



(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 331 407**

(51) Int. Cl.:

C12Q 1/68 (2006.01)

G01N 33/50 (2006.01)

G01N 33/53 (2006.01)

G01N 33/564 (2006.01)

A61K 38/17 (2006.01)

A61K 48/00 (2006.01)

A61P 31/10 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **01942050 .4**

(96) Fecha de presentación : **05.06.2001**

(97) Número de publicación de la solicitud: **1290227**

(97) Fecha de publicación de la solicitud: **12.03.2003**

(54) Título: **Composiciones, kits y procedimientos de identificación y modulación de diabetes tipo I.**

(30) Prioridad: **05.06.2000 US 209703 P**

(45) Fecha de publicación de la mención BOPI:
04.01.2010

(45) Fecha de la publicación del folleto de la patente:
04.01.2010

(73) Titular/es: **Genetics Institute, L.L.C.**
87 Cambridge Park Drive
Cambridge, Massachusetts 02140, US
THE GENERAL HOSPITAL CORPORATION

(72) Inventor/es: **Byrne, Michael, C.;**
Hill, Andrew, A. y
Wilson, S., Brian

(74) Agente: **Carpintero López, Mario**

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

Composiciones, kits y procedimientos de identificación y modulación de diabetes tipo I.

5 La presente solicitud reivindica la prioridad de la solicitud Provisional de los EEUU No. 60/209,703, presentada el 5 de junio de 2000.

La diabetes mellitus es un síndrome con componentes metabólicos, vasculares y neuropáticos interrelacionados. El componente metabólico, caracterizado generalmente por hiperglicemia, comprende alteraciones en el metabolismo de los hidratos de carbono, las grasas y las proteínas provocadas por la ausencia o una reducción marcada de la secreción de insulina y/o una acción ineficaz de la insulina. El componente vascular incluye anomalías en los vasos sanguíneos que producen complicaciones cardiovasculares, retinales y renales. Las anomalías en los sistemas nerviosos periférico y autónomo comprenden un tercer componente del síndrome diabético.

15 Existen dos tipos de la diabetes mellitus: tipo I y tipo II. A la diabetes tipo I también se la llama diabetes “insulino-dependiente”, debido a que los individuos que sufren este trastorno no pueden sintetizar su propia insulina y por lo tanto deben inyectarse insulina periódicamente. Quienes sufren de la diabetes tipo II, por su parte, son capaces de sintetizar la insulina, pero esta insulina es insuficiente para sus necesidades o este no la utiliza eficazmente. La diabetes tipo II (diabetes no insulino-dependiente) generalmente se controla mediante medicación oral.

20 Al igual que muchas enfermedades autoinmunes, la diabetes mellitus tipo I (IDDM) es un trastorno con una etiología muy compleja, que se cree implica “desencadenantes” ambientales que interactúan con una susceptibilidad genética poligénica. La identificación más temprana de un locus genético que confiere susceptibilidad a la IDDM tuvo lugar hace unos treinta años, cuando se descubrió la asociación de ciertos alelos del complejo mayor de histocompatibilidad (MHC) con la diabetes (Eisenbarth (1986) *N. Engl. J. Med.* 214(21): 1360-1368; Wicker *et al.* (1995) *Annu. Rev. Immunol.* 13: 179-200; y Todd *et al.* (1987) *Nature* 329: 599-604). Estudios epidemiológicos posteriores han demostrado concluyentemente que el riesgo de contraer diabetes está fuertemente relacionado con la herencia de ciertos alelos de MHC de clase II. Si bien el locus MHC es el factor de riesgo genético más significativo para la diabetes, representa menos del 50% de este riesgo genético, lo que indica claramente que los locus no MHC dispersados por el genoma también realizan una contribución fundamental a la susceptibilidad y gravedad de la enfermedad (Davies *et al.*) (1994) *Nature* 371: 130-136).

35 Hasta la fecha, los barridos genómicos exhaustivos han localizado 18 locus de susceptibilidad diferentes para la IDDM (Becker (1999) *Diabetes* 48(7): 1353-1358). Estos supuestos locus de susceptibilidad deben ser evaluados con cuidado, puesto que muchos de ellos poseen valores estadísticos significativos menores que el estándar común (logaritmo de impares [LOD $\geq$ 3] y solo una parte de ellos ha sido replicada. Curiosamente, 15 de los 18 locus candidatos se superponen con locus de susceptibilidad establecidos anteriormente para otras enfermedades autoinmunes como el lupus eritematoso sistémico (SLE), la esclerosis múltiple (MS, por sus siglas en inglés), la artritis reumatoide, la espondilitis anquilosante y enfermedad celíaca (Becker (1999), *Diabetes* 48(7): 1353-1358). Si bien no es definitivo, este resultado podría sugerir que los locus de susceptibilidad candidatos pueden contener genes que tienen un papel fundamental en la función y regulación inmune normal. Sin embargo, la identificación de el/los gen(es) en los supuestos locus de susceptibilidad responsables de esta regulación inmune ha resultado ser mucho más difícil que la identificación de los propios locus. Cada uno de estos locus puede abarcar distancias genéticas enormes de hasta 10-30 cM de tamaño y pueden contener cientos, sino miles de genes (Becker (1999) *Diabetes* 48(7): 1353-1358), muchos de los cuales tienen una función no definida y pueden actuar en forma combinatoria.

En un abordaje de la identificación de la base molecular que subyace a la diabetes tipo I, complementario a los barridos genómicos, también se ha estudiado la participación de diferentes células inmunes en la enfermedad. Se sabe que las células T invariantes CD161⁺V214J α 281 (células NKT) son importantes en la regulación de la célula T auxiliar (Th) Th1/Th2 bias (Bendelac *et al.* (1997) *Annu. Rev. Immunol.* 15: 535-562) y se ha demostrado que las células NKT están presentes en números reducidos y que su frecuencia se reduce aún más antes del comienzo de la enfermedad, en varios modelos murinos de autoinmunidad (Takeda y Dennert (1993) *J. Exp. Med.* 177: 155-164; Mieza *et al.* (1996) *J. Immunol.* 156: 4035-4040; y Baxter *et al.* (1997) *Diabetes* 46: 572-582). Cuando se transfirió esta población celular de donantes diabéticos no obesos o donantes diabéticos no obesos/V14J α 281-transgénicos a animales prediabéticos, los receptores estaban protegidos contra la diabetes (Baxter *et al.* (1997) *Diabetes* 46: 572-582; Lehuen *et al.* (1998) *J. Exp. Med.* 188:1831-1839). Esta transferencia de protección se inhibió significativamente mediante la administración conjunta de anticuerpos anti-IL-4 (Hammond *et al.* (1998) *J. Exp. Med.* 187: 1047-1056).

Los seres humanos tienen una población de células T CD161⁺V α 24J α Q homóloga invariante (es decir, sin adiciones de regiones N) cuyo elemento restrictivo, como el de las células murinas CD161⁺V α 14J α 281, es la molécula no polimórfica de clase Ib CD1d (Exley *et al.* (1997) *J. Exp. Med.* 186(1): 109-20). Se ha demostrado que en cinco juegos de gemelos y trillizos monocigóticos discordantes para la diabetes tipo 1, las células T V α 24J α Q estaban presentes en frecuencias significativamente más altas en los hermanos no diabéticos (Wilson *et al.* (1998) *Nature* 391 (6663):177-81). Además, los clones de células V α 24J α Q T de los hermanos no diabéticos segregaron tanto IL-4 como IFN- γ , mientras que aquellos derivados de los hermanos diabéticos tenían una dificultad extrema en su capacidad para segregar IL-4.

En WO-A-93/19174 se divulga un marcador diferenciador de células T humano llamado "HT6", un anticuerpo dirigido a este marcador, así como un procedimiento para detectar la expresión de este marcador. HT6 es una proteína expresada por los linfocitos T humanos, que en particular ofrece la posibilidad de detectar células Ti/CD3+, CD4+ y/o CD45RA.

WILSON B *ET AL*: "Extreme Th1 bias of invariant Valpha 24Jalpha4T cells in type 1 diabetes" NATURE, vol. 391, 8 de enero 1998 (1998-01-08), páginas 177-81, y WO-A-99/34209 se dirigen a un nivel reducido de secreción de IL-4 de las células NKT diabéticas (activadas) en comparación con células no diabéticas (activadas), que se relaciona con el posible desarrollo de la diabetes tipo I.

En una realización, la invención proporciona un procedimiento para evaluar si un individuo sufre de la diabetes tipo I. Como se define en la reivindicación 1, en una realización preferida, el marcador corresponde a un polinucleótido transcrito o una parte de este, en el que el polinucleótido incluye el marcador. En una realización particularmente preferida el nivel de expresión del marcador en la muestra difiere con respecto al nivel normal de expresión del marcador en un individuo que no sufre de la diabetes tipo I, por un factor de al menos dos, y en una realización más preferida aún, los niveles de expresión difieren por un factor de al menos cinco.

Como se define en la reivindicación 18 la muestra incluye células NKT obtenidas del individuo. En otra realización preferida, el nivel de expresión del marcador en la muestra se evalúa detectando en la muestra la presencia de una proteína correspondiente al marcador. En una realización especialmente preferida, la presencia de la proteína se detecta utilizando un reactivo que específicamente se une con la proteína. En una realización más preferida aún, el reactivo se selecciona del grupo de reactivos que incluye un anticuerpo, un derivado de un anticuerpo, y un fragmento de un anticuerpo. En otra realización preferida, el nivel de expresión del marcador en la muestra se evalúa detectando en la muestra la presencia del polinucleótido transcrito o una parte del este, en el que el polinucleótido transcrito incluye el marcador. En una realización particularmente preferida, el polinucleótido transcrito es un ARNm o un ADNc. En otra realización particularmente preferida, la etapa de detección comprende además amplificar el polinucleótido transcrito.

En otra realización preferida, el nivel de expresión del marcador en la muestra se evalúa detectando en la muestra la presencia del polinucleótido transcrito que se hibrida con el marcador o se hibrida con una parte de un polinucleótido en condiciones rigurosas de hibridación, en el que el polinucleótido incluye al marcador. En otra realización preferida, el nivel de expresión en la muestra de cada uno de los marcadores de un conjunto, seleccionados independientemente de los marcadores enumerados en las Tablas 1 y 4 se compara con el nivel normal de expresión de cada uno de los marcadores del conjunto en muestras del mismo tipo obtenidas de individuos de control que no sufren diabetes tipo I, en el que el nivel de expresión de uno o más de los marcadores está significativamente alterado, con respecto con los niveles normales de expresión correspondientes de los marcadores, es una indicación de que el individuo sufre de la diabetes tipo I. En una realización particularmente preferida, el conjunto incluye dos o más marcadores. En una realización aún más preferida, el conjunto incluye al menos 5 de los marcadores establecidos en las Tablas 1 y 4.

En otra realización, la invención proporciona un procedimiento para controlar la progresión de la diabetes tipo I como se define en la reivindicación 18. En una realización preferida, el marcador se selecciona del grupo que incluye los marcadores enumerados en las Tablas 1 y 4 y sus combinaciones. En otra realización preferida, el marcador corresponde a un polinucleótido transcrito o una parte de este, en el que el polinucleótido incluye al marcador. La muestra incluye células NKT obtenidas del individuo. En una realización particularmente preferida, las células NKT se recogen de tejido pancreático o sanguíneo.

En otra realización, la invención proporciona un procedimiento para evaluar la eficacia de un compuesto de prueba para inhibir la diabetes tipo I como se define en la reivindicación 23. En una realización preferida, la primera y segunda muestra son partes de una única muestra obtenida del individuo. En otra realización preferida, la primera y segunda muestra son partes de una muestra combinadas obtenidas del individuo.

En otra realización, la invención proporciona un procedimiento para evaluar la eficacia de una terapia para inhibir la diabetes tipo I como se define en la reivindicación 22.

En otra realización, la invención proporciona un procedimiento para seleccionar una composición para inhibir la diabetes tipo I como se define en la reivindicación 26.

En otra realización, la invención proporciona el uso de un kit para llevar a cabo los procedimientos de la invención como se define en la reivindicación 28.

En otra realización, la invención proporciona un procedimiento para evaluar el potencial de un compuesto de prueba para desencadenar la diabetes tipo I como se define en la reivindicación 27.

En la siguiente descripción detallada y en las reivindicaciones resultarán evidentes otras características y ventajas de la invención.

Breve descripción de los dibujos

La Fig. 1 es una representación gráfica de la metodología utilizada en la invención para determinar la expresión relativa de genes marcadores en células diabéticas y no diabéticas.

En la Fig. 2 se ilustra la expresión discordante de episodios regulados por la quinasa PI3 en clones de células T (NKT) IL-4⁺ e IL-4 nula V α 24J α Q. En la Fig. 2 A se muestran gráficos de barras que indican los niveles de secreción de IFN- γ e IL-4, medidos mediante el ensayo ELISA, de clones de células T V α 24J α Q GW4 (no diabéticas) y ME10 (células diabéticas) activadas con anti-CD3 unido a placa o un control Ig, en presencia o ausencia de varios inhibidores o estimulantes de la senda de transducción de señales. Los inhibidores utilizados incluyen wortmannin ("wort", 10 mM), LY294002 ("LY", 10 μ M), PD98059 ("PD", 50 μ M), S8203580 ("S8", 50 μ M). Los estimulantes utilizados fueron el éster de forbol PMA (1 ng/ml), ionomicina ("iono", 1 ng/ml), y ciclosporina A ("CsA", 5 ng/ml). Se recogieron los puntos de datos por triplicado y los datos que se muestran son representativos de cuatro experimentos independientes. La Fig. 2B ilustra una gráfica de la fluorescencia detectada en el tiempo por citometría de flujo (utilizando un instrumento Cytomation MoFlo) de los clones de células T V α 24J α Q CW4 (IL-4⁺) y ME10 (IL-4-nula) marcadas con indo-1 (10 μ M) y estimuladas con anti-CD3 (10 μ g/ml). Al final del experimento, se agregó la ionomicina a una concentración final de 1 μ g/ml para determinar un flujo máximo. Se presenta la proporción de fluorescencia de Indo-1 a 410/490 nm (410: Ca²⁺ unido; 490: Ca²⁺ libre) después de la estimulación en un par de clones representativos.

La Fig. 3 es una representación gráfica de genes expresados diferencialmente en clones de células NKT derivadas de un par gemelo diabético/no diabético. Se trataron clones de ME10 (diabéticos) y GW4 (no-diabéticos) con IgG de control (designado R=Reposo) o anti-CD3 (designado A=Activado) durante cuatro horas, tras lo que se aisló el ARN y se lo analizó en chips genéticos que controlan la expresión de 6800 genes humanos de la colección Unigene. Se eligieron genes cuya expresión estaba modulada al menos dos veces en cualquiera de los clones para un análisis de conglomerados utilizando el algoritmo del mapa autoorganizado (Tamayo *et al.* (1999)). Este procedimiento se utilizó para aglomerar genes en seis grupos distintos, basado en patrones de expresión diferencial entre ME10 y GW4, independiente de la magnitud de expresión. El primer grupo muestra los seis patrones representados cuando se utilizan todos los genes que cumplen con el criterio de doble cambio. Los otros agrupamientos revelan los patrones de expresión diferencial de clases funcionales de genes seleccionados.

La invención se refiere, en parte, a correlaciones recientemente descubiertas entre la expresión de marcadores seleccionados en células NKT y la presencia de la diabetes tipo I en un individuo. Se ha encontrado que los niveles relativos de expresión de estos marcadores, tanto solos como combinados, son indicativos de una predisposición en el individuo a la diabetes tipo I y/o del diagnóstico de la presencia o posible presencia de la diabetes tipo I en un individuo. La invención proporciona paneles de marcadores, procedimientos para detectar la presencia o ausencia de la diabetes tipo I en una muestra o individuo, y procedimientos para predecir la incidencia de la diabetes tipo I.

La presente invención se basa, al menos en parte, en la identificación de un número de marcadores genéticos, establecidos en las Tablas 1 y 4, que se expresan diferencialmente en células NKT activadas de un individuo diabético con respecto a un individuo no diabético. Se analizó un grupo de 6800 genes conocidos en busca de la expresión en células NKT diabéticas activadas frente a células NKT no diabéticas activadas tomadas de gemelos idénticos discordantes para la diabetes tipo I (ver el Ejemplo 2). En las Tablas 1 y 4 se identifican esos genes con al menos dos diferencias entre las células activadas enfermas y normales.

Se observaron seis patrones de expresión diferentes en células diabéticas activadas en comparación con células NKT no diabéticas. En la Tabla 1 (representativa de la fila 1, columna 1 en cada uno de los conglomerados establecidos en la Fig. 3) se enumera cada uno de los genes con una mayor expresión en células NKT diabéticas activadas y que no cambian o que aumentan en menor grado en su expresión en células NKT no diabéticas activadas, en relación con células de control en reposo apropiadas. En la Tabla 2 (representativa de la fila 1, columna 2 en cada uno de los conglomerados establecidos en la Fig. 3; fuera del ámbito de la invención) se enumera cada uno de los genes con expresión incambiada en células NKT diabéticas activadas en relación con células de control en reposo, pero tienen mayor expresión en células NKT no diabéticas activadas, en relación con células de control en reposo. En la Tabla 3 (representativa de la fila 1, columna 3 en cada uno de los conglomerados establecidos en la Fig. 3) se enumeran los genes con una mayor expresión tanto en células NKT diabéticas y no diabéticas activadas en relación con células de control en reposo apropiadas. En la Tabla 4 (representativa de la fila 2, columna 1 en cada uno de los conglomerados establecidos en la Fig. 3; fuera del ámbito de la invención) se enumera aquellos genes con una menor expresión en células NKT no diabéticas activadas en relación con células de control en reposo, pero que no cambian o disminuyen en un menor grado en su expresión en células NKT diabéticas activadas, en relación con células de control en reposo. En la Tabla 5 (representativa de la fila 2, columna 2 en cada uno de los conglomerados establecidos en la Fig. 3; fuera del ámbito de la invención) se enumeran los genes con una mayor expresión en células NKT no diabéticas activadas en relación con células de control en reposo, pero que disminuyen en su expresión en células NKT diabéticas activadas, en relación con células de control en reposo. En la Tabla 6 (representativa de la fila 2, columna 1 en cada uno de los conglomerados establecidos en la Fig. 3, fuera del ámbito de la invención) se enumeran los genes con una menor expresión en células NKT diabéticas activadas en relación con células de control en reposo, pero que no cambian o disminuyen en menor medida en su expresión en células NKT no diabéticas en relación con células de control en reposo.

Se generaron clones de células T NKT, CD4, y CD8 y se los estimuló con anti-CD3 durante 2, 4, 8, 24 o 48 horas. En la Tabla 9 (fuera del ámbito de la invención) se establecen los genes que se identificaron en una búsqueda que requería incrementar al menos tres veces los niveles de ARNm al menos en un momento en muestras de células NKT. En la Tabla 10 (fuera del ámbito de la invención) se establecen los genes que se identificaron en una búsqueda que requería incrementar al menos tres veces los niveles de ARNm al menos en un momento para las tres replicas del experimento en muestras de células CD4. En la Tabla 11 (fuera del ámbito de la invención) se establecen los genes que se identificaron en una búsqueda que requería incrementar al menos tres veces los niveles de ARNm al menos en un momento para las tres replicas del experimento en muestras de células CD8.

Se generaron clones de células T NKT, CD4, y CD8. En la Tabla 12 (fuera del ámbito de la invención) se establecen los genes que se identificaron en una búsqueda que requería al menos tres cambios en los niveles de ARNm para las tres replicas del experimento en muestras de células CD4 en reposo con respecto a muestras de células NKT en reposo. En la Tabla 13 (fuera del ámbito de la invención) se establecen los genes que se identificaron en una búsqueda que requería al menos tres cambios en los niveles de ARNm para las tres replicas del experimento en muestras de células CD8 en reposo con respecto a muestras de células NKT en reposo.

Se identificaron varios genes que fueron regulados diferencialmente en células NKT diabéticas en reposo (por ejemplo, inactivas) en relación con células NKT no diabéticas. Estos genes se establecen en la Tabla 8 (fuera del ámbito de la invención).

En la técnica se conocen varios genes implicados en la diabetes tipo I, que se establecen en la Tabla 7. No se pretende que estos genes se utilicen solos en los métodos, composiciones y kits de la invención, pero se los puede utilizar en combinación con los marcadores de la invención que se establecen en las Tablas 1 y 4.

Por consiguiente, la presente invención se refiere al uso de los genes establecidos en las Tablas 1 y 4 (por ejemplo, el ADN o ADNc), los transcritos de ARNm correspondientes y los polipeptidos codificados como marcadores para la presencia o el riesgo de desarrollar diabetes tipo I. Estos marcadores son útiles además para correlacionar el grado y/o gravedad de la enfermedad. Se pueden arreglar convenientemente paneles de los marcadores para utilizar en kits o en soportes sólidos. Los marcadores también pueden ser útiles para evaluar la eficacia de un tratamiento para la diabetes tipo I.

En un aspecto, se proporcionan los marcadores cuya cantidad o actividad se correlaciona con la presencia de la diabetes tipo I. Los marcadores pueden ser moléculas de ácido nucleico (por ejemplo ADN, ADNc, o ARN) o polipéptidos. Estos marcadores aumentan o disminuyen en cantidad o actividad en las células NKT diabéticas en comparación con células NKT no diabéticas. Por ejemplo, el gen IFN- γ (fuera del ámbito de la invención) (número de acceso J00219) tiene un mayor nivel de expresión en células NKT no diabéticas activadas pero no en células NKT diabéticas activadas (Tabla 2), mientras que el gen linfotóxica-beta (llamado "LT- β ") (número de acceso U89922) tiene un mayor expresión en células NKT diabéticas activadas pero no en células NKT no diabéticas activadas (Tabla 1). Tanto el aumento como la disminución de ARNm para estos genes (y para otros genes establecidos en las Tablas 1 y 4), y también los mayores o menores niveles de productos proteicos de estos genes (y otros genes establecidos en las Tablas 1 y 4) sirven como marcadores de la diabetes tipo I. Preferentemente, los mayores o menores niveles de los marcadores son aumentos y disminuciones de una magnitud que es estadísticamente significativa en comparación con muestras de control apropiadas (por ejemplo, muestras no afectadas por la diabetes tipo I o una afección asociada al NKT). En realizaciones particularmente preferidas, el marcador aumenta o disminuye con respecto a muestras de control en al menos 2, 3, 4, 5, 6, 7, 8, 9 ó 10 veces o más. En forma similar, un entendido en la técnica será consciente de que una metodología de detección preferida es aquella en la que los valores de detección resultantes están por encima del límite de detección mínimo de la metodología.

Se puede medir la cantidad relativa de ARN o marcador de proteína de la invención mediante cualquier procedimiento conocido en la técnica (ver, por ejemplo., Sambrook, J., Fritsh, E. F., y Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; y *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Las metodologías típicas para la detección de ARN incluyen la extracción de ARN de una muestra de célula o tejido, seguido de la hibridación de una sonda marcada (por ejemplo, una molécula de ácido nucleico complementaria) específica para el ARN objetivo con el ARN extraído y la detección de la sonda (por ejemplo, Northern blotting). Las metodologías típicas para la detección de proteína incluyen la extracción de proteína de una muestra de célula o tejido, seguido de la hibridación de una sonda marcada (por ejemplo, un anticuerpo) específica para la proteína objetivo con la proteína de muestra y la detección de la sonda. El grupo marcador puede ser un radioisótopo, un compuesto fluorescente, una enzima o un cofactor de una enzima. También se puede evaluar la detección de moléculas de ácido nucleico y proteínas específicas mediante electroforesis en gel, cromatografía en columna, secuencia directa o PCR cuantitativo (en el caso de moléculas de ácido nucleico) entre otras varias técnicas conocidas por los entendidos en la técnica.

En ciertas realizaciones, los propios genes (por ejemplo, el ADN o ADNc) pueden servir como marcadores para la diabetes tipo I. Por ejemplo, la ausencia de ácidos nucleicos correspondientes a un gen (por ejemplo un gen de la Tabla 1), por ejemplo mediante eliminación de todo o parte del gen, se puede relacionar con la enfermedad. De forma similar, un aumento de ácido nucleico correspondiente a un gen, por ejemplo mediante duplicación de un gen, también se puede relacionar con la enfermedad.

La detección de la presencia de un número de copias de todo o parte de un gen marcador se puede realizar utilizando cualquier procedimiento conocido en la técnica. Generalmente, es conveniente evaluar la presencia y/o cantidad de un ADN o ADNc mediante análisis Southern, en el que se extrae todo el ADN de una muestra de célula o de tejido, se hibrida con una sonda marcada (por ejemplo una molécula de ADN complementario) y se detecta la sonda. El grupo marcador puede ser un radioisótopo, un compuesto fluorescente, una enzima o un cofactor de una enzima. Otros procedimientos útiles para la detección de ADN y/o cuantificación incluyen secuencia directa, electroforesis en gel, cromatografía en columna, y PCR cuantitativa, conocida por los entendidos en la técnica.

También se describen las moléculas de proteínas y ácido nucleico que son estructuralmente diferentes de las moléculas descritas anteriormente (por ejemplo, que tienen una secuencia de ácido nucleico o aminoácido levemente alterada), pero que tienen las mismas propiedades que las moléculas anteriores (por ejemplo, secuencia de aminoácidos codificado, o que cambian solamente en los residuos de aminoácido no esenciales). Dichas moléculas incluyen variantes alélicas y se las describe más detalladamente en la subsección I.

En otro aspecto, se describen los marcadores cuya cantidad o actividad se relaciona con la gravedad de la diabetes tipo I (ver, por ejemplo, el Ejemplo 3). Estos marcadores tienen una mayor o menor cantidad o actividad en células NKT diabéticas de tal manera que se relaciona positiva o negativamente con el grado de gravedad de la diabetes tipo I. En otro aspecto más, se describen los marcadores cuya cantidad o actividad se relaciona con un riesgo en un individuo para desarrollar la diabetes tipo I. Estos marcadores aumentan o disminuyen su actividad o cantidad en directa relación con la posibilidad de que un individuo desarrolle la diabetes tipo I.

Cada marcador se puede considerar individualmente, aunque también se proporcionan combinaciones de dos o más marcadores para utilizar en los procedimientos y composiciones de la invención a fin de aumentar la confianza del análisis. En otro aspecto, también se proporcionan paneles de los marcadores. En las Tablas 1 y 4 se establecen los paneles de los marcadores de la invención. Los entendidos en la técnica apreciarán que los procedimientos de la invención pueden practicarse con cualquiera de los paneles establecidos en las Tablas 1 y 4, o cualquier parte o combinación de estos.

