

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 546 310**

51 Int. Cl.:

**C07K 14/47**

(2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.04.2011** **E 11720057 (6)**

97 Fecha y número de publicación de la concesión europea: **03.06.2015** **EP 2563806**

54 Título: **Leucolectinas humanas y usos de las mismas**

30 Prioridad:

**29.04.2010 GB 201007202**

45 Fecha de publicación y mención en BOPI de la  
traducción de la patente:

**22.09.2015**

73 Titular/es:

**LEUKOLECT AS (100.0%)**

**C. Sundts gate 51**

**5004 Bergen, NO**

72 Inventor/es:

**WALTHER, BERNT TH y**

**MIFTARI, MIRUSHE**

74 Agente/Representante:

**ISERN JARA, Jorge**

ES 2 546 310 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

Leucolectinas humanas y usos de las mismas

5 La presente invención se refiere a un polipéptido, particularmente una lectina, a su secuencia de ácido nucleico codificante y a anticuerpos contra el polipéptido, y se refiere al uso de dicho polipéptido y molécula de ácido nucleico en diversas aplicaciones médicas.

10 El sistema de defensa inmunitario celular es fundamental para la supervivencia frente a exposiciones a microbios y parásitos. Este sistema también es el más importante para el manejo de las células defectuosas en la carcinogénesis, después de infecciones víricas y en enfermedades autoinmunitarias.

15 El reconocimiento intercelular es fundamental para estas funciones, pero el origen de dicho fenómeno no se entiende bien. Después de siglos de diversificación de la biota sexual primordial, los genes se han visto obligados a evolucionar para asegurar el reconocimiento específico de especie de las células sexuales ("fertilización"). Si no fuera así, se habría visto amenazada la viabilidad de las especies sexuales. Evolucionaron varias moléculas de reconocimiento celular. Durante la explosión de la era Cámbrica, la biota invadió biotopos terrestres, favoreciendo cada vez más dicho biotopos áridos una restricción de la fertilización sexual a entornos protegidos con menos posibilidad de confusión de especies. Por lo tanto, los genes para el reconocimiento de gametos específicos de especie tenderían a volverse redundantes, y de esta forma estarían libres para adquirir nuevas funcionalidades en la evolución.

25 Una nueva función evidente podría ser la capacidad de los organismos de distinguir sus propias células (propias) de células extrañas (no propias). Se propone que este mecanismo proporciona un fundamento para el origen del sistema de inmunidad celular, un fenómeno que de otra forma está poco explicado. Dicha distinción por inmunidad innata (Medzhitov & Janeway, 1997, Cell, 91(3), p295-298) es esencial para organismos multicelulares para prevenir su canibalización por parásitos y microbios. Un funcionamiento erróneo del sistema para distinguir lo "propio" de lo "no propio" puede constituir la base de algunas enfermedades autoinmunitarias.

30 Las respuestas inmunitarias mal dirigidas que se conocen como autoinmunidad pueden demostrarse por la presencia de autoanticuerpos o linfocitos T reactivos con antígenos del hospedador. Aunque normalmente esto es inocuo y probablemente un fenómeno universal de la vida de los vertebrados, la autoinmunidad puede ser la causa de un amplio espectro de enfermedades humanas, conocidas como enfermedades autoinmunitarias. Este concepto de autoinmunidad como causa de enfermedades humanas es relativamente nuevo y no se aceptó en la corriente principal de pensamiento médico hasta los años 50 y 60. Las enfermedades autoinmunitarias se definen como enfermedades en las que se produce la progresión desde una autoinmunidad benigna a una autoinmunidad patológica. Esta progresión se determina tanto por influencias genéticas como por desencadenantes ambientales.

40 Las enfermedades autoinmunitarias son un riesgo importante para la salud. Hay más de ochenta enfermedades producidas por autoinmunidad. Existe un riesgo especial para las mujeres; aproximadamente el 75% de los pacientes son mujeres. Las enfermedades autoinmunitarias se encuentran entre las diez causas principales de muerte entre las mujeres en todos los grupos de edad hasta los 65 años.

45 Las enfermedades autoinmunitarias afectan a muchas partes diferentes del cuerpo incluyendo la piel (por ejemplo, alopecia areata, pénfigo buloso, epidermolisis bulosa adquirida, pénfigo foliáceo, pénfigo vulgar, vitiligo, psoriasis y acné), el riñón (por ejemplo, nefropatía por IgA), la sangre (por ejemplo, anemia aplásica, anemias hemolíticas autoinmunitarias, púrpura trombocitopénica idiopática y anemia perniciosa), las articulaciones (por ejemplo, espondilitis anquilosante), los músculos (por ejemplo, polimiositis/dermatomiositis), el oído (por ejemplo, pérdida auditiva autoinmunitaria y síndrome de Meniere), los ojos (por ejemplo, úlcera de Mooren, síndrome de Reiter y enfermedad de Vogt-Koyanagi-Harada), el corazón (por ejemplo, miocarditis autoinmunitaria, síndrome de Churg-Strauss, arteritis de células gigantes, enfermedad de Kawasaki, poliarteritis nodosa, arteritis de Takayasu y granulomatosis de Wegener), el sistema endocrino (por ejemplo, enfermedad de Addison, hipoparatiroidismo autoinmunitario, hipofisitis autoinmunitaria, ooforitis autoinmunitaria, orquitis autoinmunitaria, enfermedad de Graves, tiroiditis de Hashimoto, síndrome autoinmunitario poliglandular de tipo 1 (PAS-1), síndrome autoinmunitario poliglandular de tipo 2 (PAS-2), síndrome autoinmunitario poliglandular de tipo 3 (PAS-3) y diabetes mellitus de tipo 1), el sistema gastroentérico (por ejemplo, hepatitis autoinmunitaria, enfermedad celíaca, enfermedad inflamatoria del intestino y cirrosis biliar primaria), el sistema nervioso (por ejemplo, polineuropatía desmielinizante inflamatoria crónica, síndrome de Guillain-Barre, esclerosis múltiple y miastenia grave) o pueden afectar al cuerpo sistémicamente (por ejemplo, síndrome antifosfolípido, linfoproliferación autoinmunitaria, poliendocrinopatía autoinmunitaria, enfermedad de Bechet, síndrome Goodpasture, artritis reumatoide, sarcoidosis, esclerodermia, síndrome de Sjogren y lupus eritematoso sistémico).

65 Las enfermedades autoinmunitarias pueden atacar a cualquier parte del cuerpo y, por lo tanto, los síntomas varían ampliamente y con frecuencia son difíciles el diagnóstico y el tratamiento. Algunas enfermedades autoinmunitarias pueden poner en peligro la vida a menos que se diagnostiquen y se traten de forma apropiada. Los trastornos autoinmunitarios crónicos tales como la artritis reumatoide producen minusvalías en el paciente y también crean

cargas pesadas sobre las familias de los pacientes. Algunos tipos de uveítis pueden producir ceguera. Ciertas enfermedades tales como la esclerodermia requieren un tratamiento eficaz de por vida. Otras enfermedades autoinmunitarias, incluyendo la enfermedad de Graves y la tiroiditis crónica, pueden tratarse de forma satisfactoria si se diagnostican correctamente, pero frecuentemente esto no es así debido a su sutil inicio.

La inflamación es un proceso normal en el que los glóbulos blancos del cuerpo y los agentes químicos protegen al cuerpo de una infección y de sustancias extrañas tales como bacterias y virus. Sin embargo, en algunas enfermedades el sistema inmunitario del cuerpo desencadena de forma inapropiada una respuesta inflamatoria sin que haya sustancias extrañas con las que luchar.

De esta manera, la inflamación es una enfermedad autoinmunitaria común, pero no todos los trastornos inflamatorios son respuestas autoinmunitarias. Las enfermedades inflamatorias también pueden producirse por una amplia diversidad de agentes que atacan directamente al cuerpo tales como microorganismos (virus y hongos), toxinas bacterianas, ciertos agentes farmacéuticos (antibióticos y esteroides antiinflamatorios) y agentes químicos (sales biliares, agentes químicos domésticos tóxicos). Las enfermedades que están asociadas con la inflamación incluyen artritis (que es un término general que describe la inflamación en las articulaciones), inflamación del corazón (miocarditis), inflamación de los pequeños conductos que transportan aire a los pulmones (que puede producir un ataque asmático), inflamación de los riñones (nefritis) e inflamación del intestino grueso (colitis).

Los trastornos inflamatorios gastrointestinales tienen un interés particular (por ejemplo, enfermedades inflamatorias gástricas (tales como úlcera gástrica, úlcera duodenal y gastritis) y enfermedades inflamatorias intestinales (incluyendo enfermedad de Crohn, enfermedad inflamatoria del intestino, esprúe tropical y no tropical, enteritis infecciosa, colitis, colitis ulcerosa, colitis pseudomembranosa, diverticulitis y enfermedades inflamatorias alérgicas y radiológicas).

Las enfermedades inflamatorias gastrointestinales se caracterizan por inflamación, específicamente por la presencia de edema, células inflamatorias características (es decir, leucocitos, histiocitos y macrófagos) y, en algunos casos, necrosis y ulceración del epitelio de la superficie. Es conocido que estas enfermedades inflamatorias se producen por una amplia diversidad de agentes presentes en el tracto gastrointestinal que se sabe que atacan a las superficies del mismo, produciendo la respuesta de enfermedad inflamatoria. Estos agentes incluyen microorganismos (virus y hongos), toxinas bacterianas, ciertos agentes farmacéuticos (antibióticos y esteroides antiinflamatorios) y agentes químicos (sales biliares, agentes químicos domésticos tóxicos). De hecho, el propio ácido gástrico es capaz de atacar al revestimiento del estómago y producir un estado inflamatorio.

Las medicaciones usadas actualmente para tratar la inflamación incluyen fármacos antiinflamatorios no esteroideos (AINE - tales como aspirina, ibuprofeno o naproxeno), corticosteroides (tales como prednisona), medicaciones antimaláricas (tales como hidroxiquina) y otras medicaciones tales como metotrexato, sulfasalazina, leflunomida, medicaciones anti-TNF, ciclofosfamida y micofenolato. Algunas enfermedades gastrointestinales, específicamente enfermedades gástricas, pueden tratarse por inhibición de la secreción del ácido gástrico que produce la inflamación, tal como mediante la neutralización de los efectos del ácido (por ejemplo, administración de antiácidos), o mediante la administración de un agente farmacológico eficaz para inhibir la secreción de ácido gástrico.

La piel dañada es vulnerable a las infecciones y puede pasar inadvertida y/o producir dolor o molestias. Algunos tipos de heridas son resistentes a la curación en condiciones fisiológicas normales, por ejemplo, úlceras crónicas. Cualquier solución farmacéutica para reparar la piel dañada, particularmente para reparar heridas es muy deseable.

Sigue existiendo la necesidad de tratamientos adecuados para tratar trastornos y afecciones autoinmunitarias e inflamatorias o lesiones de la piel con efectos secundarios mínimos.

Después de muchos años de exploración de procesos potencialmente implicados en la evolución de los eucariotas, los inventores identificaron una nueva proteína (denominada en el presente documento leucolectina), no solo en gametos y en cigotos, sino también en las primeras fases embrionarias (en células especializadas denominadas lectocitos) y durante la ontogenia de células sanguíneas y, finalmente, en los leucocitos.

Inicialmente, la proteína se identificó y se purificó a partir de peces (véanse los ejemplos y el número de registro de la base de datos BT058707). La proteína tiene 255 aminoácidos. Ésta es la forma propéptido de la lectina, que contiene un péptido N-terminal de 19 aminoácidos que sugiere que se dirige al lisosoma para la secreción posterior (es decir, en el espacio perivitelino).

La secuencia de aminoácidos de la lectina permitió el desarrollo de anticuerpos específicos de epítipo, lo cual a su vez permitió la identificación de muchas (2-8) isoformas parecidas de la proteína (Figura 5), dependiendo del tejido analizado. Se han aislado al menos dos ARNm a partir de salmón (véase el Ejemplo 11) que contienen diferencias de secuencias minoritarias que dan como resultado únicamente 7 cambios a nivel del polipéptido (Figura 17). También se han identificado formas truncadas de la proteína a partir de leucocitos de salmón (véase la SEC ID N°: 2) y el pez cebra, en el que se ha identificado una forma secretada (como se ha descrito anteriormente, denominada

sLL) y una forma truncada que carece de los primeros 32 aminoácidos del extremo N, denominada tLL (véase el Ejemplo 9).

La proteína tiene alguna similitud en dominios pequeños con las taquilelectinas.

Se han identificado proteínas relacionadas en diversos animales, incluyendo el pez cebra, el bacalao, la trucha arco iris, *Oikopleura dioica* y también en pollo y en seres humanos (véase, por ejemplo, el número de registro de la base de datos FJ643619). Se identificaron sus ADNc (véanse los Ejemplos) y se descubrió que estaban muy bien conservados. Las proteínas codificadas en las especies que se han examinado muestran una variación mínima. Esto indica una función esencial para estas proteínas.

Es importante destacar que este nuevo gen se expresa específicamente en leucocitos humanos a pesar del hecho de que este gen no puede encontrarse en secuencias publicadas del genoma humano. Solo pueden detectarse segmentos muy cortos de este gen pero extendidos en múltiples cromosomas humanos (véanse los Ejemplos). La resolución de este enigma todavía no está cerca.

Finalmente, se ha descubierto que la entidad molecular se secreta por células, lo cual es coherente con la forma propeptídica de la proteína. Se han aislado tanto la forma nativa como la forma recombinante del producto génico.

Como se desvela en el documento WO 2010/049688, se ha descubierto que, sorprendentemente, estas proteínas tienen efectos pronunciados sobre trastornos autoinmunitarios e inflamatorios, particularmente de la piel, y otras afecciones cutáneas. Aunque no hay deseo de limitarse por ninguna teoría, se cree que la leucolectina se une a receptores específicos que se establecen durante los procesos normales e intrínsecos de curación del cuerpo humano. A partir de los efectos observados, los presentes inventores proponen que la leucolectina participa en el proceso normal de inmunidad celular. La leucolectina parece estar ausente en muchas patologías, de forma que su introducción farmacéutica desencadena otras respuestas defensivas del sistema inmunitario para comportarse de forma normal frente a diversos desafíos: la presencia de células extrañas (microbios, parásitos) y de células alteradas del propio organismo, lo cual puede asociar la leucolectina a la etiología de enfermedades autoinmunitarias, es decir, dichas afecciones pueden producirse por la ausencia de suficiente leucolectina.

Los leucocitos humanos contienen y expresan proteínas de leucolectina extremadamente conservadas. Hasta ahora, estas proteínas eran proteínas desconocidas en las células sanguíneas. La leucolectina parece ser un nuevo miembro de la familia de proteínas conocida como taquilelectina.

En un primer aspecto, la presente invención proporciona una composición farmacéutica que comprende (i) un polipéptido capaz de tratar lesiones en la piel de un animal, que comprende una secuencia de aminoácidos como se indica en una o cualquiera de las Secuencias nº 16-18 o una secuencia funcionalmente equivalente que tiene una identidad de al menos el 95 % con dicha secuencia, donde dicha comparación se realiza a lo largo de toda la longitud de la secuencia; o (ii) una molécula de ácido nucleico que comprende una secuencia de nucleótidos como se indica en la Secuencia nº 19; y uno o más excipientes y/o diluyentes farmacéuticamente aceptables para su uso en terapia.

Los polipéptidos particularmente preferidos para su uso de acuerdo con la invención comprenden una secuencia de aminoácidos como la indicada en las Secuencias nº 18 o 16, o una secuencia que tiene una identidad de al menos el 95 % con dicha secuencia, donde dicha comparación se realiza a lo largo de toda la longitud de la secuencia.

En el presente documento también se desvelan polipéptidos que comprenden una secuencia de aminoácidos como la indicada en una o cualquiera de las Secuencias nº 1-8, 16-18 o 20-23

Los "polipéptidos", como se mencionan en el presente documento, son moléculas con más de 200 o 250 restos y/o menos de 500, 400 o 300 restos o un intervalo seleccionado de los mismos. Una "parte" puede comprender al menos 30, 40, 50, 60, 70, 80, 90, 100, 150, 200 o más aminoácidos de la secuencia de la que procede. Dicha parte puede obtenerse a partir de una parte central, N-terminal o C-terminal de la secuencia. Una parte puede obtenerse a partir del extremo N-terminal, por ejemplo, de los primeros 50, 100 o 150 restos del polipéptido. Pueden producirse truncamientos de dichos polipéptidos en el extremo N-terminal, por ejemplo, para retirar un péptido señal o una parte ausente en variantes naturales, y pueden tener una longitud de 1 a 50, por ejemplo, de 1 a 10, 20, 30 o 40, o de 5 a 40, por ejemplo, de 10 a 35 restos.

Preferentemente, dicha secuencia tiene una identidad de al menos el 96, 97, 98 o 99% con la secuencia con la que se compara.

La identidad de secuencia puede determinarse, por ejemplo, usando el banco de datos de secuencias de proteínas SWISS-PROT usando FASTA pep-cmp con un factor pam variable, y una penalización de creación de espacios en blanco establecida a 12,0 y una penalización de extensión de espacios en blanco establecida a 4,0, y una ventana de 2 aminoácidos. Dicha comparación se realiza en toda la longitud de la secuencia.

Los polipéptidos relacionados por identidad de secuencia son funcionalmente equivalentes a los polipéptidos que se indican en los números de Secuencias mencionados. Dichos polipéptidos de funcionalidad equivalente pueden tener la forma de derivados como se indica más adelante. De forma similar, los polipéptidos con secuencias indicadas en los nº de Secuencia pueden modificarse sin afectar a la secuencia del polipéptido como se describe más adelante.

Además, las "partes" descritas en el presente documento pueden ser equivalentes funcionales. Estas partes satisfacen las condiciones de identidad (con respecto a una región comparable) mencionadas en el presente documento.

Como se indica en el presente documento, para conseguir "equivalencia funcional" el polipéptido puede mostrar alguna eficacia reducida en la presentación de la función médica con respecto a la molécula parental (es decir, la molécula a partir de la que se obtuvo, por ejemplo, por sustitución de aminoácidos), pero preferentemente es igual de eficaz o es más eficaz. De esta manera, la equivalencia funcional se refiere a un polipéptido que es eficaz para tratar una enfermedad como se menciona en el presente documento, es decir, para reducir uno o más síntomas del paciente, por ejemplo, inflamación o el aspecto de la piel como se describe más adelante en el presente documento. Esto puede ensayarse por comparación de los efectos del polipéptido derivado con respecto al polipéptido del que procede de una manera cualitativa o cuantitativa, por ejemplo, realizando los análisis *in vivo* mencionados en los Ejemplos. Cuando son posibles resultados cuantitativos, el derivado es al menos un 30, 50, 70 o 90 % tan eficaz como el polipéptido parental. Como alternativa, el ensayo *in vitro* puede realizarse, por ejemplo, por análisis de la unión a células dendríticas o efectos en cultivos celulares *in vitro*.

Pueden obtenerse proteínas funcionalmente equivalentes que están relacionadas o proceden de la proteína natural mediante la modificación de la secuencia de aminoácidos nativa por sustitución, adición y/o delección de un solo o múltiples aminoácidos (siempre que satisfagan los requisitos de identidad de secuencia mencionados anteriormente), pero sin destruir la función de la molécula. Preferentemente, la secuencia nativa tiene menos de 10 sustituciones, adiciones o delecciones, por ejemplo menos de 5, 4, 3, 2, o 1 de dichas modificaciones. Dichas proteínas se codifican por "moléculas de ácido nucleico funcionalmente equivalentes" que se generan mediante la sustitución, adición y/o delección apropiada de una o más bases.

Los equivalentes funcionales preferidos son variantes de "adición" en los que se generan polipéptidos o proteínas de fusión amino y/o carboxi terminales, que comprenden una proteína o polipéptido adicional fusionado al polipéptido parental.

Las variantes funcionalmente equivalentes particularmente preferidas son variaciones biológicas naturales (por ejemplo, variantes alélicas o variaciones geográficas dentro de una especie o, como alternativa, en diferentes géneros, por ejemplo, plantas, animales o bacterias) y derivados preparados usando técnicas conocidas. Por ejemplo, pueden introducirse moléculas de ácido nucleico que codifican proteínas funcionalmente equivalentes por síntesis química o en forma recombinante usando las técnicas conocidas de mutagénesis dirigida incluyendo delección, mutagénesis aleatoria o escisión enzimática y/o ligamiento de ácidos nucleicos.

También se desvelan secuencias de nucleótidos que codifican polipéptidos descritos en el presente documento.

La presente invención de esta manera proporciona una molécula de ácido nucleico para su uso en la invención, que comprende una secuencia de nucleótidos como la indicada en la Secuencia nº 19. Dicha molécula de ácido nucleico codifica un polipéptido como se ha indicado anteriormente en el presente documento.

También se desvelan en el presente documento moléculas de ácido nucleico que comprenden una secuencia de nucleótidos como se indica en una o cualquiera de las secuencias nº 9-15.

Las "moléculas de ácido nucleico" como se mencionan en el presente documento con moléculas con más 450, 600 o 750 bases y/o menos de 1.500, 1.200, 900 o 600 bases o un intervalo seleccionado entre los mismos. Una "parte" puede comprender al menos 90, 120, 150, 180, 210, 240, 270, 300, 450 o 600 bases de nucleótidos de la secuencia de la que procede. Dichas partes pueden codificar péptidos N-terminales, centrales o C-terminales como se ha descrito anteriormente en el presente documento. Como se ha analizado anteriormente en relación con los polipéptidos, los truncamientos dan como resultado la eliminación de restos del extremo N-terminal de los polipéptidos mencionados. En las secuencias de nucleótidos codificantes, con respecto a las secuencias presentadas en el presente documento, un truncamiento puede tener una longitud de 1 a 150, por ejemplo de 1 a 30, 60, 90 o 120, o de 13 a 120, por ejemplo, de 28 a 105 bases.

La identidad de secuencia puede determinarse, por ejemplo, por búsqueda FASTA usando paquetes GCG, con valores por defecto y un factor pam variable, y una penalización de creación de espacios en blanco establecida a 12,0 y una penalización de extensión de espacios en blanco establecida a 4.0 con una ventana de 6 nucleótidos.

Las "partes", como se describen en el presente documento, pueden ser equivalentes funcionales.

Las moléculas de ácido nucleico desveladas en el presente documento pueden ser ARN, ADNc o ADN monocatenario o bicatenario, preferentemente ADN. Sin embargo, idealmente las moléculas son ADN o ADNc.

Los polipéptidos como se han descrito anteriormente incluyen los que están modificados sin afectar a la secuencia del polipéptido, por ejemplo, por modificación química, incluyendo por desglicosilación o glicosilación. Dichos polipéptidos pueden prepararse por modificación después de la síntesis/aislamiento del polipéptido sin afectar a la funcionalidad, por ejemplo, cierta glicosilación, metilación etc. de restos particulares.

Los polipéptidos desvelados en el presente documento también pueden tomar la forma de peptidomiméticos que pueden considerarse derivados en los que se conservan las características funcionales del polipéptido pero se presentan en el contexto de una estructura diferente, por ejemplo, no peptídica. Dichos peptidomiméticos se han desarrollado satisfactoriamente y usado para otras aplicaciones médicas particulares.

Los peptidomiméticos, particularmente las moléculas no peptídicas pueden generarse a través de diversos procesos, incluyendo el diseño de fármacos de base conformacional, exploración, diseño de bibliotecas focalizadas y la química medicinal clásica. No solo pueden usarse oligómeros de aminoácidos no naturales u otros bloques de construcción orgánicos, sino también carbohidratos, compuestos heterocíclicos o macrocíclicos o cualquier molécula orgánica que comprenda elementos estructurales y una conformación que proporcione una superficie electrostática molecular que imita las mismas propiedades de la conformación tridimensional del péptido, por métodos conocidos en la técnica.

De esta manera, los peptidomiméticos pueden tener un parecido pequeño o nulo con una cadena principal de péptido. Los peptidomiméticos pueden comprender una forma no peptídica completamente sintética (por ejemplo, basada en una cadena principal de carbohidrato con sustituyentes apropiados) o pueden conservar uno o más elementos del péptido en el que están basados, por ejemplo, por derivatización de uno o más aminoácidos o reemplazo de uno o más aminoácidos con componentes no peptídicos alternativos. Los moldes de tipo péptido incluyen seudopéptidos y péptidos cíclicos. Los elementos estructurales considerados redundantes para la función del péptido pueden minimizarse para conservar una función de almacén únicamente o eliminarse cuando sea apropiado.

Cuando los peptidomiméticos conservan uno o más elementos peptídicos, es decir, más de un aminoácido, dichos aminoácidos pueden reemplazarse con un análogo no convencional o un análogo estructural del mismo.

Los aminoácidos conservados en las secuencias también pueden derivatizarse o modificarse (por ejemplo, marcarse, glicosilarse o metilarse) siempre que se conserven las propiedades funcionales de los polipéptidos. Los peptidomiméticos se consideran "derivables de" una cierta secuencia polipeptídica. Por esto se entiende que el peptidomimético está diseñado con referencia a una secuencia polipeptídica definida, de tal forma que retiene las características estructurales del péptido que son esenciales para su función. Estas pueden ser las cadenas laterales particulares del polipéptido, o la capacidad de formar enlaces de hidrógeno de la estructura. Dichas características pueden proporcionarse por componentes no peptídicos o uno o más de los restos de aminoácido o los enlaces que unen dichos restos de aminoácido del polipéptido pueden modificarse para mejorar ciertas funciones del polipéptido tales como la estabilidad o la resistencia a proteasas, mientras se conservan las características estructurales del polipéptido que son esenciales para su función.

Son ejemplos de aminoácidos análogos estructurales o no convencionales que pueden usarse D aminoácidos, isómeros de amida (tales como N-metil amida, amida retro-inversa, tioamida, tioéster, fosfonato, cetometileno, hidroximetileno, fluorovinilo, (E)-vinilo, metilnamino, metilendio o alcanos), L-N metilaminoácidos, D- $\alpha$  metilaminoácidos, D-N-metilaminoácidos. En la Tabla 1 se presentan ejemplos de aminoácidos no convencionales.

TABLA 1

Aminoácido no convencional	Código	Aminoácido no convencional	Código
ácido $\alpha$ -aminobutírico	Abu	L-N-metilalanina	Nmala
$\alpha$ -amino- $\alpha$ -metilbutirato	Mgab	L-N-metilarginina	Nmarg
aminociclopropano-carboxilato	Cpro	L-N-metilasparagina	Nmasn
		ácido L-N-metilaspártico	Nmasp
ácido aminoisobutírico		L-N-metilcisteína	Nmcys
aminonorbornil-carboxilato	Norb	L-N-metilglutamina	Nmgln
		ácido L-N-metilglutámico	Nmglu
ciclohexilalanina	Chexa	L-N-metilhistidina	Nmhis
ciclopentilalanina	Cpen	L-N-metilisoleucina	Nmile
D-alanina	Dal	L-N-metil-leucina	Nmleu
D-arginina	Darg	L-N-metil-lisina	Nmlys
ácido D-aspártico	Das	L-N-metilmetionina	Nmmet
D-cisteína	Dcys	L-N-metilnorleucina	Nmnle
D-glutamina	Dgln	L-N-metilnorvalina	Nmnva
ácido D-glutámico	Dglu	L-N-metilornitina	Nmorn
D-histidina	Dhis	L-N-metilfenilalanina	Nmphe
D-isoleucina	Dile	L-N-metilprolina	Nmpro

Aminoácido no convencional	Código	Aminoácido no convencional	Código
D-leucina	Dleu	L-N-metilserina	Nmser
D-lisina	Dlys	L-N-metiltreonina	Nmthr
D-metionina	Dmet	L-N-metiltryptófano	Nmtrp
D-ornitina	Dorn	L-N-metil tirosina	Nmtyr
D-fenilalanina	Dphe	L-N-metilvalina	Nmval
D-prolina	Dpro	L-N-metiletilglicina	Nmetg
D-serina	Dser	L-N-metil-t-butilglicina	Nmtbug
D-treonina	Dthr	L-norleucina	Nle
D-tryptófano	Dtrp	L-norvalina	Nva
D-tirosina	Dtyr	$\alpha$ -metil-aminoisobutirato	Maib
D-valina	Dval	$\alpha$ -metil- $\gamma$ -aminobutirato	Mgab
D- $\alpha$ -metilalanina	Dmala	$\alpha$ -metilciclohexilalanina	Mchexa
D- $\alpha$ -metilarginina	Dmarg	$\alpha$ -metilciclopentilalanina	Mcpen
D- $\alpha$ -metilasparagina	Dmasn	$\alpha$ -metil- $\alpha$ -naftilalanina	Manap
D- $\alpha$ -metilaspartato	Dmasp	$\alpha$ -metilpenicilamina	Mpen
D- $\alpha$ -metilcisteína	Dmcys	N-(4-aminobutil)glicina	Nglu
D- $\alpha$ -metilglutamina	Dmgln	N-(2-aminoetil)glicina	Naeg
D- $\alpha$ -metilhistidina	Dmhis	N-(3-aminopropil)glicina	Norn
D- $\alpha$ -metilisoleucina	Dmile	N-amino- $\alpha$ -metilbutirato	Nmaabu
D- $\alpha$ -metil-leucina	Dmleu	$\alpha$ -naftilalanina	Anap
D- $\alpha$ -metil-lisina	Dmlys	N-bencilglicina	Nphe
D- $\alpha$ -metilmetionina	Dmmet	N-(2-carbamiletil)glicina	Ngln
D- $\alpha$ -metilornitina	Dmorn	N-(carbamiletil)glicina	Nasn
D- $\alpha$ -metilfenilalanina	Dmphe	N-(2-carboxietil)glicina	Nglu
D- $\alpha$ -metilprolina	Dmpro	N-(carboximetil)glicina	Nasp
D- $\alpha$ -metilserina	Dmser	N-ciclobutilglicina	Ncbut
D- $\alpha$ -metiltreonina	Dmthr	N-cicloheptilglicina	Nchep
D- $\alpha$ -metiltryptófano	Dmtrp	N-ciclohexilglicina	Nchex
D- $\alpha$ -metiltirosina	Dmtyr	N-ciclododecilglicina	Ncdec
D- $\alpha$ -metilvalina	Dmval	N-ciclododecilglicina	Ncdod
D-metilalanina	Dnmala	N-ciclooctilglicina	Ncoct
D-N-metilarginina	Dnmarg	N-ciclopropilglicina	Ncpro
D-N-metilasparagina	Dnmasn	N-cicloundecilglicina	Ncund
D-N-metilaspartato	Dnmasp	N-(2,2-difeniletil)glicina	Nbhm
D-N-metilcisteína	Dnmcys	N-(3,3-difenilpropil)glicina	Nbhe
D-N-metilglutamina	Dnmglu	N-(3-guanidinopropil)glicina	Narg
D-N-metilglutamato	Dnmglu	N-(1-hidroxietil)glicina	Nthr
D-N-metilhistidina	Dnmhis	N-(hidroxietil)glicina	Nser
D-N-metilisoleucina	Dnmile	N-(imidazoliletal)glicina	Nhis
D-N-metil-leucina	Dnmleu	N-(3-indoliletal)glicina	Nhtrp
D-N-metil-lisina	Dnmlys	N-metil- $\gamma$ -aminobutirato	Nmgabu
N-metilciclohexilalanina	Nmchexa	D-N-metilmetionina	Dnmmt
D-N-metilornitina	Dnmorn	N-metilciclopentilalanina	Nmcpen
N-metilglicina	Nala	D-N-metilfenilalanina	Dnmphe
N-metilaminoisobutirato	Nmaib	D-N-metilprolina	Dnmpro
N-(1-metilpropil)glicina	Nile	D-N-metilserina	Dnmser
N-(2-metilpropil)glicina	Nleu	D-N-metiltreonina	Dnmthr
D-N-metiltryptófano	Dnmtrp	N-(1-metiletil)glicina	Nval
D-N-metiltirosina	Dnmtyr	N-metil- $\alpha$ -naftilalanina	Nmanap
D-N-metilvalina	Dnmval	N-metilpenicilamina	Nmpen
ácido $\gamma$ -aminobutírico	Gabu	N-(p-hidroxifenil)glicina	Nhtyr
L-t-butilglicina	Thug	N-(tiometil)glicina	Ncys
L-etilglicina	Etg	Penicilamina	Pen
L-homofenilalanina	Hphe	L- $\alpha$ -metilalanina	Mala
L- $\alpha$ -metilarginina	Marg	L- $\alpha$ -metilasparagina	Masn
L- $\alpha$ -metilaspartato	Masp	L- $\alpha$ -metil-t-butilglicina	Mtbug
L- $\alpha$ -metilcisteína	Mcys	L-metiletilglicina	Metg
L- $\alpha$ -metilglutamina	Mgln	L- $\alpha$ -metilglutamato	Mglu
L- $\alpha$ -metilhistidina	Mhis	L- $\alpha$ -metilhomofenilalanina	Mhphe
L- $\alpha$ -metilisoleucina	Mile	N-(2-metiltioetil)glicina	Nmet
L- $\alpha$ -metil-leucina	Mleu	L- $\alpha$ -metil-lisina	Mlys
L- $\alpha$ -metilmetionina	Mmet	L- $\alpha$ -metilnorleucina	Mnle
L- $\alpha$ -metilnorvalina	Mnva	L- $\alpha$ -metilornitina	Morn
L- $\alpha$ -metilfenilalanina	Mphe	L- $\alpha$ -metilprolina	Mpro
L- $\alpha$ -metilserina	Mser	L- $\alpha$ -metiltreonina	Mthr

Aminoácido no convencional	Código	Aminoácido no convencional	Código
L- $\alpha$ -metiltryptófano	Mtrp	L- $\alpha$ -metiltirosina	Mtyr
L- $\alpha$ -metilvalina	Mval	L-N-metilhomofenilalanina	Nmhph
N-(N-(2,2-difeniletil)carbamilmetil)glicina	Nnbhm	N-(N-(3,3-difenilpropil)carbamilmetil)glicina	Nnbhe
1-carboxi-1-(2,2-difeniletilamino)ciclopropano	Nmbc	L-O-metil serina	Omser
		L-O-metil homoserina	Omhsr

Los aminoácidos no convencionales que pueden usarse incluyen análogos restringidos conformacionalmente, tales como Tic (para reemplazar F), Aib (para reemplazar A) o ácido pipecólico (para reemplazar Pro).

- 5 Los polipéptidos y moléculas de ácido nucleico analizadas anteriormente también incluyen derivados que se han modificado, por ejemplo, para facilitar su uso en aplicaciones farmacéuticas (analizadas más adelante), por ejemplo, mediante la adición de grupos de dirección o funcionales, por ejemplo, para mejorar la lipofilia, ayudar al transporte celular, solubilidad y/o estabilidad. De esta manera, pueden conjugarse oligosacáridos, ácidos grasos, alcoholes grasos, aminoácidos, péptidos o polipéptidos con los polipéptidos o moléculas de ácido nucleico mencionadas anteriormente. Las moléculas de ácido nucleico pueden estar presentes en un vehículo viral como se describe más adelante en el presente documento.

15 Los polipéptidos también incluyen derivados en forma de "profármacos" o "propéptidos" de tal forma que el componente añadido pueda retirarse por escisión una vez administrado, por ejemplo, por escisión de un sustituyente añadido mediante esterificación que puede retirarse por la acción de esterasas. Dichos profármacos incluyen precursores nativos de las proteínas naturales que se escinden, por ejemplo, por proteólisis para producir el polipéptido de interés. Dichos precursores pueden ser inactivos en la forma precursora pero pueden activarse por escisión proteolítica. De esta manera, las moléculas de ácido nucleico de forma similar incluyen moléculas que codifican dichos profármacos o precursores. Los polipéptidos o moléculas de ácido nucleico modificadas descritas anteriormente pueden ensayarse para asegurar que conservan la actividad funcional con respecto a la molécula no modificada mediante la determinación de si tienen el mismo efecto médico o efectos médicos similares.

25 Las moléculas de ácido nucleico descritas anteriormente pueden unirse operativamente a una secuencia de control de la expresión, o un vehículo o vector de clonación de ADN recombinante que contiene dicha molécula de ADN recombinante. Esto permite la expresión intracelular del polipéptido como un producto génico, cuya expresión se dirige por el gen o los genes introducidos en las células de interés. La expresión génica se dirige desde un promotor activo en las células de interés y puede insertarse en cualquier forma de vector de ADN lineal o circular para la incorporación en el genoma o para la replicación independiente o transfección/expresión transitoria. Las técnicas de transformación o transfección adecuadas están bien descritas en la bibliografía. Como alternativa, la molécula de ADN desnudo puede introducirse directamente en la célula para los usos descritos en el presente documento.

35 Los vectores de expresión apropiados incluyen secuencias de control apropiadas tales como, por ejemplo, elementos de control de la traducción (por ejemplo, codones de inicio y terminación, sitios de unión a ribosomas) y elementos de control de la transcripción (por ejemplo, regiones de promotor-operador, secuencias stop de terminación), unidas en la fase de lectura correspondiente a las moléculas de ácido nucleico necesarias para la realización del método descrito en el presente documento más adelante. Los vectores apropiados pueden incluir plásmidos y virus (incluyendo virus tanto bacteriófagos como eucarióticos). Los vectores virales adecuados incluyen baculovirus y también adenovirus, virus adenoasociados, virus del herpes y vaccinia/pox virus. En la técnica se describen muchos otros vectores virales. Los vectores preferidos incluyen vectores de expresión en bacterias y mamíferos pGEX-KG, pEF-neo y pEF-HA. La molécula de ácido nucleico convenientemente puede fusionarse con ADN que codifica un polipéptido adicional, por ejemplo, glutatión-S-transferasa, para producir una proteína de fusión tras la expresión.

45 De esta manera, en el presente documento también se desvela un vector, preferentemente un vector de expresión, que comprende una molécula de ácido nucleico como se ha definido anteriormente.

Otros aspectos desvelados en el presente documento incluyen métodos para preparar moléculas de ácido nucleico recombinantes, que comprenden insertar secuencias de nucleótidos que codifican los polipéptidos desvelados en el presente documento en un ácido nucleico de vector.

55 Los polipéptidos pueden administrarse a una célula por transfección de una célula con una molécula de ácido nucleico. Las moléculas de ácido nucleico que comprenden una secuencia que codifica los polipéptidos descritos en el presente documento pueden usarse en métodos descritos en el presente documento. Preferentemente, dichas moléculas de ácido nucleico están contenidas en un vector, por ejemplo, un vector de expresión.

