



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 308 815**

51 Int. Cl.:  
**C07K 14/47** (2006.01)  
**A61K 38/17** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **98946476 .3**  
96 Fecha de presentación : **23.09.1998**  
97 Número de publicación de la solicitud: **1017718**  
97 Fecha de publicación de la solicitud: **12.07.2000**

54 Título: **Gen novedoso defectuoso en APECED y su uso.**

30 Prioridad: **23.09.1997 FI 973762**

45 Fecha de publicación de la mención BOPI:  
**01.12.2008**

45 Fecha de la publicación del folleto de la patente:  
**01.12.2008**

73 Titular/es: **FIT Biotech Oy**  
**Biokatu 8**  
**33520 Tampere, FI**

72 Inventor/es: **Krohn, Kai;**  
**Heino, Maarit;**  
**Peterson, Pärt;**  
**Scott, Hamish;**  
**Antonarakis, Stylianos;**  
**Lalioti, Maria;**  
**Shimizu, Nobuyoshi y**  
**Kudoh, Jun**

74 Agente: **Carpintero López, Mario**

ES 2 308 815 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

Gen novedoso defectuoso en APECED y su uso.

5 **Campo de la invención**

La presente invención se refiere a un gen novedoso, una proteína novedosa codificada por dicho gen, una forma mutada del gen y a usos diagnósticos y terapéuticos del gen o la forma mutada del mismo. Más específicamente, la presente invención se refiere a un gen novedoso defectuoso en el síndrome de poliendocrinopatía autoinmune de tipo I (APS I), también llamado poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED) (MIM No. 240.300).

**Antecedentes**

15 El síndrome autoinmune poliglandular de tipo I (APS I), también conocido como poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED), es una enfermedad hereditaria recesiva rara (MIM No. 240.300) que es más frecuente entre ciertas poblaciones aisladas, tales como poblaciones de finlandeses, de sardos y judíos iraníes. La incidencia de la enfermedad entre los fineses y los judíos iraníes se estima que es 1:25000 y 1:9000, respectivamente, mientras solamente unos pocos casos en otras partes del mundo se encuentran cada año.

20 APECED es uno de los dos principales síndromes de poliendocrinopatía autoinmunes. El factor causante de APECED no ha sido todavía identificado. El síndrome se caracteriza por la por la carencia de tolerancia a numerosos auto-antígenos y se pueden por lo tanto considerar como un prototipo de enfermedades autoinmunes específicas de órgano. En APECED, el paciente desarrolla candidiasis mucocutánea crónica inmediatamente después del nacimiento, y más tarde varias enfermedades autoinmunes específicas de órgano, principalmente se produce hipoparatiroidismo, enfermedad de Addison, gastritis atrófica crónica, con o sin anemia perniciosa, y disfunción gonadal de la pubertad [Ahonen P, Clin. Genet. 27 (1985) 535-542]. Un criterio aceptado para la diagnosis de APECED es la presencia de al menos dos de los tres síntomas principales, enfermedad de Addison, hipoparatiroidismo y candidiasis, en pacientes [Neufeld, M. *et al.*, Medicine 60 (1981) 355-362]. Inmunológicamente, los hallazgos principales son la presencia de anticuerpos de suero de alta titulación contra los órganos afectados, anticuerpos contra *Candida albicans*, y respuestas bajas o carentes de células T hacia antígenos de Candida [Blizzard, R. M. y Kyle M., J. Clin. Invest. 42 (1963) 1653-1660; Arulanantham, K. *et al.*, New Eng. J. Med. 300 (1979) 164-168; Krohn, K. *et al.*, Lancet 339 (1992) 770-773; Uibo R. *et al.*, J. Clin. Endocrinol. Metab. 78 (1994) 323-328]. La enfermedad usualmente se produce en la juventud, pero pueden aparecer nuevos síntomas específicos de tejidos a lo largo de la vida [Ahonen, P. *et al.*, New Engl. J. Med. 322 (1990) 1829-1836]. APECED no está asociado a un haplotipo HLA particular, y tanto varones como mujeres están igualmente afectados consistente con el modo recesivo autonómico de herencia.

40 El locus para el gen APECED se ha mapeado en el cromosoma 21q22.3 entre los marcadores génicos D21S49 y D21S171 basándose en los análisis de unión de familias finesas [Aaltonen, J. *et al.*, Nature Genet. 8 (1994) 83-87]. Recientemente, Börses *et al.* reseñaron una puntuación máxima de LOD de 10,23 con el marcador D21S1912 justo próximo al gen PFKL, y de este modo mediante desequilibrio de unión de la región crítica para APECED se pueden considerar que son de menos de 500 kb entre marcadores D21S1912 y D21S171. La heterogeneidad del locus no se reveló mediante análisis de unión de familias no finesas [Björnses, P. *et al.*, Am. J. Hum. Genet. 59 (1996) 879-886].

45 Para el gen de APECED, el nombre “regulador autoinmune” o “AIRE” se ha adoptado por la comunidad científica después de la fecha de prioridad de la presente solicitud. De manera similar, la proteína codificada por el gen AIRE se denomina ahora la “proteína AIRE”.

50 Los mapas físicos de cromosoma humano 21q22.3 se han desarrollado usando YACs, y bacterias basándose en grandes vectores de clonación insertados [Chumakov *et al.*, Nature 359 (1992) 380; Stone *et al.*, Genome Res. 6 (1996) 218], y muchos laboratorios han contribuido a la construcción de un mapa de transcripción del cromosoma entero y 21q22.3 en particular [Chen *et al.*, Genome Res. 6 (1996) 747-760; Yaspo *et al.*, Hum. Mol. Genet. 4 (1995) 1291-1304]. Numerosos exones atrapados en los cósmidos específicos del cromosoma 21 y también secuencias contiguas físicas de la región crítica de APECED se han identificado y caracterizado parcialmente. Además, un número de EST del proyecto del genoma humano internacional se han mapeado para la región crítica de APECED.

60 Recientemente, como parte de los esfuerzos internacionales de generación de la secuencia entera del cromosoma 21 humano y acuerdos internacionales sobre la capacidad inmediata de este tipo de datos de secuencias, la secuencia particular de la región crítica del gen de APECED se hizo disponible en GenBank mediante El Centro de genoma Humano de Stanford que actualmente lleva a cabo la secuenciación de 1,0 Mb alrededor de la región crítica del gen APECED.

65 Sin embargo, la localización y secuencia precisa del gen de APECED y la naturaleza del producto génico no se han aclarado hasta ahora. De este modo la presencia de la diagnosis de APECED se basa principalmente en el desarrollo de los síntomas clínicos y hallazgos clínicos típicos, por ejemplo, la presencia de anticuerpos contra las enzimas del córtex adrenal o esteroadrenérgicas P450c17 y/o P450sc. El análisis de unión se usa raramente. Además, no están fácilmente disponibles medios para la diagnosis natal o presintomática de la enfermedad, ya que el análisis de unión proporciona solamente datos indirectos mediante los marcadores de genes conocidos y requiere muestras de varios

## ES 2 308 815 T3

miembros de la familia en varias generaciones. De manera adicional, el análisis de unión es tedioso y se puede realizar solamente en laboratorios especializados por personal altamente cualificado.

5 También el mapeado de los vehículos del gen de la enfermedad se basa actualmente en el análisis de unión y de este modo no es fácilmente disponible.

### Resumen de la invención

10 Los inventores han identificado ahora un gen novedoso que codifica una proteína novedosa de dedos de cinc, que se denomina un regulador autoinmune 1 o AIR-1, que está mutado en APECED. El gen y proteína novedoso permiten además el desarrollo de la diagnosis y terapia de enfermedades relacionadas con la maduración inmune y regulación de la respuesta inmune de auto y no auto anticuerpo, tal como APECED.

15 El objeto de la invención es proporcionar medios que son útiles en un procedimiento diagnóstico y un procedimiento terapéutico génico en la diagnosis y tratamiento de enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a un auto y no auto, tal como APECED.

20 Otro objeto de la invención es proporcionar un procedimiento novedosos para la diagnosis de las enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a un auto y no auto, tal como APECED, incluyendo la diagnosis pre- y post-natal y el mapeado de los vehículos, siendo el procedimiento fácil y seguro de realizar.

25 La presente invención se refiere a una secuencia de ADN aislado que tiene la SEQ ID NO: 1 o un fragmento funcional de la misma que codifica una proteína asociada a poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED), o su forma mutada que incluye un APECED responsable del gen defectuoso, la secuencia de ADN que está asociada a las enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto, tal como APECED.

30 La presente invención también se refiere a una proteína que tiene la SEQ ID NO: 2 de aminoácidos o un fragmento funcional de la misma asociada a poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED) o sus variantes asociadas a APEDED.

35 Dicha proteína tiene distintos motivos estructurales, incluyendo el motivo de dedos PHD (PHD), el motivo LXXLL (L), la región rica en prolina (PRR), y la región rica en cisteína (CRR).

40 La presente invención además se refiere a un procedimiento *in vitro* para la diagnosis de enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a un auto y no auto, tal como APECED, que comprende la detección en un espécimen biológico de la presencia de un fragmento funcional de una secuencia de ADN que tiene la SEQ ID NO: 1 que codifica una proteína asociada a APECED, o una forma mutada de una secuencia de ADN que tiene la SEQ ID NO: 1 que incluye un gen defectuoso responsable de APECED.

45 La presente invención además se refiere al uso de las secuencias de ADN identificadas anteriormente en la diagnosis *in vitro* de las enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto, tal como APECED.

50 La presente invención además se refiere a un procedimiento *in vitro* para la diagnosis de enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto, tal como APECED, que comprende la detección en un espécimen biológico de la ausencia de una proteína que tiene la SEQ ID NO: 2 o la presencia de un fragmento funcional de la misma asociada a enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto, tal como APECED, o sus variantes asociadas a APECED.

55 La presente invención además se refiere al uso de la proteína anteriormente mencionada o su fragmento funcional en la diagnosis *in vitro* de enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto, tal como APECED.

60 La presente invención además se refiere al uso de las secuencias de ADN identificadas anteriormente para la preparación de una preparación farmacéutica útil en un procedimiento de terapia génica de las enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a un auto y no auto, tal como APECED.

### Breve descripción de los dibujos

65 La Figura 1 muestra un mapa físico del locus del gen APECED en el cromosoma 21 q22.3. Los cósmidos D1G8, D40G11, D9G11, D28B11, y D4G11, que se solapan con los clones usados para la secuenciación genómica [Kudoh, J. *et al.*, ADN Res. 4 (1997) 45-52] se indican mediante líneas horizontales. El gen de APECED localizado justo proximal al extremo 5' del gen vecino de PFKL se indica mediante una flecha entera de color negro. N indica los sitios NotI. El marcador de ADN D21S1912 se muestra como casilla en blanco.

La Figura 2 muestra las estructuras del gen APECED y las proteínas AIR. (A) estrategia de clonación de los ADNc de APECED y el orden de los exones en el gen de APECED. Fragmentos de ADN amplificados mediante la PCR y 3'-y 5'-RACE se indican por las líneas. Exón 1' es el exón no codificante en 5'- de AIR-2 y AIR-3. un ajuste adicional de alternativa de AIR-3 en el exón 10, da como resultado un cambio de aminoácidos en su cadena abajo, se indica mediante líneas verticales. Cada exón, excepto el exón 1', está en el borde de la secuencia de consenso del sitio de ajuste común, ag:gt. Las mutaciones en el exón 2 y exón 6 se indican por las flechas. (B) presentación esquemática de las tres proteínas AIR que muestran los distintos motivos estructurales, que incluyen el motivo de dedos PHD (PHD), el motivo LXXLL (L), la región rica en prolina (PRR), y la región rica en cisteína (CRR).

La Figura 3 muestra electroferogramas que muestran la secuencia que rodean las mutaciones en el gen de APECED. (A) Análisis de mutación e una familia suiza de APECED. Los padres son heterocigóticos para el alelo (normal "C" y anormal "T"). El chico y chica afectados muestran la transición "C" a "T" que da como resultado la mutación no codificante "Arg" a "parada" en la posición de aminoácidos 257. (B) Análisis de mutación de dos pacientes fineses de APECED. El paciente MP es homocigoto para el alelo mutante (izquierda), NP es homocigoto para el alelo (derecha). (C) El paciente NP muestra la transversión "A" a "G" que da como resultado la mutación no codificante "Lys" a "Glu" en la posición de aminoácidos 42. FLEB es un control normal.

La Figura 4 muestra el resultado de un ensayo de digestión de la enzima de restricción TaqI que muestra la mutación de parada de R257. Se muestran cuatro pacientes de APECED [HP1 (calle 1), HP2 (calle 2), NP (calle 6), y MP (calle 8)], las madres de dos familias [HM (calle 5) y NM (calle 7)], dos parientes sanos [HN1 (calle 3) y HN2 (calle 4)] de la familia H y controles normales [C1, C2 y C3 (calles 9-11)]. Los pacientes de APECED HP1, HP2 y MP son homocigotos para la mutación de parada R257. El paciente NP de APECED es heterocigótico para la mutación de parada de R257 pero lleva una mutación en una mutación diferente distinta del alelo del gen de APECED (mostrado anteriormente en la Fig. 3C). Tanto las madres (HM y NM) como dos parientes sanos (HN1 y HN2) son heterocigóticos para la mutación de parada R257 y por lo tanto vehículos de APECED pero no tienen la enfermedad. Dos controles (C1 y C2) son ambos homocigóticos para alelos normales. Los alelos normales producen un fragmento de menos de 225 pares de bases, el fragmento mutado es la banda superior a 285 pares de bases.

