

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 745 298**

51 Int. Cl.:

**C12Q 1/6881** (2008.01)

**C12Q 1/6883** (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **20.03.2014 PCT/EP2014/055667**

87 Fecha y número de publicación internacional: **02.10.2014 WO14154569**

96 Fecha de presentación y número de la solicitud europea: **20.03.2014 E 14711284 (1)**

97 Fecha y número de publicación de la concesión europea: **19.06.2019 EP 2978858**

54 Título: **Método in vitro para la evaluación pronóstica de las perspectivas de éxito de una implantación y/o un trasplante**

30 Prioridad:

**27.03.2013 DE 102013205516**  
**28.01.2014 DE 102014201528**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**28.02.2020**

73 Titular/es:

**TETEC TISSUE ENGINEERING TECHNOLOGIES AG (100.0%)**  
**Aspenhastrasse 18**  
**72770 Reutlingen, DE**

72 Inventor/es:

**MOLLENHAUER, JÜRGEN;**  
**GAISSMAIER, CHRISTOPH y**  
**BENZ, KARIN**

74 Agente/Representante:

**TOMAS GIL, Tesifonte Enrique**

ES 2 745 298 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

Método *in vitro* para la evaluación pronóstica de las perspectivas de éxito de una implantación y/o un trasplante

## CAMPO DE APLICACIÓN Y ESTADO DE LA TÉCNICA

5 [0001] La presente invención se refiere a un método *in vitro* para la evaluación pronóstica de las perspectivas de éxito de una implantación y/o un trasplante.

[0002] La elaboración de perfiles de expresión génica o el análisis del transcriptoma se han establecido con el establecimiento de la tecnología de micromatrices como una herramienta importante en la ciencia biomédica.

10 [0003] Particularmente, el desarrollo de métodos de secuenciación de ARN de segunda generación (Next Generation Sequencing, NGS) no solo ha llevado a una disminución drástica de los costes para la realización de un análisis transcriptómico, sino que también aumenta la precisión de la identificación de actividades génicas hasta ahora desconocidas. Ejemplos para campos de aplicación de perfiles de expresión génica son el diagnóstico y el pronóstico de enfermedades, el análisis de seguimiento de los tratamientos, el análisis de las predisposiciones genéticas, el análisis de los mecanismos de acción farmacológicos, así como el análisis cualitativo y cuantitativo de los procesos de crecimiento y diferenciación de células y tejidos.

15 [0004] Un método usual para la evaluación de datos de expresión génica lo constituye el análisis diferencial con el que se examinan tanto la expresión de genes conocidos como también se puede efectuar la detección de genes desconocidos. En este método se verifican o comparan los datos de expresión de la muestra que se va a examinar con los patrones de expresión génica de muestras de referencia o también con los datos de expresión de genes seleccionados. En el caso del análisis de la expresión de genes patofisiológicamente relevantes se comparan, por  
20 ejemplo, los datos de expresión de tejido sano (muestra de referencia) con los datos de expresión de tejido enfermo (muestra de medida) como, por ejemplo, tejido tumoral. En base a esta comparación se pueden hacer declaraciones de la expresión cualitativa (respuesta sí/no) o cuantitativa (aumento o reducción de la expresión) de genes seleccionados y estos clasificarse a su vez en un estado determinado, por ejemplo, un estado patológico.

25 [0005] En la DE 10 2010 033 565 A1 se divulgan diversos marcadores para la determinación *in vitro* de la identidad farmacéutica, pureza o potencia de los condrocitos (células de cartílago), mediante los que se comprueba la aptitud de los condrocitos para un trasplante de condrocitos exitoso. El establecimiento de estos marcadores se impulsó por el hecho de que los condrocitos, en cuanto a su aptitud para utilizarse como células autólogas para una implantación para la regeneración de cartílago, pueden variar mucho, no solo los condrocitos de un donante en  
30 relación a los condrocitos de otro donante, sino también los condrocitos de un mismo donante. Además, se consideró que el cultivo de condrocitos puede cambiar sus características de tal manera que dejan de ser tan adecuados para una implantación como directamente después del aislamiento del donante.

[0006] De DE 102 43 131 A1 se conoce un método para la identificación de un cultivo de condrocitos que está en posición de servir como trasplante de cartílago, así como un método para la determinación del estado de expresión de uno o varios genes en un cultivo de condrocitos.

35 [0007] El objeto de WO 2011/127561 A1 es un método para la evaluación de las perspectivas de éxito en pacientes con fibrosis pulmonar, en el que se determina el nivel de expresión génica de diversos genes.

[0008] Además se conoce que existen diferencias en el fenotipo, así como en la expresión génica entre condrocitos humanos primarios y células madre mesenquimales durante la diferenciación condrogénica en cultivos de pellets tridimensionales (P. Bernstein et al.: "Expression pattern differences between osteoarthritic chondrocytes and  
40 mesenchymal stem cells during chondrogenic differentiation", Osteoarthritis and Cartilage (2010), 18, páginas 1596-1607).

[0009] Además se comparó la expresión de genes marcadores en condrocitos de pacientes que habían recibido un trasplante de condrocitos autólogos, con la expresión de genes marcadores en condrocitos de pacientes enfermos de artritis. A tal objeto se determinó el perfil de expresión génica de colágeno tipo I, colágeno tipo II, colágeno tipo X, agregano, IL-1 $\beta$ , así como quinasa-1 similar a la activina (Stoop et al.: "Comparison of marker gene expression  
45 in chondrocytes vom patients receiving autologous chondrocyte transplantation versus osteoarthritis patients", Arthritis Research & Therapy (2007) 9/3: R60).

[0010] De WO 2008/061804 A2 se conoce un método *in vitro* para la determinación de la facultad de una población celular para la producción de cartílago hialino estable. El método se basa en la determinación de la expresión de

un conjunto de al menos tres genes marcadores, donde el conjunto muestra FRZB, ALK1 y uno o varios marcadores seleccionados del grupo consistente en PEDF, COL11, COL2, FGFR3, OPN, BMP-2 y PLA-RASF.

5 [0011] Aunque un análisis selectivo de unos pocos genes, particularmente implicados en el metabolismo celular, puede conducir ciertamente a resultados potentes en el aseguramiento de la calidad de las células que se van a trasplantar, no obstante, los resultados de un tal método son limitados en su significación estadística, tanto más porque la diferenciación condrocítica únicamente representa un parámetro para la evaluación de un éxito de curación.

## OBJETIVO Y SOLUCIÓN

10 [0012] Ante estos antecedentes, la presente invención tiene, por lo tanto, la tarea de proporcionar un método que evite los inconvenientes conocidos del estado de la técnica y permita particularmente un pronóstico individual más válido.

[0013] Esta tarea se resuelve mediante un método *in vitro* con las características de la reivindicación independiente 1. Se indican formas de realización preferidas del método en las reivindicaciones dependientes 2 a 5. El texto de todas las reivindicaciones se incorpora por la presente por referencia explícita al contenido de la descripción.

15 [0014] Además se divulga un método *in vitro* para la evaluación pronóstica o el pronóstico de la capacidad regenerativa tisular y/o la potencia celular y/o las perspectivas de éxito de una implantación, preferiblemente una implantación celular, y/o trasplante, especialmente para la evaluación pronóstica o el pronóstico del fracaso de un implante y/o de un trasplante.

20 [0015] El método se caracteriza especialmente por el hecho de que el transcriptoma de células, particularmente células de un paciente (células del paciente), y/o la expresión génica de células que se origina en el transcriptoma o basada en el transcriptoma, particularmente de células de un paciente (células del paciente), se analizarán/analizará *in vitro*.

25 [0016] En el análisis transcriptómico y/o el análisis de la expresión génica basada en el transcriptoma pueden/puede detectarse con particular ventaja el comportamiento de transcripción y/o de traducción de todos los genes de una célula y no solo -como se conoce del estado de la técnica- la transcripción o la traducción de unos pocos, particularmente por medio de su importancia para el metabolismo celular, genes seleccionados específicamente. Mediante el análisis transcriptómico proporcionado se puede generar particularmente un perfil metabólico completo, expresado en la actividad de expresión génica de las células en cuestión, preferiblemente células del paciente.

30 [0017] Sorprendentemente se ha hallado ahora que a los perfiles de expresión transcriptómica y/o génica obtenidos de células del paciente en el marco del análisis transcriptómico se les puede asignar un valor terapéutico en cuanto a la capacidad regenerativa tisular, así como a las perspectivas de éxito de una medida implantológica y/o trasplantológica en el o los pacientes en cuestión. Los inventores pudieron verificar esto en el marco de un seguimiento clínico retrospectivo con el ejemplo de un trasplante de condrocitos autólogos guiado por matriz o por soporte (MACT).

35 [0018] Puesto que los perfiles obtenidos se basan en un análisis *in vitro* del transcriptoma y/o en una expresión génica basada en el transcriptoma, es posible un pronóstico individual más válido en comparación con métodos genéricos, es decir, un pronóstico individual con una significación estadística más alta.

[0019] En otras palabras, se divulga un método para pronosticar la regeneración de tejido, la potencia celular, de un éxito o fracaso de la implantación y/o un éxito o fracaso del trasplante.

40 [0020] La expresión "implantación celular" se refiere en el sentido de la presente divulgación a una implantación aplicando un implante cargado o inoculado con células.

45 [0021] La expresión "transcriptoma" comprende en el sentido de la presente divulgación al menos la suma de los genes transcritos de ADN en ARNm (messenger RNA o ARN mensajero) en un momento determinado en una célula. sin embargo, la expresión "transcriptoma" comprende preferiblemente en el sentido de la presente divulgación la suma de los genes transcritos de ADN en ARN en un momento determinado en una célula, es decir, la totalidad de todas las moléculas de ARN producidas en una célula.

[0022] La expresión "perfil transcriptómico" (o patrón del transcriptoma) muestra en el sentido de la presente divulgación el perfil (o patrón) de todos los transcritos detectables de las células preferiblemente mediante métodos

basados en hibridación y/o basados en secuencias, particularmente métodos de secuenciación de segunda generación.

5 [0023] La expresión "expresión génica" comprende en el sentido de la presente invención la síntesis que tiene lugar en el transcurso de la transcripción de ARN, particularmente de ARNm (producto génico primario), ARN regulador y/o otros tipos de ARN, y/o la traducción a proteínas (productos génicos secundarios) establecida sobre secuencias de ARNm maduras. Como ejemplos de ARN regulador se citan el microARN (miARN), el ARN de interferencia pequeño (ARNip) y/o el ARN nuclear pequeño (small nuclear RNA). El ARN ribosómico (ARNr) y/o el ARN de transferencia (ARNt) representan ejemplos de otros tipos de ARN mencionados además en este párrafo, que se cuentan también entre los productos génicos primarios.