60 Las moléculas de ácido nucleico desveladas en el presente documento, preferentemente contenidas en un vector, pueden introducirse en una célula por cualquier medio apropiado. En la bibliografía están bien descritas técnicas de transformación o transfección adecuadas. Se conoce una diversidad de técnicas y pueden usarse para introducir dichos vectores en células procariotas o eucariotas para la expresión. Las células hospedadoras para este fin incluyen líneas de células de insecto, líneas de células eucariotas o *E. coli*, tales como la cepa BL21/DE3. También



se desvelan en el presente documento células hospedadoras procariotas o eucariotas transformadas o transfectadas que contienen una molécula de ácido nucleico, particularmente un vector como se ha definido anteriormente.

5 En el presente documento también se desvela un método para preparar un polipéptido como se ha definido anteriormente en el presente documento, que comprende cultivar una célula hospedadora que contiene una molécula de ácido nucleico como se ha definido anteriormente, en condiciones en las que se expresa dicho polipéptido y recuperar dicha molécula producida de esta manera.

10 También se desvelan en el presente documento anticuerpos (monoclonales o policlonales) y sus fragmentos de unión a antígenos (por ejemplo, fragmentos F(ab)<sub>2</sub>, Fab y Fv, es decir, fragmentos de la región "variable" del anticuerpo, que comprende el sitio de unión a antígeno) dirigidos específicamente a los polipéptidos como se han definido anteriormente en el presente documento.

15 Los anticuerpos descritos en el presente documento pueden usarse *in vitro* para identificar la presencia o cantidad del polipéptido descrito en el presente documento y pueden usarse como diagnóstico para identificar trastornos o lesiones cutáneas como se describe en el presente documento asociadas con niveles aberrantes de dicho polipéptido, es decir una variación con respecto a los niveles normales, por ejemplo, niveles elevados o reducidos de dicho polipéptido.

20 De esta manera, también se desvela un método para determinar la presencia o cantidad de un polipéptido o una parte del mismo en una muestra, donde un anticuerpo como se describe en el presente documento se pone en contacto con dicha muestra y el grado de unión del anticuerpo es indicativo de la presencia o cantidad de dicho polipéptido o parte del mismo.

25 La muestra que se ensaya puede ser cualquier muestra conveniente, por ejemplo, una muestra de sangre o de tejido o una muestra no biológica. Preferentemente, la muestra procede de un animal como se describe en el presente documento, preferentemente un ser humano.

30 En el presente documento también se desvela un método para diagnosticar un trastorno o lesión de la piel como se describe en el presente documento, que comprende al menos las etapas de determinar la presencia o cantidad de un polipéptido como se describe en el presente documento en una muestra de dicho animal, donde dicha presencia o cantidad sirve como diagnóstico de dicho trastorno o lesión de la piel. El polipéptido puede detectarse, por ejemplo, usando los anticuerpos descritos anteriormente en el presente documento. La muestra y el animal preferentemente es como se describe en el presente documento. La cantidad es preferentemente una reducción en el nivel de dicho polipéptido en comparación con niveles normales. El diagnóstico puede conseguirse por comparación con tablas patrón de sujetos normales en comparación con los sujetos que tienen el trastorno o lesión cutánea bajo investigación.

40 Los polipéptidos o moléculas de ácido nucleico usadas en las composiciones y usos de la invención, como se describe más adelante en el presente documento, pueden obtenerse o derivar de fuentes naturales o pueden generarse total o parcialmente de forma sintética.

45 Convenientemente, los polipéptidos y moléculas de ácido nucleico se aíslan de acuerdo con los protocolos descritos en los ejemplos.

- Cuando el polipéptido para uso en la invención se obtiene a partir de leucocitos, se obtiene en forma no modificada.

50 Los polipéptidos o moléculas de ácido nucleico descritas en el presente documento preferentemente carecen sustancialmente de cualquier componente contaminante derivado del material o materiales originales usados en el procedimiento de aislamiento o en su preparación sintética. De una forma especialmente preferida, el compuesto se purifica hasta un grado de pureza mayor del 50 o 60 %, por ejemplo, >70, 80 o 90 %, preferentemente mayor del 95 o 99 %, según se evalúa p/p (peso en seco). Dichos niveles de pureza corresponden a la molécula específica de interés, pero incluyen sus productos de degradación. Cuando sea apropiado, pueden usarse preparaciones enriquecidas que tienen menos pureza, por ejemplo, que contienen más de un 1, 2, 5 o 10 % de la molécula de interés, por ejemplo, más de 20 o 30%. Los polipéptidos para su uso de acuerdo con la invención pueden purificarse, por ejemplo, por cromatografía (por ejemplo, HPLC, exclusión molecular, intercambio iónico, afinidad, interacción hidrófoba, fase inversa) o electroforesis capilar.

60 Los polipéptidos para su uso de acuerdo con la invención pueden generarse de forma sintética, por ejemplo, por ligamiento de péptidos generados de forma sintética más pequeños o, más convenientemente, mediante expresión recombinante de una molécula de ácido nucleico que codifica dicho polipéptido como se ha descrito anteriormente en el presente documento.

Las moléculas de ácido nucleico para su uso de acuerdo con la invención pueden generarse de forma sintética, por ejemplo, por amplificación de una secuencia de ácido nucleico como se describe en el presente documento, tal como a partir de una biblioteca de ADNc apropiada.

5 Los polipéptidos, moléculas de ácido nucleico o anticuerpos descritos en el presente documento pueden usarse *in vitro*, por ejemplo, en un cultivo de células o de órganos, para afectar a las funciones inmunitarias en las células.

10 Como alternativa, los polipéptidos, moléculas de ácido nucleico o anticuerpos pueden usarse *ex vivo*, en partes o productos animales, por ejemplo órganos o sangre recogida, células o tejidos, particularmente cuando se contempla que estos se reintroducirán en el cuerpo del que proceden.

Sin embargo, los polipéptidos, moléculas de ácido nucleico y anticuerpos se prefieren para el uso *in vivo* como se analiza con más detalle más adelante.

15 Los polipéptidos y moléculas de ácido nucleico como se describen en el presente documento tienen aplicaciones para el tratamiento de diversos trastornos o afecciones como se describen más adelante en el presente documento. La presente invención, por lo tanto, se amplía a una composición farmacéutica que comprende un polipéptido o molécula de ácido nucleico como se define en las reivindicaciones adjuntas y uno o más excipientes y/o diluyentes farmacéuticamente aceptables para uso en terapia.

20 Por "farmacéuticamente aceptable" o "fisiológicamente aceptable" se entiende que el ingrediente tiene que ser compatible con otros ingredientes en la composición, así como fisiológicamente aceptable para el receptor.

25 El ingrediente activo para la administración puede modificarse de forma apropiada para su uso en una composición farmacéutica. Por ejemplo, los compuestos usados de acuerdo con la invención pueden estabilizarse frente a la degradación mediante el uso de derivados como se ha descrito anteriormente.

30 El ingrediente activo también puede estabilizarse en las composiciones, por ejemplo, mediante el uso de aditivos apropiados tales como sales o no electrólitos, acetato, SDS, EDTA, tampones citrato o acetato, manitol, glicina, HSA o polisorbato.

La molécula de ácido nucleico o polipéptido puede estar presente en dichas composiciones como único ingrediente activo o puede combinarse con otros ingredientes, particularmente otros ingredientes activos, por ejemplo, para aumentar el efecto terapéutico o para hacer la composición más apetecible para el consumidor.

35 La composición que comprende el polipéptido también puede comprender zonasa o una enzima relacionada. Como se menciona en el presente documento, la zonasa es una preparación que contiene una enzima, donde dicha preparación presenta una banda proteica en el análisis de SDS-PAGE con un peso molecular de aproximadamente 28 kDa, que puede obtenerse con un método que comprende las etapas de:

- 40
- a) suspender huevos de salmón en un volumen mínimo de agua;
  - b) inducir la eclosión rápida sincronizada de dichos huevos de salmón;
  - c) filtrar los huevos del salmón eclosionados para obtener fluido de eclosión;
  - 45 d) añadir urea sólida a dicho fluido de eclosión para permitir la disociación de los fragmentos de la cáscara de huevo de salmón y someter dicho fluido a centrifugación de baja velocidad;
  - e) purificar adicionalmente dicha zonasa sometiendo el sobrenadante de centrifugación a filtración en gel; y
  - f) purificar adicionalmente dicha zonasa por cromatografía de afinidad en una columna de Superose 6B® modificada con benzamidina, donde dicha cromatografía de afinidad se realiza realizando lavados con sal concentrada seguidos de elución con dioxano, en solución salina concentrada, para extraer las zonas unidas a la matriz de cromatografía o a estructuras macromoleculares;
- 50

donde dicha zonas tiene las siguientes propiedades:

- 55
- a) escinde la cromozima X;
  - b) se inhibe por benzamidina;
  - c) escinde los enlaces peptídicos con arginina;
  - d) permanece activa en presencia de urea 8 M, concentraciones molares de sal, agua destilada y disolventes orgánicos, preferentemente dioxano o propanol; y
  - 60 e) retiene la actividad enzimática en solución a temperatura ambiente durante 50 días.

La zonasa puede prepararse como se describe en los ejemplos y añadirse a la composición o puede representar una "impureza" después de la preparación de un polipéptido desvelado en el presente documento a partir de fuentes naturales. En composiciones que comprenden tanto el polipéptido desvelado en el presente documento como zonasa, el polipéptido puede estar presente en el intervalo del 1 al 100% de su peso total combinado y la zonasa puede constituir hasta un 0-99%. Preferentemente, el polipéptido está presente en un intervalo del 50 al 100%, por

ejemplo, >80, 90, 95, 96, 97, 98 o 99 y la zonasa está presente en un 0-50%, por ejemplo < 20, 10, 5, 4, 3, 2 o 1 % del peso combinado.

Las composiciones descritas en el presente documento son para su uso en terapia.

Como se ha mencionado anteriormente, los polipéptidos y moléculas de ácido nucleico presentan propiedades terapéuticas en el tratamiento de diversos trastornos autoinmunitarios e inflamatorios, particularmente de la piel, así como en el tratamiento de la piel dañada, por ejemplo, por el sol, el frío, una irradiación (por ejemplo, rayos X, tal como los usados para tratar cánceres) o como resultado de una herida.

Como se menciona en el presente documento, un "trastorno" se refiere a una alteración patológica subyacente en un organismo sintomático o asintomático con respecto a un organismo normal, que puede producirse, por ejemplo, como resultado de la infección de una imperfección genética congénita o adquirida.

Como se indica en los ejemplos, se ha ilustrado la utilidad de la leucolectina para tratar una diversidad de afecciones inflamatorias, autoinmunitarias y otras afecciones cutáneas. Puede esperarse una eficacia similar para tratar otras afecciones. Por ejemplo, la inflamación gastrointestinal crónica puede explicarse por una reacción excesiva al (o quizás en algunos casos como un) fallo de la reacción por las células dendríticas contra agentes nocivos. Por lo tanto, sería de esperar que la introducción de leucolectina, por ejemplo, por vía oral, tratara la inflamación crónica del tracto gastrointestinal, de la misma manera que la piel inflamada responde favorablemente a la introducción de leucolectina.

Las enfermedades autoinmunitarias pueden producirse como resultado de la expresión somática específica de versiones mutadas de antígenos normales de la superficie celular. Sin el deseo de imitarse por ninguna teoría, las leucolectinas pueden servir para proteger las células diana de un ataque.

Como se menciona en el presente documento, un "trastorno inflamatorio" es un trastorno en el que se observa inflamación en algún punto durante la progresión del trastorno y puede ser el único síntoma o uno de varios síntomas. El trastorno inflamatorio puede ser agudo o crónico y puede ser como se ha descrito anteriormente en el presente documento o como se describirá posteriormente. Los trastornos inflamatorios incluyen inflamación cardiovascular (por ejemplo, aterosclerosis, ictus), inflamación gastrointestinal (incluyendo úlceras, tales como úlceras gástricas o duodenales), trastornos inflamatorios hepáticos, inflamación pulmonar (por ejemplo, asma, lesión pulmonar inducida por ventilador), inflamación renal, inflamación ocular (por ejemplo, uveítis), inflamación pancreática, inflamación genitourinaria, trastornos neuroinflamatorios (por ejemplo, esclerosis múltiple, enfermedad de Alzheimer), alergia (por ejemplo, rinitis alérgica/sinusitis, alergias y trastornos cutáneos (por ejemplo, urticaria/irritación, angioedema, dermatitis atópica, dermatitis de contacto, psoriasis), alergias alimentarias, alergias a fármacos, alergias a insectos, mastocitosis), inflamación esquelética (por ejemplo, artritis, osteoartritis, artritis reumatoide, espondiloartropatías), infección (por ejemplo, infecciones bacterianas o víricas; trastornos inflamatorios orales (por ejemplo, periodontitis, gingivitis o estomatitis); llagas en las membranas mucosas; y trasplante (por ejemplo, rechazo de aloinjertos o xenoinjertos o tolerancia fetomaterna).

Los trastornos inflamatorios preferidos para el tratamiento de acuerdo con la presente invención son trastornos inflamatorios cutáneos tales como eccema, acné, rosácea, psoriasis y dermatitis de contacto, inflamación gastrointestinal y afecciones de las membranas mucosas tales como úlceras o llagas.

Como se menciona en el presente documento, un "trastorno autoinmunitario" es uno en el que la autoinmunidad benigna ha progresado a una autoinmunidad patogénica y puede ser como se describe anteriormente o más adelante en el presente documento. Las enfermedades autoinmunitarias incluyen síndrome de inmunodeficiencia adquirida (SIDA, que es una enfermedad viral con un componente autoinmunitario), alopecia areata, espondilitis anquilosante, síndrome antifosfolípido, enfermedad de Addison autoinmunitaria, anemia hemolítica autoinmunitaria, hepatitis autoinmunitaria, enfermedad autoinmunitaria del oído interno (EAOI), síndrome linfoproliferativo autoinmunitario (SLPA), púrpura trombocitopénica autoinmunitaria (PTA), enfermedad de Behcet, cardiomiopatía, esprúe celíaco-dermatitis herpetiforme; síndrome de de fatiga crónica y disfunción inmune (SFCDI), polineuropatía desmielinizante inflamatoria crónica (PDIC), penfigoide cicatricial, enfermedad de crioglobulinas, síndrome de Crest, enfermedad de Crohn, enfermedad de Degos, dermatomiositis juvenil, lupus discoide, crioglobulinemia mixta esencial, fibromialgia-fibromiositis, enfermedad de Graves, síndrome de Guillain-Barre, tiroiditis de Hashimoto, fibrosis pulmonar idiopática, púrpura trombocitopénica idiopática (PTI), nefropatía por IgA, diabetes mellitus dependiente de insulina, artritis crónica juvenil (enfermedad de Still), artritis reumatoide juvenil, enfermedad de Meniere, enfermedad del tejido conectivo mixta, esclerosis múltiple, miastenia grave, anemia perniciosa, poliarteritis nodosa, policondritis, síndromes poliglandulares, polimialgia reumática, polimiositis y dermatomiositis,  $\alpha$ -gamma-globulinemia primaria, cirrosis biliar primaria, psoriasis, artritis psoriásica, fenómeno de Raynaud, síndrome de Reiter, fiebre reumática, artritis reumatoide, sarcoidosis, esclerodermia (esclerosis sistémica progresiva (ESP), también conocida como esclerosis sistémica (SS)), síndrome de Sjogren, síndrome de la persona rígida, lupus eritematoso sistémico, arteritis de Takayasu, arteritis temporal/arteritis de células gigantes, colitis ulcerosa, uveítis, vitíligo y granulomatosis de Wegener.

Las "lesiones de la piel", como se menciona en el presente documento, incluyen la piel que se ha dañado por influencias externas tales como calor, irradiación (por luz de diversas longitudes de onda tales como rayos X o UV), frío, fricción, desgarro o una herida, por ejemplo, como resultado de un accidente o una operación quirúrgica. Como alternativa, la piel dañada puede producirse como resultado de una infección, o enfermedad o anomalía genética subyacente. La lesión puede presentarse como grietas, enrojecimiento, picor, inflamación, piel córnea, escamas, etc.

La presente invención proporciona el uso de un polipéptido, molécula de ácido nucleico o composición farmacéutica como se define en las reivindicaciones en la preparación de un medicamento para tratar o prevenir un trastorno autoinmunitario, un trastorno inflamatorio o lesiones de la piel en un animal.

En una alternativa adicional, la invención proporciona un polipéptido, molécula de ácido nucleico o composición farmacéutica como se define en las reivindicaciones para tratar o prevenir un trastorno autoinmunitario, un trastorno inflamatorio o lesiones de la piel en un animal.

En un aspecto preferido, la invención proporciona dichos usos, polipéptidos, moléculas de ácido nucleico o composiciones farmacéuticas para tratar o prevenir un trastorno autoinmunitario de la piel, un trastorno inflamatorio de la piel o lesiones de la piel, donde se administra a la piel, preferentemente de forma tópica, un polipéptido, molécula de ácido nucleico o composición farmacéutica como se describe en las reivindicaciones.

Preferentemente, dicho trastorno es eccema, acné, psoriasis, inflamación gastrointestinal (tal como en la enfermedad de Crohn, colitis ulcerosa y otra inflamación crónica), gingivitis e inflamación de la cavidad oral y esófago y dicha piel dañada está irradiada o inflamada, agrietada por el frío, quemada por el sol, dañada por calor o irradiación o como resultado de una herida.

Como se usa en el presente documento, "tratamiento" se refiere a la reducción, alivio o eliminación, preferentemente a niveles normales, de uno o más de los síntomas o efectos de dicho trastorno o lesión, por ejemplo, presencia o grado de lesión de la piel, por ejemplo, en relación con el tamaño de la herida, inflamación, enrojecimiento, picor, dolor etc. con respecto a los síntomas o efectos presentes en una parte diferente del cuerpo de dicho individuo no sometido a dicho tratamiento o en un individuo correspondiente no sometido a dicho tratamiento.

"Prevención" se refiere a la prevención absoluta, o reducción o alivio del grado de o tiempo (retraso) de la aparición de ese síntoma o efecto. Esto puede conseguirse, por ejemplo, por métodos de terapia génica, por ejemplo, el uso de anti-ARN o secuencias sin sentido.

El tratamiento o prevención de acuerdo con la invención, ventajosamente, puede combinarse con la administración de uno o más ingredientes activos que son eficaces para tratar o prevenir los trastornos o la lesión de la piel. De esta manera, las composiciones farmacéuticas para su uso de acuerdo con la invención pueden contener además uno o más de dichos ingredientes activos.

De acuerdo con otro aspecto adicional de la invención, se proporcionan productos que contienen uno o más polipéptidos o moléculas de ácido nucleico como se definen en las reivindicaciones y uno o más ingredientes activos adicionales como una preparación combinada para el uso simultáneo, separado o secuencial en terapia humana o animal.

Las composiciones para su uso de acuerdo con la invención pueden formularse de una manera convencional con uno o más vehículos, excipientes y/o diluyentes fisiológicamente aceptables, de acuerdo con técnicas bien conocidas en este campo usando ingredientes fácilmente adquiribles.

De esta manera, el ingrediente activo puede incorporarse, opcionalmente junto con otras sustancias activas como una preparación combinada, con uno o más vehículos, diluyentes y/o excipientes convencionales, para producir preparaciones galénicas convencionales tales como comprimidos, píldoras, polvos, grageas, bolsitas, obleas, elixires, suspensiones (como fluidos de inyección o infusión), emulsiones, soluciones, jarabes, aerosoles (como un sólido o en un medio líquido), pomadas, cápsulas de gelatina blanda y dura, supositorios, soluciones inyectables estériles, polvos envasados estériles y similares. También pueden usarse polímeros biodegradables (tales como poliésteres, polianhídridos, ácido poliláctico o ácido poliglicólico) para implantes sólidos. Las composiciones pueden estabilizarse mediante el uso de liofilización, subenfriamiento o Permazyme.

Son excipientes, vehículos o diluyentes adecuados lactosa, dextrosa, sacarosa, sorbitol, manitol, almidones, goma arábiga, fosfato cálcico, carbonato cálcico, lactosa cálcica, almidón de maíz, alginatos, tragacanto, gelatina, silicato cálcico, celulosa microcristalina, polivinilpirrolidona, celulosa, jarabe acuoso, agua, agua/etanol, agua/glicol, agua/polietilenglicol, propilenglicol, metilcelulosa, metilhidroxibenzoatos, propilhidroxibenzoatos, talco, estearato de magnesio, aceite mineral o sustancias grasas tales como grasa hidrogenada o mezclas adecuadas de los mismos. También pueden usarse agentes para obtener formulaciones de liberación sostenida, tales como carboxipolimetileno, carboximetilcelulosa, acetato ftalato de celulosa, o polivinilacetato.

Las composiciones pueden incluir además agentes lubricantes, agentes humectantes, agentes emulsionantes, agentes para aumentar la viscosidad, agentes de granulación, agentes disgregantes, agentes aglutinantes, agentes con acción osmótica, agentes de suspensión, agentes conservantes, agentes edulcorantes, agentes saporíferos, potenciadores de la adsorción (por ejemplo, agentes de penetración en la superficie o para administración nasal, por ejemplo, sales biliares, lecitinas, tensioactivos, ácidos grasos, quelantes), agentes bronceadores, disolvente orgánico, antioxidantes, agentes estabilizadores, emolientes, silicona, alfa-hidroxiácido, delmulcentes, agentes antiespuma, agentes hidratantes, vitaminas, aromas, espesantes iónicos o no iónicos, tensioactivos, cargas, espesantes iónicos o no iónicos, secuestrantes, polímeros, propelentes, agentes alcalinizantes, acidificantes u opacificantes, agentes colorantes y compuestos grasos y similares.

Las composiciones para su uso de acuerdo con la invención pueden formularse para proporcionar la liberación rápida, sostenida o retrasada del ingrediente activo después de la administración en el cuerpo, empleando técnicas bien conocidas en este campo.

Las composiciones pueden estar en cualquier forma de dosificación apropiada para permitir la liberación o para la dirección particular a células o tejidos, por ejemplo, como una emulsión o en liposomas, niosomas, microesferas, nanopartículas o similares en las que puede absorberse, adsorberse, incorporarse o unirse el ingrediente activo. Esto puede convertir de forma eficaz el producto en una forma insoluble. Estas formas en partículas pueden solucionar tanto los problemas de estabilidad (por ejemplo degradación) como los problemas de liberación.

Estas partículas pueden llevar moléculas de superficie apropiadas para mejorar el tiempo de circulación (por ejemplo, componentes del suero, tensioactivos, polioxamina908, PEG, etc.) o restos para la dirección específica, tales como ligandos para receptores celulares particulares. En este campo son bien conocidas técnicas apropiadas para la liberación de fármacos y para la dirección, y se describen el documento WO99/63215.

Se prefiere el uso de soluciones, suspensiones, geles y emulsiones, por ejemplo, el ingrediente activo puede transportarse en agua, un gas, un líquido acuoso, un aceite, un gel, una emulsión, una emulsión de aceite en agua o de agua en aceite, una dispersión o una mezcla de los mismos.

Las composiciones pueden ser para administración tópica (por ejemplo, en la piel), oral o parenteral, como por ejemplo por inyección. Como se ha mencionado anteriormente en el presente documento, las leucolectinas muestran muy poca variación entre diferentes especies. De hecho, se ha descubierto que son constantes en una población humana sana y, por lo tanto, esto permitiría el uso de leucolectina, por ejemplo, mediante la administración por inyección en sujetos, sin desencadenar una respuesta humoral de anticuerpos, de una manera similar a la proteína insulina que también es constante en la población humana. Por lo tanto, puede usarse una terapia sistémica por inyección de leucolectina para tratar las afecciones descritas en el presente documento, particularmente trastornos autoinmunitarios.

Sin embargo se prefieren las composiciones y la administración tópica, e incluyen geles, cremas, pomadas, pulverizaciones, lociones, salvas, barras, jabones, polvos, películas, aerosoles, gotas, espumas, soluciones, emulsiones, suspensiones, dispersiones, por ejemplo, dispersiones de vesículas no iónicas, leches y cualquier otras formas farmacéuticas convencionales de la técnica.

Las pomadas, geles y cremas pueden formularse, por ejemplo, con una base acuosa u oleosa con la adición de agentes espesantes y/o gelificantes adecuados. Las lociones pueden formularse con una base acuosa u oleosa y, en general, también contendrán uno o más agentes emulsionantes, dispersantes, de suspensión, espesantes o colorantes. Los polvos pueden formarse con la ayuda de cualquier base en polvo adecuada. Las gotas y soluciones pueden formularse con una base acuosa o no acuosa que también comprende uno o más agentes dispersantes, solubilizantes o de suspensión. Las pulverizaciones de aerosol convenientemente se suministran a partir de envases a presión, con el uso de un propulente adecuado.

Como alternativa, las composiciones pueden proporcionarse en una forma adaptada para administración oral o parenteral. De esta manera, las formas farmacéuticas alternativas incluyen comprimidos recubiertos o no, cápsulas, suspensiones y soluciones que contienen el componente activo opcionalmente junto con uno o más vehículos y/o diluyentes inertes convencionales, por ejemplo, con almidón de maíz, lactosa, sacarosa, celulosa microcristalina, estearato de magnesio, polivinilpirrolidona, ácido cítrico, ácido tartárico, agua, agua/etanol, agua/glicerol, agua/sorbitol, agua/polietilenglicol, propilenglicol, alcohol estearílico, carboximetilcelulosa o sustancias grasas tales como grasa hidrogenada o mezclas adecuadas de los mismos.

La concentración de ingrediente activo en las composiciones para su uso de acuerdo con la invención depende de la naturaleza del compuesto usado (es decir, la molécula de polipéptido o ácido nucleico), el modo de administración, el curso de tratamiento, la edad y peso del paciente, la indicación médica, el cuerpo o área corporal a tratar, y puede variarse o ajustarse de acuerdo con la elección. Sin embargo, en general, los intervalos de concentración para el compuesto descrito en el presente documento son del 0,0001, 0,0005, 0,001 o 0,01 al 25%, por ejemplo, del 0,0005 al 15%, por ejemplo, del 0,01 al 10%, tal como del 0,1 al 5, por ejemplo del 1 al 5% (p/p de la preparación final para administración, particularmente para administración tópica). Dichas concentraciones se determinan por referencia a

la cantidad del propio compuesto y, de esta manera, deberían realizarse las adaptaciones apropiadas para tener en cuenta la pureza de la composición. Las dosis individuales eficaces pueden estar en el intervalo de 0,1 a 100 mg/día, preferentemente de 2 a 10 mg/día, dependiendo del animal que se esté tratando, tomadas como una sola dosis.

La administración puede realizarse por cualquier método adecuado conocido en las técnicas medicinales, incluyendo por ejemplo la administración oral, parenteral (por ejemplo intramuscular, subcutánea, intraperitoneal o intravenosa), percutánea, bucal, rectal o tópica o administración por inhalación. Las formas de administración preferidas se administrarán por vía oral, o más preferentemente por vía tópica. Como se apreciará, la administración oral tiene sus limitaciones si el ingrediente activo se puede digerir. Para solucionar estos problemas, los ingredientes pueden estabilizarse como se ha mencionado previamente.

Se apreciará que como el ingrediente activo para realizar la invención toma una diversidad de formas, por ejemplo, molécula de ácido nucleico (que puede estar en un vector) o polipéptido, la forma de la composición y la vía de liberación variarán. Preferentemente, sin embargo, deben emplearse soluciones líquidas, cremas o suspensiones, particularmente, por ejemplo, para liberación oral o administración tópica.

Las moléculas de polipéptido o ácido nucleico pueden usarse para las indicaciones médicas mencionadas anteriormente. En los últimos métodos de terapia génica, las moléculas de ácido nucleico preferentemente se proporcionan en vectores que son adecuados para transfección/transformación como se ha descrito anteriormente, por ejemplo, vectores virales tales como adenovirus usando métodos de terapia génica conocidos en la técnica para aplicaciones médicas.

Los animales a los que pueden aplicarse o administrarse las composiciones incluyen mamíferos, reptiles, aves, insectos y peces (por ejemplo, salmón o bacalao). Preferentemente, los animales a los que se aplican las composiciones para su uso de acuerdo con la invención son mamíferos, particularmente primates, animales domésticos, ganado y animales de laboratorio. De esta manera, los animales preferidos incluyen ratones, ratas, conejos, cobayas, gatos, perros, monos, cerdos, vacas, cabras, ovejas y caballos. De una forma especialmente preferida, las composiciones se aplican o se administran a seres humanos.

Los siguientes ejemplos se proporcionan solo a modo de ilustración y las Figuras mencionadas en ellos son las siguientes:

La Figura 1 muestra (A) la filtración en gel de zonasa de salmón parcialmente purificada para producir leucolectina. Una columna Superdex 16/60 (GE Healthcare) se conectó a un sistema de FPLC y se usó con un flujo de tampón de 1 ml/min para 120 ml. La elución de las proteínas de la columna se supervisó por UV 280 nm (derecha, eje y) y se recogió en fracciones de 1 ml. La actividad de la zonasa se midió por escisión de Chromozym X a DO406. El peso molecular nativo de la zonasa abarca un amplio intervalo con un pico de aproximadamente 50 kDa. PM de la lectina = 30 kDa, y (B) identificación de (leuko-) lectina de salmón por inmunotransferencia. Análisis de inmunotransferencia de proteínas de fracciones seleccionadas obtenidas a partir de la filtración en gel anterior usando anticuerpo anti-(leuko-) lectina basado en epítipo (véase la Fig. 3). Los números de la fracción se muestran encima de las manchas de inmunotransferencia. A la izquierda están indicados los marcadores de peso molecular (kDa). El peso molecular estimado de la lectina corresponde al obtenido por otros métodos (véase la Fig. 3).

La Figura 2 muestra la identificación de células productoras de lectina (lectocitos) en embriones de salmón atlántico donde los paneles A, B y C representan secciones transversales del cuerpo de los embriones de salmón atlántico que se habían fijado con formalina e incluido en bloques de parafina. La presencia de leucolectina se determinó por medio de un anticuerpo anti-LL policlonal de conejo usando un método de tinción con inmunoperoxidasa (A) o inmunofluorescencia indirecta (B y C). Las células que reaccionan con anticuerpo antilectina pueden verse como oscuras (A) o claras (B y C) con respecto al efecto de fondo, respectivamente. Se ven numerosas células individuales inmunorreactivas con características citológicas específicas en la epidermis sin restricciones espaciales a lo largo del eje embrionario antero-posterior. Las flechas en A y B indican células relativamente grandes (lectocitos), fácilmente distinguibles de las células epidérmicas circundantes. Sus núcleos son basales, mientras que los citoplasmas están rellenos de leucolectina de tal manera que las células tienen unas protuberancias con forma de seta en el espacio perivitelino. El Panel 2C muestra sus contenidos granulares.

La Figura 3 muestra el alineamiento de leucolectinas de ser humano (SEC ID N°: 5), salmón, pollo, bacalao y dos traducciones de los dos cóntigos que constituyen la familia UniGene Ssa.23163, una colección EST de salmón. De los dos cóntigos, el cóntigo 1 es el que más se parece a la estructura primaria de las leucolectinas. Sin embargo, la diferencia entre los dos cóntigos no es grande, solo de aproximadamente un 80 %.

La Figura 4 muestra (A) dominios SMART de secuencias similares a la leucolectina humana. Se presentan las secuencias de máxima puntuación en una búsqueda BLASTP usando WU-BLAST 2. Las secuencias se presentan en orden desde la más parecida a la menos parecida. a - leucolectina-1 humana (SEC ID N°: 5), b - leucolectina embrionaria de salmón, c - leucolectina leucocítica de salmón (fragmento), d - leucolectina de pollo

(fragmento), e -leucolectina de bacalao (fragmento), f - lectina de huevo de carpa común (Ac#:P68512), g - ranaspumina-6 de rana tungara (Ac#:B5DCK6), h - Zgc de pez cebra: proteína 173443 (Ac#: A8E4Z1). Obsérvese que las secuencias de las leucolectinas leucocítica de salmón, leucocítica de pollo y leucocítica de bacalao son solo parciales, y no secuencias de longitud completa. La versión de longitud completa de estas secuencias probablemente se parecerá a la estructura total encontrada en la leucolectina de los leucocitos humanos (a) y embriones de salmón (b). Además, el patrón de leucolectina, con 5 dominios TECPR consecutivos, aún se conserva para la secuencia de rana tungara, pero no para la secuencia del pez cebra donde solo se encuentran 4 dominios TECPR, y (B) una representación gráfica de leucolectina con 5 dominios TECPR. Los dominios TECPR muy probablemente están compuestos de 4 cadenas beta que generan 2 láminas beta. Hay TRES enlaces disulfuro: uno es interno en TECPR N° 4, los otros dos están en bucles de conexión, como se indica. En general, los sitios de reconocimiento de azúcar se encuentran en dichas lectinas en las áreas entre los dominios TECPR. La lectina más relacionada con la leucolectina tiene una especificidad de azúcar conocida principalmente por N-Acetilglucosamina. La supuesta especificidad de reconocimiento de azúcar para la leucolectina aún no se ha establecido.

La Figura 5 muestra la presencia de lectina perivitelina en leucocitos. En la primera dimensión se usó una tira de IPG que abarcaba el intervalo de pH de 4-7, mientras que durante la electroforesis de la segunda dimensión se utilizó un gel de poliácridamida al 12,5%.

A: En los leucocitos de salmón, se observó que solo dos manchas reaccionaban positivamente con el anticuerpo policlonal anti-lectina. Su peso molecular era de aproximadamente 26 kDa, con un pI de aproximadamente 6,5. La leucolectina también puede identificarse en preparaciones de leucocitos en otras especies.

B: En una preparación de salmón, se encontraron varias manchas inmunorreactivas (con los números 1-8), con un peso molecular de aproximadamente 30 kDa, que abarcaba un intervalo de pH de 4,9 a 6,5.

La Figura 6 muestra la coexpresión de leucolectina en el linaje mieloide, específicamente en el linaje de monocitos-macrófagos en el pez cebra. (A) Perfil de expresión del gen de lectina y L-plastina. Se observó que el pez cebra expresaba la misma lectina que se encuentra en salmón. Se usó una sonda de ARNm de lectina marcada con fluoresceína junto con una sonda marcada con DIG para la sonda de ARNm de L-plastina, que es un marcador específico del linaje de monocitos-macrófagos. La hibridación *in situ* con doble marcaje revela la coexpresión del ARNm de lectina y L-plastina en las mismas células durante la embriogénesis del pez cebra. La edad de los embriones en horas trascurridas desde la fertilización (pf) se muestran en las esquinas inferiores izquierdas. Las señales más claras indican la expresión del gen de la L-plastina; las señales más oscuras indican la expresión de lectina.

El Panel A (vista lateral) muestra la coexpresión del ARNm de lectina y L-plastina en las mismas células, que forma una población axial y paraxial dispersa en la parte anterior y lateral de la yema. Con más aumentos (A2) se revela la morfología típica de los leucocitos de estas células. Los datos en A1 demuestran la expresión de estos dos genes en bandas bilaterales de células en el mesodermo de placa lateral anterior (flechas). El Panel B revela los mismos hallazgos a las 21 h.

El Panel C (24 h pf) demuestra la coexpresión de los genes de L-plastina y lectina en la masa celular intermedia posterior (dos flechas superiores), en la región de la vena ventral (flecha inferior), en un número significativo de células dispersadas a lo largo del cuerpo y la cabeza del embrión. A las 24 h pf (C2) la mayoría de las células coexpresan estos dos genes (véanse las flechas), pero muchas células parecen expresar únicamente un gen o el otro. Las flechas en C3 indican células que expresan solo el ARNm de la L-plastina y (B) la colocalización celular de la proteína lectina y el ARNm de L-plastina. Se analizaron embriones de pez cebra (22 h pf) por inmunohistoquímica combinada con un ensayo de hibridación *in situ* para la proteína lectina (señal más oscura) y el ARNm de L-plastina (señal más clara). La hibridación *in situ* usó ARNc marcado con DIG específico para el gen de la L-plastina, junto con inmunohistoquímica con anticuerpo policlonal de conejo anti-lectina.

Panel A: vista frontal del embrión. La flecha apunta a un grupo de células marcadas por expresión común de proteína lectina y ARNm de L-plastina. A1: más aumentos del grupo de células que demuestra la coexpresión del gen de la L-plastina y la proteína lectina en la misma célula (flechas negras). Obsérvese unas pocas células, de pequeño tamaño con forma redondeada que se parecen a los linfocitos maduros cerca de otra célula con un núcleo segmentado, una característica morfológica típica de los leucocitos neutrófilos maduros. Los Paneles B, C, D y E ilustran poblaciones de células dispersas a lo largo de todo el embrión del pez cebra que coexpresan proteína lectina y el ARNm de la L-plastina (flecha).

La Figura 7 muestra la familia de la leucolectina. Se ensamblaron proteínas leucolectina disponibles procedentes de ser humano (SEC ID N°: 5), salmón, pollo y bacalao (alineamiento weblogo). La correspondencia total de restos de AA en una posición se representa por letras grandes, indicando la reducción del tamaño un menor grado de correspondencia. Múltiples letras indican variaciones.

La Figura 8 muestra fracciones teñidas con azul de coomassie después de la purificación de leucolectina recombinante por cromatografía en columna. Las calles son 1 - sobrenadante, 2 - flujo eluido, 3 - lavado, 4 - F1, 5 - F2, 6- mezcla de F3, F4 y F5, 7- F6 y 8 - patrón.

La Figura 9 muestra (A) experimentos para identificar y establecer una secuencia génica para la leucolectina de salmón.

El Panel A ilustra dos manchas duplicadas positivas identificadas en el filtro nº 6, paneles nº 5 y 2, que corresponden a los números de placa exactos 257 y 176, y coordenadas de pocillo de los clones L-7 y F-7 respectivamente.

El Panel B demuestra una mancha duplicada positiva identificada en el filtro nº 3, panel nº 6, que corresponde al número de placa exacto 144 (basándose en la hoja localizadora de placas), y coordenadas de pocillos del clon A-12, y

(B) muestra la secuencia genómica deducida del gen de la leucolectina (2323 pb). Se indica el sitio de los 5 exones y, como consecuencia, también los intrones. Se indica la localización del inicio de la transcripción para este gen (predicción del promotor, basado en BDGP Neural Network Promoter Prediction, "P", inicio de la transcripción previsto "ST", caja TATA "T") y las secuencias que especifican la poliadenilación ("PA") en el producto génico.

La Figura 10 muestra (A) un mapa que representa los cromosomas humanos, las posiciones de los aciertos BLASTP (véase la tabla) se indican por flechas, (B) mapa de cobertura de leyendas: Una visualización de la posición de los aciertos BLASTP (barras) y la cobertura total de la secuencia problema (Leucolectina-1 humana). Aparentemente, la suma de secuencias encontradas cubren la secuencia entera de la leucolectina, pero los datos claramente no implican que este sea el gen de la leucolectina, (C) tabla que representa los aciertos BLASTP y su posición en el genoma humano usando la leucolectina-1 humana como secuencia problema.