Los entendidos en la técnica apreciarán también que los paneles de marcadores pueden proporcionarse convenientemente en soportes sólidos. Por ejemplo, los polinucleótidos como oligonucleótidos o ADNc, se pueden acoplar a un conjunto (por ejemplo, un conjunto GeneChip para el análisis de hibridación), a una resina (por ejemplo, una resina con la que se puede rellenar una columna de cromatografía en columna) o una matriz (por ejemplo, una matriz de nitrocelulosa para el análisis Northern blot). La inmovilización de moléculas complementarias de los marcadores, ya sea o no covalentemente, permite un análisis discreto de la presencia o actividad de cada marcador en una muestra. En un conjunto, por ejemplo, los polinucleótidos complementarios de cada miembro de un panel de marcadores pueden estar unidos individualmente a ubicaciones conocidas, diferentes en el conjunto. El conjunto se puede hibridar con, por ejemplo, polinucleótidos extraídos de una muestra de sangre de un individuo. Se puede detectar la hibridación de polinucleótidos de la muestra con el conjunto en cualquier ubicación en el conjunto, y de ese modo se puede determinar la presencia o cantidad del marcador en la muestra. En una realización preferida, se utiliza un conjunto "GeneChip" (Affymetrix). En forma similar, los análisis Western se pueden realizar en anticuerpos inmovilizados específicos para diferentes marcadores polipeptídicos hibridados a una muestra de proteína de un individuo.

Los entendidos en la técnica podrán apreciar también que no se necesita que toda la proteína marcadora o molécula de ácido nucleico se conjugue con el soporte; una parte del marcador de longitud suficiente a efectos de la detección (por ejemplo, para hibridación), por ejemplo, una parte de un marcador que tiene 7,10,15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 100 ó más nucleótidos o aminoácidos de longitud puede ser suficiente a efectos de la detección.

El ácido nucleico y los marcadores proteicos divulgados se pueden aislar de cualquier tejido o célula de un individuo. Las células son células NKT. Sin embargo, los entendidos en la técnica podrán apreciar que otras muestras de tejido, incluidos fluidos corporales (por ejemplo, orina, bilis, suero, linfa, saliva, mucosidad y pus), entre otras muestras de tejido, también pueden servir como fuentes a partir de las que se pueden aislar los marcadores, o en los que la presencia, actividad y/o cantidad de marcadores se puede evaluar. Las muestras de tejido que contienen uno o más de los propios marcadores pueden ser útiles en los procedimientos divulgados y un entendido en la técnica conocerá de los procedimientos mediante los cuales las muestras se pueden obtener, almacenar y/o preservar convenientemente.

Antes de la invención se sabía que varios marcadores estaban asociados a la diabetes. Estos marcadores se establecen en la Tabla 7. Estos marcadores no están incluidos con los marcadores de la invención. Sin embargo, estos marcadores se pueden utilizar convenientemente en combinación con los marcadores de la invención en los métodos, paneles y usos de la invención.

La invención también proporciona el uso de kits para evaluar la presencia de células NKT diabéticas en una muestra (por ejemplo, una muestra de un individuo en riesgo de la diabetes tipo I), el kit que comprende un anticuerpo, en el que el anticuerpo específicamente se une con una proteína correspondiente a un marcador seleccionado del grupo constituido por los marcadores enumerados en las Tablas 1 y 4.

ES 2 331 407 T3

La invención también proporciona el uso de kits para evaluar la presencia de células NKT diabéticas de tipo 1 en una muestra (por ejemplo, un individuo en riesgo de la diabetes tipo I), el kit comprende una sonda de ácido nucleico, en el que específicamente se une con un polinucleótido transcripto correspondiente a un marcador seleccionado del grupo constituido por los marcadores enumerados en las Tablas 1 y 4.

La invención proporciona además el uso de kits para evaluar la adecuación de cada uno de varios compuestos para inhibir la diabetes tipo I en un individuo. Dichos kits incluyen varios compuestos que se evaluarán y un reactivo para evaluar la expresión de un marcador seleccionado del grupo constituido por uno o más marcadores establecidos en las Tablas 1 y 4.

Un entendido en la técnica apreciará fácilmente las modificaciones de las composiciones y procedimientos de la invención antes descritos, de conformidad con las técnicas estándar.

Para facilitar la comprensión de la presente invención, a continuación se definen varios términos y expresiones:

Tal como se utiliza en la presente memoria los términos “polinucleótido” y “oligonucleótido” se utilizan de forma intercambiable e incluyen formas poliméricas de nucleótidos de cualquier longitud, ya sea desoxirribonucleótidos o ribonucleótidos, o sus análogos. Los polinucleótidos pueden tener cualquier estructura tridimensional y pueden realizar cualquier función, conocida o desconocida. Los siguientes son ejemplos no taxativos de polinucleótidos: un gen o fragmento de gen, exones, intrones, ARN mensajero (ARNm), ARN de transferencia, ARN ribosómico, ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, plásmidos, vectores, ADN aislado de cualquier secuencia, ARN aislado de cualquier secuencia, sondas de ácido nucleico y cebadores. Un polinucleótido puede comprender nucleótidos modificados, como nucleótidos metilados y análogos de nucleótidos. Si estuvieran presentes, las modificaciones de la estructura del nucleótido se pueden impartir antes o después de ensamblar el polímero. La secuencia de nucleótidos puede interrumpirse mediante componentes no nucleótidos. Se puede modificar aún más un polinucleótido después de la polimerización, por ejemplo mediante la conjugación con un componente marcador. El término también incluye moléculas de doble hebra o de hebra única. A menos que se especifique lo contrario o se requiera otra cosa, cualquier realización de esta invención que sea un polinucleótido abarca tanto a la forma de doble hebra como a cada una de dos formas complementarias de hebra única que se sabe o supone que formarían una forma de doble hebra.

Un polinucleótido está compuesto de una secuencia específica de cuatro bases de nucleótidos: adenina (A); citosina (C); guanina (G); timina (T); y uracilo (U) para guanina cuando el polinucleótido es ARN. El término “secuencia de polinucleótidos” es la representación alfabética de una molécula de polinucleótidos. Esta representación alfabética se puede ingresar en bases de datos en una computadora que tenga una unidad de procesamiento central y utilizarse para aplicaciones bioinformáticas como la genómica funcional y la búsqueda de homología.

Un “gen” incluye un polinucleótido que contiene al menos un marco abierto de lectura que es capaz de codificar un polipéptido o proteína particular después de ser transcripto y traducido. Cualquiera de las secuencias de polinucleótidos aquí descritas se puede utilizar para identificar fragmentos más grandes o secuencias codificantes de longitud completa del gen con que están asociados. Los entendidos en la técnica conocen procedimientos para aislar secuencias de fragmentos más grandes, algunos de los que se describen aquí.

Un “producto génico” incluye un aminoácido (por ejemplo, péptido o polipéptido) generado cuando se transcribe y traduce un gen.

Tal como se utiliza en la presente memoria un “polinucleótido se corresponde con” otro (un primer) polinucleótido si se relaciona con el primer polinucleótido mediante cualquiera de las siguientes relaciones:

1) El segundo polinucleótido comprende el primer polinucleótido y el segundo polinucleótido codifica un producto génico.

2) El segundo polinucleótido es 5' o 3' del primer polinucleótido en ADNc, ARN, ADN genómico, o fragmentos de cualquiera de estos polinucleótidos. Por ejemplo, un segundo polinucleótido puede ser un fragmento de un gen que incluye el polinucleótido primero y segundo. Los polinucleótido primero y segundo están relacionados por ser componentes del gen que codifica un producto génico, como una proteína o anticuerpo. Sin embargo, no es necesario que el segundo polinucleótido comprenda o se superponga con el primer polinucleótido incluido en la definición de “correspondiente a” que se utiliza en la presente. Por ejemplo, el primer polinucleótido puede ser un fragmento de una región 3' no traducida del segundo polinucleótido. Los polinucleótidos primero y segundo pueden ser fragmentos de un gen que codifica un producto génico. El segundo polinucleótido puede ser un exón del gen mientras que el primer polinucleótido puede ser un intrón del gen.

3) El segundo polinucleótido es el complemento del primer polinucleótido.

Una “sonda”, cuando se utiliza en el contexto de manipulación de polinucleótidos, incluye un oligonucleótido que se proporciona como reactivo para detectar un objetivo presente en una muestra de interés mediante la hibridación con el objetivo. Generalmente, una sonda comprenderá un marcador o una forma de adjuntar un marcador, ya sea antes

o después de la reacción de hibridación. Las marcas apropiadas incluyen, aunque no taxativamente, radioisótopos, fluorocromos, compuestos quimioluminiscentes, tintes y proteínas, incluidas las enzimas.

Un “cebador” incluye un polinucleótido corto, generalmente con un grupo libre de 3'-OH que se une a un objetivo o “patrón” presente en una muestra de interés mediante hibridación con el objetivo, promoviendo después la polimerización de un polinucleótido complementario al objetivo. Una “reacción en cadena de la polimerasa” (“PCR”, por sus siglas en inglés) es una reacción en la que se hacen copias replicadas de un polinucleótido objetivo utilizando un “par de cebadores” o “juego de cebadores” que consisten en un cebador “aguas arriba” y un cebador “aguas abajo”, y un catalizador de polimerización, como una ADN polimerasa y generalmente una enzima polimerasa termoestable. Los procedimientos de PCR son conocidos en la técnica y se enseñan, por ejemplo, en MacPherson *et al.*, IRL Press en Oxford University Press (1991)). En la presente, todos los procedimientos para producir copias replicadas de un polinucleótido, como PCR o clonación de genes, se denominan colectivamente “replicación”. Un cebador también se puede utilizar como sonda en reacciones de hibridación, como en análisis Southern o Northern blot (ver, por ejemplo, Sambrook, J., Fritsh, E. R, y Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989*).

El término “ADNc” incluye ADN complementario, es decir moléculas ARNm presentes en una célula u organismo transformado en ADNc con una enzima como transcriptasa inversa. Una “biblioteca de ADNc” incluye una colección de moléculas de ARNm presentes en una célula u organismo, convertida en moléculas de ADNc con la enzima transcriptasa inversa, que se inserta luego en “vectores” (otras moléculas de ADN que pueden seguir replicando después de agregar ADN extraño). Los ejemplos de vectores para bibliotecas incluyen bacteriófagos, virus que infectan bacterias (por ejemplo, fago lambda). Se puede sondear la biblioteca en busca del ADNc específico (y por tanto el ARNm) de interés.

Un “vehículo de suministro de genes” incluye una molécula que es capaz de insertar uno o más polinucleótidos en una célula receptora. Son ejemplos de vehículos de suministro de genes los liposomas, polímeros biocompatibles, incluidos polímeros naturales y polímeros sintéticos; lipoproteínas; polipéptidos; polisacáridos; lipopolisacáridos, envolturas virales artificiales; partículas de metal; y bacterias, virus y vectores virales, como baculovirus, adenovirus y retrovirus, bacteriófagos, cósmidos, plásmidos, vector fúngicos y otros vehículos de recombinación utilizados generalmente en la técnica, que se han descrito para la replicación y/o expresión en varios huéspedes eucariotas y procariotas. Los vehículos de suministro de genes se pueden utilizar para la replicación del polinucleótido insertado, la terapia de genes y la expresión de polipéptidos y proteínas simplemente.

Un “vector” incluye una molécula de ácido nucleico auto-replicante que transfiere un polinucleótido insertado en y/o entre las células receptoras. El término incluye vectores que funcionan principalmente para la inserción de una molécula de ácido nucleico en una célula, vectores de replicación que funcionan principalmente para la replicación de ácido nucleico y vectores de expresión que funcionan para la transcripción y/o traducción de ADN o ARN. También comprende vectores que proporcionan más de una de las funciones anteriores.

Una “célula receptora” incluye cualquier célula individual o cultivo celular que puede ser o ha sido un receptor de vectores o de la incorporación de moléculas de ácido nucleico exógenas, polinucleótidos y/o proteínas. También comprende la progenie de una única célula. La progenie no tiene necesariamente que ser completamente idéntica (en morfología o en el complemento genómico o de ADN total) a la célula parental original debido a la mutación natural, accidental o deliberada. Las células pueden ser procariotas o eucariotas, e incluyen, aunque no taxativamente, células bacterianas, células de levadura, células de insecto, células de animales y células de mamíferos, por ejemplo, células murinas, de rata, simio o ser humano.

El término “genéticamente modificado” incluye una célula que contiene y/o expresa un gen o una secuencia de ácido nucleico extraños que a su vez modifica el genotipo o fenotipo de la célula o de su progenie. Este término incluye cualquier adición, eliminación o interrupción de los nucleótidos endógenos de una célula.

Como se usa en la presente, “expresión” incluye el proceso mediante el que se transcriben los polinucleótidos a ARNm y se los traduce a péptidos, polipéptidos o proteínas. Si el polinucleótido deriva de ADN genómico, la expresión puede incluir el corte y empalme del ARNm, si se selecciona un receptor eucariota apropiado. Los elementos reguladores necesarios para la expresión incluyen secuencias de promotores para unir la ARN polimerasa y las secuencias de iniciación de transcripción para la unión del ribosoma. Por ejemplo, un vector de expresión vectorial incluye un promotor como el promotor lac y para la iniciación de la transcripción la secuencia Shine-Dalgarno y el codón de inicio AUG (Sambrook, J., Fritsh, E. E, y Maniatis, T. *Molecular Cloning: A Laboratory Manual. 2nd, ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989*). De forma similar, un vector de expresión eucariota incluye un promotor heterólogo u homólogo para la ARN polimerasa II, una señal de poliadenilación aguas abajo, el codón de inicio AUG y un codón de terminación para desengancharse del ribosoma. Dichos vectores se pueden obtener comercialmente o ensamblarse mediante las secuencias descritas en procedimientos que son conocidos en la técnica, por ejemplo, los procedimientos descritos a continuación para construir vectores en general.

“Expresado diferencialmente”, cuando se aplica a un gen, incluye la producción diferencial de ARNm transcrito desde un gen o un producto proteico codificado por un gen. un gen expresado diferencialmente puede estar sobreexpresado o subexpresado en comparación con el nivel de expresión de una célula normal o de control. En un aspecto,

incluye un diferencial que es 2,5 veces, preferentemente 5 veces o preferentemente 10 veces más alto o más bajo que el nivel de expresión detectado en una muestra de control. El término “expresado diferencialmente” también incluye secuencias de nucleótidos en una célula o tejido que se expresan mientras que en la célula de control están en silencio o que no se expresan mientras que en la célula de control sí lo hacen.

5

El término “polipéptido” incluye un compuesto de dos o más subunidades de aminoácidos, análogos de aminoácidos o peptidomiméticos. Las subunidades pueden estar unidas por enlaces peptídicos. En otra realización, la subunidad puede estar unida por otros enlaces, por ejemplo, éster, éter, etc. Como se utiliza en la presente, el término “aminoácido” incluye aminoácidos naturales y/o aminoácidos no naturales o sintéticos, incluidos la glicina y tanto el isómero óptico D como el L y análogos de aminoácidos y peptidomiméticos. Comúnmente a un péptido de tres o más aminoácidos se lo llama un oligopéptido. A las cadenas peptídicas mayores a tres o más aminoácidos se las llama un polipéptido o una proteína.

10

La “hibridación” incluye una reacción en la que uno o más polinucleótidos reaccionan para formar un complejo que se estabiliza mediante un enlace de hidrógeno entre las bases de los residuos nucleotídicos. El enlace por puente de hidrógeno puede ocurrir mediante emparejamiento de bases Watson-Crick, de Hoogsteen o mediante cualquier otra forma específica de secuencia. El complejo puede comprender dos hebras que forman una estructura dúplex, tres o más hebras que forman un complejo de hebras múltiples, una hebra única autohibridante o cualquier combinación de estas. Una reacción de hibridación puede constituir una etapa de un proceso más extenso, como el inicio de una reacción PCR o la escisión enzimática de un polinucleótido por una ribozima.

15

20

Las reacciones de hibridación se pueden realizar bajo condiciones de diferente “rigurosidad”. La rigurosidad de una reacción de hibridación incluye la dificultad con la que dos moléculas de ácido nucleico cualquiera se hibridarán una con otra. Bajo condiciones rigurosas, las moléculas de ácido nucleico que son idénticas entre sí en un 60%, 65%, 70% o 75% permanecen hibridadas una a la otra, mientras que las moléculas con porcentajes menores de identidad no pueden permanecer hibridadas. Un ejemplo preferido y no taxativo de condiciones de hibridación altamente rigurosas son hibridaciones en cloruro de sodio/citrato de sodio (SSC) 6X a aproximadamente 45°C, seguido de uno o más lavados en 0,2xSSC, 0,1% SDS a 50°C, preferentemente a 55°C, más preferentemente a 60°C; y más preferentemente aún a 65°C.

25

30

Cuando la hibridación ocurre en una configuración antiparalela entre dos polinucleótidos de hebra única, la reacción se llama “apareamiento” y esos polinucleótidos se los describe como “complementarios”. Un polinucleótido de doble hebra puede ser “complementario” u “homólogo” a otro polinucleótido, si la hibridación puede ocurrir entre una de las hebras del primer polinucleótido y el segundo. La “complementariedad” o la “homología” (el grado en que un polinucleótido es complementario con otro) es cuantificable en función de la proporción de bases en hebras opuestas que se prevé se enlazarán mediante un enlace de hidrógeno una con la otra, de conformidad con reglas de emparejamiento de bases generalmente aceptadas.

35

Un “anticuerpo” incluye una molécula de inmunoglobulina capaz de unir un epítipo presente en un antígeno. Como se utiliza en la presente, el término abarca no solo moléculas de inmunoglobulina intactas como anticuerpos monoclonales y policlonales, sino también anticuerpos antiidiotípicos, mutantes, fragmentos, proteínas de fusión, anticuerpos bi-específicos, proteínas humanizadas y modificaciones de la molécula de inmunoglobulina que comprende un sitio de reconocimiento de antígeno de la especificidad requerida.

40

Como se utiliza en la presente, el término “diabetes tipo I” incluye un trastorno no contagioso en el que un individuo no es capaz de producir insulina. Los síntomas de la diabetes tipo I incluyen pérdida de peso, irritabilidad, micción frecuente, sed excesiva, hambre extrema, debilidad o fatiga y náuseas o vómitos.

45

Como se utiliza en la presente, el término “célula NKT” incluye células que se identifican por la expresión tanto de marcadores de células asesinas naturales (NK, por sus siglas en inglés) como de un receptor de células T invariante. Dichas células son CD16i⁺, y también se las conoce como células V α 24J α Q T.

50

Como se utiliza aquí, el término “tejido diabético” o “células diabética” o “tejido diabético tipo I” o “célula diabética tipo I” incluye un tejido o célula de un individuo que sufre diabetes tipo I, en el que el tejido mismo está implicado en la sintomatología de la diabetes tipo I. Un ejemplo de tejido diabético tipo I son las células beta pancreáticas (células productoras de insulina) o sangre. Un “tejido no diabético” o una “célula no diabética” incluye un tejido o célula que pertenece a un individuo que no sufre diabetes tipo I o de un individuo que sufre diabetes tipo I, pero en el que los propios tejido o célula no están implicados en la sintomatología de la enfermedad y/o no está afectado por la presencia de la enfermedad. Por lo tanto, una “célula diabética NKT” es una célula NKT tomada de un individuo diabético. Una “célula no diabética NKT” es una célula NKT tomada de un individuo no diabético.

55

60

Como se utiliza aquí, el término “marcador” incluye un polinucleótido o molécula polipeptídica que está presente o ausente o aumentada o disminuida en cantidad o actividad en individuos que sufren de la diabetes tipo I, o en células implicadas en la diabetes tipo I. El cambio relativo en cantidad o actividad del marcador se relaciona con la incidencia o el riesgo de incidencia de la diabetes tipo I.

65

Como se utiliza en la presente, el término “panel de marcadores” incluye un grupo de marcadores, en el que cantidad o actividad de cada miembro se relaciona con la incidencia o riesgo de incidencia de la diabetes tipo I.

En ciertas realizaciones, un panel de marcadores puede incluir solamente los marcadores cuya cantidad o actividad es mayor o menor en individuos que sufren de la diabetes tipo I o células implicadas en este trastorno. En otras realizaciones, un panel de marcadores puede incluir solamente los marcadores presentes en un tipo específico de tejido que se relaciona con la incidencia o riesgo de incidencia de la diabetes tipo I.

En las siguientes subsecciones se describen en mayor detalle varios aspectos:

I. Moléculas de ácido nucleico aisladas

Un aspecto se refiere a moléculas de ácido nucleico aisladas que son ellas mismas los marcadores genéticos (por ejemplo, ARNm) de la invención, o codifican los marcadores polipeptídico de la invención o fragmentos de los mismos. Otro aspecto se refiere a fragmentos de ácido nucleico aislados suficientes como para utilizarlos como sondas de hibridación para identificar las moléculas de ácido nucleico que codifican los marcadores divulgados en una muestra, así como también fragmentos nucleotídico para utilizar como cebadores PCR para la amplificación o mutación de las moléculas de ácido nucleico que codifican a los marcadores divulgados. Como se utiliza en la presente, el término “molécula de ácido nucleico” incluye moléculas de ADN (por ejemplo, ADNc o ADN genómico) y moléculas de ARN (por ejemplo ARNm) y análogos de ADN o ARN generados utilizando análogos de nucleótidos. La molécula de ácido nucleico puede ser de hebra única o de doble hebra, pero preferentemente es ADN de doble hebra.

El término “molécula de ácido nucleico aislada” incluye moléculas de ácido nucleico que están separadas de otras moléculas de ácido nucleico presentes en la fuente natural del ácido nucleico. Por ejemplo, con respecto al ADN genómico, el término “aislado” incluye moléculas de ácido nucleico que están separadas del cromosoma al que el ADN genómico está asociado naturalmente. Preferentemente, un ácido nucleico “aislado” está libre de secuencias que naturalmente flanquean al ácido nucleico (es decir, secuencias localizadas en los extremos 5' y 3' del ácido nucleico) en el ADN genómico del organismo del que deriva el ácido nucleico. Por ejemplo, en varias realizaciones, la molécula marcadora de ácido nucleico aislada o molécula de ácido nucleico que codifica un polipéptido marcador puede contener menos de aproximadamente 5 kb, 4kb, 3kb, 2kb, 1kb, 0,5kb ó 0,1kb de secuencias de nucleótidos que naturalmente flanquean a la molécula de ácido nucleico en el ADN genómico de la célula de la que deriva el ácido nucleico. Además, una molécula de ácido nucleico “aislada”, por ejemplo una molécula de ADNc, puede estar sustancialmente libre de otro material celular o medio de cultivo cuando se la produce mediante técnicas recombinantes o sustancialmente libre de precursores químicos u otros productos químicos cuando se la sintetiza químicamente.

Se puede aislar una molécula de ácido nucleico divulgada, por ejemplo, una molécula de ácido nucleico que tiene la secuencia de nucleótidos de uno de los genes establecidos en las Tablas 1-13 o una parte de estos, utilizando técnicas de biología molecular estándar y la información de secuencia que aquí se proporciona. Utilizando la totalidad o parte de la secuencia de ácido nucleico de uno de los genes establecidos en las Tablas 1-13 como una sonda de hibridación, se puede aislar un gen marcador o una molécula de ácido nucleico que codifica un polipéptido marcador, utilizando técnicas de hibridación y clonación estándar (por ejemplo, como se describe en Sambrook, J., Fritsh, E. R, y Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989).

Se puede amplificar un ácido nucleico divulgado utilizando ADNc, ARNm o alternativamente, ADN genómico, como patrón y cebadores oligonucleótidos apropiados de conformidad con técnicas de amplificación PCR estándar. El ácido nucleico así amplificado se puede clonar en un vector apropiado y caracterizar mediante el análisis de la secuencia de ADN. Además, los oligonucleótidos correspondientes a las secuencias de nucleótidos marcadoras o secuencias de nucleótidos que codifican un marcador divulgado se pueden preparar mediante técnicas sintéticas estándar, por ejemplo, utilizando un sintetizador de ADN automático.

En otra realización preferida, una molécula de ácido nucleico aislada comprende una molécula de ácido nucleico que es un complemento de la secuencia de nucleótidos de un marcador divulgado (por ejemplo, un gen establecido en las Tablas 1-13) o una parte de cualquiera de estas secuencias de nucleótidos. Una molécula de ácido nucleico que es complementaria de dicha secuencia de nucleótidos es aquella que es suficientemente complementaria a la secuencia de nucleótidos como para poderse hibridar con la secuencia nucleotídica, formando de ese modo un dúplex estable.

Además, la molécula de ácido nucleico puede comprender solamente una parte de la secuencia de ácido nucleico de un ácido nucleico marcador divulgado o un gen que codifica un polipéptido marcador divulgado, por ejemplo, un fragmento que puede ser utilizado como sonda o cebador. La sonda/el cebador generalmente comprende oligonucleótidos sustancialmente purificados. El oligonucleótido generalmente comprende una región de la secuencia de nucleótidos que se hibrida bajo condiciones rigurosas con al menos aproximadamente 7 ó 15, preferentemente aproximadamente 20 o 25, más preferentemente aproximadamente 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 400 ó más nucleótidos consecutivos de un ácido nucleico marcador o un ácido nucleico que codifica un polipéptido marcador divulgado.

Se pueden utilizar sondas basadas en la secuencia de nucleótidos de un gen marcador o de una molécula de ácido nucleico que codifica un polipéptido marcador divulgado para detectar secuencias transcritas o genómicas correspondientes a el/los genes marcadores y/o polipéptidos marcadores divulgados. En realizaciones preferidas, la sonda comprende un grupo marcador unido a esta, por ejemplo, el grupo marcador puede ser un radioisótopo, un compues-

to fluorescente, una enzima o un cofactor de una enzima. Dichas sondas se pueden utilizar como parte de un kit de diagnóstico para identificar células o tejidos que expresan erróneamente (por ejemplo, subexpresan o sobreexpresan) un polipéptido marcador divulgado, o que tienen más o menos copias de un gen marcador divulgado. Por ejemplo, se puede detectar el nivel de un ácido nucleico que codifica un marcador polipéptido en una muestra de células de un individuo, se puede determinar la cantidad de ARNm transcripto de un gen que codifica un polipéptido marcador o se puede evaluar la presencia de mutaciones o supresiones de un gen marcador divulgado.

Asimismo, se describen las moléculas de ácido nucleico que difieren de las secuencias de ácido nucleico de los genes establecidos en las Tablas 1 y 4, debido a la degeneración del código genético y que por lo tanto codifican las mismas proteínas codificadas por los genes que se muestran en las Tablas 1 y 4.