10 [0024] La expresión "perfil de expresión génica" (o patrones de expresión génica) muestra en el sentido de la presente invención la interpretación de los datos generados preferiblemente mediante métodos basados en hibridación y/o basado en secuencias, particularmente métodos de secuenciación de segunda generación, como perfil (o patrón) de las actividades génicas de las células examinadas.

15 [0025] La expresión "regeneración de tejido" puede comprender en el sentido de la presente divulgación fundamentalmente la regeneración de cualquier tejido corporal o de un paciente. La expresión "regeneración de tejido" comprende sin embargo preferiblemente la regeneración del tejido de sostén, preferiblemente la regeneración del tejido cartilaginoso, de manera especialmente preferida la regeneración del cartílago articular, y/o la regeneración del tejido cartilaginoso intervertebral.

20 [0026] Correspondientemente, la expresión "capacidad regenerativa tisular" puede comprender en el sentido de la presente divulgación fundamentalmente la facultad para la regeneración de cualquier tejido corporal o de un paciente, particularmente la facultad para la regeneración del tejido de sostén, preferiblemente la regeneración del tejido cartilaginoso, de manera especialmente preferida la regeneración del cartílago articular, y/o la regeneración del tejido cartilaginoso intervertebral.

25 [0027] Bajo la expresión "potencia celular" debe entenderse en el sentido de la presente divulgación la facultad de las células tisulares para desarrollar características específicas de tejido y/o mantener o reanudar el desarrollo de características específicas de tejido, particularmente tras un cultivo *in vitro* previo. Por ejemplo, bajo la potencia de los condrocitos debe entenderse su facultad para producir matriz extracelular y/o reanudar la producción de matriz extracelular, particularmente cuando los condrocitos se implantan en una zona defectuosa que hay que tratar.

30 [0028] La expresión "implantación celular guiada por matriz o por soporte" o "trasplante celular guiado por matriz o por soporte" significa en el sentido de la presente invención la implantación o el trasplante de un implante provisto o inoculado con células autólogas.

[0029] Las células pueden ser de origen fundamentalmente humano y/o animal. En otras palabras, las células pueden tratarse de células humanas y/o animales.

[0030] Las células provienen preferiblemente de un paciente humano.

35 [0031] Particularmente, las células pueden tratarse de células propias del cuerpo o autólogas.

[0032] Preferiblemente, las células se toman de un paciente en forma de una muestra de tejido. En función del tipo o la procedencia de la muestra puede ser ventajoso preparar la muestra antes de la realización del análisis transcriptómico. Un tratamiento adecuado de la muestra puede comprender pasos como la centrifugación, la concentración, la homogeneización, la propagación *in vitro*, así como otras etapas de tratamiento conocidas fundamentalmente por la persona experta.

40

[0033] En una forma de realización preferida, las células provienen del tejido de un paciente, cuya facultad de regeneración y/o potencia celular debe evaluarse.

[0034] Las células provienen particularmente del tejido de un paciente con un defecto que debe tratarse mediante la implantación y/o el trasplante.

45 [0035] Las células provienen preferiblemente de un tejido de sostén, preferiblemente además de tejido cartilaginoso, de manera especialmente preferida de tejido cartilaginoso articular, y/o tejido intervertebral.

[0036] En otra forma de realización, las células son células de tejido de sostén, preferiblemente condrocitos (células de cartílago), y/o células precursoras de los mismos, de manera especialmente preferida condrocitos articulares, células del disco intervertebral, particularmente células del núcleo y/o del anillo, y/o células precursoras de las mismas.

5 [0037] En otra forma de realización, las células son células sanas o células que provienen de partes o áreas de tejido sanas.

[0038] En el sentido de la presente invención, la implantación es, en una forma de realización particularmente preferida, una implantación celular autóloga guiada por matriz o por soporte, preferiblemente una implantación de condrocitos autólogos guiada por matriz o por soporte (matrix-assisted autologous chondrocyte implantation, MACI).  
 10 Particularmente, los soportes de colágeno representan matrices adecuadas. Una matriz preferida o un soporte de colágeno preferido es un implante construido en capas de una membrana pericárdica y una esponja de colágeno, donde la esponja de colágeno posee poros preferiblemente columnares, que están orientados perpendiculares o esencialmente perpendiculares a la membrana pericárdica y se forman mediante una liofilización unilateral. Un soporte de colágeno de este tipo se distribuye comercialmente por el solicitante, por ejemplo, bajo la denominación  
 15 Novocart® Basic o Novocart® 3D. Con respecto a otras características y ventajas de un tal soporte de colágeno se hace referencia además a EP 1 824 420 B1, cuyo contenido de la divulgación respecto al implante descrito en la misma para la reparación de un defecto del cartílago se incorpora por la presente por referencia explícita al contenido de la presente descripción.

20 [0039] En otra forma de realización, el trasplante es en el sentido de la presente invención un trasplante celular autólogo, preferiblemente un trasplante de condrocitos autólogos.

[0040] Preferiblemente, antes de la realización del análisis transcriptómico y/o el análisis de la expresión génica que se origina en el transcriptoma, las células se cultivan y, particularmente, se propagan *in vitro*. El cultivo puede ocurrir, por ejemplo, en un medio de cultivo que está enriquecido preferiblemente con suero autólogo u homólogo.

25 [0041] Las células pueden cultivarse particularmente durante un período de 14 días a 30 días, particularmente de 17 días a 24 días, preferiblemente de 19 días a 21 días.

[0042] La evaluación pronóstica se realiza, en una forma de realización preferida, por medio de un perfil transcriptómico obtenido mediante el análisis transcriptómico y/o un perfil de expresión génica con origen en el perfil transcriptómico.

30 [0043] En una forma de realización adicional se compararán/comparará el perfil transcriptómico y/o el perfil de expresión génica con origen en el perfil transcriptómico con un perfil transcriptómico y/o un perfil de expresión génica del mismo tipo celular, que son/es indicativo(s) o específico(s) o característico(s) de una regeneración de tejido exitosa, una implantación exitosa y/o un trasplante exitoso y/o para la presencia de potencia celular.

35 [0044] Alternativa o complementariamente a la forma de realización anterior se compararán/comparará el perfil transcriptómico y/o el perfil de expresión génica con origen en el perfil transcriptómico con un perfil transcriptómico y/o un perfil de expresión génica del mismo tipo celular, que son/es indicativo(s) o específico(s) o característico(s) de una regeneración de tejido infructuosa o poco prometedora, una implantación infructuosa o poco prometedora y/o un trasplante infructuoso o poco prometedor y/o para la falta de potencia celular.

40 [0045] Por medio de perfiles transcriptómicos y/o de expresión génica indicativos o específicos o característicos mencionados en las dos formas de realización previas es posible con particular ventaja un pronóstico (más) fiable de un posible éxito de la regeneración de tejido, de la implantación o del implante y/o del trasplante o -expresado de otra manera- un posible fracaso de la regeneración de tejido, de la implantación o del implante y/o del trasplante.

45 [0046] El perfil transcriptómico y/o el perfil de expresión génica indicativo o específico o característico de una regeneración de tejido exitosa, una implantación exitosa y/o un trasplante exitoso y/o la presencia de potencia celular se determinarán/determinará preferiblemente mediante la evaluación de perfiles transcriptómicos y/o de perfiles de expresión génica con origen en el perfil transcriptómico del mismo tipo celular de pacientes en los que la regeneración de tejido, la implantación y/o el trasplante ha transcurrido con éxito, y/o en los que el tipo celular ha sido potente.

50 [0047] El perfil transcriptómico y/o perfil de expresión génica indicativo o específico o característico de una regeneración de tejido infructuosa o poco prometedora, una implantación infructuosa o poco prometedora y/o un trasplante infructuoso o poco prometedor y/o la falta de potencia celular se determinarán/determinará preferiblemente mediante la evaluación de perfiles transcriptómicos y/o de perfiles de expresión génica con origen

en el perfil transcriptómico del mismo tipo celular de pacientes en los que la regeneración de tejido, la implantación y/o el trasplante ha sido infructuoso o ha fracasado, y/o en los que el tipo celular no ha sido potente.

5 [0048] El perfil transcriptómico y/o el perfil de expresión génica indicativos mencionados en las formas de realización anteriores se determinarán/determinará preferiblemente en el marco de un seguimiento o análisis clínico retrospectivo de los resultados terapéuticos.

[0049] Además es preferible que la determinación del perfil transcriptómico y/o el perfil de expresión génica indicativos se realice mediante un algoritmo de búsqueda, preferiblemente un algoritmo de búsqueda asistido por ordenador. Esto puede ocurrir con la ayuda de cada uno de los *softwares* de evaluación disponibles (comerciales) para datos transcriptómicos, como, por ejemplo, en el ejemplo de aplicación representado.

10 [0050] Para la realización del análisis transcriptómico en una forma de realización apropiada se aísla ARN de las células. A tal objeto, las células se lisan generalmente en un entorno químico, en el que se desnaturalizan rápidamente las RNAsas (ribonucleasas). El ARN se separa a continuación de los demás componentes celulares como, por ejemplo, ADN, proteínas, azúcares, lípidos o similares. El aislamiento del ARN puede consistir en una extracción o purificación. Por ejemplo, se puede aislar ARN mediante el denominado método del tiocianato de guanidina con extracción de fenol/cloroformo sucesiva.

15

[0051] El ARN aislado puede someterse a un análisis cualitativo y/o cuantitativo. Una determinación cualitativa del ARN aislado puede ocurrir, por ejemplo, con ayuda de un fotómetro, que para la elaboración de un espectro de ácido nucleico generalmente entre 220 nm y 450 nm solo necesita en la mayoría de los casos una cantidad de muestra muy pequeña. Por un lado, se mide típicamente la proporción de absorción 260 nm/280 nm y, por otro lado, la proporción de absorción 260 nm/230 nm. La proporción de absorción 260 nm/280 nm debería estar entre 1,8 y 2,0. Esto permite particularmente deducciones sobre contaminaciones proteicas. La proporción de absorción 260 nm/230 nm debería estar sobre 1,8 e indica particularmente contaminaciones con solventes, sales, así como proteínas. El análisis electroforético representa otro método adecuado para el análisis cualitativo y/o cuantitativo, con el cual se separa ARN aislado por electroforesis capilar en un chip especial y se obtiene un denominado electroferograma de ARN.

20

25

[0052] En otra forma de realización se elimina, dentro del marco del análisis transcriptómico, ARN no codificante, particularmente ARN no codificante y no regulador.