La Figura 11 muestra los efectos de la leucolectina de salmón en epitelio humano. Cultivo de control = A; cultivo expuesto a leucolectina = B. Las flechas indican células grandes que aparecen en la capa basal solo después de la exposición a la leucolectina.

La Figura 12 muestra la expresión de la proteína leucolectina en leucocitos. El panel A muestra la PAGE 2D al 12 % de proteína purificada, ~ 2 µg, procedente de leucocitos de salmón. El panel B muestra la PAGE 2D al 15 % de proteína purificada, 0,8 µg, procedente de leucocitos humanos de Lymphoprep™. Las membranas se trataron con anticuerpo policlonal primario contra LL de salmón, antes de tratarse con anticuerpo de cabra anti-conejo y se visualizaron por un sistema de detección mejorado por ECL.

La Figura 13 muestra transcritos de leucolectina de salmón identificados en transferencias de Northern. El panel A muestra el ARN total de embriones de salmón atlántico (370 dd) fraccionado en agarosa al 1,2 % en presencia de formaldehído, y tratados con una ribosonda antisentido marcada con DIG específica para LL (720 pb). El panel B muestra el análisis de transferencia de Northern de ARNm (purificado por perlas magnéticas poliT). La hibridación usó ribosondas marcadas con DIG generadas a partir de la secuencia codificante parcial de LL. El panel C muestra la membrana de nitrocelulosa (B) que se trató con ribosonda con sentido marcada con DIG. Las flechas grises significan el transcrito presente; la flecha negra indica su ausencia. Se usó el marcador I de ARN DIG, 0,3-6,9 kb (Roche).

La Figura 14 muestra un alineamiento Clustal W de LL con proteínas de hélice enrollada β relacionadas.

La Figura 15 muestra un modelo 3-D de LL. El panel izquierdo muestra dos vistas de una representación 3D de un modelo 3D de hélice enrollada β de 5 láminas de la leucolectina basándose en la estructura de la taquilectina 1 procedente de *Tachypleus tridentatus* (Biesel *et al.*, 1999, EMBO J. 18, pág. 2313-2322). El modelo se genera usando el software PyMol v 0.99. Los restos de los péptidos del epítipo están dibujados como líneas negras gruesas en el cuerpo de la proteína. Además, las posiciones de los 5 sitios de unión a carbohidrato previstas por Biesel *et al.* (1999) están indicadas como estructuras de hexosa oscuras tenues pero sólidas.

El panel derecho muestra representaciones de los dominios de hélice enrollada previstos en LL (parte superior) y en FEL (parte inferior), en comparación con el dominio de hélice enrollada consenso.

La Figura 16 muestra la amplificación de la secuencia de ADNc de la leucolectina (LL) de pez cebra de longitud completa. El panel A muestra amplicones que se obtuvieron usando varios pares de cebadores, y se clonaron por RT-PCR. El ARNm usado para la transcripción inversa se extrajo de embriones a 24 hpf. Marcador de ADN: escala de 100 pb (New England Biolabs). El panel B muestra que la PCR 5'RACE dio como resultado la amplificación de dos productos distintos, respectivamente -500 pb y -400 pb (generados después de una segunda vuelta de amplificación por PCR usando el cebador inverso específico de gen y el cebador anidado directo GeneRacer). Los ARNm se recogieron de las fases de desarrollo indicadas. El panel C muestra que la secuencia de ADNc de longitud completa de LL del pez cebra tiene 1240 nt que incluyen una fase de lectura abierta (ORF) de 765 nt. La ORF comprende cinco exones que están dibujados a escala y se muestran como recuadros negros. La región 5'UTR se presenta como una línea continua. Se muestra la secuencia de nucleótidos de la 5'UTR. El sitio de inicio de la traducción ATG en la posición +1 y el segundo ATG potencial en la posición +94 están destacados. El panel D muestra la expresión de la LL truncada en diversas fases del desarrollo. El panel E muestra que las secuencias de salmón y de pez cebra están muy conservadas incluso en la 3'UTR.



La Figura 17 proporciona un resumen de los exones, dominios y variabilidad de secuencia de leucolectina de salmón. Se muestran los datos de secuencia para diferentes leucolectinas en relación con las posiciones de restos de AA variables, restos de cisteína (C) y dominios TECPR (de la base de datos SMART). Las secuencias de la leucolectina difieren únicamente en siete posiciones, marcadas con recuadros. Cinco de estos restos variables aparecen dentro de dominios TECPR, mientras que dos se encuentran cerca del extremo C. Debajo se muestran variaciones incompatibles con la leucolectina-1 o la leucolectina-2, como leucolectina-3, donde solo dos posiciones están definidas (nº 2 y nº 245). Aunque las otras posiciones no están marcadas, los datos indican que la mayoría de estas posiciones son casi idénticas a las de las otras leucolectinas. Como ocurre con la lectina de huevo de pez (FEL), está previsto que los 3 puentes disulfuro existentes en la LL conecten, respectivamente, con la primera y la última cisteínas, mientras que la segunda y penúltima cisteínas forman un segundo puente, y las dos cisteínas (intermedias) dentro del exón 3 forman el último puente disulfuro. Estas cisteínas están marcadas en gris y subrayadas. La cisteína en el propéptido no implicada en los puentes disulfuro está en el exón 1. Obsérvese la ausencia de bucles de interconexión entre los primeros tres dominios y la ausencia de variación en dos bucles que interconectan los dos últimos dominios TECPR. Tres dominios TECPR están codificados por un solo exón (2 y 3), mientras que dos dominios abarcan dos exones. La fuerte correlación de los extremos de los exones y los dominios es destacable. En cuatro de cinco dominios, un resto de Trp ocupa el tercer AA del dominio, mientras que en el primer dominio TECPR previsto, W es el resto inicial. Finalmente, los asteriscos marcan restos en los que se han encontrado AA variantes en la leucolectina-1, en la leucolectina-2 o en ambas.

La Figura 18 muestra la proteína LL en el fluido de eclosión de embriones de la trucha arco iris (*Oncorhynchus mykiss*) detectados en transferencias de Western. Los patrones de PM se muestran a la izquierda (BioRad 161-0373). Las proteínas del fluido de eclosión de la trucha arco iris (calle 1) y la proteína leucolectina de salmón se prepararon por cromatografía de afinidad (calle 2) utilizando como sonda anticuerpo antileucolectina. En ambos casos se encuentra una proteína de ~ 26 kDa, que corresponde al Pm de LL.

La Figura 19 A, el panel de la izquierda muestra la leucolectina de bacalao entre proteínas encontradas en fluido de eclosión de bacalao. El fluido de eclosión de bacalao se analizó por SDS PAGE al 15 %, y sus proteínas constituyentes se visualizaron por tinción con plata. Calle 1: marcador de proteína (BioRad's Dual color 161-0374). Calle 2: proteínas del fluido de eclosión. El panel de la derecha muestra el fluido de eclosión de bacalao, separado por análisis de SDS PAGE al 15 %, transferido sobre una membrana de nitrocelulosa y tratado con diluciones apropiadas de anticuerpo policlonal de conejo anti-IgG de leucolectina purificado por afinidad, como sonda. Calle 1: marcador de proteína (BioRad's Dual color 161-0374). Calle 2: alícuota de fluido de eclosión. La localización de la leucolectina se precisó usando anticuerpo secundario de cabra anti-conejo marcado con peroxidasa de rábano picante (HRP), mejorado por el sistema de detección ECL. Se detectó una proteína inmunorreactiva principal de aproximadamente 26 kDa. B, muestra la SDS PAGE equivalente y la membrana para proteínas de fluido de eclosión de *Oikopleura dioica*.

#### Ejemplo 1: Identificación y caracterización de leucolectina

##### Aislamiento de proteína

Durante el transcurso del análisis de los componentes del fluido de eclosión de salmón, se identificó una nueva proteína presente en embriones.

En el documento WO99/29836 (particularmente en el ejemplo 1 del método descrito, pero opcionalmente sin la etapa de urea) se proporciona un método para preparar zonasa parcialmente purificada que puede usarse como material de partida para el aislamiento de un polipéptido desvelado en el presente documento.

Tanto la zonasa como la leucolectina se purificaron a partir del fluido de eclosión de salmón. Para mejorar la concentración de proteína del fluido de eclosión, se transfirieron huevos de salmón a volúmenes mínimos de agua antes de la eclosión. Puede inducirse una eclosión altamente sincrónica por temperaturas elevadas (ambiente), o por desoxigenación (Oppen-Berntsen *et al.* 1990, Aquaculture, 86, pág. 417-430), que produce un pequeño volumen de preparación muy concentrada de zonasa bruta y proteínas asociadas.

La purificación inicial de la zonasa implicaba la filtración de huevos de salmón eclosionados a través de estopilla. Este filtrado puede congelarse durante años sin degradación significativa de la zonasa, antes de descongelarse y emplearse para una purificación de proteínas adicional. Este hecho simplifica en gran medida la producción de un material de partida para purificar la zonasa de salmón y las proteínas asociadas, incluyendo la leucolectina.

La siguiente etapa opcional implicaba el ajuste del filtrado de proteína a urea 4 M, para disociar fragmentos de la cáscara de huevo de salmón, que permitía su retirada junto con el desecho extraño por centrifugación a baja velocidad (15.000 g; 2 x 15 min). Este material no mostró ningún signo de oclusión de las columnas, lo cual es característico de materiales brutos preparados de forma diferente de lo que se ha descrito anteriormente. La

preparación de proteína bruta fue adecuada para la purificación por técnicas cromatográficas convencionales y la zonasa puede purificarse adicionalmente como se describe en el ejemplo 7.

La leucolectina del fluido de eclosión puede aislarse junto con la zonasa. A partir de preparaciones de zonasa parcialmente purificadas (como se ha descrito anteriormente), la leucolectina puede aislarse por cromatografía de exclusión ya que la zonasa en su forma nativa es sustancialmente más grande que la leucolectina. Para una primera separación, son suficientes columnas Superdex 16/60 con las condiciones descritas en la leyenda de la Figura 1, después de lo cual la zonasa puede retirarse por cromatografía de afinidad en columnas de Benzamidine-Sepharose.

Para preparaciones a gran escala también es adecuado el uso de ultrafiltración, ya que la zonasa en su forma nativa no atraviesa de forma significativa los ultrafiltros con una exclusión de tamaños de 100 kDa a diferencia de la leucolectina.

Los tampones usados son Tris milimolar (por ejemplo 10 mM) a un pH aproximadamente neutro o ligeramente alcalino (pH 7,5 - 8,5), que contiene NaCl 5 mM.

También se prevé la extracción a gran escala de leucocitos humanos de productos que están caducados en bancos de sangre, mediante el uso de extracción en disolventes orgánicos (por ejemplo acetona al 80 %), porque la leucolectina permanece en solución durante este procedimiento que precipita la mayoría de las otras proteínas. La purificación final y los métodos cromatográficos y de filtración son como se ha indicado anteriormente.

También se descubrió que la proteína aislada se expresaba en gametos y en el cigoto y además, también en las primeras fases embrionarias (durante la ontogenia del sistema hemopoyético) y finalmente en los leucocitos.

Se observó que la proteína previamente no identificada se copurificaba con la zonasa. El tamaño de esta nueva proteína, estimado por cromatografía en condiciones nativas, era solo de 30 kDa. (Fig. 1).

#### *Expresión de proteína*

La proteína aislada se usó para generar un anticuerpo policlonal. En las células se analizó la presencia de proteína por medio de un anticuerpo policlonal de conejo anti-LL usando un método de tinción con inmunoperoxidasa o inmunofluorescencia indirecta. Los resultados se muestran en la Figura 2.

Mediante el uso del anticuerpo, se descubrió que la nueva proteína se originaba en un tipo celular diferente de las células que producen coriolisininas (enzimas de eclosión). La morfología de estas células individuales (denominadas *lectocitos*) se muestra en la Figura y es claramente distinta de las células de la glándula de eclosión (individuales) que expresan coriolisininas.

#### *Análisis de la secuencia*

La nueva proteína se sometió a caracterización convencional tal como determinación de sus péptidos tripticos, seguido por secuenciación de Edman directa parcial. A partir del puzzle de péptidos, las construcciones de ensayo y error de cebadores degenerados para detectar un ARNm en embriones de peces para producir un ADNc, fueron al principio poco satisfactorias hasta que se estableció el cebador inverso simplemente de forma que correspondiera a una cola de poliA. De esta manera, y con el uso de múltiples series nuevas de cebadores según se iban identificando, se pudieron encontrar al menos dos ADNc de aproximadamente 1200 NT. Estos resultaron ser muy similares en secuencia, con solo unos pocos restos de AA variantes. En ambos casos, estaban ausentes algunas de las secuencias N-terminales.

Usando la metodología N-RACE posteriormente se estableció que las secuencias anteriores pertenecían a una proteína de 255 aminoácidos, con un propéptido de 19 AA y con un resto de triptófano N-terminal poco habitual en la proteína procesada. Este último resultado explicaría las grandes dificultades encontradas para obtener una secuencia N-terminal por análisis directo. Si embargo, esto se consiguió finalmente para confirmar los restos N-terminales virtuales.

Usando la secuencia N-terminal de la proteína para buscar las bases de datos EST, se identificó una secuencia de proteína que representaba claramente la nueva proteína (Fig. 3).

La búsqueda en bases de datos de proteínas parecidas a la nueva proteína identificaba solo un candidato algo similar (Leren, no publicado). Esta es una lectina de huevo de pez aislada a partir de carpa (Galliano *et al.*, 2003, Biochem., J., 376, pág. 433-440), con una similitud global de un 48 % como máximo encontrada por alineamiento manual de secuencias.

El análisis bioinformático sugirió que la nueva proteína era una lectina relacionada con las tectoninas (Leren, no publicado). Al recopilar una lista de todas las proteínas con estructura de dominios de TECPR similar, la nueva

lectina de los presentes inventores parece representar un nuevo tipo de lectina (Figs. 4A y B). La secuencia de AA virtual de la nueva lectina contenía cinco dominios TECPR que mostraban alguna similitud con las taquilectinas de los invertebrados inferiores (Shigenaga *et al.*, 1993, J. Biochem (Tokio), 114, pág. 307-316).

- 5 La forma propéptido de la nueva lectina contiene un péptido de 19 AA, lo cual sugería su dirección a los lisosomas y la posible secreción posterior (es decir, en el espacio perivitelino).

10 La secuencia de AA virtual de la nueva lectina permitió el desarrollo de anticuerpos con especificidad de epítipo, que a su vez permitieron identificar muchas (2 - 8) isoformas parecidas de la proteína (Fig. 5), dependiendo del tejido analizado.

15 Una posibilidad es que la lectina encontrada en las células de las glándulas de eclosión (Fig. 6B) se secrete en el espacio perivitelino junto con las coriolisinas, y que estas posean modificaciones post-traduccionales que aumentan su tamaño aparente y que reducen su pl aparente.

20 Como el salmón es un pez tetraploide, y como los datos de los presentes inventores indican más de un gen para la leucolectina del salmón, la serie de 8 restos de leucolectina en preparaciones de zonasa parcialmente purificadas puede indicar tres modificaciones crecientes de dos proteínas leucolectinas ligeramente diferentes, para dar un total de 4 x 2, u 8 leucolectinas diferentes. Las causas de variación necesitan verificarse empíricamente.

25 El PM estimado de la lectina es de aproximadamente 25-30 kDa. El pl estimado para la lectina de salmón correspondiente es de aproximadamente pH 6,5. Los valores de pl observados (Riste, no publicado) eran de pH 6,5 a 4,9 en las lectinas perivitelinas de salmón y de pH 6,4 a 6,6 en las lectinas leucocíticas de salmón (la lectina se identificó por técnicas de transferencias de Western).

30 *Expresión durante el desarrollo embrionario.*

Usando sondas de nucleótidos para localizar la expresión de la nueva proteína durante el desarrollo embrionario de peces, se encontró la proteína presente en tipos específicos de células epidérmicas de embriones de peces. Además, se encontró la expresión de la proteína tanto en los gametos como en el cigoto. Además, se demostró que la proteína estaba ausente en la mayoría de las células de los organismos, pero se expresaba en unas pocas células embrionarias, lo que sugería que la proteína podía estar asociada a la mielopoyesis (Figs. 5 y 6). Dicho patrón de expresión es único. De esta manera, se eligió en nombre "*leucolectina*" (abreviado LL) para la proteína.

35 *Expresión en diferentes especies*

Una investigación de la sangre desde los peces hasta seres humanos pasando por los pollos reveló que las leucolectinas están presentes en los leucocitos a lo largo de todo el linaje de los vertebrados. Su presencia también pudo detectarse en invertebrados enteros usando el anticuerpo policlonal de los presentes solicitantes. El descubrimiento de leucolectinas en leucocitos (humanos) sugirió inmediatamente posibles funciones para las leucolectinas. Las leucolectinas aparecen en células que realizan funciones inmunitarias.

40 Lo más sorprendente fue la extrema falta de variabilidad de secuencia de las leucolectinas durante la evolución (Fig. 7). Usando la tecnología de par de cebadores convencional se pudieron detectar ADNc a lo largo de todo el linaje de los vertebrados. El análisis reveló una secuencia extremadamente conservada desde los peces hasta los seres humanos pasando por los pollos, lo que indicaría funciones esenciales para dichos genes. La variación entre la secuencia de la leucolectina 1 de peces y la de seres humanos es de aproximadamente un 4 % (10 AA variantes), o una similitud del 96 %, donde la mitad de las variaciones observadas eran sustituciones de AA conservativas. Por comparación, la única proteína conocida ligeramente similar es una lectina de huevo de peces (FEL) obtenida a partir de la carpa común (con una similitud menor del 48 %).

45 Se han identificado dos secuencias de proteína leucolectina adicionales en seres humanos, la leucolectina-2 (SEC ID N°: 16) y la leucolectina-3 (SEC ID N°: 17 y 18). La secuencia de la leucolectina-2 es similar a la de la leucolectina-2 de salmón, pero la secuencia de leucolectina-2 humana comprende dos insertos únicos. En comparación con la secuencia de la leucolectina de salmón, los restos de YDLG (posiciones n° 216-20) se han reemplazado por DLRPG y el extremo C-terminal difiere de la posición n° 221, donde RLWVV se ha reemplazado por HPWVC seguido de 28 AA únicos, particularmente DHTGWSNIPMGLLMGHVDLRPGPSLGRQLQVCRH. Los últimos cuatro aminoácidos son similares al extremo C-terminal de salmón, donde VCTH se ha reemplazado por VCRH.

60 La secuencia de la leucolectina-3 humana muestra similitud con la secuencia de la leucolectina-1 de salmón. La SEC ID N°: 17 es una versión truncada de la proteína y la SEC ID N°: 18 muestra la secuencia sin el péptido señal. La secuencia de nucleótidos que codifica la secuencia de la proteína truncada se muestra en la SEC ID N°: 19.

65 Se han identificado variantes adicionales de la secuencia de la leucolectina-3 humana y se han denominado leucolectina-4 (SEC ID N°: 20) y leucolectina-5 (SEC ID N°: 21-23). La leucolectina-4 difiere de la leucolectina-3 por

un solo aminoácido en la posición nº 88, donde la leucolectina-3 comprende un resto de ácido glutámico y la leucolectina-4 comprende un resto de asparagina. La leucolectina-5 (SEC ID N°: 21) comprende varias sustituciones de aminoácidos cuando se compara con las variantes leucolectina-3 y -4. También se ha identificado una variante truncada del polipéptido de la leucolectina-5 (SEC ID N°: 23). Una parte preferida de la variante de la leucolectina-5 se presenta como SEC ID N°: 22.

El resumen de la estructura de la LL variante (Fig. 7) define la nueva familia LL de proteínas. Se observan aproximadamente un 2 % de cambios no conservativos de AA desde los peces hasta los seres humanos. Claramente, los genes de LL codifican proteínas que están sometidas a una selección evolutiva extrema para mantener la invariabilidad de la secuencia. Esto a su vez indica funciones importantes, pero aún desconocidas, que impiden que se produzcan y se mantengan la mayor parte de las variaciones aleatorias en los productos génicos durante la evolución de las especies.

#### *Preparación de leucolectina recombinante*

Se clonó ADNc de leucolectina de salmón a partir de embriones. Los cebadores de PCR se diseñaron para que contuvieran el sitio de restricción NcoI (directo) y ACC65 I (inverso). El vector pETM-60 se digirió con NcoI/ACC65I y los productos de PCR digeridos se sometieron a ligamiento durante una noche. Después de la amplificación del plásmido y la secuenciación, los vectores de expresión, pETM-60 que contenían el inserto de leucolectina con la secuencia verificada se utilizaron para transformar células competentes BL21 (DE3)p y las colonias se usaron para inocular 5 ml de LB (= caldo Luria) para preparar soluciones madre de glicerol bacterianas.

*Cultivo bacteriano y purificación de proteínas:* El nuevo vector bacteriano pETM60-Leucolectina permitió la expresión de proteínas recombinantes marcadas con His fusionadas al extremo C de NusA a través de una secuencia de reconocimiento de proteasa TEV. Se purificaron eficazmente por afinidad en metal y se recuperaron en forma soluble después de retirar el compañero de fusión.

#### *Materiales y métodos:*

Cebadores de PCR superior e inferior:

NWA15dPET(N°23): NcoI  
5'-GCA.CCA:TGG.CCA.TGG.GCT.GGG.ACT.GTC.AGG.AGG.TAG.TA  
CH.A15dPET(N°13): ACC65I  
5'-CCG.AAG.CTT.GGT.ACC.ATG.TGT.GCA.CAC.CAT.GGT.GAC

#### *Mezcla de PCR:*

dNTP	4 µl
Cebador nº23/nº13	0,5 µl (10 µM)
ADNc	2 µl
Tampón ADN polimerasa	10 µl
ADN polimerasa	0,5 µl
dH <sub>2</sub> O	hasta 50 µl

#### *Reacción de PCR*

98 °C	10 s
55 °C	15 s
72 °C	1 min.
30 ciclos	
70 °C	10 min.

Tanto el producto de la PCR como el plásmido pETM-60 se digirieron con las enzimas de restricción NcoI y ACC65I; y se purificaron en gel de agarosa al 1% para el ligamiento.

#### *Ligamiento de inserto de ADN en el vector por ADN ligasa de T4:*

Se usó ligasa de T4 para unir el fosfato 5' y los grupos hidroxilo 3' de moléculas de ADN bicatenarias usando la siguiente mezcla:

200 ng de ADN de vector  
500 ng de ADN de inserto  
Tampón de ligasa 10X

1 µl de ligasa de T4

Volumen hasta 10 µl

5 La incubación se realizó a temperatura ambiente durante una noche y la inactivación térmica de la ligasa se consiguió poniendo el tubo en un baño de agua a 65 °C durante 10 min.

10 La mezcla de ligamiento se introdujo por electroporación en células competentes DH5α y las colonias se usaron para inocular 5 ml de LB.

Para el análisis se seleccionaron aleatoriamente 10 clones. El análisis se realizó por digestión de pETM60-Leucolectina con NcoI/ACC65I.

15 La secuencia del inserto de leucolectinas se verificó con secuenciación de pETM60-Leucolectina con los cebadores nº13 y nº23.

#### *Traducción in vitro de leucolectina*

20 BL21 (DE3) contiene un plásmido pLysS (pLysS con resistencia a cloranfenicol) que codifica la lisozima de T7 para reducir los niveles de fondo de la polimerasa de T7 antes de la inducción. BL21 (DE3) pLysS proporciona un control preciso de la expresión de proteínas tóxicas y es una cepa para la expresión de proteínas de alto nivel y de fácil inducción.

25 El marcador de fusión usado en la expresión y purificación de la proteína recombinante fue NusA que tiene un tamaño de 495 aa (54,8 kDa) y está localizado en el extremo N y se usa para mejorar la solución de proteínas que se sobreexpresan.

30 NusA-Leucolectina se purificó eficazmente por afinidad en metal y se recuperó en forma soluble después de retirar el compañero de fusión.

#### *Materiales para la expresión y purificación de NusA-Leucolectina:*

Placa de kanamicina

35 Medio 2YTG

IPTG 0,4 mM (concentración final 0,4 mM).

40 Columna HisTrap HP (5 ml)

La inducción se realizó durante 3 horas a 26 °C.

#### *Protocolo de purificación:*

45 Actividad enzimática intrínseca del sobrenadante de NusA-Leucolectina antes de la purificación: 0,0060/min

Equilibrio de la columna con 25 ml de tampón B (PBS 1x, imidazol 30 mM, NaCl 0,4 M)

50 Aplicación de la muestra:

Se filtraron 28 ml del sobrenadante y se desgasificaron, y 25 ml se aplicaron a la columna.

55 La columna se lavó con 75 ml de tampón B. La elución se realizó con 25 ml de tampón C (PBS 1x, imidazol 0,5 M, NaCl 0,4 M), con una elución de una etapa. Se recogieron muestras de 1 ml.

#### *Resultados:*

60 Más del 98 % de la actividad proteasa de las proteínas recombinantes brutas se encontró en el eluido y el lavado. Sólo se recuperó un 1,6 % de la actividad proteasa en fracciones de columna, incluyendo las que tenían (F6 y F7) la leucolectina recombinante, identificada por análisis de SDS-PAGE (véase la Figura 8). Estas fracciones contenían un 0,3 % (niveles de fondo de actividad proteasa aplicados a la columna).

#### *Conclusión:*

65 La leucolectina recombinante carece de actividad serín proteasa y, por lo tanto, es una verdadera lectina.

Por lo tanto, las pequeñas cantidades de actividad serán proteasa encontradas en algunas preparaciones de leucolectina procedentes de fuentes naturales (fluido de eclosión) representan pequeñas cantidades de zonasa.

- 5 La leucolectina y la zonasa coexisten en el fluido de eclosión y tienden a copurificarse durante la mayor parte de los procedimientos cromatográficos, pero pueden separarse por ultrafiltración final usando filtros Biomax con un tamaño de exclusión de 100 kDa.

#### *Análisis de proteína recombinante*

- 10 La LL de salmón recombinante producida como se ha indicado anteriormente y purificada por SDS PAGE se analizó por técnicas de MS/MS. El espectro de huellas peptídicas de LL se analizó en Mascot tanto en la proteína NCBI como en las bases de datos EST para *Actinopterygii* sin encontrar ningún acierto significativo (Riste, no publicado). Estos espectros se comparaban bien con el espectro de LL obtenido a partir de restos de LL embrionaria y leucocítica eluidos del análisis de gel 2D de dichas LL, lo que confirma adicionalmente la existencia de isoformas LL observadas por métodos de gradiente de pH.

#### *Secuencia genómica*

- 20 En salmón, se ha establecido la secuencia genómica de genes para esta entidad tratando bibliotecas de Bac con una sonda de 670 NT (también se describe un protocolo completo con detalle en el Ejemplo 10). De esta manera, se disponía de una secuencia de ADNc de longitud completa para la leucolectina por el uso de cebadores degenerados diseñados basándose en la información procedente de las secuencias de AA de proteínas parciales. Las amplificaciones por PCR del ADNc de embriones de 370 dd (grados día) (es decir, embriones cerrados pero que aún no habían eclosionado) generaron el amplicón de 630 pb, que se purificó en gel y se secuenció.

- 25 Se usaron sondas específicas de gen de lectina para explorar una biblioteca BAC (cobertura 18x; tamaño de inserto medio 188 kb), que se proporcionó por el consorcio Salmon Genome Project (SGP). Las sondas (tamaño 630 pb) se prepararon incorporando digoxigenina-11-dUTP (Roche) usando el método de PCR. La exploración de la biblioteca Bac dio como resultado la identificación de 7 clones positivos caracterizados como A-12/144, L16/143, P6/76, B-5/70, O-20/85, L-7/257, F-7/176. En los paneles A y B se ejemplifican las respuestas positivas.

- 30 Las respuestas específicas de los clones seleccionados se confirmaron por amplificación por PCR usando los mismos cebadores que para la sonda marcada con ADN usada para explorar la biblioteca BAC. El tamaño del inserto dentro de la secuencia del vector se estimó por electroforesis de campo pulsado. Los clones se secuenciaron comercialmente (MWG-Biotech, Alemania) por el método de secuenciación aleatoria.

- 35 Este trabajo estableció varios genes muy relacionados para las leucolectinas en el genoma del salmón (tetraploide) (Figs. 9A y B). Una variación en las secuencias segmentarias parecería sugerir la existencia de al menos dos genes similares. Estructuralmente, las leucolectinas aparecen como genes normales con 5 exones y 4 intrones, en un gen de aproximadamente 2.300 pb desde el inicio de la transcripción hasta el sitio de poliadenilación, un tamaño bastante convencional.

- 40 Sin embargo, los datos de los presentes solicitantes permitieron determinar que ningún gen con una secuencia similar se había presentado en ningún organismo hasta la fecha. Usando sondas para exones o intrones individuales, se pudo establecer únicamente la rara similitud de un exón a partes de otras dos proteínas presentadas, que son totalmente distintas de la entidad molecular que se ha descubierto.

- 45 Las búsquedas en las bases de datos indican que no puede encontrarse ningún gen que corresponda a la proteína LL, el transcrito de ARNm de LL, o al gen genómico para la leucolectina en cromosomas humanos. Aparentemente, el gen de la leucolectina aún no se ha encontrado en ningún sitio o su secuencia aún no se ha depositado, o la secuencia del gen de LL aún no se ha detectado en seres humanos. En búsquedas adicionales se usaron elementos estructurales pequeños individuales indicados en la Tabla 2 mostrada a continuación, para localizar dichos elementos en las bases de datos.

55

TABLA 2

Exón/intrón de LL	Longitud (NT)	Aciertos
Desde inicio de T al exón I	51	Ninguno
Exón I	55	Crom.11,7,1,18,17,10
Exón I + intrón I	55+89	Crom.11,14,18, 13
Exón II	234	Región de protocadherina, enzima tiamín trifosf.
Exón II + intrón II	234 + 163	Ninguno

Exón/intrón de LL	Longitud (NT)	Aciertos
Exón III	218	Proteína de músculo liso, Receptor de células T, etc.
Exón III + intrón III	218 + 133	Región V de cadena beta de Receptor de células T
Exón IV	116	Ninguno; Crom. 20 & 21
Exón IV + intrón IV	116 + 682	Cromosoma 8 (x2)
Exón V	142	Ninguno; ¿alguno?
Exón V + intrón V	142 + 404	Cromosomas 5 y 4
Inicio de T hasta extremo poliA	2283	Ninguno; ¿alguno?
ADNc	765	Ninguno
Proteína LL	255 AA	Ninguno
Intrón I	89	Crom. 3,4,18,14
Intrón II	163	Crom. 14, X, 4
Intrón III	133	Crom. 21,2,7,4,15,13,18,12,8,1,
Intrón IV	682	Crom. 18 (x2)
Intrón V	404	Ninguno; ¿alguno?

Los resultados indican que algunas proteínas conocidas mostraban algún grado de similitud con partes de la leucolectina – se proporcionan los números de cromosomas relevantes. Las secuencias no localizadas en los cromosomas se indican como ¿"alguno"?

5 El acierto primario fue la Transgelina humana, que es una proteína 22-alfa del músculo liso (SM22-alfa) (WS3-10). La transgelina es una proteína de unión a actina y es uno de los primeros marcadores del músculo liso diferenciado. La función de esta proteína aún no se ha determinado. <http://harvester.embl.de/harvester/P378/P37802.htm>

10 Sin embargo, el alineamiento de secuencias de proteína de transgelina y leucolectina mostró secuencias consenso mínimas. Por otra parte, se encontró alguna pequeña similitud en partes de la estructura de LL con respecto a un receptor de células T y con una proteína desconocida. Es importante destacar que este nuevo gen se expresa específicamente en leucocitos humanos, a pesar del hecho de que este gen no puede encontrarse "aún" en secuencias publicadas del genoma humano. (También se ve alguna expresión en preparaciones purificadas de trombocitos humanos). Sólo pueden detectarse segmentos muy cortos de este gen en el genoma humano, pero estos segmentos están ampliamente extendidos en una multitud de cromosomas humanos (véase la Figura 10).

20 El patrón de distribución observado en los cromosomas es desconcertante. Dichas secuencias cortas pueden no estar relacionadas con el gen de LL, si el propio gen de LL humano hasta ahora no se ha secuenciado localizado en un cromosoma específico. La resolución de este problema aún no está cerca.

#### Análisis

25 Se han identificado nuevas lectinas LL en leucocitos humanos que son casi idénticos a la LL de pez y pollo. El gen de LL en salmón es de la estructura convencional pero aún no se ha encontrado en seres humanos.

Ejemplo 2: Aplicaciones médicas de leucolectina.

30 Materiales y métodos.

Los siguientes estudios se realizaron usando la proteína leucolectina de salmón (denominada LL) preparada como se describe en el Ejemplo 1. La concentración de LL fue aproximadamente 1-10 microgramos/ml. La proteasa zonasa estaba presente con LL en una relación de 1:100 en una emulsión de agua y aceite de coco (en la que está presente más de un 30 % de aceite de coco). La leucolectina es estable en presencia de zonasa.

35 Resultados.

A. Piel agrietada por frío.

40 Muchas personas experimentan un agrietamiento estacional de la piel, particularmente en las manos.

La aplicación de leucolectina en la piel durante el inicio de la estación fría pospone espectacularmente, o en muchos casos impide, la aparición de grietas por frío. Este fenómeno se ha registrado en una docena de casos. Las observaciones sobre los efectos en grietas por frío establecidas de LL son menores, pero se han revelado algunos

45 cierres de heridas establecidas por frío.

B. Piel dañada por quemaduras solares.

Una exposición solar excesiva en la piel desnuda puede producir quemaduras solares, y dar como resultado sensación de calor, enrojecimiento, picor y descamación eventual de la piel expuesta.

- 5 La aplicación de leucolectina tuvo como resultado la desaparición del enrojecimiento en minutos, el cese del picor y la desaparición eventual de la descamación de la piel. Estas observaciones sorprendentes se han observado repetidamente en muchos individuos.

C. Piel dañada por el calor.

- 10 Después de una lesión directa en la piel por el calor, la aplicación de leucolectina de salmón parecía prevenir las consecuencias normales de dichas lesiones en una medida considerable, y detener la lesión si se aplicaba algún tiempo después de la lesión original, y producir una aceleración de la recuperación completa.

15 D. Acné

Los casos subclínicos de acné han respondido en muchos jóvenes con un cese inmediato del picor y del enrojecimiento. Se ha observado una mejora del acné *per se* en un pequeño estudio piloto en aproximadamente la mitad de los casos tras la aplicación de LL en una emulsión de aceite de coco.

20 E. Psoriasis tópica

- Grandes parches de piel psoriásica han respondido a la aplicación de leucolectina de salmón con, en primer lugar, una recesión casi inmediata del enrojecimiento, seguida de una reducción drástica de la sensación de picor. El exceso de congregación de piel córnea disminuyó durante las primeras semanas, pero se volvió a producir tras el cese de la aplicación de LL.

F. Heridas de piel abiertas.

- 30 Se realizaron observaciones en voluntarios humanos. Se estudiaron casos de heridas de piel abiertas crónicas y supurantes que se habían producido por compresión de heridas en el talón después de una cirugía abierta de cadera y que se habían tratado de forma no satisfactoria por el personal sanitario durante meses y hasta un año. Las llagas observadas tenían de medio a dos centímetros de diámetro.

- 35 Tras la aplicación de preparaciones de LL esterilizadas por filtración a las heridas abiertas crónicamente en las extremidades inferiores (piernas), se observó que el líquido detenía la supuración de las heridas después de 2-3 días seguido por una disminución rápida del tamaño del área de la herida, de forma que después de 2-3 semanas, las heridas estaban desapareciendo y se estaban reemplazando por piel normal.

- 40 Dada la naturaleza de los pacientes con heridas, primero aparecían en el extremo más distal de las extremidades, y se observó que las primeras heridas en desaparecer de la forma anterior eran las heridas proximales. Esto se interpretó como las heridas establecidas más recientemente. Las últimas heridas en desaparecer eran las heridas distales próximas al talón, que son las heridas establecidas durante más tiempo. Finalmente todas las heridas desaparecieron.

45 G. ESTUDIOS IN VITRO

- Se obtuvieron cultivos de epitelio de piel humana diferenciada en SkinEthics (Niza, Francia) el día 16 después de la siembra en sustratos de crecimiento de plástico con microporos que permitían el acceso de los nutrientes al tejido epitelial desde abajo. Dichos cultivos presentan una morfología normal de la piel después de la diferenciación durante el periodo de cultivo a 37 °C. Estos cultivos se mantuvieron durante dos días *in vitro* de forma que el estrato córneo superior se expuso al aire y estrato basal al sustrato de crecimiento. Los cultivos paralelos se movieron a una atmósfera húmeda a 30 °C y se presentaron con un medio de solución salina tamponada con fosfato que contenía Ca y Mg durante 6 horas con o sin la presencia de 10 µg/ml de leucolectina de salmón. Los cultivos se fijaron en formalina y se incluyeron en parafina de acuerdo con procedimientos convencionales, y se tiñeron con hematoxilina/eosina. La rápida inducción de células basales grandes se observó en la epidermis, con signos de proliferación celular (Fig. 11). Los presentes solicitantes interpretan dichos resultados como causalmente relacionados con los efectos curativos de la leucolectina en la piel, en la que una combinación de proliferación celular y diferenciación celular es suficiente para cerrar la herida y reestablecer la integridad epitelial de la piel.

60 Análisis:

- Los presentes solicitantes no han encontrado expresión de productos del gen de LL en cultivos celulares de queratinocitos. Además, las células dendríticas son las únicas células en piel epitelial relacionadas con los leucocitos sanguíneos. El acceso normal de los leucocitos a la piel epitelial se ve obstaculizado después de algunas lesiones mecánicas o biológicas debido a una peor circulación, que es lo que ocurre en las heridas de compresión que se



producen a menudo en la piel normal. La administración de leucolectina compensa esta deficiencia y desencadena mecanismos normales de curación en la piel.

#### Ejemplo 3: expresión de proteína leucolectina en leucocitos

Se obtuvo sangre de salmón en Industrib, Univ. de Bergen. Se recogió sangre entera de pez en tubos heparinizados (Riste, no publicado), y se procesó de acuerdo con Miller *et al.* (1988, Nucleic Acids Research, 16(3) pág. 1215). Se obtuvieron muestras de sangre humana del banco de sangre del hospital de la universidad de Haukeland, Noruega.