Además de la secuencia de nucleótidos de los genes establecidos en las Tablas 1 y 4, los entendidos en la técnica podrán apreciar que los polimorfismos de la secuencia de ADN que conducen a cambios en las secuencias de aminoácidos de las proteínas codificadas por los genes establecidos en las Tablas 1 y 4 pueden existir en una población (por ejemplo, la población humana). Dicho polimorfismo genético en los genes establecidos en las Tablas 1 y 4 pueden existir entre individuos de una población debido a la variación alélica natural. Un alelo es un gen de un grupo de genes que ocurren alternativamente en un determinado locus genético. Además, se podrá apreciar que los polimorfismos de ADN que afectan los niveles de expresión de ARN también pueden existir y afectar la totalidad del nivel de expresión del gen (por ejemplo, afectando la regulación o degradación). Como se utiliza en la presente, la frase “variante alélica” incluye una secuencia de nucleótidos que ocurre en un locus determinado o un polipéptido codificado por la secuencia nucleotídica. Como se utilizan en la presente, los términos “gen” y “gen recombinante” se refieren a moléculas de ácido nucleico que incluyen un marco de lectura abierto que codifica un polipéptido marcador divulgado.

Se pueden aislar moléculas de ácido nucleico que corresponden a variantes alélicas naturales y homólogas de los genes marcadores o genes que codifican las proteínas marcadoras divulgadas, basándose en su homología con los genes establecidos en las Tablas 1 y 4, utilizando los ADNc aquí divulgados o parte de estos, como una sonda de hibridación, de conformidad con técnicas de hibridación estándar, bajo condiciones de hibridación rigurosas. Las moléculas de ácido nucleico correspondientes a variantes alélicas naturales y homólogos de los genes marcadores divulgados se pueden aislar aún más mediante el mapeo del mismo cromosoma o locus que los genes marcadores o genes que codifican las proteínas marcadoras divulgadas.

En otra realización, una molécula de ácido nucleico aislada tiene al menos 15, 20, 25, 30, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000 ó más nucleótidos en longitud y se hibrida bajo condiciones rigurosas con una molécula de ácido nucleico correspondiente a una secuencia de nucleótidos de un gen marcador o gen que codifica una proteína marcadora divulgada. Como se utiliza en la presente, el término “se hibrida bajo condiciones rigurosas” describe condiciones de hibridación y lavado bajo las que las secuencias nucleotídicas, homólogas al menos en un 60% entre sí, generalmente permanecen hibridadas una a la otra. Preferentemente, las condiciones son tales que las secuencias homólogas entre sí en al menos aproximadamente un 70%, más preferentemente al menos aproximadamente un 80%, aún más preferentemente al menos aproximadamente un 85% o 90%, generalmente permanecen hibridadas una a la otra. Dichas condiciones rigurosas son conocidas por los entendidos en la técnica y se pueden encontrar en *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1 989), 6.3.1-6.3.6. Un ejemplo preferido, no taxativo, de condiciones de hibridación rigurosas es la hibridación en cloruro de sodio/citrato de sodio (SSC) 6X a aproximadamente 45°C, seguido de uno o más lavados en 0,2xSSC, 0,1% SDS a 50°C, preferentemente a 55°C, más preferentemente a 60°C, y más preferentemente aún a 65°C. Preferentemente, una molécula de ácido nucleico aislada que se hibrida bajo condiciones rigurosas con la secuencia de uno de los genes establecidos en las Tablas 1-13 corresponde a una molécula de ácido nucleico existente naturalmente. Como se utiliza en la presente, una molécula de ácido nucleico “existente naturalmente” incluye una molécula de ADN o ARN que tiene una secuencia de nucleótidos que existe en la naturaleza (por ejemplo, codifica una proteína natural).

Además de variantes alélicas existentes naturalmente del gen marcador y el gen que codifica una secuencia de proteína marcadora divulgada que puede existir en la población, el entendido en la materia apreciará además que se pueden introducir cambios mediante mutación en las secuencias de nucleótidos de los genes marcadores o genes que codifican las proteínas marcadoras divulgadas, lo que conduce a cambios en la secuencia de aminoácidos de las proteínas codificadas, sin alterar la actividad funcional de estas proteínas. Por ejemplo, se pueden preparar sustituciones de nucleótidos que conducen a sustituciones de aminoácidos en residuos de aminoácidos “no esenciales”. Un residuo de aminoácido “no esencial” es un residuo que puede ser alterado desde la secuencia silvestre de una proteína sin alterar la actividad biológica, mientras que para la actividad biológica se requiere un residuo de aminoácido “esencial”. Por ejemplo, se predice que los residuos de aminoácidos que se conservan entre variantes alélicas u homólogos de un gen (por ejemplo, entre los homólogos de un gen de diferentes especies) son particularmente resistentes a la alteración.

Por consiguiente, otro aspecto se relaciona con moléculas de ácido nucleico que codifican una proteína marcadora divulgada que contiene cambios en los residuos de aminoácidos que no son esenciales para la actividad. Dichas proteínas difieren de las proteínas marcadoras codificadas por los genes establecidos en las Tablas 1 y 4 en la secuencia de aminoácidos y sin embargo mantienen la actividad biológica. En una realización, la proteína comprende una secuencia de aminoácidos homóloga en al menos aproximadamente un 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% ó más receptor de una proteína marcadora divulgada.

Se puede crear una molécula de ácido nucleico aislada que codifica una proteína homóloga a una proteína marcadora divulgada, introduciendo una o más sustituciones, adiciones o supresiones de nucleótidos en la secuencia de nucleótidos del gen que codifica la proteína marcadora, de modo tal que se introducen una o más sustituciones, adiciones o supresiones de aminoácidos en la proteína codificada. Las mutaciones se pueden introducir en los genes (por ejemplo, un gen establecido en las Tablas 1 y 4) mediante técnicas estándar, como mutagénesis dirigida a un sitio y mutagénesis mediada por PCR. Preferentemente, las sustituciones de aminoácidos conservadoras se hacen en uno o más residuos de aminoácidos no esenciales previstos. Una "sustitución de aminoácido conservadora" es aquella en la que el residuo de aminoácido es remplazado por un residuo de aminoácido que tiene una cadena lateral similar. Las familias de residuos de aminoácidos que tienen cadenas laterales similares han sido definidas en la técnica. Estas familias incluyen aminoácidos con cadenas laterales básicas (por ejemplo, lisina, arginina, histidina), cadenas laterales ácidas (por ejemplo ácido aspártico, ácido glutámico), cadenas laterales polares no cargadas (por ejemplo glicina, asparagina, glutamina, serina, treonina, tirosina, cisteína), cadenas laterales no polares (por ejemplo, alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptofano), cadenas laterales beta ramificadas (por ejemplo, treonina, valina, isoleucina) y cadenas laterales aromáticas (por ejemplo, tirosina, fenilalanina, triptofano, histidina). Alternativamente, las mutaciones se pueden introducir en forma aleatoria en toda o parte de una secuencia codificante de un gen, por ejemplo mediante mutagénesis de saturación, y se pueden analizar los mutantes resultantes en busca de actividad biológica para identificar los mutantes que retienen actividad. Después de la mutagénesis, la proteína codificada se puede expresar en forma recombinante y se puede determinar la actividad de la proteína.

En otra realización, las moléculas de ácido nucleico descritas se pueden modificar en la porción de base, porción de azúcar o en el esqueleto fosfato para mejorar, por ejemplo, la estabilidad, la hibridación o solubilidad de la molécula. Por ejemplo, el esqueleto de desoxirribosa y fosfato de las moléculas de ácido nucleico se puede modificar para generar ácidos nucleicos peptídicos (ver Hyrup B. *et al.* (1996) *Bioorganic & Medicinal Chemistry* 4 (1): 5-23). Como se utiliza en la presente, los términos "ácidos nucleicos peptídicos" o "PNA" (por sus siglas en inglés), se refieren a miméticos de ácido nucleico, por ejemplo, miméticos de ADN, en las que se reemplaza el esqueleto de desoxirribosa y fosfato por un esqueleto pseudopéptido y solo se retienen las cuatro nucleobases naturales. Se ha demostrado que el esqueleto de PNA neutro permite la hibridación específica con el ADN y ARN bajo condiciones de baja potencia iónica. La síntesis de oligómeros de PNA se puede realizar utilizando protocolos estándar de síntesis de péptidos en fase sólida como se describe en Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *et al.* *Proc. Natl. Acad. Sci.* 93: 14670-675.

Los PNA se pueden utilizar en aplicaciones terapéuticas y diagnósticas. Por ejemplo, los PNA se pueden utilizar como agentes antisentido o anfígenos para la modulación de secuencia específica de la expresión de genes, induciendo por ejemplo, la detención de la transcripción o traducción o inhibiendo la replicación. También se pueden utilizar los PNA de las moléculas de ácido nucleico divulgadas (por ejemplo, un gen establecido en las Tablas 1-13) en el análisis de mutaciones de un solo par de bases en un gen (por ejemplo, mediante abrazaderas PCR dirigidas por PNA); como "enzimas de restricción artificial" cuando se las utiliza en combinación con otras enzimas, (por ejemplo, S1 nucleasas (Hyrup B. (1996) *supra*)); o como sondas o cebadores para secuenciación de ADN o hibridación/Hyrup B. *et al.* (1996) *supra*; Perry-O'Keefe *supra*).

En otra realización, se pueden modificar los PNA (por ejemplo, para aumentar su estabilidad o captación celular) uniendo lipofílicos u otros grupos ayudantes al PNA, mediante la formación de quimeras de PNA-ADN o mediante el uso de liposomas u otras técnicas de administración de fármacos conocidas en la técnica. Por ejemplo, se pueden generar quimeras de las moléculas de ácido nucleico divulgadas que pueden combinar las propiedades ventajosas del PNA y el ADN. Tales quimeras permiten que las enzimas de reconocimiento de ADN, (por ejemplo, polimerasas de ARN H y ADN) interactúen con la porción de ADN mientras que la porción de PNA proporcionaría una especificidad y afinidad de unión altas. Las quimeras de PNA-ADN se pueden unir utilizando enlazadores de longitud apropiada seleccionados en términos de apilamiento de bases, el número de enlaces entre las nucleobases y la orientación (Hyrup B. (1996) *supra*). La síntesis de PNA-ADN se puede realizar como se describe en Hyrup B. (1996) *supra* y Finn P. J. *et al.* (1996) *Nucleic Acids Res.* 24 (17): 3357-63. Por ejemplo, se puede sintetizar una cadena de ADN en un soporte sólido utilizando química de acoplamiento de fosforamidita estándar y análogos nucleósidos modificados, por ejemplo, se puede utilizar 5'-(4-metoxitritil)amino-5'-desoxitimidina fosforamidita, entre el PNA y el extremo 5' del ADN (Mag, M *et al.* (1989) *Nucleic Acids Res.* 17: 5973-88). Luego se acoplan los monómeros PNA paso por paso para producir una molécula quimérica con un segmento PNA 5' y un segmento ADN 3' (Finn P. J. *et al.* (1996) *supra*). Alternativamente, las moléculas quiméricas se pueden sintetizar con un segmento de ADN 5' y un segmento de PNA 3' (Peterser, K.H. *et al.* (1975) *Bioorganic Med. Chem. Lett.* 5: 1119-11124).

En otras realizaciones, los oligonucleótidos pueden incluir otros grupos anexos como péptidos (por ejemplo, dirigidos a receptores de células receptoras *in vivo*), o agentes que faciliten el transporte a través de la membrana celular (ver, por ejemplo, Letsinger *et al.* (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; Publicación PCT No. W088/09810) o la barrera sangre-cerebro (ver por ejemplo, Publicación PCT No. W089/10134). Además, los oligonucleótidos se pueden modificar con agentes de escisión desencadenados por hibridación (Ver, por ejemplo, Krol *et al.* (1988) *Bio-Techniques* 6:958-976) o agentes intercalantes. (Ver, por ejemplo, Zon (1988) *Pharm. Res.* 5:539-549). A tales efectos, el oligonucleótido se puede conjugar con otra molécula, (por ejemplo, un péptido, un agente reticulante desencadenado por hibridación, un agente de transporte o un agente de escisión desencadenado por hibridación). Finalmente, el oligonucleótido puede marcarse detectablemente, ya sea que la marca se detecte mediante la adición de otro reactivo (por ejemplo, un sustrato para una marca enzimática) o que se detecte inmediatamente al hibridarse el nucleótido (por ejemplo, una marca radioactivo o una marca fluorescente (por ejemplo, una baliza molecular, como se describe en la Patente estadounidense 5,876,930).

II. Anticuerpos y proteínas aisladas

Un aspecto se refiere a proteínas marcadoras aisladas y partes biológicamente activas de estas, así como también fragmentos polipeptídicos apropiados para utilizar como inmunógenos para elevar los anticuerpos de proteínas anti-marcadoras. En una realización, las proteínas marcadoras nativas se pueden aislar de las fuentes de tejido o células mediante un esquema apropiado de purificación utilizando técnicas estándar de purificación de proteínas. En otra realización, las proteínas marcadoras se producen mediante técnicas de ADN recombinante. De manera alternativa a la expresión recombinante, se puede sintetizar químicamente una proteína marcadora o polipéptido marcador utilizando técnicas estándar de síntesis de péptidos.

Una proteína “aislada” o “purificada” o una parte biológicamente activa de esta está sustancialmente libre de material celular u otras proteínas contaminantes de la fuente de tejido o células de la que deriva la proteína marcadora, o está sustancialmente libre de precursores químicos u otros productos químicos cuando se la sintetiza químicamente. La expresión “sustancialmente libre de material celular” incluye preparaciones de proteína marcadora en las que la proteína se separa de componentes celulares de las células a partir de las que se aísla o produce en forma recombinante. En una realización, la expresión “sustancialmente libre de material celular” incluye preparaciones de proteínas marcadoras que tienen menos de aproximadamente un 30% (por peso en seco) de proteínas no marcadoras (también llamada en la presente “proteína contaminante”), más preferentemente menos de aproximadamente un 20% de proteínas no marcadoras, más preferentemente menos de aproximadamente un 10% de proteínas no marcadoras y más preferentemente aún menos de aproximadamente un 5% de proteínas no marcadoras. Cuando la proteína marcadora o la parte biológicamente activa de esta se produce de forma recombinante, también está preferentemente de manera sustancial libre de medio de cultivo, es decir el medio de cultivo representa menos del 20% aproximadamente, más preferentemente menos del 10% aproximadamente y más preferentemente aún menos del 5% aproximadamente del volumen de la preparación de la proteína.

La expresión “sustancialmente libre de precursores químicos u otros productos químicos” incluye preparaciones de proteínas marcadoras en las que se separa la proteína de los precursores químicos u otros productos químicos implicados en la síntesis de la proteína. En una realización, la expresión “sustancialmente libre de precursores químicos u otros productos químicos” incluye preparaciones de proteínas que tienen menos de aproximadamente 30% (por peso en seco) de precursores químicos o productos químicos no proteicos, más preferentemente menos de aproximadamente el 20% de precursores químicos o productos químicos no proteicos, más preferentemente menos de aproximadamente el 10% de precursores químicos o productos químicos no proteicos y más preferentemente aún menos de aproximadamente el 5% de precursores químicos o productos químicos no proteicos.

Como se utiliza en la presente, una “parte biológicamente activa” de una proteína marcadora incluye un fragmento de una proteína marcadora que comprende secuencias de aminoácidos suficientemente homólogas o derivadas de la secuencia de aminoácidos de la proteína marcadora, que incluyen menos aminoácidos que las proteínas marcadoras en toda su longitud y tiene al menos una actividad de la proteína marcadora. Generalmente, las partes biológicamente activas comprenden un dominio o motivo con al menos una actividad de la proteína marcadora. Una parte biológicamente activa de una proteína marcadora puede ser un polipéptido que tiene, por ejemplo, 10, 25, 50, 100, 200 o más aminoácidos de longitud. Las partes biológicamente activas de una proteína marcadora se pueden utilizar como objetivos para desarrollar agentes que modulan una actividad mediada por proteína marcadora.

En una realización preferida, la proteína marcadora está codificada por un gen establecido en las Tablas 1 y 4. En otras realizaciones, la proteína marcadora es esencialmente homóloga a una proteína marcadora codificada por un gen establecido en las Tablas 1-13 y retiene la actividad funcional de la proteína marcadora, pero difiere en la secuencia de aminoácidos debido a variaciones alélicas naturales o mutagénesis, como se describe detalladamente en la sección I anterior. Por consiguiente en otra realización, la proteína marcadora es una proteína que comprende una secuencia de aminoácidos homóloga en al menos aproximadamente el 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% o más con respecto a la secuencia de aminoácidos mediante un gen establecido en las Tablas 1 y 4.

Para determinar el porcentaje de identidad de dos secuencias de aminoácidos, se alinean las secuencias a efectos de una comparación óptima (por ejemplo, se puede introducir intervalos en una o ambas de una primera y segunda secuencia de aminoácidos o ácido nucleico para una alineación óptima y se pueden descartar las secuencias no homólogas a efectos de comparación). En una realización preferida, la longitud de una secuencia de referencia alineada a efectos de comparación es al menos un 30%, preferentemente al menos un 40%, más preferentemente al menos un 50%, aun más preferentemente al menos un 60%, aún más preferentemente un 70%, 80% o un 90% de la longitud de la secuencia de referencia. Luego se comparan los residuos de aminoácidos o nucleótidos en las correspondientes posiciones de aminoácidos o posiciones de nucleótidos. Cuando se ocupa una posición en la primera secuencia con el mismo residuo de aminoácido o nucleótido que la posición correspondiente en la segunda secuencia, entonces las moléculas son idénticas en esa posición (como se utiliza en la presente, “identidad” de aminoácido o ácido nucleico es equivalente a “homología” de aminoácido o ácido nucleico). El porcentaje de identidad entre las dos secuencias es una función del número de posiciones idénticas compartidas por las secuencias, teniendo en cuenta el número de intervalos y la longitud de cada intervalo, que es necesario introducir para una alineación óptima de las dos secuencias.

La comparación de secuencias y la determinación del porcentaje de identidad entre dos secuencias se puede lograr utilizando un algoritmo matemático. En una realización preferida, el porcentaje de identidad entre dos secuencias de

aminoácidos se determina utilizando el algoritmo Needleman y Wunsch (*J. Mol. Biol.* (48):444-453 (1970)) que se ha incorporado al programa GAP en el paquete de software GCG (disponible en <http://www.gcg.com>), utilizando ya sea una matriz Blossum 62 o una matriz PAM250 y un peso del intervalo de 16, 14, 12, 10, 8, 6 ó 4 y un peso de longitud de 1, 2, 3, 4, 5 ó 6. En otra realización preferida, el porcentaje de identidad entre dos secuencias de nucleótidos se determina utilizando el programa GAP del paquete de software GCG (disponible en <http://www.gcg.com>), utilizando una matriz NWSgapdna. CMP y un peso del intervalo de 40, 50, 60, 70 u 80 y un peso de longitud de 1, 2, 3, 4, 5 ó 6. En otra realización, el porcentaje de identidad entre dos secuencias de aminoácidos o nucleótidos se determina utilizando el algoritmo de E. Meyers y W. Miller (CABIOS, 4:11-17 (1989)) que ha sido incorporado al programa ALIGN (versión 2.0), utilizando una tabla de residuos de peso PAM120, una penalidad de longitud de intervalo de 12 y una penalidad de intervalo de 4.

Las secuencias de ácido nucleico y proteínas divulgadas pueden utilizarse también como “secuencia de búsqueda” para realizar una búsqueda en las bases de datos públicas a fin de identificar, por ejemplo, otros miembros de la familia o secuencias relacionadas. Dichas búsquedas se pueden realizar utilizando los programas NBLAST y XBLAST (versión 2.0) de Altschul, *et al.* (1990)/*Mol. Biol.* 215:403-10. Las búsquedas de nucleótidos BLAST se pueden realizar con el programa NBLAST, puntuación=100, longitud de palabra=12 para obtener secuencias de nucleótidos homólogas a las moléculas de ácido nucleico de la invención. Las búsquedas de proteínas BLAST se pueden realizar con el programa XBLAST, puntuación=50m, longitud de palabra=3 para obtener secuencias de aminoácidos homólogas a las moléculas de proteína marcadora de la invención. Para obtener alineaciones con intervalos a efectos de hacer una comparación, se puede utilizar Gapped BLAST tal como se describe en Altschul *et al.*, (1997) *Nucleic Acids Res.* 25 (17):3389-3402. Al utilizar los programas BLAST y Gapped BLAST, se pueden utilizar los parámetros predeterminados de los respectivos programas (por ejemplo, XBLAST y NBLAST). Ver <http://www.ncbi.nlm.nih.gov>.

También se proporcionan proteínas marcadoras quiméricas o de fusión. Como se utiliza en la presente, una “proteína quimérica” o “proteína de fusión” marcadora comprende un polipéptido marcador unido operativamente a un polipéptido no marcador. Un “polipéptido marcador” incluye un polipéptido que tiene una secuencia de aminoácidos codificada por un gen establecido en las Tablas 1 y 4, mientras que un “polipéptido no marcador” incluye un polipéptido que tiene una secuencia de aminoácidos correspondiente a una proteína que no es esencialmente homóloga a la proteína marcadora, por ejemplo, una proteína que es diferente a la proteína marcadora y que deriva del mismo organismo o de un organismo diferente. En una proteína de fusión marcadora el polipéptido puede corresponder a toda o una parte de la proteína marcadora. En una realización preferida, una proteína de fusión comprende al menos una parte biológicamente activa de una proteína marcadora. En la proteína de fusión, el término “unido operativamente” indica que el polipéptido marcador y el polipéptido no marcador están fusionados en la misma fase de lectura. El polipéptido no marcador puede estar fusionado al terminal N o terminal C del polipéptido marcador.

Por ejemplo, en una realización, la proteína de fusión es una proteína de fusión marcadora GST en la que las secuencias marcadoras están fusionadas al terminal C de las secuencias GST. Dichas proteínas de fusión pueden facilitar la purificación de proteínas marcadoras recombinantes.

En otra realización, la proteína de fusión es una proteína marcadora que contiene una secuencia de señal heteróloga en su terminal N. En ciertas células receptoras (por ejemplo células receptoras de mamífero), la expresión y/o secreción de proteínas marcadoras se puede aumentar a través del uso de una secuencia de señal heteróloga. Tales secuencias de señales son muy conocidas en la técnica.

Las proteínas de fusión marcadoras divulgadas se pueden incorporar en una composición farmacéutica y administrar a un individuo *in vivo*, como se describe en la presente. Las proteínas de fusión marcadoras se pueden utilizar para afectar la biodisponibilidad del sustrato de una proteína marcadora. El uso de proteínas de fusión marcadoras puede ser útil terapéuticamente para el tratamiento de trastornos (por ejemplo, diabetes tipo I o una afección asociada a la NKT) provocada, por ejemplo, por (i) modificación o mutación aberrante de un gen que codifica la proteína marcadora; (ii) desregulación de un gen que codifica la proteína marcadora; (iii) modificación post traduccional aberrante de una proteína marcadora.

Además, las proteínas de fusión marcadoras divulgadas se pueden utilizar como inmunógenos para producir anticuerpos de proteínas antimarcadoras en un individuo, para purificar los ligandos de proteína marcadora y en ensayos de análisis para identificar moléculas que inhiben la interacción de una proteína marcadora con un sustrato de proteína marcadora.

Preferentemente, una proteína quimérica o de fusión marcadora divulgada se produce mediante técnicas estándar de ADN recombinante. Por ejemplo, los fragmentos de ADN que codifican a las diferentes secuencias de polipéptidos se ligan juntos en la misma fase de lectura, de conformidad con técnicas convencionales, por ejemplo empleando terminales de extremos redondeados o terminales de extremos en punta para la ligación, digestión de una enzima de restricción para proporcionar terminales apropiados, rellenando de los extremos cohesivos según corresponda, tratamiento con fosfatasa alcalina para evitar uniones no deseadas y ligación enzimática. En otra realización, el gen de fusión se puede sintetizar mediante técnicas convencionales incluidos los sintetizadores de ADN automáticos. De manera alternativa, se puede realizar la amplificación por PCR de fragmentos de genes utilizando cebadores de anclaje que dan lugar a protuberancias complementarias entre dos fragmentos consecutivos de genes que a continuación pueden hibridarse y reamplificarse para generar una secuencia quimérica de genes (ver, por ejemplo *Current Protocols in Molecular Biology*, eds. Ausubel *et al.* John Wiley & Sons: 1992). Además, existen muchos vectores de expresión

comercialmente disponibles que ya codifican una porción de fusión (por ejemplo, un polipéptido GST). Un ácido nucleico que codifica una proteína marcada se puede clonar en un vector de expresión tal que la porción de fusión está unida en la misma fase de lectura con la proteína marcadora.

5 Una secuencia de señal se puede utilizar para facilitar la secreción y el aislamiento de la proteína segregada u otras proteínas de interés. Las secuencias de señales generalmente se caracterizan por un núcleo de aminoácidos hidrofóbicos que generalmente están escindidos de la proteína madura durante la secreción en uno o más episodios de escisión. Dichos péptidos de señal contienen sitios de procesamiento que permiten la escisión de la secuencia de señal de las proteínas maduras a medida que pasa a través de la vía de secreción. Así, se describen los polipéptidos que
10 tienen una secuencia de señal, así como los polipéptidos a partir de los cuales la secuencia de señal ha sido escindida proteolíticamente (es decir, los productos de la escisión). En una realización, una secuencia de ácido nucleótico que codifica una secuencia de señal puede unirse operativamente en un vector de expresión, a una proteína de interés, como por ejemplo una proteína que comúnmente no es secretada o que es difícil de aislar. La secuencia de señal dirige la secreción de la proteína, como por ejemplo de un huésped eucariota en el que se transforma el vector de expresión, y
15 la secuencia de señal se escinde posteriormente o al mismo tiempo. Luego la proteína se puede purificar fácilmente del medio extracelular mediante procedimientos reconocidos en la técnica. De manera alternativa, la secuencia de señal se puede unir a la proteína de interés utilizando una secuencia que facilita la purificación, por ejemplo con un dominio GST.