[0053] Se eliminarán/ eliminará preferiblemente, dentro del marco del análisis transcriptómico, ARN ribosómico (ARNr) y/o ARN de transferencia (ARNt). De este modo se consigue con particular ventaja un enriquecimiento de ARN codificante y/o ARN regulador y se permite la realización del análisis transcriptómico sin la molesta interferencia de otro ARN. En otras palabras, se prefiere que el análisis transcriptómico se realice exclusivamente por medio de ARN codificante y/o ARN regulador. Preferentemente se realiza un empobrecimiento o depleción de ARNr.

30

[0054] Para la eliminación de ARN ribosómico (ARNr) están disponibles fundamentalmente kits de preparación de diversos fabricantes. Por ejemplo, se puede usar para la eliminación de ARNr el kit RiboMinus™ Eukaryote (Firma Life Technologies), que se basa en la eliminación selectiva de moléculas grandes de ARN ribosómico frecuentemente existentes en el conjunto de ARN total. Esto se consigue mediante una hibridación de estos ARNr con sondas oligonucleotídicas marcadas con biotina específicas de secuencia. El complejo hibridado se inmoviliza y elimina entonces mediante bolitas (esferas) magnéticas revestidas con estreptavidina. El producto empobrecido en ARNr se concentra generalmente a continuación.

35

40

[0055] Después de la eliminación de ARN no codificante y particularmente no regulador, preferiblemente de ARN ribosómico (ARNr) y/o ARN de transferencia (ARNt), se puede realizar (de nuevo) un análisis cualitativo y/o cuantitativo. En este sentido se hace referencia completamente al análisis cualitativo y/o cuantitativo previamente descrito en relación con el ARN aislado.

[0056] En una forma de realización particularmente preferida se realiza el análisis transcriptómico exclusivamente por medio de ARNm (messenger ARN o ARN mensajero). El ARNm es ARN procesado, que entre otros ya ha pasado por el denominado empalme, es decir, al contrario que el pre-ARNm o el ADN natural, ya no contiene ningún intrón (segmentos no codificantes).

45

[0057] En otra forma de realización, el análisis transcriptómico comprende una fragmentación de ARN, preferiblemente de ARNm. La fragmentación puede ocurrir mediante una digestión enzimática, generalmente mediante una RNasa (ribonucleasa) como, por ejemplo, la RNasa III, y/o físicamente, por ejemplo, mediante

50

ultrasonidos. Los fragmentos adecuados para el método según la invención pueden comprender de 30 a 1000 nucleótidos. La fragmentación se realiza preferiblemente después de la eliminación de ARNr.

5 [0058] En otra forma de realización, el análisis transcriptómico comprende la realización de una transcripción inversa, es decir, la transcripción de ARN, particularmente ARNm, en ADNc (ADN complementario). La transcripción se realiza preferiblemente después de una fragmentación del ARN. La transcripción se realiza generalmente con ayuda de la enzima transcriptasa inversa. El producto primario obtenido en la transcripción inversa es una cadena de ADNc, que se hibrida con la cadena de ARN original. Esta última puede desintegrarse luego mediante la RNasa H. En otro paso se sintetiza mediante una polimerasa de ADN dependiente de ADN (a través de un cebador) una cadena de ADN complementaria a la cadena simple de ADNc ya existente para obtener ADNc bicatenario.

10 [0059] En una forma de realización adicional el análisis transcriptómico comprende la multiplicación o la amplificación de ADNc bicatenario. Preferiblemente, el ADNc se multiplica o se amplifica mediante una reacción en cadena de la polimerasa (PCR), particularmente una reacción en cadena de la polimerasa en emulsión (PCR en emulsión).

15 [0060] Con ayuda de una transcripción inversa y una amplificación subsiguiente a la misma del ADNc obtenido en el curso de la transcripción inversa se pueden elaborar con particular ventaja genotecas de ADNc.

[0061] En formas de realización apropiadas se puede realizar antes de la multiplicación o amplificación una selección de tamaños del ADNc bicatenario mediante electroforesis en gel de poliacrilamida (PAGE).

[0062] El ADNc se somete en otra forma de realización a un método de secuenciación.

20 [0063] El análisis transcriptómico comprende preferiblemente un método de micro o macromatrices basado en hibridación, también denominado método de matrices de hibridación de ADN, o un método basado en secuencias, preferiblemente un método de secuenciación de segunda generación.

[0064] Tanto el método de micro o macromatrices como también el método basado en secuencias permiten respectivamente la expansión del análisis de expresión génica a un enfoque de todo el genoma, en cuanto a que permiten en un experimento la detección simultánea de las diferencias de expresión de varios miles de genes.

25 [0065] El método de micro o macromatrices se parece a las técnicas de hibridación convencionales de la biología molecular con respecto a su principio de funcionamiento como, por ejemplo, el análisis de transferencia de Southern o Northern. Estos métodos usan la característica de los ácidos nucleicos de hibridar entre sí de manera específica de secuencia. Bajo hibridación se entiende la unión no covalente de dos cadenas simples de ácido nucleico complementarias una respecto a la otra, que se basa principalmente en la formación de puentes de hidrógeno entre  
30 las bases heterocíclicas de las moléculas de ácido nucleico.

[0066] En el método de micro o macromatrices o el método de matrices de hibridación de ADN se aplican y se inmovilizan en posiciones específicas ácidos nucleicos, cuyas secuencias son conocidas, denominadas sondas, en gran número y alta densidad, generalmente con ayuda de un robot. Estas matrices de hibridación de ADN se hibridan a continuación con ácidos nucleicos marcados. Para el marcaje pueden incorporarse en el ADNc, por  
35 ejemplo, nucleótidos marcados con radiactividad o fluorescencia durante la transcripción inversa de ARN, generalmente ARNm. Puesto que se realiza una hibridación solo entre moléculas de ácido nucleico complementarias, la intensidad de la señal medida es proporcional a la frecuencia de las hibridaciones resultantes. Puesto que cada posición de una sonda corresponde a un gen o fragmento de gen determinado, la intensidad de señal medida en esta posición proporciona una medida para el nivel de expresión relativo de este gen. En función  
40 del número de las sondas de genes disponibles y la densidad con la que estas se aplican sobre el soporte, se pueden analizar simultáneamente mediante matrices de este tipo varios miles de genes.

[0067] Los métodos de secuenciación de segunda generación ya no se basan en una separación del ADN a través de electroforesis capilar como en el denominado método Sanger, sino en un acoplamiento de fragmentos de ADNc a soportes sólidos y la unión complementaria de nucleótidos u oligonucleótidos individuales, donde la unión de los  
45 mismos se confirma mediante cámaras de alta resolución.

[0068] En una forma de realización preferida del método según la invención, el análisis transcriptómico comprende un método de secuenciación de segunda generación, seleccionado del grupo que comprende la pirosecuenciación, la secuenciación a través de la síntesis (secuenciación por síntesis) y la secuenciación a través de ligamiento (secuenciación por ligamiento).

[0069] En la pirosecuenciación, los fragmentos de ADN se hibridan sobre moléculas conectoras, generalmente en forma de adaptadores de oligopéptidos, con esferas (bolitas) (un fragmento por esfera). Los fragmentos de ADN se multiplican luego mediante una reacción en cadena de la polimerasa (PCR). A tal objeto vienen las esferas en una emulsión con reactivos de PCR. Las nuevas copias de ADN formadas a consecuencia de la reacción en cadena de la polimerasa se quedan enganchadas también a las esferas. Para la secuenciación, las esferas se distribuyen a continuación en placas de titulación idóneas, preferiblemente placas de picotitulación, en cuyas cavidades (pocillos) se encuentran las enzimas y los cebadores que hacen falta para la realización de la secuenciación. Consecutivamente se añaden entonces los cuatro nucleótidos desoxiadenosina trifosfato (dATP), desoxiguanosina trifosfato (dGTP), desoxicitidina trifosfato (dCTP) y desoxitimidina trifosfato (dTTP). En cada incorporación de nucleótido se libera pirofosfato, que, como ATP, por ejemplo, estimula la enzima luciferasa, para convertir la luciferina en oxiluciferina y luz. Las cavidades correspondientes de las placas de titulación se iluminan. Puesto que solo se añade un nucleótido por paso de secuenciación, se puede determinar de este modo por medio de la señal la secuencia de los fragmentos de ADN.

[0070] En el método de secuenciación "secuenciación a través de la síntesis" (secuenciación por síntesis) se utilizan nucleótidos de terminación reversibles. Los fragmentos de ADN que se van a secuenciar se unen a la superficie de vidrio de una célula de flujo y se multiplican mediante una reacción en cadena de la polimerasa (PCR). Las copias de la PCR se fijan alrededor del fragmento de ADN original, por lo cual surgen moléculas idénticas al grupo. La secuenciación se basa -de manera similar al método Sanger- en nucleótidos de terminación reversibles. Los reactivos de síntesis (cebadores, polimerasa de ADN y cuatro nucleótidos de terminación diferentes reversibles marcados con un colorante fluorescente) se proporcionan sobre la célula de flujo. Cuando en un fragmento de ADN se añade uno de los cuatro nucleótidos de terminación, el fluoróforo bloquea la síntesis adicional. La reacción se detiene brevemente, el colorante y el nucleótido de terminación se disocian y la señal luminosa se documenta, antes de comenzar una nueva ronda.

[0071] En el método de secuenciación "secuenciación a través de ligamiento" (secuenciación por ligamiento) se realiza la verdadera reacción de secuenciación después de una reacción en cadena de la polimerasa en emulsión (PCR en emulsión) en esferas (bolitas). En una primera ronda (de cinco en total) se proporcionan, además de cebadores de secuenciación universales (longitud  $n$ ), una mezcla de cuatro oligonucleótidos octaméricos diferentes en la reacción. En las posiciones 1 y 2 de estos octámeros hay bases definidas (cuatro de 16 emparejamientos dinucleotídicos posibles; en las cinco rondas se utilizan los 16 dinucleótidos posibles), que se codifican por uno de cuatro colorantes fluorescentes. El oligonucleótido octamérico adecuado hibrida con el fragmento de PCR y se liga con el también hibridado cebador de secuenciación. Se mide la señal fluorescente y se elimina el colorante junto a los últimos tres nucleótidos. Estos pasos se repiten varias veces, según la longitud del ADN (para 30-35 bases serían 6-7 rondas, donde en el ciclo siguiente se exploran las bases 6/7, después las 11/12). Finalmente se vuelven a eliminar todas las construcciones de cebador-oligonucleótido ligadas (reinicio). Se inicia una nueva ronda con un nuevo cebador de secuenciación de la longitud  $n-1$  y otros cuatro dinucleótidos marcados con fluorescencia. Ahora se identifican por lo tanto en la primera ronda de ciclo las bases  $n-1$  y 1, entonces la base 5/6, 10/11, etc. Tras tres rondas más (cebador  $n-2$ ,  $n-3$  y  $n-4$ ) se encuentra la secuencia, cada base se comprobó mediante dos oligonucleótidos diferentes.