Se recogió sangre humana entera en tubos de 5 ml con tampón citrato, y se procesó de acuerdo con Miller *et al.* (1988, anteriormente citado) para preparar leucocitos. Esta preparación se purificó adicionalmente por enriquecimiento Lymphoprep™ (Axis-Shield PoC, Noruega). Se usó Lymphoprep™ de acuerdo con su protocolo para preparar una fracción de leucocitos humanos. Esta fracción se usó para el análisis de inmunotransferencia, como se describe en Miftari y Walther (no publicado). El análisis por PAGE 2D siguió los procedimientos descritos por MacGillivray y Rickwood (1974, European Journal of Biochemistry, 41, pág. 181-190), con ~1 µg de proteína. Las proteínas LL de estas dos especies se visualizaron por PAGE 2D usando inmunotransferencias para la detección específica de proteínas LL y sus isoformas. Los tratamientos con anticuerpo primario policlonal contra LL de salmón seguido de anticuerpo de cabra anti-conejo permitieron la visualización de las proteínas LL por el sistema de detección potenciado por ECL. La especificidad de la inmunorreacción se ensayó en múltiples geles aplicando 0,5-5 µg de proteína, con resultados similares.

Los leucocitos de salmón presentan 2 isoformas de LL de PM ~26 kDa, con un pI similar a las formas más básicas de las 8 formas isoelectricas de LL encontradas en PVF de salmón. Es difícil estimar los PM exactos mediante extrapolación a partir de los patrones de PM en la segunda dimensión de PAGE 2D. Los leucocitos humanos contienen un antígeno de LL principal de peso molecular ~ 30 kDa, que a partir de otras observaciones está próximo a 27 kDa. El pI ácido corresponde a la media del pI del salmón de las isoformas de LL encontradas en PVF (Riste, anteriormente citado) (véase la Figura 12).

#### Ejemplo 4: transcritos de leucolectina de salmón identificados en transferencias de Northern

Se obtuvieron huevos de salmón a partir de un criadero de salmones en Bolaks (Fusa, Noruega), mantenido durante múltiples generaciones desde 1975 por selección fenotípica, y ahora parte de las reservas de salmón propagadas por Salmo Breed A/S, Noruega.

#### Aislamiento de ARN total y de ARN poliadenilado

El ARN total se aisló a partir de embriones de salmón en las fases de pre-eclosión posteriores (aproximadamente 370 dd) usando reactivo Trizol (Life Technologies) de acuerdo con las instrucciones del fabricante. La calidad y cantidad del ARNm total en términos de integridad y pureza se evaluó en un gel de agarosa y formaldehído teñido con bromuro de etidio. El ARN total se cuantificó espectrofotométricamente. La fracción poliadenilada de ARN (ARNm) se aisló a partir de 5-10 µg de ARN total usando un kit de purificación de ARNm con perlas Dynal (Invitrogen).

#### Construcción de sondas

Se generó una sonda de ARNc antisentido marcada con digoxigenina (DIG) por transcripción de moldes de ADN plasmídico digerido con SP1 con ARN polimerasa de T3. La sonda de ARNc marcada con DIG con sentido se generó por transcripción de molde de ADN plasmídico digerido con XhoI con ARN polimerasa de T7 (con sentido) usando el kit de marcaje de ARN DIG (SP6/T7) (Roche Diagnostics) de acuerdo con los protocolos del fabricante. Las sondas de ARNc con sentido y antisentido de LL marcadas con digoxigenina (DIG) generadas se extendieron ~ 650 pb.

#### Análisis de transferencia de Northern

Se sometieron a electroforesis alícuotas de 0,5 µg de ARNm puro y 1 µg de ARN total y se transfirieron sobre una membrana de nitrocelulosa cargada positivamente (Amersham). Las membranas se prehibridaron en tampón DIG Easy Hyb (Roche) a 55 °C durante 2 horas, y después se hibridaron a 55 °C durante 16 horas, con 1 µg/ml de ribosonda de LL. El lavado se realizó 2 veces con condiciones de baja rigurosidad en SSC 2x con SDS al 0,1 % a TA durante 15 minutos cada uno, y dos veces en condiciones de alta rigurosidad en 0,1x con SDS al 0,1 % a 65 °C durante 20 minutos cada uno. Las membranas se incubaron en una solución bloqueante que contenía tampón de ácido maleico (pH 7,5) y reactivo bloqueante al 1 % (Roche) y posteriormente en la solución bloqueante con anticuerpo anti-DIG conjugado con fosfatasa alcalina (AP) (Roche) diluido a 1:10.000 a TA durante 1 hora. Después de lavar con tampón de ácido maleico 100 mM (pH 7,5) que contenía NaCl 150 mM y Tween 20 al 0,5 %, la membrana se equilibró durante 2 minutos en tampón de detección (Tris HCl 0,1 M, NaCl 0,1 M, pH 9,5 a TA). Las dianas de la sonda híbrida se visualizaron por ensayo quimioluminiscente usando CSPD (Roche) como sustrato, y

exposición de las manchas de transferencia a una película de rayos X (Kodak) durante 5-30 minutos. Se usó el marcador de peso molecular de ARN marcado con DIG (Roche).

Se usó una ribosonda específica para LL marcada con DIG antisentido para estimar el tamaño de los transcritos de LL maduros. El análisis por transferencia de Northern del ARN total (Figura 13A) reveló un transcrito importante con un tamaño ~ 1250 pb y también un transcrito de banda más débil (~3200 pb). La banda dominante corresponde a un transcrito de la proteína madura (sin intrones) y el transcrito de LL traducible. El análisis por transferencia de Northern del ARNm purificado usando la misma ribosonda reveló una sola banda de ~ 1250 pb (que indica que el ADNc amplificado estaba próximo a una secuencia codificante de longitud completa (Figura 13B). Cuando las transferencias del ARN total y las muestras de ARNm se incubaron con una ribosonda específica de LL con sentido, se observó la ausencia de una señal de hibridación que indicaba que la ribosonda que se usó era suficientemente específica (Figura 13C). El transcrito minoritario (~3200 pb) es mayor que el transcrito de LL de longitud completa para salmón. Dadas las condiciones específicas para la hibridación, esta entidad probablemente refleja una entidad que contiene información de la secuencia de LL. Las señales positivas proceden solo de las sondas antisentido y no de las sondas con sentido, destacando la especificidad de los resultados. El tamaño del transcrito de longitud completa se acomoda a todos los exones de LL descritos anteriormente.

Ejemplo 5: alineamiento de LL con proteínas de hélice enrollada  $\beta$  relacionadas

Se realizaron análisis bioinformáticos en secuencias de LL de salmón en comparación con la información depositada en los bancos de secuencias públicos.

El análisis Clustal W (Figura 14) indicó que la secuencia de aminoácidos deducida de la leucolectina de salmón no compartía una similitud de secuencia significativa con ninguna proteína presentada en el banco de datos. Sin embargo, BLAST revela similitudes con proteínas relacionadas de forma más distante, incluyendo FEL, una lectina derivada de huevo de pez de *Cyprinus carpio* (Gene bank nº de referencia P68512) con una similitud total de únicamente un 48 % (valores E de 2.00 E-50). Además, una proteína hipotética del pez cebra (HPZ; nº de referencia LOC678590) presentó una similitud total del 45 % (valores E de 2.00 E-48). Las comparaciones de secuencias también demostraron que la leucolectina tenía una identidad del 38 %, 28 % y 26 % con TPL-1 (previamente lectina L6) del cangrejo de herradura (nº de referencia P82151), la tectonina I de *Physarum polycephalum* (nº de referencia 061063) y la Tectonina II (061064). (Por SwissProt, estos números tenían unas identidades de 49, 35, 32 y 29 %, respectivamente, y el valor E para FEL 1 e-51 y para L6, 1 e-40). Los datos de los presentes solicitantes sugieren una mayor similitud de LL con los vertebrados que con los homólogos invertebrados. No se incluye una identidad del 30 % de LL con la proteína que contiene repetición hélice enrollada  $\beta$  de tectonina humana (K1AA0329\_human).

Los múltiples alineamientos revelaron que de los cuatro enlaces disulfuro confirmados bioquímicamente en FEL (Galliano *et al.*, 2003, Biochem. J. 376, pág. 433-440), tres parecen haberse conservado en la leucolectina. El primer puente conecta el extremo N-terminal con el extremo C-terminal de las dos proteínas. El puente conecta las cisteínas en las posiciones 3 y 234 en LL (o 238 en FEL). El segundo y tercer puentes conectan la Cys 102 con la Cys 155 en LL (nº 157 en FEL), y la Cys 132 y Cys 137 tanto en LL como en FEL. El cuarto enlace disulfuro encontrado en FEL de carpa, que implica Cys 208 (212) y Cys 226 (230) no se encuentra en la leucolectina.

Ejemplo 6: Modelo 3D de LL (leucolectina)

La proteína leucolectina comparte un alto grado de similitud en términos de arquitectura de dominios TECPR con L6 o taquilectina 1 (*Tachypleus tridentatus*), una lectina que se une a lipopolisacáridos bacterianos procedente de hemocitos del cangrejo de herradura y otras proteínas relacionadas con la taquilectina de invertebrados. Sin embargo, se observaron diferencias en términos del número de dominios, junto con un ligero desplazamiento de los dominios hélice enrollada  $\beta$  hacia el extremo C de la proteína. Resumiendo la información de las búsquedas de similitud y los buscadores SMART, se deduce que la estructura 3-D global de la leucolectina presenta una estructura de proteína hélice enrollada  $\beta$  de 5 filos. Basándose en esto se generó el modelo 3D (Biesel *et al.*, 1999, EMBO J., 18, pág. 2313-2322) para dar un modelo de trabajo de qué aspecto puede tener la proteína leucolectina. Esto es solo una aproximación, ya que los números de restos de una proteína se establecen igual que los restos en la otra. Las búsquedas en la base de datos SMART indican una característica estructural prominente de la proteína leucolectina prácticamente traducida, que consiste en cinco repeticiones en tándem, conteniendo cada una 35-36 AA con identidades de secuencia internas de 32-61 % (Figura 15, panel derecho). Los tramos de repeticiones internas de la proteína leucolectina deducida mostraron similitudes entre sí mismos, que eran principalmente de 13 AA de longitud. Dos secuencias consenso cortas, *XWXXLPGXLKXXVGPX* y *GVNKNDDXXYXLVG*, están muy conservadas en cada repetición. Aparte de un tramo de 20 restos en la región N terminal de la leucolectina, que no se encuentra en la FEL, las dos proteínas muestran un alto grado de similitud entre el número y la arquitectura global de dominios de los dominios TECPR, mientras que comparten solo un 48 % de identidad de secuencia.

Ejemplo 7: purificación de zonasa a partir de fluido de eclosión de salmón

El fluido de eclosión de salmón del ejemplo 1 puede purificarse adicionalmente para producir una preparación de zonasa pura. Una vuelta de filtración en gel fue suficiente para separar la zonasa de los componentes moleculares

mayores en el filtrado con una purificación de 12 veces obteniéndose un rendimiento mejor del 50 %. La matriz utilizada puede variar, pero la elección habitual fue Sephacryl SR-200®. El tampón era Tris-HCl pH 8,0 o pH 8,5 (0,05 M) o Tris-acetato (0,025 M, el mismo pH). La zonasa obtenida después de los procedimientos de filtración en gel representaba los restos de zonasa predominantes en el fluido de eclosión.

La zonasa se purificó hasta un producto proteico homogéneo por cromatografía de afinidad en columnas de benzamidina Sepharose 6B® disponibles en el mercado. Las condiciones específicas utilizadas (con columnas de 25 o 125 ml de volumen) fueron de nuevo tampón Tris-HCl 0,05 M (pH 8 o 8,5), que para retirar el material unido de forma no específica en las columnas, se ajustó a NaCl 1 M. La zonasa no se retiró en esta etapa, ya que la proteína permanecía unida fuertemente a la columna. La elución de la zonasa de la columna se consiguió usando un gradiente de dioxano del 10-33 % en NaCl 1 M en el mismo tampón Tris-HCl. Después de la purificación por afinidad, la preparación de zonasa presentó una banda proteica en el análisis de SDS-PAGE, con un peso molecular de aproximadamente 28 kDa. Aunque este producto de zonasa no tenía la pureza de calidad de secuencia estaba muy purificado.

La zonasa purificada por filtración en gel más purificada por afinidad se purificó adicionalmente para obtener una pureza de calidad de secuencia por un procedimiento cromatográfico final. Este procedimiento empleó una columna PBE94®, con un tampón de Tris-acetato (10 mM, pH 9,0), donde la elución posterior se realizó con un gradiente de sal (hasta sal NaCl 1 M) en este tampón. Esta etapa por sí misma aumentó la actividad catalítica de la zonasa 7,6 veces más, para una purificación global de 714 veces, y con un rendimiento del 28 % a partir del material de partida. Esta etapa de purificación dejó la identidad de la proteína de la zonasa intacta como un resto de 28 kDa.

#### Ejemplo 8: producción y purificación de anticuerpos de LL

Se preparó antisuero anti-LL de salmón policlonal de conejo anti-conejo como se indica a continuación: se usó la proteína leucolectina purificada con calidad de secuencia del ejemplo 1 para inducir el anticuerpo policlonal en conejo chinchilla de 4 kg como se describe por Harlow y Lane (1988, Laboratory Manual, Cold Spring Harbor, NY). Se pusieron tres inyecciones de 80, 40 y 25 mg de leucolectina de calidad de secuencia en múltiples sitios subcutáneos. La primera se emulsionó con adyuvante completo de Freund, las dos últimas con adyuvante incompleto después de 3 y 6 semanas. Ocho después del refuerzo final, se recogió sangre y se preparó suero después de la centrifugación a 3.000 rpm (rotor Sorvall SS-34) durante 15 min.

Se almacenaron alícuotas del antisuero a -80 °C. Se purificaron por afinidad anticuerpos primarios usando columnas de proteína G HiTrap de 1 ml (Amersham Pharmacia Biotech) con un caudal de 1 ml/min. El antisuero completo se filtró a través de una unidad de filtro de 45 µm antes de aplicarse a columnas de proteína G de 1 ml. El equilibrado de la columna se realizó usando 5-10 volúmenes de columna de tampón de unión (fosfato Na 20 mM, pH 7,0) seguido de varios lavados con el mismo tampón hasta que no se detectó proteína por absorbancia UV a 280 nm. La fracción de IgG se eluyó con 5 ml de tampón de elución (glicina HCl 0,1 M, pH 2,7) y se dividió en 10 tubos que contenían 25 µl de tampón de neutralización (Tris HCl 1 M, pH 9,0). La concentración se determinó espectrofotométricamente. La pureza se estimó analizando las alícuotas de las fracciones de IgG purificadas en SDS-PAGE al 12 % con la posterior tinción con plata. La sensibilidad se determinó por análisis de transferencia de Western usando anticuerpo IgG de cabra anti-conejo secundario conjugado con peroxidasa de rábano picante.

Además, se preadsorbieron anticuerpos policlonales por incubación en polvo de acetona de pez cebra. El F(ab')<sub>2</sub> de cerdo anti-conejo marcado con isotiocianato de fluoresceína (FITC) procedía de Dako (F-0054). La IgG de cabra anti-conejo marcada con Alexa Fluor 647 se obtuvo de Molecular Probes, Eugene, US (Invitrogen, nº de catálogo A-21244) y las inmunoglobulinas policlonales de cerdo anti-conejo biotiniladas procedían de Dako (nº de catálogo E0353). Las concentraciones óptimas de los anticuerpos primarios y secundarios se determinaron mediante experimentos de dilución previos. Los anticuerpos policlonales se usaron en el análisis de expresión de proteínas descrito en el ejemplo 1.

#### Ejemplo 9: Identificación y caracterización de LL a partir del pez cebra

Se aisló ARN total a partir de embrión de pez cebra a 4, 6, 12, 24, 48 hdf y 5 dde usando el procedimiento de Trizol (Life Technologies) de acuerdo con las instrucciones del fabricante. La integridad y pureza del ARN total se evaluó en un gel de agarosa y formaldehído teñido con bromuro de etidio. Solo se usaron muestras con una relación de ARNr 28S a 18S de 2-2,4:1 y ninguna contaminación de ADN detectable. El ARN se cuantificó espectrofotométricamente. El ARNm poliadenilado se aisló a partir de 5-10 µg de ARN usando kits Dynal (Invitrogen).

La transcripción inversa se realizó usando 200 U/µl de TernoScript (Invitrogen), se calentó ARNm (0,5 µg) a 72 °C durante 5 minutos, se enfrió, se sedimentó y se añadió a una mezcla de reacción de 20 µl que contenía 100 ng/µl de oligo dT, DTT 10 µM y 0,8 U/µl de inhibidor de RNasa en tampón de síntesis de ADNc 2,0 mM (Invitrogen). La incubación se realizó a 50 °C durante 1 hora. Los productos se purificaron por extracción con fenol/cloroformo/alcohol isoamílico, se pasaron a través de una columna MicroSpin S-200HR y se precipitaron con etanol. Los cebadores de LL específicos del gen se diseñaron a partir de secuencias de LL de salmón (tabla 3). La mezcla de

- 5 PCR (20  $\mu$ l) contenía 2  $\mu$ l de molde de ADNc, una concentración de 0,4  $\mu$ M de cebadores (LLF1, LLF2, LLR1, LLR2, y LLR3), tampón PCR 1x, 0,5 U de ADN polimerasa Taq (Takara) y una concentración 0,2 mM de dNTP. La amplificación por PCR usó 32 ciclos de 94 °C durante 45 s/58 °C durante 45 s/72 °C durante 90 s y una extensión final a 72 °C durante 7 min. La mezcla de reacción se analizó mediante gel de agarosa-TBE al 1,8 % con 0,5  $\mu$ g/ml de bromuro de etidio en tampón TBE 1x (pH 8,3) con un marcador de ADN de una escala de 100 pb. (New England Biolabs).

Tabla 3: Cebadores usados para amplificación de LL del pez cebra.

Tabla 6. Cebadores usados para amplificación de EL de los púas de conejo.				
RT-PCR		Tamaño esperado (pb)		
Cebador	Secuencia	Posición (nt)		
LLF <sub>1</sub> ,	5'-ATGCAGATCGATGCAGGACTGGG-3'	94-116	F1/R1 = ~720	
LLF <sub>2</sub>	5'-TGGTTCCCTGAAGCATGTCACTGT-3'	186-209	F1/R2 = ~454 F1/R3 = ~340	
LLR <sub>1</sub> ,	5'-GAAAGAGAGATCAATGAGTTGCGCA-3'	840-817	F2/R1 = ~655	
LLR <sub>2</sub>	5'-CAAAGACACTACCATCAGTTGCCAC-3'	595-571	F2/R2 = ~410	
LLR <sub>3</sub>	5'-GTCCGCGAGCTGTAGTACTTCACAG-3'	457-434	F2/R3 = ~300	
5' RACE-PCR				
Cebador	Secuencia	Cebador	Secuencia	
Cebador 5'	GeneRacer	Cebador 3'	GeneRacer	5'-GCTGTCAACGATACGCTACGTAACG-3'
Cebador anidado 5'	GeneRacer	Cebador anidado 3'	GeneRacer	5'-CGCTACGTAACGGCATGACAGTG-3'
LLR <sub>2</sub>	5'-CGACACTGACATGGACTGAAGGAGTA-3'	LLF <sub>2</sub>	5'-TGGTTCCCTGAAGCATGTCAGTGT-3'	
LLR <sub>4</sub>	5'-CAAAGACACTACCATCAGTTGCCAC-3'	LLF <sub>4</sub>	5'-GTGAGGTGGCAACTGATGGTAGTGTCT-3'	

El fragmento de PCR se ligó en un sistema 1 de vector pGEM T-Easy (Promega, Madison, WI) mediante protocolos convencionales.

La electroporación usó 40 µl de *E. coli* DH5α, y 1 µl de fragmento de PCR ligado en el vector pGEM T-Easy. Se añadió medio SOC (1 ml) antes de incubar las bacterias a 37 °C durante 1 h. Las bacterias recombinantes (200 µl) se cultivaron en placas de agar IPTG convencionales LB/ampicilina, X-gal y se incubaron a 37 °C durante toda la noche para la exploración azul/blanca. Se seleccionaron 10 clones blancos y se cultivaron en un cultivo de una noche de 5 ml en medio LB con 100 µg/ml ampicilina. La purificación del plásmido usó el kit de Minipurificación de Plásmidos (Qiagen, Chatsworth, CA). Los plásmidos se amplificaron y secuenciaron usando protocolos convencionales (DNA Sequence Lab, Bergen).

Las PCR RACE 5' y 3' de LL se realizaron usando el kit central GeneRacer (Invitrogen nº de catálogo 45-0168) de acuerdo con las instrucciones. Se usaron 2 µg de ARN total para sintetizar la primera cadena de ADNc en RT usando Superscript III (Invitrogen). La primera vuelta de amplificación por PCR 5' y 3'-RACE se realizó aplicando el cebador 5' o 3' GeneRacer y el cebador específico de gen inverso o directo LLR2 y LLF3 (Tabla 3). La amplificación 5' o 3'-RACE PCR anidada empleó cebador GeneRacer 5' o 3'-RACE PCR y cebador específico de gen inverso o directo anidado LLR4 & LLF4. Para la 5' y 3'-RACE PCR, se usaron en la primera etapa 5 ciclos, 94 °C durante 30 s y 72 °C durante 1 min; a continuación cinco ciclos a 94 °C durante 30 s y 70 °C durante 1 min; tercera etapa 30 ciclos a 94 °C durante 30 s y 68 °C durante 45 s y 72 °C durante 1 min; etapa de extensión final 72 °C durante 7 min. PCR 5' o 3'-RACE anidada en tres etapas: primera etapa 5 ciclos a 94 °C durante 30 s, 72 °C durante 2 min; siguiente etapa 5 ciclos 94 °C, 30 s, 70 °C 2 min; 3ª etapa 25 ciclos 94 °C durante 30 s, 65 °C durante 30 s y 68 °C durante 1 min; etapa de extensión final 72 °C durante 10 min. Los moldes para la PCR anidada usaron el producto de PCR de la primera vuelta de RACE PCR diluido a 1:25 en agua destilada.

Los cebadores para PCR RACE se indican en la Tabla 3. Se purificaron productos de PCR individuales usando el kit de purificación de PCR (Promega, Madison, WI, US), y se ligaron en el vector topo PCR II (Invitrogen) para la transformación de *E. coli* competentes de alta eficacia JM109 (Invitrogen). Las bacterias recombinantes se purificaron por exploración azul/blanco en placas de agar Lucía Bertani. Los plásmidos que contenían las inserciones se purificaron usando el kit de purificación de Promega y se secuenciaron con cebadores de M13. La secuencia de ARNm de LL se analizó usando los programas de búsqueda Blastx y Blastp.

Combinando tres conjuntos de cebadores (véase la Tabla 3), la amplificación dio amplicones de tamaños previstos (Figura 16A) a partir de LL de salmón. El blanco acuoso sin de ARNm de pez cebra fue negativo. El producto de mayor tamaño (720 pb) se secuenció, demostrando que LL está presente en el pez cebra. La amplificación de productos de PCR de diferentes tamaños (Figura 16A) cebados por oligonucleótidos específicos que flanquean los exones de salmón en diferentes posiciones, documentó adicionalmente la alta homología entre LL de pez cebra y de salmón. Como control positivo para las reacciones de amplificación se usó β-actina.

La PCR 5'-RACE clarificó la secuencia de la región 5' no traducida (5'-UTR) de la LL del pez cebra. Se usó ARNm (Figura 16B) procedente de huevos desde las 4 horas después de la fertilización (hdf) hasta los 5 días después de la eclosión (dde). La primera vuelta de amplificación RACE PCR produjo productos muy débiles. Dichos productos amplificados de todas las fases se usaron para la segunda vuelta de RACE PCR usando una alta temperatura de hibridación (65 °C) para aumentar la especificidad de la reacción, usando diluciones 1:20 de la primera reacción PCR RACE. Los mismos dos productos (520 y 420 pb) se generaron a partir de todas las fases (Figura 16B), incluso desde 4 hdf. A partir de estas fases del desarrollo, se obtuvo la información de la secuencia de LL de más de 20 clones. Todos los clones contenían la región cadena arriba de 29 nt (Figura 16C) detectada en LL de salmón. El extremo 3' también se secuenció por análisis múltiple y reveló una estructura muy similar (casi idéntica) a la estructura encontrada en la leucolectina de salmón. Basándose en la secuencia de ADNc de longitud completa, la proteína LL del pez cebra deducida se depositó en el Gene Bank del NCBI (nº de referencia FJ 643620). La secuencia codificante consiste en 1.213 nt con una fase de lectura abierta de 765 nucleótidos. La secuencia de aminoácidos deducida para la LL de pez cebra sugiere una proteína de 255 AA, dado un codón de inicio de la traducción en la posición de nt +1 (o nt 30-32; (Figura 16C). El ADNc de LL también tiene un codón de iniciación potencial en nt +94 (Figura 16C). En 4 de 20 clones de LL de pez cebra, el transcrito contenía los 29 nt cadena arriba del sitio +1, pero carecía de los siguientes 93 nt (la secuencia de nt se lee correctamente hasta -1 seguido por el nt a +94: Fig. 16C). De esta manera los datos sugieren que pueden traducirse dos proteínas LL diferentes. (Figura 16C).

La fuerte homología entre las proteínas LL del pez cebra y del salmón fue evidente. La LL del pez cebra es una estructura híbrida en comparación con las dos LL de salmón. De los siete criterios de clasificación de LL establecidos, 4 criterios pusieron a la LL del pez cebra en la categoría de LL-1, mientras que 2 criterios la pusieron en la categoría LL-2 (Figura 16E). El séptimo criterio es único para la LL del pez cebra (Figura 16E, área recuadrada). Además, falta una Cys en el exón 3 y se reemplaza por Arg en la secuencia "nmnTPYClts", que en el pez cebra es "nmnDTPRlts". En consecuencia, la LL del pez cebra tiene 5 cisteínas donde el salmón tiene 6 cisteínas, todas previstas en los puentes disulfuro.

En relación con los diversos exones, el exón 1 es idéntico, mientras que el exón 2 del pez cebra tiene tres AA únicos no encontrados en salmón (Gly-Arg en lugar de Ser-Gln en las posiciones nº 57 y 58, y Val en lugar de Ile en la posición nº 78). El exón 3 tiene 4 restos alterados (¿o 2?). En los exones 4 y 5 no hay diferencias aparte de los 3 sitios de clasificación. De esta manera hay 5 (o 7) AA únicos en la LL del pez cebra (255 AA en total), ignorando los sitios de clasificación. La LL del pez cebra está muy conservada en comparación con la LL de salmón difiriendo sólo en unos pocos puntos de porcentaje.

La LL del pez cebra (Figura 16D, E) es muy similar a la LL de salmón. La LL del pez cebra carece de seis Cys vistas en la LL salmón, y por lo tanto puede contener solo dos fuentes disulfuro: uno que conecta los extremos N y C, más un puente interno en el exón 4. Comparando la LL de salmón y la LL grande del pez cebra (Figura 16E), las variaciones estructurales estereotípicas observadas en el salmón se refuerzan por la estructura de LL del pez cebra. Se observa una diferencia radial (resto nº 131), pero principalmente se ven sustituciones conservadas donde se producen variaciones en la LL de salmón. En PVF, la LL puede tener funciones inmunoprotectoras. Las variaciones restringidas en la estructura de la LL pueden estar relacionadas con dichas funciones. La leucolectina truncada (tLL) carece del péptido de secreción, lo que sugiere que en la tLL (Figura 16D) no se secreta, y la tLL puede ser una proteína de hélice enrollada  $\beta$  de 4 fillos. Las diferentes proteínas LL en teoría, pueden realizar diferentes funciones celulares. La función de tLL puede existir en células distintas de los leucocitos.

Ejemplo 10: Aislamiento y caracterización del gen de LL a partir de salmón

#### *Exploración de la biblioteca BAC.*

Para obtener la estructura genómica del gen de la leucolectina, se exploró una biblioteca de BAC genómica de salmón pública (cobertura 18x tamaño de inserto medio aproximadamente 188 kb) que se puso a disposición por el Proyecto Genoma del Salmón (Oslo, Noruega). La biblioteca de cromosomas artificiales bacterianos (BAC) altamente redundante construida a partir de una cepa de acuicultura noruega de salmón atlántico (*Salmo salar*) macho se exploró para buscar clones positivos para LL. La biblioteca consiste en un número total de 305.557 clones. El tamaño medio del inserto de la biblioteca es 188 kpb, que representa una cobertura del genoma de 18 veces. Filtros de alta densidad CHORI-214 Seg. 1 (serie de filtro 007193), consistiendo cada uno en 18.432 clones que punteados por duplicado, se han producido para la exploración por hibridación.

Para explorar la biblioteca, se preparó una sonda de ADNc específica de LL de 620 pb incorporando digoxigenina-11-dUTP (Roche) por el método de PCR usando cebadores directos LL/F 5'-TACGGACACAGGTCTGAATCCCCTACTACC-3' y el cebador inverso LL/R 5'-ACAGAGAAGAGGCTAATGTGTGCAC-3'. La sonda de ADNc marcada con DIG se incubó a 95 °C durante 10 min. Inmediatamente se puso en hielo durante 5 min antes de la dicción al tampón de hibridación (SSC 5x, formamida al 50 %, SDS al 0,02 % agente bloqueante al 2 % (Roche), agua tratada con DEPC).

La hibridación se realizó en tubos de hibridación a 55 °C durante una noche. La etapa de lavado después de la hibridación se realizó dos veces a baja rigurosidad en SSC 2x con SDS al 0,1 % a TA durante 15 min cada uno, dos veces en condiciones de alta rigurosidad en SSC 0,1 x con SDS al 0,1 % a 65 °C durante 20 min cada uno. Los filtros se incubaron en una solución de bloqueo que contenía tampón de ácido maleico (pH 7,5) y reactivo bloqueante al 1 % (Roche), y posteriormente en solución bloqueante con anticuerpo anti-Dig conjugado con fosfatasa alcalina (AP) (Roche) diluido 1:10.000 durante 1 h a TA. Después de lavar con tampón de ácido maleico 0,1 M (pH 7,5) que contenía NaCl 150 mM y Tween 20 al 0,5 %, la membrana se equilibró durante 2 minutos en tampón de detección (TrisHCl 0,1 M, NaCl 0,1 M, pH 9,5 a TA). Las manchas duplicadas positivas se visualizaron por ensayo quimioluminiscente usando CSPD (Roche) como sustrato. La identificación de las señales positivas se consiguió exponiendo los filtros a una película de rayos X (Kodak). El tiempo de exposición varió de 3 a 5 min. Dos clones positivos elegidos L-7/257 y A-12/144 se cultivaron en placas de agar preparadas y se incubaron a 37 °C durante una noche. Se seleccionaron tres clones, se cultivaron durante una noche en un cultivo de agitación de 5 ml en medio LB a 37 °C con 20 µg/ml de cloranfenicol. La purificación del ADN plasmídico se realizó al día siguiente usando un kit de purificación de ADN ultrapure (Qiagen). El ADN de BAC DNA se digirió con *Not* I y se analizó por electroforesis de campo pulsado (PFGE) (Osoegawa et al., 1998, Genomics, 52, pág. 1-8). Como marcadores del tamaño del ADN se usaron marcador PFG de bajo intervalo (New England Biolabs) y el fragmento  $\lambda$ -Hind III (Takara).

#### *Verificación por PCR de los clones positivos*

Las reacciones de PCR se realizaron usando ADN de BAC purificado como molde y cebadores específicos de gen. Las reacciones de PCR se incubaron a 95 °C durante 3 min, seguido de 35 ciclos a 95 °C durante 30 s, hibridación a 54 °C durante 30 s, y extensión a 72 °C durante 30 s con una incubación final a 72 °C durante 10 min. Después de los ciclos térmicos, se analizaron 2 µl de cada reacción de PCR en un gel de agarosa al 1,5 %.

#### *Secuenciación aleatoria de clones BAC de leucolectina*

Se secuenciaron dos clones por MRW, Berlín, Alemania. Las inserciones del vector se secuenciaron hasta una redundancia estimada de 2,4 veces. Las lecturas se limpiaron de vector y elementos del hospedador. A partir de aproximadamente 180 lecturas, se estableció un total de respectivamente de 33 y 32 a partir de los dos clones BAC usando bioinformática. Los cóntigos se examinaron tanto con respecto a las secuencias de ARNm como a las secuencias de aminoácidos. El análisis identificó exones e intrones, tanto búsquedas de programa como por inspecciones manuales. El gen de LL entero se identificó a partir de cóntigos solapantes.

#### Estructura genómica

Los resultados de la exploración anterior identificaron 7 clones específicos de nominados A-12/144, L-16/143, P-6/76, B-5/70, O-20/85, L-7/257 y F-7/176 de acuerdo con la nomenclatura de la biblioteca BAC. El tamaño medio del inserto de los clones de BAC genómicos se estimó a partir del ADN purificado de todos los clones de BAC seleccionados, que posteriormente se digirió con la enzima de restricción *Not* I. Los productos de digestión de ADN que contenían vector se separaron por electroforesis en gel de campo pulsado y se visualizaron por tinción con bromuro de etidio. Todos los clones de BAC produjeron solo dos bandas, que mostraban que el tamaño del vector (pTARBAC2.1) era ~13 kpb (13,397 pb; Osoegawa *et al.*, 1998, anteriormente citado). Esto sugiere que el tamaño del inserto de todos los clones era ~25 kpb. Las secuencias de 6 de los 7 clones BAC seleccionados se verificaron una vez más por amplificación usando como molde ADN purificado de un clon BAC, junto con cebadores específicos del gen de LL. La amplificación por PCR dio un producto de ~600 pb de todos esos clones. Su identidad se confirmó adicionalmente por secuenciación directa.

#### Secuenciación aleatoria de dos clones BAC.

En todos los datos de secuencias se usaron los programas de ensamblaje de secuencias BioEdit y Dna Baser. Los análisis revelaron que los dos clones BAC de 25 kpb aislados contenían la secuencia de ARNm de leucolectina completa (después de la eliminación de los elementos de la secuencia del vector), cuando se exploraba frente a secuencias de ARNm y secuencias de proteína. Los datos de uno de los clones BAC produjeron un total de 80 lecturas, que se ensamblaron en 33 cóntigos. De éstos, al menos 6 cóntigos contenían secuencias de LL, lo cual permitió la reconstrucción del gen de LL entero. Además de las secuencias de LL, en estos cóntigos se identificaron frecuentemente otros componentes con un alto grado de certidumbre, e incluyeron transcriptasa inversa, peptidasa iónica y elementos transponibles.

#### Características de la estructura génica genómica de LL

En el primer análisis, solo uno de los clones (L7/247) permitió la identificación de un gen de LL entero, mientras que el otro clon (A12/144) reveló la mayor parte de la estructura genómica (véase la Figura 16). Los datos del genómico L7/247 sugirieron una organización genómica de la leucolectina, donde el gen está compuesto de 5 exones interrumpidos por 4 intrones, y abarca ~ 2,3 kpb a partir de la caja TATA. Las secuencias de LL genómico revelaron una caja TATA que empezaba en la posición -81 (TATAAAA) desde el codón de iniciación ATG, indicando una 5'UTR de 81 nt (elemento no traducible) del ARNm. El codón de terminación TAG está localizada aproximadamente a 1850 nt de la ATG en la secuencia genómica. Una señal de poliadenilación de 6 nt longitud AATAAA empieza aproximadamente en la posición 2.250 desde ATG. La 3'UTR es una secuencia de aproximadamente 400 nt de longitud que empieza después del codón de terminación TAG. Con una 5'UTR que consiste en 81 nt, se estima un tamaño del gen de LL de al menos ~ 2,3 kpb por definiciones convencionales, representando los intrones dos veces la secuencia de los exones. Para verificar los límites exón-intrón del gen de la leucolectina, se diseñaron cebadores de PCR a partir de las secuencias de intrones genómicos establecidas, y se usaron para amplificar cada exón a partir del ADN embrionario de salmón. Estas cinco secuencias de exones se obtuvieron por análisis directo de la secuencia de amplicones, y todas correspondían a las secuencias genómicas establecidas.

Los tamaños de los exones codificantes parecían ser bastante poco ambiguos, y variaban entre 55 pb (exón 1) y 234 pb (exón 2). Notablemente, dos intrones eran relativamente pequeños de forma que podía obtenerse una longitud exacta a partir de la lectura manual de los datos. Por el contrario, el intrón 3 tiene aproximadamente 250 nt, mientras que el intrón 4 abarca aproximadamente 700 pb. La longitud exacta del intrón 4 no podría definirse a partir de los datos disponibles. Los datos de la secuencia de exones compuestos confirman el producto previsto de traducción, particularmente una proteína de 255 AA.

Además, los datos confirman una estructura del gen LL genómico con 5 exones y 4 intrones para una de las dos variantes descritas de leucolectina. El resto de AA N-terminal de la proteína LL secretada, según se ha establecido, es triptófano por degradación de Edman de la LL purificada. Este resto Trp ocupa la posición nº 20 desde la metionina N-terminal en la proteína precursora de 255 AA, y está precedido por los aminoácidos nº 18 (His) y nº 19 (Ala). Se prevé un sitio de escisión entre Ala y Trp para una proteasa de procesamiento, para generar la proteína LL secretada. Por degradación directa de Edman, la secuencia de esta LL secretada se ha establecido como WDCQE WNIK NLMQI DAGLG Q-V, y también se verifica por múltiples secuencias de proteínas virtuales. Los exones no están todos en fase, pero permiten potencialmente algún salto de exón sin romper todas las fases de lectura. El análisis de la secuencia genómica indicó las secuencias de los límites exón/intrón. La excepción es el intrón 4, donde posiblemente nt no convencionales flanquean las secuencias de los intrones. En los intrones 1, 2 y 3 se



encontró la característica común GT//AG que flanquea el comienzo y el final de muchos intrones de vertebrados (Shapiro y Senepathy, 1987, Nucleic Acids Research, 15, pág. 7155-7174).

El exón 1 (19 AA) tiene 55 nt de longitud, con la secuencia MRATA - AVLLV - LCLLT- ISH(A) y se detienen en la primera G del codón para el resto Ala nº 19 que precede el triptófano N terminal de la leucolectina secretada. La R grande en negrita y en cursiva significa un AA variante en esta posición posiblemente asociado a un gen de leucolectina 3 (véase la Figura 17). La secuencia génica genómica continúa con los nucleótidos GT para empezar el intrón 1. El intrón 1 termina con los nucleótidos AG seguidos del exón II, que contiene al menos dos bases (-CA) del codón Ala nº 19 seguido por el codón TGG para el Trp nº 20, el extremo N terminal de la leucolectina. El intrón 1 tiene 89 nt de longitud.

El exón 2 (234 nt) contiene 77 AA ( $234 - 2 \text{ nt} / 3 = 77 \text{ AA} + 1 \text{ nt}$ ). En este caso, las secuencia genómica traduce (ala) - WDCQEVVNIK - NLMQI DAGLG - QWAT DTSQI - PYYLV G\* DKWI - RL\* PGS LKHIT - VGPAG IWGVN -KDYAI YKYVA - GNWVQ AA (G) (= 77 AA + 1 nt) antes de continuar con los nucleótidos GT (en el intrón 2). El asterisco indica un resto en el que se encuentran AA alternativos en la proteína LL por múltiples secuencias del ADNc de LL. (véase la Figura 17). G\* parece variar en la leucolectina-1, mientras que L\* parece variar en la leucolectina-2. El intrón 2 tiene una longitud de 186 nt, que termina en los nucleótidos AG.