La Figura 5 muestra una alineación de la secuencia de aminoácidos para el motivo de dedos de PHD de AIR-1, Mi-2, y TIF1. Los restos de aminoácidos de consenso conservados en el motivo de dedos de PHD se indica mediante letras en negrita debajo. Los restos que son idénticos a AIR-1 (aa 299-340) se muestran por las líneas de puntos. Números de acceso del GenBank de Mi-2 y TIF1 son X86691 y AF009353, respectivamente.

La Figura 6 es una transferencia de Western que muestra la expresión de AIR-1 en hígado fetal. Se ensayó una muestra de hígado fetal sobre PAGE, se transfirió a un filtro de nitrocelulosa y se sondó con suero como sigue: Calle 1, suero de ratón de control, calle 2 suero de control de ratón absorbido con péptido AIR-1/2 (SEQ ID NO: 25), calles 3 y 4, suero de un ratón inmunizado con péptido AIR-1/2 durante cuatro y seis semanas, respectivamente y absorbido con el péptido AIR-1/2, calles 5 y 6, suero no absorbido de un ratón inmunizado con el péptido AIR-1/2 durante cuatro y seis semanas, respectivamente. La banda superior observada en las calles 5 y 6 representa la proteína AIR-1 con un peso molecular de aproximadamente 58 kD, la banda inferior es un producto de ruptura de aproximadamente de 20 kD de la proteína AIR. Las bandas observadas en las calles son no específicas.

La Figura 7 muestra la expresión del ARNm de APECED (7A) o la proteína AIR (7B, 7C y 7D) mostrada en la hibridación *in situ* (7A) o mediante inmunohistoquímica (7B, 7C y 7D). La Figura 7A muestra las células positivas de ARNm de APECED dispersadas en la región medular del timo humano. La Figura 7B muestra células similares con la misma localización teñida ahora para proteína AIR. La Figura 7C es una amplificación mayor de 78, que muestra la localización de la proteína AIR en los núcleos. Hay que hacer notar el patrón de localización moteado en los núcleos. La Figura 7D muestra la localización citoplásmica de la proteína de AIR en unas pocas células en la médula de ganglios linfáticos.

La Figura 8 muestra la caracterización fenotípica de las células reactivas de APECED en el timo mediante inmunofluorescencia doble. La proteína AIR se observa como un color rojo en los núcleos, formando un patrón moteado típico con manchas de los núcleos. En la Figura 8A, la tinción conjunta es con un anticuerpo que reconoce los marcadores de bajo peso molecular (AE1). Las células positivas de APECED están en los dos tipos, uno expresa citoqueratina y de esta manera célula epitelial, la otra es no epitelial y no expresa de manera conjunta las citoqueratinas. En la Figura 8 B una célula positiva de APECED que expresa de manera conjunta un marcador (CD83) típico para las células que pertenecen a la línea de células dendríticas de macrófagos monocitos.

La Figura 9 muestra la expresión de la proteína de AIR, mostrado mediante inmunofluorescencia, en células dendríticas maduras, activadas de sangre periférica. La expresión de la proteína de AIR muestra manchas distintas en los núcleos de las células dendríticas.

### Descripción detallada de la invención

La presente invención se basa en estudios que pretenden la identificación y caracterización del gen defectuoso en APECED. En los estudios de secuencia, se construyó un cósmido/BAC (cromosoma artificial bacteriano) que es contiguo de 520 kb que cubre cuatro marcadores de genes D21S1460-D21S1912-PFKL-D21S154 [Kudoh, J. *et al.*, DNA Res. 4 (1997) 45-52], y se realizó la secuenciación genómica en esta región [Kawasaki, K. *et al.*, Genome Res. 7

## ES 2 308 815 T3

(1997) 250-261]. A partir de esta información de la secuencia genómica se determinó que la distancia entre D21S1912 y PFKL era aproximadamente 140 kb (Fig. 1).

5 Usando un programa de ordenador, tal como GRAIL y GENSCAN [Uberbacher, E. C. y Mural, R. J., Proc. Natl. Acad. Sci. USA 88 (1991) 11261-11265; Burge, C. y Karlin, S., J. Mol. Biol. 268 (1997) 78-94], se realizó la selección génica en los datos de secuenciación parcial dentro de esta región. GENSCAN predijo varios genes entre D21S1912 y PFKL. Uno de estos genes localizados justo próximos al gen PFKL contenían el exón HC21 EXc33 previamente atrapado [Kudoh, J. *et al.*, DNA Res. 4 (1997) 45-52] o MDC04M06 [Chen, H. *et al.*, Genome Res. 6 (1996) 747-760]. Un conjunto de cebadores para la reacción en cadena de la polimerasa (PCR) se diseñó después a partir de los exones predichos. La selección de PCR de diversas genotecas de ADNc que usan estos cebadores permitió el aislamiento de un clon de ADNc que contiene el exón HC21EXc33 (exón 13) a partir de la genoteca de ADNc de timo (Fig. 2A).

15 Una amplificación rápida en 3' de los extremos de ADNc (3'-RACE) y 5'-RACE que usa el Kit Marathon™ cDNA Amplification (Clontech Laboratories Inc, California, USA) de acuerdo con el protocolo del fabricante a partir de la genoteca de ADNc de timo se realizó usando un cebador c33F (SEQ ID NO: 7) y un cebador 1R (SEQ ID NO: 8), respectivamente.

20 El análisis de secuenciación reveló una única secuencia de 2027 pares de bases en la superposición de los productos de la PCR que contiene un marco abierto de lectura de 1635 pares de bases (ORF) desde metionina en el nt 128 a un codón de parada TAG en el nt 1763 que codifica una proteína novedosa predicha denominada AIR-1, para el regulador autoinmune 1. AIR-1 codifica una proteína de 545 aminoácidos con un punto isoeléctrico predicho de 7,32 y una masa molecular calculada de 57.723 (Fig. 2B).

25 Un A 5'-RACE del ADNc de timo que usa un cebador 4R (SEQ ID NO: 9) dio como resultado un producto de ajuste alternativo. Además, se amplificaron dos tipos de clones de ADNc con un par de cebadores 3F/c33R (SEQ ID NO: 10/SEQ ID NO: 11) y estos clones codifican las proteínas AIR-2 y AIR-3, SEQ ID NO: 4 y SEQ ID NO: 6, respectivamente (Fig. 2A) SEQ ID NO: 3 y SEQ ID NO: 5). Las proteínas AIR-2 y AIR-3 constan de 348 y 254 aminoácidos, respectivamente (Fig. 2B). Estos resultados sugieren que el gen APECED se transcribe como al menos tres tipos de ARNm mediante ajuste alternativo y/o uso de un exón alternativo en 5' dentro del gen. El análisis de RT-PCR [Griffin, H. G. y Griffin, A. M., PCR Technology. Current Innovations, CRC Press, 1994] reveló que la transcripción AIR-1 también se expresa en hígado fetal (datos no mostrados).

30 El gen APECED es de aproximadamente 13 kb de longitud y contiene 15 exones, incluyendo el exón 1' específico para AIR-2 y AIR-3. Se transcribe en la dirección de centrómero a telómero (Figs 1, 2A). Basándose en esta información, los cebadores de la PCR se diseñaron para amplificar cada exón a partir del ADN genómico y un análisis de mutación de las familias de APECED suizas y finlandesas. La comparación de secuencias identificó dos mutaciones en el gen de APECED de los pacientes (Fig. 3). La primera mutación cambia un codón de Arg (CGA) a un codón de parada (TGA) en la posición de aminoácido 257 en el exón 6. Esta mutación se denominó mutación de parada R257. La segunda mutación es una mutación no codificante que se deriva del cromosoma materno en un paciente finlandés (NP): un codón de Lys (AAG) cambia a un codón Glu (GAG) en la posición de aminoácido 42 en el exón 2. Esta mutación se denomina mutación K42E (Figs 2A, 3C).

45 La mutación de parada R257 destruye un sitio de la enzima de restricción *TaqI* y la mutación K42E introduce un sitio novedoso *TaqI*. De este modo estas dos mutaciones se pueden fácilmente demostrar en uno o ambos alelos mediante la digestión por *TaqI* o mediante digestión usando otra escisión por enzima en el sitio de reconocimiento 5'-TCGA-3' (Fig. 4).

50 La proteína AIR-1 tiene fuerte homología en ciertos dominios con los antígenos principales (Mi-2) asociados a la enfermedad autoinmune dermatomitosi [Seeig, H. P. *et al.*, Arthritis Rheum. 38 (1995) 1389-1399; Ge, Q. *et al.*, J. Clin. Invest. 96 (1995) 1730-1737], Sp140, una proteína del cuerpo, un orgánulo implicado en la patogénesis de ciertos tipos de leucemia, y que también la diana de anticuerpos en el suero de pacientes con la enfermedad autoinmune cirrosis biliar primaria [Bloch, D. B. *et al.*, J. Biol. Chem. 271 (1996) 29198-29204]. Además, las homologías se extienden a otras proteínas nucleares tal como TIF1 [Le Douarin, B. *et al.*, EMBO J. 14 (1995) 2020-2033], LYSP100 [Dent, A. L. *et al.*, Blood 88 (1996) 1423-1426], y proteínas de levadura supuesta y *C. elegans*. Las homologías de la proteína AIR-1 están principalmente en dos motivos de dedos PHD (aminoácido 299 a 340 y 434 a 475) (Fig. 5). AIR-1 también contiene una región rica en prolina (aminoácido 350 a 430) (Fig. 2B). El dedo de PHD es una estructura rica en cisteína que se distingue del dedo del dominio (C3HC4) y LIM (C2HC5) debido a que contiene un consenso de C4HC3. [Aasland, R. *et al.*, Trenes Biochem. Sci. 20 (1995) 56-59]. El motivo de dedo de PHD se encuentra en un número de proteínas asociadas a cromatina tal como HRX que está implicado en la traslocación t(11:17) en leucemia aguda [Chaplin, T. *et al.*, Blood 86 (1995) 2073-2076]. La región rica en prolina se asume que está implicada en la interacción proteína-proteína y unión de ADN. La presencia del dedo de PHD y regiones ricas en prolina indica una función para las AIR como proteínas reguladoras de la transcripción. Sin embargo, las proteínas AIR no tienen señal evidente de traslocación nuclear, y de este modo otras proteínas que contienen tal señal puede interactuar con para translocarla al núcleo. De hecho, las proteínas AIR también tienen el motivo LXXLL que es una secuencia de señal que se une a los receptores nucleares [Heery, D. M. *et al.*, Nature 387 (1997) 733-736] (Fig. 2B).

La ilustración clínica de APECED y la anomalía inmunológica observada con fuerte respuesta autoinmune a varios órganos y antígenos diana sugieren que el producto del gen APECED tiene un papel central en la maduración y regulación inmune (ontogenia) de la respuesta inmune a auto y no auto.

5 De acuerdo con el procedimiento de diagnóstico de la invención, la presencia del APECED defectuoso se puede detectar a partir de una muestra biológica mediante procedimientos de detección conocido adecuado para detectar mutaciones. Tales procedimientos incluyen el procedimiento descrito por Saiki *et al.* [Proc. Natl. Acad. Sci USA 86 (1989) 6230-6234] utilizando la hibridación a una sonda de oligonucleótidos específica de alelo, o sus modificaciones; el procedimiento descrito por Newton, C. R. *et al.* [Nucl. Acids Res. 17 (1989) 2503-2516] usando las secuencias de ADN o fragmentos de ADN de la invención como sondas; el procedimiento de miniselección de fase sólida descrito por Syvänen *et al.* [Genomics 8 (1990) 684-692] en el que se hace uso de una sonda biotinilada; o el procedimiento de ligamiento de oligonucleótidos descrito por Landegren, U. *et al.* [Science 241 (1988) 1077-1080]. Los procedimientos incluyen la desnaturalización en electroforesis de gel de gradiente (DGGE) [Fischer, S.G. y Lerman, L.S., PNAS 80 (1983) 1579-1583] o una modificación de este procedimiento, electroforesis en gel desnaturalizante constante (CDGE) [Hoving *et al.*, Genes Chromosomes Cancer 5 (1992) 97-103]. El principio de separación de mutaciones de DGGE y CDGE se basa en el comportamiento de fusión de la doble hélice de ADN de un fragmento dado.

20 Ya que las mutaciones del gen APECED implican un sitio sensible a la digestión por *TaqI*, las mutaciones se detectan preferiblemente en uno o ambos alelos mediante digestión por *TaqI* o mediante digestión usando otra enzima que escinde en el sitio de reconocimiento 5'-TCGA-3'. Se puede usar la escisión discordante química para el análisis de mutación [Grompe, M. *et al.*, Proc. Natl. Acad. Sci. USA 86 (15) (1989) 5888-5892].

25 En el procedimiento diagnóstico de la invención la muestra biológica puede ser cualquier tejido o fluido corporal que contiene células, tal como sangre, por ejemplo, sangre del cordón umbilical, células de sangre separadas, tal como linfocitos, células B, células T etc., material de biopsia, tal como biopsia de hígado fetal o de timo, esperma, saliva, etc. La muestra biológica puede ser, si es necesario, tratada previamente de una manera adecuada conocida por los expertos en la técnica.

30 Cuando la secuencia de ADN de la presente invención se usa terapéuticamente se pueden emplear cualesquiera técnicas para la terapia génica. De acuerdo con lo anterior, en la técnica conocida como terapia *ex vivo* se toman células de paciente (por ejemplo, sangre del cordón umbilical del feto) con el gen defectuoso del paciente, las secuencias de ADN que codifican el producto génico normal (sano) incorporado en un vector vehículo se transluce o se transfecta a las células y las células se devuelven al paciente. Si las técnicas conocidas como terapia *in situ* se usan, las secuencias de ADN que codifican el producto génico normal se insertan primero a un vector vehículo adecuado, y el vehículo se introduce después en el tejido afectado, tal como sangre periférica, hígado o médula ósea. El vector vehículo usado puede ser un vector de retrovirus, un vector de adeno virus, un vector de virus asociado a adeno (AAV). La terapia se puede realizar intra útero o durante la vida del adulto. Dependiendo de las células a tratar estas técnicas conducen bien a una cura transitoria, donde las células del órgano afectado se tratan, o a una cura permanente, en el caso del tratamiento de las células troncales.