[0072] Con respecto a una visión general sobre los métodos de secuenciación de alto rendimiento establecidos actualmente se hace referencia a las publicaciones de Niedringhaus et al. (Landscape of Next-Generation Sequencing Technologies, Anal. Chem. 2011, 83, 4327-4341) y Hurd et al. ((Advantages of next-generation sequencing versus the microarray in epigenetic research, BRIEFINGS IN FUNCTIONAL GENOMICS AND PROTEOMICS. VOL 8. N.º 3. 174-183), así como sobre el artículo de Hollicher (Hochleistungs-Sequenzieren, Laborjournal 2009, 4, 44-48), cuyo contenido de la divulgación respecto a los métodos de secuenciación descritos en los mismos respectivamente se incorpora por referencia explícita al contenido de la presente descripción.

[0073] El análisis transcriptómico comprende en otra forma de realización un ensamblaje o una unión del ADNc o fragmentos de ADNc secuenciado. De este modo se pueden extraer deducciones sobre parentescos funcionales o evolutivos y, por lo tanto, sobre la secuencia original. El ensamblaje puede realizarse mediante los correspondientes métodos bioinformáticos usuales para la persona experta.

[0074] En otra forma de realización se someten los fragmentos de ADNc ensamblados a una anotación de genes, que permite una identificación de secuencias portadoras de información, particularmente de genes regulados diferencialmente. La anotación de genes se apoya convenientemente en métodos bioinformáticos, mediante los cuales pueden descubrirse patrones o perfiles y relacionarse con los conocimientos conocidos, particularmente sobre redes metabólicas y reguladoras. El análisis de estos datos requiere fundamentalmente una normalización antes del verdadero tratamiento, por ejemplo, por el método de conglomerados. Si los datos están en forma de cocientes de valores medidos a través de un experimento de tratamiento y un experimento de referencia, se consigue generalmente una normalización a través de la formación de logaritmos. Otras normalizaciones se basan, por ejemplo, en la norma vectorial, la jerarquía, la variación unitaria o la denominada puntuación  $z$ . Esta última se trata de un método para decidir si un valor determinado significativo se encuentra por debajo o por encima de un

valor medio. En este caso, un valor negativo es un indicador para valores más pequeños y un valor positivo, un indicador para valores más grandes que un valor medio. El análisis de la desviación típica proporciona la significación de esta desviación. Para una visualización de estos datos hay diversos sistemas de *software* disponibles que permiten por un lado generalmente una estructuración de los datos por medio de categorías funcionales diferentes, así como, por otro lado, una visualización según la categorización ocurrida. La asignación funcional o categorización puede ocurrir fundamentalmente sobre la base de anotaciones disponibles en relación con algoritmos de búsqueda conocidos o mediante una combinación de anotaciones disponibles y algoritmos de búsqueda hallados individualmente. La base de este tipo de algoritmos de búsqueda está formada en la mayoría de casos por análisis diferenciales, en los que se comparan el patrón de expresión génica de una muestra que se va a examinar con muestras de referencia, que representan un fenotipo patofisiológico determinado. Por medio de estos datos se pueden programar entonces algoritmos de búsqueda que se adaptan especialmente a las condiciones celulares que hay que identificar.

[0075] Además se divulga la aplicación del análisis transcriptómico, particularmente de perfiles transcriptómicos y/o de perfiles de expresión génica con origen en un transcriptoma o basados en un transcriptoma, para la evaluación pronóstica de la capacidad regenerativa tisular y/o la potencia celular y/o las perspectivas de éxito de una implantación, particularmente una implantación celular, y/o un trasplante. Para evitar repeticiones innecesarias, se hace referencia completamente a la descripción precedente con respecto a otras características y ventajas.

[0076] Otras características y ventajas de la invención resultan evidentes de las formas de realización descritas a continuación por medio de las figuras, descripciones de las figuras, un ejemplo, así como las reivindicaciones dependientes. Pueden realizarse a este respecto características individuales de la invención solas o en combinación entre sí. Las formas de realización descritas sirven únicamente para la aclaración y para una mejor comprensión de la invención y de ninguna manera deben entenderse restrictivamente.

## DESCRIPCIÓN DE LAS FIGURAS

[0077]

La figura 1 muestra valores de expresión determinados mediante qRT-PCR (quantitative real time PCR; PCR en tiempo real cuantitativa) de GAPDH (A) y seis genes marcadores seleccionados (B). El diagrama de caja muestra para cada gen el rango de 25%-75% (caja gris), el rango de 5%/95% (líneas horizontales por arriba y por debajo de la caja), la mediana (línea dentro de la caja) así como los valores extremos (puntos negros). A: los valores de expresión de GAPDH se normalizaron en 3 mediciones repetidas (respectivamente aplicando las 422 muestras de pacientes) sobre la expresión de ARNm total. Las diferencias entre los tres conjuntos de datos son estadísticamente no significativas (ANOVA simple), lo que indica la fiabilidad del método de qRT-PCR. B: la expresión de los seis genes marcadores se representa en proporción a GAPDH como valor de  $\Delta Ct$  negativo. Los datos se recogieron en el marco de un control de calidad rutinario en el momento de la recolección de los cultivos celulares monocapa directamente antes de la colonización del soporte (Novocart® Basic). El valor de expresión de ARNm obtenido para cada paciente se normalizó sobre el patrón de ADNc del paciente en cuestión (según las indicaciones del fabricante del kit de síntesis de ADNc), de modo que se garantice una comparación directa de los datos de expresión para cada uno de los seis genes seleccionados. COL1: COL1A2, cadena alfa 2 de colágeno tipo I; COL2: COL2A1, cadena alfa 1 de colágeno tipo IIb; AGG: ACAN, agregano; IL1: IL1- $\beta$ , interleucina-1 $\beta$ ; BSP2: sialoproteína ósea 2; FLT-1: receptor 1 del factor de crecimiento endotelial vascular.

La figura 2 muestra la distribución del nivel de expresión de diferentes familias de proteínas en condrocitos humanos cultivados. Los valores de expresión RPKM representados en escala logarítmica se recogieron para cada una de las 20 muestras en el marco de un análisis transcriptómico. El diagrama de caja muestra para cada gen un rango de 25%-75% (caja gris), la mediana (línea dentro de la caja) así como el área total (líneas horizontales por encima y por debajo de la caja). Los triángulos marcan genes a los que se les pueden asociar características positivas (p.ej., FGF-2) o anabólicas (p.ej., ACAN) con respecto al cartílago. Los rombos marcan genes a los que se les atribuyen características indeseadas (p.ej., colágeno I) o negativas o catabólicas (p.ej., interleucina-1, ADAM-TS5). La nomenclatura de los genes corresponde a la nomenclatura NCBI.

La figura 3 muestra el análisis de correlación entre la qRT-PCR y la secuenciación de ARN (NGS, Next Generation Sequencing) en base a los datos de expresión de COL1A2, COL2A1, ACAN e IL-1 $\beta$ . Los valores de expresión para COL1A2, COL2A1, ACAN e IL-1 $\beta$  se obtuvieron por un lado mediante la qRT-PCR convencional de 422 muestras (compárese con la tabla 1) y por otro lado mediante la secuenciación de ARN basada en NGS de 20 muestras. Los valores de  $\Delta Ct$  y los valores de expresión de RPKM se representan logarítmicamente. La imagen auxiliar muestra la aplicación de los valores de expresión de RPKM antes de su conversión logarítmica, donde su muy buena correlación es reconocible de manera más clara todavía.

La figura 4 muestra la distribución de los valores numéricos para las relaciones de expresión génica de los 3114 genes expresados con más fuerza en forma de un histograma. Los valores numéricos se obtuvieron de la proporción de los valores de RPKM promediados de cada gen transcrito del grupo con buenos resultados clínicos (S1 a S10) respecto al valor medio del grupo con fracasos del implante (S11 a S20). Los valores numéricos > 1 se correlacionan con buenos resultados clínicos. Los valores numéricos < 1 se correlacionan con fracasos del implante. Cada barra representa 1/100 del rango total determinado (valor numérico más pequeño: 0,006; valor numérico máximo: 4,9).

La figura 5 muestra el agrupamiento jerárquico por medio de datos de expresión de todo el genoma de las 20 muestras (S1 a S20). El agrupamiento se basa en una correlación de Pearson entre los valores de expresión de todos los genes de una muestra aplicando el algoritmo de Neighbor-Joining. Las muestras se muestran según el resultado clínico de la implantación en el paciente en cuestión como "positivo" o "negativo". Mientras que la variación entre las muestras negativas es considerablemente más alta que entre las muestras positivas, se puede registrar una separación clara entre los resultados clínicos subyacentes. Esta separación indica que los resultados clínicos -trasplante transcurrido positivo o negativo- pueden ser asociados a un fenotipo transcriptómico de las células extraídas para la implantación.

La figura 6 muestra dos análisis de mapa de calor A y B del análisis de conglomerados de correlación de Pearson de los valores numéricos de RPMK de las muestras S1 a S20. Los valores se dispusieron según su semejanza, donde las muestras de expresión estrechamente emparentadas entre sí están próximas entre sí. A: se consideraron los genes completos con todos sus exones. B: los exones se analizaron individual e independientemente de la estructura de ARNm. Excepto de S2 y S11, con esta evaluación se pudo lograr una separación de las muestras S1 a S20 según el transcurso clínico subyacente (S1-S10: transcurso positivo; S11-S20: fracaso del implante).

## EJEMPLO

### 1. Material y métodos

#### 1.1 Estructura del estudio

[0078] Se realizó una encuesta retrospectiva sobre primeros resultados, efectos secundarios y variaciones del estado inicial con respecto al dolor, función e inflamaciones en una población de pacientes como se define a continuación. Se realizaron otros análisis para examinar la influencia de los pacientes, la producción y las características celulares biológicas sobre la seguridad y los resultados de los pacientes. Los procedimientos clínicos y quirúrgicos, entre ellos las indicaciones y la rehabilitación se definieron en los procedimientos normalizados de trabajo (SOP). Los cirujanos fueron entrenados en las técnicas quirúrgicas para la obtención de cartílago y la cirugía de implantación antes de aplicar el implante por primera vez. Después de que los pacientes hubieran concedido su consentimiento informado, la indicación de tratamiento se confirmó a través de la artroscopia. En la articulación afectada se tomaron de la fosa intercondilar (de una zona no cargada) de dos a tres piezas osteocondrales aplicando un trépano estándar estéril y validado (Aesculap AG, Tuttlingen, Alemania, diámetro de punción: 4 mm).