El exón 3 (72 AA) empieza después de 2 nucleótidos con GC (= Gly, resto nº 78) y continúa durante otros 216 nt (para un total de 218 nt) o 72 AA. La secuencia de aminoácidos del exón 3 es: (G)LL\* KQL DAGGE - QFIVG ANMN\* D TPYCLTSSAT - VGYKG PGSPL - PWTGL PGAVK - YYSCG P\* FGCW - AVNKN DDIYL - MS, donde el resto nº 150 es S. Tres letras mayúsculas, en negrita y subrayadas, significan AA variantes que sirven para distinguir entre leucolectina 1 y la leucolectina 2. Las letras con una asterisco de nuevo significan restos con microvariación (N\* en leucolectina 2; P\* significa microvariación tanto en leucolectina 1 como en leucolectina 2) en dicha posición (véase la Figura 17). Continúa el intrón 3, empezando con los nucleótidos GT y continúa durante otros ~ 250 nt.

El exón 4 (39 AA) empieza en el resto de AA nº 151 = LNQDC QNKGW - SHIE\* G KLSMI - EVATD GSVFG - VNSA\* GSVYT, donde T es el resto nº 189. De nuevo el resto subrayado en negrita sirve para distinguir la leucolectina 1 de la leucolectina 2, mientras que un asterisco significa una microvariación en esta posición (véase en la leucolectina en posición E\*, y solo en la leucolectina 2 para la posición A\*). El intrón 4 posiblemente puede no empezar y detenerse con las secuencias de nt convencionales para la definición de intrones. De esta manera los datos, en el intrón 4 son algo preliminares, y por lo tanto la longitud del intrón 4 es una estimación (~700 nt).

El exón 5 (47 AA) empieza con el codón para R (= resto nº 190) y continúa hasta una histidina C terminal (resto nº 236), que va seguida de un codón determinación TAG y la 3'UTR. La secuencia se lee: RDGIT-ASKPE-GTGWS-NIPMG\*-MLMGHVYDYL-GRLWV-VKSAV-TMVCT-H, donde un asterisco indica microvariación en esta posición en la leucolectina 1. Las letras en negrita significan diferencias entre la leucolectina 1 y la leucolectina 2. Las letras en negrita y en cursiva significan la posibilidad de existencia de leucolectina 3. Estos datos se resumen en la Figura 17.

Ejemplo 11: comparación de secuencia genómica de leucolectina 1 de salmón con secuencias de ARNm de leucolectina

Con más de una docena de experimentos de clonación y secuenciación separados del ADNc de salmón, las secuencias alineadas revelaron que solo 7 posiciones son claras alternativas de diferentes restos de aminoácido acomodados. La falta de variación de secuencia global y la limitación estereotípica de estas variaciones es destacable.

La secuencia genómica de L-7/247 corresponde a clones de leucolectina clasificados como leucolectina 1. Los otros ADNc secuenciados denominados leucolectina 2, difieren en las siete posiciones indicadas, además de estas posiciones, hay indicaciones de algunas variaciones adicionales en AA. Esto puede sugerir una microheterogeneidad adicional de los dos ADNc clasificados, pero los datos pueden sugerir la existencia de una tercera categoría de ADNc de LL, ya que, por ejemplo, en la posición nº 226, encontramos un resto de tirosina en algunos clones que nunca se observó ni el ADNc de la leucolectina 1 ni el de la leucolectina 2 (Figura 17).

En la Figura 17 se muestra un resumen de variaciones observadas. Los dos tipos de leucolectinas se caracterizan por respectivamente (E, I, Y, Y, K, A, V) frente a (N, V, F, F, N, G, G) en la siete posiciones (nº 88, 91, 101, 147, 158, 229, 230) de la secuencia de AA deducida a partir de la leucolectina madura. Estas 7 posiciones clasifican de forma inequívoca estas dos leucolectinas, que además pueden presentar microvariaciones en posiciones estereotípicas mostradas por asterisco en la Figura 17. Como indican los datos de la Figura 17 algunas secuencias de ARNm indican una sola Gly en la posición nº 2 del péptido señal de LL que no se observa en las leucolectinas 1 y 2, y que, por lo tanto, sugiere una posible leucolectina-3. Además, los presentes inventores han observado una Tyr en la posición nº 226 en la leucolectina madura donde solo se observa Ser en las leucolectinas 1 y 2. Los presentes inventores concluyen que la secuencia genómica en L7/247 corresponde a secuencias de la leucolectina 1 encontradas por múltiple secuenciación de ADNc de LL de salmón. La secuenciación adicional de clones BAC de LL revelaría la otra categoría principal de leucolectina 2 y quizás también la existencia sospechada de leucolectina 3, como se define por los criterios de restos anteriores.

Ejemplo 12: Detección de proteínas leucolectina en otras especies

El fluido de eclosión de la trucha arco iris (*Oncorhynchus mykiss*), bacalao y *Oikopleura dioica* se investigó con respecto a la presencia de proteínas leucolectinas. Se preparó proteína a partir de fluido de eclosión por cromatografía de afinidad analizada por SDS-PAGE al 15 %. El gel se tiñó con plata para mostrar la presencia de una diversidad de proteínas, o se sometió a transferencia y se sondeo con un anticuerpo de leucolectina de salmón. en cualquier caso se obtuvo una sola proteína de ~ 26 kDa que correspondía al peso molecular de la leucolectina (véanse las figuras 18 (trucha arco iris), 19A (bacalao) y 19B (*Oikopleura dioica*)).

Secuencias:

1: (polipéptido de leucolectina de embrión de salmón):

MRTTAAFLVLCLLAISHAWDCQEVVNIKNLMQIDAGLGQVVATDTSQIPYYLVGDKWIRLP  
GSLKHITVGPAGIWGVNKDYAIYKYVAGNWVQAAGLLKQLDAGGEQFIVGANMNNDTPYCLT  
SSATVGYKGPGSPLPWTGLPGAVKYYSCGPFGCWAVNKNDDIYLSLNQDCQNKGWSHI  
EGKLSMIEVATDGSVFGVNSAGSVYTRDGITASKPEGTGWSNIPMGMLMGHVTYDLGRLW  
VVSksAVTMVCTH

2: (polipéptido de leucolectina de leucocitos de salmón):

SIPYYLVGDKWIRLP GSLKHITVGPAGIWGVNKDYAIYKYVAGNWVQAAGLPKQLDAGGEQ  
FIVGANMDDTPYCLTSSATVGYKGPGSPLPWTGLPGAVKYYSCGPFGCWAVNKNDDIYLM  
SLNQDCQNNGWSHIEGKLSMIEVATDGSVFGVNSAGSVYTRDGITASKPEGTGWSNIPMC  
MLMGHVTYDLGRLWVVSksAVTMVCTH

3: (polipéptido de leucolectina de leucocitos de bacalao):

LVGDKWIRLP GSLKHITVGPAGIWGVNKDYAIYKYVAGNWVQAAGLLKQLDAGGEQFIVGA  
NMNDTPYCLTSSATVGYKGPGSPLPWTGLPGAVKYYSCGPFGCWAVNKNDDIYLSLNQ  
DCQNKGWSHIDGKLSMIEVATDGSVFGVNSAGSVYTRD

4: (polipéptido de leucolectina de leucocitos de pollo):

IPYYLVGDKWIRLP GSLKHITVGPAGIWGVNKDYAIYKYVAGNWVQAAGLLKQLDAGGEQFI  
VGANMNNDTPYCLTSSATVGYKGPGSPLPWTGLPGAVKYYSCGPFGCWAVNKNDDIYLS  
LNQDCQNKGWSHIDGKLSMIEVATDGSVFGVNSAGSVYTRDGITASKPEGTGWSNIPMG  
MLMGHVTYDLGRLWV

5: (polipéptido de leucolectina 1 de leucocitos humanos):

MRATAAVLLVLCLLTISHAWDCQEVVNIKNLMQIDAGLGQVVATDTGRIPYYLVGDKWIRLP  
GSLKHVTVPAGIWGVNKDYAIYKYVAGNWVQAAGLLKQLDAGGEQFIVGANMNNDTPYRL  
TSSATVGYKGPGSPLPWTGLPGAVKYYSCGPFGCWAVNKNDDIYLSLNQDCQNKGWSH  
IEGKLSMIEVATDGSVFGVNSAGSVYTRDGITASKPEGTGWSNIPMGMLMGHVTYDLGRL  
WVVSksGGTMVCTH

6: (polipéptido de leucolectina de pez cebra):

MGTTAAFLLVLCCLAISHAWDCQEVVNIKNLMQIDAGLGQVVATDTSQIPYYLVGDKWIRLP  
GSLKHITVGPAGIWGVNKDYAIYKYVAGNWVQAAGLLKQLDAGGNQFIVGANMNDTPYRLT  
SSATVGYKGPGSPLPWTGLPGAVKYYSCGPFGCWAVNKNDDIYLSLNQDCQNKGWSHI  
EGKLSMIEVATDGSVFGVNSAGSVYTRDGITASKPEGTGWSNIPMGMLMGHVTYDLGRLW  
VVTKSGGTMVCTH

7: (polipéptido de leucolectina 2 de salmón):

MRTTAAFLLVLCCLAISHAWDCQEVVNIKNLMQIDAGLGQVVATDTSQIPYYLVGDKWIRLP  
GSLKHITVGPAGIWGVNKDYAIYKYVAGNWVQAAGLLKQLDAGGNQFVVGANMDDTPFCL  
TSSATVGYKGPGSPLPWTGLPGAVKYYSCGHFGCWAVNKNDDIFLSLNQDCQNNGWS  
HIDGKLSMIEVATDGSVFGVNSAGSVYTRDGITASKPEGTGWSNIPMGMLMGHVTYDLGRL

WVSKSGGTMVCTH

5

8: (polipéptido de leucolectina 3 de salmón):

MGTTAAFLLVLCCLAISHAWDCQEVVNIKNLMQIDAGLGQVVATDTSQIPYYLVGDKWIRLP  
GSLKHITVGPAGIWGVNKDYAIYKYVAGNWVQAAGLLKQLDAGGEQFIVGANMNDTPYCLT  
SSATVGYKGPGSPLPWTGLPGAVKYYSCGPFGCWAVNKNDDIYLSLNQDCQNKGWSHI  
EGKLSMIEVATDGSVFGVNSAGSVYTRDGITASKPEGTGWSNIPMGMLMGHVTYDLGRLW  
VVYKSAVTMVCTH

10

9: (secuencia de ADNc de leucolectina de sangre humana):

ATGAGAGCCACTGCAGCCGTCCTATTGGTCCTCTGTCTCCTGACCATCAGTÇATGCATG  
GGACTGTCAGGAGGTAGTAAACATCAAGAATCTGATGCAGATCGATGCAGGACTGGGG  
CAAGTGGTTGCTACGGACACAGGTGCAATCCCCTACTACCTGGTAGGTGATAAATGGA  
TCCGTCTGCCTGGTTCCCTGAAGCATGTCACTGTAGGACCAGCAGGGATCTGGGGTGT  
CAACAAGGACTATGCAATCTACAAGTATGTGGCCGGTAACTGGGTTCAAGCTGCAGGC  
CTTCTGAAACAGTTGGATGCTGGAGGTGAACAGTTTATTGTGGGGGCTAACATGAACG  
ATACTCCATACCGTCTGACAAGTAGTGCCACAGTTGGCTACAAGGGTCCAGGCTCACC  
CCTTCCATGGACAGGATTGCCAGGAGCTGTGAAGTACTACAGCTGCCGACCCTTTGGG  
TGCTGGGCAGTCAACAAGAATGATGATATCTACTTAATGAGTCTGAATCAAGACTGCCA  
AAACAAGGGGTGGAGTCACATTGAAGGCAAGCTTTCCATGATTGAGGTGGCAACTGAT  
GGTAGTGTCTTTGGGGTCAACTCTGCGGGTAGTGTTTATACCAGAGACGGCATCACAG  
CCAGTAAACCAGAGGGGCACCGGATGGAGCAATATCCCAATGGGCATGCTCATGGGCCA  
CGTGACCTACGACCTGGGCCGTCTTTGGGTCTCTCCAAGTCTGGCGGCACCATGGT  
GTGCACACATTAGCCTCTTCTCTGTAGCTGAAGGCCGTTTCGGGATCTGTCTAAAGTTCA  
CTTGCGAACTCATTGATCTCTCTTTCTGGAAAAGCCTTTAGTTCATTAGTTCATAAAAAAT  
CCTTCATTTTAAACCTATTGCTCTACCTATTATTTTCAGTTCTTCAATTATCTTATTGCC  
ATTTAAAAAATATCAATGAAGATGTTATATTTTCTTGACCACTCCTTGATTAACACTTCA  
ATAGATCTTTGCCATGGAGGCTATTTAAGTGTAGTGTAAGTGGGCATGGTCATGTTG  
CTCACAATCCACATGGGTTTTGCTGTGCTTCAGAGGTCATCAATAGGATTGGATGGAAT  
CCTTGTCATTGTTTATTATCTCATTATATAAACATTTCTGCAAAAATAAAGCATTCAATT  
TGAAACTATTGTAAAAAAAAAAAAAAAAAAAAAAAAAAN

10: (secuencia del gen de leucolectina de salmón, procedente de una biblioteca de ADN genómico de salmón maduro):

TGTGCAGGGCTATAAAAGCGCAAAGTCTTCCAATGGGACAATTGAAGTCTGGTGTACAA  
 CCAAACGTATACTGTAGATCTACATAGACATCATGAGAGCCACTGCAGCCGTCCTATTG  
 GTCCTCTGTCTCCTGACCATCAGTCATGGTAAGTTACCATCATCTGAAACATGCTTGAT  
 CAACTTGGAGTTGAAGTTTTTCTTGGTATACTCTACTCATATGTCTTTGTCTCCATAGCA  
 TGGGACTGTCAGGAGGTAGTAAACATCAAGAATCTGATGCAGATCGATGCAGGACTGG  
 GGCAAGTGGTTGCTACGGACACAAGTCAAATCCCCTACTACCTGGTAGGTGATAAATG  
 GATCCGTCTGCCTGGTTCCCTGAAGCATATCACTGTAGGACCAGCAGGGATCTGGGGT  
 GTCAACAAGGACTATGCAATCTACAAGTATGTGGCCGGTAACTGGGTTCAAGCTGCAG  
 GTAAGTGGAGAGCATTACTCAATATTTATCCAGAGGACACCTGCTTATTAGCTTTCTG  
 ATACCATCAGGCTGTTGAAAAAACGATTGATGTTTTAAATTGTAAGTTGTAGGTAATTT  
 GGCAGTACTCCTTGTTTGCTTGTCTGTCTGTCTTTGTGGTCTTGGCCTTCTGAAACAGT  
 TGGATGCTGGAGGTGAACAGTTTATTGTGGGGGCTAACATGAACGATACTCCATACTGT  
 CTGACAAGTAGTGCCACAGTTGGCTACAAGGGTCCAGGCTCACCCCTTCCATGGACAG  
 GATTGCCAGGAGCTGTGAAGTACTACAGCTGCGGACCCTTTGGGTGCTGGGCAGTCAA  
 CAAGAATGATGATATCTA

CTTAATGAGTGTAAGATCTGGGAAAGAGTGGGAGAGCTGGAGTAGAGTAGTAGAGGAT  
 GGAGAGTGTGAGTTATTTTAAACTGTTTCATATTATAACTGTTGAAATTGTCCTAAAC  
 CCTGATTGTATCATTTTGTTTCCAGCTGAATCAAGACTGCCAAAACAAGGGGTGGAGTC  
 ACATTGAAGGCAAGCTTTCCATGATTGAGGTGGCAACTGATGGTAGTGTCTTTGGGGT  
 CAACTCTGCGGGTAGTGTTTATACCCAGGTAAGGTTGCTACTGAACTATGTGTATGGTC  
 CACCACCCCCCCCCCCCCAACAGTATTAAGTTGAAAATGACTTGTAATAATAACTTAGAA  
 TAATAATGGTATACCCTTTAATTATAACTCTGATCCTTACAGTACATGCTATGTGAATCTC  
 CTTACACAAAACTAAATATTGTAGGTACATAAATAAAATCAGTTAAATATAATCAGATCT  
 AAAGTTATAGGACTTATTAAGAAATGTGTAGACAGTGTATGATAAAATATGTAAAAGTTG  
 GATGTCCTGTAAAGCTACAGTTTGGGATAAAAAACAACAACCTCCCAGACACCCACCA  
 CTTGTTCTGGTAAACAGCTGAGGAATGTAGTTAGAGAAATGTAACCACTCTCACATTCA  
 TACATGGAGCTACGGATGCAAAGACACAACAATTTTTTTTATTAATAAAAAAAAAAATGTTT  
 ATATTTTCTTTTAAAGCTAAACATTGTTTGTGTTACAATAACATTGTTTACAAACAATTGAG  
 TAAAAGCTTACATTTTGGCTTCTAATGTGGTTGAATAAAGCTCAAGATGCAGAAGTTATA  
 TTCTTCAAAAATCTATGGCTATATTTAATTATTAAAGTCCAAAATGGATGTACTTAAAAA  
 AATGGATAAGCTTTAAACATGAACCCCAACCCTTTCTTCAACACAGAGACGGCATCA  
 CAGCCAGTAAACCAGAGGGCACCGGATGGAGCAATATCCCAATGGGCATGCTCATGG

GCCACGTGACCTACGACCTGGGCCGTCTTTGGGTCGTCTCCAAGTCTGGCGGCACCAT  
GGTGTGCACACATTAGCCTCTTCTCTGTAGCTGAAGGCCGTTTCGGGATCTGTCTAAAGT  
TCACTTGCGAACTCATTGATCTCTCTTTCTGGAAAAGCCTTTAGTTCATTAGTTCATAAA  
AATCCTTCATTTTAAACCTATTGCTCTACCTATTATTTTCAGTTCTTCAATTATCTTATTG  
CCATTTAAAAAATATCAATGAAGATGTTATATTTTCTTGACCACTCCTTGATTAACACTT  
CAATAGATCTTTGCCATGGAGGCTATTTAAGTGTAGTGTAAGTGGGACGGTCATGT  
TGCTCACAATCCACATGGGTTTTGCTGTGCTTCAGAGGTCATCAATAGGATTGACGGA  
ATCCTTGTCATTGTTTATTATCTCATTATATAATCATTTCTGCAAAAATAAA

11: (secuencia de ADNc de leucolectina de embriones de salmón de 200-450 días de edad - exones solo):

ATGAGAGCCA CTGCAGCCGT CCTATTGGTC CTCTGTCTCC TGACCATCAG  
TCATGCATGG GACTGTCAGG AGGTAGTAAA CATCAAGAAT CTGATGCAGA  
TCGATGCAGG ACTGGGGCAA GTGGTTGCTA CGGACACAAG TCAAATCCCC  
TACTACCTGG TAGGTGATAA ATGGATCCGT CTGCCTGGTT CCCTGAAGCA  
TATCACTGTA GGACCAGCAG GGATCTGGGG TGTCACAAG GACTATGCAA  
TCTACAAGTA TGTGGCCGGT AACTGGGTTT AAGCTGCAGG CCTTCTGAAA  
CAGTTGGATG CTGGAGGTGA ACAGTTTATT GTGGGGGCTA ACATGAACGA  
TACTCCATAC TGTCTGACAA GTAGTGCCAC AGTTGGCTAC AAGGGTCCAG  
GCTCACCCT TCCATGGACA GGATTGCCAG GAGCTGTGAA GTACTACAGC  
TGCGGACCCT TTGGGTGCTG GGCAGTCAAC AAGAATGATG ATATCTACTT  
AATGAGTCTG AATCAAGACT GCCAAAACAA GGGGTGGAGT CACATTGAAG  
GCAAGCTTTC CATGATTGAG GTGGCAACTG ATGGTAGTGT CTTTGGGGTC  
AACTCTGCGG GTAGTGTTTA TACCAGAGAC GGCATCACAG CCAGTAAACC  
AGAGGGCACC GGATGGAGCA ATATCCCAAT GGGCATGCTC ATGGGCCACG  
TGACCTACGA CCTGGGCCGT CTTTGGGTCT TCTCCAAGTC TGCGGGCACC  
ATGGTGTGCA CACAT

5

12: (secuencia del gen de leucolectina de leucocitos de salmón):

ACTGGGGCAAGTGGTTGCTACGGACACAAGTCAAATCCCCTACTACCTGGTAGGTGAT  
 AAATGGATCCGTCTGCCTGGTTCCCTGAAGCATATCACTGTAGGACCAGCAGGGATCT  
 GGGGTGTCAACAAGGACTATGCAATCTACAAGTATGTGGCCGGTAACTGGGTTCAAGC  
 TGCAGGCGTTCCGAAACAGTTGGATGCTGGAGGTAACCAAGTTTGTGTGGGGGCTAAC  
 ATGGACGATACTCCATTTTGTCTGACAAGTAGTGCCACAGTTGGCTACAAGGGTCCAG  
 GCTCACCCCTTCCATGGACAGGATTGCCAGGAGCTGTGAAGTACTACAGCTGCGGACA  
 CTTTGGGTGCTGGGCAGTCAACAAGAATGATGATATTTTCTTAATGAGTCTGAATCAAG  
 ACTGCCAAAACAACGGGTGGAGTCACATTGATGGCAAGCTTTCCATGATTGAGGTGGC  
  
 AACTGATGGTAGTGTCTTTGGGGTCAACTCTGCGGGTAGTGTTTATACCAGAGACGGC  
 ATCACAGCCAGTAAACCAGAGGGCACC GGATGGAGCAATATCCCAATGGGCATGCTCA  
 TGGGCCACGTGACCTACGACCTGGGCCGTCTTTGGGTGCTCTCCAAGTCTGGCGGCA  
 CCATGGTGTGCACACATTAGCCTCTTCTCTGTAGCTGAAGGCCGT

13: (secuencia del gen de leucolectina de bacalao):

CAATCCCCTACTACCTGGTAGGTGATAAATGGATCCGTCTGCCTGGTTCCCTGAAGCAT  
 ATCACTGTAGGACCAGCAGGGATCTGGGGTGTCAACAAGGACTATGCAATCTACAAGT  
 ATGTGGCCGGTAACTGGGTTCAAGCTGCAGGCCTTCTGAAACAGTTGGATGCTGGAGG  
 TGAACAGTTTATTGTGGGGGCTAACATGAACGATACTCCATACTGTCTGACAAGTAGTG  
 CCACAGTTGGCTACAAGGGTCCAGGCTCACCCCTTCCATGGACAGGATTGCCAGGAGC  
 TGTGAAGTACTACAGCTGCGGACCCTTTGGGTGCTGGGCAGTCAACAAGAATGATGAT  
 ATCTACTTAATGAGTCTGAATCAAGACTGCCAAAACAAGGGGTGGAGTCACATTGAAGG  
 CAAGCTTTCCATGATTGAGGTGGCAACTGATGGTAGTGTCTTTGGGGTCAACTCTGCG  
 GGTAGTGTTTATACCAGAGAC

5

14: (secuencia del gen de leucolectina de pollo):

CAGGACTGGGGCAAGTGGTTGCTACGGACACAAGTCAAATCCCCTACTACCTGGTAGG  
 TGATAAATGGATCCGTCTGCCTGGTTCCCTGAAGCATATCACTGTAGGACCAGCAGGG  
 ATCTGGGGTGTCAACAAGGACTATGCAATCTACAAGTATGTGGCCGGTAACTGGGTTT  
 AAGCTGCAGGCCTTCTGAAACAGTTGGATGCTGGAGGTGAACAGTTTATTGTGGGGGC  
 TAACATGAACGATACTCCATACTGTCTGACAAGTAGTGCCACAGTTGGCTACAAGGGTC  
 CAGGCTCACCCCTTCCATGGACAGGATTGCCAGGAGCTGTGAAGTACTACAGCTGCGG  
 ACGCTTTGGGTGCTGGGCAGTCAACAAGAATGATGATATCTACTTAATGAGTCTGAATC  
 AAGACTGCCAAAACAAGGGGTGGAGTCACATTGAAGGCAAGCTTTCCATGATTGAGGT  
 GGCAACTGATGGTAGTGTCTTTGGGGTCAACTCTGCGGGTAGTGTTTATACCAGAGAC  
 GGCATCACAGCCGTAAACCAGAGGGCACC GGATGGAGCAATATCCCAATGGGCATGCT  
 CATGGGCCACGTGACCTACGACCTGGGCCGTCTTTGGGTGCTCT

10

15: (secuencia genómica parcial de leucolectina de pez cebra):

NNAAAAANTAAAATANNGTAGGTACANNNNAAATCAGTTAAATATAATCAGATNTAAANTT  
NTAGGACTATTAAGAAATGTGTAGACAGCGTACGATAAAATATGTAAAAGTTGGATGTCC  
TGTAAGCTACAGTTTGGGATAAAAAACAACAACCTNCCAGACACCCCACCACTTGTTNT  
GGTAAACAGNTGAGGAATGTAGTTAGAGAAATGTAACCANTCTCACATTCATACATGGAG  
TTACGGATGCAAAGACACAACAATTTTTTTGTTTACATTNTTTTAAACATGTTTTTTAAAGCA  
ATACACATTGTTTGTTTACAATAACATTGTTTACAAACAATTGAGTAAAAGCTTACATTTTG

GCTTCTGTGTGGTTGAAATAAAGCTCAAGAGGCAGAAGTTATATTCTTCAAAAATCAATG  
GCTATATTTAATTATTAAAGTTCCAAAAGGATGTACTTAATAAAATGGATAAGCTTTAAAA  
CATGAACCCCAACCCTTTCTTCAACACAGAGACGGCATCACAGCCAGTAAACCAGAGGG  
CACCGGATGGAGCAATATCCCAATGTGTATGCTCATGGGCCACGTGACCTACGACCTGG  
GCCGTCTTTGGGTCGTCTCCAAGTCTGCCGTACCATGGTGTGCACACATTAGCCTCTT  
CTCTGTAGCTGAAGGCCGTTCCGGATCCGTCCAAAGTTCCCTGGCGAACTCATTGATCT  
CTCTTTCTGGAAAAGCCTTTAGTTTATAAAAAATCCTTCATTTTAAACCTATTGCTCTACCT  
ATTATTTTCAGTTCTTCAATGATCTTATTGACATTTAAAAAAAATATCATTGAAGATTTTATA  
TTTTCTTGACAACTCCTAGATTAACACTTCAATAGACCTTTGCCATGGAGGCTATTTAAGT  
GTAGTGTAAGTAACTAGGGCACGGTCATGTTGCTCACAATCCACATGGGTTTTGCTGTGCTTC  
AAAGGTCATCAATAAATCACTAGTGCGGCCGCTGCAGGTGACCATATGGGAGAGCTC  
CCAACGCGTNGGATGCATAAG

16: (polipéptido de leucolectina 2 de leucocitos humanos):

5

WDCQEVVNIKMLMQIDAGLGQVVATDTSQIPYYLVGDKWIRLPGSLKHITVGPAGIWGVNK  
DYAIYKYVAGNWVQAAGLLKQLDAGGNQFVVGANMDDTPYCLTSSATVGYKGPGSPLPW  
YGLPGAVKYYSCGPFGCWAVNKNDDIYLSLNQDCQNKGWSHIEGKLSMIEVATDGSVFG  
VNSAGSVYTRDGITASKPEGTGWSNIPMGLLMGHVDLRPGHPWVCDHTGWSNIPMGLLM  
GHVDLRPGPSLGRLQVCRH

17: (polipéptido de leucolectina-3 de leucocitos humanos - truncado):

GSLKHITVGPAGIWGVNKDYAIYKYVAGNWVQAAGLLKQLDAGGEQFIVGANMNDTPYCLT  
SSATVGYKGPGSPLPWTGLPGAVKYYSCGPFGCWAVNKNDDIYLSLNQDCQNKGWSHI  
EGKLSMIEVATDGSVFGVNSAGSVYTRDGITASKPEGTGWSNIPMGLLMGHVDLRPGPSL  
GRLQVCRHPWVCTH

10

18: (polipéptido de leucolectina 3 de leucocitos humanos - sin péptido señal):



WDCQEVVNIKNLMQIDAGLGQVVATDTSQIPYYLVGDKWIRLPGSLKHITVGPAGIWGVNK  
DYAIYKYVAGNWWQAAGLLKQLDAGGEQFIVGANMNDTPYCLTSSATVGYKGPGSPLPWT  
GLPGAVKYYSCGPFGCWAVNKNDDIYLSLNQDCQNKGWSHIEGKLSMIEVATDGSVFGV  
NSAGSVYTRDGITASKPEGTGWSNIPMGLLMGHVDLRPGPSLGRLQVCRHPWVCTH

19: (secuencia de ADNc de leucolectina 3 de sangre humana - truncada):

GGTTCCCTGAAGCATATCACTGTAGGACCAGCAGGGATCTGGGGTGTCAACAAGGACT  
ATGCAATCTACAAGTATGTGGCCGTAAGTGGGTTCAAGCTGCAGGCCTTCTGAAACA  
GTTGGATGCTGGAGGTGAACAGTTTATTGTGGGGGCTAACATGAACGATACTCCATACT

GTCTGACAAGTAGTGCCACAGTTGGCTACAAGGGTCCAGGCTCACCCCTTCCATGGAC  
AGGATTGCCAGGAGCTGTGAAGTACTACAGCTGCGGACCCTTTGGGTGCTGGGCAGT  
CAACAAGAATGATGATATCTACTTAATGAGTCTGAATCAAGACTGCCAAAACAAGGGGT  
GGAGTCACATTGAAGGCAAGCTTTCCATGATTGAGGTGGCAACTGATGGTAGTGTCTTT  
GGGGTCAACTCTGCGGGTAGTGTATACCAGAGACGGCATCACAGCCAGTAAACCAG  
AGGGCACCGGATGGAGCAATATCCCAATGGGCCTGCTCATGGGCCACGTGGACCTAC  
GACCTGGGCCGTCTTTGGGTGCTCTCCAAGTCTGCCGTCACCCATGGGTGTGCACACA  
T

5

20: (polipéptido de leucolectina 4 procedente de humano - sin péptido señal):

WDCQEVVNIKNLMQIDAGLGQVVATDTSQIPYYLVGDKWIRLPGSLKHITVGPAGIWGVNK  
DYAIYKYVAGNWWQAAGLLKQLDAGGNQFIVGANMNDTPYCLTSSATVGYKGPGSPLPWT  
GLPGAVKYYSCGPFGCWAVNKNDDIYLSLNQDCQNKGWSHIEGKLSMIEVATDGSVFGV  
NSAGSVYTRDGITASKPEGTGWSNIPMGLLMGHVDLRPGPSLGRLQVCRHPWVCTH

10

21: (polipéptido de leucolectina-5 procedente de humano - sin péptido señal):

WDCQEVVNIKNLMQIDAGLGQVVATDTSQIPYYLVGDKWIRMPGYLKMITVGPGGIWGVNK  
DYAIYKYVAGNWWQAAGLLKQLDAGGNQFVVGANMDDTPFCLTSSATVGYKGPGSPLPW  
TGLPGAVKYYSCGHFGCWAVNKNDDIFLSLNQDRQNNGWSHIDGKLSMIEVATDGSVFG  
VNSAGSVYTRDGITASKPEGTGWSNIPMGLLMGHVDLRPGPSLGRLQVCRHPWVCTH

15

22: (polipéptido de leucolectina 5 procedente de la parte humana sin extremo C-terminal):

WDCQEVVNIKNLMQIDAGLGQVVATDTSQIPYYLVGDKWIRMPGYLKMITVGPGGIWGVNK  
DYAIYKYVAGNWWQAAGLLKQLDAGGNQFVVGANMDDTPFCLTSSATVGYKGPGSPLPW  
TGLPGAVKYYSCGHFGCWAVNKNDDIFLSLNQDRQNNGWSHIDGK

20

23: (polipéptido de leucolectina 5 procedente de ser humano - variante truncada):

WDCQEVVNIKNLMQIDAGLGQVWATDTSQIPYYLVGDKWIRMPGYLKMITVGPGGIWGVNK  
DYAIYKYVAGNWWQAAGLLKRLDAGGDQFVVGLT

LISTADO DE SECUENCIAS

5	<110> LEUKOLECT AS
	<120> Leucolectinas humanas y usos de las mismas
	<130> 42.20.105534/01
10	<150> GB1007202.3
	<151> 29-04-2010
	<160> 46
15	<170> PatentIn versión 3.5
	<210> 1
	<211> 255
20	<212> PRT
	<213> <i>Salmo salar</i>
	<400> 1

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

<210> 2  
 <211> 208

<212> PRT  
 <213> *Salmo salar*

<400> 2

5

```

Ser Ile Pro Tyr Tyr Leu Val Gly Asp Lys Trp Ile Arg Leu Pro Gly
1           5           10           15

Ser Leu Lys His Ile Thr Val Gly Pro Ala Gly Ile Trp Gly Val Asn
          20           25           30

Lys Asp Tyr Ala Ile Tyr Lys Tyr Val Ala Gly Asn Trp Val Gln Ala
          35           40           45

Ala Gly Leu Pro Lys Gln Leu Asp Ala Gly Gly Glu Gln Phe Ile Val
          50           55           60

Gly Ala Asn Met Asp Asp Thr Pro Tyr Cys Leu Thr Ser Ser Ala Thr
65           70           75           80

Val Gly Tyr Lys Gly Pro Gly Ser Pro Leu Pro Trp Thr Gly Leu Pro
          85           90           95

Gly Ala Val Lys Tyr Tyr Ser Cys Gly Pro Phe Gly Cys Trp Ala Val
          100          105          110

Asn Lys Asn Asp Asp Ile Tyr Leu Met Ser Leu Asn Gln Asp Cys Gln
          115          120          125

Asn Asn Gly Trp Ser His Ile Glu Gly Lys Leu Ser Met Ile Glu Val
          130          135          140

Ala Thr Asp Gly Ser Val Phe Gly Val Asn Ser Ala Gly Ser Val Tyr
145           150           155           160

Thr Arg Asp Gly Ile Thr Ala Ser Lys Pro Glu Gly Thr Gly Trp Ser
          165          170          175

Asn Ile Pro Met Cys Met Leu Met Gly His Val Thr Tyr Asp Leu Gly
          180          185          190

Arg Leu Trp Val Val Ser Lys Ser Ala Val Thr Met Val Cys Thr His
          195          200          205

```

<210> 3  
 <211> 158  
 <212> PRT  
 <213> *Gadus morhua*

10

$\langle 400 \rangle$  3

Leu Val Gly Asp Lys Trp Ile Arg Leu Pro Gly Ser Leu Lys His Ile  
1 5 10 15

Thr Val Gly Pro Ala Gly Ile Trp Gly Val Asn Lys Asp Tyr Ala Ile  
20 25 30

Tyr Lys Tyr Val Ala Gly Asn Trp Val Gln Ala Ala Gly Leu Leu Lys  
35 40 45

Gln Leu Asp Ala Gly Gly Glu Gln Phe Ile Val Gly Ala Asn Met Asn  
50 55 60

Asp Thr Pro Tyr Cys Leu Thr Ser Ser Ala Thr Val Gly Tyr Lys Gly  
65 70 75 80

Pro Gly Ser Pro Leu Pro Trp Thr Gly Leu Pro Gly Ala Val Lys Tyr  
85 90 95

Tyr Ser Cys Gly Pro Phe Gly Cys Trp Ala Val Asn Lys Asn Asp Asp  
100 105 110

Ile Tyr Leu Met Ser Leu Asn Gln Asp Cys Gln Asn Lys Gly Trp Ser  
115 120 125

His Ile Asp Gly Lys Leu Ser Met Ile Glu Val Ala Thr Asp Gly Ser  
130 135 140

Val	Phe	Gly	Val	Asn	Ser	Ala	Gly	Ser	Val	Tyr	Thr	Arg	Asp
145					150					155			

5

$\langle 210 \rangle$  4

<211> 196

<212> PRT

<213> *Gallus gallus*

10

<400> 4

Ile Pro Tyr Tyr Leu Val Gly Asp Lys Trp Ile Arg Leu Pro Gly Ser  
1 5 10 15

Leu Lys His Ile Thr Val Gly Pro Ala Gly Ile Trp Gly Val Asn Lys  
20 25 30

Asp Tyr Ala Ile Tyr Lys Tyr Val Ala Gly Asn Trp Val Gln Ala Ala  
35 40 45

Gly Leu Leu Lys Gln Leu Asp Ala Gly Gly Glu Gln Phe Ile Val Gly  
50 55 60

Ala Asn Met Asn Asp Thr Pro Tyr Cys Leu Thr Ser Ser Ala Thr Val  
65 70 75 80

Gly Tyr Lys Gly Pro Gly Ser Pro Leu Pro Trp Thr Gly Leu Pro Gly  
85 90 95

Ala Val Lys Tyr Tyr Ser Cys Gly Pro Phe Gly Cys Trp Ala Val Asn  
100 105 110

Lys Asn Asp Asp Ile Tyr Leu Met Ser Leu Asn Gln Asp Cys Gln Asn  
115 120 125

Lys Gly Trp Ser His Ile Asp Gly Lys Leu Ser Met Ile Glu Val Ala  
130 135 140

Thr Asp Gly Ser Val Phe Gly Val Asn Ser Ala Gly Ser Val Tyr Thr  
145 150 155 160

Arg Asp Gly Ile Thr Ala Ser Lys Pro Glu Gly Thr Gly Trp Ser Asn  
165 170 175

Ile Pro Met Gly Met Leu Met Gly His Val Thr Tyr Asp Leu Gly Arg  
180 185 190

Leu Trp Val Val  
195

<210> 5

<211> 255

<212> PRT

<213> *Homo sapiens*

<400> 5

Met Arg Ala Thr Ala Ala Val Leu Leu Val Leu Cys Leu Leu Thr Ile  
1 5 10 15

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

<210> 6  
 <211> 255  
 <212> PRT  
 <213> *Danio rerio*

<400> 6

[illegible]

5           <210> 7  
              <211> 255  
              <212> PRT  
              <213> *Salmo salar*