20 Se describen variantes de las proteínas marcadoras divulgadas que funcionan ya sea como agonistas (miméticos) o como antagonistas de las proteínas marcadoras. Se pueden generar variantes de las proteínas marcadoras mediante mutagénesis, por ejemplo, mutación de punto discreto o truncamiento de una proteína marcadora. Un agonista de las proteínas marcadoras puede retener esencialmente las mismas, o un subconjunto, de las actividades biológicas de una proteína marcadora en su forma de existencia natural. Un antagonista de una proteína marcadora puede inhibir una o
25 más de las actividades de la proteína marcadora en la forma en que existe naturalmente, por ejemplo, modular competitivamente una actividad de una proteína marcadora. Por lo tanto, se pueden provocar efectos biológicos específicos mediante el tratamiento con una variante de función limitada. En una realización, el tratamiento de un individuo con una variante que tiene un subconjunto de las actividades biológicas de la proteína en la forma en que existe naturalmente, tiene menos efectos secundarios en un individuo, comparado con un tratamiento con la proteína marcadora en
30 la forma en que existe naturalmente.

Las variantes de una proteína marcadora que funcionan ya sea como agonistas de una proteína marcadora (miméticos) o como antagonistas de la proteína marcadora se pueden identificar mediante el análisis de bibliotecas combinatorias de mutantes, por ejemplo, mutantes de truncamiento de una proteína marcadora para la actividad agonista o
35 antagonista de la proteína marcadora. En una realización, se genera una biblioteca variegada de variantes de proteínas marcadoras mediante mutagénesis combinatoria en el nivel de ácido nucleico y se codifica mediante una biblioteca de genes variegada. Se puede producir una biblioteca variegada de variantes de proteínas marcadoras, por ejemplo, ligando enzimáticamente una mezcla de oligonucleótidos sintéticos con secuencias de genes de tal manera que un conjunto degenerado de posibles secuencias de proteínas marcadoras es expresable como polipéptidos individuales, o de manera alternativa, como un conjunto de proteínas de fusión más grandes (por ejemplo, para expresión en fago) que
40 contiene el conjunto de secuencias de proteínas marcadoras en el mismo. Existen varios procedimientos que se pueden utilizar para producir bibliotecas de variantes de posibles proteínas marcadoras de una secuencia de oligonucleótidos degenerada. La síntesis química de una secuencia de genes degenerada se puede realizar en un sintetizador de ADN automático y después ligar el gen sintético en un vector de expresión apropiado. El uso de un conjunto de genes degenerados permite proporcionar, en una mezcla, todas las secuencias que codifican el conjunto de secuencias de proteínas
45 marcadoras posibles deseado. Los procedimientos para sintetizar oligonucleótidos degenerados son conocidos en la técnica (ver, por ejemplo., Narang, S. A. (1983) *Tetrahedron* 39:3; Itakura *et al.* (1984) *Annu. Rev. Biochem.* 53:323; Itakura *et al.* (1984) *Science* 198:1056; Ike *et al.* (1983) *Nucleic Acids Res.* 11:477).

50 Además, se pueden utilizar las bibliotecas de fragmentos de una secuencia de codificación de una proteína correspondientes a una proteína marcadora divulgada para generar una población variegada de fragmentos de proteína marcadora a efectos de buscar y posteriormente seleccionar variantes de una proteína marcadora. En una realización, se puede generar una biblioteca de fragmentos de secuencia de codificación tratando un fragmento PCR de doble hebra de una secuencia de codificación de una proteína marcadora con una enzima nucleasa en condiciones en las que se
55 producen muescas solamente una vez por molécula aproximadamente, desnaturalizando el ADN de doble hebra, renaturalizando el ADN para formar ADN de doble hebra que puede incluir pares sentido/antisentido a partir de diferentes productos con muescas, retirando las porciones de hebra única de los dúplex reformados mediante el tratamiento con nucleasa SI y ligando la biblioteca de fragmentos resultante en un vector de expresión. Mediante este método, se puede derivar una biblioteca de expresión que codifica fragmentos de terminal N, de terminal C y fragmentos internos de
60 varios tamaños de la proteína marcadora.

En la técnica se conocen varias técnicas para analizar productos génicos en busca de bibliotecas combinatorias preparadas mediante mutaciones de punto o truncamiento, y para analizar bibliotecas de ADNc para productos génicos que tienen una propiedad seleccionada. Las técnicas de uso más extendido, que son susceptibles al análisis de alto
65 rendimiento, para buscar en grandes bibliotecas de genes, generalmente incluyen clonar la biblioteca de genes en vectores de expresión replicables, transformar células apropiadas con la biblioteca de vectores resultante y expresar los genes combinatorios en condiciones en las que la detección de una actividad deseada facilita el aislamiento del vector que codifica al gen cuyo producto se detectó. Se puede utilizar la mutagénesis por ensamblaje recursiva (REM,

por sus siglas en inglés), una técnica nueva que aumenta la frecuencia de mutantes funcionales en las bibliotecas, en combinación con los ensayos de análisis para identificar variantes marcadoras (Arkin y Yourvan (1992) *Proc. Natl. Acad. Sci. USA* 89:7811-7815; Delgrave *et al.* (1993) *Protein Engineering* 6(3):327-331).

Una proteína marcadora aislada o una parte o fragmento de esta se puede utilizar como un inmunogen para generar anticuerpos que unen proteínas marcadoras utilizando técnicas estándar para la preparación de anticuerpos policlonales u monoclonales. Se puede utilizar una proteína marcadora en toda su longitud o, de manera alternativa, se proporcionan fragmentos de péptidos antigénicos de estas proteínas para utilizarlos como inmunogenes. El péptido antigénico de una proteína marcadora comprende al menos 8 residuos de aminoácido de una secuencia de aminoácidos codificada por un gen establecido en las Tablas 1-13 y abarca un epítipo de una proteína marcadora tal, que un anticuerpo dirigido contra el péptido forma un complejo inmune específico con la proteína marcadora. Preferentemente, el péptido antigénico comprende al menos 10 residuos de aminoácido, incluso más preferentemente al menos 15 residuos de aminoácido, más preferentemente aún al menos 20 residuos de aminoácido y más preferentemente todavía al menos 30 residuos de aminoácido.

Los epítopos preferidos abarcados por el péptido antigénico son regiones de la proteína marcadora que se ubican en la superficie de la proteína, por ejemplo, regiones hidrófilas, así como regiones con antigenicidad alta.

Un inmunogen de proteína marcadora generalmente se utiliza para preparar anticuerpos inmunizando un individuo apropiado (por ejemplo, conejo, cabra, ratón u otro mamífero) con el inmunogen. Una preparación inmunogénica apropiada puede contener, por ejemplo, una proteína marcadora expresada en forma recombinante o un polipéptido marcador sintetizado químicamente. La preparación puede incluir además un adyuvante, por ejemplo un adyuvante completo o incompleto de Freund o un agente inmunoestimulante similar. La inmunización de un individuo apropiado con una preparación de proteínas marcadoras inmunogénicas induce una respuesta de anticuerpo de una proteína antimarcadora policlonal.

Por consiguiente, otro aspecto se refiere a anticuerpos de proteínas antimarcadoras. El término “anticuerpo” tal como se utiliza en la presente, incluye moléculas de inmunoglobulina y partes inmunológicamente activas de moléculas de inmunoglobulina, es decir, moléculas que contienen un sitio de unión al antígeno que se une específicamente (inmunorreacciona con) un antígeno como, por ejemplo una proteína marcadora. Los ejemplos de partes inmunológicamente activas de moléculas de inmunoglobulina incluyen fragmentos F(ab) y F(ab')₂ que se pueden generar tratando el anticuerpo con una enzima como pepsina. Se proporcionan anticuerpos policlonales y monoclonales que se unen a proteínas marcadoras. En la presente el término “anticuerpo monoclonal” o “composición de anticuerpo monoclonal”, incluye una población de moléculas de anticuerpos que contienen una sola especie de un sitio de unión al antígeno capaz de inmunorreaccionar con un epítipo particular. Por lo tanto una composición de anticuerpos monoclonal, exhibe generalmente una única afinidad de unión para una proteína marcadora particular con la que inmunorreacciona.

Los anticuerpos de proteínas antimarcadoras policlonales se pueden preparar como se describió anteriormente, inmunizando un individuo apropiado con la proteína marcadora divulgada. La titulación de anticuerpos de proteínas antimarcadoras en el individuo inmunizado se puede controlar a través del tiempo mediante técnicas estándar, por ejemplo con un ensayo por inmunoabsorción ligado a enzimas (ELISA, por sus siglas en inglés) utilizando una proteína marcadora inmovilizada. Si se desea, las moléculas de anticuerpo dirigidas a las proteínas marcadoras se pueden aislar del mamífero (por ejemplo, de la sangre) y después purificarse mediante técnicas conocidas, como la cromatografía de afinidad por proteína A, para obtener la fracción IgG. En un momento apropiado después de la inmunización, por ejemplo, cuando las titulaciones de anticuerpos de proteína antimarcadora son más altas, se pueden obtener células productoras de anticuerpos del individuo y utilizarlas para preparar anticuerpos monoclonales mediante técnicas estándar, por ejemplo la técnica del hibridoma originalmente descrita por Kohler y Milstein (1975) *Nature* 256:495-497) (ver también, Brown *et al.* (1981) *J. Immunol* 127:539-46; Brown *et al.* (1980) *JBiol. Chem.* 255:4980-83; Yeh *et al.* (1976) *Proc. Natl. Acad. Sci. USA* 76:2927-31; y Yeh *et al.* (1982) *Int. J Cancer* 29:269-75), la técnica más reciente del hibridoma de la célula B humana (Kozbor *et al.* (1983) *Immunol Today* 4:72), la técnica EBV-hibridoma (Cole *et al.* (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pág. 77-96) o técnicas trioma. La tecnología para producir hibridomas de anticuerpos monoclonales es muy conocida (ver generalmente R. H. Kennel, en *Monoclonal Antibodies: A New Dimension In Biological Analyses*, Plenum Publishing Corp., New York, N.Y. (1980); E. A. Lerner (1981) *Yale J Biol. Med.*, 54:387-402; M. L. Gefler *et al.* (1977) *Somatic Cell Genet.* 3:231-36). Brevemente, una línea celular inmortal (generalmente un mieloma) se fusiona a linfocitos (generalmente esplenocitos) de un mamífero inmunizado con un inmunogen de proteína marcadora como se describió anteriormente, y se analizan los sobrenadantes de cultivo de las células hibridoma resultantes para identificar un hibridoma que produce un anticuerpo monoclonal que se une a una proteína marcadora de la invención.

Cualquiera de los protocolos conocidos utilizados para fusionar linfocitos y líneas celulares immortalizadas se puede aplicar para generar un anticuerpo monoclonal de una proteína antimarcadora (ver, por ejemplo, G. Galfre *et al.* (1977) *Nature* 266:55052; Gefler *et al Somatic Cell Genet.*, citado antes; Lerner, *Yale J Biol. Med.*, citado antes; Kennel, *Monoclonal Antibodies*, citado antes). Además, el entendido en la técnica apreciará que hay muchas variaciones de dichos procedimientos que también serían útiles. Generalmente, la línea celular inmortal (por ejemplo una línea celular de mieloma) se deriva de esta especie de mamífero que los linfocitos. Por ejemplo, se pueden preparar hibridomas murinos fusionando linfocitos de un ratón inmunizado con una preparación inmunogénica de la presente invención con una línea celular de ratón immortalizada. Las líneas celulares inmortales preferidas son líneas celulares de mieloma de ratón que son sensibles al medio de cultivo que contiene hipoxantina, aminopterina y timidina (“medio

HAT", por sus siglas en inglés). Se puede utilizar cualquiera de un conjunto de líneas celulares de mieloma como compañero de fusión de conformidad con técnicas estándar, por ejemplo, las líneas de mieloma P3-NSI/I-Ag4-1, P3-x63-Ag8.653 o Sp2/O-Ag14. Estas líneas de mieloma están disponibles de ATCC. Generalmente, las células de mieloma de ratón sensibles al HAT se fusionan con esplenocitos de ratón utilizando polietilenglicol ("PEG"). Las células hibridoma resultantes de la fusión se seleccionan entonces utilizando medio HAT, que mata las células de mieloma no fusionadas y fusionadas improductivas (los esplenocitos no fusionados mueren después de varios días porque no se transforman). Las células de hibridoma que producen un anticuerpo monoclonal de la invención se detectan analizando los sobrenadantes del cultivo de hibridoma en busca de anticuerpos que se unen a una proteína marcadora, por ejemplo, utilizando un ensayo ELISA estándar.

De manera alternativa a la preparación de hibridomas que secretan anticuerpos monoclonales, se puede identificar un anticuerpo de proteína antimarcadora monoclonal y se lo puede aislar analizando una biblioteca de inmunoglobulina combinatoria recombinante (por ejemplo, una biblioteca de expresión en fago de anticuerpos) con una proteína marcadora para aislar así miembros de una biblioteca de inmunoglobulina que se unen a una proteína marcadora. Están comercialmente disponibles, paquetes para generar y analizar bibliotecas de expresión de fago (por ejemplo., The Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; y The Stratagene *SurfZAP™ Phage Display Kit*, Catalog No. 240612). Además, se pueden encontrar ejemplos de procedimientos y reactivos particularmente susceptibles para utilizar en la generación y el análisis de bibliotecas de expresión de anticuerpos, por ejemplo, en Ladner *et al.* Pat. EE.UU No. 5,223,409; Kang *et al.* publicación internacional PCT No. WO 92/18619; Dower *et al.* publicación internacional PCT No. WO 91/17271; Winter *et al.* publicación Internacional PCT WO 92/20791; Markland *et al.* publicación internacional PCT No. WO 92/15679; Breilling *et al.* publicación Internacional PCT WO 93/01288; McCafferty *et al.* publicación internacional PCT No. WO 92/01047; Garrard *et al.* publicación internacional PCT No. WO 92/09690; Ladner *et al.* publicación internacional PCT No. WO 90/02809; Fuchs *et al.* (1991) *Bio/Technology* 9:1370-1372; Hay *et al.* (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse *et al.* (1989) *Science* 246:1275-1281; Griffiths *et al.* (1993) *EMBO J* 12:725-734; Hawkins *et al.* (1992) *J. Mol Biol.* 226:889-896; Clarkson *et al.* (1991) *Nature* 352:624-628; Gram *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:3576-3580; Garrard *et al.* (1991) *Bio/Technology* 9:1373-1377; Hoogenboom *et al.* (1991) *Nuc. Acids Res.* 19:4133-4137; Barbas *et al.* (1991) *Proc. Natl. Acad. Sci. USA* 88:7978-7982; y McCafferty *et al.* *Nature* (1990) 348:552-554.

Además, los anticuerpos de proteínas antimarcadoras recombinantes, como los anticuerpos monoclonales quiméricos y humanizados, que comprenden tanto partes humanas como no humanas, se pueden preparar utilizando técnicas de ADN recombinante estándar. Dichos anticuerpos monoclonales quiméricos y humanizados se pueden producir mediante técnicas de ADN recombinante conocidas en la técnica, por ejemplo, utilizando los procedimientos descritos en Robinson *et al.* solicitud internacional No. PCT/US86/02269; Akira, *et al.* solicitud de patente europea 184,187; Taniguchi, M., solicitud de patente europea 171,496; Morrison *et al.* solicitud de patente europea 173,494; Neuberger *et al.* publicación internacional PCT No. WO 86/01533; Cabilly *et al.* Pat. EEUU No. 4,816,567; Cabilly *et al.* solicitud de patente europea 125,023; Better *et al.* (1988) *Science* 240:1041-1043; Liu *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu *et al.* (1987) *J. Immunol.* 139:3521-3526; Sun *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura *et al.* (1987) *Cane. Res.* 47:999-1005; Wood *et al.* (1985) *Nature* 314:446-449; y Shaw *et al.* (1988) *J. Natl. Cancer. Inst.* 80:1553-1559; Morrison, S. L. (1985) *Science* 229:1202-1207; Oi *et al.* (1986) *Bio Techniques* 4:214; Winter U.S. Pat. No. 5,225,539; Jones *et al.* (1986) *Nature* 321:552-525; Verhoeven *et al.* (1988) *Science* 239:1534; y Beidler *et al.* (1988) *J. Immunol.* 141:4053-4060.

Los anticuerpos completamente humanos son particularmente deseables para el tratamiento terapéutico de seres humanos. Estos anticuerpos se pueden producir utilizando ratones transgénicos que son incapaces de expresar genes de cadenas livianas y pesadas de inmunoglobulina endógena, pero que pueden expresar genes de cadenas livianas y pesadas humanas. Los ratones transgénicos se inmunizan de manera normal con un antígeno seleccionado, por ejemplo, una parte o la totalidad de un polipéptido correspondiente a un marcador de la invención. Los anticuerpos monoclonales dirigidos contra el antígeno se pueden obtener utilizando tecnología de hibridoma convencional. Los transgenes de inmunoglobulina humana albergados por los ratones transgénicos se acomodan durante la diferenciación celular B y posteriormente atraviesan la mutación somática y de cambio de clase. Así, al utilizar esta técnica, es posible producir anticuerpos IgG, IgA e IgE terapéuticamente útiles. Para una sinopsis de esta tecnología para producir anticuerpos humanos, ver Lonberg y Huszar (1995) *Int. Rev. Immunol.* 13:65-93. Para un análisis detallado de esta tecnología para producir anticuerpos humanos y anticuerpos monoclonales humanos y los protocolos para producir dichos anticuerpos, ver, por ejemplo., pat. EEUU Nos. 5,625,126; 5,633,425; 5,569,825; 5,661, 016; y 5,545,806. Además, empresas como Abgenix, Inc. (Freemont, Calif.), pueden proporcionar anticuerpos humanos dirigidos contra un antígeno seleccionado, utilizando una tecnología similar a la aquí descrita.

Los anticuerpos completamente humanos que reconocen un epítipo seleccionado se pueden generar utilizando una técnica llamada "selección guiada". En este abordaje, se utiliza un anticuerpo monoclonal no humano seleccionado, por ejemplo, un anticuerpo murino, para guiar la selección de un anticuerpo completamente humano que reconoce el mismo epítipo (Jespersen *et al.*, 1994, *Bio/technology* 12:899-903).

Un anticuerpo de proteína antimarcadora (por ejemplo, un anticuerpo monoclonal) se puede utilizar para aislar una proteína marcadora divulgada, mediante técnicas estándar, como cromatografía de afinidad o inmunoprecipitación. Un anticuerpo de proteína antimarcadora puede facilitar la purificación de proteínas marcadoras naturales de células y de proteínas marcadoras producidas recombinantemente, expresadas en células receptoras. Además, se puede utilizar un anticuerpo de proteína antimarcadora para detectar la proteína marcadora (por ejemplo, en un lisado

celular o sobrenadante celular) a efectos de evaluar la abundancia y el patrón de expresión de la proteína marcadora. Los anticuerpos de proteínas antimarcadoras se pueden utilizar en forma de diagnóstico para controlar los niveles de proteína en tejidos, como parte de un procedimiento de prueba clínico, por ejemplo, para determinar la eficacia de un régimen de tratamiento determinado. La detección se puede facilitar mediante el acoplamiento (es decir, unir físicamente) del anticuerpo con una sustancia detectable. Los ejemplos de sustancias detectables incluyen varias enzimas, grupos prostéticos, materiales fluorescentes, materiales luminiscentes, materiales bioluminiscentes y materiales radioactivos. Los ejemplos de enzimas apropiadas incluyen peroxidasa de rábano picante, fosfatasa alcalina, galactosidasa o acetilcolinesterasa; los ejemplos de complejos de grupos prostéticos adecuados incluyen estreptavidina/biotina y avidina/biotina; los ejemplos de materiales fluorescentes apropiados incluyen umbeliferona, fluoresceína, isotiocianato de fluoresceína, rodamina, diclorotriazinilamina fluoresceína, cloruro de dansilo o ficoeritrina; un ejemplo de material luminiscente incluye luminol; los ejemplos de materiales bioluminiscentes incluyen luciferasa, luciferina y aecuorina, y los ejemplos de material radioactivo incluyen ^{125}I , ^{131}I , ^{35}S o ^3H .

15 III. Células receptoras y vectores de expresión recombinantes

Otro aspecto se refiere a vectores, preferentemente vectores de expresión, que contienen un ácido nucleico que codifica una proteína marcadora de la invención (o una parte de esta). Como se usa en la presente, el término “vector” incluye una molécula de ácido nucleico capaz de transportar otro ácido nucleico al que se lo ha unido. Un tipo de vector es un “plásmido”, que incluye un bucle de ADN de doble hebra circular al que se pueden ligar segmentos adicionales de ADN. Otro tipo de vector es un vector viral, al que se pueden ligar segmentos adicionales de ADN en el genoma viral. Ciertos vectores son capaces de replicación autónoma en una célula receptora dentro de la que se introducen (por ejemplo, vectores bacteriales que tienen un origen bacterial de replicación y vectores episomales de mamífero). Otros vectores (por ejemplo, vectores de mamífero no episomal) se integran al genoma de una célula receptora al introducirlos en esta y de ese modo se replican junto con el genoma del receptor. Además, ciertos vectores son capaces de dirigir la expresión de los genes a los cuales están unidos operativamente. En la presente dichos vectores se denominan “vectores de expresión”. En general, los vectores de expresión de utilidad en las técnicas de ADN recombinante están a menudo en forma de plásmidos. En la presente memoria descriptiva, “plásmido” y “vector” se pueden utilizar de forma intercambiable puesto que el plásmido es la forma de vector usada más comúnmente. Sin embargo, se describen también otras formas de vectores de expresión, como vectores virales (por ejemplo, retrovirus defectuosos en replicación, adenovirus y virus adenoasociados), que realizan funciones equivalentes.

Los vectores de expresión recombinantes descritos comprenden un ácido nucleico divulgado en una forma apropiada para la expresión del ácido nucleico en una célula receptora, lo que significa que los vectores de expresión recombinantes incluyen una o más secuencias reguladoras, seleccionadas sobre la base de las células receptoras que se van a utilizar para la expresión, las que están unidas operativamente a la secuencia de ácido nucleico que se ha de expresar. En un vector de expresión recombinante, “unido operativamente” significa que la secuencia de nucleótidos de interés está unida a las secuencias reguladoras en una forma que permite la expresión de la secuencia de nucleótidos (por ejemplo en un sistema de transcripción/traducción *in vitro* o en una célula receptora cuando se introduce el vector en la célula receptora). El término “secuencia reguladora” incluye promotores, aumentadores y otros elementos de control de expresión (por ejemplo, señales de poliadenilación). Dichas secuencias reguladoras se describen, por ejemplo, in Goeddel; *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990). Las secuencias reguladoras incluyen aquellas que dirigen la expresión constitutiva de una secuencia de nucleótidos en muchos tipos de células receptoras y a aquellas que dirigen la expresión de la secuencia de nucleótidos solamente en ciertas células receptoras (por ejemplo, secuencias reguladoras específicas de tejido). Los entendidos en la técnica apreciarán que el diseño del vector de expresión puede depender de factores tales como la elección de la célula receptora a transformar, el nivel de expresión de proteína deseada, y similar. Los vectores de expresión divulgados se pueden introducir en células receptoras para producir así proteínas o péptidos, incluidas proteínas de fusión o péptidos, codificados por ácidos nucleicos como se describe en la presente (por ejemplo, proteínas marcadoras, formas mutantes de proteínas marcadoras, proteínas de fusión y similares).

Los vectores de expresión recombinantes divulgados se pueden diseñar para que expresen las proteínas marcadoras en células procariotas o eucariotas. Por ejemplo, las proteínas marcadoras se pueden expresar en células bacterianas como *E. coli*, células de insectos (utilizando vectores de expresión de baculovirus), células de levadura o células de mamífero. En Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990) se analizan las células receptoras de manera más detallada. En forma alternativa, el vector de expresión recombinante se puede transcribir y traducir *in vitro*, por ejemplo utilizando secuencias reguladoras promotoras de T7 y polimerasa T7.

La expresión de proteínas en procariotas se realiza más a menudo en *E. coli* con vectores que contienen promotores constitutivos o inducibles que dirigen la expresión de las proteínas de fusión o no fusión. Los vectores de fusión agregan varios aminoácidos a una proteína allí codificada, generalmente al terminal amino de la proteína recombinante. Dichos vectores de fusión sirven generalmente para tres propósitos: 1) aumentar la expresión de la proteína recombinante; 2) aumentar la solubilidad de la proteína recombinante; y 3) ayudar en la purificación de la proteína recombinante actuando como un ligando en la purificación por afinidad. A menudo, en los vectores de expresión de fusión se introduce un sitio de escisión proteolítica en la unión de la porción de fusión y la proteína recombinante para posibilitar la separación de la proteína recombinante de la porción de fusión, después de la purificación de la proteína de fusión. Estas enzimas y sus secuencias de reconocimiento cognato, incluyen al Factor Xa, trombina y enteroqui-

nasa. Los vectores de expresión de fusión típicos incluyen pGEX (Pharmacia Biotech Inc; Smith, D. B. y Johnson, K. S. (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, Mass.) y pRITS (Pharmacia, Piscataway, N.J.) que fusionan glutatión S-transferasa (GST), la proteína de unión de maltosa E o la proteína A, respectivamente, con la proteína recombinante objetivo.

Las proteínas de fusión purificadas se pueden utilizar en ensayos de actividad marcadora, (por ejemplo, ensayos directos o ensayos competitivos que se describen en detalle posteriormente) o para generar anticuerpos específicos para las proteínas marcadoras, por ejemplo.

Los ejemplos de vectores de expresión de no fusión inducibles en *E. coli*, apropiados incluyen pTrc (Amann *et al.*, (1988) *Gene* 69:301-315) y pET lid (Studier *et al.*, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990) 60-89). La expresión de un gen objetivo del vector pTrc se basa en la transcripción de ARN polimerasa huésped de un promotor de fusión trp-lac híbrido. La expresión de un gen objetivo del vector pET lid se basa en la transcripción de un promotor de fusión T7 gn 10-lac mediado por una ARN polimerasa viral coexpresada (T7 gnl). Esta polimerasa viral es suministrada por cepas receptoras BL21(DE3) o HMS174(DE3) de un profago residente que alberga a un gen T7gnl bajo el control transcripcional del promotor lac UV5.

Una estrategia para maximizar la expresión de proteína recombinante en *E. coli* es expresar la proteína en una bacteria huésped con una capacidad alterada para escindir proteolíticamente la proteína recombinante (Gottesman, S., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, Calif. (1990) 119-128). Otra estrategia es alterar la secuencia de ácido nucleico del ácido nucleico que se va a insertar en el vector de expresión de modo que los codones individuales para cada aminoácido sean aquellos utilizados preferentemente en *E. coli* (Wada *et al.*, (1992) *Nucleic Acids Res.* 20:2111-2118). Tal alteración de las secuencias de ácido nucleico divulgadas se puede realizar mediante técnicas de síntesis de ADN estándar.

