

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 509 567**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.08.2010 E 10748206 (9)**

97 Fecha y número de publicación de la concesión europea: **30.07.2014 EP 2470673**

54 Título: **Identificación de variación genética en tejidos afectados**

30 Prioridad:

28.08.2009 US 237908 P

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

17.10.2014

73 Titular/es:

**CELLULAR DYNAMICS INTERNATIONAL, INC.
(100.0%)
University Research Park 525 Science Drive,
Suite 200
Madison, WI 53711, US**

72 Inventor/es:

**THOMSON, JAMES y
SEAY, NICK**

74 Agente/Representante:

ARIAS SANZ, Juan

ES 2 509 567 T3

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

DESCRIPCIÓN

Identificación de variación genética en tejidos afectados

ANTECEDENTES DE LA INVENCION

5 La presente solicitud reivindica el beneficio de prioridad a la solicitud provisional de EE.UU. n.º de serie 61/237.908 presentada el 28 de agosto de 2009.

1. Campo de la invención

La presente invención se refiere generalmente al campo del diagnóstico molecular. Más particularmente, se refiere al uso de células madre pluripotentes inducidas (células iPS) en la determinación y caracterización de variación genética.

2. Descripción de la técnica relacionada

10 La patología producida por defectos en genes humanos es normalmente altamente específica de tejido. En enfermedades hereditarias, esto sugiere que las funciones espaciotemporales específicas de los genes implicados están alteradas debido a mutaciones de la línea germinal. Aunque se ha mostrado que los genes de enfermedad generalmente tienden a expresarse en un número limitado de tejidos, todavía no está claro en muchos casos cómo los patrones de expresión específicos de tejido o estructura genética de genes de enfermedad se correlacionan con sus manifestaciones patológicas o anormales y sigue siendo difícil determinar o caracterizar los genes anormales o la estructura genética anormal en
15 ciertos tipos de tejidos o células.

En un aspecto más amplio, el entendimiento de la variación genética es un objetivo clave de la genética humana, que engloba susceptibilidad a enfermedad, respuesta variable a fármacos y por último tratamiento y salud pública.

20 El cuerpo humano está compuesto por más de 200 tipos de células presentes en una diversidad de tipos de tejido. Para identificar defectos genéticos desconocidos o determinar variación genética, necesitan aislarse y caracterizarse tipos de células específicas. Sin embargo, algunos tipos de células pueden ser difíciles de obtener de un sujeto humano *in vivo*, tales como las células del epitelio pigmentario de la retina (RPE) altamente especializadas. Kennan A y col. (Human Molecular Genetics, vol. 11, n.º: 5, 2002) desvelan el análisis de la expresión génica diferencial en tejido de la retina. Así, sigue existiendo una necesidad de desarrollar procedimientos más convenientes para proporcionar tipos de células específicas para analizar la variación genética específica de tejido.
25

RESUMEN DE LA INVENCION

La presente invención vence una deficiencia importante en la materia proporcionando procedimientos de diagnóstico de variación genética específica de tejido relevante para el uso de células iPS, proporcionando, específicamente, tipos de células relevantes para enfermedad específicas que pueden ser exigentes de aislar u obtener. En una primera realización,
30 se proporciona un procedimiento para determinar la presencia de una variación genética específica de tejido en un sujeto de prueba con respecto a una estructura genética de control seleccionada, que comprende: a) obtener material genético de una célula diferenciada de un tejido seleccionado, habiéndose preparado la célula diferenciando células madre pluripotentes inducidas (iPS) obtenidas reprogramando una célula somática del sujeto de prueba; y b) probar el material genético de la célula diferenciada para comparar una o más estructuras genéticas del material genético con una o más
35 estructuras genéticas de control para determinar la presencia de una variación genética tal en células del tejido del sujeto de prueba. En ciertas realizaciones, el tejido seleccionado puede ser retina, tejido neural o tejido de músculo cardíaco.

En ciertos aspectos, el sujeto puede tener o se sospecha que tiene una anomalía genética o enfermedad genética. La enfermedad genética puede ser una enfermedad específica de tejido con defectos genéticos difíciles de caracterizar, tales como una enfermedad de la retina. Para la caracterización de la enfermedad de la retina, tipos de células de la retina específicas, tales como una célula neural de la retina o una célula de epitelio pigmentario de la retina (RPE), podrían proporcionarse por ciertos aspectos de los presentes procedimientos, específicamente diferenciando células iPS reprogramadas de una célula somática derivada del sujeto de prueba.
40

El material genético de la célula diferenciada, que puede comprender nucleótidos tales como ARN o ADN, podría ponerse a prueba por diversas formas conocidas en la técnica, tales como secuenciación de nucleótidos, por ejemplo, secuenciación de ARN o ADN, o análisis de micromatrices. La estructura genética puede ser secuencia de nucleótidos primaria, estructura secundaria, estructura epigenética, estructura cromosómica y similares. En otros aspectos de la invención, la estructura genética de la célula diferenciada puede representarse por un perfil de expresión, que puede reflejar las consecuencias funcionales de las regiones de control de la transcripción dentro de secuencias de ADN, que no son evidentes de la simple secuenciación de nucleótidos sola.
45

50 Ciertos aspectos de los presentes procedimientos pueden usarse para identificar variación genética, que puede ser variación de una o más secuencias codificantes, o variación de una o más secuencias no codificantes, tales como

5 elementos reguladores de genes. La variación genética determinada por aquellas etapas anteriores podría ser cualesquiera tipos de mutaciones, tales como delección, inserción, desapareamientos, translocalización, duplicación, inversión, pérdida de heterocigosidad; o polimorfismo, tal como polimorfismo de un solo nucleótido (SNP); o representarse por expresión diferencial. La variación genética también puede ser una diferencia en la abundancia de transcritos resultante de diferencias en el funcionamiento de elementos de control de la expresión en los tejidos del paciente.

10 Los presentes procedimientos pueden ser particularmente útiles para determinar la presencia de variación genética tal como polimorfismos en regiones no codificantes, como intrones o regiones reguladoras que incluyen, pero no se limitan a, promotores, potenciadores, silenciadores, elementos sensibles o secuencias reguladoras sobre ARN, tales como 5'- o 3'-UTR (regiones sin traducir), o cualquier elemento regulador sin caracterizar. La variación genética también puede comprender variación en ARN no codificante, tal como microARN, o variación epigenética, tal como variación en la modificación de la estructura genética, por ejemplo, metilación.

15 Para determinar la presencia de una variación genética tal puede usarse una estructura del gen de control seleccionada, tal como una sin la variación genética seleccionada. En un cierto aspecto, la estructura del gen de control puede estar comprendida en material genético de un tejido normal, preferencialmente un tejido normal tal puede ser de un sujeto de control tal como un hermano u otro miembro de la familia del sujeto de prueba. El sujeto de control puede no tener la variación genética, especialmente en el tipo de tejido del tejido seleccionado del sujeto de prueba. En otro aspecto, la estructura del gen de control puede estar comprendida en material genético, en el que el material genético es de una célula normal seleccionada preparada diferenciando una célula iPS obtenida reprogramando una célula somática normal de un sujeto de control, tal como un sujeto que no tiene la variación genética. Preferentemente, la célula normal seleccionada o el tejido normal pueden ser del mismo tipo de tejido que el tejido seleccionado del sujeto de prueba para la comparación de estructura genética específica de tejido. En ciertos aspectos, el sujeto de control puede ser un miembro de la familia emparentado con el sujeto de prueba.

25 Las realizaciones tratadas en el contexto de procedimientos y/o composiciones de la invención pueden emplearse con respecto a cualquier otro procedimiento o composición descrito en el presente documento. Así, una realización referente a un procedimiento o composición también puede aplicarse a otros procedimientos y composiciones de la invención.

Como se usa en el presente documento, los términos “codificar” o “que codifica” con referencia a un ácido nucleico se usan para hacer la invención fácilmente entendible por el experto; sin embargo, estos términos pueden usarse indistintamente con “comprender” o “que comprende”, respectivamente.

30 Como se usa en el presente documento, la especificación “un” o “uno” puede significar uno o más. Como se usa en el presente documento en la(s) reivindicación (reivindicaciones), cuando se usa conjuntamente con la palabra “que comprende”, las palabras “un” o “una” pueden significar uno o más de uno.

35 El uso del término “o” en las reivindicaciones se usa para significar “y/o”, a menos que se indique explícitamente para referirse a alternativas solo o las alternativas sean mutuamente excluyentes, aunque la divulgación soporte una definición que se refiere a solo alternativas e “y/o”. Como se usa en el presente documento, “otro” puede significar al menos un segundo o más.

En toda la presente solicitud, el término “aproximadamente” se usa para indicar que un valor incluye la variación inherente de error para el dispositivo, el procedimiento que se emplea para determinar el valor, o la variación que existe entre los sujetos del estudio.

40 Otros objetivos, características y ventajas de la presente invención llegarán a ser evidentes a partir de la siguiente descripción detallada. Debe entenderse, sin embargo, que la descripción detallada y los ejemplos específicos, aunque indican realizaciones preferidas de la invención, se facilitan a modo de ilustración solo.

La invención se define por las reivindicaciones.

DESCRIPCIÓN DE REALIZACIONES ILUSTRATIVAS

I. Introducción

45 La presente divulgación, según ciertas realizaciones, se refiere generalmente a procedimientos para determinar variación genética específica de tejido probando material genético de tipos de células específicas que pueden derivarse de células madre pluripotentes inducidas. El material genético puede comprender cualquier material nuclear (cromosómico, ADN, ARN, etc.) y citoplásmico (por ejemplo, mitocondrial, proteínas) que podría desempeñar una función fundamental en determinar la naturaleza de sustancias celulares, estructuras celulares y/o efectos celulares. La variación genética que va a determinarse puede incluir, por ejemplo, defectos genéticos o polimorfismo genético.

Los defectos genéticos, que pueden producir una enfermedad o trastorno que es heredado genéticamente, comprenden

anomalías en las estructura genéticas, tales como un gen ausente o defectuoso o una aberración cromosómica. Aquellos defectos genéticos determinados por ciertos aspectos de la presente invención podrían aplicarse para el diagnóstico de enfermedad genética o incluso terapia.

5 El polimorfismo genético en tipos de tejidos o células específicos también puede determinarse por ciertos aspectos de los presentes procedimientos. Por ejemplo, el polimorfismo genético se refiere a la aparición de más de un alelo o marcador genético en el mismo locus, produciéndose el alelo o marcador de menor frecuencia más frecuentemente que lo que puede explicarse por mutación solo. Los polimorfismos genéticos pueden incluir polimorfismos de un solo nucleótido (SNP); polimorfismos en la longitud de fragmentos de restricción (RFLP); polimorfismos en el número de copias; polimorfismos en la transcripción o expresión, tales como variación en los niveles de ARN, coordinación de tiempo o distribución tisular; polimorfismos en regiones no codificantes, que incluyen intrones y elementos reguladores, tales como promotores, silenciadores, potenciadores, elementos sensibles, factores de transcripción, regiones sin traducir, o ARN no codificante; polimorfismos de repetición de trinucleótidos, por ejemplo, el número, dinámica y/o distribución de trinucleótidos, tales como repeticiones de CAG; polimorfismos en estado epigenético, por ejemplo, estados de metilación; sensibilidad diferencial a cambios del metabolismo, factores de crecimiento (por ejemplo, insulina), agentes químicos o farmacéuticos, procedimientos o tratamientos terapéuticos, o similares; susceptibilidad a una enfermedad o trastorno; variación en la respuesta a fármacos, metabolismo de fármacos o toxicidad de fármacos, etc. Por ejemplo, la determinación de los polimorfismos genéticos puede ser útil para el diagnóstico o pronóstico de enfermedad para su correlación con la aparición de defectos genéticos o con el estado del portador.