#### 1.2 Recogida de datos clínicos

[0079] Se pidió al cirujano de cada paciente que rellenara un formulario de recogida de datos que incluyera la anamnesis familiar (etiología), los datos básicos demográficos (edad, sexo) y el período de tiempo entre la intervención quirúrgica y el último contacto con el paciente. Junto al paciente se valoraron la intensidad del dolor, la función y las inflamaciones sobre una escala de 10 puntos tanto durante la conversación antes de la intervención como también tras la intervención. La escala de valoración de los resultados se modificó a partir de la escala analógica visual (VAS). Los valores más altos indican mejores resultados (es decir, 10 significa "sin dolor", "sin inflamación", "sin limitación funcional"). Esta gradación de los resultados con 10 puntos se realizó por cada cirujano en una conversación con los pacientes y se usó como indicador precoz para el transcurso clínico posterior. También se les pidió a los cirujanos involucrados que indicasen todos los efectos secundarios aparecidos, así como un tratamiento necesario subsiguiente. El cuestionario no examinó si los pacientes respondieron insuficientemente a un método de reparación del cartílago artroscópico u otro método quirúrgico anterior. En caso extremo, el efecto indeseado pudo representar un fracaso del implante a través de la no curación o un desgarro.

#### 1.3 Población del estudio

[0080] La encuesta se realizó en sesenta y un centros en Alemania, en los cuales se trataron a 433 pacientes. Se reportaron datos para 422 pacientes (97,4 %) en total. Los demás 11 pacientes se excluyeron del estudio posterior, así como de los análisis de ARN. Entre los 422 pacientes había 140 mujeres y 282 hombres. Su edad media era

5 33,4 años (mínima: 14,7 años, máxima: 66,3 años). La pluralidad de los defectos del cartílago primarios (daños de grado III a IV según la clasificación de la International Cartilage Repair Society (ICRS)) se detectaron en el cóndilo femoral medial de la rodilla (68,5 %), seguido del cóndilo lateral (14,9 %), la zona retrorrotuliana (10,2 %), la tróclea (5 %) y la tibia (0,2 %). En seis pacientes (1,4 %) se detectó el defecto en el astrágalo. La pluralidad de los pacientes mostró un único defecto (93 %). El tamaño del defecto osciló entre 1 y 20 cm<sup>2</sup>. El tamaño intermedio fue 5,9 cm<sup>2</sup>. La duración media del seguimiento de los pacientes fue 6,9 meses. En total en 83 pacientes transcurrió al menos un año desde la intervención quirúrgica hasta el día de la última visita anotada por los cirujanos (tabla 1).

Tabla 1: población del estudio (422 pacientes)

		Número absoluto	Porcentaje
Sexo	Masculino	282	66,8%
	Femenino	140	33,2%
Edad	Valor medio	33,4	
	SD	10,0	
	Rango	14,7 - 66,3	
Duración del seguimiento de los pacientes en días	Valor medio	210	
	SD	185	
	Rango	0 - 921	
Diagnóstico	Daño crónico	24	5,7%
	Defecto degenerativo	144	34,1%
	Osteocondrosis disecante	123	29,1%
	Defecto traumático	137	32,5%
Localización del defecto primario	Cóndilo femoral lateral	63	14,9%
	Cóndilo femoral medial	289	68,5%
	Rótula	43	10,2%
	Astrágalo	6	1,4%
	Tibia	1	0,2%
	Tróclea	21	5,0%
Tamaño del defecto primario (cm <sup>2</sup> )	Valor medio	5,9	
	SD	3,1	
	Rango	1 - 20	
Número de defectos tratados	1	391	92,7%
	2	29	6,9%
	Datos ausentes	2	0,5%

10 [0081] La tabla 1 muestra la población del estudio con las características de los pacientes correspondientes. Las características se asocian a números de pacientes absolutos y, si es aplicable, a proporciones porcentuales. En un paciente fue posible un diagnóstico múltiple, donde los defectos afectaban a más de una región.

#### 1.4 Cultivo

15 [0082] Después de la eliminación de huesos, de cartílago mineralizado y superficial se aislaron condrocitos del cartílago restante por extracción mecánica y enzimática, ajustada según procedimientos estándares y se propagaron *in vitro* como cultivo primario en un dispositivo aprobado según buenas prácticas de fabricación (GMP) (TETEC AG, Reutlingen, Alemania). Los pasos incluían, tras el aislamiento celular, una proliferación celular en medio de cultivo, que estaba enriquecido con suero autólogo, la recolección a través de una tripsinización y la colonización en el andamio (Novocart Basic), lo que se llevó a cabo 21 días tras la artroscopia para la obtención de cartílago. El andamio fue colonizado con una dosis de células media de  $1,36 \times 10^6$  células por cm<sup>2</sup>. Típicamente, 20 las células se aplicaron directamente del cultivo primario. En circunstancias especiales (p. ej., enfermedad de los

pacientes) las células se crioconservaron antes de una aplicación y se administraron a partir de cultivos secundarios (menos de un 10 % de los casos).

#### 1.5 Secuenciación de ARN de alto rendimiento

5 [0083] Se analizaron 20 muestras de ARN de las cargas descritas en la sección anterior, de las cuales diez vinieron de tratamientos clínicos exitosos y diez se obtuvieron de pacientes con resultados negativos. Cada muestra se asoció a una secuencia de código de barras unívoca y alícuotas de las muestras provistas con código de barras se combinaron en dos grupos según el resultado de los pacientes (resultados positivos, resultados negativos).

10 [0084] De cada muestra se usó de 1 a 10 µg del ARN total de condrocitos como material inicial. El ARN total se analizó con un Agilent RNA 6000 Nano Chip (5067-1511) en un aparato del tipo Bioanalyzer 2100 con respecto a la cantidad y la integridad. Se realizó un empobrecimiento o depleción (ARNr) de ARN ribosómico aplicando 2-3 pasos con el RiboMinusTMEukaryote Kit for RNA-Seq (A10837-08, Life Technologies) según el protocolo del fabricante. Se realizó una concentración con el RiboMinus Concentration Module (K1550-05, Life Technologies). La eficiencia de la eliminación del ARNr se examinó y se determinó en un paso de una alícuota en un Agilent RNA 6000 Nano Chip (5067-1511) en un aparato del tipo Bioanalyzer 2100, donde se mostró una eliminación completa de los picos de ARNr.

15

20 [0085] Se elaboró una genoteca aplicando el SOLiDTM Total RNA-Seq Kit (4445373, Life Technologies) según el protocolo del fabricante. En resumen, de 10 a 100 ng de ARN después de la depleción de ribosomas se fragmentaron en primer lugar mediante una digestión de 10 minutos de duración a 37 °C con RNasa III y se depuraron de nuevo con el RiboMinus Concentration Module (K1550-05, Life Technologies) y se eluyeron en 12 µl de agua libre de RNasa. El tamaño de los fragmentos se examinó en el rango de tamaño óptimo para el aparato SOLiD4 (de 150 a 200 nucleótidos), donde a su vez se realizó un paso de una alícuota en un Agilent RNA 6000 Nano Chip (5067-1511) en un aparato del tipo Bioanalyzer 2100.

25 [0086] Para la estructura de genotecas de ADNc, los fragmentos de ARN se enlazaron de manera específica de cadena con adaptadores y se transportaron en genotecas de ADNc bicatenario, donde se aplicaron una transcripción inversa seguida de una selección de tamaños con PAGE y una amplificación a través de la PCR. Durante la amplificación por PCR se incluyeron secuencias de código de barras, donde el cebador adaptador 3' llevaba un extremo protuberante específico de códigos de barra específico. Este método evita distorsiones de los códigos de barras, que se reportaron cuando los adaptadores de códigos de barras se enlazan directamente con la genoteca. El tamaño de los fragmentos se examinó de nuevo en el rango de 150 a 200 nucleótidos en un paso de una alícuota en un Agilent DNA 1000 Chip (5067-1504) en un aparato del tipo Bioanalyzer 2100.

30

35 [0087] Las mediciones de concentración del bioanizador se usaron para el cálculo y la mezcla equimolar de las muestras con código de barras S1-S10 (buenos resultados clínicos) y S11-S20 (resultados clínicos negativos) (respectivamente, los códigos de barra 1-10). Se realizó respectivamente para cada genoteca combinada una PCR en emulsión (emPCR) en una escala E20 según las recomendaciones del fabricante con una concentración final de la genoteca de 0,5 pM. La rotura de la emulsión y el lavado de esferas, así como el enriquecimiento de esferas se realizaron a mano según el protocolo del fabricante. La calidad y la cantidad de las esferas molde se midieron en un paso de análisis del flujo de trabajo (WFA) de 15 millones de esferas molde enriquecidas en un aparato SOLiD4.0. La relación señal a ruido fue en ambas genotecas de un 4 % y la proporción P2/P1 fue del 100%, lo que indica un porcentaje alto de esferas monoclonales. Se aplicaron respectivamente 158 millones de esferas molde de las genotecas en un campo Quad (1/4) de un soporte SOLiD-4,0 y se secuenciaron a través del ligamiento del adaptador P1, de modo que se obtuvo un registro de datos de lectura de fragmentos de 50 pb.

40

#### 1.6 Preparación de datos

45 [0088] Se recopilaron datos de los efectos secundarios y resultados de los pacientes y se registraron en una hoja de cálculo Excel. Se registraron datos de control de calidad de la producción, así como de covariables (características demográficas, de pacientes y de liberación de producto) en la misma base de datos. Se examinó la conformidad de todos los datos introducidos.