45 La presente invención proporciona medios para una diagnosis fácil y más rápida de las enfermedades relacionadas con la maduración y regulación inmune de la respuesta a auto y no auto, tal como APECED, y, específicamente, permite la diagnosis prenatal y diagnosis de vehículo. Además, proporciona un antecedente para la terapia.

La invención ahora se esclarece mediante los siguientes ejemplos no limitantes.

#### Ejemplo 1

##### 50 Localización del gen APECED

La secuenciación genómica de los ADN de cósmico se realizó mediante el procedimiento de escopeta descrito por Kawasaki, K. *et al.*, Genome Res. 7 (1997) 250-261. Cosmids D1G8, D40G11, D9G11, D28B11, y D4G11 y marcador génico D21S1912 se describen por Kudoh, J. *et al.*, DNA Res. 4 (1997) 45-52].

##### 55 Clonación de ADNc

60 Los AND de fago preparados a partir de genoteca de ADNc de timo humano (Clontech, HL1127a) se usaron como un molde de la PCR. 20 ng de ADN de fago que representa aproximadamente  $4 \times 10^8$  fagos se añadió a 10 ml de mezcla de reacción que contiene tampón 1 x [16 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 50 mM Tris-HCl, pH 9,2, 1,75 mM MgCl<sub>2</sub>, 0,001% (p/v) gelatina], 0,2 mM cada uno de los dNTPs, 1 M Betaína (Sigma), 0,35 U de Tap y Pwo ADN polimerasa (EXpand Long Template PCR System, Boehringer Mannheim), y 0,5 mM de cada uno de los cebadores, 2F y c33R, 2F y 4R, y 2F' y 2R', respectivamente.

65 El fragmento de ADNc se amplificó mediante la PCR usando las siguientes condiciones: 94°C durante 3 min., 35 ciclos de 94°C durante 30 seg, 60°C durante 30 seg en 2F/c33R y 2F/\_4R o 65°C durante 30 seg en 2F'/2R', y 68°C fragmentos 90 seg. 3'- y 5'-RACE se llevaron a cabo mediante el Kit Marathon ADNc Amplification (Human Thymus; Clontech). La reacción de la PCR se realizó en un volumen de 10  $\mu$ l que contenía tampón 1 x (50 mM KCl, 10 mM

## ES 2 308 815 T3

Tris-HCl, pH 8,3, 1,5 mM MgCl<sub>2</sub>, 0,001% (p/v gelatina), 0,2 mM cada uno de los dNTPs, 0,25 U de AmpliTaq Gold polimerasa (Perkin-Elmer), y 0,5 mM de cada uno de los cebadores específicos de exón. El producto de 3'-RACE se amplificó mediante la PCR con las siguientes condiciones: 95°C durante 9 min., 35 ciclos de 94°C durante 30 seg, 60°C durante 30 seg, y 72°C durante 30 sec.

5 Los fragmentos de ADNc se secuenciaron mediante el procedimiento de secuenciación de ciclo terminador de tinte desoxi (de acuerdo con el protocolo P/N 402078 del Kit ABI PRISM Dye Terminator Cycle Sequencing Ready Reaction, Perkin Elmer Corporation, California) usando cebadores específicos, 2F y c33R, y AmpliTaq/FS ADN polimerasa (Perkin-Elmer), y después se analizó mediante el uso de un secuenciador de ADN (Applied Biosystems 10 377). Se usaron las secuencias de cebadores

1 R: 5'-GTTCCCGAGTGGGAAGGCGCTGC-3' (SEQ ID NO: 8)

15 2F: 5'-GGATTCAGACCATGTCAGCTTCA-3' (SEQ ID NO: 12)

3F: 5'-GAGTTCAGGTACCCAGAGATGCTG-3' (SEQ ID NO: 10)

c33R: 5'-CTCGCTCAGAAGGGACTCCA-3' (SEQ ID NO: 11)

20 4R: 5'-AGGGGACAGGCAGGCCAGGT-3' (SEQ ID NO: 9)

2F': 5'-GTGCTGTTCAAGGACTACAAC-3' (SEQ ID NO: 13)

2R': 5'-TGGATGAGGATCCCCTCCACG-3' (SEQ ID NO: 14)

25 AP1: 5'-CCATCCTAATACGACTCACTATAGGGC-3' (SEQ ID NO: 15) y

c33F: 5'-GATGACACTGCCAGTCACGA-3' (SEQ ID NO: 7).

### 30 Ejemplo 2

#### *Análisis de mutación del gen APECED*

35 Para el análisis de mutación las muestras de ADN se purificaron a partir de células de sangre periférica de pacientes con APECED y de vehículos sospechosos de APECED y de controles sanos normales (de acuerdo con Sambrook *et al.* 1989, Molecular Cloning. A Laboratory Manual. CSH Press) y se sometieron a la PCR usando cebadores específicos para todos los exones identificados.

40 Para secuenciar los exones mutados, los fragmentos de PCR, 6F/6R en el exón 6 y 49300F/49622R in exón 2, se amplificaron mediante la PCR con las siguientes condiciones: 95°C para 9 min., 35 ciclos de 94°C durante 30 seg, 60°C durante 30 seg y 72°C durante 30 seg, y 94°C durante 3 min., 35 ciclos de 94°C durante 30 seg, 60°C durante 30 seg, y 68°C durante 30 seg, respectivamente. Los productos de la PCR se secuenciaron usando cebadores específicos

45 6F: 5'-TGCAGGCTGTGGGAAGTCCA-3' (SEQ ID NO: 16)

6R: 5'-AGAAAAGAGCTGTACCCTGTG-3' (SEQ ID NO: 17)

3R: 5'-TGCAAGGAAGAGGGGCGTCAGC-3' (SEQ ID NO: 18)

50 49300F: 5'-TCCACCACAAGCCGAGGAGAT-3' (SEQ ID NO: 19)

y 49622R: 5'-ACGGGCTCCTCAAACACCACT-3' (SEQ ID NO: 20).

55 En el análisis de mutación mediante secuenciación, dos pacientes suizos y tres finlandeses (HP1, HP2 y MP) con APECED eran homocigóticos para el alelo de parada de R257, mientras un paciente finlandés (NP) era heterocigótico para esta mutación (Fig. 3A, B). la mutación de parada de R257 de NP se derivó del cromosoma paterno. La segunda mutación, K42E mutación, se encontró en un paciente finlandés (NP): un codón de Lys (AAG) cambia a un codón de Glu (GAG) en la posición de aminoácido 42 en el exón 2. (Figs 2A, 3C). Esta mutación se deriva del cromosoma materno.

### 60 Ejemplo 3

#### *Análisis de enzima de restricción TaqI de dos mutaciones en los exones 2 y 6 del gen APECED*

65 El análisis de los sitios de mutación en los exones 2 y 6 en grandes series de individuos se realizó usando la enzimas de restricción TaqI. La digestión TaqI para los exones 2 y 6 se realizó como sigue. Diez microlitros del producto de amplificación se incubó a 65°C durante 1 hora en 20 µl de mezcla de reacción que contenía 1x de tampón de digestión

## ES 2 308 815 T3

TaqI (New England Biolabs, NY, 100  $\mu$ l/ml de BSA y 10 U de enzima TaqI (New England Biolabs, NY). Después se separaron los fragmentos de digestión en gel de agarosa al 1,5% y se visualizaron mediante tinción por EtBr.

Para el exón 2, el fragmento que contenía el sitio de mutación K42E se amplificó con los cebadores GR1/2F y GR1/2R con las siguientes condiciones: 95°C durante 3 min., 35 ciclos de 94°C durante 30 seg, 62°C durante 30 seg y 72°C durante 1 min. La mezcla de reacción 1x usada contenía 50 mM KCl, 10 mM Tris-HCl, pH 8,3, 1,5 mM MgCl<sub>2</sub>, 0,001% (p/v) gelatina, 0,2 mM cada uno de los dNTP, 0,25 U de Dynazyme (Finnzymes, Finlandia), y 0,5 mM de cada uno de los cebadores específicos de exón. El alelo normal produce un fragmento de 312 pares de bases mientras el alelo mutado proporciona un fragmento de 133 pares de bases y de 179 pares de bases. Las secuencias de cebadores para GR1/2F y GR1/2R son 5'-TGGAGATGGGCAGGCCGCAGGGTG SEQ ID NO: 21) y 5'-CAGTCCAGCTGGGCTGAGCAGGTG (SEQ ID NO: 22), respectivamente.

Para el exón 6, el fragmento que contenía el sitio de mutación de parada R257 se amplificó con los cebadores GR1/51F y GR1/51R con las mismas condiciones descritas para el exón 2 (véase anteriormente). El alelo normal produce un fragmento de 225 pares de bases mientras el alelo mutado produce un fragmento de 285 pares de bases. Las secuencias de cebador para GR1/51F y GR1/51R son 5'-GCGGCTCCAAGAAGTGCATCCAGG (SEQ ID NO: 23) y 5'-CTCCACCCTGCAAGGAAGAGGGGC (SEQ ID NO: 24), respectivamente.

La selección de 50 individuos sanos finlandeses y 50 suizos no revelan mutaciones de parada R257 o K42E mediante digestión por TaqI. De manera similar, se realizó el análisis de la PCR de 20 japoneses no afectados y no se encontraron mutaciones en ninguna de estas posiciones. Estos resultados demuestran que el gen APECED es responsable de la patogénesis de APECED.

Se encontraron mutaciones en las transcripciones de AIR-1 pero no en las transcripciones de AIR-2 y AIR-3 de los pacientes de APECED ensayados. Dos pacientes suizos y tres finlandeses (HP1, HP2 y MP) son homocigóticos para la mutación de parada R257 que carece completamente de la proteína funcional de AIR-1 pero todavía tiene intactas las proteínas AIR-2 y AIR-3.

Una mutación común parece responsable del defecto genético en aproximadamente 90% de los casos de APECED finlandeses y un análisis de haplotipo con los marcadores D21 S141, D21S1912 y PFKL muestra que la mutación de parada de R257 es probable que sea esta mutación común [Björnses. P. *et al.*, Am. J. Hum. Genet. 59 (1996) 879-886].

### Ejemplo 4

#### 35 *Análisis de la expresión de la proteína AIR*

En este ejemplo, los péptidos sintéticos que representan secuencias de aminoácidos de la proteína AIR-1, se usaron para generar un antisuero de ratón polivalente contra la proteína AIR-1.

Para la síntesis de péptidos, se eligieron dos péptidos de acuerdo con la predicción de antigenicidad mediante el programa Pepsort (paquete GCC, Wisconsin, Estados Unidos). Los péptidos AIR-1/2 y AIR-1/6 (TLHLKE KEGCPQAFH, SEQ ID NO: 25 y GKNKARSSSGPKPLV, SEQ ID NO: 26, respectivamente) que representan los exones 2 y 6, respectivamente, del gen APECED se sintetizaron en un núcleo de lisina ramificado (resina Fmoc8-Lys4-Lys2-Lys-betaAla-Wang, Calbiochem-Novabiochem, La Jolla, Ca, USA) que da como resultado un péptido antigénico múltiple octamérico (MAP) [Tam, J. P. *et al.*, Proc. Natl. Acad. Sci. Estados Unidos 85 (1988) 5409-5413; Adermann, K. *et al.*, in Solid Phase Synthesis, Biological and Biomedical Applications, p. 429-432, Ed. R. Epton, Mayflower Worldwide Ltd., Birmingham, 1994]. Se realizaron las síntesis mediante química de Fmoc (N-(9-fluorenil) methoxycarbonilo) sobre un sintetizador de péptidos múltiples simultáneo (SMPS 350, Zinsser Analytic, Frankfurt, Alemania). Se analizó la pureza de MAPS mediante HPLC de fase inverse (System Gold, Beckman Instruments Inc, Fullerton, CA, USA).

Para obtener anticuerpos policlonales murinos, ratones Balb/c de ocho semanas de edad se inmunizaron con una inyección intraperitoneal de 25 microgramos de cada péptido en 0,4 ml de una mezcla 1:1 se adyuvante completo de Freund (Difco Laboratories, Detroit, MI, Estados Unidos) y solución salina fisiológica (NaCl, 0,15 M). un mes más tarde los animales se volvieron a inmunizar con una inyección intramuscular de 35 microgramos de antígenos en adyuvante incompleto de Freund y solución salina (1:1) (0,2 ml se distribuyeron en cuatro sitios). Tres semanas después se administraron los péptidos en una dosis de 50 microgramos/ratón por vía intravenosa y se obtuvieron sueros 7 días después.

Para la producción de células B transformadas de EBV, se obtuvieron leucocitos de sangre a partir de personas control saans. Las células B se transformaron con EBV (virus de Epstein-Barr) usando protocolo convencional, y se mantuvieron las líneas celulares en RPMI 1640, suplementado con 10% de FCS (suero de ternera fetal). Se estimuló una alcuota de células durante 12 Horas con 10 mg/ml de fitohemaglutinina (PHA) para obtener células T obtenidas por mitógeno.

Se obtuvieron muestras de tejido se fetos nacidos muertos a los seis meses de edad de gestación. Se homogeneizaron hígado, bazo, timo y ganglios linfáticos fetales, los homogenatos se aclararon mediante centrifugación (20 000 rpm durante 20 minutos) y las muestras se usaron para análisis de transferencia de Western.