20 En ciertos aspectos, la prueba de material genético puede comprender determinar el polimorfismo de la expresión de ARN o secuencia de ARN de los tipos de células específicas. La caracterización de ARN específica de célula puede ser ventajosa en dar información que incluso completa la secuenciación del genoma de ADN o completa lo que la secuenciación proteómica no pueda dar. Por ejemplo, como las consecuencias funcionales de las variaciones en las regiones de control de ADN todavía no se entienden muy bien, el examinar el ARN en el tejido diana o tipos de células (análisis completo del transcriptoma) puede ser capaz de identificar variaciones funcionales en los elementos reguladores o regiones que podrían no identificarse teniendo una secuencia de genoma entera secuenciada y puede proporcionar información no disponible por la secuenciación del genoma completo. Particularmente, el nivel de expresión génica (por ejemplo, transcripción) es una medida cuantitativa que está directamente ligada a la variación genética y puede normalmente representar polimorfismos en elementos reguladores. Si un polimorfismo o mutación en los elementos reguladores está relacionado con una enfermedad específica de tejido o anomalía, puede representarse por cambios en los niveles de expresión de ARN, localización de ARN, coordinación de tiempo de ARN o niveles relativos de variantes de ARN cortadas y empalmadas en el tejido relevante. La secuenciación de ARN de los tipos de células específicas puede también revelar la variación específica mucho más rápido que la secuenciación de ADN, ya que los autores de la invención examinan específicamente el material genético relacionado con ese tipo de células sin requerir secuencias de ADN del genoma completo. Este análisis se haría mejor en comparación por pares con hermanos sin afectar, por tener la mayor posibilidad de identificar las diferencias de transcritos asociadas a la condición de la enfermedad.

40 La presente invención podría ser un prelude para tratar la enfermedad específica de tejido por trasplante de células diferenciadas derivadas de células iPS genéticamente corregidas, o un diagnóstico de variación genética específica de tejido o de célula sola. Adicionalmente, ciertos aspectos de la presente invención reducen espectacularmente la complejidad de la identificación de la causa genética o el marcador genético para trastornos genéticos sin caracterizar, ya que los autores de la invención pueden empezar con los polimorfismos representados por diferencia entre los perfiles de expresión de sujetos de control y de prueba en vez de con el genoma o transcriptoma completo.

II. Variación genética y enfermedad genética

45 Ciertos aspectos proporcionan procedimientos para determinar la presencia de variación genética, que puede proporcionar diagnóstico de enfermedades genéticas o de anomalía genética. La variación genética puede desempeñar una función clave en moldear la diversidad fenotípica entre individuos, en la que los mecanismos subyacentes pueden incluir polimorfismos o anomalías que alteran secuencias codificantes de proteínas o cambios en las secuencias reguladoras que afectan la función o expresión de un gen o de redes de genes relacionadas. La variación genética puede resultar de la presencia de diferentes genotipos en la población y asociarse a, crear susceptibilidad a o ser la causa subyacente para enfermedad o anomalía genética.

50 Las enfermedades genéticas pueden producirse por mutaciones de la línea germinal que, a pesar de la presencia tisular amplia, frecuentemente conducen a patología específica de tejido, o a anomalía en la operación de elementos reguladores, que también puede tener un fenotipo específico de tejido.

A. Variación genética

55 La variación genética puede influir en la expresión génica en tipos de tejidos o de células específicos. En ciertos aspectos de los presentes procedimientos, asociando la expresión específica de tejido o célula con la estructura genética, la

variación genética podría identificarse o determinarse. Por ejemplo, la variación genética puede incluir polimorfismos tales como polimorfismo de un solo nucleótido (SNP), que son variaciones de la secuencia de ADN que se producen cuando un único nucleótido - A, T, C o G - en el genoma (u otra secuencia compartida) difiere entre miembros de una especie (o entre cromosomas apareados en un individuo). El análisis de secuencias diploides también ha mostrado que la variación no SNP da cuenta de mucha más variación genética humana que la diversidad de un solo nucleótido. Esta variación no SNP incluye variación del número de copias y puede resultar de deleciones, inversiones, inserciones y duplicaciones. Se estima que aproximadamente el 0,4 % de los genomas de personas sin emparentar normalmente se diferencian con respecto al número de copias.

Algunos de los polimorfismos pueden estar en regiones no codificantes del genoma mientras que algunos pueden residir en las regiones codificantes, por ejemplo, polimorfismos en regiones reguladoras o elementos reguladores, o variantes exónicas que alteran la estabilidad o corte y empalme de los transcritos.

La variación genética puede incluir polimorfismos reguladores genéticos que actúan en *cis* cambiando la secuencia codificante de proteínas o en el nivel de ARN: afectando la transcripción (activación o inhibición mediante sitios reguladores o estructura de elementos reguladores), el procesamiento de ARNm, el corte y empalme de pre-ARNm, los potenciadores del corte y empalme exónico (ESE), el salto de exón, la estabilidad de ARNm, el tráfico de ARNm o los ARN reguladores. La variación epigenética puede imitar polimorfismos reguladores a cierto grado (Johnson y col., 2005).

Los polimorfismos en regiones reguladoras o elementos reguladores pueden ser uno de los efectos primarios clave que contribuyen a la variación fenotípica en seres humanos y así son importantes para el análisis molecular de la variación fenotípica humana, alguna de la cual puede estar ligada a trastornos genéticos.

Por ejemplo, hay cada vez más pruebas de que los polimorfismos reguladores desempeñan una función importante en la determinación de susceptibilidad individual a rasgos de enfermedad compleja (Knight 2005). Tales polimorfismos reguladores incluyen, pero no se limitan a, variación en el número variable de repeticiones en tándem (es decir, VNTR) de INS, que codifica insulina, en diabetes tipo 1; polimorfismo de CTLA4, que codifica antígeno de linfocitos T citotóxicos, en enfermedad autoinmunitaria; polimorfismo de la proteína de unión a Duffy en malaria; polimorfismo de CCR5, que codifica receptor 5 de quimiocinas, en infección por el VIH-1. La variación genética puede también operar a otros niveles de control de la expresión génica y la modulación del corte y empalme en PTPRC, que codifica el receptor tipo C de la proteína tirosina fosfatasa y de la eficiencia traduccional en F12, que codifica el factor XII.

B. Trastornos genéticos

Una enfermedad genética (es decir, trastorno genético) es una enfermedad producida por anomalías en genes o cromosomas, o en elementos reguladores de genes. Las anomalías pueden oscilar de una pequeña mutación en un único gen a la adición o resta de un cromosoma entero o conjunto de cromosomas. Algunas enfermedades, tales como el cáncer, se deben en parte a trastornos genéticos y también se producen en parte por factores medioambientales. Algunos tipos de trastornos de los genes recesivos confieren una ventaja en el estado heterocigótico en ciertos entornos. Una célula haploide tiene solo un conjunto de cromosomas. Una célula diploide tiene dos conjuntos de cromosomas. En ser humano, las células somáticas son diploides y los gametos son haploides.

La variedad de enfermedades genéticas es tan amplia en la naturaleza que los individuos afectados pueden encontrarse en prácticamente todas las prácticas médicas. Sencillamente, el código genético influye en la salud desde el momento de la concepción hasta la muerte. Basándose en un estudio de población de trastornos genéticos aparentes a la edad de 25 años (Baird y col., 1988), aproximadamente el 0,4 % de la población tiene un trastorno de un único gen (mendeliano), el 0,2 % tienen una anomalía cromosómica y el 4,6 % tienen una condición multifactorial. Otro 0,1 % tienen una anomalía genética obvia de herencia desconocida y el 0,3 % tienen problemas congénitos que no son de naturaleza genética. Un trastorno de un único gen es el resultado de un único gen mutado. Se estima que hay más de 4000 enfermedades humanas producidas por defectos de un único gen. Los trastornos de un único gen pueden pasarse a generaciones posteriores de varias formas. Sin embargo, el sellado genómico y la disomía uniparental pueden afectar los patrones de herencia. Las divisiones entre tipos recesivos y dominantes no son "estrictas" aunque sí las divisiones entre tipos autosómicos y ligados al cromosoma X (ya que los últimos tipos se distinguen puramente basándose en la localización cromosómica del gen). Por ejemplo, la acondroplasia se considera normalmente un trastorno dominante, pero los niños con dos genes para acondroplasia tienen un grave trastorno esquelético del que los acondroplásicos podrían ser considerados portadores. La anemia de células falciformes también se considera una condición recesiva, pero los portadores heterocigóticos tienen elevada inmunidad a la malaria en la primera infancia, que podría describirse como una afección dominante relacionada.

Solo una copia mutada del gen será necesaria para que una persona esté afectada por un trastorno autosómico dominante. Cada persona afectada normalmente tiene un padre afectado. Hay un 50 % de probabilidad de que un niño herede el gen mutado. Las afecciones que son autosómicas dominantes frecuentemente tienen baja penetrancia, los que significa que aunque solo se necesita una copia mutada, una proporción relativamente pequeña de aquellos que heredan

esa mutación pueden continuar desarrollando la enfermedad. Ejemplos de este tipo de trastorno son enfermedad de Huntington, neurofibromatosis 1, síndrome de Marfan, cáncer colorrectal hereditario no polipósico y exostosis múltiple hereditaria, que es un trastorno dominante autosómico altamente penetrante. Los defectos del nacimiento también se llaman anomalías congénitas.

5 Dos copias del gen deben mutarse para que una persona esté afectada por un trastorno recesivo autosómico. Una persona afectada normalmente tiene padres sin afectar que portan cada uno una única copia del gen mutado (y se denominan portadores). Dos personas sin afectar que portan cada una una copia del gen mutado tienen un 25 % de probabilidad con cada embarazo de tener un niño afectado por el trastorno. Ejemplos de este tipo de trastorno son fibrosis quística, enfermedad de células falciformes (también enfermedad parcial de células falciformes), enfermedad de Tay-Sachs, enfermedad de Niemann-Pick, atrofia muscular espinal y cerumen seco (conocido de otro modo como cerumen "tipo arroz").

10 Los trastornos dominantes ligados al cromosoma X se producen por mutaciones en genes sobre el cromosoma X. Solo algunos trastornos tienen este patrón de herencia, siendo un excelente ejemplo el raquitismo hipofosfatémico ligado al cromosoma X. Hombres y mujeres están ambos afectados en estos trastornos, estando los hombres normalmente más gravemente afectados que las mujeres. Algunas afecciones dominantes ligadas al cromosoma X tales como síndrome de Rett, incontinencia pigmentaria tipo 2 y el síndrome de Aicardi son normalmente letales en hombres tanto en el útero como poco después del nacimiento y por tanto, se observan predominantemente en mujeres. Excepciones a este hallazgo son casos extremadamente raros en los que niños con síndrome de Klinefelter (47, XXY) también heredan una afección dominante ligada al cromosoma X y presentan síntomas más similares a los de una mujer en términos de gravedad de la enfermedad. La probabilidad de transmitir un trastorno dominante ligado al cromosoma X se diferencia entre hombres y mujeres. Los hijos de un hombre con un trastorno dominante ligado al cromosoma X estarán todos sin afectar (ya que reciben el cromosoma Y de su padre) y sus hijas heredarán todas la afección. Una mujer con un trastorno dominante ligado al cromosoma X tiene un 50 % de probabilidad de tener un feto afectado con cada embarazo, aunque debe observarse que en casos tales como incontinencia pigmentaria solo la descendencia de las mujeres es generalmente viable. Además, aunque estas afecciones no alteran en sí la fertilidad, los individuos con síndrome de Rett o síndrome de Aicardi raramente se reproducen.

25 Los trastornos recesivos ligados al cromosoma X también se producen por mutaciones en genes sobre el cromosoma X. Los hombres están más frecuentemente afectados que las mujeres y la probabilidad de transmitir el trastorno se diferencia entre hombres y mujeres. Los hijos de un hombre con un trastorno recesivo ligado al cromosoma X no se afectarán y sus hijas portarán una copia del gen mutado. Una mujer que es una portadora de un trastorno recesivo ligado al cromosoma X (XR_{Xr}) tiene un 50 % de probabilidad de tener hijos que estén afectados y un 50 % de probabilidad de tener hijas que porten una copia del gen mutado y por tanto sean portadoras. Ejemplos de este tipo de trastorno son hemofilia A, distrofia muscular de Duchenne, ceguera a los colores rojo-verde, distrofia muscular y alopecia androgenética.

