC GAC GGC CTT TGC GAC TTC CCT TTG CGT TCG CAC CTC GAG Tyr His Cys Glu Gly Leu Cys Asp Phe Pro Leu Arg Ser His Leu Glu 55 60 65 70	1056

```

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) INFORMACIÓN PARA LA SEC ID Nº: 28

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 411 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 28

```

Met Val Ala Gly Thr Arg Cys Leu Leu Ala Leu Leu Leu Pro Gln Val
-282      -280                                -275                                -270

Leu Leu Gly Gly Ala Ala Gly Leu Val Pro Glu Leu Gly Arg Arg Lys
-265      -260                                -255

Phe Ala Ala Ala Ser Ser Gly Arg Pro Ser Ser Gln Pro Ser Asp Glu
-250      -245                                -240                                -235

Val Leu Ser Glu Phe Glu Leu Arg Leu Leu Ser Met Phe Gly Leu Lys
-230      -225                                -220

Gln Arg Pro Thr Pro Ser Arg Asp Ala Val Val Pro Pro Tyr Met Leu
-215      -210                                -205

Asp Leu Tyr Arg Arg His Ser Gly Gln Pro Gly Ser Pro Ala Pro Asp
-200      -195                                -190

His Arg Leu Glu Arg Ala Ala Ser Arg Ala Asn Thr Val Arg Ser Phe
-185      -180                                -175

His His Glu Glu Ser Leu Glu Glu Leu Pro Glu Thr Ser Gly Lys Thr
-170      -165                                -160                                -155

Thr Arg Arg Phe Phe Phe Asn Leu Ser Ser Ile Pro Thr Glu Glu Phe
-150      -145                                -140

Ile Thr Ser Ala Glu Leu Gln Val Phe Arg Glu Gln Met Gln Asp Ala
-135      -130                                -125

Leu Gly Asn Asn Ser Ser Phe His His Arg Ile Asn Ile Tyr Glu Ile
-120      -115                                -110

Ile Lys Pro Ala Thr Ala Asn Ser Lys Phe Pro Val Thr Arg Leu Leu
-105      -100                                -95

Asp Thr Arg Leu Val Asn Gln Asn Ala Ser Arg Trp Glu Ser Phe Asp
-90      -85                                -80                                -75
    