En otra realización, el vector de expresión de la proteína marcadora es un vector de expresión de levadura. Los ejemplos de vectores de expresión en la levadura *S. cerevisiae* incluyen pYepSecl (Baldari, *et al.*, (1987) *Embo J.* 6:229-234), pMFa (Kurjan y Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz *et al.*, (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, Calif.) y picZ (Invitrogen Corp, San Diego, CA).

De manera alternativa, las proteínas marcadoras divulgadas se pueden expresar en células de insecto utilizando vectores de expresión de baculovirus. Los vectores de baculovirus disponibles para la expresión de proteínas en células de insecto cultivadas (por ejemplo, células Sf9) incluyen las series pAC (Smith *et al.* (1983) *Mol. Cell Biol.* 3:2156-2165) y la serie pVL (Lucklow y Summers (1989) *Virology* 170:31-39).

En otra realización adicional, se expresa un ácido nucleico divulgado en células de mamífero utilizando un vector de expresión de mamífero. Los ejemplos de vectores de expresión de mamífero incluyen pCDMS (Seed, B. (1987) *Nature* 329:840) y pMT2PC (Kaufman *et al.* (1987) *EMBO J.* 6:187-195). Cuando se los utiliza en células de mamífero, las funciones de control de los vectores de expresión son suministradas a menudo por elementos reguladores virales. Por ejemplo, los promotores utilizados comúnmente derivan de poliovirus, Adenovirus 2, citomegalovirus y virus de simio 40. Para otros sistemas de expresión apropiados tanto para células procariotas como eucariotas ver capítulos 16 y 17 de Sambrook, J., Fritsh, E. E. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*. 2nd, ed, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989.

En otra realización, el vector de expresión mamífero recombinante es capaz de dirigir la expresión del ácido nucleico, preferentemente en un tipo particular de células (por ejemplo, se utilizan elementos reguladores específicos de tejido para expresar el ácido nucleico). Los elementos reguladores específicos de tejido son conocidos en la técnica. Los ejemplos no taxativos de promotores específicos de tejido apropiados incluyen al promotor de albúmina (específico para el hígado; Pinkert *et al.* (1987) *Genes Dev.* 1:268-277), promotores linfoides específicos (Calame y Eaton (1988) *Adv. Immunol.* 43:235-275), en particular promotores de receptores de células T (Winoto y Baltimore (1989) *EMBO J.* 8:729-733) e inmunoglobulinas (Banerji *et al.* (1983) *Cell* 33:729-740; Queen y Baltimore (1983) *Cell* 33:741-748), promotores específicos de neuronas (por ejemplo, el promotor de neurofilamento; Byrne y Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), promotores específicos de páncreas (Edlund *et al.*, (1985) *Science* 230:912-916), y promotores específicos de la glándula mamaria (por ejemplo., promotor de suero lácteo; Pat. EE.UU No. 4,873,316 y solicitud de patente europea No. 264,166). También abarca los promotores regulados por el desarrollo, por ejemplo los promotores de caja murina Kessel y Gruss (1990) *Science* 249:374-379) y el promotor de alfa fetoproteína (Campes y Tilghman (1989) *Genes Dev.* 3:537-546).

Se proporciona además un vector de expresión recombinante que comprende una molécula de ADN divulgada clonada en el vector de expresión en una orientación antisentido. Es decir, la molécula de ADN está unida operativamente a una secuencia reguladora de forma tal que permite la expresión (mediante transcripción de la molécula de ADN) de una molécula de ARN que es el antisentido de un ARNm correspondiente a un gen descrito (por ejemplo, un gen establecido en las Tablas 1-13). Se puede elegir las secuencias reguladoras unidas operativamente a un ácido nucleico clonado en una orientación antisentido, que dirigen la expresión continua de la molécula de ARN antisentido en varios tipos de células, por ejemplo promotores virales y/o aumentadores, o se puede elegir aquellas secuencias reguladoras que dirigen la expresión de ARN antisentido constitutiva, específica de tejido o de una célula específica. El vector de expresión antisentido puede estar en la forma de un plásmido recombinante, un virus fagémido o atenuado en el que los ácidos nucleicos antisentido se producen bajo el control de una región reguladora altamente eficiente, cuya activi-

dad se puede determinar mediante el tipo de célula en la que se introduce el vector. Para un estudio de la regulación de la expresión de genes utilizando genes anti sentido, ver Weintraub, H. *et al.*, Antisense RNA as a molecular tool for genetic analysis, *Reviews-Trends in Genetics*, Vol. 1(1) 1986.

5 Otro aspecto se refiere a células receptoras dentro de las que se introduce una molécula de ácido nucleico divulgado, por ejemplo, un gen establecido en las Tablas 1-13 dentro de un vector de expresión recombinante o una molécula de ácido nucleico divulgada que contiene secuencias que le permiten recombinarse homológamente en un sitio específico del genoma de la célula receptora. Los términos “célula receptora”, “célula receptora recombinante” se utilizan de manera intercambiable en la presente. Se entiende que dichos términos se refieren no solo a la célula objeto particular
10 sino también a la progenie o progenie posible de dicha célula. Puesto que en generaciones futuras pueden ocurrir ciertas modificaciones debido ya sea a la mutación o a las influencias ambientales, dicha progenie no puede ser, de hecho, idéntica a la célula madre, pero aún así, tal como se utiliza en la presente, está incluida dentro del alcance del término.

15 Una célula receptora puede ser cualquier célula procariota o eucariota. Por ejemplo, una proteína marcadora divulgada se puede expresar en células bacterianas como *E. coli*, células de insecto, levadura o células de mamífero (como en células de ovario de hámster chino (CHO) o células COS). Los entendidos en la técnica conocen otras células receptoras apropiadas.

20 Un vector de ADN se puede introducir en células procariotas o eucariotas mediante técnicas de transformación o transfección convencionales. Como se utiliza en la presente, los términos “transformación” y “transfección” se refieren a una variedad de técnicas reconocidas en la técnica para introducir ácido nucleico extraño (por ejemplo ADN) en una célula receptora, incluida la co-precipitación de fosfato de calcio o de cloruro de calcio, la transfección mediada por DEAE-dextrano, la lipofección o la electroporación. En Sambrook, *et al.* (*Molecular Cloning: A Laboratory Manual*,
25 *2nd, ed*, Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989) y otros manuales de laboratorio se pueden encontrar procedimientos apropiados para transformar o transfectar células receptoras.

Para la transfección estable de células de mamífero, se sabe que en función del vector de expresión y de la técnica
30 de transfección utilizada, solo una pequeña fracción de células puede integrar el ADN extraño a su genoma. A efectos de identificar y seleccionar estos integrantes, generalmente se introduce un gen que codifica marcadores seleccionables (por ejemplo, resistencia a antibióticos) en las células receptoras, junto con el gen de interés. Los marcadores seleccionables preferidos incluyen aquellos que le confieren resistencia a los fármacos, como G418, higromicina y metotrexato. El ácido nucleico que codifica un marcador seleccionable se puede introducir en una célula receptora
35 en el mismo vector que codifica una proteína marcadora o se lo puede introducir en un vector separado. Las células transfectadas establemente con el ácido nucleico introducido se pueden identificar mediante selección de fármaco (por ejemplo, las células que han incorporado el gen marcador seleccionable sobrevivirán, mientras que las otras células morirán).

40 Una célula receptora divulgada, como una célula receptora procariota o eucariota en cultivo, se puede utilizar para producir (es decir, expresar) una proteína marcadora. Por consiguiente, se proporcionan los procedimientos para producir una proteína marcadora utilizando las células receptoras divulgadas. En una realización, el procedimiento comprende cultivar la célula receptora (dentro de la que se ha introducido un vector de expresión recombinante que codifica una proteína marcadora) en un medio apropiado en el que se produzca una proteína marcadora. En otra
45 realización, el procedimiento comprende además aislar una proteína marcadora del medio o de la célula receptora.

Las células receptoras también se pueden utilizar para producir animales transgénicos no humanos. Por ejemplo, en una realización, una célula receptora es un ovocito fertilizado o una célula madre embrionaria dentro de la que se han introducido secuencias codificadoras de una proteína marcadora. Dichas células receptoras se pueden utilizar entonces
50 para crear animales transgénicos no humanos en los que se han introducido en sus genomas secuencias exógenas que codifican una proteína marcadora divulgada, o animales recombinantes homólogos en los que las secuencias endógenas que codifican las proteínas marcadoras divulgadas han sido alteradas. Dichos animales son útiles para estudiar la función y/o actividad de una proteína marcadora y para identificar y/o evaluar moduladores de la actividad de la proteína marcadora. Como se utiliza en la presente, un “animal transgénico” es un animal no humano, preferentemente un mamífero, más preferentemente un roedor como una rata o ratón, en el que una o más células del animal incluyen un transgén. Otros ejemplos de animales transgénicos incluyen primates no humanos, ovejas, perros, vacas, cabras, pollos, anfibios y similares. Un transgén es ADN exógeno que está integrado dentro del genoma de una célula a partir de la que se desarrolla un animal transgénico y que permanece en el genoma del animal maduro, dirigiendo de ese modo la expresión de un producto génico codificado en uno o más tipos de células o tejidos del animal transgénico. Como se
60 utiliza en la presente, un “animal recombinante homólogo” es un animal no humano, preferentemente un mamífero, más preferentemente un ratón, en el que se ha alterado un gen endógeno (por ejemplo, un gen establecido en las Tablas 1-13) mediante recombinación homóloga entre el gen endógeno y una molécula de ADN exógena introducida en una célula del animal, por ejemplo, una célula embrionaria del animal, antes de que el animal se desarrolle.

65 Un animal transgénico se puede crear introduciendo un ácido nucleico que codifica un marcador en el pronúcleo macho de un ovocito fertilizado, por ejemplo, mediante microinyección, infección retroviral, y permitiendo que el ovocito se desarrolle en un animal adoptivo hembra pseudopreñado. También se pueden incluir secuencias intrónicas y señales de poliadenilación en el transgén para aumentar la eficiencia de expresión del transgén. Unas o más secuencias

reguladoras específicas de tejido se pueden unir operativamente a un transgén para dirigir la expresión de una proteína marcadora a células particulares. Los procedimientos para generar animales transgénicos mediante la manipulación de embriones y microinyección, particularmente animales como ratones, se han vuelto convencionales en la técnica y se describen por ejemplo, en las Pats. EEUU Nos. 4,736,866 y 4,870,009, ambas de Leder *et al.*, la Pat. EEUU No. 4,873,191 de Wagner *et al.* y en Hogan, B., *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Se utilizan procedimientos similares para producir otros animales transgénicos. Un animal fundador transgénico se puede identificar basándose en la presencia de un transgén divulgado en su genoma y/o la expresión de ARNm correspondiente a un gen de la invención en tejidos o células de los animales. Luego, un animal fundador transgénico se puede utilizar para criar animales adicionales que portan el transgén. Además, los animales transgénicos que portan un transgén que codifica una proteína marcadora se pueden cruzar además con otros animales transgénicos portadores de otros transgenes.

Para crear un animal recombinante homólogo, se prepara un vector que contiene al menos una parte de un gen divulgado dentro del cual se ha introducido una adición, supresión o sustitución para alterar así al gen, por ejemplo, trastornarlo funcionalmente. El gen puede ser un gen humano, pero más preferentemente, es un homólogo no humano de un gen humano divulgado (por ejemplo, un gen establecido en las Tablas 1-13). Por ejemplo, un gen de ratón se puede utilizar para construir una molécula de ácido nucleico de recombinación homóloga, por ejemplo, un vector, apropiado para alterar un gen endógeno en el genoma de ratón. En una realización preferida, la molécula de ácido nucleico de recombinación homóloga está diseñada de manera tal que, con la recombinación homóloga, el gen endógeno se trastorna funcionalmente (es decir, ya no codifica a una proteína funcional, también llamado vector “knock out”). De manera alternativa, la molécula de ácido nucleico de recombinación homóloga se puede diseñar de manera tal que, con la recombinación homóloga, el gen endógeno muta o se altera de algún otro modo pero todavía codifica la proteína funcional (por ejemplo, la región reguladora aguas arriba se puede alterar para alterar así la expresión de la proteína marcadora endógena). En la molécula de ácido nucleico de recombinación homóloga, la parte alterada del gen está flanqueada en sus extremos 5' y 3' por secuencias de ácido nucleico adicionales del gen para permitir que ocurra la recombinación homóloga entre el gen exógeno portado por la molécula de ácido nucleico de recombinación homóloga y un gen endógeno en una célula, por ejemplo, una célula madre embrionaria. La secuencia de ácido nucleico adicional que flanquea es lo suficientemente larga como para que se produzca una recombinación homóloga exitosa con el gen endógeno. Generalmente, se incluyen varias kilobases de ADN flanqueador (en ambos extremos 5' y 3') en la molécula de ácido nucleico de recombinación homóloga (ver, por ejemplo., Thomas, K. R. y Capecchi, M. R. (1987) *Cell* 51:503 para una descripción de vectores de recombinación homóloga). Se introduce la molécula de ácido nucleico de recombinación homóloga en una célula, por ejemplo, una línea celular madre embrionaria (por ejemplo, mediante electroporación) y células en las que el gen introducido se ha recombinado homológamente con el gen endógeno seleccionado (ver, por ejemplo, Li, E. *et al.* (1992) *Cell* 69:915). Entonces se puede inyectar las células seleccionadas en un blastocito de un animal (por ejemplo un ratón) para formar quimeras de acumulación (ver, por ejemplo, Bradley, A. in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987) pág. 113-152). Luego se puede implantar un embrión quimérico en un animal adoptivo hembra pseudopreñado y llevar a término el embrión. La progenie que alberga el ADN recombinado homológamente en sus células germinales se puede utilizar para criar animales en los que todas las células del animal contienen el ADN recombinado homológamente mediante la transmisión de la línea germinal del transgén. Los procedimientos para construir moléculas de ácido nucleico de recombinación homóloga, por ejemplo, vectores o animales recombinantes homólogos se describen más detalladamente en Bradley, A. (1991) *Current Opinion in Biotechnology* 2:823-829 y en la publicación internacional PCT Nos.: WO 90/11354 por Le Mouellec *et al.*; WO 91/01140 por Smithies *et al.*; WO 92/0968 por Zijlstra *et al.*; y WO 93/04169 por Berns *et al.*

En otra realización, se pueden producir animales transgénicos no humanos que contengan sistemas seleccionados que permitan la expresión regulada del transgén. Un ejemplo de este sistema es el sistema de recombinación *cre/loxP* del bacteriófago P1. Para una descripción del sistema de recombinación *cre/loxP*, ver, por ejemplo, Lakso *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Otro ejemplo de sistema de recombinación es el sistema de recombinación FLP de *Saccharomyces cerevisiae* (O'Gorman *et al.* (1991) *Science* 251:1351-1355. Si se utiliza un sistema de recombinación *cre/loxP* para regular la expresión del transgén, se necesitan animales que contengan transgenes que codifican tanto al Cre de recombinación como a una proteína seleccionada. Dichos animales se pueden proporcionar a través de la construcción de animales “doblemente” transgénicos, por ejemplo, cruzando dos animales transgénicos, uno que contiene un transgén que codifica una proteína seleccionada y el otro que contiene un transgén que codifica una recombinasa.

Los clones de los animales transgénicos no humanos aquí descritos también se pueden producir de conformidad con los procedimientos descritos en Wilmut, I. *et al.* (1997) *Nature* 385:810-813 y la publicación internacional PCT Nos. WO 97/07668 y WO 97/07669. En resumen, una célula, por ejemplo, una célula somática, del animal transgénico se puede aislar e inducir a que salga del ciclo de crecimiento y entre en la fase G₀. La célula quiescente puede entonces fusionarse, por ejemplo, a través del uso de impulsos eléctricos, con un ovocito enucleado de un animal de esta especie de la que se aisló la célula quiescente. Luego se cultiva el ovocito reconstruido de modo tal que desarrolla en una mórula o blastocito y después se transfiere a un animal adoptivo hembra pseudopreñado. Las crías nacidas de este animal adoptivo hembra serán un clon del animal del cual se aisló la célula, por ejemplo, la célula somática.

V. Soporte magnético y matrices

Se describen también soportes magnéticos que comprenden uno o más marcadores divulgados. Como se utiliza aquí, “soporte magnético” incluye un medio que se puede leer y al que se puede acceder directamente a través de un ordenador. Tal medio incluye, aunque no taxativamente: medios de almacenamiento magnético, como disquetes, medios de almacenamiento en discos duros y cintas magnéticas; medios de almacenamiento ópticos como CD-ROM; medios de almacenamiento eléctrico como RAM y ROM; e híbridos de estas categorías como medios de almacenamiento magnético/óptico. El entendido en la materia apreciará fácilmente cómo cualquiera de los soportes magnéticos conocidos actualmente se puede utilizar para crear un dispositivo que comprende soportes magnéticos en los que se graba un marcador divulgado.

Como se utiliza en la presente, “grabado” incluye un procedimiento para almacenar información en medios informáticos legibles. Los entendidos en la técnica podrán adoptar fácilmente cualquiera de los procedimientos conocidos actualmente para grabar información en soportes magnéticos para generar dispositivos que comprendan los marcadores divulgados.

Se puede utilizar una variedad de formatos y programas de procesamiento de datos para almacenar la información del marcador en soportes magnéticos. Por ejemplo, la secuencia de ácido nucleico correspondiente a los marcadores se puede representar en un archivo de texto de un procesador de texto, con el formato del software comercialmente disponible como Word Perfect y Microsoft Word, o se puede representar en forma de una archivo ASCII, almacenado en una aplicación de una base de datos, como DB2, Sybase, Oracle o similares. Cualquier número de formatos estructurantes de un procesador de datos (por ejemplo, archivo de texto o base de datos) se puede adaptar a efectos de obtener soportes magnéticos en los que se graban los marcadores divulgados.

Mediante los marcadores divulgados en medios magnéticos, se puede acceder rutinariamente a la información de la secuencia marcadora para varios propósitos. Por ejemplo, un entendido en la técnica puede utilizar las secuencias de nucleótidos o de aminoácidos divulgadas en forma computarizada legible para comparar una secuencia objetivo o un motivo estructural objetivo con la información de secuencia almacenada en los medios de almacenamiento de información. Se utilizan medios de búsqueda para identificar los fragmentos o regiones de las secuencias que se corresponden con una secuencia objetivo particular o un motivo objetivo.

También se proporciona una matriz que comprende uno o más marcadores divulgados. La matriz se puede utilizar para ensayar la expresión de uno o más genes en la matriz. En una realización, la matriz se puede utilizar para ensayar la expresión de un gen en un tejido para asegurar la especificidad de tejido de los genes en la matriz. de este modo, se puede ensayar la expresión de hasta 8600 genes simultáneamente. Esto permite desarrollar un perfil que muestra una batería de genes expresados específicamente en uno o más tejidos.

Además de dicha determinación cualitativa, la divulgación permite la cuantificación de la expresión de genes. Por lo tanto, no solo se puede verificar la especificidad de los tejidos, sino también el nivel de expresión de una batería de genes en el tejido. Por lo tanto, los genes se pueden agrupar basándose en su expresión en los tejidos *per se* y en el nivel de expresión en ese tejido. Esto es útil, por ejemplo, para verificar la relación de la expresión de genes entre tejidos. Por lo tanto, se puede perturbar un tejido y determinar el efecto en la expresión del gen en un segundo tejido. En este contexto, se puede determinar el efecto de un tipo de célula u otro tipo de célula en respuesta a un estímulo biológico. Tal determinación es útil, por ejemplo, para saber el efecto de la interacción célula-célula en el nivel de la expresión de genes. Si se administra terapéuticamente un agente para tratar un tipo de célula pero tiene un efecto no deseado en otro tipo de célula, se proporciona un ensayo para determinar la base molecular del efecto no deseado y de ese modo proporcionar la oportunidad de coadministrar un agente que contrarreste o trate de algún otro modo el efecto no deseado. de forma similar, aun dentro de un único tipo de célula, se pueden determinar efectos biológicos no deseados en el nivel molecular. Así, se puede verificar y contrarrestar los efectos de un agente sobre la expresión de otro gen que no sea el gen objetivo.

En otra realización, la matriz se puede utilizar para controlar el transcurso de tiempo de la expresión de uno o más genes en la matriz. Esto puede ocurrir en varios contextos biológicos, como se divulga en la presente, por ejemplo, desarrollo y diferenciación, avance de una enfermedad, procedimientos *in vitro*, como transformación celular y senescencia, procesos neurológicos y neuronales autónomos, por ejemplo, dolor y apetito, y funciones cognitivas como el aprendizaje o la memoria.

La matriz también es útil para verificar el efecto de la expresión de un gen en la expresión de otros genes en la misma célula o en células diferentes. Esto proporciona, por ejemplo, una selección de objetivos moleculares alternativos para la intervención terapéutica si el objetivo final o aguas abajo no se puede regular.

La matriz también es útil para verificar patrones de expresión diferenciales de uno o más genes en células normales y enfermas. Esto proporciona una batería de genes que podrían servir como objetivo molecular para el diagnóstico o la intervención terapéutica.

VI. Medicina predictiva

La presente divulgación se refiere al campo de la medicina predictiva en la que se utilizan ensayos diagnósticos, ensayos predictivos, farmacogenética y ensayos clínicos de control con objetivos predictivos para, de ese modo, poder tratar a un individuo profilácticamente. Por consiguiente, un aspecto se refiere a ensayos diagnósticos para determinar proteínas marcadoras y/o la expresión de ácido nucleico así como también la actividad de una proteína marcadora, en el contexto de muestras biológicas (por ejemplo, sangre, suero, células, tejido) para determinar así si un individuo sufre de una enfermedad o trastorno, o está en riesgo de desarrollar un trastorno, asociado al aumento o disminución de la actividad o expresión de una proteína marcadora. También se proporcionan ensayos predictivos para determinar si un individuo está en riesgo de desarrollar un trastorno asociado a la expresión o actividad de una proteína marcadora o un ácido nucleico. Por ejemplo, el número de copias de un gen marcador se puede ensayar en una muestra biológica. Dichos ensayos se pueden utilizar con objetivos predictivos para así tratar a un individuo profilácticamente antes del comienzo de un trastorno (por ejemplo, diabetes tipo I), caracterizado por con la expresión o actividad de una proteína marcadora o ácido nucleico o asociado con estas.

Otro aspecto se refiere a controlar la influencia de agentes (por ejemplo, fármacos, compuestos) en la expresión o actividad de marcadores en ensayos clínicos.

En las siguientes secciones se describen en mayor detalle estos y otros agentes:

1. Ensayos diagnósticos

Un procedimiento ejemplarizante para detectar la presencia o ausencia de proteínas marcadoras o ácido nucleico divulgados en una muestra biológica implica obtener una muestra biológica de un individuo de prueba y poner a la muestra biológica en contacto con un compuesto o agente capaz de detectar la proteína el o ácido nucleico (por ejemplo ARNm, ADN genómico) que codifica la proteína marcadora, de modo tal que se detecta la presencia de la proteína marcadora o el ácido nucleico en la muestra biológica. Un agente preferido para detectar el ARNm o ADN genómico correspondiente a un gen marcador o proteína divulgada es una sonda de ácido nucleico marcada capaz de hibridarse a un ARNm o ADN genómico divulgado. En la presente se describen sondas apropiadas para utilizar en los ensayos diagnósticos divulgados.

Un agente preferido para detectar una proteína marcadora es un anticuerpo capaz de unirse a la proteína marcadora, preferentemente un anticuerpo con una marca detectable. Los anticuerpos pueden ser policlonales, o más preferentemente monoclonales. Se puede utilizar un anticuerpo intacto, o un fragmento de este (por ejemplo, Fab o F(ab')₂). El término “marcado”, con respecto a la sonda o al anticuerpo, abarca tanto el marcado directo de la sonda o anticuerpo mediante acoplamiento (es decir, uniendo físicamente) de una sustancia detectable a la sonda o anticuerpo, como el marcado indirecto de una sonda o anticuerpo mediante la reactividad con otro agente que está marcado directamente. Los ejemplos de marcado indirecto incluyen la detección de un anticuerpo primario utilizando un anticuerpo secundario marcado fluorescentemente y el marcado de los extremos de una sonda de ADN con biotina de modo que se lo pueda detectar con estreptavidina marcada fluorescentemente. El término “muestra biológica” incluye tejidos, células y fluidos biológicos aislados de un individuo, así como tejidos, células y fluidos presentes en el individuo. Es decir, el procedimiento de detección divulgado se puede utilizar para detectar ARNm marcador, proteína o ADN genómico en una muestra biológica *in vitro* al igual que *in vivo*. Por ejemplo, las técnicas *in vitro* para detectar el marcador ARNm incluyen las hibridaciones Northern e hibridaciones *in situ*. Las técnicas *in vitro* para detectar proteínas marcadoras incluyen ensayos por inmunoadsorción ligada a enzimas (ELISA), Western blots, inmunoprecipitaciones e inmunofluorescencia. Las técnicas *in vitro* para detectar ADN genómico marcador incluyen hibridaciones Southern. Además, las técnicas *in vivo* para detectar proteínas marcadoras incluyen introducir en un individuo un anticuerpo antimarcador marcado. Por ejemplo, el anticuerpo se puede marcar con un marcador radioactivo cuya presencia y ubicación en un individuo se puede detectar mediante técnicas de representación de imagen estándar.

En una realización, la muestra biológica contiene moléculas de proteínas del individuo de prueba. De manera alternativa, la muestra biológica puede contener moléculas de ARNm del individuo de prueba o moléculas de ADN genómico del individuo de prueba. Una muestra biológica preferida es una muestra de suero de un individuo aislada mediante formas convencionales.

En otra realización, los procedimientos implican además obtener una muestra biológica de control (por ejemplo, tejido no diabético) de un individuo de control, poner en contacto la muestra de control con un compuesto o agente capaz de detectar proteínas marcadoras, ARNm, o ADN genómico, de modo tal que se detecte la presencia de proteínas marcadoras, ARNm o ADN genómico en la muestra biológica, y comparar la presencia de proteínas marcadoras, ARNm o ADN genómico en la muestra de control con la presencia de proteínas marcadoras, ARNm o ADN genómico en la muestra de prueba.

También se divulgan kits para detectar la presencia de un marcador en una muestra biológica. Por ejemplo, el kit puede comprender un compuesto o agente marcado capaz de detectar proteínas marcadoras o ARNm en una muestra biológica; medios para determinar la cantidad de marcador en una muestra; y medios para comparar la cantidad de marcador en una muestra con un estándar. El compuesto o agente puede estar empaquetado en un envase apropiado. El kit puede comprender además instrucciones para utilizarlo para detectar proteínas marcadoras o ácido nucleico.