30 Los trastornos ligados al cromosoma Y se producen por mutaciones sobre el cromosoma Y. Debido a que los hombres heredan un cromosoma Y de sus padres, cada hijo de un padre afectado estará afectado. Debido a que las mujeres heredan un cromosoma X de sus padres, la descendencia femenina de padres afectados nunca está afectada. Como el cromosoma Y es relativamente pequeño y contiene muy pocos genes, hay relativamente pocos trastornos ligados al cromosoma Y. Frecuentemente, los síntomas incluyen infertilidad, que puede sortearse con la ayuda de algunos tratamientos de fertilidad. Ejemplos son infertilidad masculina e hipertricosis del pabellón auricular.

35 Los trastornos genéticos también pueden ser complejos, multifactoriales o poligénicos, esto significa que es probable que estén asociados a los efectos de múltiples genes en combinación con el estilo de vida y los factores medioambientales. Los trastornos multifactoriales incluyen enfermedad cardíaca y diabetes. Aunque los trastornos complejos se agrupan frecuentemente en familias, no tienen un patrón definido de herencia. Esto dificulta determinar el riesgo de que una persona herede o transmita estos trastornos. Los trastornos complejos también son difíciles de estudiar y tratar debido a que los factores específicos que producen la mayoría de estos trastornos todavía no se han identificado.

C. Consideración de trastornos genéticos o anomalías

Las mismas consideraciones que entran en añadir un posible trastorno genético a un diagnóstico diferencial de un neonato o niño se aplican a adultos. ¿Es ese trastorno genético o adquirido? ¿Cómo de significativo es el componente genético frente al componente medioambiental? ¿Cuál es el patrón de herencia? ¿Hay posibilidades de que otros miembros de la familia del paciente se afecten actualmente o en el futuro? ¿Cómo puede confirmarse el diagnóstico? Hay algunas pistas que sugerirán un diagnóstico genético. Los antecedentes familiares podrían indicar un trastorno genético. Numerosos trastornos de aparición en la adultez son dominantes en la naturaleza. Es decir, una mutación en solo uno de un par de genes es necesaria para producir un problema genético (esto contrasta con la herencia recesiva en la que ambos genes de un par deben tener una mutación). Es muy probable que los trastornos genéticos dominantes produzcan antecedentes familiares, mientras que los trastornos recesivos no. Desafortunadamente, algunas afecciones dominantes

5 tienen una alta tasa de mutaciones nuevas, de manera que su paciente podría ser perfectamente el primer caso en la familia. Alternativamente, antecedentes familiares positivos pueden no haberse apreciado previamente, quizás debido al pequeño tamaño de la familia o a la dispersión de los miembros de la familia. Los miembros de la familia que portaban una mutación del gen pueden haber fallecido antes de que se desarrollaran signos y síntomas, o un diagnóstico genético puede no haberse contemplado nunca previamente a pesar de la existencia de parientes afectados.

10 El análisis de la genealogía pueden complicarse por los fenómenos de penetrancia incompleta y expresividad variable (Harper, 1998). La penetrancia incompleta significa que algunos individuos pueden llevar una mutación genética pero no expresar ningún signo ni síntoma. Expresividad variable significa que, aún cuando todos los miembros afectados de la familia tengan la misma mutación, pueden no tener manifestaciones idénticas. La edad de aparición y gravedad de la afección puede variar considerablemente.

Cuando una enfermedad se produce a una edad mucho más joven que a la que normalmente se esperaría, ello sugiere una posible predisposición genética. Por ejemplo, el cáncer de mama en una persona de 25 años de edad es muy inusual (Langston y col., 1996). Emparejado con una historia de cáncer de mama y de ovario en parientes próximos, esto es prácticamente diagnóstico de una mutación en un gen *BRCA* (Haber, 1999).

15 Los médicos deben también considerar una causa genética en tres circunstancias adicionales: casos con participación de múltiples sistemas (por ejemplo, la sordera y nefritis asociadas a síndrome de Alport (Flinter, 1997)), con una presentación multifocal (por ejemplo, la multiplicidad de pólipos durante todo el diagnóstico de colon de poliposis adenomatosa familiar (Midgley y Kerr, 1999)) o con una combinación inusual de eventos (por ejemplo, osteoporosis de aparición temprana y sordera parcial conductiva pueden ser indicativos de osteogénesis imperfecta (Byers y Steiner, 1992)).

D. Trastornos oculares genéticos

En particular, los presentes procedimientos pueden usarse para determinar variación genética o identificar defectos genéticos novedosos en tipos de células de la retina específicas diferenciadas de células iPS, que pueden derivarse de un paciente que tiene o que se sospecha que tiene un trastorno ocular genético, tal como retinitis pigmentaria.

25 La retinitis pigmentaria (RP) es un grupo de trastornos oculares genéticos. En la progresión de síntomas para RP, la ceguera nocturna precede generalmente años o incluso décadas a la visión de túnel. Muchas personas con RP no se convierten legalmente en ciegos hasta sus 40 o 50 años y retienen algo de vista toda su vida. Otros se vuelven completamente ciegos de RP, en algunos casos ya desde la infancia. La progresión de RP es diferente en cada caso.

30 La RP es un tipo de distrofia retiniana progresiva, un grupo de trastornos heredados en los que las anomalías de los fotorreceptores (bastones y conos) o del epitelio pigmentario retiniano (RPE) de la retina conducen a pérdida visual progresiva. Los individuos afectados experimentan primero adaptación defectuosa a la oscuridad o nictalopia (ceguera nocturna), seguida de reducción del campo visual periférico (conocida como visión de túnel) y algunas veces, pérdida de la visión central posteriormente en el transcurso de la enfermedad.

35 El diagnóstico de la retinitis pigmentaria se basa en la documentación de pérdida progresiva en la función fotorreceptora por electroretinografía (ERG) y prueba del campo visual. El modo de herencia de RP se determina por los antecedentes familiares. Se conocen al menos 35 genes o loci diferentes que producen "RP no sindrómica" (RP que no es el resultado de otra enfermedad o parte de un síndrome más amplio).

40 Hay múltiples genes que, cuando mutan, pueden producir el fenotipo de retinitis pigmentosa. En 1989, se identificó una mutación del gen para rodopsina, un pigmento que desempeña una parte esencial en la cascada de transducción visual que permite la visión en condiciones de baja luz. Desde entonces, se han encontrado más de 100 mutaciones en este gen, representando el 15 % de todos los tipos de degeneración retiniana. La mayoría de aquellas mutaciones son mutaciones de sentido erróneo y se heredan principalmente de una manera dominante. El gen de rodopsina codifica una proteína principal de segmentos externos de los fotorreceptores. Estudios muestran que las mutaciones en este gen son responsables de aproximadamente el 25 % de las formas dominantes autosómicas de RP.

45 Se sabe que las mutaciones en cuatro factores de corte y empalme de pre-ARNm producen retinitis pigmentaria dominante autosómica. Estos son PRPF3, PRPF8, PRPF31 y PAP1. Estos factores se expresan ubicuamente y es todavía un rompecabezas por qué los defectos en un factor ubicuo solo producirían enfermedad en la retina.

50 Se ha informado hasta la fecha de hasta 150 mutaciones en el gen de opsina asociado a la RP ya que la mutación Pro23His en el dominio intradiscal de la proteína se informó por primera vez en 1990. Estas mutaciones se encuentran en todo el gen de opsina y se distribuyen a lo largo de los tres dominios de la proteína (los dominios intradiscal, transmembrana y citoplásmico). Una de las principales causas bioquímicas de la RP en el caso de mutaciones de rodopsina es el plegamiento erróneo de proteínas y también se han implicado en RP chaperonas moleculares. Se encontró que la mutación del codón 23 en el gen de rodopsina, en la que la prolina se cambia a histidina, da cuenta de la

mayor fracción de mutaciones de rodopsina en los Estados Unidos. Varios otros estudios han informado de otras mutaciones que también correlacionan con la enfermedad. Estas mutaciones incluyen Thr58Arg, Pro347Leu, Pro347Ser, además de la delección de Ile-255. En el año 2000, se informó de una mutación rara en el codón 23 que causa retinitis pigmentaria dominante autosómica, en la que la prolina cambió a alanina. Sin embargo, este estudio mostró que la distrofia de la retina asociada con esta mutación fue característicamente leve en presentación y evolución. Además, hubo mayor preservación en las amplitudes de electroretinografía que con la mutación Pro23His más prevalente.

III. Diagnósticos moleculares

Aunque se están desarrollando nuevas pruebas de diagnóstico molecular y la variedad de pruebas disponibles está aumentando rápidamente, la gran mayoría de los trastornos genéticos todavía carece de una prueba genética con base genética sin caracterizar. Ciertos aspectos de los presentes procedimientos proporcionan análisis específicos de tejidos o de células de diferencias en la estructura genética que comprenden estructura genómica, que pueden representarse por el perfil de expresión entre células derivadas de un sujeto de prueba y una célula normal, por ejemplo, una célula normal del mismo tipo de célula o de tejido.

Una prueba genética es el análisis de ADN humano, ARN humano, cromosomas humanos, proteínas humanas o ciertos metabolitos humanos con el fin de detectar alteraciones relacionadas con un trastorno hereditario. Esto puede llevarse a cabo examinando directamente el ADN o ARN que constituye un gen (prueba directa), mirando en marcadores coheredados con un gen causante de enfermedad (prueba de vinculación), ensayando ciertos metabolitos (prueba bioquímica), o examinando los cromosomas (prueba citogenética). La prueba genética es frecuentemente la mejor forma para confirmar un diagnóstico en un paciente con signos o síntomas sugerentes de una enfermedad genética. La técnica elegida depende tanto de la cuestión clínica como del valor predictivo de las pruebas disponibles.

A. Secuenciación de ARN

El ARN es menos estable en la célula y también más propenso al ataque por nucleasas experimentalmente. Como el ARN se genera por transcripción de ADN, la información ya está presente en el ADN de la célula. Sin embargo, es algunas veces deseable secuenciar moléculas de ARN. En particular, en eucariotas las moléculas de ARN no son necesariamente co-lineales con su molde de ADN, ya que se escinden intrones. Para secuenciar ARN, el procedimiento usual es transcribir primero de forma inversa la muestra para generar fragmentos de ADN. Esto puede luego secuenciarse como secuenciación de ADN.

La tecnología de micromatrices y los procedimientos de amplificación múltiple variados han mostrado que el ARN es válido como diana para diagnósticos moleculares rutinarios y para la realización futura de pruebas de diagnóstico. La detección de defectos genéticos en el perfil de ARN en enfermedades o trastornos específicos de tejido puede dar más exactitud y objetividad y proporcionar más información al procedimiento de diagnóstico que los procedimientos de detección basados en ADN. El reto principal usando proteína como una diana para diagnósticos rutinarios ha sido baja sensibilidad, reproducibilidad y especificidad. Sin embargo, ARN y no ADN como una diana para diagnósticos rutinarios puede dar la información de actividad clínica, regulación o procedimientos, además de mayor o igual sensibilidad, reproducibilidad y especificidad en comparación con ADN como diana. Durante más de 10 años se han desarrollado nuevos procedimientos de aislamiento, purificación y estabilización de ARNm para diagnósticos rutinarios haciendo el ARN mucho más adecuado como marcador para el desarrollo de nuevos procedimientos de diagnóstico e incluso fármacos.

Los cuatro aspectos que son relevantes para el diagnóstico molecular o función génica son niveles de expresión de ARN, especificidad tisular de la expresión de ARN, coordinación de tiempo de la expresión de ARN y la secuencia primaria de ARN. Solo examinando el ARN en el tejido diana o en las células diana con ciertos aspectos de la presente invención, será mucho más fácil identificar un defecto específico que tenga una secuencia genómica entera, ya que todavía no se entienden las consecuencias funcionales de los cambios en las regiones de control de ADN, pero tales diferencias afectarían los niveles de ARN, la coordinación de tiempo de ARN o la distribución tisular de ARN.