```

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) INFORMACIÓN PARA LA SEC ID Nº: 29

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1203 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE CADENA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(vii) FUENTE INMEDIATA:

- (B) CLON: MV1 murino

(ix) CARACTERÍSTICA:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 2..721

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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	GCC	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 Gly Gly Gly Thr Ala X06 Glu Glu Arg Ala Leu Leu Val Ile Ser Ser 35 40 45	142
CGT ACG CAA AGG AAA GAG AGT CTG TTC CGG GAG ATC CGA GCC CAG GCC Arg Thr Gln Arg Lys Glu Ser Leu Phe Arg Glu Ile Arg Ala Gln Ala 50 55 60	190
CGT GCT CTC CGG GCC GCT GCA GAG CCG CCA CCG GAT CCA GGA CCA GGC Arg Ala Leu Arg Ala Ala Ala Glu Pro Pro Pro Asp Pro Gly Pro Gly 65 70 75	238
GCT GGG TCA CGC AAA GCC AAC CTG GGC GGT CGC AGG CGG CAG CGG ACT Ala Gly Ser Arg Lys Ala Asn Leu Gly Gly Arg Arg Arg Gln Arg Thr 80 85 90 95	286
GCG CTG GCT GGG ACT CGG GGA GNG NAG GGA AGC GGT GGT GGC GGC GGT Ala Leu Ala Gly Thr Arg Gly X07 X08 Gly Ser Gly Gly Gly Gly Gly 100 105 110	334
GGC GGT GGC GCA Gly Gly Ala 115 120 125	382
GGC AGG GGC CAC GGG CGC AGA GGC CGG AGC CGC TGC GGT CGC AAG TCA Gly Arg Gly His Gly Arg Arg Gly Arg Ser Arg Cys Gly Arg Lys Ser 130 135 140	430
CTG CAC GTG GAC TTT AAG GAG CTG GGC TGG GAC GAC TGG ATC ATC GCG Leu His Val Asp Phe Lys Glu Leu Gly Trp Asp Trp Ile Ile Ala 145 150 155	478
CCA TTA GAC TAC GAG GCA TAC CAC TGC GAG GGC GTT TGC GAC TTT CCT Pro Leu Asp Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe Pro 160 165 170 175	526
CTG CGC TCG CAC CTG GAG CCT ACC AAC CAC GCC ATC ATT CAG ACG CTG Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr Leu 180 185 190	574
CTC AAC TCC ATG GCG CCC GAC GCT GCG CCA GCC TCC TGC TGC GTG CCC Leu Asn Ser Met Ala Pro Asp Ala Ala Pro Ala Ser Cys Cys Val Pro 195 200 205	622
GCA AGG CTC AGT CCC ATC AGC ATT CTC TAC ATC GAT GCC GCC AAC AAC Ala Arg Leu Ser Pro Ile Ser Ile Leu Tyr Ile Asp Ala Ala Asn Asn 210 215 220	670
GTG GTC TAC AAG CAG TAC GAA GAC ATG GTG GTG GAG GCC TGC GGC TGC Val Val Tyr Lys Gln Tyr Glu Asp Met Val Val Glu Ala Cys Gly Cys 225 230 235	718
AGG TAGCATGCGG TCTGGGGAGG GTCTGGCCGC CCAGGACCCT AGCTCAAGAG Arg 240	771