#### 1.7 Estadística y análisis covariante de datos de parámetros individuales

50 [0089] Se utilizó estadística sumaria con valores medios, desviaciones estándar, rangos e intervalos de confianza para mostrar las sumas de los efectos secundarios, los resultados de eficacia y las características de los pacientes. Las características de liberación de cada producto se compararon con los datos de las encuestas, para identificar asociaciones entre las características de la producción de los implantes y los resultados de los pacientes. Se elaboró una serie de pruebas de chi cuadrado y modelos de regresión para comprobar las asociaciones entre las

características de liberación o de producción y las cinco mediciones de resultados en la encuesta de seguridad de los pacientes incluían: efectos secundarios relativos a los implantes (definidos como la combinación de fracasos del implante, el desprendimiento, la hipertrofia, la artrofibrosis, las adhesiones, la condromalacia, las infecciones de las articulaciones y la aparición de cuerpos articulares libres), reoperaciones (cualquier motivo) y variaciones del estado inicial con respecto a los dolores, la función y las inflamaciones. Las regresiones para el modelado de la asociación entre los biomarcadores y un resultado interesante se ajustaron simultáneamente tanto para cada biomarcador y para todos los biomarcadores. Se aplicó un modelo de regresión logística para determinar la influencia de las covariantes independientes sobre la aparición de efectos secundarios relativos a los implantes y la aparición de reoperaciones. Se aplicaron modelos lineales para modelar las variaciones del estado inicial con respecto a los dolores, la función y las inflamaciones. Los valores de p se consideraron significativos a 0,05. Todos los modelos de regresión incluían el período de tiempo de la intervención hasta la última cita del médico con los pacientes, puesto que se supone que este es un valor predictivo independiente significativo para los resultados.

### 1.8 Análisis estadístico de todos los datos transcriptómicos

[0090] Se representaron secuencias de lectura con el *software* Whole Transcriptome Analysis Pipeline der BioScope 1.32 (Applied Biosystems) sobre el genoma humano (UCSC versión HG 19). Partiendo de las observaciones del UCSC refSeq (descargado el 7 de noviembre de 2010) se calcularon los valores de RPKM de la expresión para los genes. En resumen, los valores de RPKM (Reads Per Kilobase exon model per Million mapped) representan el número de segmentos secuenciados que se representan sobre los exones de una transcripción dada, normalizados por la profundidad de secuenciación por muestra (número total de segmentos) y la longitud (pb) de todos los exones. Estos valores se usaron para la determinación inicial. El método de normalización de la escala descrito por Bullard et al. (J.H. Bullard, E. Purdom, K.D. Hansen, S. Dudoit, Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments. BMC Bioinformatics 11: 94 (2010)) se utilizó para eliminar distorsiones técnicas específicas de las muestras y los valores de expresión obtenidos se representaron sobre una escala logarítmica. Las muestras se agruparon con el algoritmo de Neighbor-Joining (N. Saitou and M. Nei, The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4 (4): 406-425 (1987)), que se basa en la distancia de correlación de Pearson entre todos sus perfiles de expresión del transcriptoma. El método de Rank-Product no paramétrico (R. Breitling, P. Armengaud, A. Amtmann, P. Herzyk, Rank products: a simple, yet powerful, new method to detect differentially regulated genes. FEBS Letters 573 (1): 83-92 (2004)) se utilizó para comprobar la expresión diferencial, de la que se obtienen valores de pfp (porcentaje de valores falsos positivos, una medida comparable con valores de p corregidos por FDR) y los transcritos con pfp < 0,05 y un valor absoluto del cambio múltiplo mayor de 0,9 (es decir, expresión duplicada o dividida por la mitad) se consideraron una expresión diferencial significativa. Se usaron los transcritos codificantes de proteínas expresados diferencialmente para comprobar un enriquecimiento de categorías funcionales específicas (análisis de sobrerrepresentación, prueba hipergeométrica, valores de p corregidos para la tasa de errores de clasificación). Para todos los análisis se usó la técnica descrita por Mayday (F. Battke, S. Symons, K. Nieselt, Mayday -Integrative analytics for expression data. BMC Bioinformatics 11 (1): 121 (2010)). Se examinaron en detalle los segmentos de genes individuales de interés con el Integrative Genomics Viewer IGV (J.T. Robinson, H. Thorvaldsdottir, W. Winckler, M. Se Gutt, E.S. Lander, G. Getz, J.P. Mesirov, Integrative Genomics Viewer. Nat. Biotech. 29: 24-26 (2011)).

## 2. Resultados

### 2.1 Variaciones del estado inicial en los resultados de los pacientes

[0091] Antes de una intervención quirúrgica y en las sucesivas visitas a los pacientes, los médicos determinaron los dolores articulares, la función y las inflamaciones sobre una escala de 10 puntos, donde los valores más altos indican mejores resultados. Los pacientes que dieron respuestas tanto a las preguntas antes de la intervención como también después de la intervención lograron mejoras significativas frente al estado inicial ( $p < 0,0001$ , prueba de los rangos con signo de Wilcoxon) con las tres mediciones. La duración media desde la intervención quirúrgica fue 6,9 meses (rango: de 2 a 30 meses), lo que es un horizonte temporal notablemente más corto que los indicados en la mayoría de otros resúmenes de resultados ACI. El estudio retrospectivo aportó aspectos esenciales para el conocimiento actual relativo a la reparación de cartílago articular. En el subgrupo de 83 pacientes con al menos 12 meses desde su intervención quirúrgica se reportó que presentaron variaciones más fuertes de media en las tres mediciones de resultados que la población de pacientes total (tabla 2).

Tabla 2: medición de resultados individual

Medición de resultados	Todos los pacientes (N=422)					Período de análisis >1 año (N=83)				
	Mín.	Media	Máx.	SD	N	Mín.	Media	Mín.	SD	N
Dolor preoperatorio	1	3,4	10	1,3	412	2	3,4	9	1,2	82

Medición de resultados	Todos los pacientes (N=422)					Período de análisis >1 año (N=83)					
	Mín.	Media	Máx.	SD	N	Mín.	Media	Mín.	SD	N	
postoperatorio	1	7,0	10	2,0	382	2	7,2	10	1,9	80	
	Cambio	-5	3,8	8	2,1	376	2	3,9	8	2,1	80
Inflamación	preoperatoria	1	5,3	10	2,4	411	1	5,1	10	2,4	82
	postoperatoria	1	7,7	10	1,9	382	4	8,1	10	1,6	80
	Cambio	-7	2,5	9	2,6	375	-3	3,1	8	2,2	80
Función	preoperatoria	1	4,2	10	1,8	412	2	4,4	10	2,0	82
	postoperatoria	2	7,3	10	1,8	380	2	7,8	10	1,6	80
	Cambio	-7	3,2	8	2,3	374	-7	3,6	8	2,4	80

## 2.2 Apariciones de efectos secundarios

[0092] La tabla 3 muestra la aparición de efectos secundarios reportados. La aparición del fracaso del implante fue de un 3,1 % en la población de pacientes total y de un 6 % en el subgrupo de pacientes en los que al menos habían transcurrido 12 meses desde su intervención. Los números de casos de complicaciones relativas a los implantes reportados fueron generalmente bajos en la población de pacientes total. Se observaron desprendimientos (delaminaciones), artrofibrosis e hipertrofia en un 1,7%, un 2,4 % o un 0,7 %. En total 36 pacientes (8,5 %) requirieron una reoperación y/o corrección. Los efectos secundarios más frecuentes reportados en los pacientes que requirieron reoperaciones fueron fracasos del implante (13, las diez muestras examinadas adicionalmente mediante el análisis transcriptómico procedían de estos casos), desprendimiento (6), artrofibrosis (7), sinovitis (7), adhesiones (5) y dolores (6). El subgrupo de 83 pacientes en los que habían transcurrido al menos 12 meses desde su intervención presentaba una tasa de reoperaciones de un 13,3 %. La pluralidad de las reoperaciones se realizó de modo artroscópico.

Tabla 3: complicaciones reportadas

Complicación	Todos los pacientes (N=422)				Período de análisis >1 año (N=83)			
	Casos	%	95%	IC**	Casos	%	95%	IC**
complicaciones relativas a los implantes***								
Fracasos del implante	13	3,1%	1,7%	5,2%	5	6,0%	2,0%	13,5%
Delaminación	7	1,7%	0,7%	3,4%	2	2,4%	0,3%	8,4%
Hipertrofia	3	0,7%	0,2%	2,1%	1	1,2%	0,0%	6,5%
Artrofibrosis	10	2,4%	1,1%	4,3%	2	2,4%	0,3%	8,4%
Adherencias	7	1,7%	0,7%	3,4%	2	2,4%	0,3%	8,4%
Cuerpos articulares libres	1	0,2%	0,0%	1,3%	1	1,2%	0,0%	6,5%
Infección (articular) profunda	3	0,7%	0,2%	2,1%	0	-	-	-
Condromalacia	2	0,5%	0%	1,7%	0	-	-	-
Otras complicaciones								
Efusión	32	7,6%	5,2%	10,5%	6	7,2%	2,7%	15,1%
Dolor	29	6,9%	4,7%	9,7%	7	8,4%	3,5%	16,6%
Sinovitis	14	3,3%	1,8%	5,5%	3	3,6%	0,8%	10,2%
Hematoma/hemartrosis	7	1,7%	0,7%	3,4%	5	6,0%	2,0%	13,5%
Rigidez	1	0,2%	0,0%	1,3%	1	1,2%	0,0%	6,5%
Infección superficial	2	0,5%	0,0%	1,7%	0	-	-	-

## 2.3 Asociación entre resultados y factores de riesgo independientes

5 [0093] El tamaño de todos los defectos tratados no mostró correlación con uno de los resultados de los pacientes medidos. Para el cambio del estado inicial en la función de los pacientes se halló una asociación significativa con la edad del paciente (es decir, a menores edades se mostraron de media mayores mejoras,  $p = 0,004$ , no mostrado). La tabla 4 indica los resultados de las pruebas de chi cuadrado para la asociación de la posición del defecto primario sobre la aparición de efectos secundarios relativos a los implantes. Se consideró que tanto la aparición de efectos secundarios relativos a los implantes como de las reoperaciones es en gran parte independiente de la posición del defecto primario, con excepción de los defectos rotulianos. En todo caso, la probabilidad de un tal acontecimiento para un afectado fue mayor cuando el defecto primario se encontraba en la rótula ( $p < 0,0001$ ).

Complicaciones* relativas a los implantes en función de la localización del defecto							
Localización	Defecto en esta posición			Defecto no en esta posición			Valor de p
	Casos	N	%	Casos	N	%	
Cóndilo femoral medial	22	289	7,6%	14	133	10,5%	0,32
Cóndilo femoral lateral	3	63	4,8%	33	359	9,2%	0,25
Tróclea	1	21	4,8%	35	401	8,7%	0,53
Rótula	11	43	25,6%	25	379	6,6%	<0,0001
Tibia	0	1	0,0%	36	421	8,6%	0,76
Astrágalo	0	6	0,0%	36	416	8,7%	0,45

Reoperaciones relativas a los implantes en función de la localización del defecto**							
Localización	Defecto en esta posición			Defecto no en esta posición			Valor de p
	Casos	N	%	Casos	N	%	
Cóndilo femoral medial	20	289	6,9%	16	133	12,0%	0,08
Cóndilo femoral lateral	3	63	4,8%	33	359	9,2%	0,25
Tróclea	1	21	4,8%	35	401	8,7%	0,53
Rótula	12	43	27,9%	24	379	6,3%	<0,0001
Tibia	0	1	0,0%	36	421	8,6%	0,76
Astrágalo	0	6	0,0%	36	416	8,7%	0,45

10 Tabla 4: reoperaciones y complicaciones relativas al implante en función de la localización del defecto

15 [0094] La tabla 5 indica los resultados de las pruebas de chi cuadrado para la asociación entre la clasificación de lesiones del cartílago y la aparición de efectos secundarios relativos a los implantes y reoperaciones. En caso de efectos secundarios relativos a los implantes, la probabilidad de un tal acontecimiento para un afectado fue mayor cuando el defecto del cartílago se clasificó como degenerativo ( $p = 0,005$ ). Por el contrario, la probabilidad de un tal acontecimiento para los pacientes fue menor cuando existía un defecto del cartílago debido a la osteocondritis disecante ( $p = 0,04$ ).