En ciertos aspectos, la secuenciación de ARN puede usarse para identificar el defecto genético en los tipos de células específicos derivados de células iPS. ARN-Sec, también llamada "secuenciación al azar del transcriptoma completo" ("WTSS") y "una herramienta revolucionaria para transcriptómica" se refiere al uso de tecnologías de secuenciación de alto rendimiento para secuenciar ADNc con el fin de conseguir información sobre un contenido de ARN de una muestra, una técnica que se está convirtiendo rápidamente en inapreciable en el estudio de enfermedades genéticas (Denoeud y col., 2008). Gracias a la profunda cobertura y resolución de nivel base proporcionadas por los instrumentos de secuenciación de la siguiente generación, el ARN-Sec proporciona formas eficaces para medir datos de transcriptomas experimentalmente y para conseguir información sobre cómo se expresan alelos diferentes de un gen, cómo detectar mutaciones post-transcripcionales o cómo identificar fusiones de genes.

B. Micromatriz

Una micromatriz de ADN es una tecnología múltiplex usada en biología molecular y en medicina. Consiste en una serie de matrices de miles de puntos microscópicos de oligonucleótidos de ADN, llamados características, cada una de las cuales contiene picomoles de una secuencia de ADN específica. Esto puede ser una sección corta de un gen u otro elemento de ADN que se usa como sondas para hibridarse con una muestra de ADNc o ARNc (llamada diana) en condiciones de alta rigurosidad. La hibridación de sonda-diana se detecta normalmente y se cuantifica por detección de dianas marcadas con fluoróforo, plata o quimioluminiscencia para determinar la abundancia relativa de secuencias de ácidos nucleicos en la diana.

En micromatrices convencionales, las sondas están unidas a una superficie sólida por un enlace covalente a una matriz química (mediante epoxi-silano, amino-silano, lisina, poliacrilamida u otros). La superficie sólida puede ser vidrio o un chip de silicón, en cuyo caso se conoce comúnmente como chip de gen o coloquialmente chip Affy cuando se usa un chip Affymetrix. Otras plataformas de micromatriz, tales como Illumina, usan perlas microscópicas, en lugar del gran soporte sólido. Las matrices de ADN son diferentes de otros tipos de micromatriz solo porque tanto miden ADN como usan ADN como parte de su sistema de detección.

Las micromatrices de ADN pueden usarse para medir cambios en los niveles de expresión, para detectar polimorfismos de un solo nucleótido (SNP), en genotipado o en resecuenciación de genomas mutantes para la determinación de variación genética en ciertos aspectos. Más específicamente, las micromatrices de ADN pueden usarse para detectar ADN (como en hibridación genómica comparativa), o para detectar ARN (lo más comúnmente como ADNc después de la transcripción inversa) que puede o puede no traducirse en proteínas. El procedimiento de medición de la expresión génica mediante ADNc se llama análisis de expresión o perfilado de la expresión. En otros aspectos, el defecto genético en los tipos de células específicos derivados de células iPS puede identificarse por perfilado de la expresión génica, la medición de la actividad (la expresión) de miles de genes de una vez, para crear un cuadro global de función celular. Muchos experimentos de este tipo miden un genoma entero simultáneamente, es decir, cada gen presente en una célula particular. Ejemplos de perfilado de la expresión génica incluyen, pero no se limitan a, micromatriz de ADN, SAGE o qPCR.

La tecnología de micromatrices de ADN mide la actividad relativa de genes diana previamente identificados. Esto puede ser una sección corta de un gen u otro elemento de ADN que se usa como sondas para hibridar una muestra de ADNc o ARNc (llamada diana) en condiciones de alta rigurosidad. Las técnicas basadas en secuencia, como análisis en serie de la expresión génica (SAGE, SuperSAGE) también pueden usarse para el perfilado de la expresión génica. SuperSAGE es especialmente preciso y puede medir cualquier gen activo, no solo un conjunto predeterminado. La reacción en cadena de la polimerasa en tiempo real, también llamada reacción en cadena de la polimerasa en tiempo real cuantitativa (Q-PCR/qPCR) o reacción en cadena de la polimerasa cinética, es una técnica de laboratorio basada en la reacción en cadena de la polimerasa, que se usa para amplificar y cuantificar simultáneamente una molécula de ADNc elegida como diana. Permite tanto la detección como la cuantificación (como número de copias absoluto o cantidad relativa cuando se normaliza a la entrada de ADNc o genes de normalización adicionales) de una secuencia específica en una muestra de ADNc.

C. Análisis de la vinculación

Con referencia a familias con un trastorno genético en el que las mutaciones no pueden localizarse, si el gen se clona o su localización en el genoma se identifica por mapeo genético, puede ofrecerse una forma limitada de la prueba genética que se basa en el análisis de la vinculación. Esto implica el uso de secuencias polimórficas de ADN o ARN que están estrechamente ligadas a o que se encuentran dentro del gen para rastrear la secuencia mutante a través de la familia. La técnica puede usarse si dos o más generaciones de miembros de la familia afectada están disponibles para el estudio y pueden necesitarse amplios análisis de la familia para marcadores informativos. Este enfoque también puede limitarse por posibles errores debidos a recombinación genética o heterogeneidad genética (si más de un sitio del gen es posiblemente responsable del mismo trastorno en diferentes familias). A pesar de estas limitaciones, la realización de prueba basada en vinculación ha sido muy útil para asesorar familias con enfermedades genéticas tales como distrofia muscular de Duchenne, hemofilia, atrofia muscular espinal y muchos otros trastornos.

D. Procedimientos de identificación de polimorfismos

La medición de variantes de secuencia, principalmente polimorfismos de un solo nucleótido (SNP), proporciona las unidades fundamentales para vincular la secuencia genética con rasgos. La mayoría de los genes albergan variaciones de múltiples secuencias (por ejemplo, SNP, repeticiones, INDEL) que muestran un amplio intervalo de frecuencias y desequilibrio de la vinculación entre ellos. La mayoría de los polimorfismos son no funcionales y pueden servir como marcadores para alelos funcionales. Además de usar polimorfismos individuales, frecuentemente también se hacen asociaciones con el uso de haplotipos, bloques de polimorfismos ligados, que pueden delimitar regiones de secuencia cis significativas de rasgos. Ahora están apareciendo en línea procedimientos de genotipado de SNP de alto rendimiento, tales como SNPlex, que pueden cribar miles de SNP en muchas muestras (Wenz, 2004). Tales procedimientos se han usado para establecer mapas de haplotipos sobre una base de genoma completo, que incluye genes que participan en el

metabolismo de los fármacos, a densidad de marcador significativa (Kamatani y col., 2004).

En ciertos aspectos de la presente invención, la expresión de ARNm medida por micromatrices puede combinarse con análisis de vinculaciones del genoma completo, tomando el nivel de expresión de cada gen como el fenotipo medido restringido a tipo de tejido o célula específico. La heredabilidad de fenotipos de la expresión génica puede explorarse mediante el genotipado familiar y la realización de pruebas del desequilibrio de transmisión en familias nucleares (Spielman & Ewens, 1996) o de pruebas del desequilibrio de la genealogía en genealogías mayores (Martin y col., 2000). Usando tejidos diana (derivados de células iPS) de miembros de la familia, este tipo de análisis puede distinguir entre factores genéticos que actúan en cis y trans y muestra una abundancia de sitios genómicos funcionales y una preponderancia de efectos que actúan en trans, como es de esperar.

Un enfoque alternativo implica el análisis de la expresión específica de alelo en un tejido diana relevante, que también puede prepararse a partir de células iPS; cada alelo experimenta su propia regulación en el mismo entorno celular, sirviendo el otro alelo (para genes autosómicos) de control interno. Como un resultado, el procedimiento controla las condiciones de tejido, factores que actúan en trans y otras influencias medioambientales. Así, los SNP en regiones exónicas y sin traducir de mensajero, pueden servir de marcadores para los niveles de expresión de alelos en individuos heterocigóticos para estos marcadores. Tomando la familia 15 de transportadores de soluto humanos (transportador de H+/péptidos), gen del miembro 2 (hPepT2) como ejemplo, se ha descrito recientemente un procedimiento para la medición específica de alelo de la expresión de ARNm mediante incorporación de la extensión del cebador de sondas que terminan en didesoxi-nucleótido fluorescente después de la amplificación por RT-PCR (Pinsonneault y col., 2004). Diferencias significativas en la abundancia relativa de cada alelo en ARNm de tejidos de riñón demostraron la presencia de factores que actúan en cis funcionales. La reacción de extensión del cebador puede multiplexarse (Bray y col., 2004) de manera que sea posible buscar polimorfismos que actúan en cis funcionales en un gran número de genes (Yan y col., 2002). Pueden lograrse resultados similares mediante procedimientos empleados en otras plataformas (Wojnowski & Brockmoller, 2004) que incluyen el uso de espectroscopía de dispersión por desorción/ionización por láser asistida por matriz (MALDI-TOF; Ding y col., 2004) y metodologías de RT-PCR específicas de alelos (Zhang y col., 2004). Estas técnicas pueden extenderse a ARN nuclear heterogéneo (ARNnh) sin procesar si marcadores exónicos y sin traducir no están disponibles y si el ARNnh es suficientemente abundante en las muestras diana (Hirota y col., 2004).

IV. Células madre

En ciertas realizaciones de la invención, se desvelan procedimientos de uso de células madre pluripotentes inducidas (células iPS) para la determinación de variación genética específica de tejido. Aquellas células iPS pueden prepararse reprogramando células somáticas y podrían ser idénticas a células madre embrionarias en diversos aspectos como se describe más adelante. El entendimiento de las características de células madre embrionarias podría ayudar a seleccionar células madre pluripotentes inducidas. La reprogramación de factores conocidos de estudios de reprogramación de células madre podría usarse para estos procedimientos novedosos. Se contempla adicionalmente que estas células madre pluripotentes inducidas podrían usarse posiblemente para sustituir células madre embrionarias para aplicaciones terapéuticas y de investigación debido a la carga ética de usar las últimas.

A. Células madre

Las células madre son células encontradas en la mayoría, si no en todos, los organismos pluricelulares. Se caracterizan por la capacidad de renovarse a sí mismas mediante división celular mitótica y diferenciarse en una diversa gama de tipos de células especializadas. Los dos amplios tipos de células madre de mamíferos son: células madre embrionarias que se encuentran en blastocistos y células madre adultas que se encuentran en tejidos adultos. En un embrión en desarrollo, las células madre pueden diferenciarse en todos los tejidos embrionarios especializados. En organismos adultos, las células madre y las células progenitoras actúan de sistema de reparación para el cuerpo, reforzando las células especializadas, pero también manteniendo la renovación normal de órganos regenerativos, tales como sangre, piel o tejidos intestinales.

Como las células madre pueden cultivarse y transformarse en células especializadas con características de acuerdo con células de diversos tejidos tales como músculos o nervios mediante cultivo celular, se ha propuesto su uso en terapias médicas. En particular, líneas celulares embrionarias, células madre embrionarias autólogas generadas mediante clonación terapéutica y células madre adultas altamente plásticas de la sangre del cordón umbilical o la médula ósea están solicitadas como candidatos prometedoros. Más recientemente, la reprogramación en células adultas en células madre pluripotentes inducidas tiene enormes posibilidades de sustituir a las células madre embrionarias.

B. Células madre embrionarias

Las líneas de células madre embrionarias (líneas de células ES) son cultivos de células derivadas del tejido epiblasto de la masa celular interna (ICM) de un blastocisto o de embriones en el estado de mórula temprana. Un blastocisto es un embrión en el estado temprano - aproximadamente cuatro a cinco días de edad en seres humanos y que consiste en 50-150 células. Las células ES son pluripotentes y dan lugar durante el desarrollo a todos los derivados de las tres capas germinales primarias: ectodermo, endodermo y mesodermo. En otras palabras, pueden desarrollarse en cada uno de los

más de 200 tipos de células del cuerpo adulto cuando se da estimulación suficiente y necesaria para un tipo de célula específica. No contribuyen a las membranas extra-embrionarias o la placenta.