CAGGTGTCAT CAGGCCCGAG GGACGGCGGA CTATGGCCTC TGCCAGCACA GAGGAGAGCA	831
CACAGTTAAC ACTCACATTT ACACACTCCT TCACTCACGC ACATGTTTAC CGTGGACGGC	891
AGCCGCTAAA AGCCTTGCTT APTTGCTACC ATTGATACAA ACCTCTGTCC TTTTCGGGAG	951
AGGGAAGGGC ATCTGTGTTT ATGTTGCAGT AATTGGCACT AAATCCAAGT AGAARTGGGT	1011
TAGCATTGGA TTCTCCTTTT AGTTGGAGGC GGTGTGGCTG GATTCTGAC GTTGCATATG	1071
GAGTGCACCTG CAGGGCTGGG ATACCCAGAT TCTCTGGAGT GGGCATTGGG AACCTTCAAA	1131
AGTAAGGAGC CACTGGGGCT TGGGAGGGAG CACCCGGTTC CTAACAAGT CTGATGTGTA	1191
CTGCTCAGTT TG	1203

(2) INFORMACIÓN PARA LA SEC ID Nº: 30

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 240 aminoácidos

(B) TIPO: aminoácido

5

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 30

```

Lys Phe Cys Leu Val Leu X01 X02 Val Thr Ala Ser Glu Ser X03 X04
 1           5           10
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 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) INFORMACIÓN PARA LA SEC ID Nº: 31

10

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1046 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(iii) HIPOTÉTICA: NO

(iv) ANTI-SENTIDO: NO

(vii) FUENTE INMEDIATA:

5 (B) CLON: MV2 murino

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 2..790

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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		

10

GAG GGC GTG TGC GAC TTT CCG CTG CGC TCG CAC CTT GAG CCC ACT AAC Glu Gly Val Cys Asp Phe Pro Leu Arg Ser His Leu Glu Pro Thr Asn 195 200 205	622
CAT GCC ATC ATT CAG ACG CTG ATG AAC TCC ATG GAC CCG GGC TCC ACC His Ala Ile Ile Gln Thr Leu Met Asn Ser Met Asp Pro Gly Ser Thr 210 215 220	670
CCG CCT AGC TGC TGC GTT CCC ACC AAA CTG ACT CCC ATT AGC ATC CTG Pro Pro Ser Cys Cys Val Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu 225 230 235	718
TAC ATC GAC GCG GGC AAT AAT GTN GTC TAC AAG CAG TAT GAG GAC ATG Tyr Ile Asp Ala Gly Asn Asn X02 Val Tyr Lys Gln Tyr Glu Asp Met 240 245 250 255	766
GTG GTG GAG TCC TGC GGC TGT AGG TAGCGGTGCT GTCCCGCCAC CTGGGCCAGG Val Val Glu Ser Cys Gly Cys Arg 260	820
GACCATGGAG GGAGGCCTGA CTGCCGAGAA AGGAGCAGGA GCTGGCCTTG GAAGAGGCCA	880
CAGGTGGGGG ACAGCCTGAA AGTAGGAGCA CAGTAAGAAG CAGCCCAGCC TTCCAGAAC	940
CTTCCAATCC CCCAACCAG AAGCAGCTAA GGGGTTTCAC AACTTTGGC CTTGCCAGCC	1000
TGGAAGACT AGACAAGAGG GATTCTTCTC TTTTATTAT GGCTTG	1046