&lt;400&gt; 7

Met Arg Thr Thr Ala Ala Phe Leu Leu Val Leu Cys Leu Leu Ala Ile  
1 5 10 15

Ser His Ala Trp Asp Cys Gln Glu Val Val Asn Ile Lys Asn Leu Met  
20 25 30

Gln Ile Asp Ala Gly Leu Gly Gln Val Val Ala Thr Asp Thr Ser Gln  
35 40 45

Ile Pro Tyr Tyr Leu Val Gly Asp Lys Trp Ile Arg Leu Pro Gly Ser  
50 55 60

Leu Lys His Ile Thr Val Gly Pro Ala Gly Ile Trp Gly Val Asn Lys  
65 70 75 80

Asp Tyr Ala Ile Tyr Lys Tyr Val Ala Gly Asn Trp Val Gln Ala Ala  
85 90 95

Gly Leu Leu Lys Gln Leu Asp Ala Gly Gly Asn Gln Phe Val Val Gly  
100 105 110

Ala Asn Met Asp Asp Thr Pro Phe Cys Leu Thr Ser Ser Ala Thr Val  
115 120 125

Gly Tyr Lys Gly Pro Gly Ser Pro Leu Pro Trp Thr Gly Leu Pro Gly  
130 135 140

Ala Val Lys Tyr Tyr Ser Cys Gly His Phe Gly Cys Trp Ala Val Asn  
145 150 155 160

Lys Asn Asp Asp Ile Phe Leu Met Ser Leu Asn Gln Asp Cys Gln Asn  
165 170 175

Asn Gly Trp Ser His Ile Asp Gly Lys Leu Ser Met Ile Glu Val Ala  
180 185 190

Thr Asp Gly Ser Val Phe Gly Val Asn Ser Ala Gly Ser Val Tyr Thr  
195 200 205

Arg Asp Gly Ile Thr Ala Ser Lys Pro Glu Gly Thr Gly Trp Ser Asn  
210 215 220

Ile Pro Met Gly Met Leu Met Gly His Val Thr Tyr Asp Leu Gly Arg  
225 230 235 240

Leu Trp Val Val Ser Lys Ser Gly Gly Thr Met Val Cys Thr His  
245 250 255

5

<210> 8  
<211> 255  
<212> PRT  
<213> *Salmo salar*

<400> 8

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

<210> 9  
 <211> 1214  
 <212> ADN

<213> *Homo sapiens*

<220>

<221> misc\_feature

<222> (1214)..(1214)

<223> n e s a, c, g o t

<400> 9

atgagagcca ctgcagccgt cctattggtc ctctgtctcc tgaccatcag tcatgcatgg	60
gactgtcagg aggtagtaaa catcaagaat ctgatgcaga togatgcagg actggggcaa	120
gtgggttgcta cggacacagg tgaatcccc tactacctgg taggtgataa atggatccgt	180
ctgcctgggtt ccctgaagca tgtcactgta ggaccagcag ggatctgggg tgtcaacaag	240
gactatgcaa tctacaagta tgtggccgggt aactgggttc aagctgcagg ccttctgaaa	300
cagttggatg ctggaggtga acagtttatt gtgggggcta acatgaacga tactccatac	360
cgtctgacaa gtagtgccac agttggctac aagggtccag gctcaccctt tccatggaca	420
ggattgccag gagctgtgaa gtactacagc tgcggaccct ttgggtgctg ggcagtcaac	480
aagaatgatg atatctactt aatgagtctg aatcaagact gccaaaacaa ggggtggagt	540
cacattgaag gcaagctttc catgattgag gtggcaactg atggtagtgt ctttggggtc	600
aactctgcgg gtagtgttta taccagagac ggcatcacag ccagtaaacc agagggcacc	660
ggatggagca atatcccaat gggcatgctc atggggccacg tgacctacga cctggggcgt	720
ctttgggtcg tctccaagtc tggcggcacc atgggtgtgca cacattagcc tcttctctgt	780
agctgaaggc cgttcgggat ctgtctaaag ttcacttgcg aactcattga tctctcttc	840
tggaaaagcc tttagttcat tagttcataa aaatccttca ttttaaaacc tattgctcta	900
cctattatatt tcagttcttc aattatctta ttgccattta aaaaaatata aatgaagatg	960
ttatatatttc ttgaccactc cttgattaac acttcaatag atctttgcca tggaggctat	1020
ttaagtgtag tgtaaaactag ggcatgggtca tggtgctcac aatccacatg ggttttgctg	1080
tgcttcagag gtcacataa ggattggatg gaatccttgt cattgtttat tatctcatta	1140
tataaacatt tcctgcaaaa ataaagcatt ctttttgaaa ctattgtaaa aaaaaaaaaa	1200
aaaaaaaaaa aaan	1214

<210> 10

<211> 2323

<212> ADN

<213> *Salmo salar*

<400> 10

tgtgcagggc tataaaagcg caaagtcttc caatgggaca attgaagtct ggtgtacaac	60
caaacgtata ctgtagatct acatagacat catgagagcc actgcagccg tcctattggt	120
cctctgtctc ctgaccatca gtcattggtta gttaccatca tctgaaacat gcttgatcaa	180
cttggagttg aagtttttct tgggtatactc tactcatatg tctttgtctc catagcatgg	240
gactgtcagg aggtagtaaa catcaagaat ctgatgcaga tcgatgcagg actggggcaa	300
gtggttgcta cggacacaag tcaaattccc tactacctgg taggtgataa atggatccgt	360
ctgcctggtt ccctgaagca tatcaactgta ggaccagcag ggatctgggg tgtcaacaag	420
gactatgcaa tctacaagta tgtggccggg aactgggttc aagctgcagg taagtggaga	480
gcattactca atatttatcc agaggacacc tgcttattag ctttcctgat accatcaggc	540
tgttgaaaaa aacgattgat gttttaaat gtaacttgta ggtaatttgg cagtactcct	600
tgtttgcttg tctgtctgtc tttgtggtct tggccttctg aaacagttgg atgctggagg	660
tgaacagttt attgtggggg ctaacatgaa cgatactcca tactgtctga caagtagtgc	720
cacagttggc tacaagggtc caggctcacc ccttccatgg acaggattgc caggagctgt	780
gaagtactac agctgaggac cctttgggtg ctgggcagtc aacaagaatg atgatatcta	840
cttaatgagt gtaagatctg ggaaagagtg ggagagctgg agtagagtag tagaggatgg	900
agagtgtcag ttattttaaa actgtttcat attataactg ttgaaattgt cctaaaaccc	960
tgattgtatc attttgtttc cagctgaatc aagactgcc aacaaggggg tggagtcaca	1020
ttgaaggcaa gctttccatg attgaggtgg caactgatgg tagtgtcttt ggggtcaact	1080
ctgcgggtag tgtttatacc caggtaaggt tgctactgaa ctatgtgtat ggtccaccac	1140
ccccccccc ccaacagtat taacttgaaa atgacttgta ataataactt agaataataa	1200
tggatatacc ttttaattata actotgatcc ttacagtaca tgctatgtga atctccttac	1260
acaaaaacta aatattgtag gtacataaat aaaatcagtt aaatataatc agatctaaac	1320
ttataggact tattaagaaa tgtgtagaca gtgtatgata aaatatgtaa aagttggatg	1380
tcctgtaaag ctacagtttg ggataaaaaa caacaacttc ccagacaccc caccacttgt	1440
tctggtaaac agctgaggaa tgtagttaga gaaatgtaac cactctcaca ttcatacatg	1500
gagctacgga tgcaagaca caacaatttt tttattaaaa aaaaaaaat gtttatattt	1560
tcttttaaaag ctaaacattg tttgtttaca ataacattgt ttacaaacaa ttgagtaaaa	1620

gcttacattt tggcttctaa tgtggttgaa taaagctcaa gatgcagaag ttatatcttt 1680  
caaaaatcta tggctatatt taattattaa agtccaaaaa tggatgtact taaaaaaaat 1740  
ggataagctt taaaacatga accccaaccc tttcttcaac acagagacgg catcacagcc 1800  
agtaaaccag agggcaccgg atggagcaat atcccaatgg gcatgctcat gggccacgtg 1860  
acctacgacc tgggccgtct ttgggtcgtc tccaagtctg gcggcaccat ggtgtgcaca 1920  
cattagcctc ttctctgtag ctgaaggccg ttcgggatct gtctaaagtt cacttgcgaa 1980  
ctcattgatc tctctttctg gaaaagcctt tagttcatta gttcataaaa atccttcatt 2040  
ttaaaaccta ttgctctacc tattattttc agttcttcaa ttatcttatt gccatttaaa 2100  
aaaatatcaa tgaagatggt atattttctt gaccactcct tgattaacac ttcaatagat 2160  
ctttgccatg gaggtatatt aagtgtagt taaactaggg cagggtcatg ttgctcacia 2220  
tccacatggg ttttgctgtg cttcagaggt catcaatagg atttgacgga atccttgtca 2280  
ttgtttatta tctcattata taatcatttc ctgcaaaaat aaa 2323

<210> 11  
<211> 765  
<212> ADN  
<213> *Salmo salar*

<400> 11

atgagagcca ctgcagccgt cctattggtc ctctgtctcc tgaccatcag tcatgcatgg 60  
gactgtcagg aggtagtaaa catcaagaat ctgatgcaga tcgatgcagg actggggcaa 120  
gtggttgcta cggacacaag tcaaatcccc tactacctgg taggtgataa atggatccgt 180  
ctgcctgggt cctgaagca tatcactgta ggaccagcag ggatctgggg tgtcaacaag 240  
gactatgcaa tctacaagta tgtggccggt aactgggttc aagctgcagg ccttctgaaa 300  
cagttggatg ctggaggtga acagtattatt gtgggggcta acatgaacga tactccatac 360  
tgtctgacaa gtagtgccac agttggctac aagggtccag gctcaccctc tccatggaca 420  
ggattgccag gagctgtgaa gtactacagc tgccggaccct ttgggtgctg ggcagtcaac 480  
aagaatgatg atatctactt aatgagtctg aatcaagact gccaaaacaa ggggtggagt 540  
cacattgaag gcaagctttc catgattgag gtggcaactg atggtagtgt ctttggggtc 600  
aactctgcgg gtagtgttta taccagagac ggcacacag ccagtaaacc agagggcacc 660  
ggatggagca atatcccaat gggcatgctc atggggccacg tgacctacga cctgggccgt 720  
ctttgggtcg tctccaagtc tggcggcacc atggtgtgca cacat 765

<210> 12  
<211> 683  
<212> ADN  
<213> *Salmo salar*

<400> 12

actggggcaa gtggttgcta cggacacaag tcaaatcccc tactacctgg taggtgataa 60  
atggatccgt ctgcctgggt ccctgaagca tatcactgta ggaccagcag ggatctgggg 120  
tgtcaacaag gactatgcaa tctacaagta tgtggccgggt aactgggttc aagctgcagg 180  
cgttccgaaa cagttggatg ctggaggtaa ccagtttggt gtgggggcta acatggacga 240  
tactccattt tgtctgacaa gtagtgccac agttgggtac aagggtccag gctcaccct 300  
tccatggaca ggattgccag gagctgtgaa gtactacagc tgcggacact ttgggtgctg 360  
ggcagtcaac aagaatgatg atattttctt aatgagtctg aatcaagact gccaaaacaa 420  
cgggtggagt cacattgatg gcaagcttcc catgattgag gtggcaactg atggtagtgt 480  
ctttggggtc aactctgcgg gtagtggtta taccagagac ggcatcacag ccagtaaacc 540  
agagggcacc ggatggagca atatcccaat gggcatgctc atgggccacg tgacctacga 600  
cctgggccgt ctttgggtcg tctccaagtc tggcggcacc atggtgtgca cacattagcc 660  
tcttctctgt agctgaaggc cgt 683

<210> 13  
<211> 488  
<212> ADN  
<213> *Gadus morhua*

<400> 13

caatccccta ctacctggtta ggtgataaat ggatccgtct gcctgggttc ctgaagcata 60  
tactgtagg accagcagg atctgggggtg tcaacaagga ctatgcaatc tacaagtatg 120  
tggccggtaa ctgggttcaa gctgcaggcc ttctgaaaca gttggatgct ggaggtgaac 180  
agtttattgt gggggctaac atgaacgata ctccatactg tctgacaagt agtgccacag 240  
ttggctacaa ggggccaggc tcacccttc catggacagg attgccagga gctgtgaagt 300  
actacagctg cggacccttt ggggtgctggg cagtcaacaa gaatgatgat atctacttaa 360  
tgagtctgaa tcaagactgc caaaacaagg ggtggagtca cattgaaggc aagctttcca 420  
tgattgaggt ggcaactgat ggtagtgtct ttgggggtcaa ctctgcgggt agtggtttata 480  
ccagagac 488

<210> 14  
<211> 626  
<212> ADN  
<213> *Gallus gallus*

<400> 14

caggactggg gcaagtgggtt gctacggaca caagtcaaat cccctactac ctggtaggtg 60  
ataaatggat ccgtctgcct gggtccctga agcatatcac tgtaggacca gcagggatct 120  
ggggtgtcaa caaggactat gcaatctaca agtatgtggc cggttaactgg gttcaagctg 180  
caggccttct gaaacagttg gatgctggag gtgaacagtt tattgtgggg gctaacatga 240  
acgatactcc atactgtctg acaagtagtg ccacagttgg ctacaagggt ccaggctcac 300  
cccttccatg gacaggattg ccaggagctg tgaagtacta cagctgcgga cgctttgggt 360  
gctgggcagt caacaagaat gatgatatct acttaatgag tctgaatcaa gactgccaaa 420  
acaaggggtg gagtacatt gaaggcaagc tttccatgat tgaggtggca actgatggta 480  
gtgtctttgg ggtcaactct gcgggtagtg tttataccag agacggcatc acagccgtaa 540  
accagagggc accggatgga gcaatatccc aatgggcatg ctcatggggc acgtgacct 600  
cgacctgggc cgtctttggg tcgtct 626

5 <210> 15  
<211> 1089  
<212> ADN  
<213> *Danio rerio*

10 <220>  
<221> misc\_feature  
<222> (1)..(2)  
<223> n es a, c, g o t

15 <220>  
<221> misc\_feature  
<222> (7)..(8)  
<223> n es a, c, g o t

20 <220>  
<221> misc\_feature  
<222> (16)..(17)  
<223> n es a, c, g o t

25 <220>  
<221> misc\_feature  
<222> (27)..(30)  
<223> n es a, c, g o t

30 <220>  
<221> misc\_feature  
<222> (54)..(54)  
<223> n es a, c, g o t

35 <220>  
<221> misc\_feature  
<222> (59)..(59)  
<223> n es a, c, g o t

40 <220>  
<221> misc\_feature  
<222> (62)..(62)  
<223> n es a, c, g o t

45 <220>  
<221> misc\_feature  
<222> (158)..(158)



	<223> n es a, c, g o t
	<220>
5	<221> misc_feature
	<222> (180)..(180)
	<223> n es a, c, g o t
	<220>
10	<221> misc_feature
	<222> (191)..(191)
	<223> n es a, c, g o t
	<220>
15	<221> misc_feature
	<222> (221)..(221)
	<223> n es a, c, g o t
	<220>
20	<221> misc_feature
	<222> (280)..(280)
	<223> n es a, c, g o t
	<220>
25	<221> misc_feature
	<222> (1038)..(1038)
	<223> n es a, c, g o t
	<220>
30	<221> misc_feature
	<222> (1078)..(1078)
	<223> n es a, c, g o t
	<400> 15

nnaaaannta aaatanngta ggtacannnn aaatcagtta aatataatca gatntaaant	60
tntaggaacta ttaagaaatg tgtagacagc gtacgataaa atatgtaaaa gttggatgtc	120
ctgtaaagct acagtttggg ataaaaaaca acaacttncc agacacccca ccacttgtn	180
tggtaaacag ntgaggaatg tagttagaga aatgtaacca ntctcacatt catacatgga	240
gttacggatg caaagacaca acaatttttt gtttacattn tttttaacat gttttttaa	300
gcaatacaca ttgtttgttt acaataacat tgtttacaaa caattgagta aaagcttaca	360
ttttggcttc tgtgtggttg aaataaagct caagaggcag aagttatatt cttcaaaaat	420
caatggctat atttaattat taaagttcca aaaaggatgt acttaataaa atggataagc	480
tttaaaacat gaaccccaac cttttcttca acacagagac ggcatcacag ccagtaaacc	540
agagggcacc ggatggagca atatccaat gtgtatgctc atgggccacg tgacctacga	600
cctgggcagt ctttgggtcg tctccaagtc tgccgtcacc atggtgtgca cacattagcc	660
tcttctctgt agctgaaggc cgttcgggat ccgtccaaag ttccctggcg aactcattga	720
tctctcttctc tggaaaagcc tttagttcat aaaaatcctt cattttaaaa cctattgctc	780
tacctattat tttcagttct tcaatgatct tattgacatt taaaaaaat atcattgaag	840
attttatatt ttcttgacaa ctctagatt aacacttcaa tagaccttg ccattggaggc	900
tatttaagtg tagtgtaaac tagggcacgg tcatgttgct cacaatccac atgggttttg	960
ctgtgcttca aaggatcatca ataaatcact agtgcgcccg cctgcaggtc gaccatatgg	1020
gagagctccc aacgcgtngg atgcataagc gaccatatgg gagagctccc aacgcgtngg	1080
atgcataag	1089

<210> 16  
 <211> 258  
 <212> PRT  
 <213> *Homo sapiens*  
 <400> 16

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

His Val Asp Leu Arg Pro Gly Pro Ser Leu Gly Arg Leu Gln Val Cys  
245 250 255

Arg His

5

<210> 17  
<211> 194  
<212> PRT  
<213> *Homo sapiens*  
<400> 17

Gly Ser Leu Lys His Ile Thr Val Gly Pro Ala Gly Ile Trp Gly Val  
1 5 10 15

Asn Lys Asp Tyr Ala Ile Tyr Lys Tyr Val Ala Gly Asn Trp Val Gln  
20 25 30

Ala Ala Gly Leu Leu Lys Gln Leu Asp Ala Gly Gly Glu Gln Phe Ile  
35 40 45

Val Gly Ala Asn Met Asn Asp Thr Pro Tyr Cys Leu Thr Ser Ser Ala  
50 55 60

Thr Val Gly Tyr Lys Gly Pro Gly Ser Pro Leu Pro Trp Thr Gly Leu  
65 70 75 80

Pro Gly Ala Val Lys Tyr Tyr Ser Cys Gly Pro Phe Gly Cys Trp Ala  
85 90 95

Val Asn Lys Asn Asp Asp Ile Tyr Leu Met Ser Leu Asn Gln Asp Cys  
100 105 110

Gln Asn Lys Gly Trp Ser His Ile Glu Gly Lys Leu Ser Met Ile Glu  
115 120 125

Val Ala Thr Asp Gly Ser Val Phe Gly Val Asn Ser Ala Gly Ser Val  
130 135 140

Tyr Thr Arg Asp Gly Ile Thr Ala Ser Lys Pro Glu Gly Thr Gly Trp  
145 150 155 160

Ser Asn Ile Pro Met Gly Leu Leu Met Gly His Val Asp Leu Arg Pro  
165 170 175

Gly Pro Ser Leu Gly Arg Leu Gln Val Cys Arg His Pro Trp Val Cys  
180 185 190

Thr His

<210> 18  
<211> 237  
<212> PRT  
<213> *Homo sapiens*

<400> 18

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

<210> 19  
 <211> 582

<212> ADN  
<213> *Homo sapiens*

<400> 19

5

```

ggttccctga agcatatcac ttaggacca gcagggatct ggggtgtcaa caaggactat      60
gcaatctaca agtatgtggc cggttaactgg gttcaagctg caggccttct gaaacagttg      120
gatgctggag gtgaacagtt tattgtgggg gctaacaatga acgataactcc atactgtctg      180
acaagtagtg ccacagttgg ctacaagggt ccaggctcac cccttccatg gacaggattg      240
ccaggagctg tgaagtacta cagctgcgga ccctttgggt gctgggcagt caacaagaat      300
gatgatatact acttaatgag tctgaatcaa gactgccaaa acaaggggtg gagtcacatt      360
gaaggcaagc tttccatgat tgaggtggca actgatggta gtgtctttgg ggtcaactct      420
gcgggtagtg tttataccag agacggcatc acagccagta aaccagaggg caccggatgg      480
agcaatatcc caatgggcct gctcatgggc cagctggacc tacgacctgg gccgtctttg      540
ggtcgtctcc aagtctgccg tcacccatgg gtgtgcacac at                          582

```

<210> 20  
<211> 237  
<212> PRT  
<213> *Homo sapiens*

10

<400> 20

```

Trp Asp Cys Gln Glu Val Val Asn Ile Lys Asn Leu Met Gln Ile Asp
1          5          10          15

Ala Gly Leu Gly Gln Val Val Ala Thr Asp Thr Ser Gln Ile Pro Tyr
20          25          30

Tyr Leu Val Gly Asp Lys Trp Ile Arg Leu Pro Gly Ser Leu Lys His
35          40          45

Ile Thr Val Gly Pro Ala Gly Ile Trp Gly Val Asn Lys Asp Tyr Ala
50          55          60

Ile Tyr Lys Tyr Val Ala Gly Asn Trp Val Gln Ala Ala Gly Leu Leu
65          70          75          80

Lys Gln Leu Asp Ala Gly Gly Asn Gln Phe Ile Val Gly Ala Asn Met
85          90          95

Asn Asp Thr Pro Tyr Cys Leu Thr Ser Ser Ala Thr Val Gly Tyr Lys
100         105         110

```

15

Gly Pro Gly Ser Pro Leu Pro Trp Thr Gly Leu Pro Gly Ala Val Lys  
115 120 125

Tyr Tyr Ser Cys Gly Pro Phe Gly Cys Trp Ala Val Asn Lys Asn Asp  
130 135 140

Asp Ile Tyr Leu Met Ser Leu Asn Gln Asp Cys Gln Asn Lys Gly Trp  
145 150 155 160

Ser His Ile Glu Gly Lys Leu Ser Met Ile Glu Val Ala Thr Asp Gly  
165 170 175

Ser Val Phe Gly Val Asn Ser Ala Gly Ser Val Tyr Thr Arg Asp Gly  
180 185 190

Ile Thr Ala Ser Lys Pro Glu Gly Thr Gly Trp Ser Asn Ile Pro Met  
195 200 205

Gly Leu Leu Met Gly His Val Asp Leu Arg Pro Gly Pro Ser Leu Gly  
210 215 220

Arg Leu Gln Val Cys Arg His Pro Trp Val Cys Thr His  
225 230 235

<210> 21

<211> 237

<212> PRT

<213> *Homo sapiens*

<400> 21

Trp Asp Cys Gln Glu Val Val Asn Ile Lys Asn Leu Met Gln Ile Asp  
1 5 10 15

Ala Gly Leu Gly Gln Val Val Ala Thr Asp Thr Ser Gln Ile Pro Tyr  
20 25 30

Tyr Leu Val Gly Asp Lys Trp Ile Arg Met Pro Gly Tyr Leu Lys Met  
35 40 45

Ile Thr Val Gly Pro Gly Gly Ile Trp Gly Val Asn Lys Asp Tyr Ala  
50 55 60

Ile Tyr Lys Tyr Val Ala Gly Asn Trp Val Gln Ala Ala Gly Leu Leu  
65 70 75 80

Lys Gln Leu Asp Ala Gly Gly Asn Gln Phe Val Val Gly Ala Asn Met  
85 90 95



Asp	Asp	Thr	Pro	Phe	Cys	Leu	Thr	Ser	Ser	Ala	Thr	Val	Gly	Tyr	Lys
			100					105					110		
Gly	Pro	Gly	Ser	Pro	Leu	Pro	Trp	Thr	Gly	Leu	Pro	Gly	Ala	Val	Lys
		115					120					125			
Tyr	Tyr	Ser	Cys	Gly	His	Phe	Gly	Cys	Trp	Ala	Val	Asn	Lys	Asn	Asp
	130					135					140				
Asp	Ile	Phe	Leu	Met	Ser	Leu	Asn	Gln	Asp	Arg	Gln	Asn	Asn	Gly	Trp
145					150					155					160
Ser	His	Ile	Asp	Gly	Lys	Leu	Ser	Met	Ile	Glu	Val	Ala	Thr	Asp	Gly
				165					170					175	
Ser	Val	Phe	Gly	Val	Asn	Ser	Ala	Gly	Ser	Val	Tyr	Thr	Arg	Asp	Gly
			180					185					190		
Ile	Thr	Ala	Ser	Lys	Pro	Glu	Gly	Thr	Gly	Trp	Ser	Asn	Ile	Pro	Met
	195						200					205			
Gly	Leu	Leu	Met	Gly	His	Val	Asp	Leu	Arg	Pro	Gly	Pro	Ser	Leu	Gly
	210					215					220				
Arg	Leu	Gln	Val	Cys	Arg	His	Pro	Trp	Val	Cys	Thr	His			
225					230					235					

<210> 22  
 <211> 166  
 <212> PRT  
 <213> *Homo sapiens*  
 <400> 22

Trp Asp Cys Gln Glu Val Val Asn Ile Lys Asn Leu Met Gln Ile Asp  
1 5 10 15

Ala Gly Leu Gly Gln Val Val Ala Thr Asp Thr Ser Gln Ile Pro Tyr  
20 25 30

Tyr Leu Val Gly Asp Lys Trp Ile Arg Met Pro Gly Tyr Leu Lys Met  
35 40 45

Ile Thr Val Gly Pro Gly Gly Ile Trp Gly Val Asn Lys Asp Tyr Ala  
50 55 60

Ile Tyr Lys Tyr Val Ala Gly Asn Trp Val Gln Ala Ala Gly Leu Leu  
65 70 75 80

Lys Gln Leu Asp Ala Gly Gly Asn Gln Phe Val Val Gly Ala Asn Met  
85 90 95

Asp Asp Thr Pro Phe Cys Leu Thr Ser Ser Ala Thr Val Gly Tyr Lys  
100 105 110

Gly Pro Gly Ser Pro Leu Pro Trp Thr Gly Leu Pro Gly Ala Val Lys  
115 120 125

Tyr Tyr Ser Cys Gly His Phe Gly Cys Trp Ala Val Asn Lys Asn Asp  
130 135 140

Asp Ile Phe Leu Met Ser Leu Asn Gln Asp Arg Gln Asn Asn Gly Trp  
145 150 155 160

Ser His Ile Asp Gly Lys  
165

<210> 23

<211> 95

<212> PRT

<213> *Homo sapiens*

<400> 23

Trp Asp Cys Gln Glu Val Val Asn Ile Lys Asn Leu Met Gln Ile Asp  
1 5 10 15

Ala Gly Leu Gly Gln Val Val Ala Thr Asp Thr Ser Gln Ile Pro Tyr  
20 25 30

Tyr Leu Val Gly Asp Lys Trp Ile Arg Met Pro Gly Tyr Leu Lys Met  
35 40 45

Ile Thr Val Gly Pro Gly Gly Ile Trp Gly Val Asn Lys Asp Tyr Ala  
50 55 60

Ile Tyr Lys Tyr Val Ala Gly Asn Trp Val Gln Ala Ala Gly Leu Leu  
65 70 75 80

Lys Arg Leu Asp Ala Gly Gly Asp Gln Phe Val Val Gly Leu Thr  
85 90 95

<210> 24

<211> 38

<212> ADN

<213> Secuencia artificial

<220>

<223> NWA15dPET(nº 23): NcoI

<400> 24

gcaccatggc catgggctgg gactgtcagg aggtagta 38

<210> 25

<211> 36

<212> ADN

<213> Secuencia artificial

<220>

<223> CH.A15dPET(nº 13): ACC65I

<400> 25

ccgaagcttg gtaccatgtg tgcacacat ggtgac 36

<210> 26

<211> 17

<212> PRT

<213> Secuencia artificial

<220>

<223> Leucolectina consenso 1

<220>

<221> misc\_feature

<222> (1).. (1)

<223> Xaa puede ser cualquier aminoácido de origen natural

<220>

<221> misc\_feature

<222> (3)..(4)

<223> Xaa puede ser cualquier aminoácido de origen natural

<220>

<221> misc\_feature  
 <222> (8)..(8)  
 <223> Xaa puede ser cualquier aminoácido de origen natural

5

<220>  
 <221> misc\_feature  
 <222> (11)..(13)  
 <223> Xaa puede ser cualquier aminoácido de origen natural

10

<220>  
 <221> misc\_feature  
 <222> (17)..(17)  
 <223> Xaa puede ser cualquier aminoácido de origen natural

15

<400> 26

**Xaa Trp Xaa Xaa Leu Pro Gly Xaa Leu Lys Xaa Xaa Xaa Val Gly Pro**  
**1 5 10 15**

**Xaa**

20

<210> 27  
 <211> 13  
 <212> PRT  
 <213> Secuencia artificial

25

<220>  
 <223> Leucolectina consenso 2

30

<220>  
 <221> misc\_feature  
 <222> (7)..(8)  
 <223> Xaa puede ser cualquier aminoácido de origen natural

35

<220>  
 <221> misc\_feature  
 <222> (10)..(10)  
 <223> Xaa puede ser cualquier aminoácido de origen natural

<400> 27

**Gly Val Asn Lys Asn Asp Xaa Xaa Tyr Xaa Leu Val Gly**  
**1 5 10**

40

<210> 28  
 <211> 23  
 <212> ADN  
 <213> Secuencia artificial

45

<220>  
 <223> Pez cebra LLF1

50

<400> 28  
 atgcagatcg atgcaggact ggg 23

55

<210> 29  
 <211> 24  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> Pez cebra LLF2

	<400> 29 tgggtccctg aagcatgtca ctgt	24
5	<210> 30 <211> 24 <212> ADN <213> Secuencia artificial	
10	<220> <223> Pez cebra LLR1	
15	<400> 30 gaaagagaga tcaatgagtt cgca	24
20	<210> 31 <211> 25 <212> ADN <213> Secuencia artificial	
25	<220> <223> Pez cebra LLR2	
25	<400> 31 caaagacact accatcagtt gccac	25
30	<210> 32 <211> 24 <212> ADN <213> Secuencia artificial	
35	<220> <223> Pez cebra LLR3	
35	<400> 32 gtccgcagct gtagtacttc acag	24
40	<210> 33 <211> 23 <212> ADN <213> Secuencia artificial	
45	<220> <223> cebador 5' GeneRacer	
50	<400> 33 cgactggagc acgaggacac tga	23
50	<210> 34 <211> 26 <212> ADN <213> Secuencia artificial	
55	<220> <223> cebador anidado 5' GeneRacer	
60	<400> 34 ggacactgac atggactgaa ggagta	26
60	<210> 35 <211> 25 <212> ADN <213> Secuencia artificial	
65	<220> <223> Pez cebra LLR4	

# ES 2 546 310 T3

	<400> 35 agcctggacc ctgtagcca actgt	25
5	<210> 36 <211> 25 <212> ADN <213> Secuencia artificial	
10	<220> <223> cebador 3' GeneRacer	
15	<400> 36 gctgtcaacg atacgtacg taacg	25
20	<210> 37 <211> 23 <212> ADN <213> Secuencia artificial	
	<220> <223> cebador anidado 3' GeneRacer	
25	<400> 37 cgctacgtaa cggcatgaca gtg	23
30	<210> 38 <211> 27 <212> ADN <213> Secuencia artificial	
	<220> <223> Pez cebra LLF4	
35	<400> 38 gtgaggtggc aactgatggt agtgtct	27
40	<210> 39 <211> 29 <212> ADN <213> Secuencia artificial	
45	<220> <223> LL/F de salmón	
	<400> 39 tacggacaca ggtcgaatcc cctactacc	29
50	<210> 40 <211> 25 <212> ADN <213> Secuencia artificial	
55	<220> <223> LL/R de salmón	
	<400> 40 acagagaaga ggctaattgtg tgcac	25
60	<210> 41 <211> 22 <212> PRT <213> <i>Salmo salar</i>	
65	<400> 41	

Trp Asp Cys Gln Glu Val Val Asn Ile Lys Asn Leu Met Gln Ile Asp  
1 5 10 15

Ala Gly Leu Gly Gln Val  
20

<210> 42  
<211> 18  
<212> PRT  
<213> *Salmo salar*

<400> 42

Met Arg Ala Thr Ala Val Leu Leu Val Leu Cys Leu Leu Thr Ile Ser

1 5 10 15

His Ala

<210> 43  
<211> 78  
<212> PRT  
<213> *Salmo salar*

<400> 43

Trp Asp Cys Gln Glu Val Val Asn Ile Lys Asn Leu Met Gln Ile Asp  
1 5 10 15

Ala Gly Leu Gly Gln Val Val Ala Thr Asp Thr Ser Gln Ile Pro Tyr  
20 25 30

Tyr Leu Val Gly Asp Lys Trp Ile Arg Leu Pro Gly Ser Leu Lys His  
35 40 45

Ile Thr Val Gly Pro Ala Gly Ile Trp Gly Val Asn Lys Asp Tyr Ala  
50 55 60

Ile Tyr Lys Tyr Val Ala Gly Asn Trp Val Gln Ala Ala Gly  
65 70 75

<210> 44  
<211> 73  
<212> PRT  
<213> *Salmo salar*

<400> 44

# ES 2 546 310 T3

Gly Leu Leu Lys Gln Leu Asp Ala Gly Gly Glu Gln Phe Ile Val Gly  
1 5 10 15

Ala Asn Met Asn Asp Thr Pro Tyr Cys Leu Thr Ser Ser Ala Thr Val  
20 25 30

Gly Tyr Lys Gly Pro Gly Ser Pro Leu Pro Trp Thr Gly Leu Pro Gly  
35 40 45

Ala Val Lys Tyr Tyr Ser Cys Gly Pro Phe Gly Cys Trp Ala Val Asn  
50 55 60

Lys Asn Asp Asp Ile Tyr Leu Met Ser  
65 70

<210> 45  
<211> 39  
<212> PRT  
<213> *Salmo salar*

<400> 45

Leu Asn Gln Asp Cys Gln Asn Lys Gly Trp Ser His Ile Glu Gly Lys  
1 5 10 15

Leu Ser Met Ile Glu Val Ala Thr Asp Gly Ser Val Phe Gly Val Asn  
20 25 30

Ser Ala Gly Ser Val Tyr Thr  
35

<210> 46  
<211> 47  
<212> PRT  
<213> *Salmo salar*

<400> 46

Arg Asp Gly Ile Thr Ala Ser Lys Pro Glu Gly Thr Gly Trp Ser Asn  
1 5 10 15

Ile Pro Met Gly Met Leu Met Gly His Val Thr Tyr Asp Leu Gly Arg  
20 25 30

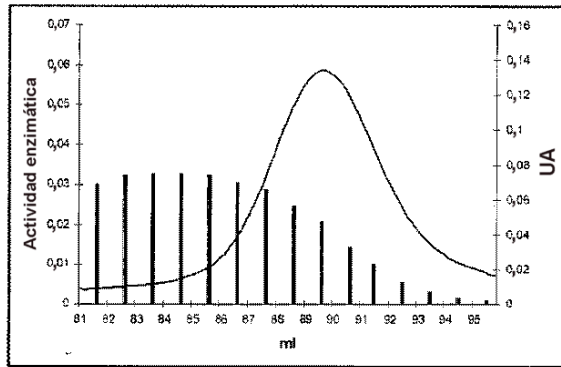
Leu Trp Val Val Ser Lys Ser Ala Val Thr Met Val Cys Thr His  
35 40 45



## REIVINDICACIONES

1. Una composición farmacéutica que comprende (i) un polipéptido capaz de tratar lesiones de la piel en un animal, que comprende una secuencia de aminoácidos como la indicada en una cualquiera de las secuencias nº 16 a 18, o una secuencia funcionalmente equivalente que tiene una identidad de al menos un 95 % con dicha secuencia, donde dicha comparación se realiza a lo largo de toda la longitud de la secuencia; o (ii) una molécula de ácido nucleico que comprende una secuencia de nucleótidos como la indicada en la secuencia nº 19; y uno o más excipientes y/o diluyentes farmacéuticamente aceptables para su uso en terapia.
2. El uso de un polipéptido, molécula de ácido nucleico o composición farmacéutica como se describe en la reivindicación 1, en la preparación de un medicamento para tratar o prevenir un trastorno autoinmunitario, un trastorno inflamatorio o una lesión de la piel en un animal.
3. Un polipéptido, molécula de ácido nucleico o composición farmacéutica como se describe en la reivindicación 1, para su uso en el tratamiento o prevención de un trastorno autoinmunitario, un trastorno inflamatorio o una lesión de la piel en un animal.
4. Un uso de acuerdo con la reivindicación 2, en el que dichos trastornos a tratar o prevenir son trastornos autoinmunitarios de la piel, trastornos inflamatorios de la piel o afecciones de las membranas mucosas o lesiones de la piel.
5. Un polipéptido, molécula de ácido nucleico o composición farmacéutica para su uso de acuerdo con la reivindicación 3, en el que dichos trastornos a tratar o prevenir son trastornos autoinmunitarios de la piel, trastornos inflamatorios de la piel o afecciones de las membranas mucosas o lesiones de la piel.
6. Un uso de acuerdo con la reivindicación 4, donde dicho polipéptido, molécula de ácido nucleico o composición es para administración tópica u oral.
7. Un polipéptido, molécula de ácido nucleico o composición farmacéutica para su uso de acuerdo con la reivindicación 5, donde dicho polipéptido, molécula de ácido nucleico o composición es para administración tópica u oral.
8. Un uso de acuerdo con la reivindicación 4 o 6, donde dichos trastornos a tratar o prevenir son llagas, úlceras, eccema, acné, psoriasis, inflamación gastrointestinal, gingivitis, inflamación de la cavidad oral y el esófago o lesiones de la piel.
9. Un polipéptido, molécula de ácido nucleico o composición farmacéutica para su uso de acuerdo con la reivindicación 5 o 7, donde dichos trastornos a tratar o prevenir son llagas, úlceras, eccema, acné, psoriasis, inflamación gastrointestinal, gingivitis, inflamación de la cavidad oral y el esófago o lesiones de la piel.
10. Un uso de acuerdo con la reivindicación 8, en el que (i) dicha inflamación gastrointestinal es enfermedad de Crohn, colitis ulcerosa u otra inflamación crónica; o (ii) dicha piel lesionada está irritada o inflamada, agrietada por el frío, quemada por el sol, irradiada, dañada por calor o es una herida.
11. Un polipéptido, molécula de ácido nucleico o composición farmacéutica para su uso de acuerdo con la reivindicación 9, donde (i) dicha inflamación gastrointestinal es enfermedad de Crohn, colitis ulcerosa u otra inflamación crónica; o (ii) dicha piel lesionada está irritada o inflamada, agrietada por el frío, quemada por el sol, irradiada, dañada por calor o es una herida.
12. Un producto que contiene uno o más polipéptidos o moléculas de ácido nucleico como se definen en la reivindicación 1, y uno o más ingredientes activos adicionales como una preparación combinada para el uso simultáneo, separado o secuencial en terapia humana o animal.
13. Un uso de acuerdo con una cualquiera de las reivindicaciones 2, 4, 6, 8 o 10, donde dicho animal es un mamífero.
14. Un polipéptido, molécula de ácido nucleico, composición farmacéutica o producto para su uso de acuerdo con una cualquiera de las reivindicaciones 3, 5, 7, 9, 11 o 12, donde dicho animal es un mamífero.
15. Un polipéptido, molécula de ácido nucleico, composición farmacéutica o producto para su uso de acuerdo con la reivindicación 14, donde dicho mamífero es un ser humano.