## ES 2 308 815 T3

Para el análisis de sueros policlonales, se realizaron análisis de Elisa y transferencia de Western. Placas de microvaloración de ELISA (Maxisorp, Nunc, Roskilde, Dinamarca) se revistieron con los péptidos (1 microgramos/pocillo en PBS, pH 7,5) a 4°C durante toda una noche y se bloquearon con 2% de BSA en PBS. Después las placas se incubaron con suero de ratón valorado y suero normal (control) a temperatura ambiente durante 4 h. Finalmente los anticuerpos específicos de péptido unidos se detectaron mediante el uso de las inmunoglobulinas marcadas con HRP de ratón (Dako A/S, Dinamarca) esencialmente como se ha descrito anteriormente [Ovod, V. A. *et al.*, AIDS 6 (1992) 25-34].

Para la transferencia de Western, homogenatos de tejidos, células B transformadas por EBV o células T activadas por PHA se sometieron a ebullición durante 10 minutos en tampón de muestra 2x (para homogenatos de tejido: 100 microlitros de homogenato mezclado con 100 microlitros de tampón de muestra; para las células: un millón de células/100 ml de tampón) y se analizaron en transferencia de Western como se describe en Ovod, V. A. *et al.*, *supra*.

Los antisueros así producidos se hicieron reaccionar con una baja cantidad de la proteína AIR-1 en bazo, timo y linfocitos fetales así como, en células b transformadas por EBV y en células T activadas por PHA. En el ensayo de ELISA a los péptidos inmunogénicos, los cuatro ratones proporcionaron una fuerte reactividad para el péptido usado para la inmunización. En el análisis de transferencia de Western que usa o bien los homogenatos de tejidos o células T estimuladas o células B establecidas, una banda fuerte de aproximadamente 60 kD de peso molecular se observó en hígado fetal (Fig. 6), mientras se observaron bandas más débiles del mismo tamaño en las otras muestras.

### Ejemplo 5

#### *Identificación de la expresión de APECED en timo y otros órganos linfoides*

La hibridación de ARNm *in situ* e inmunológica se usó para identificar las células que expresan APECED en diversos tejidos fetales y humanos de adultos normales. Las muestras de timo se obtuvieron junto con cirugía correctora a partir de pacientes de 2-19 años de edad. Otras muestras de tejido se obtuvieron a partir de biopsia quirúrgica o de otro material de biopsia. Esto fue aprobado por el Comité de Ética del Hospital en el Hospital de la Universidad de Tampere y Hospital Central de la Universidad de Helsinki. El material de tejido se almacenó congelado o se fijó con formaldehído y se incrustó en parafina hasta su uso.

Para la hibridación de ARNm *in situ*, tres fragmentos de ADNc para las ribosondas se amplificaron mediante RT-PCR a partir de ARNm de timo (Clontech) con pares de cebadores: 5'-ATG GCG ACG GAC GCG GCG CTA CGC-3' (SEQ ID NO: 27) y 5'-CCT GGA TGT ACT TCT TGG AGC CGC-3' (SEQ ID NO: 28), 5'-GAG CCC GAG GGG CCG TGG AGG GGA-3' (SEQ ID NO: 29) y 5'-GGC TGC ACC TCC TGG ACT GTT GCC-3' (SEQ ID NO:30), y 5'-GAT CCT GCT CAG GAG ACG TGA CCC-3' (SEQ ID NO: 31) y 5'-CAC CAG GCA AGG AGA GGC TCC CGG-3' (SEQ ID NO: 32). Diseñados para amplificar los fragmentos que se extienden sobre los nucleótidos 137-812, 738-1185 y 1554-2009 de la SEQ ID NO: 1, respectivamente. Los fragmentos amplificados se subclonaron en un vector PCRII-TOPO vector (Invitrogen).

Para la transcripción *in vitro* de los plásmidos se linealizaron y se sintetizaron sondas codificantes y no codificantes con digoxigenina-UTP como se ha descrito (Boehringer Mannheim Nonradioactive *in situ* Hybridization Application Manual). Las sondas marcadas se purificaron con columnas MicroSpinG-50 (Pharmacia Biotech). El tratamiento previo e hibridación se secciones de tejidos fijados con formaldehído, incrustados en parafina se realizaron como se ha descrito por H. Breitschopf y G. Sucharek. (Boehringer Mannheim Nonradioactive *in situ* Hybridization Application Manual, Detection of ARNm on paraffin embedded material of the central nervous system with DIG-labeled RNA probes, p 136-138.)

Para la preparación de los anticuerpos para la proteína AIR, el ADNc de APECED (secuencias 137-1774 de la SEQ ID NO: 1) que contenía una región codificante completa se amplificó a partir de ADNc de timo humano Marathon (Clontech) con los cebadores ExF y ExR2. Las secuencias de cebadores para ExF y ExR2 eran 5'-CCA CCC CAT GGC GAC GGA CG-3' (SEQ ID NO: 33) y 5'-GGA ATT CGG AGG GGA AGG GGG CCG CCG GA-3' (SEQ ID NO: 34) El ADNc amplificado se digirió con NcoI y EcoRI y se clonó (pHPA(RE) en el vector pET32a (Novagen). La proteína se expresó en *E. coli* y se purificó mediante His-tag como se describe por el fabricante (QiaExpress Type IV Kit, Cat No 32149, Qiagen, USA).

Para obtener anticuerpos policlonales murinos, ratones Balb/c se inmunizaron esencialmente como se ha descrito en el Ejemplo 4 usando 100 microgramos de la proteína AIR expresada bacteriológicamente con dosis de refuerzo de 25 y 25 microgramos.

Conejos blancos japoneses se inmunizaron con un péptido sintético que representa los aminoácidos 526-545 (DGILQWAIQSMARPAAPFPS, SEQ ID NO 36) de SEQ ID NO: 2 Las especificidades de los antisueros se comprobaron con ELISA y Transferencia de Western usando procedimientos convencionales.

Para la inmunocitoquímica, secciones congeladas de muestras de tejido se fijaron durante 20 min en paraformaldehído al 4%. El anticuerpo de AIR (conejo o ratón) en una dilución apropiada se incubó durante 30 min a 37°C, con un anticuerpo secundario de biotina conjugado de anti-ratón o anti-conejo (Vector, CA, USA). El anticuerpo biotinilado se reveló que mediante la incubación con Texas Red-avidin (Vector, CA, ESTADOS UNIDOS) durante 30 min a 37°C.

Con la hibridación *in situ*, se observó una señal positiva en unas pocas células en médula de timo (Fig. 7A). Las células positivas *in situ* de APECED eran infrecuentes y se dispersaron como células individuales en la médula, pero ocasionalmente una o dos células que APECED se observaron adyacentes a u ocultas en los corpúsculos de Hazla que representan conglomerados de células epiteliales medulares. En las células positivas, el ARNm de APECED se localizó predominantemente en el núcleo de las células. En los tejidos de ganglios linfáticos adultos humanos, las células infrecuentes expresaron el ARNm de APECED en la médula y ocasionalmente en la región paracortical, también (Fig. 7B) No se observó ninguna señal de hibridación en los centros germinales.

La inmunohistoquímica con antisueros policlonales de ratón y de conejo para la proteína AIR mostró una fuerte reactividad con las células seleccionadas en médula de timo, ganglios linfáticos e hígado fetal (Fig. 7C y 7D) la comparación del patrón de reacción obtenido mediante inmunohistoquímica con la obtenida mediante hibridación *in situ* estableció claramente que las células específicas, raras en médula de timo y médula de ganglios linfáticos y para cortex expresan el ARNm de APECED y la proteína AIR proteína. Mediante cualquier procedimiento, ni el ARNm ni la proteína se detectó en otros tejidos de adultos, incluyendo los órganos diana para la destrucción de tejidos en APECED (glándulas adrenales, glándulas paratiroides, gónadas). En los tejidos fetales humanos, se observaron células positivas de APECED, aunque extremadamente infrecuentes, en el estroma de las vellosidades coriónicas de la placenta y en el área sinusoidal del hígado . En el hígado fetal, las células positivas de APECED se localizaron a menudo como imágenes por pares en el espejo, lo que sugiere que las células se encuentran en mitosis. Las células raras que expresan APECED también se encontraron en timo fetal pero no se observó la expresión en otros tejidos fetales.

A nivel subcelular, la proteína AIR localizada como puntos nucleares pequeños en el timo adulto, proporcionando un patrón característico de puntos (Fig. 7C; y Fig. 8A y 8B), pero localizado en el citoplasma de las células en los ganglios linfáticos. En las células positivas raras en el hígado fetal, muchas de las cuales son mitóticas, la proteína de AIR se localiza en el citoplasma.

#### Ejemplo 6

##### *Caracterización del fenotipo de las células positivas de APECED en timo*

Se usó doble tinción de con los anticuerpos para caracterizar adicionalmente el tipo de célula que expresa el gen APECED. En vista del hecho que las células dendríticas (DC) y epitelio de timo están ambos implicados en la regulación de la maduración inmune, se estudió la expresión de marcadores para estas células.

Para la detección por inmunofluorescencia doble la tinción de AIR se realizó como se ha descrito en el in ejemplo 5 con suero anti conejo-AIR. Los portaobjetos después se incubaron con un segundo anticuerpo primario [AE1 (Neomarkers, CA, ESTADOS UNIDOS), AE3 (Neomarkers, CA, ESTADOS UNIDOS), CD11c (Immunotech, Francia), o CD83 (immunotech, Francia)] en una dilución apropiada durante 30 min a 37°C, y la reacción se reveló mediante incubación con un anticuerpo anti ratón secundario conjugado a FITC (Vector, CA, ESTADOS UNIDOS) durante 30 min a 37°C.

Los anticuerpos que reaccionan con citoqueratinas básicas de bajo peso molecular (AE1) o ácidas de alto peso molecular (AE3) tiñeron el timo de una manera reticular, y las células positivas de APECED se observaron o bien ocultas en su red o en estrecha yuxtaposición con las células positivas de queratina. La microscopía confocal demostró claramente que algunas de las células positivas de APECED eran positivas de citoqueratina mientras que algunas permanecían negativas (Figura 8 A). Una localización conjunta era más fuerte con AE1 que con AE3. La distribución de las células que expresan APECED epiteliales (AE1 positivo) y no epiteliales variaban pero en la mayoría de los preparados de timo más de la mitad eran epiteliales.

Menos de la mitad de las células que expresan APECED en timo teñidas con los marcadores CD11c y CD83 que reaccionan con las células de la línea celular monocito-macrófago-dendrítica. En la mayoría de los casos, la reacción de tinción era débil pero unas pocas células mostraron una tinción intensa con el marcador dado (Fig. 8B). CD83 se tiñó conjuntamente 5 a 40% de las células positivas de APECED El anticuerpo CD11c, se reseñó que era específico para las células dendríticas maduras, se hicieron reaccionar con hasta 5-10% de las células positivas de APECED. Sin embargo todas las células positivas de APECED eran fuertemente positivas para la tinción de HLA-DR, (datos no mostrados).

Estos resultados sugieren que en el timo el gen de APECED está de hecho expresado en dos poblaciones de células distintas, una epitelial y otra no epitelial. El último tipo de célula es probablemente el que expresa el gen de APECED en tejidos linfoides fuera del timo.

#### Ejemplo 7

##### *Expresión de APECED en células dendríticas *in vitro**

Para mostrar una expresión de APECED en células dendríticas derivadas de monocitos de sangre periférica que son precursores de DC, estas células se cultivaron en la presencia de citoquinas usando las condiciones que se conocen que conducen a la expresión de y maduración de células dendríticas.

## ES 2 308 815 T3

Se aislaron células mononucleares de sangre periférica mediante centrifugación de Ficoll-Hypaque, y se separaron las células adherentes y se cultivaron en la presencia de GM-CSF recombinante (1000 unidades/ml) y rhlL-4 (1000 unidades/ml, ambos de R&D Systems), como se ha descrito [Schuler, G. y Romani, N., Adv. Exp. Med. Biol. 417 (1997) 7-13]. Las células se cultivaron después durante tres días con 1/4 V/V de medio condicionado de macrófagos.

5 Las células se recogieron a intervalos de dos días y las muestras se prepararon para RT-PCR. Para RT-PCR se purificó ARN total a partir de DC mediante el uso de un kit comercial de Clontech (ESTADOS UNIDOS) (Nucleospin RNA Kit) de acuerdo con las instrucciones del fabricante. Se transfirió una alícuota de ARN en un ADNc con un kit comercial de Pharmacia (Suecia) (kit First-strand Synthesis) y se realizó la PCR para esta muestra de ADN. Para la PCR el fragmento se amplificó con los cebadores 5'-GAT CCT GCT CAG GAG ACG TGA CCC-3' (SEQ ID NO: 31; 1554-1577 de la SEQ ID NO: 1) y 5'-GGA CTG AGG AAG GAG GTG TCC TTC-3' (SEQ ID NO: 35; 1818-1841 de la SEQ ID NO: 1) con las siguientes condiciones: 35 ciclos de 95°C durante 1 min., 62°C durante 30 seg y 72°C durante 1 min. La mezcla de reacción 1x contenía 50 mM KCl, 10mM Tris-HCl, pH 8,3, 1,5 mM MgCl<sub>2</sub>, 0,001% (p/v) gelatina, 0,2 mM de cada uno de los dNTP, 0,25 U de Dynazyme (Finnzymes, Finlandia). Un fragmento de 287 pares de bases se detectó mediante electroforesis de azarosa al 1,5%.

15

Las preparaciones de citospina se realizaron además mediante inmunohistoquímica.