(2) INFORMACIÓN PARA LA SEC ID Nº: 32

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 263 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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) INFORMACIÓN PARA LA SEC ID Nº: 33

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1345 pares de bases

5 (B) TIPO: ácido nucleico

(C) TIPO DE CADENA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genómico)

(iii) HIPOTÉTICA: NO

10 (iv) ANTI-SENTIDO: NO

(vii) FUENTE INMEDIATA:

(B) CLON: V1-1 HUMANO

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: CDS

15 (B) LOCALIZACIÓN: 138..1301

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: mat\_peptide

(B) LOCALIZACIÓN: 990..1301

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC ID Nº: 33



CAA CAG GCC TGG CTC CCA CAT CGA AGA CAG CTG GGC CAT TTG CTG TTA Gln Gln Ala Trp Leu Pro His Arg Arg Gln Leu Gly His Leu Leu Leu -255 -250 -245	266
GGA GGC CCC GCG CTG ACA GTG TGC AGG ATT TGC TCT TAC ACA GCT CTT Gly Gly Pro Ala Leu Thr Val Cys Arg Ile Cys Ser Tyr Thr Ala Leu -240 -235 -230	314
TCT CTC TGT CCC TGC CGG TCC CCC GCA GAC GAA TCG GCA GCC GAA ACA Ser Leu Cys Pro Cys Arg Ser Pro Ala Asp Glu Ser Ala Ala Glu Thr -225 -220 -215	362
GGC CAG AGC TTC CTG TTC GAC GTG TCC AGC CTT AAC GAC GCA GAC GAG Gly Gln Ser Phe Leu Phe Asp Val Ser Ser Leu Asn Asp Ala Asp Glu -205 -200 -195	410
GTG GTG GGT GCC GAG CTG CGC GTG CTG CGC CGG GGA TCT CCA GAG TCG Val Val Gly Ala Glu Leu Arg Val Leu Arg Arg Gly Ser Pro Glu Ser -190 -185 -180	458
GGC CCA GGC AGC TGG ACT TCT CCG CCG TTG CTG CTG CTG TCC ACG TGC Gly Pro Gly Ser Trp Thr Ser Pro Pro Leu Leu Leu Leu Ser Thr Cys -175 -170 -165	506
CCG GGC GCC GCC CGA GCG CCA CGC CTG CTG TAC TCG CCG GCA GCT GAG Pro Gly Ala Ala Arg Ala Pro Arg Leu Leu Tyr Ser Arg Ala Ala Glu -160 -155 -150	554
CCC CTA GTC GGT CAG CGC TGG GAG GCG TTC GAC GTG GCG GAC GCC ATG Pro Leu Val Gly Gln Arg Trp Glu Ala Phe Asp Val Ala Asp Ala Met -145 -140 -135 -130	602
AGG CGC CAC CGT CGT GAA CCG CGC CCC CCC CGC GCG TTC TGC CTC TTG Arg Arg His Arg Arg Glu Pro Arg Pro Pro Arg Ala Phe Cys Leu Leu -125 -120 -115	650
CTG CGC GCA GTG GCA GGC CCG GTG CCG AGC CCG TTG GCA CTG CGG CGA Leu Arg Ala Val Ala Gly Pro Val Pro Ser Pro Leu Ala Leu Arg Arg -110 -105 -100	698
CTG GGC TTC GGC TGG CCG GGC GGA GGG GGC TCT GCG GCA GAG GAG CGC Leu Gly Phe Gly Trp Pro Gly Gly Gly Ser Ala Ala Glu Glu Arg -95 -90 -85	746
GCG GTG CTA GTC GTC TCC TCC CGC ACG CAG AGG AAA GAG AGC TTA TTC Ala Val Leu Val Val Ser Ser Arg Thr Gln Arg Lys Glu Ser Leu Phe -80 -75 -70	794
CGG GAG ATC CGC GCC CAG GCC CGC GCG CTC GGG GCC GCT CTG GCC TCA Arg Glu Ile Arg Ala Gln Ala Arg Ala Leu Gly Ala Ala Leu Ala Ser -65 -60 -55 -50	842
GAG CCG CTG CCC GAC CCA GGA ACC GGC ACC GCG TCG CCA AGG GCA GTC Glu Pro Leu Pro Asp Pro Gly Thr Gly Thr Ala Ser Pro Arg Ala Val -45 -40 -35	890
ATT GGC GGC CGC AGA CGG AGG AGG ACG GCG TTG GCC GGG ACG CGG ACA Ile Gly Gly Arg Arg Arg Arg Thr Ala Leu Ala Gly Thr Arg Thr -30 -25 -20	938
GCG CAG GGC AGC GGC GGG GGC GCG GGC CCG GGC CAC GGG CGC AGG GGC Ala Gln Gly Ser Gly Gly Gly Ala Gly Arg Gly His Gly Arg Arg Gly -15 -10 -5	986
CGG AGC CGC TGC AGC CGC AAG CCG TTG CAC GTG GAC TTC AAG GAG CTC Arg Ser Arg Cys Ser Arg Lys Pro Leu His Val Asp Phe Lys Glu Leu 1 5 10 15	1034