A)



B)

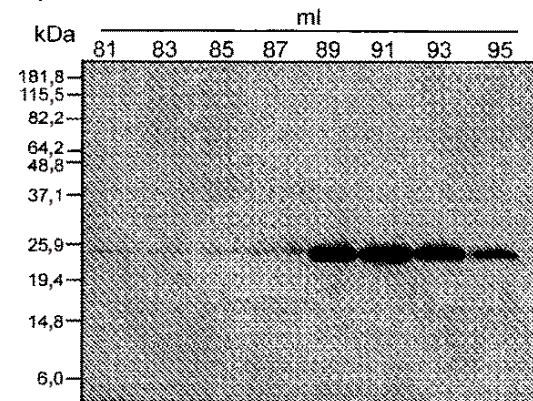


Figura 1

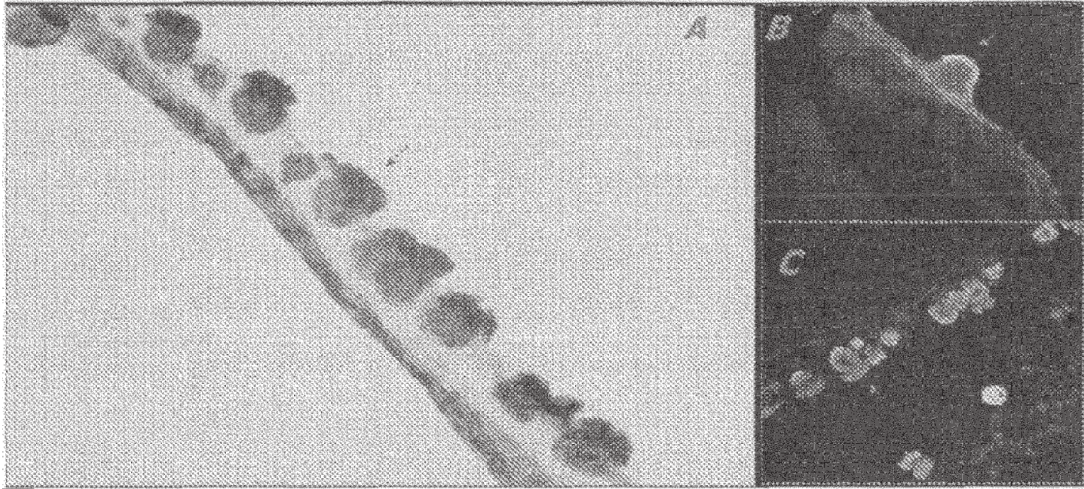
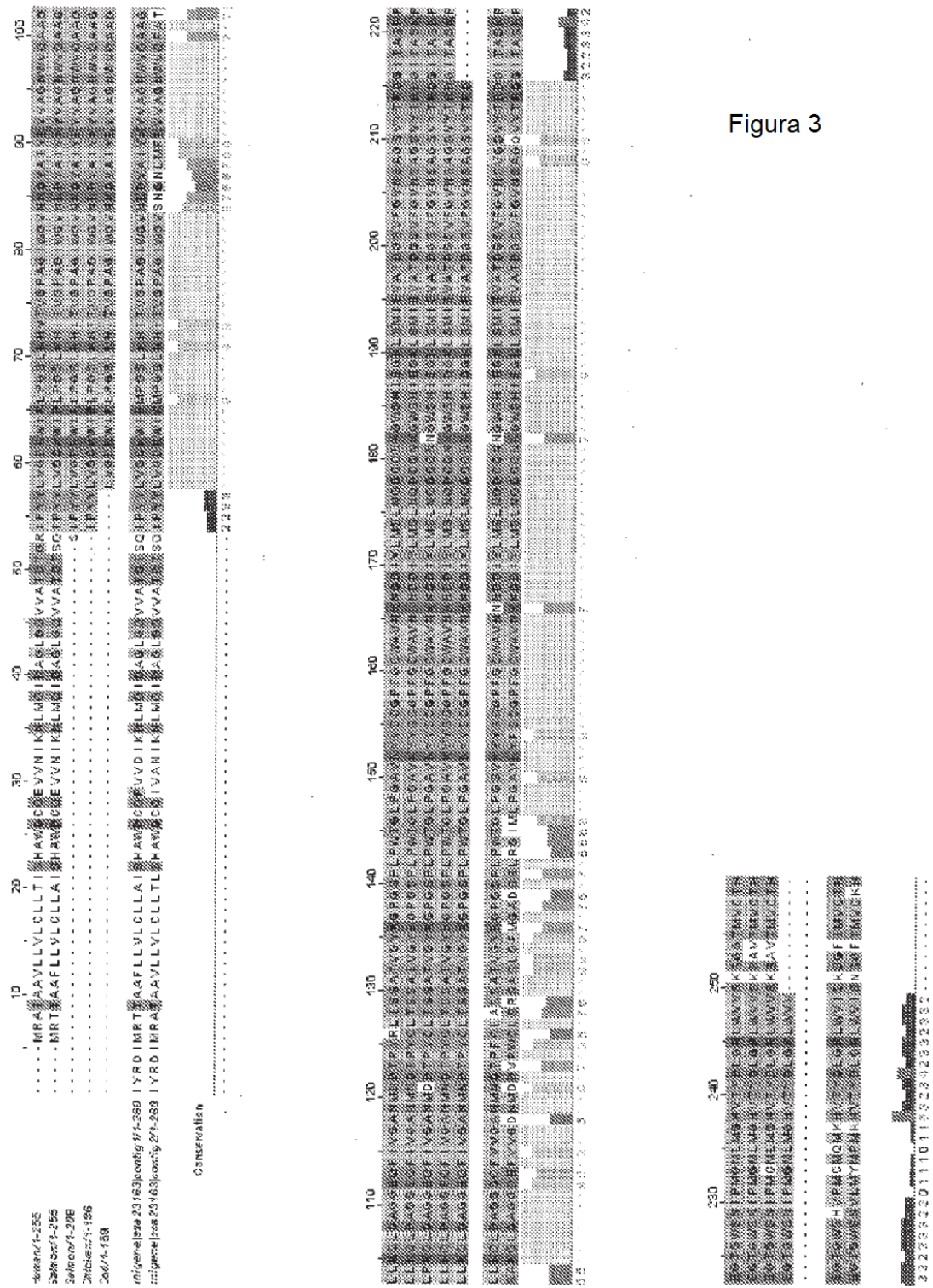
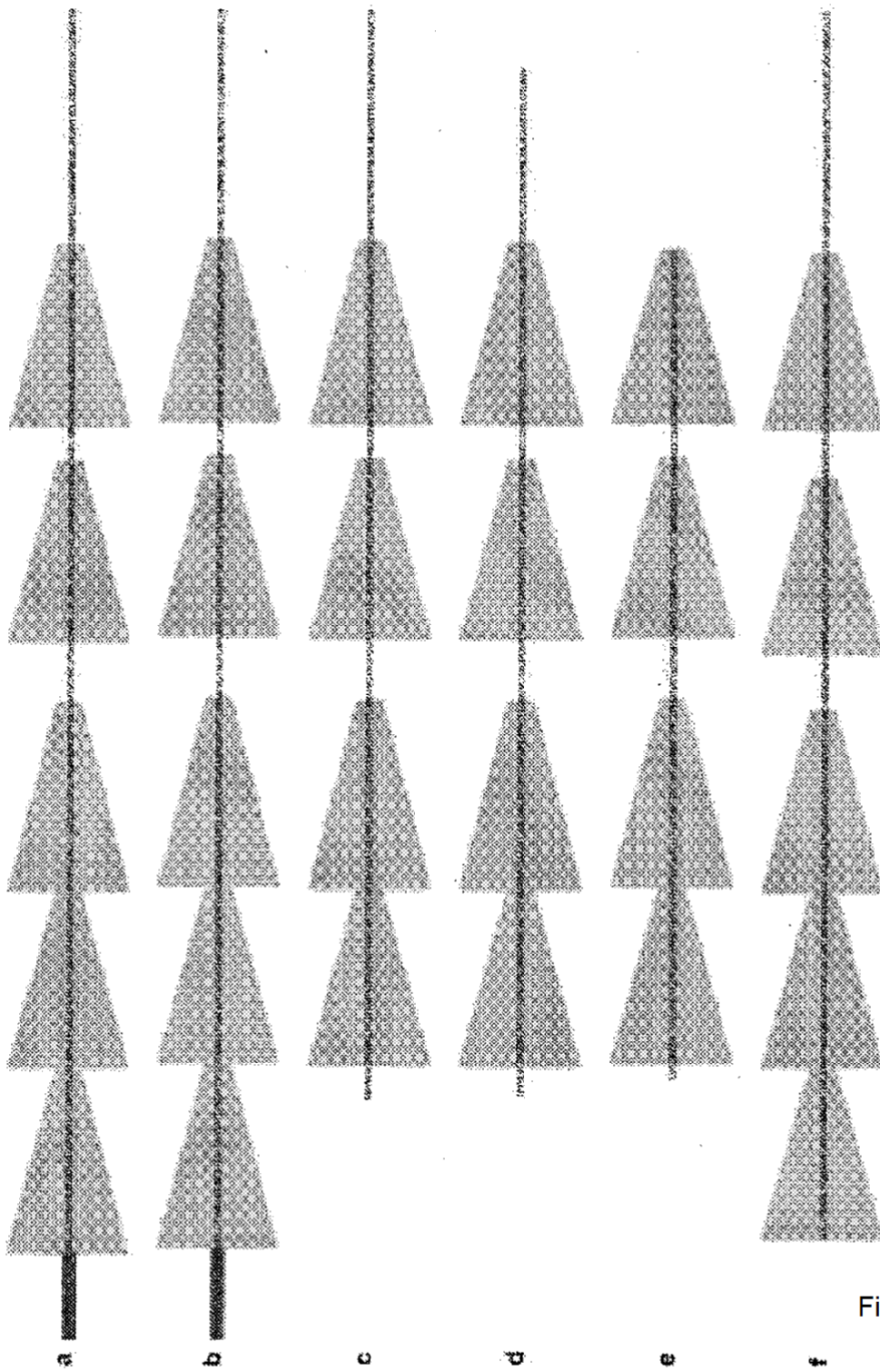


Figura 2







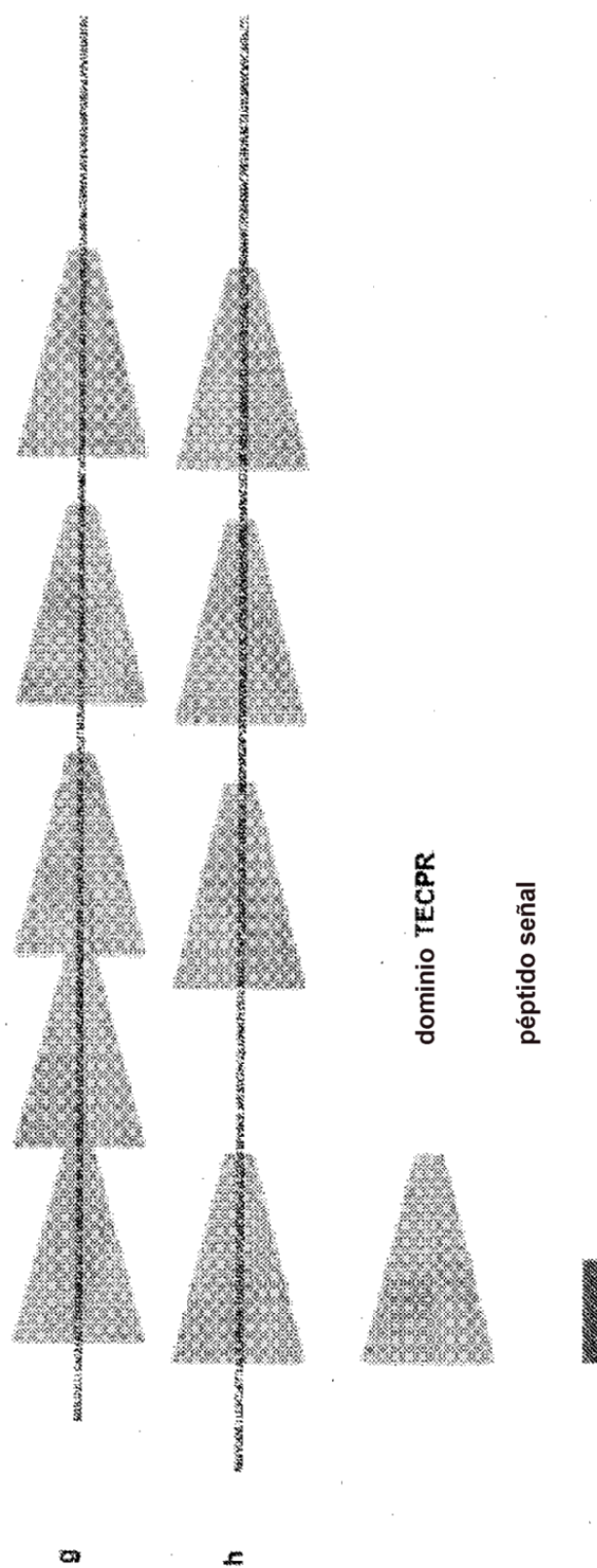


Figura 4A  
cont.

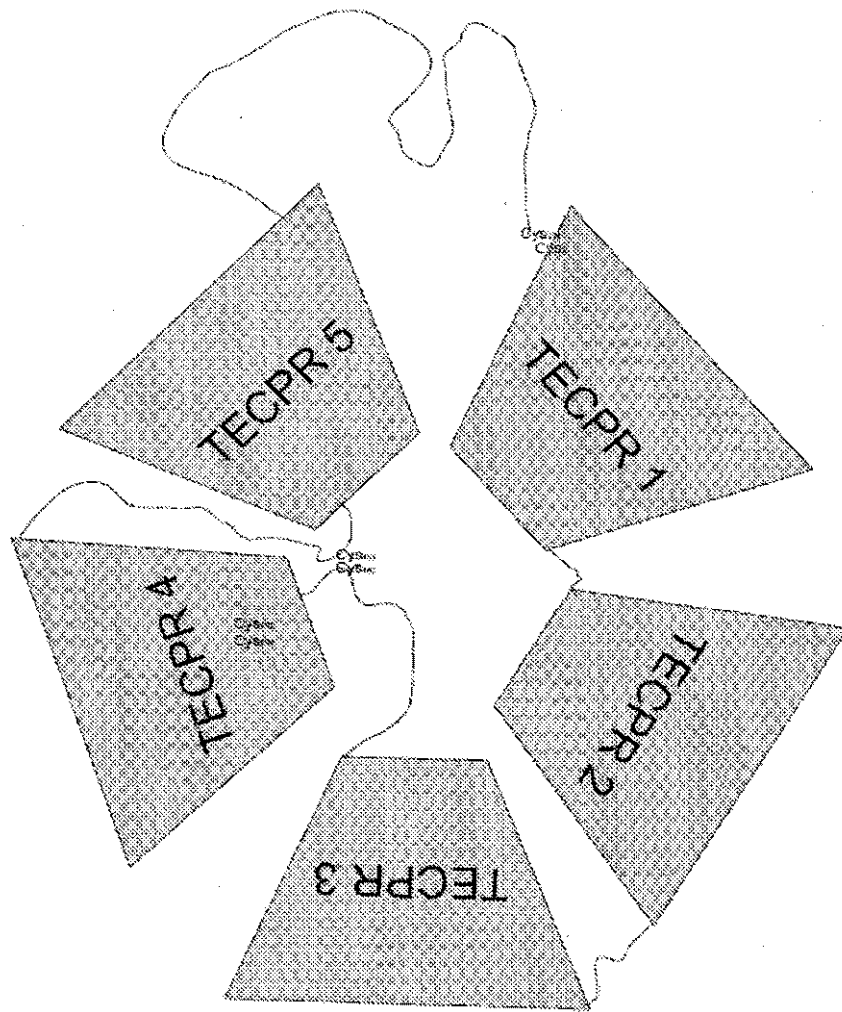


Figura 4B

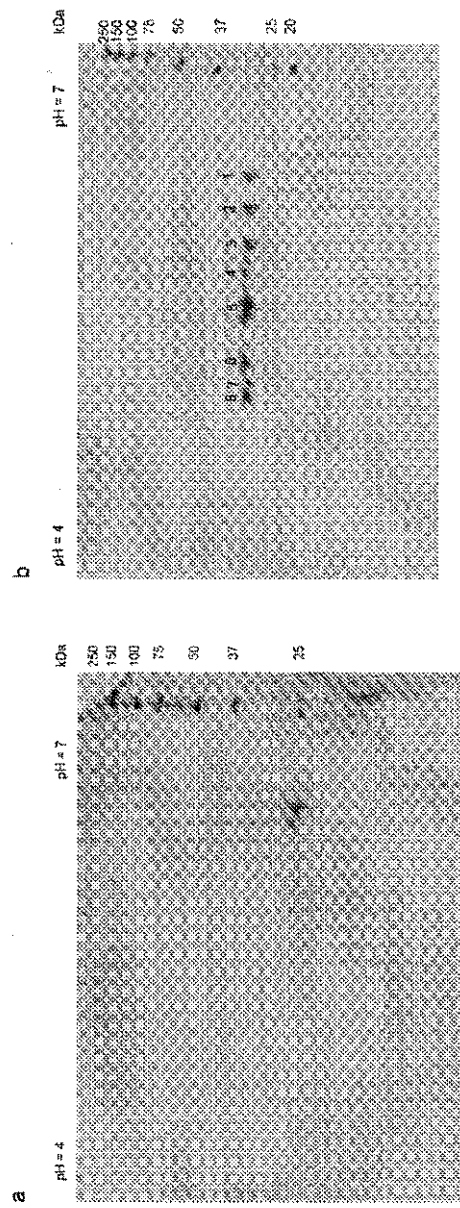


Figura 5



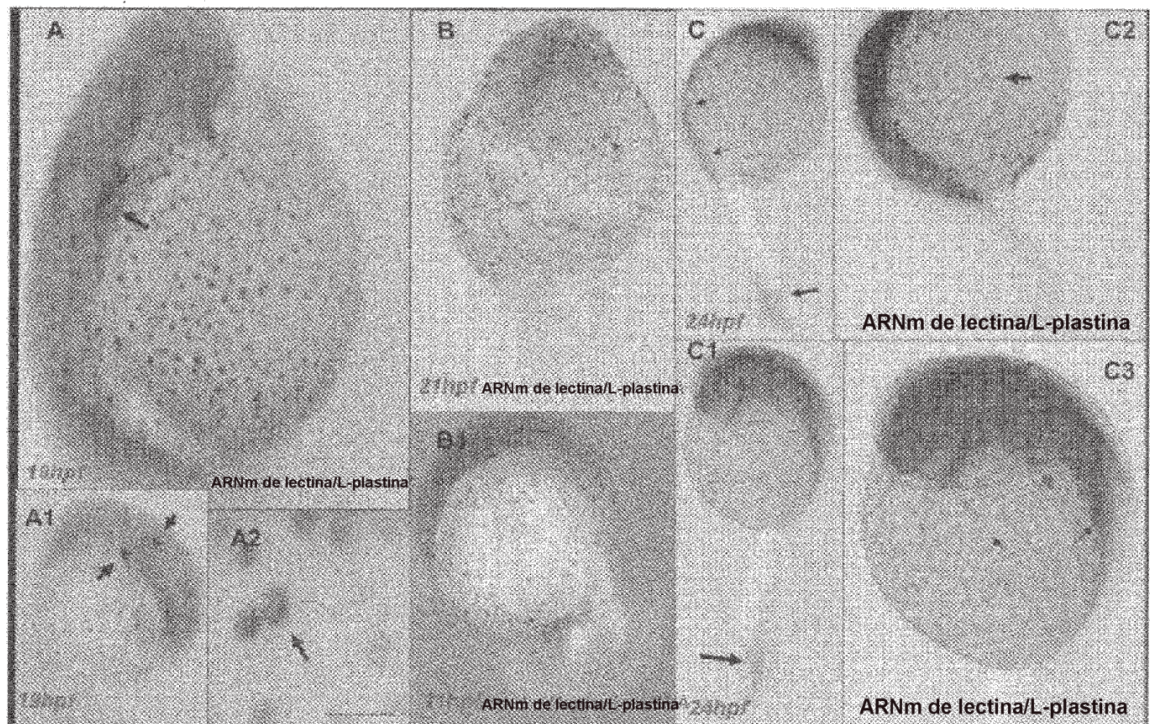


Figura 6A



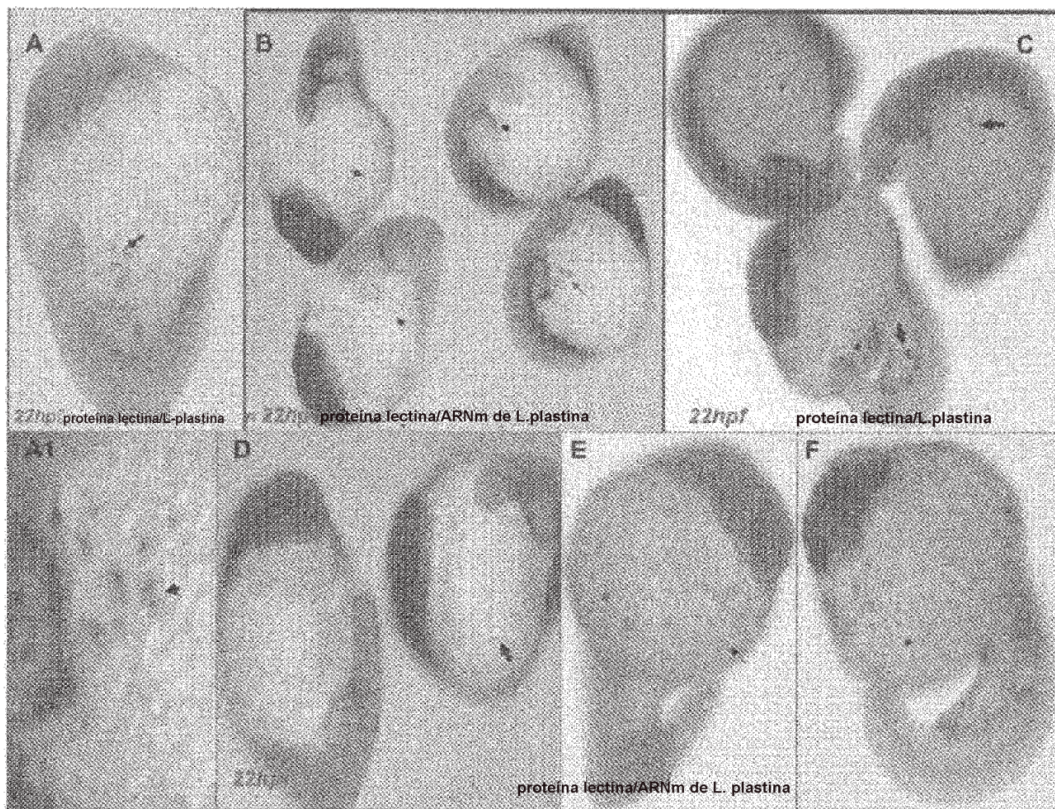


Figura 6B

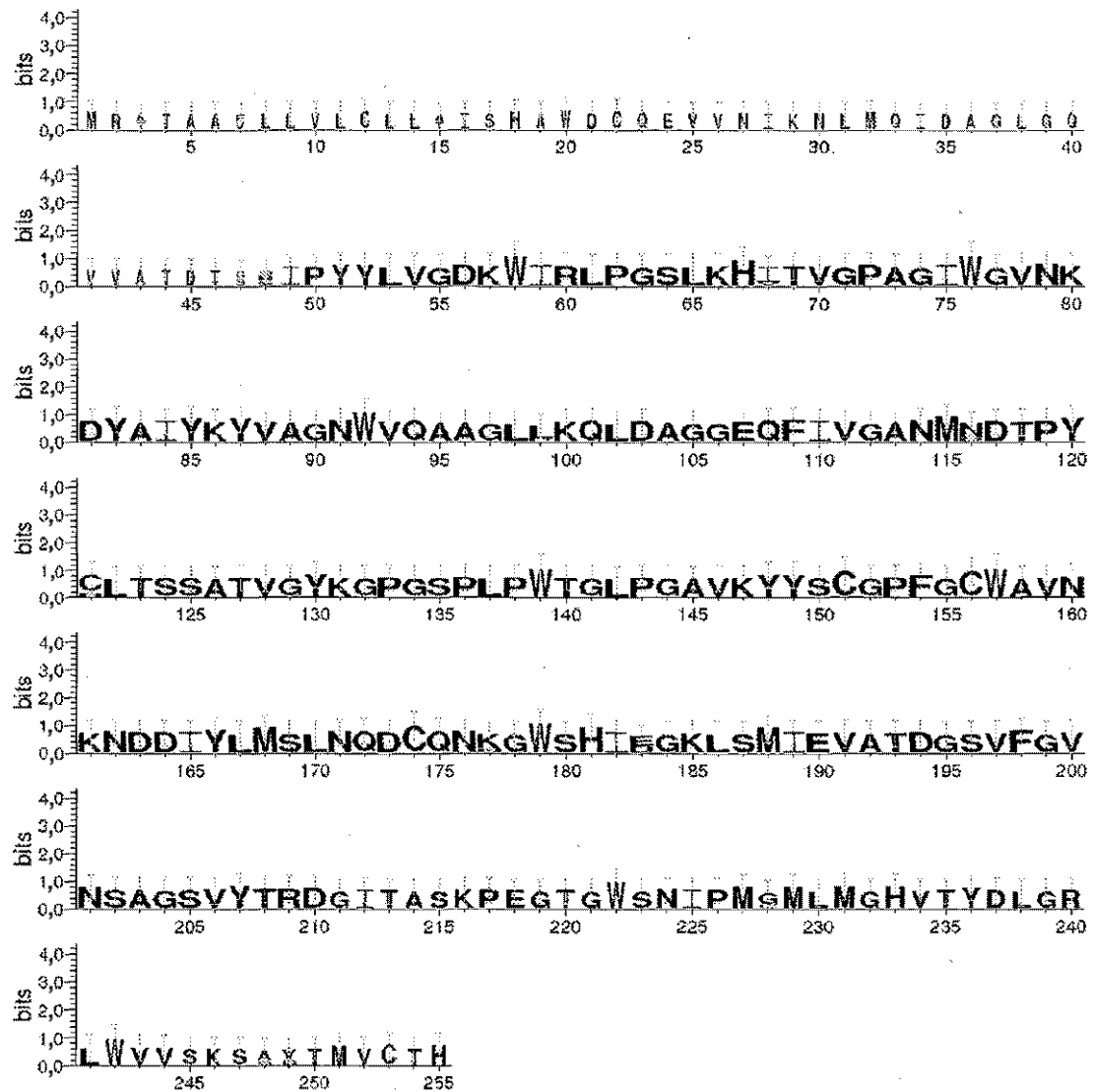


Figura 7

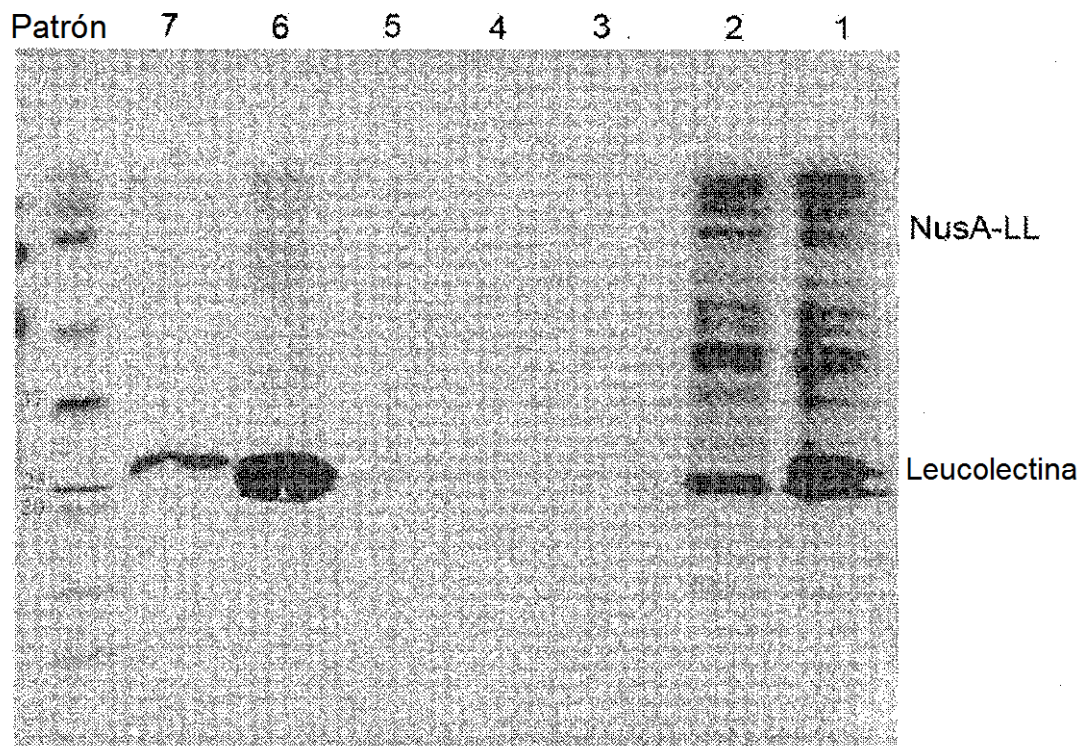


Figura 8



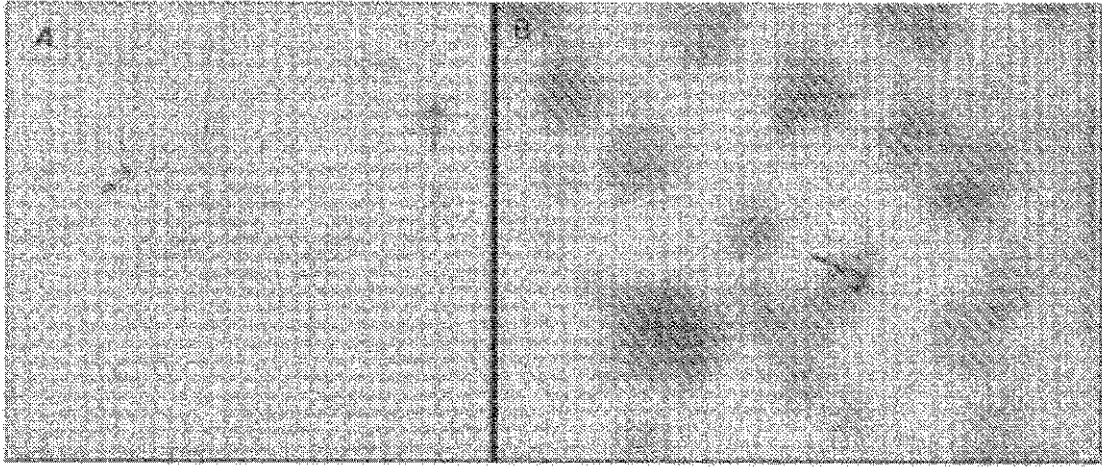
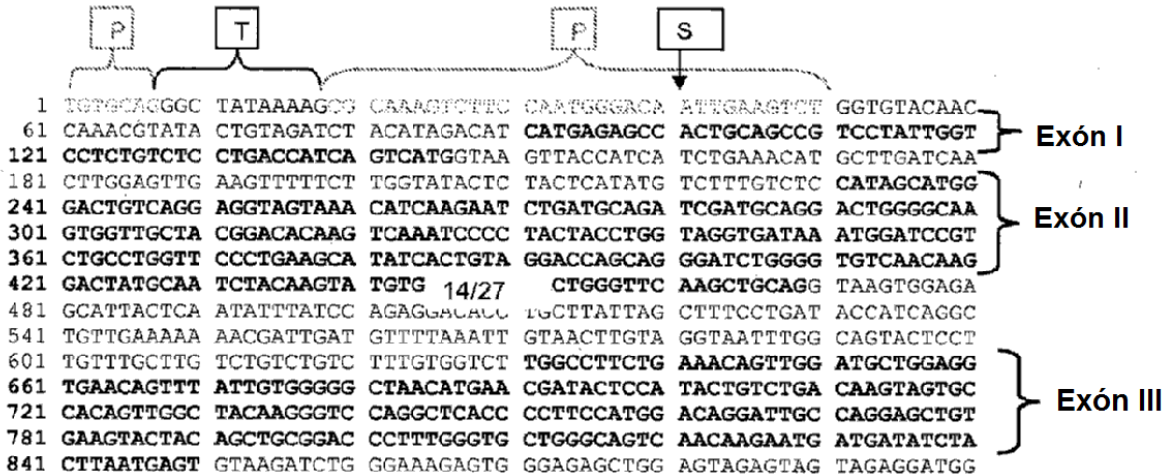


Figura 9A

TGTGCAGGGCTATAAAAGCGCAAAGTCTTCCATGGGACAATTGAAGTCTGGTGTACAAC  
CAAACGTATACTGTAGATCTACATAGACATCATGAGAGCCACTGCAGCCGTCCTATTGGT  
CCTCTGTCTCCTGACCATCAGTCATGGTAAAGTTACCATCACTGAAACATGCTTGATCAA  
CTTGGAGTTGAAGTTTTCTTGGTATACTCTACTCATATGCTTTTGTCTCCATAGCATGG  
GACTGTCAGGAGGTAGTAAACATCAAGAATCTGATGCAGATCGATGCAGGACTGGGGCAA  
GTGGTTGCTACGGACACAAGTCAAATCCCCCTACTACCTGGTAGGTGATAAATGGATCCGT  
CTGCCTGGTTCCTGAAGCATATCACTGTAGGACAGCAGGGATCTGGGGTGTCAACAAG  
GACTATGCAATCTACAAGTATGTGGCCGGTAACTGGGTTCAAGCTGCAGGTAAGTGGAGA  
GCATTACTCAATATTTATCCAGAGGACACCTGCTTATTAGCTTTCCTGATACCATCAGGC  
TGTTGAAAAAACGATTGATGTTTTAAATTGTAAGTCTGTAGGTAATTTGGCAGTACTCCT  
TGTTTGTCTGTCTGTCTTTTGGTCTTGGCCTTCTGAAACAGTTGGATGCTGGAGG  
TGAACAGTTTATTGTGGGGGCTAACATGAACGATACTCCATACTGTCTGACAAGTAGTGC  
CACAGTTGGCTACAAGGGTCCAGGCTCACCCTTCCATGGACAGGATTGCCAGGAGCTGT  
GAAGTACTACAGCTCGGGACCCCTTGGGTGCTGGGCAGTCAACAAGAATGATGATATCTA  
CTTAATGAGTGTAAGATCTGGGAAAGAGTGGGAGAGCTGGAGTAGAGTAGTAGAGGATGG  
AGAGTGTGAGTTATTTAAACTGTTTCATATTATAACTGTTGAAATTGTCTTAAACCC  
TGATTGTATCATTTTGTTCAGCTGAATCAAGACTGCCAAAACAAGGGGTGGAGTCACA  
TTGAAGGCAAGCTTTCATGATTGAGGTGGCAACTGATGGTAGTGTCTTTGGGGTCAACT  
TGCGGGTAGTGTGTTTATACCCAGGTAAGGTTGCTACTGAACATATGTGATGCTCCACCAC  
CCCCCCCCCCCCAACAGTATTAACTTGAAAATGACTTGTAATAATAACTTAGAATAATAA  
TGGTATACCCCTTAAATTATAACTCTGATCCTTACAGTACATGCTATGTGAATCTCCTTAC  
ACAAAACTAAATATTGTAGGTACATAAATAAATCAGTTAAATATAATCAGATCTAAAC  
TTATAGGACTTATTAAAGAAATGTGTAGACAGTGTATGATAAATAATGTAAGTTGGATG  
TCCTGTAAAGCTACAGTTTGGGATAAAAAACAACAACCTTCCAGACACCCCACTTGT  
TCTGGTAAACAGCTGAGGAATGTAGTTAGAGAAATGTAACCACTCTCACATTCATACATG  
GAGCTACGGATGCAAGACACAACAATTTTTTATTAAAAAATAAATGTTTATATTT  
TCTTTTAAAGCTAAACATTGTTTGTGTTACAATAACATTGTTTACAAACAATTGAGTAAAA  
GCTTACATTTTGGCTTCTAATGTGGTTGAATAAGCTCAAGATGCAGAAGTTATATTCTT  
CAAAAATCTATGGCTATATTTAATTATTAAGTCCAAAAATGGATGTACTTAAAAAATAA  
GGATAAGCTTTAAACATGAACCCCAACCCCTTCTTCAACACAGAGACGGCATCACAGCC  
AGTAAACCAGAGGGGACCGGATGGAGCAATATCCCAATGGGCATGCTCATGGGCCACGTG  
ACCTACGACCTGGGCCGCTTTGGGTGCTCTCCAAGTCTGGCGGCACCATGGTGTGCACA  
CATTAGCCTCTTCTCTGTAGCTGAAGGCCGTTCCGGATCTGTCTAAAGTTCACCTGCGAA  
CTCATTGATCTCTCTTCTGGAAAAGCCTTTAGTTCATTAGTTCATAAAAAATCCTTCATT  
TTAAACCTATTGCTCTACCTATTATTTTCAGTCTTCAATTATCTTATTGCCATTTAAA  
AAAATATCAATGAAGATGTTATATTTCTTGACCACTCCTTGATTAACACTTCAATAGAT  
CTTTGCCATGGAGGCTATTTAAGTGTAGTGTAACTAGGGCAGGTCATGTTGCTCACAA  
TCCACATGGGTTTTGCTGTGCTTCAGAGGTCATCAATAGGATTTGACGGAATCCTTGTCA  
TTGTTTATTATCTCATTATATAATCATTTCTGCAAAAAATAAA