Durante este período de cultivo de 7 a 10 días aproximadamente la mitad de las células desarrolló la morfología vellosa característica de DC y sus marcadores celulares fenotípicos (CD11c y CD83) correspondían a los DC maduros (Figura 9). La expresión de APECED se estudió mediante RT-PCR e inmunocitoquímica a intervalos de tres días. En el material de partida, es decir, el conjunto de células adherentes de la sangre periférica, no se encontró ninguna expresión de APECED. Después de siete días de cultivo en la presencia de GM-CSF e IL-4, RT-PCR mostró la expresión de ARNm de APECED y la inmunofluorescencia mostró unas pocas manchas nucleares específicas AIR. Después un cultivo de tres días adicionales con medio acondicionado a partir de cultivos de macrófagos se observó un patrón de puntos fuerte de la expresión de AIR (Figura 9A). El análisis de RT-PCR del cultivo maduro (10 días) confirmó la expresión de la proteína AIR.

30

35

40

45

50

55

60

65

## ES 2 308 815 T3

### REIVINDICACIONES

- 5 1. Una secuencia de ADN aislada, **caracterizada** porque tiene la SEQ ID NO: 1 o su fragmento funcional que codifica una proteína asociada a poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED), o su forma mutada que incluye un gen defectuoso responsable APECED.
- 10 2. Una secuencia de ADN aislada de acuerdo con la reivindicación 1, **caracterizada** porque está asociada a las enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto.
- 15 3. Una secuencia de ADN aislada de acuerdo con la reivindicación 2, **caracterizada** porque está asociada a APECED.
- 20 4. Una secuencia de ADN de acuerdo con la reivindicación 1, **caracterizada** porque tiene la secuencia de acuerdo con la SEQ ID NO: 1 o su fragmento funcional que tiene la secuencia de acuerdo con la SEQ ID NO: 3 o SEQ ID NO: 5.
- 25 5. Una proteína, **caracterizada** porque tiene la SEQ ID NO: 2 de aminoácidos o su fragmento funcional asociada a poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED) o sus variantes asociadas a APECED.
- 30 6. Una proteína de acuerdo con la reivindicación 5, **caracterizada** porque está asociada a APECED.
- 35 7. Una proteína de acuerdo con la reivindicación 5 ó 6, **caracterizada** porque tiene la SEQ ID NO: 2 de aminoácidos, o su fragmento funcional que tiene la secuencia de acuerdo con la SEQ ID NO: 4, o su fragmento funcional que tiene la secuencia de acuerdo con la SEQ ID NO: 6.
- 40 8. Una proteína de acuerdo con cualquiera de las reivindicaciones 5 a 7, **caracterizada** porque tiene distintos motivos estructurales, incluyendo el motivo de dedo PHD (PHD), el motivo LXXLL (L), la región rica en prolina (PRR), y región rica en cistina (CRR).
- 45 9. Un procedimiento *in vitro* para la diagnosis de enfermedades relacionadas con la maduración y regulación inmune de respuesta inmune a auto y no auto, **caracterizado** por la detección en un espécimen biológico la presencia de un fragmento funcional de una secuencia de ADN que tiene la SEQ ID NO: 1 que codifica una proteína asociada a poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED), o una forma mutada de una secuencia de ADN que tiene la SEQ ID NO: 1 que incluye un gen defectuoso responsable de APECED.
- 50 10. Un procedimiento *in vitro* de acuerdo con la reivindicación 9, **caracterizado** porque la secuencia de ADN está asociada a APECED.
- 55 11. Un procedimiento *in vitro* de acuerdo con la reivindicación 9, **caracterizado** porque el gen defectuoso a detectar incluye una transición “C” a “T” que da como resultado la mutación no codificante “Arg” a “parada” no codificante en la posición de aminoácido 257 y/o una transversión “A” a “G” que da como resultado la mutación de codificación errónea “Lys” a “Glu” en la posición de aminoácido 42.
- 60 12. Un procedimiento *in vitro* de acuerdo con una cualquiera de las reivindicaciones 9 a 11, **caracterizado** porque ADN las técnicas se usan para la detección.
- 65 13. Un procedimiento *in vitro* de acuerdo con una cualquiera de las reivindicaciones 9 a 12, **caracterizado** porque la detección se aprovecha de TaqI u otra enzima que se escinde en el sitio de reconocimiento digestión de 5'-TCGA-3'.
14. Un procedimiento *in vitro* de acuerdo con una cualquiera de las reivindicaciones 9 a 13, **caracterizado** porque la enfermedad es poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED).
15. Un procedimiento *in vitro* para la diagnosis de enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto, **caracterizado** por la detección en un espécimen biológico de la ausencia de una proteína que tiene la SEQ ID NO: 2 o la presencia de su fragmento funcional asociado a poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED), o sus variantes asociadas a APECED.
16. Un procedimiento *in vitro* de acuerdo con la reivindicación 15, **caracterizado** porque tiene la secuencia de aminoácidos de la SEQ ID NO: 2, o su fragmento funcional que tiene la secuencia de acuerdo con la SEQ ID NO: 4, o su fragmento funcional que tiene la secuencia de acuerdo con la SEQ ID NO: 6.
17. Un procedimiento *in vitro* de acuerdo con la reivindicación 15 ó 16, **caracterizado** porque la proteína está asociada a poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED).
18. Un procedimiento *in vitro* de acuerdo con una cualquiera de las reivindicaciones 15 a 17, **caracterizado** porque la enfermedad es poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED).

## ES 2 308 815 T3

19. El uso de la secuencia de ADN *sequence* de acuerdo con una cualquiera de las reivindicaciones 1 a 4 en la diagnosis *in vitro* de las enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto.

5 20. El uso de la reivindicación 19, **caracterizado** porque la enfermedad es poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED).

10 21. El uso de la proteína de acuerdo con una cualquiera de las reivindicaciones 5 a 8 en la diagnosis *in vitro* de las enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto.

22. El uso de la reivindicación 21, **caracterizado** porque la enfermedad es poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED).

15 23. El uso de la secuencia de ADN de acuerdo con una cualquiera de las reivindicaciones 1 a 4 para la preparación de un medicamento útil en un procedimiento de terapia génica de las enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto.

20 24. El uso de la reivindicación 23, **caracterizado** porque la enfermedad es poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED).

25 25. El uso de la secuencia de ADN de acuerdo con una cualquiera de las reivindicaciones 1 a 4 para la preparación de un medicamento útil en el tratamiento de enfermedades relacionadas con la maduración y regulación inmune de la respuesta inmune a auto y no auto.

26. El uso de la reivindicación 25, **caracterizado** porque la enfermedad es poliendocrinopatía autoinmune-candidiasis-distrofia ectodérmica (APECED).

30 27. Anticuerpos que reaccionan de manera específica con la secuencia de ADN de acuerdo con una cualquiera de las reivindicaciones 1 a 4 o la proteína de una cualquiera de las reivindicaciones 5 a 8.