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) INFORMACIÓN PARA LA SEC ID Nº: 34

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 388 aminoácidos

5 (B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC 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	



CTG CGG CTC TTT CGC CAG GCG CCC TCA GCG CCC TGG GGG CCA CCA GCC Leu Arg Leu Phe Arg Gln Ala Pro Ser Ala Pro Trp Gly Pro Pro Ala -170 -165 -160 -155	142
GGG CCG CTC CAC GTG CAG CTC TTC CCT TGC CTT TCG CCC CTA CTG CTG Gly Pro Leu His Val Gln Leu Phe Phe Cys Leu Ser Pro Leu Leu Leu -150 -145 -140	190
GAC GCG CCG ACC CTG GAC CCG CAG GGG GCG CCG CCG GCC GGC TGG GAA Asp Ala Arg Thr Leu Asp Pro Gln Gly Ala Pro Pro Ala Gly Trp Glu -135 -130 -125	238
GTC TTC GAC GTG TGG CAG GGC CTC CCG CAC CAG CCC TGG AAG CAG CTG Val Phe Asp Val Trp Gln Gly Leu Arg His Gln Pro Trp Lys Gln Leu -120 -115 -110	286
TGC TTG GAG CTG CCG GCC GCA TGG GGC GAG CTG GAC GCC GGG GAG CCC Cys Leu Glu Leu Arg Ala Ala Trp Gly Glu Leu Asp Ala Gly Glu Ala -105 -100 -95	334
GAG GCG CCG GCG CCG GGA CCC CAG CAA CCG CCG CCC CCG GAC CTG CCG Glu Ala Arg Ala Arg Gly Pro Gln Gln Pro Pro Pro Pro Asp Leu Arg -90 -85 -80 -75	382
AGT CTG GGC TTC GGC CCG AGG GTG CCG CCT CCC CAG GAG CCG GCC CTG Ser Leu Gly Phe Gly Arg Arg Val Arg Pro Pro Gln Glu Arg Ala Leu -70 -65 -60	430
CTG GTG GTA TTC ACC AGA TCC CAG CCG AAG AAC CTG TTC GCA GAG ATG Leu Val Val Phe Thr Arg Ser Gln Arg Lys Asn Leu Phe Ala Glu Met -55 -50 -45	478
CGC GAG CAG CTG GGC TCG GCC GAG GCT GCG GGC CCG GGC GCG GGC GCC Arg Glu Gln Leu Gly Ser Ala Glu Ala Ala Gly Pro Gly Ala Gly Ala -40 -35 -30	526
GAG GGG TCG TGG CCG CCC CCG TCG GGC GGC CCG GAT GCC AGC CCT TGG Glu Gly Ser Trp Pro Pro Pro Ser Gly Ala Pro Asp Ala Arg Pro Trp -25 -20 -15	574
CTG CCC TCG CCC GGC CCG CCG CCG CCG CCG ACC GCC TTC GGC ACT CCC Leu Pro Ser Pro Gly Arg Arg Arg Arg Arg Thr Ala Phe Ala Ser Arg -10 -5 1 5	622
CAT GCC AAG CCG CAC GGC AAG AAG TCC AGG CTA CCG TGC AGC AAG AAG His Gly Lys Arg His Gly Lys Lys Ser Arg Leu Arg Cys Ser Lys Lys 10 15 20	670
CCC CTG CAC GTG AAC TTC AAG GAG CTG GGC TGG GAC GAC TGG ATT ATC Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile 25 30 35	718
GCG CCC CTG GAG TAC GAG GCC TAT CAC TGC GAG GGT GTA TGC GAC TTC Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe 40 45 50	766
CCG CTG CCG TCG CAC CTG GAG CCC ACC AAC CAC GCC ATC ATC CAG ACC Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr 55 60 65 70	814
CTG ATG AAC TCC ATG GAC CCC GGC TCC ACC CCG CCC AGC TGC TGC GTG Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val 75 80 85	862
CCC ACC AAA TTG ACT CCC ATC AGC ATT CTA TAC ATC GAC GCG GGC AAT Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn 90 95 100	910