2323 pb



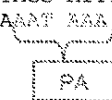
901	AGAGTGTCTAG	TTATTTTAAA	ACTGTTTCAT	ATTATAACTG	TTGAAATTGT	CCTAAAACCC	} Exón IV
961	TGATTGTATC	ATTTTGTTC	CAGCTGAATC	AAGACTGCCA	AAACAAGGGG	TGGAGTCACA	
1021	<b>TTGAAGGCAA</b>	<b>GCTTTCCATG</b>	<b>ATTGAGGTGG</b>	<b>CAACTGATGG</b>	<b>TAGTGTCTTT</b>	<b>GGGGTCAACT</b>	
1081	<b>CTGCGGGTAG</b>	<b>TGTTTATACC</b>	CAGGTAAGGT	TGCTACTGAA	CTATGTGTAT	GGTCCACCAC	
1141	CCCCCCCCC	CCAACAGTAT	TAACCTGAAA	ATGACTTGTA	ATAATAACTT	AGAATAATAA	
1201	TGGTATACCC	TTTAATTATA	ACTCTGATCC	TTACAGTACA	TGCTATGTGA	ATCTCCTTAC	
1261	ACAAAACTA	AATATTGTAG	GTACATAAAT	AAAATCAGTT	AAATATAATC	AGATCTAAAC	
1321	TTATAGGACT	TATTAAGAAA	TGTGTAGACA	GTGTATGATA	AAATATGTAA	AAGTTGGATG	
1381	TCCTGTAAAG	CTACAGTTTG	GGATAAAAAA	CAACAACCTC	CCAGACACCC	CACCACTTGT	
1441	TCGGGTAAAC	AGCTGAGGAA	TGTAGTTAGA	GAAATGTAA	CACTCTCACA	TTCATACATG	
1501	GAGCTACGGA	TGCAAAGACA	CAACAATTTT	TTTATTAAAA	AAAAAAAAT	GTTTATATTT	
1561	TCTTTTAAAG	CTAAACATTG	TTTGTTTACA	ATAACATTGT	TTACAAACAA	TTGAGTAAAA	
1621	GCTTACATTT	TGGCTTCTAA	TGTGGTTGAA	TAAAGCTCAA	GATGCAGAAG	TTATATTCTT	
1681	CAAAAATCTA	TGGCTATATT	TAATTATTAA	AGTCCAAAAA	TGGATGTACT	TAAAAAAAAT	
1741	GGATAAGCTT	TAAACATGA	ACCCCAACCC	TTTCTTCAAC	<b>ACAGAGACGG</b>	<b>CATCACAGCC</b>	} Exón V
1801	<b>AGTAAACCAG</b>	<b>AGGGCACCGG</b>	<b>ATGGAGCAAT</b>	<b>ATCCCAATGG</b>	<b>GCAATGCTCAT</b>	<b>GGGCCACGTG</b>	
1861	<b>ACCTACGACC</b>	<b>TGGGCCGTCT</b>	<b>TTGGGTCGTC</b>	<b>TCCAAGTCTG</b>	<b>GCGGCACCAT</b>	<b>GGTGTGCACA</b>	
1921	<b>CATTAGCCTC</b>	<b>TTCTCTGTAG</b>	<b>CTGAAGGCCG</b>	<b>TTCCGGGATCT</b>	<b>GTCTAAAGTT</b>	<b>CACCTGCGAA</b>	
1981	CTCATTGATC	TCTCTTTCTG	GAAAAGCCTT	TAGTTTATTA	GTTTATAAAA	ATCCTTCATT	
2041	TTAAAACCTA	TTGCTCTACC	TATTAATTTT	AGTTCTTCAA	TTATCTTATT	GCCATTTAAA	
2101	AAAAATATCA	TGAAGATGTT	ATATTTTCTT	GACCACTCCT	TGATTAACAC	TTCAATAGAT	
2161	CTTTGCCATG	GAGGCTATTT	AAGTGTAGTG	TAACTAGGG	CACGGTCATG	TTGCTCACAA	
2221	TCCACATGGG	TTTGCTGTG	CTTCAGAGGT	CATCAATAGG	ATTTGACGGA	ATCCTTGTC	
2281	TTGTTTATTA	TCTCATTATA	TAATCATTTT	CTGCAAAAT	AAA		
//							
							

Figura 9B  
continuación

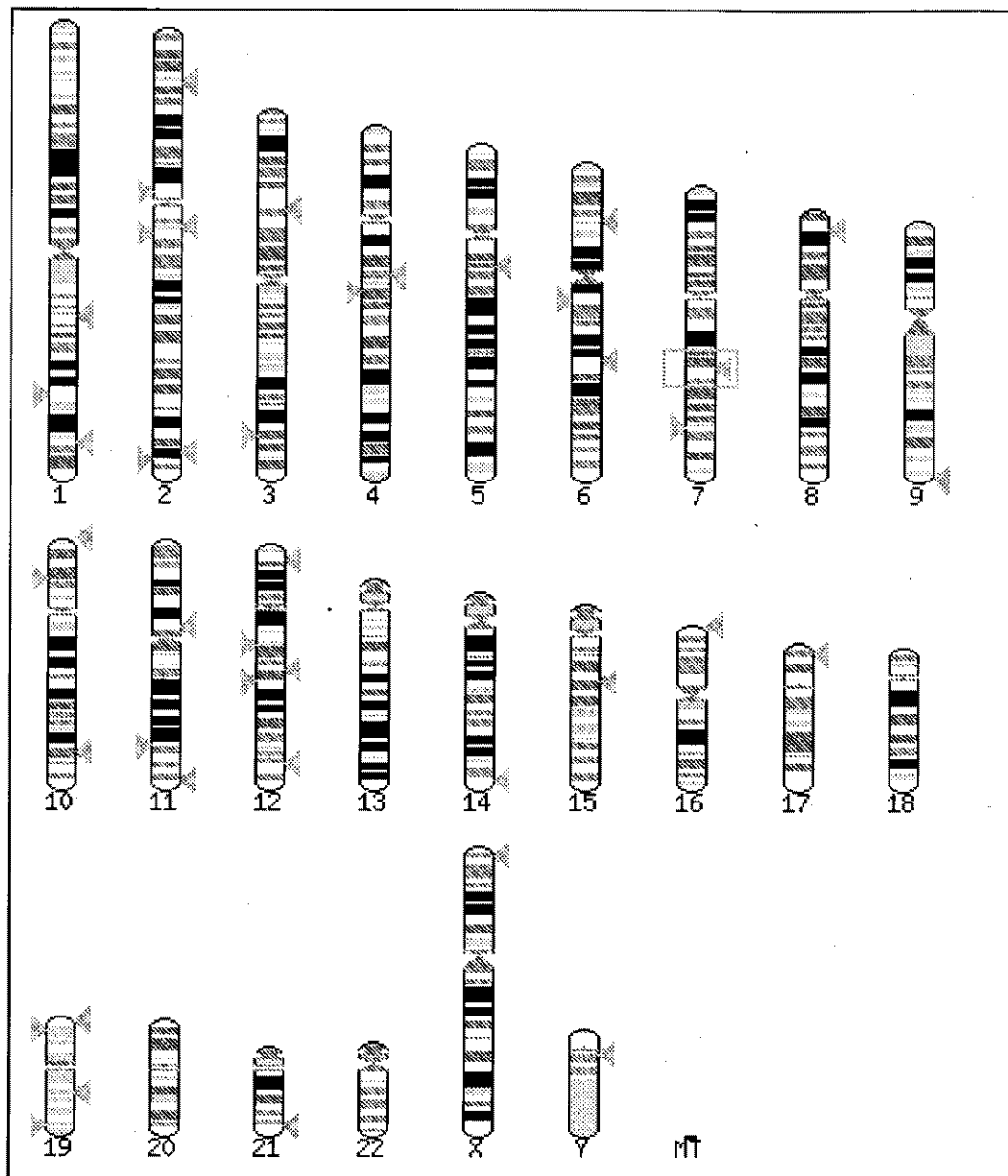


Figura 10A



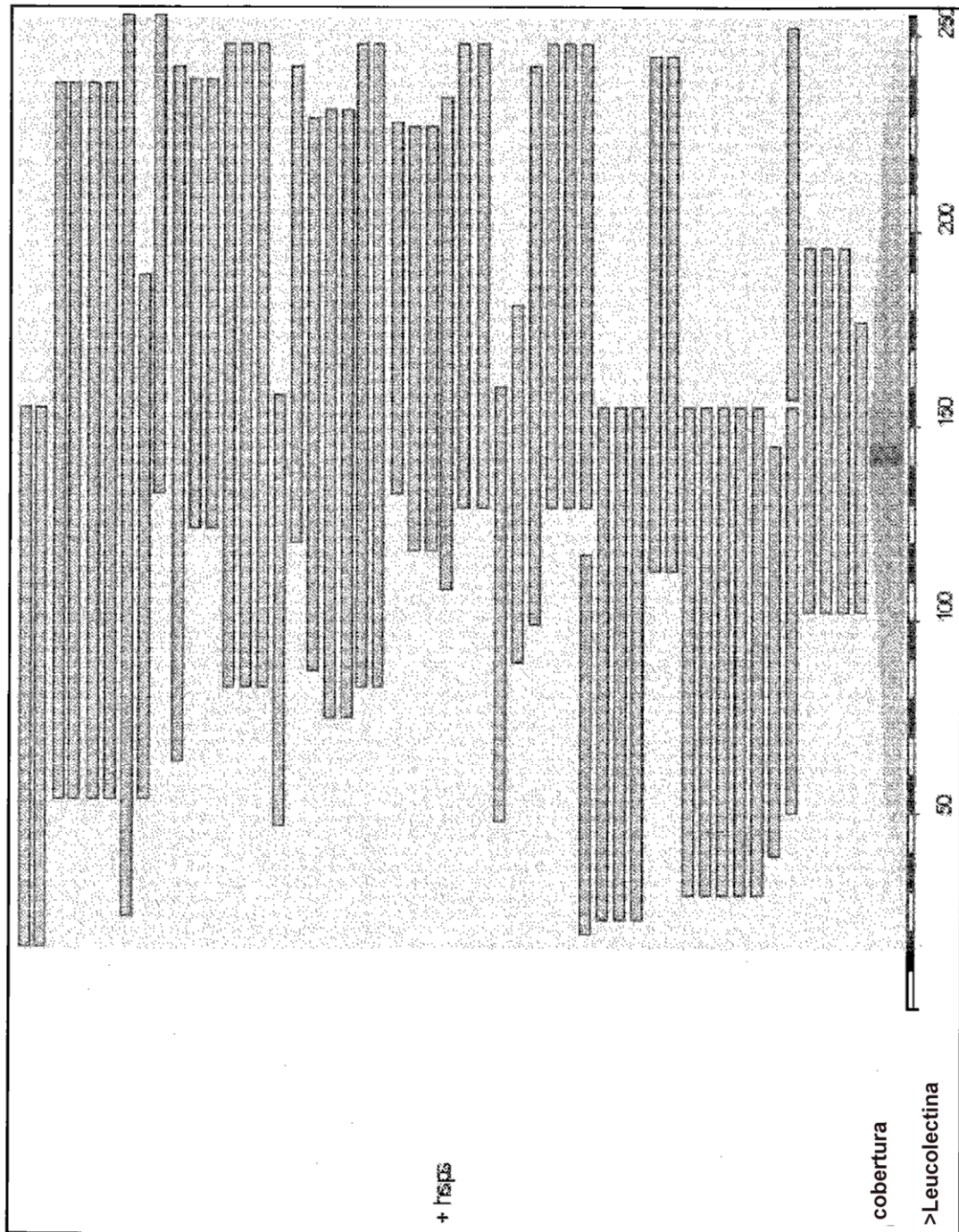


Figura 10B

Figura 10C

[illegible]

Figura 10C  
continuación



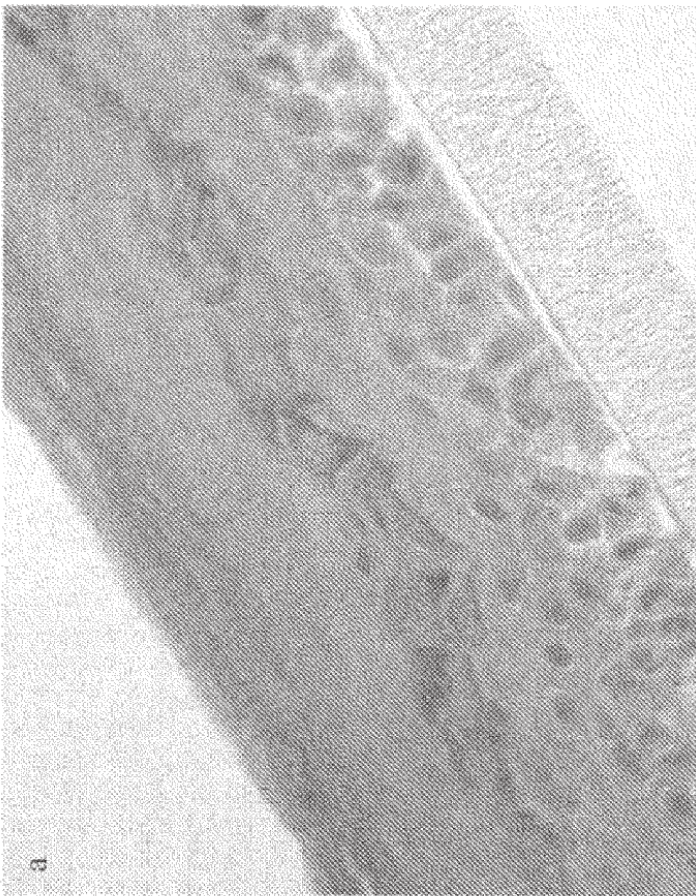
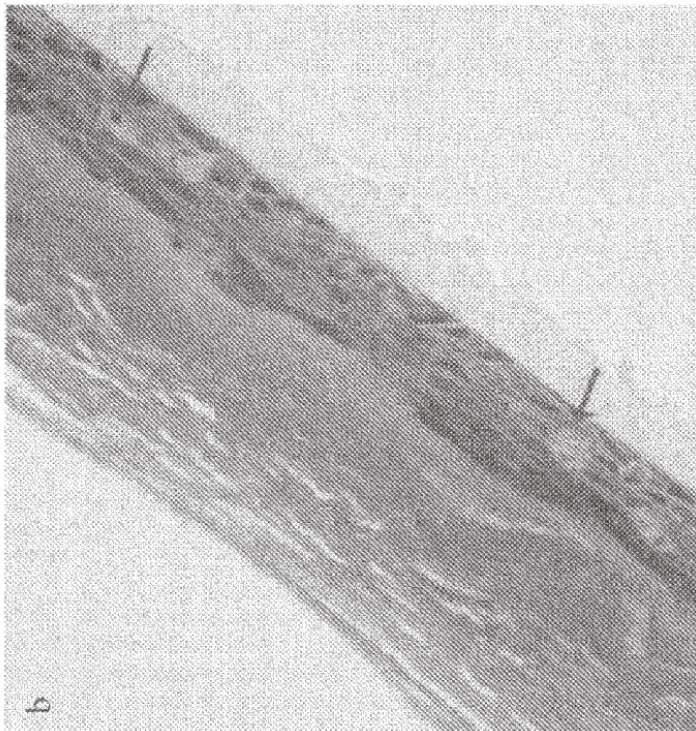


Figura 11



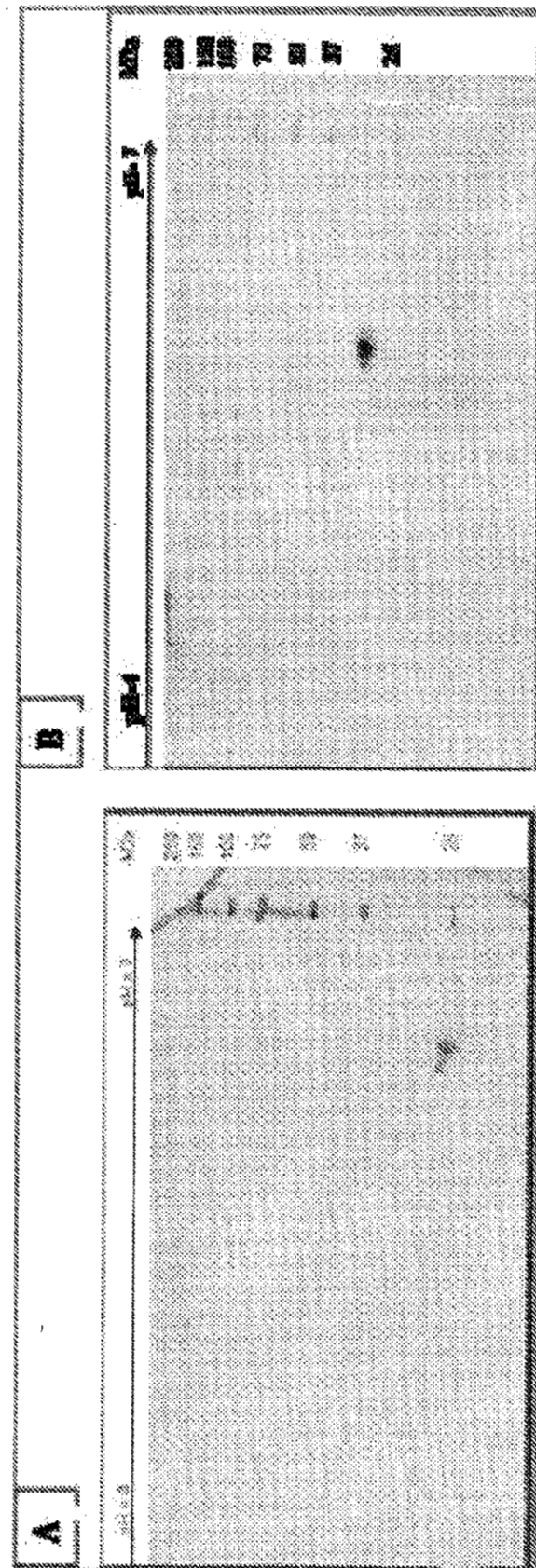


Figura 12

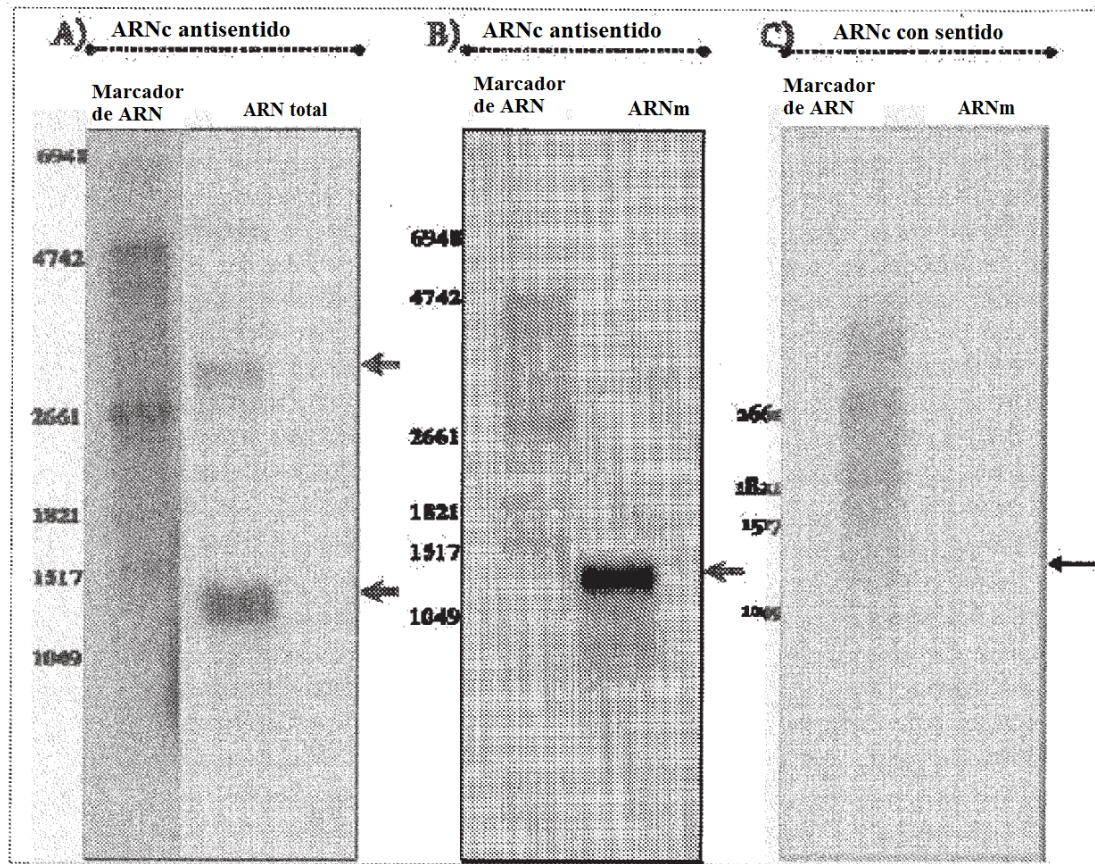
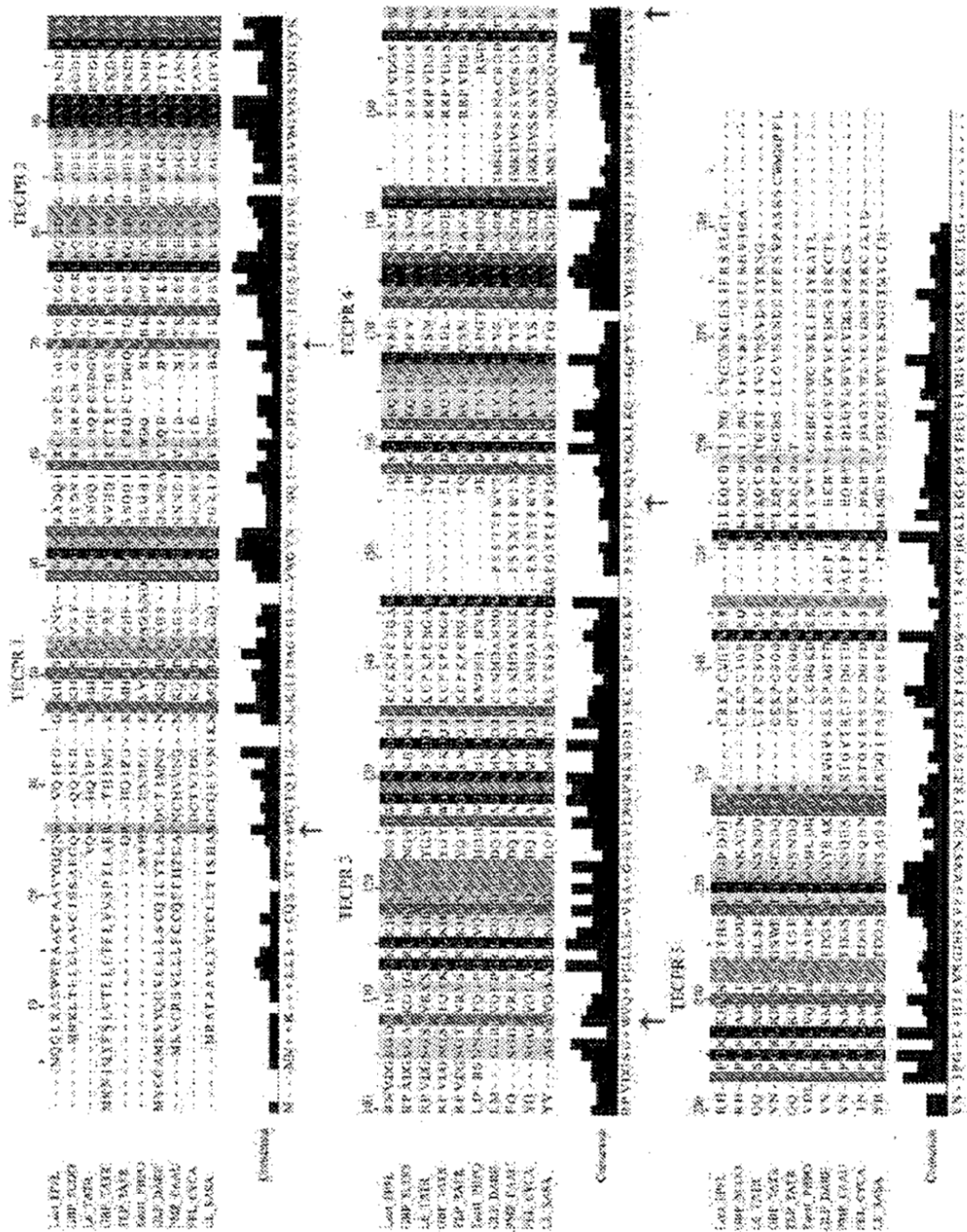


Figura 13





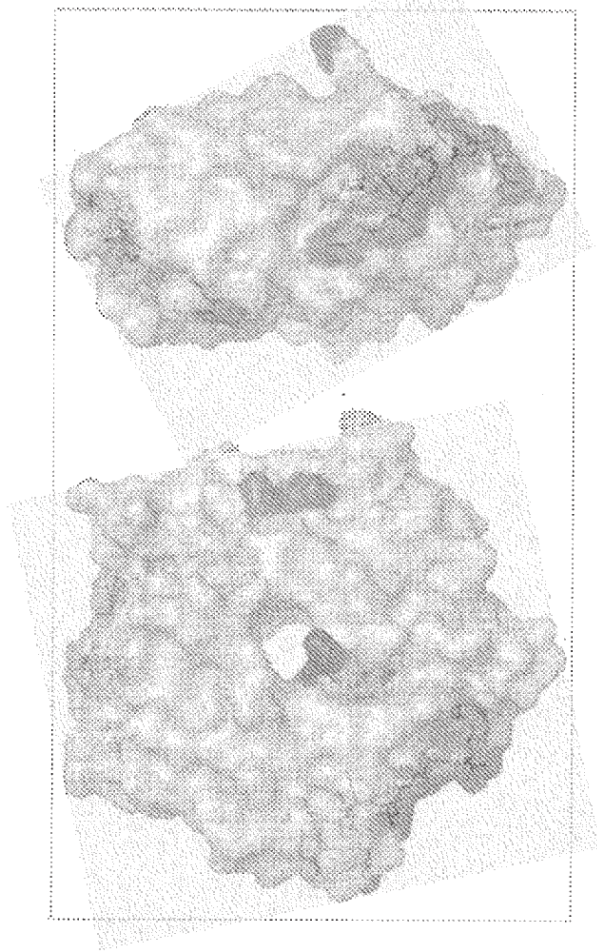
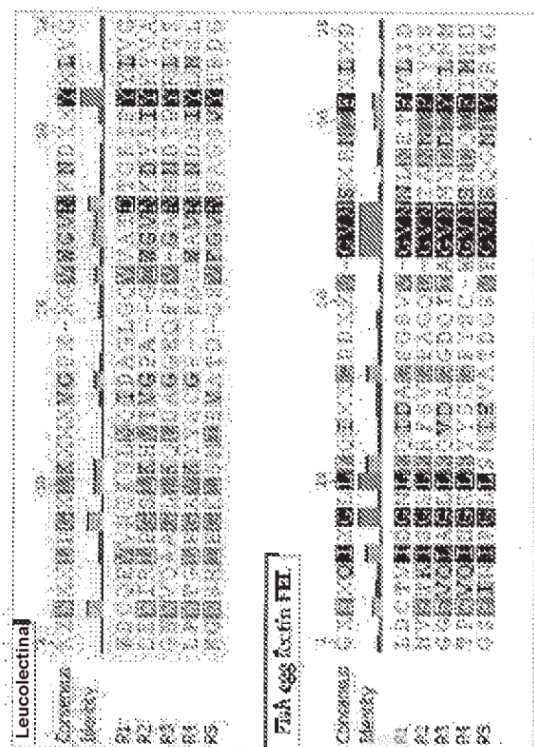


Figura 15



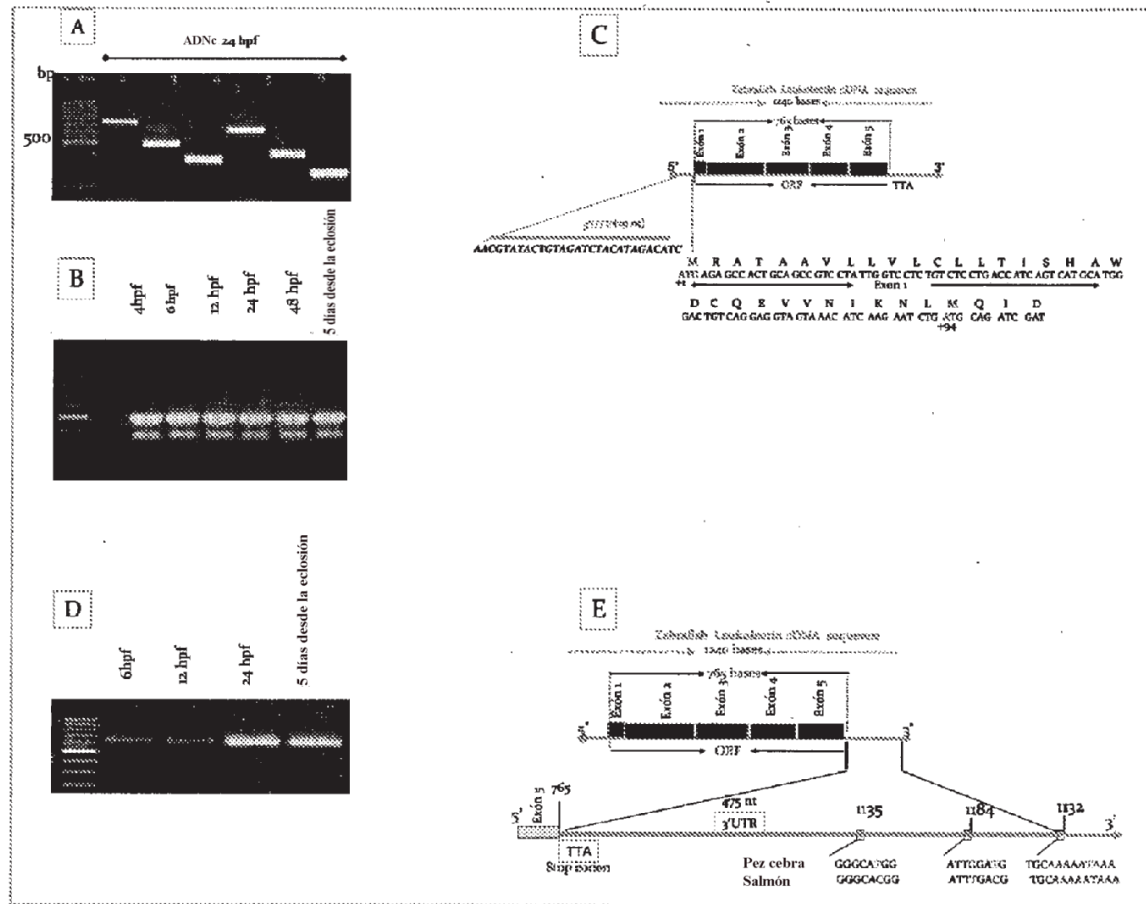


Figura 16

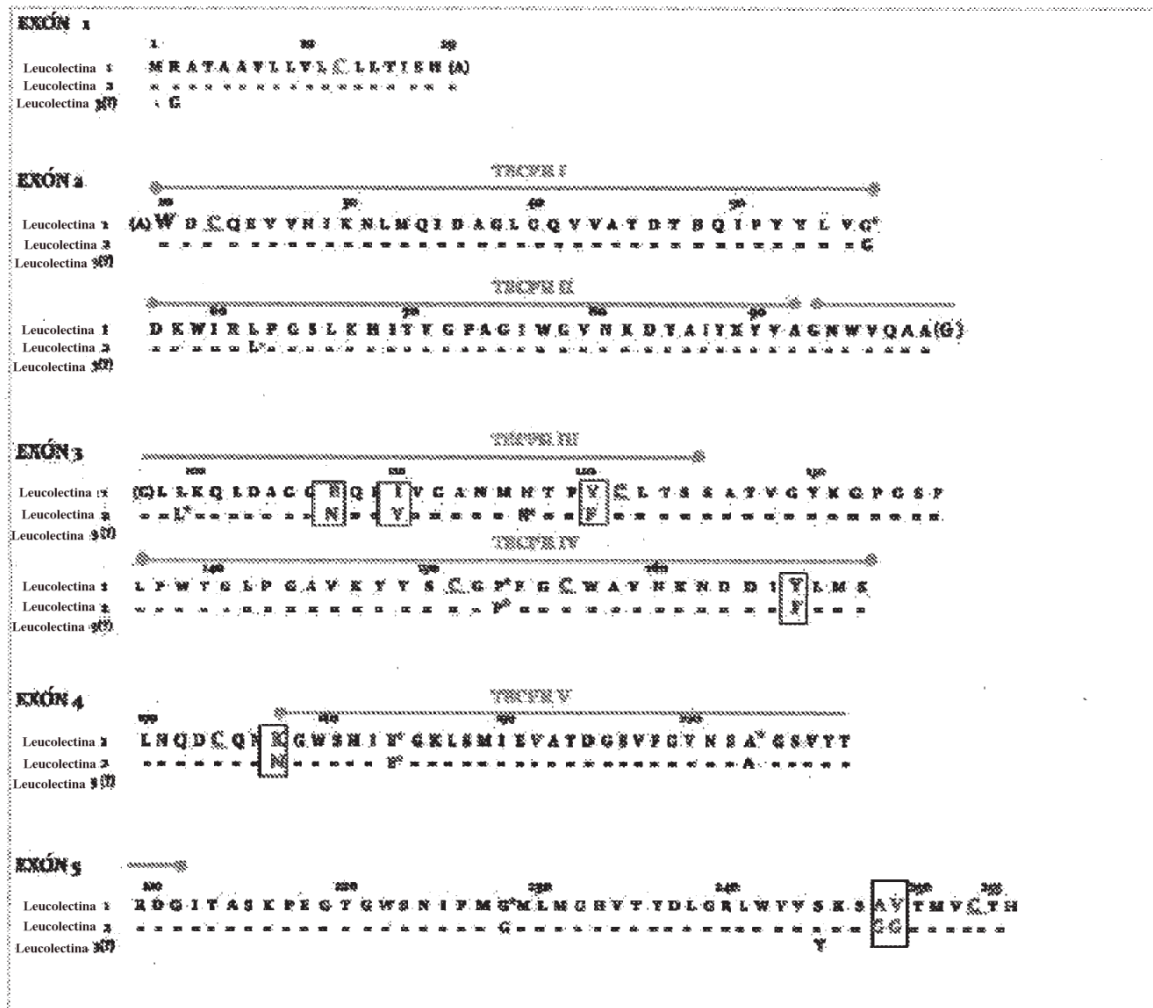


Figura 17

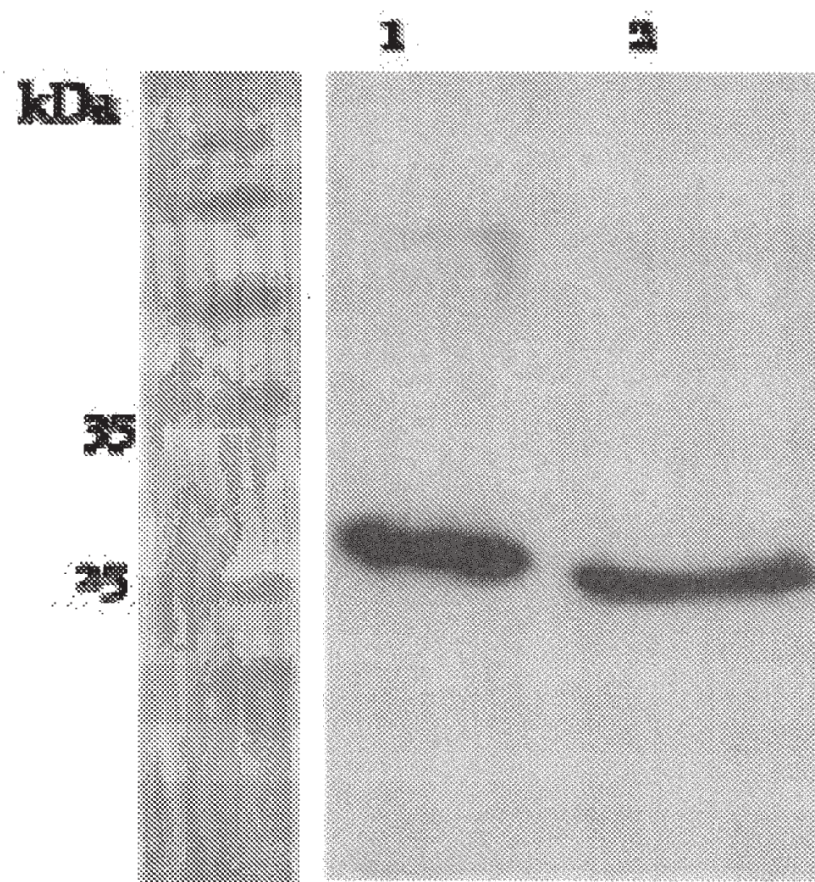
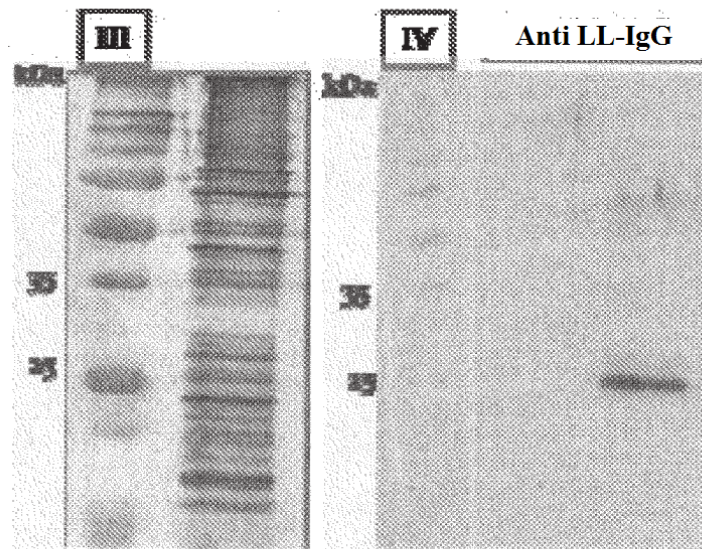


Figura 18

A



B

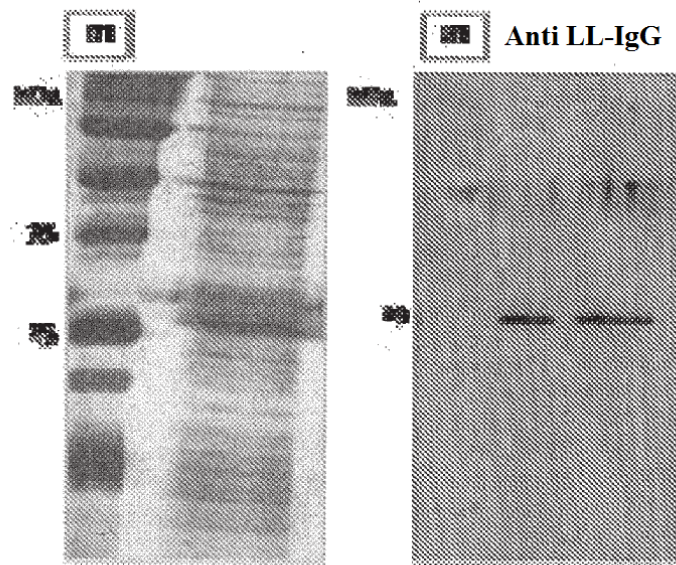


Figura 19