2. Ensayos predictivos

Los procedimientos diagnósticos aquí descritos pueden utilizarse además para identificar individuos que tienen o están en riesgo de desarrollar una enfermedad o trastorno asociado a la expresión o actividad marcadora aberrante. Como se utiliza en la presente, el término “aberrante” incluye una expresión marcadora o actividad que se desvía de la expresión o actividad marcadora silvestre. La expresión o actividad aberrante incluye aumentar o disminuir la expresión o actividad, así como la expresión o actividad que no sigue el patrón de desarrollo silvestre de expresión o el patrón subcelular de expresión. Por ejemplo, la expresión o actividad marcadora aberrante incluye los casos en que una mutación en el gen marcador provoca que el gen marcador esté subexpresado o sobreexpresado y situaciones en las que las mutaciones resultan en una proteína marcadora no funcional o una proteína que no funciona del modo silvestre, por ejemplo, una proteína que no interactúa con un ligando marcador A o uno que interactúa con un ligando proteico no marcador.

Los ensayos aquí descritos, como los ensayos diagnósticos precedentes o los ensayos siguientes, se pueden utilizar para identificar a un individuo que tiene o está en riesgo de desarrollar un trastorno asociado con la desregulación de la actividad de proteínas marcadoras o la expresión de ácido nucleico, como la diabetes tipo I. De manera alternativa, los ensayos predictivos se pueden utilizar para identificar a un individuo que tiene o está en riesgo de desarrollar un trastorno asociado con la desregulación de la actividad de proteínas marcadoras o la expresión de ácido nucleico, como la diabetes tipo I. Por lo tanto, se proporciona un procedimiento para identificar una enfermedad o trastorno asociado a la actividad o expresión marcadora aberrante, en el que se obtiene una muestra de prueba de un individuo y se detecta una proteína marcadora o ácido nucleico (por ejemplo, ARNm o ADN genómico), en el que la presencia de una proteína marcadora o ácido nucleico diagnostica a un individuo que tiene o está en riesgo de desarrollar una enfermedad o trastorno asociado con la expresión o actividad marcadora aberrante. Como se utiliza en la presente, una “muestra de prueba” incluye una muestra biológica obtenida del individuo en cuestión. Por ejemplo, una muestra de prueba puede ser un fluido biológico (por ejemplo, sangre), muestra celular o tejido (por ejemplo, tejido pancreático).

Además, los ensayos predictivos aquí descritos se pueden utilizar para determinar si se le puede administrar a un individuo un agente (por ejemplo, un agonista, antagonista, péptido mimético, proteína, péptido, ácido nucleico, molécula pequeña, u otro fármaco candidato) para tratar una enfermedad o trastorno asociado al aumento o disminución de la actividad o expresión marcadora. Por ejemplo, dichos procedimientos se pueden utilizar para determinar si un individuo puede ser tratado efectivamente con un agente para un trastorno como la diabetes tipo I. Por lo tanto, se proporcionan procedimientos para determinar si se puede tratar a un individuo efectivamente con agentes por un trastorno asociado al aumento o disminución de la actividad o expresión marcadora, en los que se obtiene una muestra de prueba y se detecta la expresión o actividad de una proteína marcadora o ácido nucleico (por ejemplo, en el que la abundancia de la actividad o expresión de proteínas marcadoras o ácido nucleico diagnostica que se le puede administrar a un individuo, el agente para tratar un trastorno asociado al aumento o disminución de la actividad o expresión marcadora).

Los procedimientos divulgados también se pueden utilizar para detectar alteraciones genéticas en un gen marcador, determinando de ese modo si un individuo con el gen alterado está en riesgo de tener un trastorno que se caracteriza por la desregulación de la actividad de las proteínas marcadoras o la expresión de ácido nucleico, como la diabetes tipo I. En realizaciones preferidas, los procedimientos incluyen detectar, en una muestra de células de un individuo, la presencia o ausencia de una alteración genética caracterizada por al menos una alteración que afecta la integridad de un gen que codifica una proteína marcadora, o la mala expresión de un gen marcador. Por ejemplo, tales alteraciones genéticas se pueden detectar asegurando la existencia de al menos uno de 1) una supresión de uno o más nucleótidos de un gen marcador; 2) una adición de uno o más nucleótidos a un gen marcador; 3) una sustitución de uno o más nucleótidos a un gen marcador; 4) una redistribución cromosómica de un gen marcador; 5) una alteración en el nivel de una transcripción de ARN mensajero de un gen marcador; 6) la modificación aberrante de un gen marcador, como el patrón de metilación del ADN genómico; 7) la presencia de un patrón de empalme no silvestre de una transcripción de ARN mensajero de un gen marcador; 8) el nivel no silvestre de una proteína marcadora; 9) pérdida alélica de un gen marcador; y 10) la modificación post-traducciona inapropiada de una proteína marcadora. Como se describe en la presente, hay un gran número de ensayos conocidos en la técnica que se pueden utilizar para detectar alteraciones en un gen marcador. Una muestra biológica preferida es un tejido (por ejemplo, tejido pancreático) o una muestra de sangre de un individuo aislada mediante formas convencionales.

En ciertas realizaciones, la detección de la alteración implica el uso de un/a cebador o sonda en una reacción en cadena de la polimerasa (PCR, por sus siglas en inglés) (ver, por ejemplo., las pats. EEUU Nos. 4,683,195 y 4,683,202), como PCR de anclaje o PCR RACE o, alternativamente, en una reacción en cadena de la ligasa (LCR, por sus siglas en inglés) (ver, por ejemplo, Landegran *et al.* (1988) *Science* 241:1077-1080; y Nakazawa *et al.* (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), éste último puede ser particularmente útil para detectar mutaciones de punto en los genes marcadores (ver Abravaya *et al.* (1995) *Nucleic Acids Res.* 23:675-682). Este procedimiento puede incluir los pasos de recolectar una muestra de células de un individuo, aislar el ácido nucleico (por ejemplo genómico, ARNm o ambos) de las células de la muestra, poner en contacto la muestra de ácido nucleico con uno o más cebadores que específicamente se hibridan a un gen marcador en condiciones tales que se produce la hibridación y amplificación del gen marcador (si está presente), y detectar la presencia o ausencia de un producto de amplificación y comparar la longitud con una muestra de control. Se prevé que puede ser deseable utilizar PCR y/o LCR como etapa preliminar de amplificación en conjunción con cualquiera de las técnicas utilizadas para detectar las mutaciones que aquí se describen.

Los procedimientos de amplificación alternativos incluyen: replicación de secuencia autosostenida (Guatelli, J.C *et al.*, (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), sistema de amplificación transcripcional (Kwoh, D. Y. *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi, P. M. *et al.* (1988) *Bio-Technology* 6:1197), o cualquier otro procedimiento de amplificación del ácido nucleico, seguido de la detección de las moléculas amplificadas utilizando técnicas conocidas por los entendidos en la técnica. Estos esquemas de detección son especialmente útiles para la detección de moléculas de ácido nucleico si dichas moléculas están presentes en números muy bajos.

En una realización alternativa, las mutaciones en un gen marcador de una muestra celular se pueden identificar por alteraciones en los patrones de escisión de la enzima de restricción. Por ejemplo, el ADN de muestra y el de control son aislados, amplificados (opcionalmente), digeridos con una o más endonucleasas de restricción, y se determinan los tamaños de longitud de los fragmentos mediante electroforesis en gel y se comparan. Las diferencias en los tamaños de la longitud de los fragmentos entre el ADN de la muestra y el de control indican mutaciones en el ADN de muestra. Además, las ribozimas específicas de secuencias (ver, por ejemplo, pat. EEUU No. 5,498,531) se pueden utilizar para llevar la cuenta de la presencia de mutaciones específicas mediante el desarrollo o pérdida de un sitio de escisión de una ribozima.

En otras realizaciones, las mutaciones genéticas en un gen marcador o un gen que codifica una proteína marcadora divulgada se pueden identificar hibridando ácidos nucleicos de muestra y de control, por ejemplo, ADN o ARN, en conjuntos de alta densidad que contengan cientos o miles de sondas de oligonucleótidos (Cronin, M. T *et al.* (1996) *Human Mutation* 7: 244-255; Kozal, M. J. *et al.* (1996) *Nature Medicine* 2: 753-759). Por ejemplo, las mutaciones genéticas en un marcador se pueden identificar en matrices en dos dimensiones que contengan sondas de ADN generadas por la luz como se describe en Cronin, M. T. *et al supra*. Brevemente, se puede utilizar una primera matriz de sondas de hibridación para buscar en largos tramos de ADN en una muestra y en el control para identificar cambios de base entre las secuencias haciendo matrices lineales de sondas secuenciales superpuestas. Este paso permite la identificación de mutaciones de punto. A este paso le sigue una segunda matriz de hibridación que permite la caracterización de mutaciones específicas utilizando matrices de sondas especializadas más pequeñas, complementarias a todas las variantes o mutaciones detectadas. Cada conjunto de mutaciones está compuesto de conjuntos de sondas paralelas, una complementaria al gen silvestre y otra complementaria al gen mutante.

En otra realización, se puede utilizar cualquiera de las muchas reacciones de secuenciación conocidas en la técnica para secuenciar directamente el gen marcador y detectar mutaciones comparando la secuencia del marcador de muestra con la correspondiente secuencia silvestre (de control). Los ejemplos de reacciones de secuenciación incluyen aquellos que se basan en técnicas desarrolladas por Maxam y Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) o Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). También se contempla que se puede utilizar cualquiera de una variedad de procedimientos de secuenciación automática cuando se realizan los ensayos diagnósticos ((1995) *Biotechniques* 19:448), inclusive secuenciación mediante espectrometría de masa (ver, por ejemplo., publicación internacional PCT No. WO 94/16101; Cohen *et al.* (1996) *Adv. Chromatogr.* 36:127-162; y Griffin *et al.* (1993) *AppL. Biochem. Biotechnol.* 38:147-159).

Otros procedimientos para detectar mutaciones en el gen marcador o el gen que codifica una proteína marcadora divulgada incluyen procedimientos en los que la protección de los agentes de escisión se utiliza para detectar bases mal apareadas en heterodúplex ARN/ARN o ARN/ADN h (Myers *et al.* (1985) *Science* 230:1242). En general, la técnica conocida de "escisión del mal apareamiento flog" comienza por proporcionar heterodúplex formados hibridando ARN o ADN (marcado) que contiene secuencias marcadoras silvestres con ARN o ADN potencialmente mutante obtenido de una muestra de tejido. Se trata los dúplex de doble hebra con un agente que escinde las regiones de hebra única del dúplex como las que existirían debido a los malos apareamientos en el par de bases entre las hebras de control y las de la muestra. Por ejemplo, los dúplex de ARN/ADN se pueden tratar con Ribonucleasa y los híbridos ADN/ADN se pueden tratar con nucleasa SI para digerir enzimáticamente las regiones con malos apareamientos. En otras realizaciones, se puede tratar tanto a los dúplex ADN/ADN como a los ARN/ADN con hidroxilamina o tetróxido de osmio y con piperidina a efectos de digerir regiones con malos apareamientos. Después de la digestión de las regiones con malos apareamientos, se separa el material resultante por tamaño en geles poliácridamida desnaturizante para determinar el sitio de mutación. Ver, por ejemplo, Cotton *et al.* (1988) *Proc. NatlAcad Sci USA* 85:4397; Saleeba *et al.* (1992) *Methods Enzymol.* 217:286-295. En una realización preferida, el ADN o ARN de control puede marcarse para su detección.

En otra realización más, la reacción de escisión equivocada emplea una o más proteínas que reconocen pares de bases mal apareados en ADN de doble hebra (llamadas enzimas de "reparación de un mal apareamiento de ADN" en sistemas guarecidos para detectar y mapear mutaciones de punto en ADNc marcadores obtenidos de muestras de células. Por ejemplo, la enzima mutY de *E. coli* escinde al A en las G/A mal apareadas y la timidina ADN glicosilasa de células HeLa escinde al T en las G/T mal apareadas (Hsu *et al.* (1994) *Carcinogenesis* 15:1657-1662). De conformidad con una realización ejemplarizante, se hibrida una sonda basada en una secuencia marcadora, por ejemplo, una secuencia marcadora silvestre, con un ADNc u otro producto de ADN de unas muestras celulares. Se trata el dúplex con una enzima reparadora de malos apareamientos de ADN y los productos de la escisión, de haberlos, se pueden detectar con protocolos de electroforesis o similares. Ver, por ejemplo, pat. EEUU No. 5,459,039.

En otras realizaciones, las alteraciones en la movilidad electroforética se utilizarán para identificar mutaciones en genes marcadores o genes que codifican una proteína marcadora divulgada. Por ejemplo, los polimorfismos de conformación de hebra única (SSCP) se pueden utilizar para detectar diferencias en la movilidad electroforética entre

ácidos nucleicos silvestres y mutantes. (1989) *Proc Natl. Acad. Sci USA*: 86:2766, ver también Cotton (1993) *Mutat. Res.* 285:125-144; y Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Los fragmentos de ADN de hebra única de ácido nucleico de muestra y de control serán desnaturalizados y se les permitirá renaturalizarse. La estructura secundaria de ácidos nucleicos de hebra única varía según la secuencia, la alteración en la movilidad electroforética resultante posibilita la detección de hasta un solo cambio de base. Los fragmentos de ADN se pueden marcar o detectar con sondas marcadas. La sensibilidad del ensayo puede aumentarse utilizando ARN (antes que ADN), en el que la estructura secundaria es más sensible a un cambio de secuencia. En una realización preferida, el procedimiento en cuestión utiliza el análisis de heterodúplex para separar moléculas de heterodúplex de doble hebra sobre la base de cambios en la movilidad electroforética (Keen *et al.* (1991) *Trends Genet* 7:5).

En otra realización adicional se ensaya el movimiento de fragmentos mutantes o silvestres en geles de poliacrilamida que contienen un gradiente de desnaturalización utilizando electroforesis en gel con gradiente de desnaturalización (DGGE) (Myers *et al.* (1985) *Nature* 313:495). Cuando se utiliza el DGGE como procedimiento de análisis, el ADN se modificará para asegurar que no desnaturaliza completamente, por ejemplo agregando una abrazadera GC de aproximadamente 40 bp de ADN rico en GC con un punto de fusión elevado mediante PCR. En otra realización, se utiliza una temperatura gradiente en lugar de un gradiente de desnaturalización para identificar las diferencias en la movilidad del ADN de control y de la muestra (Rosenbaum y Reissner (1987) *Biophys Chem* 265:12753).

Los ejemplos de otras técnicas para detectar mutaciones de punto incluyen, aunque no taxativamente, hibridación selectiva de oligonucleótidos, amplificación selectiva, o extensión selectiva de cebadores. Por ejemplo, se pueden preparar cebadores oligonucleótidos en los que la mutación conocida se coloca centralmente y luego se hibrida al ADN objetivo en condiciones que permiten la hibridación únicamente si se encuentra un apareamiento perfecto (Saiki *et al.* (1986) *Nature* 324:163); Saiki *et al.* (1989) *Proc. Natl. Acad. Sci USA* 86:6230). Dichos oligonucleótidos específicos de alelos se hibridan con el ADN objetivo amplificado por PCR o varias mutaciones diferentes cuando los oligonucleótidos están unidos a la membrana de hibridación y se hibridan con ADN objetivo marcado.

De manera alternativa, la tecnología de amplificación específica de alelos que depende de la amplificación PCR selectiva se puede utilizar junto con la presente invención. Los oligonucleótidos utilizados como cebadores para la amplificación específica pueden portar la mutación de interés en el centro de la molécula (para que la amplificación dependa de la hibridación diferencial) (Gibbs *et al.* (1989) *Nucleic Acids Res.* 17:2437-2448) o en el extremo 3' de un cebador donde, en condiciones apropiadas, los malos apareamientos pueden impedir o reducir la extensión de la polimerasa (Prossner (1993) *Tibtech* 11:238). Además, puede ser deseable introducir un nuevo sitio de restricción en la región de la mutación para crear la detección basada en la escisión (Gasparini *et al.* (1992) *Mol. Cell Probes* 6:1). Se prevé que en ciertas realizaciones la amplificación también se puede realizar utilizando Taq ligasa para la amplificación (Barany (1991) *Proc. Natl. Acad. Sci USA* 88:189) En tales casos, la ligación ocurrirá solamente si hay un apareamiento perfecto en el extremo 3' de la secuencia 5' con lo que se hace posible detectar la presencia de una mutación conocida en un sitio específico buscando la presencia o ausencia de amplificación.

Los procedimientos aquí descritos se pueden realizar, por ejemplo, utilizando kits diagnósticos empaquetados previamente que comprenden al menos una sonda de ácido nucleico o anticuerpo reactivo aquí descritos, que puede utilizarse convenientemente, por ejemplo, en escenarios clínicos para diagnosticar individuos que exhiben síntomas o una historia familiar de una enfermedad o afección que involucra un gen marcador.

Además, cualquier tipo de célula o tejido en el que se expresa el marcador se puede utilizar en los ensayos predictivos aquí descritos.

3. Control de efectos durante los ensayos clínicos

El control de la influencia de agentes (por ejemplo, fármacos) sobre la expresión o actividad de una proteína marcadora (por ejemplo, la modulación de la diabetes tipo I) se puede aplicar no sólo en el análisis de fármacos básicos, sino también en ensayos clínicos. Por ejemplo, la efectividad de un agente determinado mediante un ensayo de selección como se describe en la presente, para aumentar la expresión de los genes marcadores, los niveles de proteína o sobreregular la actividad marcadora, se puede controlar en ensayos clínicos de individuos que exhiben una expresión de genes marcadores, un nivel de proteínas reducido o la actividad de marcadores subregulada. De manera alternativa, la efectividad de un agente determinado mediante un ensayo de búsqueda como se describe en la presente para disminuir la expresión de los genes marcadores, los niveles de proteína, o subregular la actividad marcadora, se puede controlar en ensayos clínicos de individuos que exhiben una expresión de genes marcadores, un nivel de proteínas aumentados o una actividad de marcadores sobreregularada. En dichos ensayos clínicos, la expresión o actividad de un gen marcador, y preferentemente otros genes que están implicados en, por ejemplo, un trastorno asociado a los marcadores (por ejemplo la diabetes tipo I) se puede utilizar como una "lectura" o marcadores del fenotipo de una célula particular.

A modo de ejemplo no limitante, se pueden identificar los genes, incluidos los genes marcadores y genes que codifican una proteína marcadora divulgados, que están modulados en células mediante el tratamiento con agentes (por ejemplo, un compuesto, fármaco o pequeña molécula) que modulan la actividad marcadora (por ejemplo, identificada en un ensayo de selección como se describe en la presente). Por lo tanto, para estudiar los efectos de los agentes en trastornos asociados a los marcadores (por ejemplo diabetes tipo I), por ejemplo, en un ensayo clínico, las células se

pueden aislar y se puede preparar ARN y analizar los niveles de expresión de un marcador y otros genes implicados en el trastorno asociado al marcador, respectivamente. Se pueden cuantificar los niveles de expresión de genes (por ejemplo, un patrón de expresión de un gen) mediante análisis Northern blot o PCR en transcripción inversa (RT-PCR, por sus siglas en inglés), como se describe en la presente, o alternativamente, midiendo la cantidad de proteína producida, por uno de los procedimientos aquí descritos, o midiendo los niveles de actividad de un marcador u otros genes. de este modo, el patrón de expresión de un gen puede servir como marcador, indicador de la respuesta fisiológica de las células al agente. Por consiguiente, este estado de respuesta se puede determinar antes y en varios momentos durante el tratamiento del individuo con el agente.

En una realización preferida, se proporciona un procedimiento para controlar la efectividad del tratamiento de un individuo con un agente (por ejemplo, un agonista, antagonista, péptido mimético, proteína, péptido, ácido nucleico, pequeña molécula u otro fármaco candidato identificado por los ensayos de búsqueda aquí descritos) incluidos las etapas de (i) obtener una muestra del individuo antes de la administración del agente; (ii) detectar el nivel de expresión de una proteína marcadora, ARNm, o ADN genómico en la muestra antes de la administración; (iii) obtener una o más muestras del individuo posteriores a la administración; (iv) detectar el nivel de expresión o actividad de la proteína marcadora, ARNm, o ADN genómico en las muestras después de la administración; (V) comparar el nivel de expresión o actividad de la proteína marcadora, ARNm o ADN genómico en las muestras anteriores a la administración con la proteína marcadora, ARNm o ADN genómico en las muestras posteriores a la administración; y (vi) alterar la administración del agente al individuo como corresponda. Por ejemplo, puede ser deseable aumentar la administración del agente para elevar la expresión o actividad de un marcador a niveles más altos que los detectados, es decir aumentar la efectividad del agente. Alternativamente puede ser deseable disminuir la administración del agente para disminuir la expresión o actividad de un marcador a niveles más bajos que los detectados, es decir disminuir la efectividad del agente. De conformidad con tal realización, la expresión o actividad de los marcadores se puede utilizar como un indicador de la efectividad de un agente, aún en la ausencia de una respuesta fenotípica observable.

Farmacogenómica

Las moléculas de proteínas marcadoras y ácido nucleico divulgadas, así como agentes o moduladores que tienen un efecto estimulante o inhibidor de la actividad de las proteínas marcadoras (por ejemplo, la expresión de genes marcadores) tal como las identifica el ensayo de búsqueda aquí descrito se puede administrar a individuos que se desee tratar (en forma profiláctica o terapéutica) trastornos asociados a los marcadores (por ejemplo, diabetes tipo I o una condición asociada a la NKT) asociados con la actividad aberrante de la proteína marcadora. En conjunto con dicho tratamiento, se puede considerar la farmacogenómica (es decir, el estudio de la relación entre el genotipo del individuo y la respuesta de ese individuo a un compuesto o fármaco extraño). Las diferencias en el metabolismo de las terapéuticas pueden llevar a una toxicidad severa o al fracaso terapéutico al alterar la relación entre la dosis y la concentración en sangre del fármaco farmacológicamente activo. Por lo tanto, un médico puede considerar aplicar el conocimiento obtenido en estudios farmacogenómicos relevantes para determinar si conviene administrar o no una molécula marcadora o un modulador marcador, así como también confeccionar la dosificación y/o régimen terapéutico del tratamiento con una molécula marcadora o modulador marcador.

La farmacogenómica trata con variaciones hereditarias clínicamente significativas en respuesta a fármacos debido a una disposición a los fármacos alterada y a la acción anormal en personas afectadas. Ver, por ejemplo, Eichelbaum, M. *et al* (1996) *Clin. Exp. Pharmacol. Physiol.* 23(10-11) :983-985 y Linder, M. W. *et al.* (1997) *Clin. Chem.* 43 (2):254-266. Por lo general, se pueden diferenciar dos tipos de trastornos farmacogénéticos. Condiciones genéticas transmitidas como un factor único que altera la forma en que los fármacos actúan en el cuerpo (acción alterada de los fármacos) o trastornos genéticos transmitidos como factores únicos que alteran la forma en que el cuerpo actúa sobre los fármacos (metabolismo alterado de los fármacos). Estos trastornos farmacogénéticos pueden ocurrir ya sea como defectos genéticos raros o como polimorfismos existentes naturalmente. Por ejemplo, la deficiencia de glucosa-6-fosfato deshidrogenasa (G6PD por sus siglas en inglés) es una enzimopatía hereditaria común en la que la principal complicación clínica es la hemólisis después de la ingestión de fármacos oxidantes (anti maláricos, sulfonamidas, analgésicos, nitrofuranos) y el consumo de habas.

Un enfoque farmacogenómico para identificar los genes que predicen la respuesta al fármaco, conocido como “asociación del genoma completo”, se basa principalmente en un mapa de alta resolución del genoma humano que consiste en marcadores relacionados con genes, ya conocidos, (por ejemplo, un mapa marcador de genes “bialélicos” que consiste en 60.000-100.000 sitios polimórficos o variables en el genoma humano, cada uno de los cuales tiene dos variantes). Tal mapa genético de alta resolución se puede comparar con un mapa del genoma de cada uno de un número estadísticamente significativo de individuos que participen de un ensayo de fármacos de Fase II/III para identificar marcadores asociados a una respuesta a un fármaco particularmente observada o un efecto secundario. Alternativamente, dicho mapa de alta resolución se puede generar a partir de la combinación de alrededor de diez millones de polimorfismos nucleotídicos únicos conocidos (SNP) en el genoma humano. Como se utiliza en la presente, un “SNP” es una alteración común que ocurre en una base nucleotídica simple en un tramo de ADN. Por ejemplo, un SNP puede ocurrir una vez cada 1000 bases de ADN. Aunque SNP puede estar involucrado en un proceso de enfermedad, la gran mayoría no se puede asociar a la enfermedad. En función un mapa genético basado en la ocurrencia de tales SNP, los individuos se pueden agrupar en categorías genéticas según un patrón particular de SNP en su genoma individual. De tal forma, se pueden confeccionar los regímenes de tratamiento para grupos de individuos genéticamente similares, tomando en cuenta los rasgos que pueden ser comunes entre estos individuos genéticamente similares.

De manera alternativa, se puede utilizar un procedimiento llamado “enfoque del gen candidato” para identificar los genes que predicen la respuesta a los fármacos. De conformidad con este método, si un gen que codifica al objetivo de un fármaco es conocido (por ejemplo, una proteína marcadora divulgada), todas las variantes comunes de ese gen pueden ser identificadas con bastante facilidad en la población y se puede determinar si tener una versión del gen a diferencia de otra, está asociado con alguna respuesta particular al fármaco.

Como realización ilustrativa, la actividad de las enzimas que metabolizan el fármaco es un determinante principal tanto de la intensidad como de la duración de la acción del fármaco. El descubrimiento de polimorfismos genéticos de enzimas metabolizadoras de fármacos (por ejemplo, N-acetiltransferasa 2 (NAT 2) y las enzimas citocromo P450, CYP2D6 y CYP2C19) han proporcionado una explicación de porqué algunos individuos no obtienen los efectos esperados de un fármaco o presentan una respuesta exagerada al fármaco y una seria toxicidad después de tomar dosis estándar y seguras de un fármaco. Estos polimorfismos se expresan en dos fenotipos en la población, el metabolizador rápido (EM, por sus siglas en inglés) y el metabolizador lento (PM, por sus siglas en inglés). La prevalencia del PM es diferente en las diferentes poblaciones. Por ejemplo, el gen que codifica al CYP2D6 es altamente polimórfico y se han identificado varias mutaciones en el PM, que todas conducen a la ausencia de CYP2D6 funcional. Los metabolizadores lentos de CYP2D6 y CYP2C 19 con frecuencia experimentan una respuesta exagerada a los fármacos y efectos secundarios cuando reciben dosis estándar. Si un metabolito es la porción terapéutica activa, el PM no muestra respuesta terapéutica, como lo demuestra el efecto analgésico de la codeína mediado por su metabolito de morfina formada con CYP2D6. El otro extremo son los llamados metabolizadores ultra rápidos quienes no responden a dosis estándar. Recientemente, se identificó que la base molecular del metabolismo ultra rápido se debe a una amplificación del gen CYP2D6.