35

40

45

50

55

60

65

# Cromosoma humano 21

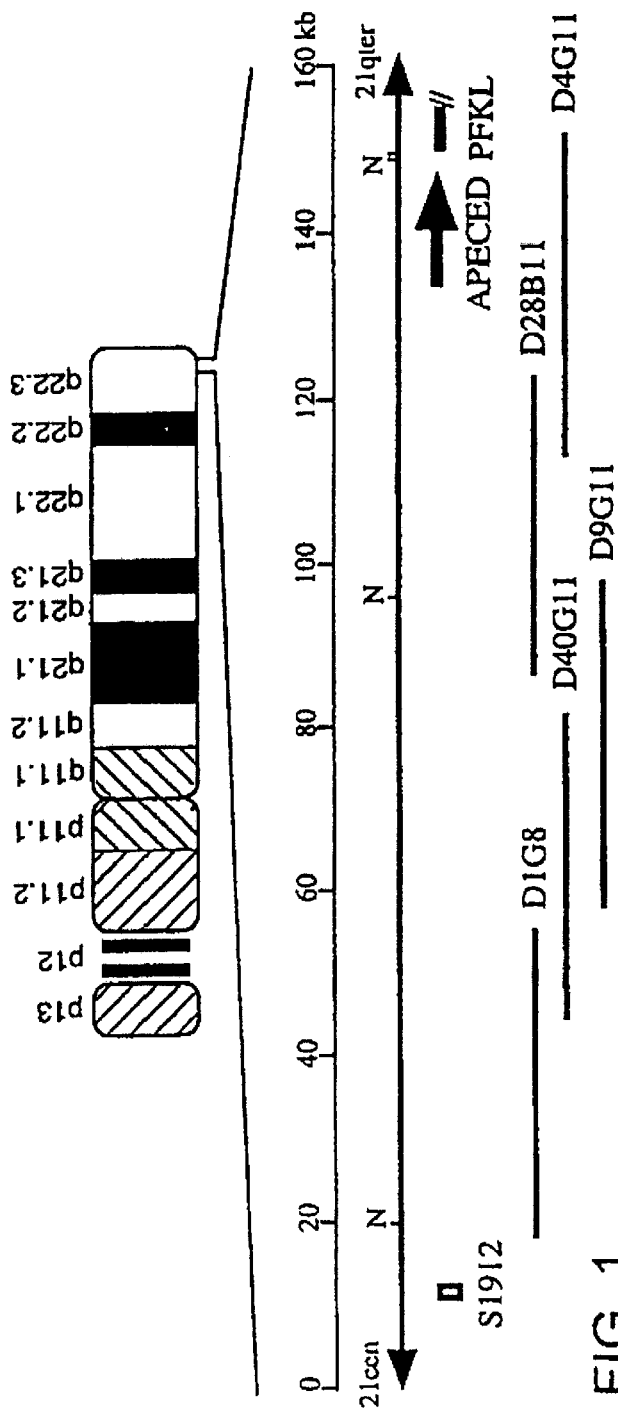


FIG. 1

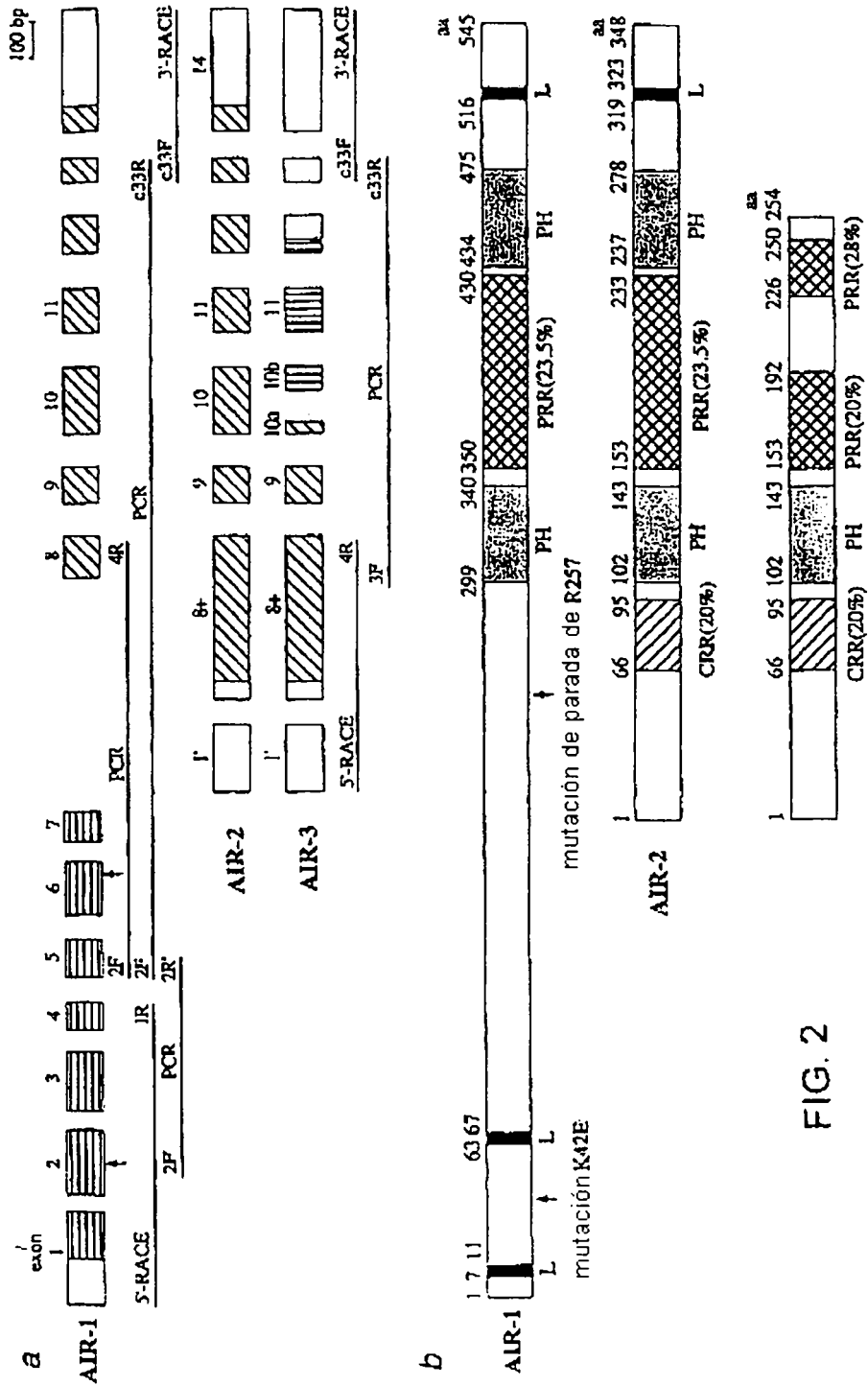


FIG. 2

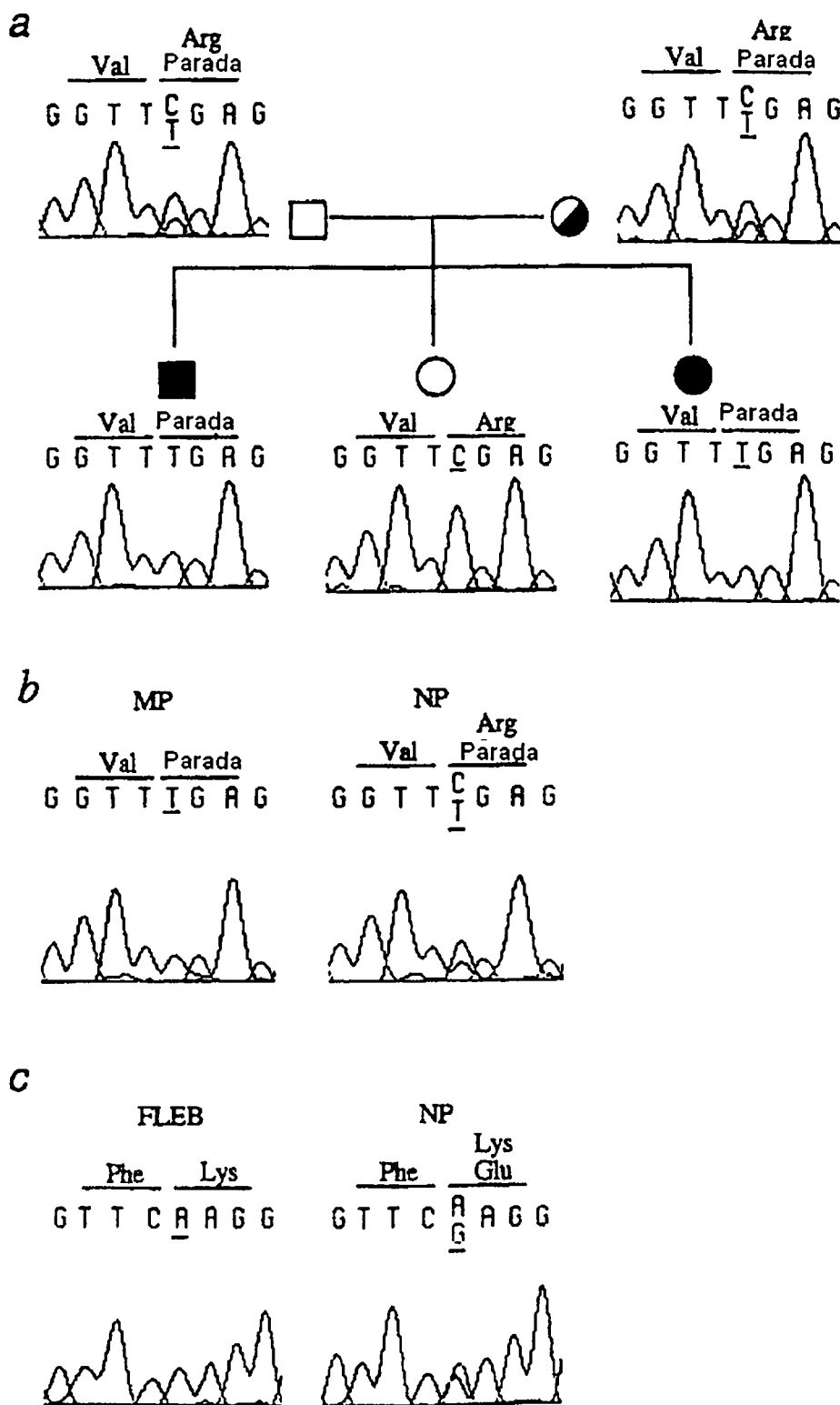


FIG. 3

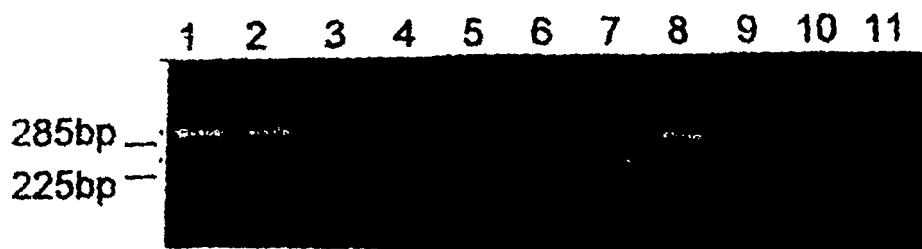


FIG. 4

|                 |     |  |     |
|-----------------|-----|--|-----|
| AIR-1:          | 299 | CAVCRDGGELICCDGCPRAFHLACLSPPLREIPSGTWRCSSC | 340 |
| AIR-1:          | 434 | .G..G..TDVLR.TH.AA...WR.HF.AGTSR.GTGL..R.. | 475 |
| Mi-2 :          | 373 | .E..QQ...I.L..T...Y.MV..D.DMEKA.E.K.S.PH.  | 414 |
| Mi-2 :          | 452 | .R..K.....T..SSY.IH..N...P...N.E.L.PR.     | 493 |
| TIF1 :          | 791 | ....QN.....EK..KV...S.HV.T.TNF...E.I.TF.   | 832 |
| <u>consenso</u> |     | C C C C H C C C C                          |     |

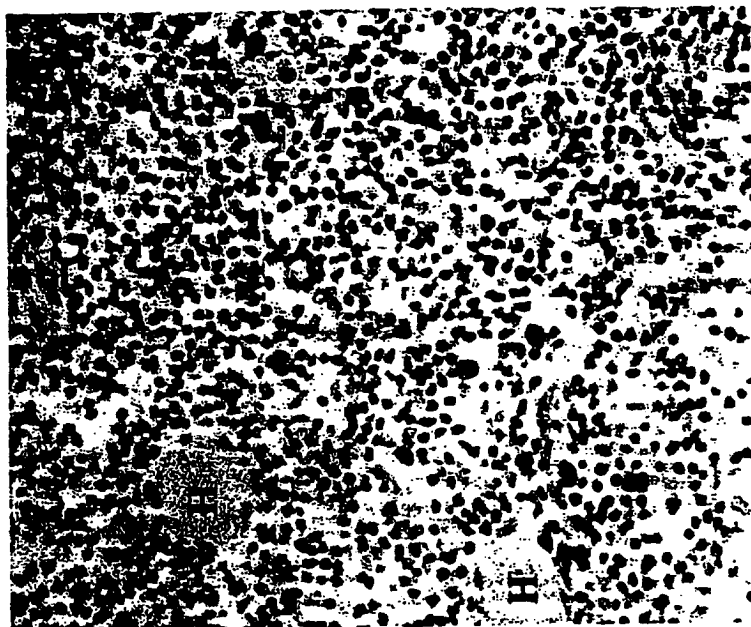
FIG. 5

**1 2 3 4 5 6**



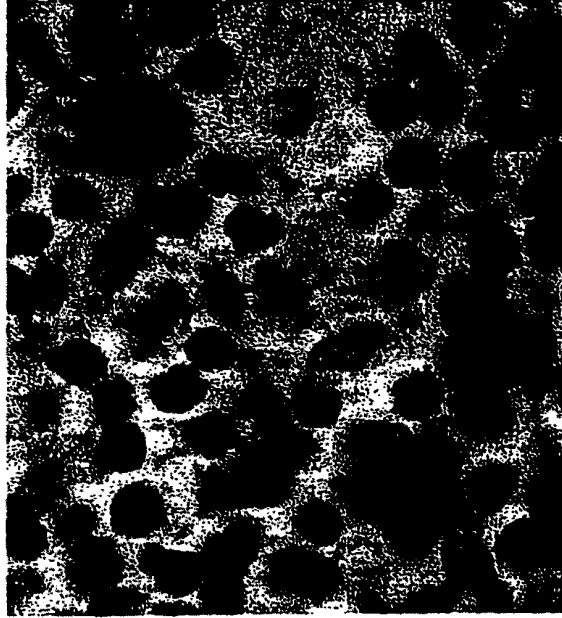
**FIG. 6**

**FIG. 7B**



**FIG. 7A**

**FIG. 7D**



**FIG. 7C**

**FIG. 8A**



**FIG. 8B**





**FIG. 9**

# ES 2 308 815 T3

## LISTA DE SECUENCIAS

### (1) INFORMACIÓN GENERAL:

- 5 (i) SOLICITANTE:
- (A) NOMBRE: Kai Krohn *et al.*
  - (B) CALLE: Iltarusko, Salmentaantie 751
  - 10 (C) CIUDAD: 36450 Salmentaka
  - (E) PAÍS: Finlandia
  - (F) CÓDIGO POSTAL (ZIP): ninguno
- 15 (ii) TÍTULO DE LA INVENCIÓN: gen novedoso
- (iii) NÚMERO DE SECUENCIAS: 26
- (iv) FORMA LEGIBLE DE ORDENADOR:
- 20 (A) TIPO DE MEDIO: Disco flexible
  - (B) ORDENADOR: IBM PC compatible
  - (C) SISTEMA OPERATIVO: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release nº 1.0, Version nº 1.30 (EPO)

### (2) INFORMACIÓN PARA LA SEQ ID NO: 1:

- 30 (i) CARACTERÍSTICAS DE SECUENCIA:
- (A) LONGITUD: 2036 pares de bases
  - (B) TIPO: ácido nucleico
  - (C) TIPO DE HEBRA: individual
  - (D) TOPOLOGÍA: lineal
- 35 (ix) CARACTERÍSTICA:
- (A) NOMBRE/CLAVE: CDS
  - (B) LOCALIZACIÓN: 137.. 1774
  - 40 (D) OTRA INFORMACIÓN: /producto = "AIR-1"
- (ix) CARACTERÍSTICA:
- (A) NOMBRE/CLAVE: mat\_péptido
  - 45 (B) LOCALIZACIÓN: 137..1771
  - (D) OTRA INFORMACIÓN: /producto = "AIR-1"
- (xi) DESCRIPCIÓN DE SECUENCIAS: SEQ ID NO: 1:

50 **AGACCGGGGA GACGGGCGGG CGCACAGCCG GCGCGGAGGC CCCACAGCCC CGCCGGGACC 60**

**CGAGGCCAAG CGAGGGGCTG CCAGTGTCCC GGGACCCACC GCGTCCGCCC CAGCCCCGGG 120**

55 **TCCCCGCGCC CACCCC ATG GCG ACG GAC GCG GCG CTA CGC CGG CTT CTG 169**

**Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu**

**1 5 10**

60

65

# ES 2 308 815 T3

|    |  |     |
|----|--|-----|
| 5  | AGG CTG CAC CGC ACG GAG ATC GCG GTG GCC GTG GAC AGC GCC TTC CCA<br>Arg Leu His Arg Thr Glu Ile Ala Val Ala Val Asp Ser Ala Phe Pro<br>15                                  20                                  25   | 217 |
| 10 | CTG CTG CAC GCG CTG GCT GAC CAC GAC GTG GTC CCC GAG GAC AAG TTT<br>Leu Leu His Ala Leu Ala Asp His Asp Val Val Pro Glu Asp Lys Phe<br>30                                  35                                  40   | 265 |
| 15 | CAG GAG ACG CTT CAT CTG AAG GAA AAG GAG GGC TGC CCC CAG GCC TTC<br>Gln Glu Thr Leu His Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe<br>45                                  50                                  55   | 313 |
| 20 | CAC GCC CTC CTG TCC TGG CTG CTG ACC CAG GAC TCC ACA GCC ATC CTG<br>His Ala Leu Leu Ser Trp Leu Leu Thr Gln Asp Ser Thr Ala Ile Leu<br>60                                  65                                  70                                  75     | 361 |
| 25 | GAC TTC TGG AGG GTG CTG TTC AAG GAC TAC AAC CTG GAG CGC TAT GGC<br>Asp Phe Trp Arg Val Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Gly<br>80                                  85                                  90   | 409 |
| 30 | CGG CTG CAG CCC ATC CTG GAC AGC TTC CCC AAA GAT GTG GAC CTC AGC<br>Arg Leu Gln Pro Ile Leu Asp Ser Phe Pro Lys Asp Val Asp Leu Ser<br>95                                  100                                  105                                       | 457 |
| 35 | CAG CCC CGG AAG GGG AGG AAG CCC CCG GCC GTC CCC AAG GCT TTG GTA<br>Gln Pro Arg Lys Gly Arg Lys Pro Pro Ala Val Pro Lys Ala Leu Val<br>110                                  115                                  120                                      | 505 |
| 40 | CCG CCA CCC AGA CTC CCC ACC AAG AGG AAG GCC TCA GAA GAG GCT CGA<br>Pro Pro Pro Arg Leu Pro Thr Lys Arg Lys Ala Ser Glu Glu Ala Arg<br>125                                  130                                  135                                      | 553 |
| 45 | GCT GCC GCG CCA GCA GCC CTG ACT CCA AGG GGC ACC GCC AGC CCA GGC<br>Ala Ala Ala Pro Ala Ala Leu Thr Pro Arg Gly Thr Ala Ser Pro Gly<br>140                                  145                                  150                                  155 | 601 |
| 50 | TCT CAA CTG AAG GCC AAG CCC CCC AAG AAG CCG GAG AGC AGC GCA GAG<br>Ser Gln Leu Lys Ala Lys Pro Pro Lys Lys Pro Glu Ser Ser Ala Glu<br>160                                  165                                  170                                      | 649 |
| 55 | CAG CAG CGC CTT CCA CTC GGG AAC GGG ATT CAG ACC ATG TCA GCT TCA<br>Gln Gln Arg Leu Pro Leu Gly Asn Gly Ile Gln Thr Met Ser Ala Ser<br>175                                  180                                  185                                      | 697 |
| 60 | GTC CAG AGA GCT GTG GCC ATG TCC TCC GGG GAC GTC CCG GGA GCC CGA<br>Val Gln Arg Ala Val Ala Met Ser Ser Gly Asp Val Pro Gly Ala Arg<br>190                                  195                                  200                                      | 745 |
| 65 | GGG GCC GTG GAG GGG ATC CTC ATC CAG CAG GTG TTT GAG TCA GGC GGC<br>Gly Ala Val Glu Gly Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Gly<br>205                                  210                                  215                                      | 793 |
| 70 | TCC AAG AAG TGC ATC CAG GTT GGC GGG GAG TTC TAC ACT CCC AGC AAG<br>Ser Lys Lys Cys Ile Gln Val Gly Gly Glu Phe Tyr Thr Pro Ser Lys<br>220                                  225                                  230                                  235 | 841 |