Complicaciones* relativas a los implantes en función del tipo de defecto							
Tipo de defecto	Defecto de este tipo			Defecto no de este tipo			Valor de p
	Casos	N	%	Casos	N	%	
Daño crónico	1	24	4,2%	35	398	8,8%	0,43
Defecto degenerativo	20	144	13,9%	16	278	5,8%	0,005
Defecto traumático	11	137	8,0%	25	285	8,8%	0,80
Osteocondrosis disecante	5	123	4,1%	31	299	10,4%	0,04

Reoperaciones relativas a los implantes en función del tipo de defecto							
Tipo de defecto	Defecto en esta posición			Defecto no en esta posición			Valor de p
	Casos	N	%	Casos	N	%	
Daño crónico	1	24	4,2%	35	398	8,8%	0,43
Defecto degenerativo	17	144	11,8%	19	278	6,8%	0,08
Defecto traumático	13	137	9,5%	23	285	8,1%	0,63
Osteocondrosis disecante	5	123	4,1%	31	299	10,4%	0,04

Tabla 5: reoperaciones y complicaciones relativas al implante en función del tipo de defecto

[0095] La tabla 6 muestra asociaciones entre los resultados de los pacientes y las variables de los cultivos celulares por un lado y, por otro lado, los valores de expresión de ARNm de seis genes marcadores seleccionados (compárese con el punto 2.4.1) como valores de p multivariantes. Se destacan las relaciones significativas y las tendencias limítrofes. Una crioconservación temporal de las células antes de una implantación no mostró una asociación con una mayor aparición de complicaciones relativas a los implantes (CRI) o con otra de las mediciones de resultados (reoperación RO, dolor, función, inflamación). Respecto a la viabilidad celular se observaron los mismos resultados. Sin embargo, un número pequeño de células administradas estaba asociado significativamente con una reoperación (RO) ( $p = 0,06$ ; tabla 6).

Tabla 6: relación entre los resultados de los pacientes y las variables de los cultivos celulares o valores de expresión de ARNm de seis genes marcadores seleccionados

Variables de los cultivos celulares	CRI	RO	Dolor	Función	Inflamación
Crioconservación	0,98	0,98	0,59	0,78	0,13
Viabilidad celular en la recolección	0,56	0,13	0,55	0,38	0,76
Número de células (log 10)	0,30	0,06	0,47	0,37	0,36
Medición de la expresión por PCR					
Agrecano	0,14	0,44	0,61	0,50	0,62
BSP-2	0,11	0,14	0,76	0,78	0,90
Colágeno I	0,49	0,70	0,53	0,79	0,27
Colágeno II	0,08	0,08	0,42	0,90	0,29
Interleucina-1 $\beta$	0,08	0,20	0,26	0,19	0,24
FLT-1	0,02	0,03	0,21	0,67	0,81

## 2.4 Datos de expresión

### 2.4.1 Expresión génica específica

[0096] Mediante una PCR en tiempo real cuantitativa (qRT-PCR) se determinó la expresión de seis genes marcadores seleccionados: agrecano (ACAN), un componente integral de la matriz de cartílago extracelular; colágeno tipo I (COL1A2), que no ocurre habitualmente en el tejido cartilaginoso; colágeno tipo II específico del cartílago (COL2A1); interleucina-1 $\beta$  (IL-1 $\beta$ ), una citoquina inflamatoria; FLT-1, una isoforma de un receptor del factor de crecimiento endotelial vascular; y sialoproteína ósea BSP-2, un factor de crecimiento óseo. Los dos últimos mencionados se seleccionaron porque las complicaciones pueden resultar particularmente de una vascularización del tejido cartilaginoso implantado o de una formación de osteofitos dentro del tejido cartilaginoso implantado.

[0097] Los valores de expresión (dCt) de estos seis genes marcadores se representan en la figura 1 y muestran diferencias significativas. Los niveles de ARNm más estables, que varían en el rango de un orden de magnitud entre las 422 muestras, se hallaron para el agrecano. Las variaciones máximas, en el rango de tres órdenes de magnitud, con un nivel de detección general muy bajo, se encontraron con FLT-1. Para el colágeno tipo I, por ejemplo, pudieron calcularse por término medio 1000 transcritos de ARNm por célula, mientras que, en el caso de FLT-1, se calcularon únicamente 10 transcritos de ARNm por 1000 células. Para comparar, las células endoteliales venosas humanas cultivadas *in vitro* contienen, por ejemplo, al menos 500 veces más transcritos de ARNm de FLT-1.

[0098] No obstante, a pesar de estas variaciones se pudieron observar correlaciones significativas o al menos limítrofes entre algunos de estos genes marcadores y el resultado clínico (tablas 6 y 7). Así pudo detectarse una asociación significativa entre complicaciones relativas a los implantes (CRI) y los niveles de expresión aumentados de FLT-1 en condrocitos propagados ( $p = 0,02$ , tabla 6). Por el contrario, solo pudo determinarse una relación limítrofe para los niveles aumentados de IL-1 $\beta$  y los niveles de colágeno tipo II reducidos (ambos  $p = 0,08$ , tabla 6). Considerando todos los genes marcadores en los modelos matemáticos pudo hallarse sin embargo una relación significativa o limítrofe en relación con las CRI para IL-1 $\beta$  ( $p = 0,02$ , tabla 6) y FLT-1 ( $p = 0,08$ , tabla 6).

[0099] Según la tabla 7, que muestra para los seis genes marcadores además de los coeficientes de regresión los valores de p multivariantes considerando el tiempo transcurrido desde la intervención, no pudo clasificarse como significativa la expresión de ninguno de los genes marcadores en relación con el dolor, la función y la inflamación.

Los coeficientes para los resultados de los pacientes CRI y RO (reoperación) se calcularon mediante regresión logística, los coeficientes para el dolor, la función y la inflamación mediante regresión lineal. Los valores de p (segunda fila) indican la significación de la relación en cuestión, donde se destacan las relaciones significativas y las tendencias limítrofes.

5 Tabla 7: análisis multivariable de la relación entre los resultados de los pacientes y el tiempo desde la intervención o los valores de expresión de ARNm de seis genes marcadores seleccionados.

Variable	CRI	RO	Dolor	Función	Inflamación
Tiempo desde la intervención	1,057	1,064	0,0084	0,055	0,060
	<i>0,037</i>	<i>0,017</i>	<i>0,644</i>	<i>0,0049</i>	<i>0,0059</i>
Medición de la expresión por PCR					
Agregano	1,141	0,962	0,105	-0,099	0,087
p=	<i>0,492</i>	<i>0,836</i>	<i>0,343</i>	<i>0,397</i>	<i>0,506</i>
BSP-2	1,086	1,078	-0,018	-0,013	-0,0085
p=	<i>0,161</i>	<i>0,177</i>	<i>0,607</i>	<i>0,730</i>	<i>0,837</i>
Colágeno I	0,901	0,973	0,015	0,049	0,057
p=	<i>0,426</i>	<i>0,835</i>	<i>0,849</i>	<i>0,564</i>	<i>0,556</i>
Colágeno II	1,095	1,112	-0,034	0,015	-0,054
p=	<i>0,166</i>	<i>0,107</i>	<i>0,347</i>	<i>0,698</i>	<i>0,216</i>
Interleucina-1 $\beta$	1,151	1,112	-0,044	-0,049	-0,054
p=	<i>0,029</i>	<i>0,083</i>	<i>0,197</i>	<i>0,177</i>	<i>0,179</i>
FLT-1	1,106	1,102	-0,033	-0,0096	-0,0041
p=	<i>0,062</i>	<i>0,070</i>	<i>0,222</i>	<i>0,739</i>	<i>0,901</i>

#### 2.4.2 Expresión de todo el genoma

10 [0100] La totalidad de los datos de expresión recogidos incluyeron aproximadamente 30.000 genes con una gran variedad de datos de exones y dio más de 200 millones de segmentos de datos individuales. Para los grupos S1-10 (buenos resultados clínicos) se representaron 42.716.671 segmentos (75,3 %) de modo unívoco, 3.828.740 (6,7 %) eran uniones que abarcan el exón, 3.791.685 de las uniones del exón eran conocidas y 37.055 eran nuevas. Para el grupo S11-20 (fracaso del implante) se representaron en total 43.429.958 segmentos (76,0 %) de modo unívoco, 3.991.606 segmentos (7,0 %) eran uniones que abarcan el exón, 3.958.041 de las uniones del exón eran conocidas y 33.565 eran nuevas. Los segmentos representados se usaron con Bioscope 1.3 para el cálculo de valores de RPKM (valores de Reads Per Kilobase per Million mapped reads) para cada exón contenido en la base de datos de refSeq UCSC (descargada el 07.11.2010). Una primera encuesta de estos RPKM se realizó aplicando familias génicas, donde miembros de las familias se han considerado "típicos" de los condrocitos, por ejemplo, colágeno (con colágeno de los tipos II, IX y XI como "colágeno de cartílago") o proteoglicano (con el agregano (ACAN) como el representante típico) (Fig. 2). Se representan gráficamente los datos, donde las desviaciones estándar indican la variabilidad de los exones. Además, se extendieron los valores de expresión sobre órdenes de magnitud y solo pudieron representarse en escala logarítmica, como en la figura 2. Los resultados ya muestran que el fenotipo "condrocítico" clásico no domina los resultados. Otro análisis de otras familias génicas, entre ellas, por ejemplo, genes relacionados con IGF o con PDGF/VEGF, confirma esta suposición. Una expresión de la citocina interleucina estaba diversificada. La expresión de estos factores de crecimiento fue más compleja todavía cuando se usaron proteínas de unión complementarias o los receptores correspondientes. La transcripción del extenso grupo de las metoloproteinasas es igualmente muy amplia (Fig. 2).