5 Casi toda la investigación hasta la fecha ha tenido lugar usando células madre embrionarias de ratón (mES) o células madre embrionarias humanas (hES). Ambas tienen las características de células madre esenciales, pero requieren entornos muy diferentes con el fin de mantener un estado sin diferenciar. Las células ES de ratón pueden cultivarse sobre una capa de gelatina y requieren la presencia de factor inhibidor de leucemia (LIF). Las células ES humanas podrían cultivarse sobre una capa de células nodrizas de fibroblastos embrionarios de ratón (MEF) y frecuentemente requieren la presencia de factores de crecimiento de fibroblastos básicos (bFGF o FGF-2). Sin condiciones de cultivo óptimas o manipulación genética (Chambers y col., 2003), las células madre embrionarias se diferenciarán rápidamente.

10 Una célula madre embrionaria humana también puede definirse por la presencia de varios factores de transcripción y proteínas de la superficie celular. Los factores de transcripción Oct4, Nanog y Sox2 forman la red reguladora central que garantiza la supresión de genes que conducen a la diferenciación y el mantenimiento de pluripotencia (Boyer y col., 2005). Los antígenos de superficie de la célula más comúnmente usados para identificar células hES incluyen los glucolípidos SSEA3 y SSEA4 y los antígenos de sulfato de queratano Tra-1-60 y Tra-1-81.

15 Después de veinte años de investigación, no hay tratamientos autorizados o ensayos humanos que usen células madre embrionarias. Las células ES, que son células pluripotentes, requieren señales específicas para la correcta diferenciación – si se inyectan directamente en el cuerpo, las células ES se diferenciarán en muchos tipos de células diferentes, causando un teratoma. La diferenciación de células ES en células útiles mientras que se evita el rechazo de trasplante son solo algunos de los obstáculos a los que todavía se enfrentan los investigadores de las células madre embrionarias.
 20 Muchas naciones tienen actualmente una moratoria sobre tanto la investigación de células ES como la producción de nuevas líneas de células ES. Debido a sus capacidades combinadas de expansión ilimitada y pluripotencia, las células madre embrionarias siguen siendo una fuente teóricamente posible para la medicina regeneradora y la sustitución tisular después de lesión o enfermedad. Sin embargo, una forma de sortear estas cuestiones es inducir el estado pluripotente en células somáticas por reprogramación directa.

25 **C. Células madre pluripotentes inducidas y factores de reprogramación**

Las células madre pluripotentes inducidas, comúnmente abreviadas células iPS o iPSC, son un tipo de células madre pluripotentes artificialmente derivadas de una célula no pluripotente, normalmente una célula somática adulta. Se cree que las células madre pluripotentes inducidas son idénticas a las células madre pluripotentes naturales, tales como
 30 células madre embrionarias en muchos aspectos, tales como en términos de la expresión de ciertos genes de células madre y proteínas, patrones de metilación de cromatina, tiempo de duplicación, formación de cuerpos embrioides, formación de teratomas, formación de quimeras viables y potencia y diferenciabilidad, pero el grado completo de su relación con respecto a células madre pluripotentes naturales todavía está siendo evaluado.

Las células iPS se produjeron primero en 2006 (Takahashi y col., 2006) a partir de células de ratón y en 2007 a partir de células humanas (Takahashi y col., 2007; Yu y col., 2007). Esto se ha citado como un avance importante en la
 35 investigación de las células madre, ya que puede permitir a los investigadores obtener células madre pluripotentes, que son importantes en la investigación y posiblemente tienen usos terapéuticos, sin el controvertido uso de embriones.

La generación de células iPS es crucial en los factores de reprogramación usados para la inducción. Los siguientes factores o combinación de los mismos podrían usarse en los procedimientos desvelados en la presente invención. En ciertos aspectos, los ácidos nucleicos que codifican Sox y Oct (preferentemente Oct3/4) se incluirán en el vector de
 40 reprogramación. Por ejemplo, uno o más vectores de reprogramación pueden comprender casetes de expresión que codifican Sox2, Oct4, Nanog y opcionalmente Lin28, o casetes de expresión que codifican Sox2, Oct4, Klf4 y opcionalmente c-Myc, o casetes de expresión que codifican Sox2, Oct4 y opcionalmente Esrrb, o casetes de expresión que codifican Sox2, Oct4, Nanog, Lin28, Klf4, c-Myc y opcionalmente el antígeno T grande del SV40. Los ácidos nucleicos que codifican estos factores de reprogramación pueden estar comprendidos en el mismo casete de expresión, diferentes
 45 casetes de expresión, el mismo vector de reprogramación o diferentes vectores de reprogramación.

Se han identificado Oct4 y ciertos miembros de la familia de genes Sox (Sox1, Sox2, Sox3 y Sox15) como reguladores de la transcripción cruciales implicados en el procedimiento de inducción cuya ausencia hace imposible la inducción. Sin embargo, se han identificado genes adicionales, que incluyen ciertos miembros de la familia Klf (Klf1, Klf2, Klf4 y Klf5), la familia Myc (c-Myc, L-Myc y N-Myc), Nanog y Lin28 que aumentan la eficiencia de inducción.

50 Oct4 (Pou5f1) es uno de la familia de los factores de transcripción octámeros (“Oct”) y desempeña una función crucial en el mantenimiento de la pluripotencia. La ausencia de Oct4 en células Oct4⁺, tales como blastómeros y células madre embrionarias, conduce a la diferenciación espontánea de trofoblastos y así la presencia de Oct4 da lugar a la pluripotencia y posible diferenciación de células madre embrionarias. Diversos otros genes en la familia “Oct”, que incluyen parientes cercanos de Oct4, Oct1 y Oct6, fallan en provocar la inducción, demostrándose así la exclusividad de Oct-4 para el
 55 procedimiento de inducción.

La familia Sox de genes está asociada con el mantenimiento de la pluripotencia similar a Oct4, aunque está asociada con células madre multipotentes y unipotentes a diferencia de Oct4, que se expresa exclusivamente en células madre pluripotentes. Mientras que Sox2 fue el gen inicial usado para la inducción por Takahashi y col. (2006), Wernig y col. (2007) y Yu y col. (2007), se ha encontrado que otros genes en la familia Sox también funcionan en el procedimiento de inducción. Sox1 da células iPS con una eficiencia similar a Sox2 y los genes Sox3, Sox15 y Sox18 también generan células iPS, aunque con eficiencia disminuida.

En células madre embrionarias, Nanog, junto con Oct4 y Sox2, es necesario en la promoción de pluripotencia. Por tanto, fue sorprendente cuando Takahashi y col. (2006) informaron que Nanog era innecesario para la inducción, aunque Yu y col. (2007) han informado que es posible generar células iPS con Nanog como uno de los factores.

Lin28 es una proteína de unión a ARNm expresada en células madre embrionarias y células de carcinoma embrionarias asociadas a diferenciación y proliferación. Thompson y col. demostraron que es un factor en la generación de iPS, aunque es innecesario.

Klf4 de la familia Klf de genes se identificó inicialmente por Takahashi y col. (2006) y se confirmó por Wernig y col. (2007) como un factor para la generación de células iPS de ratón y se demostró por Takahashi y col. (2007) como un factor para la generación de células iPS humanas. Sin embargo, Yu y col. (2007) informaron de que Klf4 era innecesario para la generación de células iPS humanas y de hecho falló en generar células iPS humanas. Se encontró que Klf2 y Klf4 eran factores que pueden generar células iPS y los genes relacionados Klf1 y Klf5 también, aunque con eficiencia reducida.

La familia Myc de genes son proto-oncogenes que participan en el cáncer. Takahashi y col. (2006) y Wernig y col. (2007) demostraron que c-Myc es un factor implicado en la generación de células iPS de ratón y Takahashi y col. (2007) demostraron que era un factor implicado en la generación de células iPS humanas. Sin embargo, Yu y col. (2007) y Takahashi y col. (2007) informaron de que c-Myc era innecesario para la generación de células iPS humanas. El uso de la familia "Myc" de genes en la inducción de células iPS es preocupante para la eventualidad de células iPS como terapias clínicas, ya que el 25 % de los ratones trasplantados con células iPS inducidas con c-Myc desarrollaron teratomas letales. Se ha identificado que N-Myc y L-Myc se inducen en lugar de c-myc con eficiencia similar. El antígeno mayor del SV40 puede usarse para reducir o evitar la citotoxicidad que puede producirse cuando se expresa c-Myc.

Las proteínas de reprogramación usadas en la presente invención pueden sustituirse con homólogos de proteínas con aproximadamente las mismas funciones de reprogramación. Los ácidos nucleicos que codifican aquellos homólogos también podrían usarse para la reprogramación. Se prefieren sustituciones de aminoácidos conservadoras – es decir, por ejemplo, aspártico-glutámico como aminoácidos ácidos polares; lisina/arginina/histidina como aminoácidos básicos polares; leucina/isoleucina/metionina/valina/alanina/glicina/prolina como aminoácidos no polares o hidrófobos; serina/treonina como aminoácidos polares o hidrófilos sin carga. La sustitución de aminoácidos conservadora también incluye agrupaciones basadas en cadenas laterales. Por ejemplo, un grupo de aminoácidos que tiene cadenas laterales alifáticas es glicina, alanina, valina, leucina e isoleucina; un grupo de aminoácidos que tiene cadenas laterales de hidroxilo alifáticas es serina y treonina; un grupo de aminoácidos que tiene cadenas laterales que contienen amida es asparagina y glutamina; un grupo de aminoácidos que tienen cadenas laterales aromáticas es fenilalanina, tirosina y triptófano; un grupo de aminoácidos que tienen cadenas laterales básicas es lisina, arginina e histidina; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina. Por ejemplo, es razonable esperar que la sustitución de una leucina con una isoleucina o valina, un aspartato con un glutamato, una treonina con una serina, o una sustitución similar de un aminoácido con un aminoácido estructuralmente relacionado no tenga un efecto importante sobre las propiedades del polipéptido resultante. Puede determinarse fácilmente si un cambio de aminoácidos da como resultado un polipéptido funcional ensayando la actividad específica del polipéptido.

V. Expresión de factores de reprogramación

En ciertos aspectos de la presente invención, células iPS se preparan a partir de reprogramar células somáticas usando factores de reprogramación. La célula somática en la presente invención puede ser cualquier célula somática que pueda inducirse para pluripotencia, tal como un fibroblasto, un queratinocito, una célula hematopoyética, una célula mesenquimática, una célula hepática, una célula del estómago o una célula β . En un aspecto preferido, los linfocitos T pueden usarse como fuente de células somáticas para reprogramación (véase la solicitud de EE.UU. n.º: 61/184.546).

Los factores de reprogramación pueden expresarse a partir de casetes de expresión comprendidos en uno o más vectores, tales como un vector integrante o un vector episómico. En otro aspecto, las proteínas de reprogramación podrían introducirse directamente en células somáticas por transducción de proteínas (véase la solicitud de EE.UU. n.º: 61/172.079).

A. Vectores integrantes

Las células iPS pueden derivarse por transfección de ciertos ácidos nucleicos o genes que codifican proteínas de reprogramación en células no pluripotentes, tales como linfocitos T, en la presente invención. La transfección se consigue

normalmente mediante vectores virales de integración en la práctica actual, tales como retrovirus. Los genes transfectados pueden incluir los reguladores de la transcripción maestros Oct4 (Pou5f1) y Sox2, aunque se sugiere que otros genes potencian la eficiencia de inducción. Después de un periodo crítico, pequeños números de células transfectadas pueden empezar a llegar a ser morfológicamente y bioquímicamente similares a células madre pluripotentes y podrían aislarse mediante selección morfológica, tiempo de duplicado o mediante un gen indicador e infección antibiótica.