De manera alternativa, un procedimiento llamado el “perfil de expresión génica” se puede utilizar para identificar los genes que predicen la respuesta a los fármacos. Por ejemplo, la expresión génica de un animal al que se le administró una dosis de un fármaco (por ejemplo una molécula marcadora o modulador marcador divulgado) puede dar una indicación de si se han activado pasajes génicos relacionados con la toxicidad.

La información generada a partir de más de uno de los abordajes farmacogenómicos mencionados se puede utilizar para determinar la dosificación apropiada y los regímenes de tratamiento para el tratamiento profiláctico o terapéutico de un individuo. Este conocimiento, cuando se lo aplica a la dosificación o selección de fármacos, puede evitar reacciones adversas o fracasos terapéuticos y de ese modo aumentar la eficiencia terapéutica o profiláctica cuando se trata un individuo con una molécula marcadora o modulador marcador, por ejemplo un modulador identificado por uno de los ensayos de búsqueda aquí descritos.

Esta invención se ilustra aún más mediante los siguientes ejemplos que no se deben interpretar como taxativos.

Ejemplos

Ejemplo 1

Estudios inhibidores

Una diferencia conocida entre las células NKT en individuos diabéticos respecto de individuos no diabéticos es que mientras que se sabe que ambas células secretan IFN- γ al activarse, las células NKT diabéticas no secretan IL-4 con la activación, mientras que las células NKT no diabéticas sí lo hacen. Para identificar cual de las cascadas de señales conocidas iniciadas por la unión del receptor de antígeno T tenía un rol dominante en la secreción de IL-4, se realizaron una serie de estudios inhibidores. (Wilson *et al.* (2000) *PNA* 97:7411-7416).

Se cultivaron distribuciones de células únicas de V α 24-positiva, CD4/8 negativa en alimentadores alogénicos irradiados a razón de 50.000 células por pozo con 5.000 células por pozo irradiadas (5.000 rads) 721.221 células linfoblastoides con 1 μ g/ml PHA-P, IL-2 y IL-7 cada uno a 10 unidades/ml (Boehringer Mannheim) y se las propagó como se describe (Wilson *et al.* (1998) *Nature* 391: 177-181). Se ensayaron clones positivos de V α 24 y NKR-PIA mediante citometría de flujo y una secuencia receptora de antígeno de células T y V α 24J α Q CDR3 para ver su secreción de citoquina en experimentos de restricción de C1R/CD1s. Para los estudios de secreción e inhibición de citoquina y inhibidores, se activaron clones de células V α 24J α Q T GW4 (no-diabéticas) y ME10 (diabéticas) a 5x10⁴ células por pozo, con anti-CD3 unido a las placas o Ig control a 1 μ g/ml. Se ensayaron IL-4 y IFN- γ secretados, mediante ELISA después de 4 h de activación como se describió (Wilson *et al.* (1998) *Nature* 391: 177-181). Previamente se determinaron las concentraciones óptimas de inhibidores mediante experimentos de inhibición de la respuesta a la dosis. Las concentraciones de inhibidores utilizadas fueron 10 nM wortmannin; 10 μ M LY294002, 50 μ M PD98059, un inhibidor de quinasa proteica activada con mitógeno y 50 μ M SB203580, un inhibidor de quinasa p38. Las concentraciones de éster de forbol y ionóforo de calcio utilizadas fueron 1 ng/ml forbol 12-miristato 13-acetato (PMA) y 1 μ g/ml ionomicina. Se utilizó ciclosporina A (CsA) a 5 ng/ml como control negativo, debido a su capacidad para impedir la activación de células T. Se determinó el flujo de calcio cargando las células con indo-1 según las indicaciones del fabricante (Molecular Probes), seguido de la activación con anti-CD3 a 10 μ g/ml. El flujo máximo de calcio se determinó agregando ionomicina.

Ambos inhibidores de fosfoinositida-3-OH quinasa (PI3 quinasa) wortmannin y LY294002 bloquearon la secreción de IL-4 indicada por anti-CD3 a partir del clon IL-4⁺, pero no tuvieron efecto en la secreción de IFN- γ de los clones IL-4⁺ ni del IL-4-nulo (Fig. 2). En contraste, la inhibición de la quinasa MEK por PD98059 o JNK y las cascadas p38 con SB203580 (Rincon y Flavell (1997) *Curr. Biol.* 7(11):R729-32; Dumont *et al.* (1998) *J. Immunol.* 160(6):2579-89) no tuvieron efecto. Después de la inhibición de la quinasa PI3 la secreción IL-4 se pudo rescatar en el clon IL-4⁺ mediante la inclusión del éster de forbol PMA o ionomicina de ionóforo de calcio. Ninguna de estas sustancias solas ni en combinación reparó el defecto de la secreción de IL-4 del clon ME10 derivado diabético. Además, los clones de las células V α 24J α Q T derivadas de individuos diabéticos tenían una capacidad disminuida para acumular calcio intracelular después de la estimulación con anti-CD3 (Fig. 2). Estos datos sugieren que el fenotipo IL-4 discordante observado que se vio después de la ligación del receptor de antígeno de células T no puede ubicarse simplemente aguas arriba de la quinasa PI3 y que posiblemente haya varias diferencias genéticas/reguladoras implicadas, incluidas las diferencias en las proteínas que regulan el flujo de calcio.

Ejemplo 2

Identificación y caracterización de marcador de ADNc

A. Citometría de flujo

Se tiñeron leucocitos de sangre periférica con anticuerpos monoclonales conjugados fluorescentemente. Se analizaron las células teñidas en un citómetro FACScan (Beckton Dickinson) y se realizaron las distribuciones de célula única y flujo de calcio utilizando un citómetro MoFlo (Citomation, Fort Collins, N.J.) como se describió (Wilson *et al.* (1998) *Nature* 391: 177-181).

B. Cultivo celular

Se cultivaron distribuciones de células únicas de V α 24-positiva, CD4/8 negativa en alimentadores alogénicos irradiados a 50.000 células por pozo con 5.000 células por pozo irradiadas (5.000 rads) 721.221 células linfoblastoides con 1 μ g/ml PHA-P, IL-2 e IL-7 cada uno a 10 unidades/ml (Boehringer Mannheim) y se las propagó como se describe (Wilson *et al.* (1998) *Nature* 391: 177-181).

C. Expresión del ARN mensajero

Se activaron clones de células V α 24J α Q T GW4 y ME10 (1x10⁷ células) durante 4 h con 10 μ g/ml de anti-CD3 soluble o control IgG. Se eligió un tiempo de 4 h debido a que se lo había utilizado en un análisis previo de secreción de citoquina en clones derivados de pares de gemelos monocigóticos discordantes para la diabetes tipo 1 (Wilson *et al.* (1998) *Nature* 391(6663): 177-81). Previamente se determinaron las concentraciones óptimas de anti CD-3 mediante experimentos de respuesta a dosis midiendo la secreción de citoquina. Se aisló todo el ARN con kits Qiagen Rneasy. Después se convirtió todo el ARN en ADNc de doble hebra cebándolo con un cebador oligo (dT) que incluía un sitio promotor de polimerasa ARN T7 en el extremo 5'. El ADNc fue utilizado directamente en una reacción de transcripción *in vitro* en presencia de nucleótidos biotinilados para producir ARNc marcado (ARN antisentido), que se hibridó durante la noche con Genechips (Affymetrix, San Jose, Calif.). Después de teñir con ficoeritrina-estreptavidina, se cuantificó la fluorescencia del ARN unido utilizando un lector GeneChip (un microscopio confocal modificado; Affymetrix), utilizando protocolos estándar.

D. Análisis de los datos

El número de genes con expresión detectable ya sea antes o después de la estimulación era casi idéntico para clones secretores IL-4 nulo e IL-4- (1.523 y 1.558, respectivamente). La frecuencia de expresión de la mayoría de las transcripciones no cambió con la activación. El número de genes cuya expresión después de la estimulación con anti-CD3 aumentó o disminuyó en al menos dos veces con respecto con los genes no estimulados fue de 86 (6%) y 226 (15%) en los clones IL-4-nulo e IL-4, respectivamente.

Para analizar más concienzudamente las diferencias en la expresión génica entre los clones IL-4-nulo y IL-4 secretores, se agruparon los genes en seis diseños de expresión distintos, utilizando el algoritmo del mapa autoorganizado (Figura 3) (Tamayo *et al.* (1999) *Proc. Natl. Acad. Sci. USA* 96: 2907-2912). Se conglomeraron todos los genes modulados al menos 2 veces con la estimulación anti-CD3 ya sea en clones IL-4-secretadores o IL-4-nulo de conformidad al comportamiento relativo de cada gen en los dos clones. El primer panel de la Fig. 3 muestra los resultados para todos los genes que cumple el criterio doble, y los otros 11 paneles muestran los resultados de clases funcionales específicas. El patrón dominante que emergió se representa en la fila 1, columna 2, y contiene genes que se sobreexpresaron con la activación en el clon secretor de IL-4 pero que no respondieron a la estimulación en el clon de IL-4 nulo. Este hallazgo se aplicó a todas las clases funcionales examinadas, lo que indica un defecto profundo en la inducción transcripcional de un gran número de genes en el clon IL-4-nulo. Sin embargo, el examen de los otros cinco conglomerados reveló que la desregulación transcripcional en el clon IL-4-nulo es más compleja que una mera falta de respuesta, como lo

demuestra un grupo de genes que fueron inducidos en este clon pero no en el clon secretor de IL-4- (fila 1, columna 1 (Tabla 1)) y por un grupo que contenía genes que fueron subregulados en el clon IL-4-nulo pero sobreexpresados en el clon secretor de IL-4 (fila 2, columna 2 (Tabla 5)). Por lo tanto, el clon IL-4-nulo es capaz de responder a la estimulación a través del receptor de células T.

Se observaron seis patrones de expresión diferentes. En la Tabla 1 (representativa de la fila 1, columna 1, en cada uno de los conglomerados establecidos en la Fig. 3) se enumera cada uno de los genes que se observó aumentaron en la expresión en células NKT diabéticas activadas y que no cambiaban o que aumentaban en menor grado en su expresión en células NKT no diabéticas activadas, en relación con células de control en reposo apropiadas. En la Tabla 2 (representativa de la fila 1, columna 2, en cada uno de los conglomerados establecidos en la Fig. 3) se enumera cada uno de los genes que se observó no cambiaban en la expresión en células NKT diabéticas activadas, en relación con células de control en reposo, pero que aumentan en su expresión en células NKT no diabéticas activadas, en relación con células de control en reposo. En la Tabla 3 (representativa de la fila 1, columna 3 en cada uno de los conglomerados establecidos en la Fig. 3) se enumera cada uno de los genes que se observó que aumentaron en la expresión tanto en células NKT diabéticas como no diabéticas activadas en relación con células de control en reposo apropiadas. En la Tabla 4 (representativa de la fila 2, columna 1 en cada uno de los conglomerados establecidos en la Fig. 3) se enumera aquellos genes con una menor expresión en células NKT no diabéticas activadas en relación con células de control en reposo, pero que no cambiaban en su expresión en células NKT diabéticas activadas, en relación con células de control en reposo. En la Tabla 5 (representativa de la fila 2, columna 2 en cada uno de los conglomerados establecidos en la Fig. 3) se enumera aquellos genes con una mayor expresión en células NKT no diabéticas activadas en relación con células de control en reposo, pero que disminuyen en su expresión en células NKT diabéticas activadas, en relación con células de control en reposo. En la Tabla 6 (representativa de la fila 2, columna 3 en cada uno de los conglomerados establecidos en la Fig. 3) se enumeran aquellos genes con una menor expresión en células NKT diabéticas activadas en relación con células de control en reposo, pero que no cambiaban o disminuían en un menor grado en su expresión en células NKT no diabéticas activadas, en relación con células de control en reposo.

Se observaron cambios significativos de miembros de la familia citoquina/quimioquina entre los clones IL-4+ y IL-4-nulo y se los confirmó en el nivel proteico mediante ensayo ELISA. Las diferencias de expresión significativas también se detectaron en otros genes importantes para la supervivencia celular, la secreción de citoquina y el flujo de calcio, que en parte se activan a través de la señal de quinasa PI3-, como BCLxL, IAP, PLCgamma y la familia de quinasas tec, Itk, transcritos que se encontraron en gran abundancia en el clon IL-4+. También se observaron diferencias en la expresión del ARNm que codifica los factores de transcripción y moduladores de señal importantes para la secreción de citoquina y el fenotipo Th, incluidos GATA3, STAT1, STAT4, JunB, JunD, y NFAT4. Algunos de estos genes tenían una mayor expresión en el clon IL-4+ (GATA3, JunB y JunD), mientras que los restantes aumentaron su expresión en el clon IL-4-nulo.

Ejemplo 3

40 Expresión génica en clones de células T NKT, CD4 y CD8

A. Enfoque

Se generaron clones de células T NKT, CD4 y CD8 a partir de un único donante utilizando los procedimientos antes descritos. Se seleccionó un único clon de cada tipo y se lo estimuló con anti-CD3 durante 2, 4, 8, 24 ó 48 horas para crear una serie de activación cinética. Se realizaron tres experimentos de replicación para CD4 y CD8, pero se realizó sólo una replicación para NKT. La búsqueda de una replicación de NKT requirió tres cambios de filtro. No se necesitó ningún cambio de filtro para los CD4 y CD8, sino que el filtro que se impuso fue que los genes tenían que aumentar o disminuir con respecto a 0 horas para las tres réplicas. Se aisló y analizó el ARN en chips que controlaban 12.000 genes humanos conocidos. Se buscó la información de expresión para genes con patrones de expresión coherentes entre las tres repeticiones, y con tres cambios entre los diferentes subconjuntos de células T.

B. Resultados

Los genes que se identificaron en una búsqueda que requería incrementar en al menos tres veces los niveles de ARNm en al menos un punto de tiempo en la muestra de células NK y que no tenían un patrón de expresión o magnitud diferente a las muestras celulares de CD4 y CD8 se establecen en la Tabla 9. Los genes que se identificaron en una búsqueda que requiere un aumento de los niveles de ARNm en al menos un punto de tiempo para las tres repeticiones del experimento en muestras de células CD4 y que tenían un patrón de expresión o magnitud diferentes a las muestras de células NKT y CD8 se establecen en la Tabla 10. Los genes que se establecen en la Tabla 9 se excluyeron de la Tabla 10. Los genes que se identificaron en una búsqueda que requería un aumento de los niveles de ARNm en al menos un punto de tiempo de las tres repeticiones del experimento en muestras de células CD8 y tenían un patrón de expresión o magnitud diferente a las muestras de células NKT o CD4 se establecen en la Tabla 11. Los genes establecidos en las Tablas 9 y 10 se excluyeron de la Tabla 11.

Ejemplo 4

Expresión génica en clones de células T NKT, CD4 y CD8 no estimuladas

5 A. Enfoque

Se generaron clones de células T NKT, CD4 y CD8 a partir de un único donante utilizando los procedimientos antes descritos. Se seleccionó un único clon de cada tipo y se observaron las diferencias en la expresión génica en células en reposo. Se realizaron tres experimentos de replicación para CD4 y CD8 y se realizaron dos replicaciones para NKT. Se aisló y analizó el ARN en chips que controlaba 12.000 genes humanos conocidos. Se buscó la información de expresión para genes con patrones de expresión coherentes entre las tres repeticiones y con tres cambios entre los diferentes subconjuntos de células T.

15 B. Resultados

Los genes que se identificaron en una búsqueda que requería un cambio en los niveles de ARNm para las tres replicaciones del experimento en muestras celulares de CD4 en reposo y que tenían un patrón de expresión o magnitud diferente a las muestras celulares de NKT en reposo se establecen en la Tabla 12. Los genes que se identificaron en una búsqueda que requería un cambio en los niveles de ARNm para las tres replicaciones del experimento en muestras celulares de CD8 en reposo y que tenían un patrón de expresión o magnitud diferente a las muestras celulares de NKT en reposo se establecen en la Tabla 13.

25

TABLA 1

<u>Citoquina</u>	<u>Quinasa/fosfatasa</u>	<u>Transducción de señal</u>
U89922 LT- β (1,1)	X79510 PTP DI (1,1)	HT5108 TRAP-3 (1,1)
		X80200 MLN62 (1,1)
<u>Receptor de superficie</u>	<u>Metabolismo RNA</u>	
U38276 Semaphorin III (1,1)	D38251 RNP B5 (1,1)	
U82169 Frizzled (1,1)	U90547 RNP homólogo (1,1)	
<u>Citoesqueleto</u>		
U80184 Flightless Thom. (1,1)		
<u>Proteína nuclear</u>		
U73477 Nuclear pp32 (1,1)		

50

55

60

65

ES 2 331 407 T3

TABLA 2

	<u>Citoquina</u>	<u>Quinasa/fosfatasa</u>	<u>Factor de transcripcion</u>
5	J00219 IFN- γ (1,2)	L10717 ITK (1,2)	M69043 I κ B α (1,2)
	V00536 LFN- γ (1,2)	X60673 AK3 (1,2)	X58072 GATA-3 (1,2)
	M13207 GM-CSF (1,2)	X85545 PKX-1 (1,2)	U43185 STAT-5A (1,2)
	M16441 TNF-cL (1,2)	D13720 LYK (1,2)	X51345 Jun-B (1,2)
10	X02910 TNF-cL (1,2)	HT1153 Nm23-H2S	X56681 Jun-D (1,2)
	X04688 IL-5 (1,2)	(1,2)	U15460 B-ATF (1,2)
	U311201L-13(1,2)	M30448 CK II 3 (1,2)	HT4899 C-myc (1,2)
15	M37435 M-CSF (1,2)	M90299 Glucokinase	L00058 C-myc (1,2)
	U02020 PBEF (1,2)	(1,2)	M13929 C-myc (1,2)
	U37518 TRAJL (1,2)	U08316 ISPK-1 (1,2)	U26173 NF-IL3A (1,2)
20	U46461 Dishevelled hom	X80910 PPP1CB (1,2)	M97796 Id-2 (1,2)
	(1,2)	X93 920 DUSP-6 (1,2)	M96843 Id-2B (1,2)
			D14826 CREM (1,2)
			S68271 CREM(1,2)
25	<u>Receptor de superficie</u>	<u>Metabolismo ARN</u>	J03827 Y Box BP (1,2)
	M32315 TNF-R (1,2)	X17567 RNP B (1,2)	U09412 ZNF134 (1,2)
	U03397 4-1BB (1,2)	M29064 RNP B1 (1,2)	U13044 NRF-2c (1,2)
30	S77812 VEGF-R (1,2)	HT110 RNP AJB (1,2)	U22431 H1F-1x (1,2)
	X01057 IL-2RcL (1,2)	Z23064 RNP 0(1,2)	X78925 HZF-2 (1,2)
	Y00285 IGF—R 11(1,2)	HT3238 RNP K (1,2)	Z47727 RNA POL2K (1,2)
	L08096 CD27 (1,2)	X52979 RNP SmB (1,2)	
35	Z30426 CD69 (1,2)	U15009 RNP SmD3	<u>Transducción de señal</u>
	U76764 CD97 (1,2)	(1,2)	U20158 SLP-76 (1,2)
	U60800 CD100 (1,2)	X85372 RNP Sm F (1,2)	U26710 Cbl-b (1,2)
40	M24283 Rhinovirus-R (1,2)	U30827 SP SRp40 (1,2)	D78132 RHEB (1,2)
	U19906 Arg. Vasopressin-R	X70944 SF (PTP-assoc.)	M63'573 SCYLP (1,2)
	(1,2)	(1,2)	M75099 FK506 BP (1,2),
	Z48042 p137 (1,2)	M60858 Nucleolin (1,2)	Z35227 TTF(1,2)
45		U10323 NF45 (1,2)	
	<u>Citoesqueleto</u>	U38846 Stim. of TAR	<u>Metabolismo de Proteína</u>
	X00351 3-Actin (1,2)	(1,2)	D28473 ILE-tRNA Synth.
50	U20582 Actin-like pep. (1,2)	X59417 PROS-27 (1,2)	(1,2)
	X82207 13-centractin (1,2)	X59892 IFN-Ind. γ 2 (1,2)	U09510 GLY-tRNA Synth.
	X98534 VASP (1,2)	X66899 EWS (1,2)	(1,2)
55		X71428 fus (1,2)	L25085 Sec61-3 (1,2)
	<u>Proteína nuclear</u>	X72727 Tuncp (1,2)	X74801 Chaperonin Cctg (1,2)
	U62962 Int-6 (1,2)	X75755 PR264 (1,2)	X77584 Thioredoxin (1,2)
		Z24724 Poly A site (1,2)	Y00281 Ribophorin 1(1,2)
60		<u>Quimioquina</u>	
		M23 178 MIP-1c (1,2)	<u>Apoptosis</u>
		J04130 lMIP-113 (1,2)	Z23115 Bcl-XL(1,2)
65		M69203 MCP-1 (1,2)	U45878 IAP-1 (1,2)
			U11821 Fas Ligand (1,2)
			S819141EX-1 (1,2)
			U37546 MIHC (1,2)

ES 2 331 407 T3

TABLA 3

5	<u>Receptor de superficie</u> D79206 Ryudocan (1,3) HT3 125 CD44 (1,3)	<u>Quinasa/fosfatasa</u> U24 152 PAK-1 (1,3) D11327 PTPN7 (1,3) U15932 DUSP-5 (1,3)	<u>Factor de transcripción</u> J04076 EGR-2 (1,3) D61380 DJ-1 (1,3) HT4567 PC4 (1,3)
	<u>Proteína nuclear</u> L2593 1 Lamin B Rec. (1,3)	<u>Metabolismo ARN</u> L28010 RNP F (1,3) HT4788 RNP 1(1,3) L03 532 M4 (1,3)	<u>Transducción de señal</u> U19261 EBV-Ind. (1,3)
			<u>Metabolismo de proteína</u> YI 0807 ARG-metiltrans. (1,3) D13748 EIF-4A1 (1,3)
25		<u>Quimioquina</u> L19686 MIF (1,3)	

TABLA 4

35	<u>Receptor de superficie</u> L39064 IL-9R (2,1) X14046 CD37 (2,1) L31584EB1-1 (2,1) X97267 LPAP (2,1)	<u>Quinasa/fosfatasa</u> L16862 GRK-6 (2,1) L27071 TXK(2,1)	<u>Factor de transcripción</u> HT4921 BTF-3 hom. (2,1)
	<u>Citoesqueleto</u> D83735 Calponin (2,1)		<u>Metabolismo de proteína</u> X55733 EIF-4B (2,1)

TABLA 5

55	<u>Receptor de superficie</u> M33680 TAPA-1 (2,2) M63 175 AMFR (2,2) U60975 gp250 (2,2) Z50022 C2lorf3 (2,2)	<u>Quinasa/fosfatasa</u> J03805 PPP2CB (2,2)	<u>Transducción de señal</u> M28209 RAB-1 (2,2)

ES 2 331 407 T3

TABLA 6

5	<u>Citoquina</u> M90391 IL-16 (2,3)	<u>Quinasa/fosfatasa</u> HT3678 CLK-1 (2,3) U66464 HPK-1 (2,3) X62535 DAG Quinasa (2,3) M31724 PTP-1B (2,3)	<u>Factor de Transcripción</u> L41067 NFÁT-4C (2,3) L78440 STAT-4 (2,3) M82882 ELF-1 (2,3) M83667 NF-IL6 (2,3)
	<u>Receptor de superficie</u> U90546 Butyrophilin BT4 (2,3) U90552 Butyrophilin BT5 (2,3) X96719 AICL (2,3)	<u>Metabolismo ARN</u> U69546 RNA BP (2,3)	<u>Transducción de señal</u> D78577 14-3-3-Eta (2,3) X89399 Ins(1345)P4 BP (2,3)
	<u>Citoesqueleto</u> J00314 -βtubu1in (2,3) M21812 Myosin LC (2,3) X98411 Myosin-IE (2,3)		<u>Metabolismo de proteína</u> X76648 Glutaredoxin (2,3)
	<u>Proteína nuclear</u> M17733 Thymosin-β4 (2,3)		

TABLA 7

<u>Número de acceso</u>	<u>Nombre del Gen</u>	<u>Cita</u>
	PGE-2 synthase	Litherland <i>et al.</i> (1999) J <i>Clin. Invest.</i> 104: 515-523
50 NM 005191	CD80	Takahashi <i>et al.</i> (1998)
NM 006889 AF142665	CD86 CD1a	<i>J Immunol.</i> 161: 2629-2635
55 NM005214	CTLA4	

TABLA 8

65	(138700) NKR-P1A (NP 009330) STAT1
----	---------------------------------------

ES 2 331 407 T3

TABLA 9

	<u>Número de acceso</u>	<u>Gen</u>
5	J00219	IFNG
	M11717	HSPA1A
	L05424	CD44
	X51757	HSPA6
10	M28130	IL-8
	X00695	IL-2
	J00219	IFNG
15	U61836	FLJ20746
	M59040	CD44
	L19779	H2 AFO
	D63789	SCYC2
20	M59830	HSPA1B
	A1885852	H2 AFO
	M91196	ICS BPI
25	AF008915	EVI5
	X81851	IL-4
	S82692	IL-2
30	X71661	LMAN1
	M27533	CD80
	M17017	IL-8
	X51757	HSPA6
35	U10550	GEM
	AF079221	BNIP 3L
	AF060568	ZNF145
40	W29115	FLJ20500
	D11466	PIGA
	X51956	ENO2
	U84371	AK2
45	M73255	VCAM1
	U48730	STAT 5B
	U31120	IL-13
50	U43185	STAT5A
	M30257	VCAM1
	L31584	CCR7
55	AJ00 1684	KLRC2

60

65

ES 2 331 407 T3

TABLA 10

	<u>Número de acceso</u>	<u>Gen</u>
5	U16720	IL-10
	AF002668	DEGS
	M97936	STAT1
10	X76220	MAL
	X04430	IL-6
	M69199	G0S2
15	D86324	CMAH
	AF015287	SPUVE
	U64 197	SCYA20
20	L05072	IRF1
	U83171	SCYA22