CTG CGG CTC TTT CGC CAG GCG CCC TCA GCG CCC TGG GGG CCA CCA GCC Leu Arg Leu Phe Arg Gln Ala Pro Ser Ala Pro Trp Gly Pro Pro Ala -170 -165 -160 -155	142
GGG CCG CTC CAC GTG CAG CTC TTC CCT TGC CTT TCG CCC CTA CTG CTG Gly Pro Leu His Val Gln Leu Phe Pro Cys Leu Ser Pro Leu Leu Leu -150 -145 -140	190
GAC GCG CGG ACC CTG GAC CCG CAG GGG GCG CCG CCG GCC GGC TGG GAA Asp Ala Arg Thr Leu Asp Pro Gln Gly Ala Pro Pro Ala Gly Trp Glu -135 -130 -125	238
GTC TTC GAC GTG TGG CAG GGC CTG CGC CAC CAG CCC TGG AAG CAG CTG Val Phe Asp Val Trp Gln Gly Leu Arg His Gln Pro Trp Lys Gln Leu -120 -115 -110	286
TGC TTG GAG CTG CCG GCC GCA TGG GGC GAG CTG GAC GCC GGG GAG GCC Cys Leu Glu Leu Arg Ala Ala Trp Gly Glu Leu Asp Ala Gly Glu Ala -105 -100 -95	334
GAG GCG CGC GCG CCG GGA CCC CAG CAA CCG CCG CCC CCG GAC CTG CCG Glu Ala Arg Ala Arg Gly Pro Gln Gln Pro Pro Pro Pro Asp Leu Arg -90 -85 -80 -75	382
AGT CTG GGC TTC GGC CCG AGG GTG CCG CCT CCC CAG GAG CCG GCC CTG Ser Leu Gly Phe Gly Arg Arg Val Arg Pro Pro Gln Glu Arg Ala Leu -70 -65 -60	430
CTG GTG GTA TTC ACC AGA TCC CAG CGC AAG AAC CTG TTC GCA GAG ATG Leu Val Val Phe Thr Arg Ser Gln Arg Lys Asn Leu Phe Ala Glu Met -55 -50 -45	478
CGC GAG CAG CTG GGC TCG GCC GAG GCT GCG GGC CCG GGC GCG GGC GCC Arg Glu Gln Leu Gly Ser Ala Glu Ala Ala Gly Pro Gly Ala Gly Ala -40 -35 -30	526
GAG GGG TCG TGG CCG CCG CCG TCG GGC GCC CCG GAT GCC AGG CCT TGG Glu Gly Ser Trp Pro Pro Pro Ser Gly Ala Pro Asp Ala Arg Pro Trp -25 -20 -15	574
CTG CCC TCG CCC GGC CGC CCG CCG CCG CCG CCG ACG GCC TTC GCC AGT CGC Leu Pro Ser Pro Gly Arg Arg Arg Arg Arg Thr Ala Phe Ala Ser Arg -10 -5 1 5	622
CAT GGC AAG CCG CAC GGC AAG AAG TCC AGG CTA CGC TGC AGC AAG AAG His Gly Lys Arg His Gly Lys Lys Ser Arg Leu Arg Cys Ser Lys Lys 10 15 20	670
CCC CTG CAC GTG AAC TTC AAG GAG CTG GGC TGG GAC GAC TGG ATT ATC Pro Leu His Val Asn Phe Lys Glu Leu Gly Trp Asp Asp Trp Ile Ile 25 30 35	718
GCG CCC CTG GAG TAC GAG GCC TAT CAC TGC GAG GGT GTA TGC GAC TTC Ala Pro Leu Glu Tyr Glu Ala Tyr His Cys Glu Gly Val Cys Asp Phe 40 45 50	766
CCG CTG CGC TCG CAC CTG GAG CCC ACC AAC CAC GCC ATC ATC CAG ACG Pro Leu Arg Ser His Leu Glu Pro Thr Asn His Ala Ile Ile Gln Thr 55 60 65 70	814
CTG ATG AAC TCC ATG GAC CCC GGC TCC ACC CCG CCC AGC TGC TGC GTG Leu Met Asn Ser Met Asp Pro Gly Ser Thr Pro Pro Ser Cys Cys Val 75 80 85	862
CCC ACC AAA TTG ACT CCC ATC AGC ATT CTA TAC ATC GAC GCG GGC AAT Pro Thr Lys Leu Thr Pro Ile Ser Ile Leu Tyr Ile Asp Ala Gly Asn 90 95 100	910