En noviembre de 2007 se alcanzó un hito creando iPS de fibroblastos humanos adultos a partir de dos estudios de equipos de investigación independientes (Yuy col., 2007; Takahashi y col., 2007). Con el mismo principio usado anteriormente en modelos de ratón, Takahashi y col. (2007) habían transformado satisfactoriamente fibroblastos humanos en células madre pluripotentes usando los mismos cuatro genes fundamentales: Oct4, Sox2, Klf4 y c-Myc con un sistema retroviral pero c-Myc es oncogénico. Yu y col. (2007) usaron Oct4, Sox2, NANOG y un gen LIN28 diferente usando un sistema lentiviral que evita el uso de c-Myc.

Como se ha descrito anteriormente, la inducción de células madre pluripotentes a partir de fibroblastos dérmicos humanos se ha conseguido usando retrovirus o vectores lentivirales para la expresión ectópica de genes de reprogramación. Retrovirus recombinantes tales como el virus de la leucemia murina de Moloney tienen la capacidad de integrarse en el genoma del huésped de un modo estable. Contienen una transcriptasa inversa que permite la integración en el genoma del huésped. Los lentivirus son una subclase de retrovirus. Están ampliamente adaptados como vectores gracias a su capacidad para integrarse en el genoma de células que no se dividen así como en el de células que se dividen. El genoma viral en forma de ARN se transcribe de forma inversa cuando el virus entra en la célula para producir ADN, que luego se inserta en el genoma en una posición al azar por la enzima integrasa viral.

B. Vectores episómicos

Estos procedimientos de reprogramación también pueden hacer uso de vectores extra-cromosómicamente replicantes (es decir, vectores episómicos), que son vectores que pueden replicarse episómicamente para hacer células iPS esencialmente libres de vector exógeno o elementos virales (véase la solicitud de EE.UU. n.º: 61/058.858; Yu y col., 2009). Varios virus de ADN, tales como adenovirus, virus vacuolantes de los simios 40 (SV40) o virus del papiloma bovino (BPV), o plásmidos que contienen ARS (secuencias de replicación autónoma) de levadura en gemación se replican extra-cromosómicamente o episómicamente en células de mamífero. Estos plásmidos episómicos están intrínsecamente libres de todas estas desventajas (Bode y col., 2001) asociadas a los vectores integrantes. Por ejemplo, uno basado en virus herpes linfotróficos incluyendo el virus de Epstein Barr (EBV) como se ha definido anteriormente puede replicarse extra-cromosómicamente y ayudar a administrar los genes de reprogramación a células somáticas.

Por ejemplo, el enfoque basado en plásmidos usado en la invención puede extraer elementos fuertes necesarios para la replicación y el mantenimiento exitosos de un sistema basado en elementos del EBV sin comprometer la adaptabilidad del sistema en un entorno clínico como se describe en detalle más adelante. Los elementos de EBV esenciales son OriP y EBNA-1 o sus variantes o equivalentes funcionales. Una ventaja adicional de este sistema es que estos elementos exógenos se perderán con el tiempo después de ser introducidos en células, conduciendo a células iPS auto-sostenidas esencialmente libres de elementos exógenos.

El uso de vectores extra-cromosómicos basados en plásmidos o en liposomas, por ejemplo, vectores basados en oriP, y/o vectores que codifican un derivado de EBNA-1, permiten que se introduzcan grandes fragmentos de ADN en una célula y se mantengan extracromosómicamente, se repliquen una vez por ciclo celular, se dividan entre las células hijas eficazmente y no provoquen sustancialmente ninguna respuesta inmunitaria. En particular, EBNA-1, la única proteína viral requerida para la replicación del vector de expresión basado en oriP, no provoca una respuesta inmunitaria celular debido a que ha desarrollado un mecanismo eficaz para evitar el procesamiento requerido para la presentación de sus antígenos sobre moléculas de MHC de clase I (Levitskaya y col., 1997). Además, EBNA-1 puede actuar en *trans* para potenciar la expresión del gen clonado, induciendo la expresión de un gen clonado hasta 100 veces en algunas líneas celulares (Langle-Rouault y col., 1998; Evans y col., 1997). Finalmente, la fabricación de tales vectores de expresión basados en oriP es económica.

Otros vectores extra-cromosómicos incluyen otros vectores basados en virus de herpes linfotróficos. Virus de herpes linfotrófico es un herpesvirus que se replica en un linfoblasto (por ejemplo, un linfoblasto B humano) y se convierte en un plásmido durante una parte de su ciclo de vida natural. El virus del herpes simple (VHS) no es un virus del herpes "linfotrófico". Virus del herpes linfotrófico a modo de ejemplo incluyen, pero no se limitan a EBV, virus del herpes del sarcoma de Kaposi (VHSK); virus del herpes saimiri (HS) y virus de la enfermedad de Marek (VEM). También se contemplan otras fuentes de vectores basados en episomas, tales como ARS de levadura, adenovirus, SV40 o BPV.

Para sortear los posibles problemas de la administración de genes virales, dos grupos informaron este año de una colaboración que ha tenido éxito en enfoques basados en transposones para producir pluripotencia en células humanas sin usar vectores virales (Woltjen y col., 2009; Kaji y col., 2009). Se produjeron células iPS estables en tanto fibroblastos

humanos como de ratón usando secuencias de péptidos 2A derivados de virus para crear un vector multicistrónico que incorpora los factores de reprogramación, administrado a la célula por el vector de transposón piggyBac. Entonces se eliminaron los factores de reprogramación ligados a 2A, no requeridos en las líneas de células iPS establecidas. Estas estrategias podrían aplicarse similarmente a reprogramar linfocitos T en ciertos aspectos de la presente invención.

5 **C. Transducción de proteínas**

Una forma posible de evitar introducir modificaciones genéticas exógenas a células diana sería administrar las proteínas de reprogramación directamente en células, más que depender de la transcripción de genes administrados. Estudios anteriores han demostrado que las diversas proteínas pueden administrarse a células *in vitro* e *in vivo* conjugándolas con un péptido corto que media en la transducción de proteínas, tal como tat del VIH y poli-arginina. Un estudio reciente
10 demostró que los fibroblastos murinos pueden reprogramarse completamente en células madre pluripotentes por administración directa de proteínas de reprogramación recombinantes (Zhou y col., 2009). Más detalles de los procedimientos para reprogramar células con transducción de proteínas se han desvelado en la solicitud de EE.UU. n.º: 61/172.079.

En ciertos aspectos de la presente invención, los dominios de transducción de proteína podrían usarse para introducir proteínas de reprogramación directamente en linfocitos T. La transducción de proteínas podría ser un procedimiento para potenciar la administración de proteínas de reprogramación en células. Por ejemplo, una región de la proteína TAT que se deriva de la proteína Tat del VIH puede fusionarse con una proteína diana permitiendo la entrada de la proteína diana en la célula. Las ventajas de usar fusiones de estos dominios de transducción es que la entrada de la proteína es rápida, dependiente de la concentración y parece funcionar con diferentes tipos de células.

En otro aspecto de la presente invención, la secuencia de localización nuclear también puede usarse para facilitar la entrada nuclear de proteínas de reprogramación. Se han descrito señales de localización nuclear (NLS) para diversas proteínas. El mecanismo de transporte de la proteína al núcleo es mediante la unión de una proteína diana que contiene una señal de localización nuclear a la subunidad alfa de carioferina. Esto va seguido del transporte del complejo de proteína diana:carioferina a través del poro nuclear y dentro del núcleo. Sin embargo, las proteínas de reprogramación son frecuentemente factores de transcripción que pueden tener secuencias de localización nuclear endógenas. Por tanto, las secuencias de localización nuclear pueden no ser necesarias.

La introducción directa de proteínas de reprogramación en células somáticas puede usarse en la presente invención, con proteínas de reprogramación operativamente ligadas a un dominio de transducción de proteínas (PTD), tanto creando una proteína de fusión que comprende un dominio tal como reticulando químicamente la proteína de reprogramación y PTD
30 mediante grupos funcionales sobre cada molécula.

Los procedimientos de ácidos nucleicos recombinantes convencionales pueden usarse para expresar una o más proteínas de reprogramación transducibles usadas en el presente documento. En una realización, una secuencia de ácidos nucleicos que codifica la proteína transducible se clona en un vector de expresión de ácido nucleico, por ejemplo, con secuencias señal y de procesamiento apropiadas y secuencias reguladoras para la transcripción y traducción. En otra
35 realización, la proteína puede sintetizarse usando procedimientos sintéticos orgánicos automatizados.

Además, ha habido varios procedimientos que también pueden ayudar en el transporte de proteínas a células, uno o más de los cuales pueden usarse solos o en combinación con los procedimientos que usan los dominios de transducción de proteínas, que incluyen, pero no se limitan a, microinyección, electroporación y el uso de liposomas. La mayoría de estos procedimientos puede necesitar una preparación de proteína purificada. La purificación de proteínas recombinantes se facilita frecuentemente por la incorporación de una marca de afinidad en la construcción de expresión, haciendo la etapa de purificación rápida y eficaz.
40

VI. Selección, cultivo y diferenciación de células IPS

En ciertos aspectos de la invención, después de introducirse uno o más factores de reprogramación en células somáticas de un sujeto de prueba, las células se cultivarán para expansión (opcionalmente seleccionadas por la presencia de
45 elementos de vector como selección positiva o marcador seleccionable para concentrar células transfectadas). Los vectores de reprogramación pueden expresar factores de reprogramación en estas células y replicarse y dividirse junto con la división celular. Alternativamente, las proteínas de reprogramación podrían entrar en estas células y su progenie reforzando el medio que contiene las proteínas de reprogramación. Estos factores de reprogramación reprogramarán el genoma de la célula somática para establecer un estado pluripotente autosuficiente y mientras tanto o después de
50 eliminar la selección positiva de la presencia de vectores, los elementos genéticos exógenos se perderán gradualmente, o no hay necesidad de añadir proteínas de reprogramación.

Estas células madre pluripotentes inducidas podrían seleccionarse de células de progenie basándose en características de células madre embrionarias debido a que se espera que sean sustancialmente idénticas a células madre embrionarias pluripotentes. Una etapa de selección negativa adicional podría también emplearse para acelerar o ayudar en la selección

de células iPS esencialmente libres de elementos genéticos exógenos probando la ausencia de ADN de vector de reprogramación o usando marcadores de selección, tales como indicadores.

Después de seleccionar o aislar células iPS, la diferenciación en tipos de células específicas puede inducirse a partir de células iPS para la determinación de la variación genética específica de tejidos. Los tipos de células específicas pueden estar comprendidos en un tejido seleccionado, tal como retina.

A. Selección de células IPS

Las iPSC exitosamente generadas de estudios anteriores fueron extraordinariamente similares a células madre pluripotentes naturalmente aisladas (tales como células madre embrionarias de ratón y humanas, mESC y hESC, respectivamente) en los siguientes aspectos, confirmando así la identidad, autenticidad y pluripotencia de iPSC con células madre pluripotentes naturalmente aisladas. Así, las células madre pluripotentes inducidas generadas a partir de los procedimientos desvelados en esta invención podrían seleccionarse basándose en una o más de las siguientes características de células madre embrionarias.

i. Propiedades biológicas celulares

Morfología: las iPSC son morfológicamente similares a las ESC. Cada célula puede tener forma redonda, nucléolos duales o nucléolo grande y poco citoplasma. Las colonias de iPSC podrían también ser similares a las de ESC. Las iPSC humanas forman colonias de bordes afilados, planas, estrechamente empaquetadas similares a hESC y las iPSC de ratón forman colonias similares a mESC, colonias menos planas y más agregadas que las de hESC.

Propiedades de crecimiento: el tiempo de duplicación y la actividad mitótica son pilares de ESC, ya que las células madre deben auto-renovarse como parte de su definición. Las iPSC podrían ser mitóticamente activas, activamente auto-renovables, proliferantes y que se dividen a una tasa igual a ESC.