TABLA 11

	<u>Número de acceso</u>	<u>Gen</u>
30	J04988	HSPCB
	U12595	TRAP1
	L77886	PTPRK
35	X17620	NME1
	U24152	PAK1
	U24152	PAK1
40	U43899	STAM
	X03541	NTRK1
	X55504	NOL1
	X73066	NME1
45	M64231	SRM
	AC005546	UNK _A C005546
	Y10805	HRM1L2
50	ABO11083	ADCY3
	W28612	UNK _W 28612
	D13626	KIAA0001
	XI 5306	NEFH
55	Y13834	ZMPSTE24
	M57506	SCYA1
	D31887	KIAA0062
60	M16660	HSP CB
	U59151	DKC1
	Y12065	NOP56
65	AF059531	PRMT3
	AL050205	UNK _A L050205

ES 2 331 407 T3

	U78525	EIF3S9
	D50914	KIAA0124
5	AJ001014	RAMP1
	AB024301	RUVBL2
	AF026166	CCT2
10	AB028965	KIAA1042
	D78130	SQLE
	Y18643	METTL1
	U86602	P40
15	U13737	CASP3
	U23 143	SHMT2
	AF008442	RPA40
20	S85655	PHB
	D21262	P130
	L33 842	IMPDH2
	X03541	NTRK1
25	D25218	KIAA0112
	AB002359	PFAS
	AL049422	UNK _A L049422
30	U94317	RPP40
	X95263	PWP2H
	M69039	C1QBP
35	U16799	ATP1B1
	M22382	HSPD1
	X06323	MRPL3
	X80200	TRAF4
40	026488	KIAA00007
	X54199	GART
	AL050159	DKFZP586A0522
45	082348	ATIC
	AL038662	NME1
	A1912041	HSPE1
	M38690	CD9
50	A3007398	DKFZP564M1 82
	D87432	SLC7A6
	W52003	K1AA1237
55	U31382	GNG4
	M31516	DAF
	ABO 14547	MTMR4
60	D43 950	CCT5
	AL080119	PAI-RBP1
	AL080119	PAI-RBP1
	U75686	PABPC4
65	AF081280	NPM3

ES 2 331 407 T3

	X74801	CCT3
	U04953	IARS
	AB018293	KIAA0750
5	U51166	TDG
	AA926957	FLJ10534
	AL050022	DKPZP564D1 16
10	AA772359	PSMA2
	A1553745	HSPC111
	A1816034	NOLA2
15	L36720	BYSL
	U28042	DDX10
	Y10256	MAP3K14
20		
25		
30		
35		
40		
45		
50		
55		
60		
65		

ES 2 331 407 T3

TABLA 12

	<u>Número de acceso</u>	<u>Gen</u>	<u>Niveles de CD4 relativos a los niveles NKT</u>
5			
	S78187	CDC25B	aumento
10	X76771	FEN1	aumento
	L25876	CDKN3	aumento
	U73379	UBCH10	aumento
15	M3 1303	LAP 18	aumento
	X51688	CCNA2	aumento
	X61079	UNK X61	aumento
	X63 547	USP6	aumento
20	S68134	CREM	disminución
	AF015254	STK12	aumento
	Y14768	UNK Y14	aumento
25	M25753	CCNBI	aumento
	D14678	KNSL2	aumento
	AB17430	KNSL4	aumento
30	AF067656	ZWINT	aumento
	D90144	SCYA3	disminución
	U63743	KNSL6	aumento
	AB000115	GS3686	disminución
35	U30872	CENPF	aumento
	D14657	KIAA0101	aumento
	D79987	KIAA0165	aumento
40	W60864	TYROBP	disminución
	U05340	CDC20	aumento
	M63 928	TNFRSF7	aumento
	X13444	CD8BI	aumento
45	X16665	HOXB2	disminución
	X14850	H2AFX	aumento
	X70944	SFPQ	aumento
50	U37426	KNSLI	aumento
	M30894	TRG	disminución
	X65550	MKI67	aumento
55	M86699	TTK	aumento
	L47276	UNK L47	aumento
	XI 6302	IGFBP2	aumento
	AA100961	PECAMI	aumento
60	M94345	CAPG	aumento
	V00599	TUBB	aumento
	AE000659	UNK AE	disminución
65	AJ00 1684	KLRC2	disminución

ES 2 331 407 T3

	<u>Número de acceso</u>	<u>Gen</u>	<u>Niveles de CD4 relativos a los niveles NKT</u>
5	M2 1121	SCYA5	disminución
	J04765	SPP 1	aumento
	AE000659	UNK AE	disminución
10	D63789	SCYC2	disminución
	M57703	PMCH	aumento
	A00 1685	KLRC3	disminución
15	AJ001684	KLRC2	disminución
	M57888	GZMB	disminución
	M283 93	PRF 1	disminución
	AF052124	SPP 1	aumento
20	U11276	KLRB 1	disminución
	AF004230	LILRB 1	aumento
	D90144	SCYA3	disminución
25	S75168	MATK	disminución
	S69 115	NKG7	disminución
	ABO 13924	TSC4O3	aumento
	Z22576	CD69	disminución
30	AFO3 1137	D6S49E	disminución
	W60864	TYROBP	disminución
	ALO3 1983	GABBR1	aumento
35	AF01361 1	CTSW	disminución
	AI651806	CRIM1	disminución
	AB 023209	K1AA0992	aumento
40	M30894	TRGα	disminución
	J04430	ACP5	aumento
	L08 177	EBI2	aumento

ES 2 331 407 T3

TABLA 13

	D86096	UNK D86	aumento
5	AF053306	BUB1B	aumento
	AF032862	HMMR	aumento
	A13 75913	TOP2A	aumento
10	AL080146	CCNB2	aumento
	M25753	CCNB1	aumento
	D11139	TIMP1	aumento
	J04088	TOP2A	aumento
15	X65550	MKI67	aumento
	X68 742	ITGA1	aumento
	D26361	KIAA0042	aumento
20	AJ001685	KLRC3	aumento
	M57888	GZMB	disminución
	X68742	ITGA1	aumento
	D88357	CDC2	aumento
25	AC004142	FLJ11129	aumento
	AB023209	K1AA0992	aumento
	U16954	AFIQ	aumento
30	AB024704	C200RF1	aumento
	X05360	CDC2	aumento
	D00596	TYMS	aumento
35	M12824	CD8A	aumento
	U74612	FOXM1	aumento
	AA926959	CKS1	aumento
	AA22530	FLJ20500	disminución
40	U14518	CENPA	aumento

45

50

55

60

65

REIVINDICACIONES

1. Un procedimiento para evaluar si un individuo sufre de la diabetes tipo I, en el que el procedimiento comprende
5 comparar:
 - a) el nivel de expresión de un marcador en una muestra de células NKT de un individuo, en el que el marcador se
selecciona del grupo que consiste en LT- β , Semaphorin III, Frizzled, Flightless I hom., pp32 nuclear, PTP D1, RNP
B5, RNP homólogo, TRAP-3, MLN62, IL-9R, CD37, EBI-1, LPAP, Calponin, GRK-6, TXK, BTF-3 hom., EIF-4B, y
10 b) el nivel normal de expresión del marcador en una muestra de control de células NKT, en el que una diferencia
significativa entre el nivel de expresión del marcador en la muestra de células NKT del individuo y el nivel normal
indica que el individuo sufre de la diabetes tipo I.
- 15 2. El procedimiento de la reivindicación 1, en el que el marcador corresponde a un polinucleótido transcrito o una
parte del este, en el que el polinucleótido comprende el marcador.
3. El procedimiento de la reivindicación 1, en el que las células NKT se recogen del tejido pancreático.
- 20 4. El procedimiento de la reivindicación 1, en el que las células NKT se recogen del tejido sanguíneo.
5. El procedimiento de la reivindicación 1, en el que el nivel de expresión del marcador en la muestra de células
NKT difiere del nivel normal de expresión del marcador en un individuo que no sufre de la diabetes tipo I, por un
factor de al menos 2.
- 25 6. El procedimiento de la reivindicación 1, en el que el nivel de expresión del marcador en la muestra difiere del
nivel normal de expresión del marcador en un individuo que no sufre de la diabetes tipo I, por un factor de al menos 5.
7. El procedimiento de la reivindicación 1, en el que el nivel de expresión del marcador en la muestra se evalúa
30 detectando en la muestra la presencia de una proteína correspondiente al marcador.
8. El procedimiento de la reivindicación 7, en el que la presencia de la proteína se detecta utilizando un reactivo
que específicamente se une con la proteína.
- 35 9. El procedimiento de la reivindicación 8, en el que el reactivo se selecciona del grupo de reactivos que consiste
en un anticuerpo, un derivado de un anticuerpo y un fragmento de un anticuerpo.
10. El procedimiento de la reivindicación 1, en el que el nivel de expresión del marcador en la muestra se evalúa
40 detectando en la muestra la presencia de un polinucleótido transcrito o una parte del este, en el que el polinucleótido
transcrito comprende el marcador.
11. El procedimiento de la reivindicación 10, en el que el polinucleótido transcrito es un ARNm.
12. El procedimiento de la reivindicación 10, en el que el polinucleótido transcrito es un ADNc.
- 45 13. El procedimiento de la realización 10, en el que la etapa de detección comprende además amplificar el polinu-
cleótido transcrito.
14. El procedimiento de la reivindicación 1, en el que el nivel de expresión del marcador en la muestra se evalúa
50 detectando en la muestra la presencia del polinucleótido transcrito que se hibrida con el marcador o se hibrida
con una parte de un polinucleótido, en el que el polinucleótido comprende el marcador, en condiciones rigurosas de
hibridación.
15. El procedimiento de la reivindicación 1 que comprende comparar:
55 a) el nivel de expresión en la muestra de células NKT de cada uno de los marcadores independientes seleccionados
del grupo constituido por
LT- β , Semaphorin III, Frizzled, Flightless I hom., pp32 nuclear, PTP D1, RNP B5, RNP homólogo, TRAP-3,
60 MLN62, IL-9R, CD37, EBI-1, LPAP, Calponin, GRK-6, TXK, BTF-3 hom., EIF-4B, y
b) el nivel normal de expresión de cada uno de los marcadores en muestras del mismo tipo obtenidas de individuos
de control que no sufre de la diabetes tipo I, en el que el nivel de expresión de uno o más marcadores está significati-
vamente alterado, con respecto a los niveles normales de expresión correspondientes de los marcadores, lo que indica
65 que el individuo sufre de la diabetes tipo I.
16. El procedimiento de la reivindicación 15, en el que el conjunto de marcadores comprende dos o más marcado-
res.

17. El procedimiento de la reivindicación 15, en el que el conjunto de marcadores comprende al menos cinco de los marcadores.

18. Un procedimiento para controlar el avance de la diabetes tipo I en un individuo que comprende:

a) detectar en un muestra de células NKT de un individuo en un primer momento, la expresión de un marcador, en el que el marcador se selecciona del grupo que consiste en

LT- β , Semaphorin III, Frizzled, Flightless I hom., pp32 nuclear, PTP D1, RNP B5, RNP homólogo, TRAP-3, MLN62, IL-9R, CD37, EBI-1, LPAP, Calponin, GRK-6, TXK, BTF-3 hom., EIF-4B y combinaciones de estos;

b) repetir la etapa a) en un momento posterior; y c) comparar el nivel de expresión detectado en los pasos a) y b), y a partir de ello controlar la progresión de la diabetes tipo I en el individuo.

19. El procedimiento de la reivindicación 18, en el que el marcador corresponde a un polinucleótido transcrito o una parte del este, en el que el polinucleótido comprende el marcador.

20. El procedimiento de la reivindicación 18, en el que las células NKT se recogen del tejido pancreático.

21. El procedimiento de la reivindicación 18, en el que las células NKT se recogen del tejido sanguíneo.

22. Un procedimiento para evaluar la eficacia de una terapia para inhibir la diabetes tipo I, en el que el procedimiento comprende comparar:

a) la expresión de un marcador en una primera muestra de células NKT obtenidas de un individuo antes de proporcionar al menos parte de la terapia al individuo, en el que el marcador se selecciona de un grupo constituido por

LT- β , Semaphorin III, Frizzled, Flightless I hom., pp32 nuclear, PTP D1, RNP B5, RNP homólogo, TRAP-3, MLN62, IL-9R, CD37, EBI-1, LPAP, Calponin, GRK-6, TXK, BTF-3 hom., EIF-4B, y

b) la expresión del marcador en una segunda muestra de células NKT obtenidas del individuo después de haberle proporcionado una parte de la terapia, en el que una disminución significativa en el nivel de expresión de un marcador en la segunda muestra, con respecto con la primera muestra, es un indicador de que la terapia es eficaz para inhibir la diabetes tipo I en un individuo.

23. Un procedimiento para evaluar la eficacia de un compuesto de prueba para inhibir la diabetes tipo I en un individuo, en el que el procedimiento comprende comparar:

a) la expresión de un marcador en una primera muestra de células NKT obtenidas de un individuo y expuestas o mantenidas en presencia del compuesto de prueba, en el que el marcador se selecciona de un grupo constituido por

LT- β , Semaphorin III, Frizzled, Flightless I hom., pp32 nuclear, PTP D1, RNP B5, RNP homólogo, TRAP-3, MLN62, IL-9R, CD37, EBI-1, LPAP, Calponin, GRK-6, TXK, BTF-3 hom., EIF-4B, y

b) la expresión del marcador en una segunda muestra de células NKT obtenidas del individuo, en el que la segunda muestra no está expuesta al compuesto de prueba, en el que una disminución significativa en el nivel de expresión de un marcador en la primera muestra, con respecto con la segunda muestra, es un indicador de que el compuesto de prueba es eficaz para inhibir la diabetes tipo I en un individuo.

24. El procedimiento de la reivindicación 23, en el que la primera y segunda muestras son partes de una única muestra de células NKT obtenidas del individuo.

25. El procedimiento de la reivindicación 23, en el que la primera y segunda muestras son partes de muestras combinadas de células NKT obtenidas del individuo.

26. Un procedimiento para seleccionar una composición para inhibir la diabetes tipo I en un individuo en el que el procedimiento comprende:

a) mantener separadamente alícuotas de una muestra que comprende células NKT obtenidas de un individuo en presencia de un conjunto de composiciones de prueba;

b) comparar la expresión de un marcador en cada una de las alícuotas, en el que el marcador se selecciona del grupo constituido por

LT- β , Semaphorin III, Frizzled, Flightless I hom., pp32 nuclear, PTP D1, RNP B5, RNP homólogo, TRAP-3, MLN62, IL-9R, CD37, EBI-1, LPAP, Calponin, GRK-6, TXK, BTF-3 hom., EIF-4B, y

ES 2 331 407 T3

c) seleccionar una de las composiciones de prueba que induce un nivel más bajo de expresión del marcador en la alícuota que contiene la composición de prueba, con respecto a otras composiciones de prueba.

27. Un procedimiento para evaluar el potencial de un compuesto de prueba para desencadenar la diabetes tipo I en un individuo, en el que el procedimiento comprende:

a) mantener alícuotas separadas de células NKT en presencia y ausencia del compuesto de prueba; y

b) comparar la expresión de un marcador en cada una de las alícuotas, en el que el marcador se selecciona del grupo constituido por

LT- β , Semaphorin III, Frizzled, Flightless I hom., pp32 nuclear, PTP D1, RNP B5, RNP homólogo, TRAP-3, MLN62, IL-9R, CD37, EBI-1, LPAP, Calponin, GRK-6, TXK, BTF-3 hom., EIF-4B, y

En el que un aumento significativo del nivel de expresión del marcador en la alícuota que se mantiene en presencia del compuesto de prueba, con respecto con la alícuota que se mantiene en ausencia del compuesto de prueba es una indicación de que el compuesto de prueba posee el potencial para desencadenar la diabetes tipo I en una célula.

28. El uso de un kit para realizar los procedimientos de conformidad con cualquiera de las reivindicaciones 1, 23, 26 y 27, en el que el kit comprende al menos un reactivo para evaluar la expresión de un marcador en una muestra de células NKT, en el que el marcador se selecciona del grupo que consiste en

LT- β , Semaphorin III, Frizzled, Flightless I hom., pp32 nuclear, PTP D1, RNP B5, RNP homólogo, TRAP-3, MLN62, IL-9R, CD37, EBI-1, LPAP, Calponin, GRK-6, TXK, BTF-3 hom., EIF-4B.

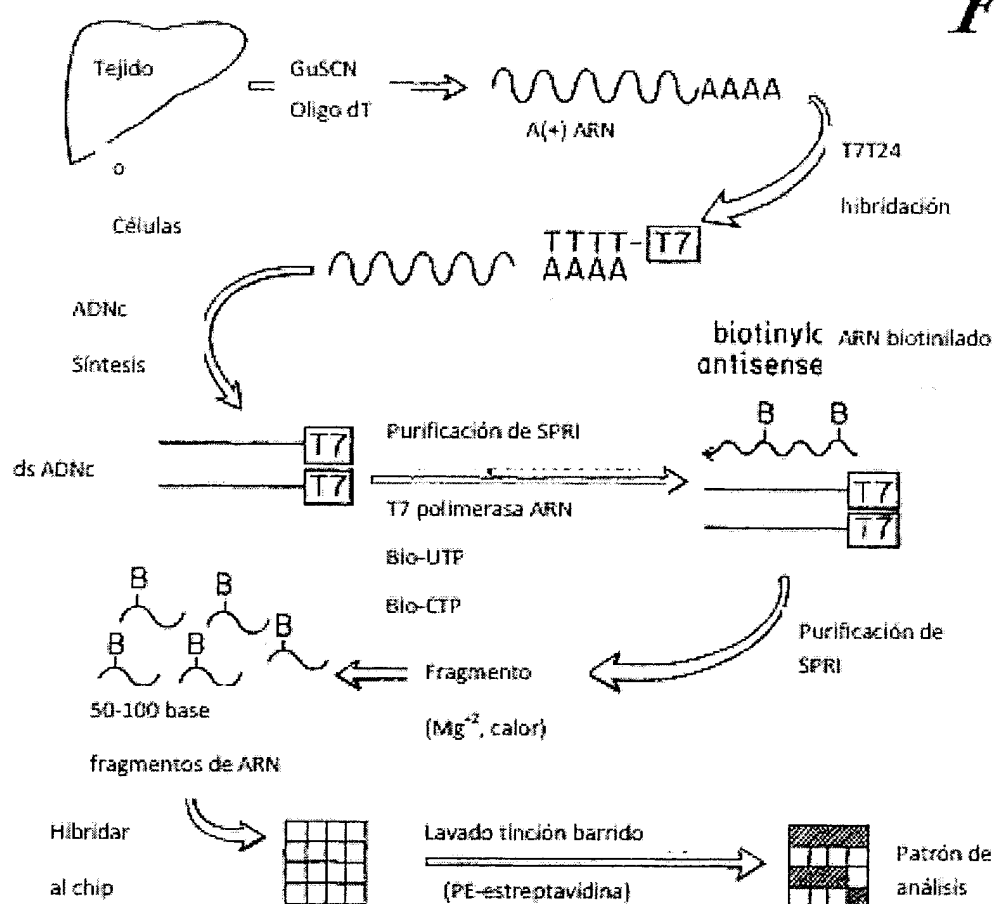
29. El uso según la reivindicación 28, en el que el reactivo es una sonda de ácido nucleico que se une específicamente con un polinucleótido transcripto correspondiente a un marcador seleccionado del grupo que consiste en

LT- β , Semaphorin III, Frizzled, Flightless I hom., pp32 nuclear, PTP D1, RNP B5, RNP homólogo, TRAP-3, MLN62, IL-9R, CD37, EBI-1, LPAP, Calponin, GRK-6, TXK, BTF-3 hom., EIF-4B.

30. El uso según la reivindicación 28, en el que el reactivo es un anticuerpo que se une específicamente con una proteína correspondiente a un marcador seleccionado del grupo que consiste en

LT- β , Semaphorin III, Frizzled, Flightless I hom., pp32 nuclear, PTP D1, RNP B5, RNP homólogo, TRAP-3, MLN62, IL-9R, CD37, EBI-1, LPAP, Calponin, GRK-6, TXK, BTF-3 hom., EIF-4B.

FIG. 1



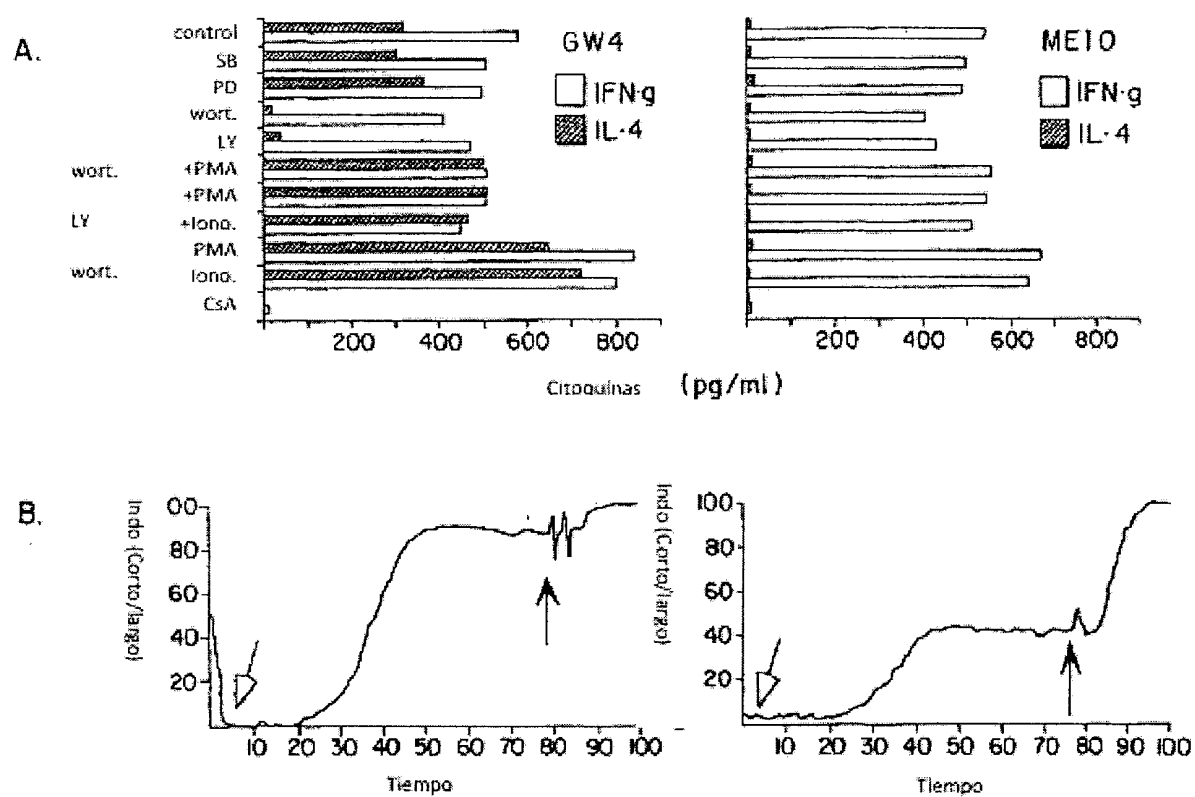


FIG. 2

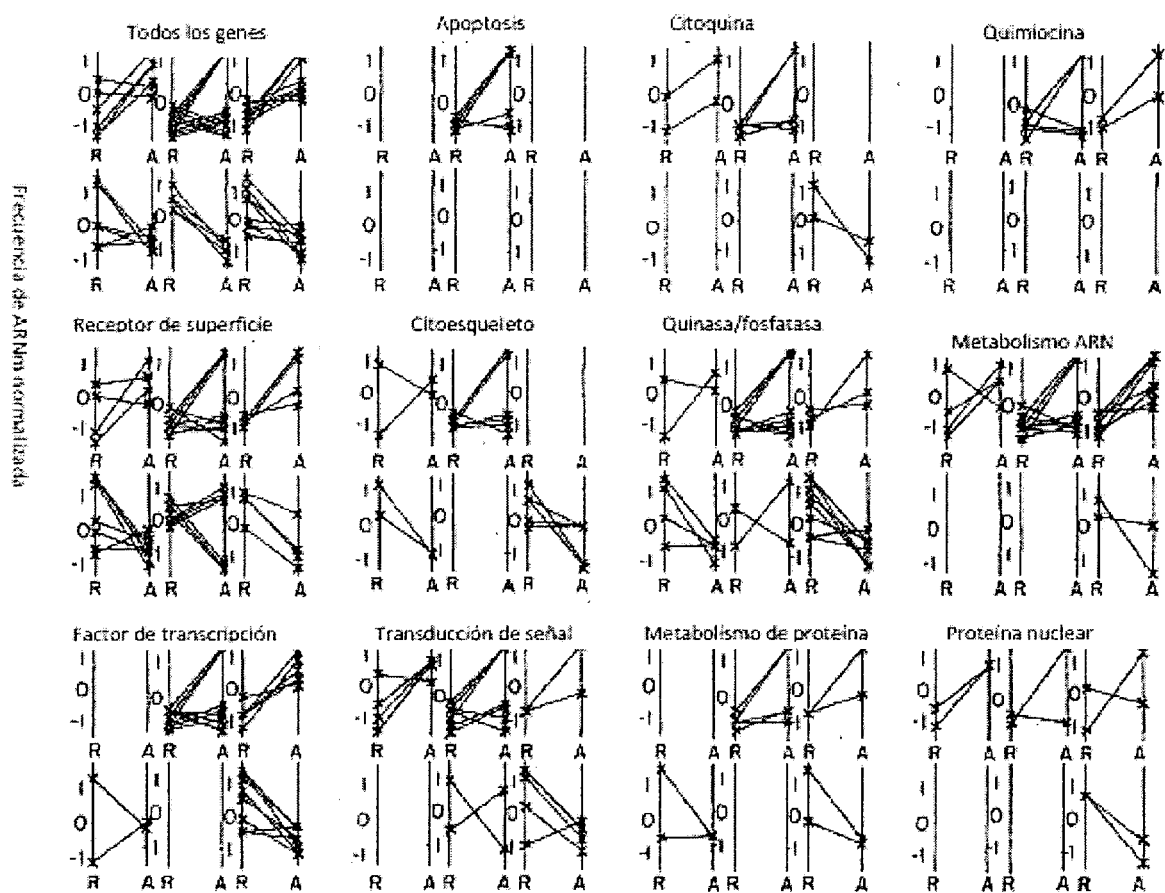


FIG. 3