ES 2 308 815 T3

CTG CGG TGT ACT CAC TGC GCC GCT GCC TTC CAC TGG CGC TGC CAC TTC 1513  
 Leu Arg Cys Thr His Cys Ala Ala Ala Phe His Trp Arg Cys His Phe  
 445 450 455

5 CCA GCC GGC ACC TCC CGG CCC GGG ACG GGC CTG CGC TGC AGA TCC TGC 1561  
 Pro Ala Gly Thr Ser Arg Pro Gly Thr Gly Leu Arg Cys Arg Ser Cys  
 460 465 470 475

10 TCA GGA GAC GTG ACC CCA GCC CCT GTG GAG GGG GTG CTG GCC CCC AGC 1609  
 Ser Gly Asp Val Thr Pro Ala Pro Val Glu Gly Val Leu Ala Pro Ser  
 480 485 490

15 CCC GCC CGC CTG GCC CCT GGG CCT GCC AAG GAT GAC ACT GCC AGT CAC 1657  
 Pro Ala Arg Leu Ala Pro Gly Pro Ala Lys Asp Asp Thr Ala Ser His  
 495 500 505

20 GAG CCC GCT CTG CAC AGG GAT GAC CTG GAG TCC CTT CTG AGC GAG CAC 1705  
 Glu Pro Ala Leu His Arg Asp Asp Leu Glu Ser Leu Leu Ser Glu His  
 510 515 520

25 ACC TTC GAT GGC ATC CTG CAG TGG GCC ATC CAG AGC ATG GCC CGT CCG 1753  
 Thr Phe Asp Gly Ile Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro  
 525 530 535

30 GCG GCC CCC TTC CCC TCC TGA CCCCAGATGG CCGGGACATG CAGCTCTGAT 1804  
 Ala Ala Pro Phe Pro Ser \*

35 GAGAGAGTGC TGAGAAGGAC ACCTCCTTCC TCAGTCCTGG AAGCCGGCCG GCTGGGATCA 1864

40 AGAAGGGGAC AGCGCCACCT CTGTGTCAGTG CTCGGCTGTA AACAGCTCTG TGTCTCTGGG 1924

GACACCAGCC ATCATGTGCC TGGAAATTAA ACCCTGCCCC ACTTCTCTAC TCTGGAAGTC 1984

CCC GGGAGCC TCTCCTTGCC TGGTGACCTA CTAAAAATAT AAAAATTAGC TG 2036

45 (2) INFORMACIÓN PARA LA SEQ ID NO: 2:

(i) CARACTERÍSTICAS DE SECUENCIA:

- (A) LONGITUD: 545 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

55 (xi) DESCRIPCIÓN DE SECUENCIAS: SEQ ID NO: 2:

Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu Arg Leu His Arg Thr  
 1 5 10 15

60 Glu Ile Ala Val Ala Val Asp Ser Ala Phe Pro Leu Leu His Ala Leu  
 20 25 30

Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln Glu Thr Leu His  
 35 40 45

# ES 2 308 815 T3

|    |   |     |     |     |
|----|---|-----|-----|-----|
|    | Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His Ala Leu Leu Ser | 50  | 55  | 60  |
| 5  | Trp Leu Leu Thr Gln Asp Ser Thr Ala Ile Leu Asp Phe Trp Arg Val | 65  | 70  | 75  |
|    |   |     |     | 80  |
| 10 | Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Gly Arg Leu Gln Pro Ile |     | 85  | 90  |
|    |   |     |     | 95  |
| 15 | Leu Asp Ser Phe Pro Lys Asp Val Asp Leu Ser Gln Pro Arg Lys Gly | 100 | 105 | 110 |
|    |   |     |     |     |
| 20 | Arg Lys Pro Pro Ala Val Pro Lys Ala Leu Val Pro Pro Pro Arg Leu | 115 | 120 | 125 |
|    |   |     |     |     |
| 25 | Pro Thr Lys Arg Lys Ala Ser Glu Glu Ala Arg Ala Ala Ala Pro Ala | 130 | 135 | 140 |
|    |   |     |     |     |
| 30 | Ala Leu Thr Pro Arg Gly Thr Ala Ser Pro Gly Ser Gln Leu Lys Ala | 145 | 150 | 155 |
|    |   |     |     | 160 |
| 35 | Lys Pro Pro Lys Lys Pro Glu Ser Ser Ala Glu Gln Gln Arg Leu Pro | 165 | 170 | 175 |
|    |   |     |     |     |
| 40 | Leu Gly Asn Gly Ile Gln Thr Met Ser Ala Ser Val Gln Arg Ala Val | 180 | 185 | 190 |
|    |   |     |     |     |
| 45 | Ala Met Ser Ser Gly Asp Val Pro Gly Ala Arg Gly Ala Val Glu Gly | 195 | 200 | 205 |
|    |   |     |     |     |
| 50 | Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Gly Ser Lys Lys Cys Ile | 210 | 215 | 220 |
|    |   |     |     |     |
| 55 | Gln Val Gly Gly Glu Phe Tyr Thr Pro Ser Lys Phe Glu Asp Ser Gly | 225 | 230 | 235 |
|    |   |     |     | 240 |
| 60 | Ser Gly Lys Asn Lys Ala Arg Ser Ser Ser Gly Pro Lys Pro Leu Val | 245 | 250 | 255 |
|    |   |     |     |     |
| 65 | Arg Ala Lys Gly Ala Gln Gly Ala Ala Pro Gly Gly Gly Glu Ala Arg | 260 | 265 | 270 |
|    |   |     |     |     |
| 70 | Leu Gly Gln Gln Gly Ser Val Pro Ala Pro Leu Ala Leu Pro Ser Asp | 275 | 280 | 285 |
|    |   |     |     |     |
| 75 | Pro Gln Leu His Gln Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp | 290 | 295 | 300 |
|    |   |     |     |     |
| 80 | Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu | 305 | 310 | 315 |
|    |   |     |     | 320 |
| 85 | Ala Cys Leu Ser Pro Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg | 325 | 330 | 335 |

ES 2 308 815 T3

Cys Ser Ser Cys Leu Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala  
 340 345 350  
 5  
 Glu Glu Pro Arg Pro Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro  
 355 360 365  
 10  
 Gly Leu Arg Ser Ala Gly Glu Glu Val Arg Gly Pro Pro Gly Glu Pro  
 370 375 380  
 15  
 Leu Ala Gly Met Asp Thr Thr Leu Val Tyr Lys His Leu Pro Ala Pro  
 385 390 395 400  
 20  
 Pro Ser Ala Ala Pro Leu Pro Gly Leu Asp Ser Ser Ala Leu His Pro  
 405 410 415  
 25  
 Leu Leu Cys Val Gly Pro Glu Gly Gln Gln Asn Leu Ala Pro Gly Ala  
 420 425 430  
 30  
 Arg Cys Gly Val Cys Gly Asp Gly Thr Asp Val Leu Arg Cys Thr His  
 435 440 445  
 35  
 Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro Ala Gly Thr Ser  
 450 455 460  
 40  
 Arg Pro Gly Thr Gly Leu Arg Cys Arg Ser Cys Ser Gly Asp Val Thr  
 465 470 475 480  
 45  
 Pro Ala Pro Val Glu Gly Val Leu Ala Pro Ser Pro Ala Arg Leu Ala  
 485 490 495  
 50  
 Pro Gly Pro Ala Lys Asp Asp Thr Ala Ser His Glu Pro Ala Leu His  
 500 505 510  
 55  
 Arg Asp Asp Leu Glu Ser Leu Leu Ser Glu His Thr Phe Asp Gly Ile  
 515 520 525  
 60  
 Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro  
 530 535 540  
 Ser \*  
 545

(2) INFORMACIÓN PARA LA SEQ ID NO: 3:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 1545 pares de bases

(B) TIPO: ácido nucleico

# ES 2 308 815 T3

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 237..1283

(D) OTRA INFORMACIÓN: /producto= "AIR-\_2"

(ix) CARACTERÍSTICA:

(A) NOMBRE/CLAVE: mat\_péptido

(B) LOCALIZACIÓN: 237..1280

(D) OTRA INFORMACIÓN: /producto = "AIR-2"

(xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 3:

|    |   |     |
|----|---|-----|
| 20 | AGAGAAAGTG AGGTCTTCTC AGGCTCTTAA GAGCATGGCG TTTGGTCCAG GCTGTACCCG   | 60  |
| 25 | CTGCTCTCAG CTGGGCCCGT GGGTGGGCCG GCGCCCCCTG CTATAGCCAG GAGGTCAAGG   | 120 |
| 30 | ATCCACTGGG AATGCCATGC TCATCTTTCG TCCCCAGCAT GGTTTCTTAA TGGGGTAGAA   | 180 |
| 35 | GCAGGTCGGG AGAGACCTCC CTGGGCCTGG CCCCCTGCC CTGTGAGGAA GGGTTC  | 236 |
| 40 | ATG TGG TTG GTG TAC AGT TCC GGG GCC CCT GGA ACG CAG CAG CCT GCA<br>Met Trp Leu Val Tyr Ser Ser Gly Ala Pro Gly Thr Gln Gln Pro Ala<br>1 5 10 15   | 284 |
| 45 | AGA AAC CGG GTT TTC TTC CCA ATA GGG ATG GCC CCG GGG GGT GTC TGT<br>Arg Asn Arg Val Phe Phe Pro Ile Gly Met Ala Pro Gly Gly Val Cys<br>20 25 30    | 332 |
| 50 | TGG AGA CCA GAT GGA TGG GGA ACA GGT GGT CAG GGC AGA ATT TCA GGC<br>Trp Arg Pro Asp Gly Trp Gly Thr Gly Gly Gln Gly Arg Ile Ser Gly<br>35 40 45    | 380 |
| 55 | CCT GGC AGC ATG GGA GCA GGG CAG AGA CTG GGG AGT TCA GGT ACC CAG<br>Pro Gly Ser Met Gly Ala Gly Gln Arg Leu Gly Ser Ser Gly Thr Gln<br>50 55 60    | 428 |
| 60 | AGA TGC TGC TGG GGG AGC TGT TTT GGG AAG GAG GTG GCT CTC AGG AGG<br>Arg Cys Cys Trp Gly Ser Cys Phe Gly Lys Glu Val Ala Leu Arg Arg<br>65 70 75 80 | 476 |
| 65 | GTG CTG CAC CCC AGC CCA GTC TGC ATG GGC GTC TCT TGC CTG TGC CAG<br>Val Leu His Pro Ser Pro Val Cys Met Gly Val Ser Cys Leu Cys Gln<br>85 90 95    | 524 |
| 70 | AAG AAT GAG GAC GAG TGT GCC GTG TGT CGG GAC GGC GGG GAG CTC ATC<br>Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp Gly Gly Glu Leu Ile<br>100 105 110 | 572 |
| 75 | TGC TGT GAC GGC TGC CCT CGG GCC TTC CAC CTG GCC TGC CTG TCC CCT<br>Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu Ala Cys Leu Ser Pro<br>115 120 125 | 620 |

# ES 2 308 815 T3

|    |   |  |
|----|---|--|
| 5  | <p>CCG CTC CGG GAG ATC CCC AGT GGG ACC TGG AGG TGC TCC AGC TGC CTG 668<br/>           Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg Cys Ser Ser Cys Leu<br/>           130 135 140</p>      |  |
| 10 | <p>CAG GCA ACA GTC CAG GAG GTG CAG CCC CGG GCA GAG GAG CCC CGG CCC 716<br/>           Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala Glu Glu Pro Arg Pro<br/>           145 150 155 160</p>  |  |
| 15 | <p>GGA GAG GAG GTA AGA GGT CCA CCT GGG GAA CCC CTA GCC GGC ATG GAC 812<br/>           Gly Glu Glu Val Arg Gly Pro Pro Gly Glu Pro Leu Ala Gly Met Asp<br/>           180 185 190</p>      |  |
| 20 | <p>ACG ACT CTT GTC TAC AAG CAC CTG CCG GCT CCG CCT TCT GCA GCC CCG 860<br/>           Thr Thr Leu Val Tyr Lys His Leu Pro Ala Pro Pro Ser Ala Ala Pro<br/>           195 200 205</p>      |  |
| 25 | <p>CTG CCA GGG CTG GAC TCC TCG GCC CTG CAC CCC CTA CTG TGT GTG GGT 908<br/>           Leu Pro Gly Leu Asp Ser Ser Ala Leu His Pro Leu Leu Cys Val Gly<br/>           210 215 220</p>      |  |
| 30 | <p>CCT GAG GGT CAG CAG AAC CTG GCT CCT GGT GCG CGT TGC GGG GTG TGC 956<br/>           Pro Glu Gly Gln Gln Asn Leu Ala Pro Gly Ala Arg Cys Gly Val Cys<br/>           225 230 235 240</p>  |  |
| 35 | <p>GGA GAT GGT ACG GAC GTG CTG CGG TGT ACT CAC TGC GCC GCT GCC TTC 1004<br/>           Gly Asp Gly Thr Asp Val Leu Arg Cys Thr His Cys Ala Ala Phe<br/>           245 250 255</p>         |  |
| 40 | <p>CAC TGG CGC TGC CAC TTC CCA GCC GGC ACC TCC CGG CCC GGG ACG GGC 1052<br/>           His Trp Arg Cys His Phe Pro Ala Gly Thr Ser Arg Pro Gly Thr Gly<br/>           260 265 270</p>     |  |
| 45 | <p>CTG CGC TGC AGA TCC TGC TCA GGA GAC GTG ACC CCA GCC CCT GTG GAG 1100<br/>           Leu Arg Cys Arg Ser Cys Ser Gly Asp Val Thr Pro Ala Pro Val Glu<br/>           275 280 285</p>     |  |
| 50 | <p>GGG GTG CTG GCC CCC AGC CCC GCC CGC CTG GCC CCT GGG CCT GCC AAG 1148<br/>           Gly Val Leu Ala Pro Ser Pro Ala Arg Leu Ala Pro Gly Pro Ala Lys<br/>           290 295 300</p>     |  |
| 55 | <p>GAT GAC ACT GCC AGT CAC GAG CCC GCT CTG CAC AGG GAT GAC CTG GAG 1196<br/>           Asp Asp Thr Ala Ser His Glu Pro Ala Leu His Arg Asp Asp Leu Glu<br/>           305 310 315 320</p> |  |
| 60 | <p>TCC CTT CTG AGC GAG CAC ACC TTC GAT GGC ATC CTG CAG TGG GCC ATC 1244<br/>           Ser Leu Leu Ser Glu His Thr Phe Asp Gly Ile Leu Gln Trp Ala Ile<br/>           325 330 335</p>     |  |
| 65 | <p>CAG AGC ATG GCC CGT CCG GCG GCC CCC TTC CCC TCC TGA CCCAGATGG 1293<br/>           Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro Ser *<br/>           340 345</p>                         |  |