Marcadores de células madre: las iPSC pueden expresar marcadores antigénicos de la superficie celular expresados sobre ESC. Las iPSC humanas expresaron los marcadores específicos para hESC, que incluyen, pero no se limitan a, SSEA-3, SSEA-4, TRA-1-60, TRA-1-81, TRA-2-49/6E y Nanog. Las iPSC de ratón expresaron SSEA-1, pero no SSEA-3 ni SSEA-4, similarmente a mESC.

Genes de células madre: las iPSC pueden expresar genes expresados en ESC sin diferenciar, que incluyen Oct4, Sox2, Nanog, GDF3, REX1, FGF4, ESG1, DPPA2, DPPA4 y hTERT.

Actividad telomerasa: las telomerasas son necesarias para sostener la división celular sin limitar por el límite de Hayflick de ~50 divisiones celulares. Las hESC expresan actividad telomerasa alta para sostener la auto-renovación y proliferación y las iPSC también demuestran actividad alta de telomerasa y expresan hTERT (telomerasa transcriptasa inversa humana), un componente necesario en el complejo proteico telomerasa.

Pluripotencia: las iPSC serán capaces de diferenciación de un modo similar a ESC en tejidos completamente diferenciados.

Diferenciación neural: las iPSC podrían diferenciarse en neuronas, que expresan β III-tubulina, tirosina hidroxilasa, AADC, DAT, ChAT, LMX1B y MAP2. La presencia de enzimas asociadas a catecolamina puede indicar que las iPSC, como hESC, pueden ser diferenciables en neuronas dopaminérgicas. Los genes asociados a células madre se regularán por disminución después de la diferenciación.

Diferenciación cardíaca: las iPSC podrían diferenciarse en cardiomiocitos que empiezan espontáneamente a latir. Los cardiomiocitos expresan cTnT, MEF2C, MIL2A, MYHC β y NKX2.5. Los genes asociados a células madre se regularán por disminución después de la diferenciación.

Formación de teratoma: las iPSC inyectadas en ratones inmunodeficientes pueden formar espontáneamente teratomas después de un cierto tiempo, tal como nueve semanas. Los teratomas son tumores de múltiples linajes que contienen tejido derivado de las tres capas germinales endodermo, mesodermo y ectodermo; esto es a diferencia de otros tumores, que normalmente son de solo un tipo de célula. La formación de teratomas es una prueba de referencia para pluripotencia.

Cuerpo embriode: las hESC en cultivo forman espontáneamente estructuras similares a embrión similares a bola llamadas "cuerpos embriodes" que consisten en un núcleo de hESC mitóticamente activas y que se diferencian y una periferia de células completamente diferenciadas de las tres capas germinales. Las iPSC también pueden formar cuerpos embriodes y tener células diferenciadas periféricas.

Inyección de blastocistos: las hESC residen naturalmente dentro de la masa celular interna (embrioblasto) de blastocistos y en el embrioblasto, se diferencian en el embrión mientras que la envuelta del blastocisto (troboblasto) se

diferencia en tejidos extraembrionarios. El trofoblasto hueco es incapaz de formar un embrión vivo y así es necesario que las células madre embrionarias dentro del embrioblasto se diferencien y formen el embrión. Las iPSC inyectadas por micropipeta en un trofoblasto para generar un blastocisto transferido a hembras receptoras pueden producir crías de ratón vivo quimérico: ratones con derivados de iPSC incorporaron todo a través de sus cuerpos con 10 %-90 y quimerismo.

5 **ii. Reprogramación epigenética**

Desmetilación de promotores: la metilación es la transferencia de un grupo metilo a una base de ADN, normalmente la transferencia de un grupo metilo a una molécula de citosina en un sitio de CpG (secuencia de citosina/guanina adyacente). La metilación generalizada de un gen interfiere con la expresión evitando la actividad de proteínas de expresión o reclutando enzimas que interfieren con la expresión. Así, la metilación de un gen lo silencia eficazmente previniendo la transcripción. Los genes asociados a promotores de pluripotencia, que incluyen Oct4, Rex1 y Nanog, pueden desmetilarse en iPSC, mostrando su actividad promotora y la promoción y expresión activa de genes asociados a pluripotencia en iPSC.

Desmetilación de histonas: las histonas son proteínas de compactación que están estructuralmente localizadas en las secuencias de ADN que pueden efectuar su actividad mediante diversas modificaciones relacionadas con cromatina. Las histonas H3 asociadas a Oct4, Sox2 y Nanog pueden desmetilarse para activar la expresión de Oct4, Sox2 y Nanog.

15 **B. Cultivo de células iPS**

Después de introducirse las células somáticas con factores de reprogramación usando los procedimientos desvelados, estas células pueden cultivarse en un medio suficiente para mantener la pluripotencia. El cultivo de células madre pluripotentes inducidas (iPS) generadas en esta invención puede usar diversos medios y técnicas desarrollados para cultivar células madre pluripotentes de primate, más especialmente células madre embrionarias, como se describe en la solicitud de patente de EE.UU. 20070238170 y la solicitud de patente de EE.UU. 20030211603. Se aprecia que procedimientos adicionales para el cultivo y mantenimiento de células madre pluripotentes humanas, como se conocerían por un experto, pueden usarse con la presente invención.

En ciertas realizaciones pueden usarse condiciones no definidas; por ejemplo, pueden cultivarse células pluripotentes sobre células nodrizas de fibroblastos o un medio que se ha expuesto a células nodrizas de fibroblastos con el fin de mantener las células madre en un estado no diferenciado. Alternativamente, pueden cultivarse células pluripotentes y mantenerse en un estado esencialmente sin diferenciar usando sistema de cultivo independiente de nodrizas definido, tal como un medio TeSR (Ludwig y col., 2006a; Ludwig y col., 2006b). Los sistemas de cultivo y medios independientes de células nodrizas pueden usarse para cultivar y mantener células pluripotentes. Estos enfoques permiten que células madre embrionarias humanas sigan en un estado esencialmente no diferenciado sin la necesidad de "capas de células nodrizas" de fibroblastos de ratón. Como se describe en el presente documento, pueden hacerse diversas modificaciones a estos procedimientos con el fin de reducir los costes, según se desee.

Por ejemplo, al igual que las células madre embrionarias humanas (hES), las células iPS pueden mantenerse en 80 % de DMEM (Gibco n.º: 10829-018 o n.º: 11965-092), 20 % de suero bovino fetal (SBF) definido no inactivado por calor (o suero AB humano), 1 % de aminoácidos no esenciales, L-glutamina 1 mM y β-mercaptoetanol 0,1 mM. Alternativamente, las células iPS pueden mantenerse en medio libre de suero, preparado con 80 % de Knock-Out DMEM (Gibco n.º: 10829-018), 20 % de sustitución de suero (Gibco n.º: 10828-028), 1 % de aminoácidos no esenciales, L-glutamina 1 mM y β-mercaptoetanol 0,1 mM. Justo antes de uso, bFGF humano puede añadirse a una concentración final de aproximadamente 4 ng/ml (documento WO 99/20741) o puede usarse bFGF de pez cebra en su lugar como en los ejemplos.

Pueden usarse diversos componentes de la matriz en el cultivo y mantenimiento de células madre pluripotentes humanas. Por ejemplo, pueden usarse colágeno IV, fibronectina, laminina y vitronectina en combinación para recubrir una superficie de cultivo como un medio para proporcionar un soporte sólido para el crecimiento de células pluripotentes, como se describe en Ludwig y col. (2006a; 2006b).

También puede usarse Matrigel™ para proporcionar un sustrato para el cultivo y mantenimiento celular de células madre pluripotentes humanas. Matrigel™ es una mezcla de proteínas gelatinosas secretada por células tumorales de ratón y está comercialmente disponible de BD Biosciences (Nueva Jersey, EE.UU.). Esta mezcla se parece al entorno extracelular complejo encontrado en muchos tejidos y se usa por biólogos celulares como sustrato para el cultivo celular.

Las células iPS, al igual que las células ES, tienen antígenos característicos que pueden identificarse o confirmarse por inmunohistoquímica o citometría de flujo, usando anticuerpos para SSEA-1, SSEA-3 y SSEA-4 (Developmental Studies Hybridoma Bank, National Institute of Child Health and Human Development, Bethesda Md.) y TRA-1-60 y TRA-1-81 (Andrews y col., 1987). La pluripotencia de células madre embrionarias puede confirmarse inyectando aproximadamente 0,5-10 x 10⁶ células en los músculos de la pata trasera de ratones SCID macho de 8-12 semanas de edad. Se desarrollan teratomas que muestran al menos un tipo de célula de cada una de las tres capas germinales.

C. Diferenciación de células iPS

Pueden usarse diversos enfoques con la presente invención para diferenciar genéticamente células iPS en linajes celulares específicos que incluyen, pero no se limitan a, células de epitelio de la retina, células neurales de la retina, células hematopoyéticas, miocitos (por ejemplo, cardiomiocitos), neuronas, fibroblastos y células epidérmicas y tejidos u órganos derivados de los mismos. La diferenciación en células de la retina puede ser un ejemplo.

El epitelio pigmentario de la retina (RPE) es la capa de células pigmentadas justo fuera de la retina neurosensorial que alimenta células visuales de la retina y está firmemente unido a la coroides subyacente y a las células visuales de la retina que están encima. El RPE está compuesto por una única capa de células hexagonales que están densamente empaquetadas con gránulos de pigmentos. El epitelio pigmentario de la retina está implicado en la fagocitosis del segmento externo de células fotorreceptoras y también está implicado en el ciclo de la vitamina A en el que isomeriza todo-trans retinol en 11-cis retinal. El epitelio pigmentario de la retina también sirve como el factor de transporte limitante que mantiene el entorno de la retina suministrando moléculas pequeñas tales como aminoácido, ácido ascórbico y D-glucosa mientras que queda una barrera hermética a las sustancias transmitidas por la sangre coroidal. La homeostasis del entorno iónico se mantiene por un sistema de intercambio de transporte delicado. Cuando se ven desde la superficie externa, estas células son lisas y de forma hexagonal. Cuando se observan en sección, cada célula consiste en una parte no pigmentada externa que contiene un gran núcleo ovalado y una porción pigmentada interna que se extiende como una serie de procesos filiformes rectos entre los bastones, siendo esto especialmente el caso cuando el ojo se expone a la luz.

El aislamiento de células de RPE de paciente humano puede ser técnicamente desafiante y complicado. En ciertos aspectos de la presente invención, las células de RPE pueden obtenerse por diferenciación de células iPS para la identificación de defectos genéticos. Por ejemplo, células de RPE pueden diferenciarse mediante procedimientos desvelados en el documento WO 2008/129554, Osakada y col. (2008), Osakada y col. (2009) y Hiramí y col. (2009). Por ejemplo, células iPS humanas podrían disociarse en acumulaciones pequeñas de células y sembrarse en placas de Petri como cultivos en suspensión con un medio sin suero (SFEB/DL) que contiene los antagonistas de Wnt y Nodal. En estas condiciones, las células iPS podrían formar agregados similares a cuerpos embrioides. En el día 20, los agregados pueden plaquerarse sobre portaobjetos de vidrio recubiertos con poli-D-lisina, laminina y fibronectina para diferenciación de RPE adicional.

Las células iPS también pueden diferenciarse en células cardíacas o glóbulos sanguíneos en ciertos aspectos. Procedimientos a modo de ejemplo de diferenciación cardíaca de células iPS pueden incluir procedimientos de cuerpo embriode (EB) (Zhang y col., 2009), o procedimientos de células del estroma OP9 (Narazaki y col., 2008), o procedimientos de factor de crecimiento/procedimientos químicos (véanse las publicaciones de patentes de EE.UU. 20080038820, 20080226558, 20080254003 y 20090047739). Procedimientos a modo de ejemplo de diferenciación hematopoyética de células iPS pueden incluir, pero no se limitan a, procedimientos desvelados por las solicitudes de EE.UU. n.º: 61/088.054 y n.º: 61/156.304, o procedimientos basados en cuerpos embrioides (EB) (Chadwick y col., 2003; Ng y col., 2005). Los procedimientos de diferenciación de fibronectina también pueden usarse para la diferenciación del linaje de sangre, como se ejemplifica en Wang y col., 2007.