## REIVINDICACIONES

- 5 1. Polipéptido que tiene la capacidad de inducir la formación de tejido similar a tendón o similar a ligamento, en el que dicho polipéptido se codifica por una molécula de ADN que comprende una secuencia de nucleótidos seleccionada del grupo que consiste en:
- (a) los nucleótidos nº 196, nº 199, nº 208, nº 217, nº 361, nº 388, nº 493, nº 496, nº 571 o nº 577 a nº 879 o nº 882 de la SEC ID Nº: 1;
  - (b) nucleótidos que codifican los aminoácidos nº -25, nº 1 o nº 3 a nº 103 o nº 104 de la SEC ID Nº: 2;
  - (c) los nucleótidos nº 845, nº 893 o nº 899 a nº 1201 o nº 1204 de la SEC ID Nº: 3;
  - 10 (d) nucleótidos que codifican los aminoácidos nº 1 o nº 19 a nº 119 o nº 120 de la SEC ID Nº: 4;
  - (e) los nucleótidos nº 410, nº 458, nº 602, nº 605 o nº 659 a nº 961 o nº 964 de la SEC ID Nº: 25;
  - (f) nucleótidos que codifican los aminoácidos nº 1 o nº 19 a nº 119 o nº 120 de la SEC ID Nº: 26; o
  - (g) secuencias alélicas humanas naturales y secuencias de codones degenerados equivalentes de uno cualquiera de (a) a (f), o
  - 15 (h) secuencias que hibridan con cualquiera de (a) a (f) anteriores en condiciones de hibridación rigurosas,
- para uso en un procedimiento para inducir la formación de tejido similar a tendón/ligamento en un paciente que lo necesita.
2. Polipéptido de la reivindicación 1, en el que el polipéptido se codifica por una molécula de ADN que comprende la secuencia de nucleótidos de la SEC ID Nº: 1.
- 20 3. Polipéptido de la reivindicación 1, en el que el polipéptido tiene la secuencia de aminoácidos de la SEC ID Nº: 2.
4. Polipéptido de la reivindicación 1, en el que el polipéptido se codifica por una molécula de ADN que comprende la secuencia de nucleótidos de la SEC ID Nº: 25.
5. Polipéptido de la reivindicación 1, en el que el polipéptido tiene la secuencia de aminoácidos de la SEC ID Nº: 26.
- 25 6. Polipéptido de cualquiera de las reivindicaciones 1 a 5, en el que el dicho polipéptido es para uso en un procedimiento para tratar tendinitis, desgarros de tendones o ligamentos, deformidades y otros defectos de tendones o ligamentos.
7. Polipéptido de cualquiera de las reivindicaciones 1 a 6, en el que el dicho polipéptido está presente como un monómero dentro de una molécula de proteína heterodimérica, y en el que el segundo monómero es un miembro de la superfamilia de factores de crecimiento TGF- $\beta$ .
- 30 8. Composición farmacéutica que comprende un polipéptido que tiene la capacidad de inducir la formación de tejido similar a tendón o similar a ligamento, en la que dicho polipéptido se codifica por una molécula de ADN que comprende una secuencia de nucleótidos seleccionada del grupo que consiste en
- (a) los nucleótidos nº 196, nº 199, nº 208, nº 217, nº 361, nº 388, nº 493, nº 496, nº 571 o nº 577 a nº 879 o nº 882 de la SEC ID Nº: 1;
  - 35 (b) nucleótidos que codifican los aminoácidos nº -25, nº 1 o nº 3 a nº 103 o nº 104 de la SEC ID Nº: 2;
  - (c) los nucleótidos nº 845, nº 893 o nº 899 a nº 1201 o nº 1204 de la SEC ID Nº: 3;
  - (d) nucleótidos que codifican los aminoácidos nº 1 o nº 19 a nº 119 o nº 120 de la SEC ID Nº: 4;
  - (e) los nucleótidos nº 410, nº 458, nº 602, nº 605 o nº 659 a nº 961 o nº 964 de la SEC ID Nº: 25;
  - (f) nucleótidos que codifican los aminoácidos nº 1 o nº 19 a nº 119 o nº 120 de la SEC ID Nº: 26; o
  - 40 (g) secuencias alélicas humanas naturales y secuencias de codones degenerados equivalentes de uno cualquiera de (a) a (f), o
  - (h) secuencias que hibridan con cualquiera de (a) a (f) anteriores en condiciones de hibridación rigurosas,
- en la que la composición es para inducir la formación de tejido similar a tendón/ligamento en un paciente que lo necesita.
- 45 9. Composición farmacéutica de la reivindicación 8, en la que el polipéptido se codifica por una molécula de ADN que comprende la secuencia de nucleótidos de la SEC ID Nº: 1.
10. Composición farmacéutica de la reivindicación 8, en la que el polipéptido tiene la secuencia de aminoácidos de la SEC ID Nº: 2.
- 50 11. Composición farmacéutica de la reivindicación 8, en la que el polipéptido se codifica por una molécula de ADN que comprende la secuencia de nucleótidos de la SEC ID Nº: 25.

12. Composición farmacéutica de la reivindicación 8, en la que el polipéptido tiene la secuencia de aminoácidos de la SEC ID N°: 26.
- 5 13. Composición farmacéutica de cualquiera de las reivindicaciones 8 a 12, en la que la composición farmacéutica es para tratar tendinitis, desgarros de tendones o ligamentos, deformidades y otros defectos de tendones o ligamentos.
14. Composición farmacéutica de cualquiera de las reivindicaciones 8 a 13, en la que el dicho polipéptido está presente como un monómero dentro de una molécula de proteína heterodimérica, y en la que el segundo monómero es un miembro de la superfamilia de factores de crecimiento TGF- $\beta$ .
- 10 15. Composición farmacéutica de cualquiera de las reivindicaciones 8 a 14, en la que la composición farmacéutica comprende además al menos otro agente terapéuticamente útil seleccionado de proteínas BMP, BMP-1 a BMP-11.
16. Composición farmacéutica de la reivindicación 15, en la que la composición farmacéutica comprende además BMP-2.

FIG 1/1

COMPARACIÓN DE V1-1 HUMANA CON MP-52 HUMANA

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
121	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
181	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
226	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
271	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	

Homología a nivel de nucleótidos: 249/312 = 79,8%

Homología a nivel de aminoácidos: 84/104 = 80,8%