ES 2 308 815 T3

CCGGGACATG CAGCTCTGAT GAGAGAGTGC TGAGAAGGAC ACCTCCTTCC TCAGTCCTGG 1353  
 AAGCCGGCCG GCTGGGATCA AGAAGGGGAC AGCGCCACCT CTTGTCAGTG CTCGGCTGTA 1413  
 AACAGCTCTG TGTTCCTGGG GACACCAGCC ATCATGTGCC TGGAAATTAA ACCCTGCCCC 1473  
 ACTTCTCTAC TCTGGAAGTC CCCGGGAGCC TCTCCTTGCC TGGTGACCTA CTAAAAATAT 1533  
 AAAAATTAGC TG 1545

(2) INFORMACIÓN PARA LA SEQ ID NO: 4:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 348 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 4:

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

ES 2 308 815 T3

Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro Gly Leu Arg Ser Ala  
 165 170 175  
 5 Gly Glu Glu Val Arg Gly Pro Pro Gly Glu Pro Leu Ala Gly Met Asp  
 180 185 190  
 10 Thr Thr Leu Val Tyr Lys His Leu Pro Ala Pro Pro Ser Ala Ala Pro  
 195 200 205  
 15 Leu Pro Gly Leu Asp Ser Ser Ala Leu His Pro Leu Leu Cys Val Gly  
 210 215 220  
 20 Pro Glu Gly Gln Gln Asn Leu Ala Pro Gly Ala Arg Cys Gly Val Cys  
 225 230 235 240  
 25 Gly Asp Gly Thr Asp Val Leu Arg Cys Thr His Cys Ala Ala Ala Phe  
 245 250 255  
 30 His Trp Arg Cys His Phe Pro Ala Gly Thr Ser Arg Pro Gly Thr Gly  
 260 265 270  
 35 Leu Arg Cys Arg Ser Cys Ser Gly Asp Val Thr Pro Ala Pro Val Glu  
 275 280 285  
 40 Gly Val Leu Ala Pro Ser Pro Ala Arg Leu Ala Pro Gly Pro Ala Lys  
 290 295 300  
 45 Asp Asp Thr Ala Ser His Glu Pro Ala Leu His Arg Asp Asp Leu Glu  
 305 310 315 320  
 50 Ser Leu Leu Ser Glu His Thr Phe Asp Gly Ile Leu Gln Trp Ala Ile  
 325 330 335  
 Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro Ser \*  
 340 345

(2) INFORMACIÓN PARA LA SEQ ID NO: 5:

55 (i) CARACTERÍSTICAS DE SECUENCIA:

- (A) LONGITUD: 1463 pares de bases
- (B) TIPO: ácido nucleico
- 60 (C) TIPO DE HEBRA: individual
- (D) TOPOLOGÍA: lineal

(ix) CARACTERÍSTICA:

- 65 (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 237..1001
- (D) OTRA INFORMACIÓN:/producto = "AIR-3"





ES 2 308 815 T3

Met Trp Leu Val Tyr Ser Ser Gly Ala Pro Gly Thr Gln Gln Pro Ala  
 1 5 10 15  
 5 Arg Asn Arg Val Phe Phe Pro Ile Gly Met Ala Pro Gly Gly Val Cys  
 20 25 30  
 10 Trp Arg Pro Asp Gly Trp Gly Thr Gly Gly Gln Gly Arg Ile Ser Gly  
 35 40 45  
 15 Pro Gly Ser Met Gly Ala Gly Gln Arg Leu Gly Ser Ser Gly Thr Gln  
 50 55 60  
 20 Arg Cys Cys Trp Gly Ser Cys Phe Gly Lys Glu Val Ala Leu Arg Arg  
 65 70 75 80  
 25 Val Leu His Pro Ser Pro Val Cys Met Gly Val Ser Cys Leu Cys Gln  
 85 90 95  
 30 Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp Gly Gly Glu Leu Ile  
 100 105 110  
 35 Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu Ala Cys Leu Ser Pro  
 115 120 125  
 40 Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg Cys Ser Ser Cys Leu  
 130 135 140  
 45 Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala Glu Glu Pro Arg Pro  
 145 150 155 160  
 50 Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro Gly Leu Arg Ser Ala  
 165 170 175  
 55 Gly Glu Glu Pro Arg Cys Gln Gly Trp Thr Pro Arg Pro Cys Thr Pro  
 180 185 190  
 60 Tyr Cys Val Trp Val Leu Arg Val Ser Arg Thr Trp Leu Leu Val Arg  
 195 200 205  
 65 Val Ala Gly Cys Ala Glu Met Val Arg Thr Cys Cys Gly Val Leu Thr  
 210 215 220  
 Ala Pro Leu Pro Ser Thr Gly Ala Ala Thr Ser Gln Pro Ala Pro Pro  
 225 230 235 240  
 Gly Pro Gly Arg Ala Cys Ala Ala Asp Pro Ala Gln Glu Thr \*  
 245 250 255

## ES 2 308 815 T3

### (2) INFORMACIÓN PARA LA SEQ ID NO: 7:

#### (i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

#### (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 7:

GATGACTG CCAGTCACGA

20

### (2) INFORMACIÓN PARA LA SEQ ID NO: 8:

#### (i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 22 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

#### (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 8:

GTTCCGAGT GGAAGGCGCT GC

22

### (2) INFORMACIÓN PARA LA SEQ ID NO: 9:

#### (i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

#### (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 9:

AGGGGACAGG CAGGCCAGGT

20

### (2) INFORMACIÓN PARA LA SEQ ID NO: 10:

#### (i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 24 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

#### (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 10:

GAGTTCAGGT ACCCAGAGAT GCTG

24

### (2) INFORMACIÓN PARA LA SEQ ID NO: 11:

#### (i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

## ES 2 308 815 T3

(xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 11:

CTCGCTCAGA AGGGACTCCA

20

5 (2) INFORMACIÓN PARA LA SEQ ID NO: 12:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 23 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

10 (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 12:

GGATTCAGAC CATGTCAGCT TCA

23

20 (2) INFORMACIÓN PARA LA SEQ ID NO: 13:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 21 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

25 (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 13:

GTGCTGTTCA AGGACTACAA C

21

30 (2) INFORMACIÓN PARA LA SEQ ID NO: 14:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 21 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

35 (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 14:

TGGATGAGGA TCCCCTCCAC G

21

40 (2) INFORMACIÓN PARA LA SEQ ID NO: 15:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 27 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

45 (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 15:

CCATCCTAAT ACGACTCACT ATAGGGC

27

50 (2) INFORMACIÓN PARA LA SEQ ID NO: 16:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 20 pares de bases

## ES 2 308 815 T3

- (B) TIPO: ácido nucleico
- (C) TIPO DE HEBRA: individual
- (D) TOPOLOGÍA: lineal

5 (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 16:

TGCAGGCTGT GGGAACTCCA

20

10 (2) INFORMACIÓN PARA LA SEQ ID NO: 17:

(i) CARACTERÍSTICAS DE SECUENCIA:

- (A) LONGITUD: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE HEBRA: individual
- (D) TOPOLOGÍA: lineal

20 (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 17:

AGAAAAGAG CTGTACCCTG TG

22

25 (2) INFORMACIÓN PARA LA SEQ ID NO: 18:

(i) CARACTERÍSTICAS DE SECUENCIA:

- (A) LONGITUD: 22 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE HEBRA: individual
- (D) TOPOLOGÍA: lineal

35 (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 18:

TGCAAGGAAG AGGGGCGTCA GC

22

40 (2) INFORMACIÓN PARA LA SEQ ID NO: 19:

(i) CARACTERÍSTICAS DE SECUENCIA:

- (A) LONGITUD: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE HEBRA: individual
- (D) TOPOLOGÍA: lineal

50 (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 19:

TCCACCACAA GCCGAGGAGA T

21

55 (2) INFORMACIÓN PARA LA SEQ ID NO: 20:

(i) CARACTERÍSTICAS DE SECUENCIA:

- (A) LONGITUD: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE HEBRA: individual
- (D) TOPOLOGÍA: lineal

65 (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 20:

ACGGGCTCCT CAAACACCAC T

21

## ES 2 308 815 T3

- (2) INFORMACIÓN PARA LA SEQ ID NO: 21:
- (i) CARACTERÍSTICAS DE SECUENCIA:
    - (A) LONGITUD: 24 pares de bases
    - (B) TIPO: ácido nucleico
    - (C) TIPO DE HEBRA: individual
    - (D) TOPOLOGÍA: lineal
  - (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 21:  
TGGAGATGGG CAGGCCGCAG GGTG 24
- (2) INFORMACIÓN PARA LA SEQ ID NO: 22:
- (i) CARACTERÍSTICAS DE SECUENCIA:
    - (A) LONGITUD: 24 pares de bases
    - (B) TIPO: ácido nucleico
    - (C) TIPO DE HEBRA: individual
    - (D) TOPOLOGÍA: lineal
  - (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 22:  
CAGTCCAGCT GGGCTGAGCA GGTG 24
- (2) INFORMACIÓN PARA LA SEQ ID NO: 23:
- (i) CARACTERÍSTICAS DE SECUENCIA:
    - (A) LONGITUD: 24 pares de bases
    - (B) TIPO: ácido nucleico
    - (C) TIPO DE HEBRA: individual
    - (D) TOPOLOGÍA: lineal
  - (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 23:  
GCGGCTCCAA GAAGTGCATC CAGG 24
- (2) INFORMACIÓN PARA LA SEQ ID NO: 24:
- (i) CARACTERÍSTICAS DE SECUENCIA:
    - (A) LONGITUD: 24 pares de bases
    - (B) TIPO: ácido nucleico
    - (C) TIPO DE HEBRA: individual
    - (D) TOPOLOGÍA: lineal
  - (xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 24:  
CTCCACCCTG CAAGGAAGAG GGGC 24
- (2) INFORMACIÓN PARA LA SEQ ID NO: 25:
- (i) CARACTERÍSTICAS DE SECUENCIA:
    - (A) LONGITUD: 15 aminoácidos
    - (B) TIPO: aminoácido
    - (C) TIPO DE HEBRA:
    - (D) TOPOLOGÍA: lineal



## ES 2 308 815 T3

(2) INFORMACIÓN PARA LA SEQ ID NO: 30:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 24 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 30:

GGCTGCACCT CCTGGACTGT TGCC

(2) INFORMACIÓN PARA LA SEQ ID NO: 31:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 24 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 31:

GATCCTGCTC AGGAGACGTG ACCC

(2) INFORMACIÓN PARA LA SEQ ID NO: 32:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 24 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 32:

CACCAGGCAA GGAGAGGCTC CCGG

(2) INFORMACIÓN PARA LA SEQ ID NO: 33:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 20 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

(xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 33:

CCACCCCATG GCGACGGACG

(2) INFORMACIÓN PARA LA SEQ ID NO: 34:

(i) CARACTERÍSTICAS DE SECUENCIA:

(A) LONGITUD: 29 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

## ES 2 308 815 T3

(xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 34:

GGAATTCGGA GGGGAAGGGG GCCGCCGGA

5

(2) INFORMACIÓN PARA LA SEQ ID NO: 35:

(i) CARACTERÍSTICAS DE SECUENCIA:

10

(A) LONGITUD: 24 pares de bases

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: individual

(D) TOPOLOGÍA: lineal

15

(xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 35:

GGACTGAGGA AGGAGGTGTC CTTC

20

(2) INFORMACIÓN PARA LA SEQ ID NO: 36:

(i) CARACTERÍSTICAS DE SECUENCIA:

25

(A) LONGITUD: 20 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE HEBRA:

(D) TOPOLOGÍA: lineal

30

(ii) TIPO DE MOLÉCULA: péptido

(xi) DESCRIPCIÓN DE SECUENCIA: SEQ ID NO: 36:

35

**Asp Gly Ile Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro Ser**  
**1 5 10 15 20**

40

**Asp Gly Ile Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro Ser**  
**1 5 10 15 20**

45

50

55

60

65