VII. Ejemplos

Los siguientes ejemplos están incluidos para mostrar realizaciones preferidas de la invención. Debe apreciarse por aquellos expertos en la materia que las técnicas desveladas en los ejemplos que siguen representan técnicas descubiertas por el autor de la invención que funcionan bien en la práctica de la invención y así puede considerarse que constituyen modos preferidos para su práctica.

Ejemplo 1

Se preparan células iPS a partir de un paciente diagnosticado con retinitis pigmentaria. Como un control también se preparan células iPS de un miembro de la familia del paciente, preferentemente un hermano, que esté sin retinitis pigmentaria. Se obtienen líneas de células iPS humanas por la transducción mediada por lentivirus de cuatro factores de transcripción (OCT4, SOX2, NANOG y LIN28) como se ha descrito previamente (Yu y col., 2007). Las células iPS derivadas tanto del paciente como del hermano se mantienen sobre Matrigel™ y se cultivan en medio TeSR.

Se prepara una única suspensión de células incubando colonias en TrypLe, una enzima similar a tripsina recombinante (Invitrogen, Carlsbad, CA) a 37 °C durante 7 minutos, se lava dos veces con medio TeSR que contiene el inhibidor de la apoptosis H1152 e inhibidor de tripsina de soja y se resuspende en 0,5 ml de PBS frío que contiene H1152 e inhibidor de tripsina de soja. Los cultivos en suspensión se siembran sobre una placa recubierta con gelatina con medio libre de suero complementado con 100 ng/ml de Dkk-1, un antagonista de Wnt y 500 ng/ml de Lefty A, un antagonista de Nodal durante 18-20 días para formar agregados. Después los agregados se siembran sobre portaobjetos de cultivo recubiertos con poli-D-lisina, laminina y fibronectina y se incuban en medio de diferenciación (G-MEM (GIBCO) que contiene 10 % de KSR

(sustitución de suero Knock-out), aminoácidos no esenciales 0,1 mM, piruvato de sodio 1 mM, 2-mercaptoetanol 0,1 mM, 50 unidades/ml de penicilina y 50 µg/ml de estreptomicina). Las células de RPE se seleccionan y se enriquecen por morfología poligonal, pigmentación y marcadores moleculares como RPE-65 y ZO-1, un marcador de la zona de oclusión.

- 5 Se extrae ARN total de las células de RPE usando el reactivo TRIzol (Invitrogen) o TRI-Reagent (Sigma). La síntesis de ADNc se lleva a cabo usando transcriptasa inversa del virus de la leucemia murina de Moloney (M-MLV RT) y cebadores al azar, según las instrucciones del fabricante (Promega Corporation, Madison, WI). Se lleva a cabo reacción en cadena de la polimerasa (PCR) usando protocolos convencionales con ADN polimerasa Taq (Gibco-BRL). Se usan Q-PCR, secuenciación de ADNc y micromatriz para comparar las células de RPE derivadas del paciente y el miembro de la familia normal emparentado para determinar una o más variaciones genéticas.

10 Todos los procedimientos desvelados y reivindicados en el presente documento pueden prepararse y ejecutarse sin excesiva experimentación en vista de la presente divulgación. Aunque las composiciones y procedimientos de la presente invención se han descrito en términos de realizaciones preferidas, será evidente para aquellos expertos en la materia que pueden aplicarse variaciones a los procedimientos y en las etapas o en la secuencia de etapas del procedimiento descrito en el presente documento.

REFERENCIAS

- Solicitud de EE.UU. n.º de serie 61/058.858
- Solicitud de EE.UU. n.º de serie 61/088.054
- Solicitud de EE.UU. n.º de serie 61/156.304
- 20 Solicitud de EE.UU. n.º de serie 61/172.079
- Solicitud de EE.UU. n.º de serie 61/184.546
- Publicación de patente de EE.UU. 20030211603
- Publicación de patente de EE.UU. 20070238170
- Publicación de patente de EE.UU. 20080038820
- 25 Publicación de patente de EE.UU. 20080226558
- Publicación de patente de EE.UU. 20080254003
- Publicación de patente de EE.UU. 20090047739
- Andrews y col., en: *Teratocarcinomas and Embryonic Stem Cells*, Robertson (Ed.), IRL Press, 207-246, 1987.
- Baird y col., *Am. J. Hum. Genet.*, 42: 677-93, 1988.
- 30 Bode y col., *Gene Ther. Mol. Biol.*, 6: 33-46, 2001.
- Boyer y col., *Cell*, 122(6): 947-56, 2005.
- Bray y col., *Mol. Psychiatry*, 9(1): 109-114, 2004.
- Byers y Steiner, *Ann. Rev. Med.*, 43: 269-82, 1992.
- Chadwick y col., *Blood*, 102(3): 906-15, 2003.
- 35 Chambers y col., *Cell*, 113(5): 643-55, 2003.
- Denoeud y col., *Genome Biol.*, 9(12): R175, 2008.
- Ding y col., *BMC Genes.*, 5(1): 8, 2004.
- Evans y col., en: *Cancer Principles and Practice of Oncology*, Devita y col. (Eds.), Lippincot-Raven, NY, 1054-1087, 1997.
- 40 Flinter, *J. Med. Genet.*, 34: 326-30, 1997.
- Haber, *J. Clin. Oncol.*, 17(11): 3367-70, 1999.

- Harper, en: Practical genetic counselling, 5th Ed., Butterworth-Heinemann, Boston, 1998.
- Hirami y col., *Neurosci Lett.*, 458(3): 126-31, 2009
- Hirota y col., *Hun. Mol. Genet.*, 13(23): 2959-2969, 2004.
- Johnson y col., *Pharmacology & Therapeutic*, 106: 19-38, 2005.
- 5 Kaji y col., *Nature*, 458, 771-775, 2009.
- Kamatani y col., *Am. J. Hum. Genet.*, 75(2): 190-203, 2004.
- Knight, *J Mol Med.*, 83(2): 97-109, 2005.
- Langle-Rouault y col., *J. Virol.*, 72(7): 6181-6185, 1998.
- Langston y col., *N. Engl J. Med.*, 334(3): 137-42, 1996.
- 10 Levitskaya y col., *Proc. Natl. Acad. Sci. USA*, 94(23): 12616-12621, 1997.
- Ludwig y col., *Nat. Biotechnol.*, 24(2): 185-187, 2006b.
- Ludwig y col., *Nat. Methods*, 3(8): 637-46, 2006a.
- Martin y col., *Am. J. Hum. Genet.*, 67(1): 146-154, 2000.
- Midgley y Kerr, *Lanet.*, 353: 391-9, 1999.
- 15 Narazaki y col., *Circulation*, 118(5): 498-506, 2008.
- Ng y col., *Development*, 132(5): 873-84, 2005.
- Osakada y col., *Nat. Biotech.*, 26: 215-224, 2008.
- Osakada y col., *Nat. Protocols*, 4: 811-824, 2009.
- Solicitud PCT WO 2008/129554
- 20 Solicitud PCT WO 99/20741
- Pinsonneault y col., *J. Pharmacol. Exp. Ther.*, 311(3): 1088-1098, 2004.
- Spielman and Ewens, *Am. J. Hum. Genet.*, 59(5): 983-989, 1996.
- Takahashi y col., *Cell*, 126(4): 663-676, 2006.
- Takahashi y col., *Cell*, 131 (5): 861-72, 2007.
- 25 Wang y col., *Nat. Biotechnol.*, 25(3): 317-8, 2007.
- Wenz, en: A novel high-throughput SNP genotyping system utilising capillary electrophoresis detection platforms, Applied Biosystems, CA, 2004.
- Wernig y col., *Nature*, 448(7151): 318-24, 2007
- Wojnowski y Brockmoller, *Pharmacogenetics*, 14(4): 267-269, 2004.
- 30 Woltjen y col., *Nature*, 458, 766-770, 2009.
- Yan y col., *Science*, 297(5584): 1143, 2002.
- Yu y col., *Science*, 318: 1917-1920, 2007.
- Yu y col., *Science*, 324(5928): 797-801, 2009.
- Zhang y col., *J. Pharmacol. Exp. Ther.*, 311(1): 373-381, 2004.
- 35 Zhang y col., *Circ Res.*, 104(4): e30-41, 2009.
- Zhou y col., *Cell Stem Cell*, 4(5): 381-4, 2009.

REIVINDICACIONES

1. Un procedimiento de diagnóstico de enfermedades genéticas o anomalía genética que comprende determinar la presencia de una variación genética específica de tejido en un sujeto de prueba con respecto a una estructura genética de control seleccionada, que comprende:
- 5 (a) obtener material genético de una célula diferenciada de un tejido seleccionado, habiéndose preparado la célula diferenciando una célula pluripotente embrionaria inducida (iPS) obtenida reprogramando una célula somática del sujeto; y
- (b) probar el material genético de dicha célula diferenciada para comparar una o más estructuras genéticas de dicho material genético con una o más estructuras genéticas de control para determinar la presencia de una variación genética tal en células del tejido del sujeto.
- 10 2. El procedimiento de la reivindicación 1, en el que el sujeto tiene o se sospecha que tiene una anomalía genética o enfermedad genética.
3. El procedimiento de la reivindicación 1 o 2, en el que dicho tejido seleccionado es retina.
4. El procedimiento de la reivindicación 3, en el que dicha célula diferenciada es una célula neural de la retina o una célula de epitelio pigmentario de la retina (RPE).
- 15 5. El procedimiento de una cualquiera de las reivindicaciones 1 a 4, en el que dicha prueba comprende secuenciación de ARN o ADN o en el que dicha prueba emplea análisis de micromatrices.
6. El procedimiento de una cualquiera de las reivindicaciones 1 a 5, en el que dicho material genético comprende ARN o ADN.
- 20 7. El procedimiento de una cualquiera de las reivindicaciones 1 a 6, en el que dicha estructura genética es una secuencia de nucleótidos o en el que dicha estructura genética se representa por un perfil de expresión sobre el nivel de ARN o nivel de proteína.
8. El procedimiento de una cualquiera de las reivindicaciones 1 a 7, en el que dicha variación genética
- (a) es un polimorfismo;
- 25 (b) es una mutación génica;
- (c) está representada por la expresión diferencial sobre el nivel de ARN o nivel de proteína;
- (d) es una variación de una o más secuencias codificantes;
- (e) es una variación de uno o más elementos reguladores de genes, o
- (f) es una variación epigenética.
- 30 9. El procedimiento de la reivindicación 8(a), en el que dicha variación genética es una variación de polimorfismo de un solo nucleótido.
10. El procedimiento de la reivindicación 8(e), en el que dicho elemento regulador génico es un promotor, potenciador, silenciador o elemento de respuesta o microARN.
11. El procedimiento de la reivindicación 8(f), en el que dicha variación epigenética es variación en la metilación de la estructura genética.
- 35 12. El procedimiento de una cualquiera de las reivindicaciones 1 a 11, en el que la estructura genética de control
- (a) es una conocida por no tener una variación genética seleccionada;
- (b) está comprendida en material genético, en el que el material genético es de una célula normal seleccionada preparada diferenciando una célula iPS obtenida reprogramando una célula somática normal de un sujeto de control, o
- 40 (c) es de un sujeto de control que es un pariente del sujeto de prueba.
13. El procedimiento de la reivindicación 12(a), en el que la estructura genética de control está comprendida en material genético de un tejido normal.

14. El procedimiento bien de la reivindicación 12(b) o bien de la reivindicación 13, en el que el tejido normal o la célula normal seleccionada es del mismo tipo de tejido que el tejido seleccionado del sujeto de prueba.
15. El procedimiento de la reivindicación 12(c), en el que el sujeto de control es un hermano del sujeto de prueba.