#### 2.5 Calibración de datos transcriptómicos por qRT-PCR

30 [0101] Para evaluar la comparación de datos con los dos métodos, se realizó un análisis de correlación basado en los datos de expresión de COL1A2, COL2A1, ACAN e IL-1 $\beta$  (compárese con la Fig. 1). Estos genes muestran valores de expresión típicos en los condrocitos en el rango de siete potencias decimales. La correlación de los valores de los dos métodos fue exitosa y dio un coeficiente de correlación de 1,0 en una función exponencial (Fig. 3).

## 2.6 Correlación entre datos transcriptómicos y resultados clínicos

- 5 [0102] Para recibir informaciones sobre muestras específicas que están (hipotéticamente) relacionadas con los resultados clínicos, se eligieron dos métodos de evaluación. En un primer planteamiento se determinó la relación de expresión para el valor de RPKM medio de cada gen transcrito del grupo con buenos resultados clínicos (S1 a S10) respecto al valor medio del grupo con fracasos del implante (S11 a S20). Así se obtuvo una lista de valores numéricos  $> 1$ , que presentan una correlación positiva con buenos resultados, de los valores numéricos  $< 1$ , que presentan una correlación positiva con fracasos del implante. Los valores numéricos próximos a 1 señalan genes que se comportan de forma neutral respecto a los resultados. Los resultados se representaron como un histograma en la figura 4. El histograma contiene valores para los 3114 genes expresados con mayor fuerza. 1803 valores son  $> 1$ , 1311 valores son  $< 1$ . Entre ellos, 62 valores son  $> 2$ , lo que significa una expresión dos veces más fuerte en el grupo con buenos resultados S1-10, y 70 valores son  $< 0,5$ , lo que significa una expresión dos veces más fuerte en el grupo con fracasos del implante. En otras palabras, estos dos grupos de genes representan candidatos para una lista delimitada de genes de parámetros de calidad para estudios posteriores. Esta lista no contiene ARN reguladores.
- 10
- 15 [0103] En un segundo planteamiento se realizó un análisis de conglomerados de correlación de los valores numéricos de RPKM: el coeficiente de correlación de Pearson se calculó para cada pareja de muestras (S1-S20) en los genes y se aplicó un denominado "mapa de calor" de la matriz de correlación con la función de mapa de calor incorporada. La hipótesis detrás de esta correlación supone que el cálculo da una lista clasificada para los 20 resultados clínicos que separa los resultados buenos de los casos de fracasos clínicos, cuando la relación esbozada en los datos enumerados arriba es significativa estadísticamente. El resultado se representa en la figura 5. En la figura 6A se relacionaron los datos para los genes completos con todos sus exones. En la figura 6B se analizaron los exones individual e independientemente de la estructura de ARNm. En otras palabras, en la figura 6A se desprecia el empalme alternativo, mientras que se representa en la figura 6B. Esto destacó que los datos se clasificaron correspondientemente a su origen, donde se separó el grupo S1-10 del grupo S11-20. Son excepciones la muestra S2, que se clasificó en la zona de transición entre los dos grupos cuando se analizaron los genes y migró posteriormente al grupo de fracasos cuando se analizaron los exones, y la muestra S11, que permaneció estable en el grupo "falso". El mapa de calor dio en total una "predicción" correcta del resultado clínico en 18 de 20 casos o con una probabilidad de un 90 %.
- 20
- 25

**REIVINDICACIONES**

1. Método *in vitro* para la evaluación pronóstica de las perspectivas de éxito de una implantación y/o un trasplante, **caracterizado por el hecho de que** se determinan en condrocitos autólogos los valores de expresión de ARNm de un grupo consistente en COL2A1, IL-1 $\beta$ , FLT-1, BSP-2, COL1A2 y ACAN.
- 5 2. Método *in vitro* según la reivindicación 1, **caracterizado por el hecho de que** el ARNm se determina mediante PCR en tiempo real.
3. Método *in vitro* según la reivindicación 1 o 2, **caracterizado por el hecho de que** los condrocitos son condrocitos articulares.
- 10 4. Método *in vitro* según cualquiera de las reivindicaciones anteriores, **caracterizado por el hecho de que** la implantación es una implantación de condrocitos autólogos asistida por soporte.
5. Método *in vitro* según cualquiera de las reivindicaciones anteriores, **caracterizado por el hecho de que** el trasplante es un trasplante de condrocitos autólogos.

Figura 1

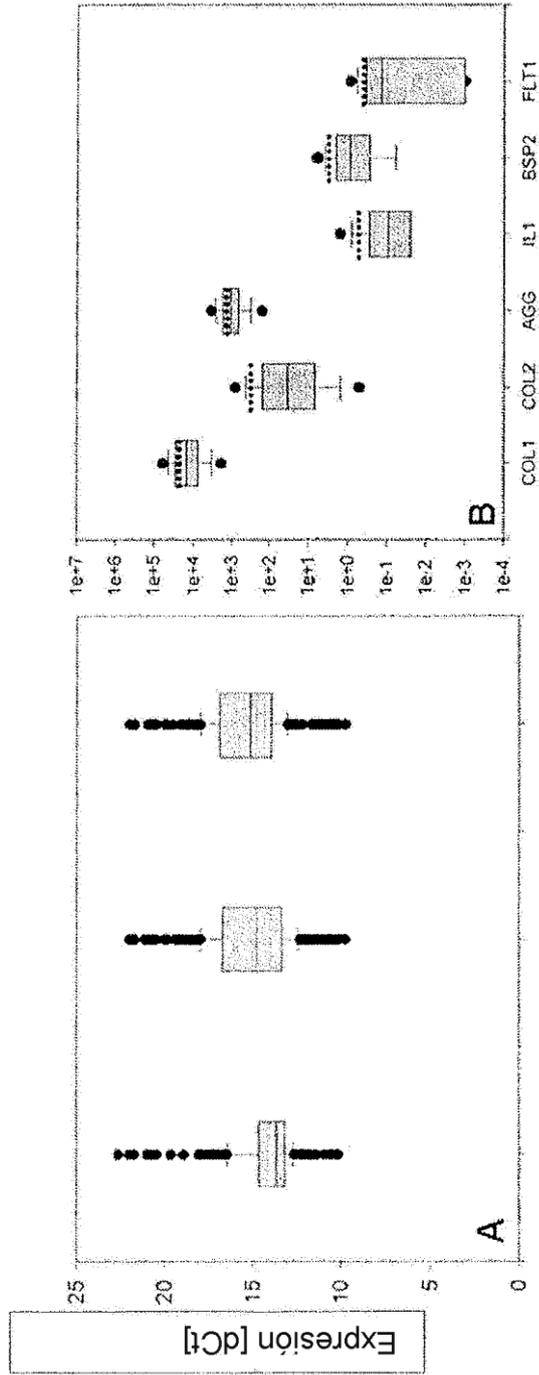
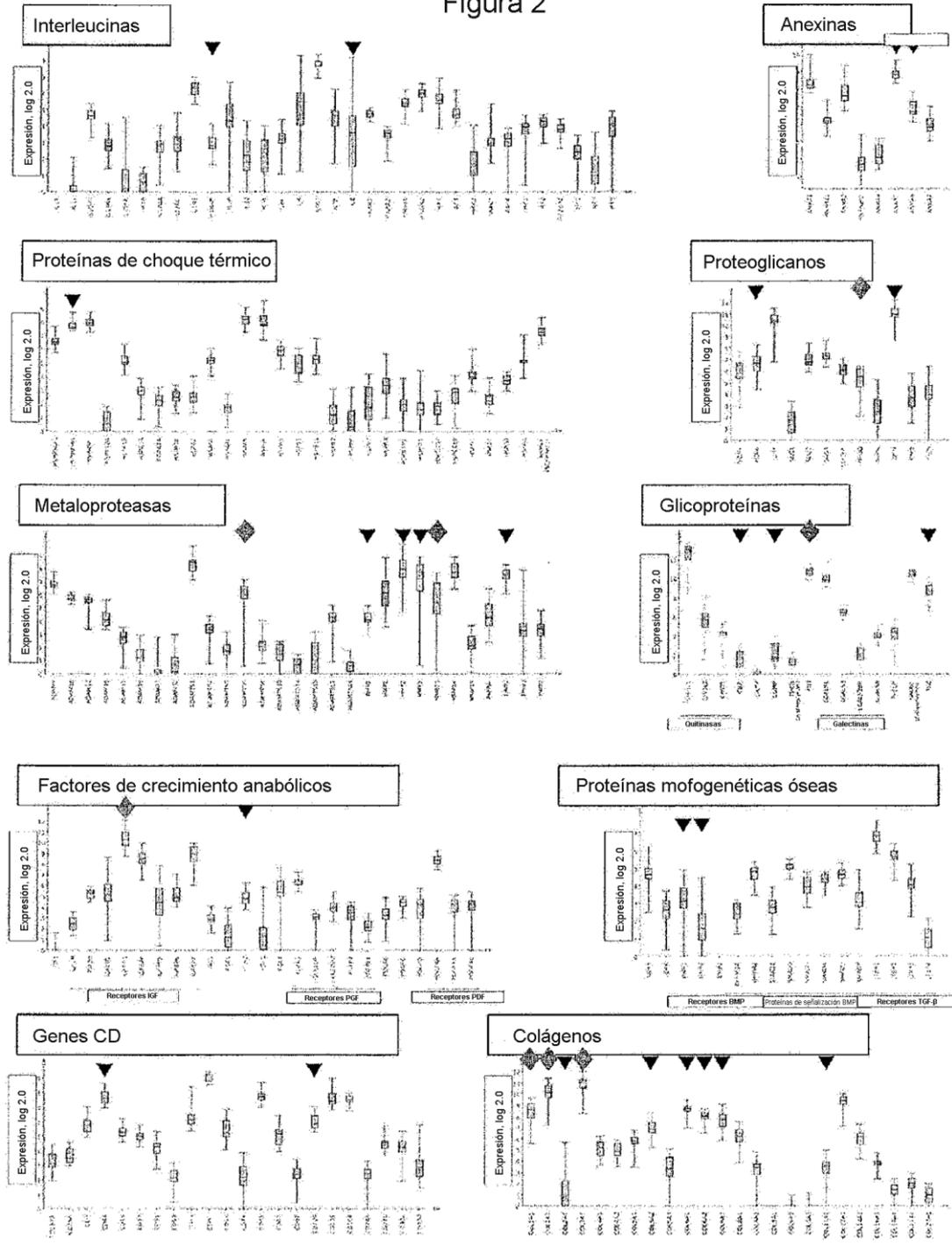


Figura 2



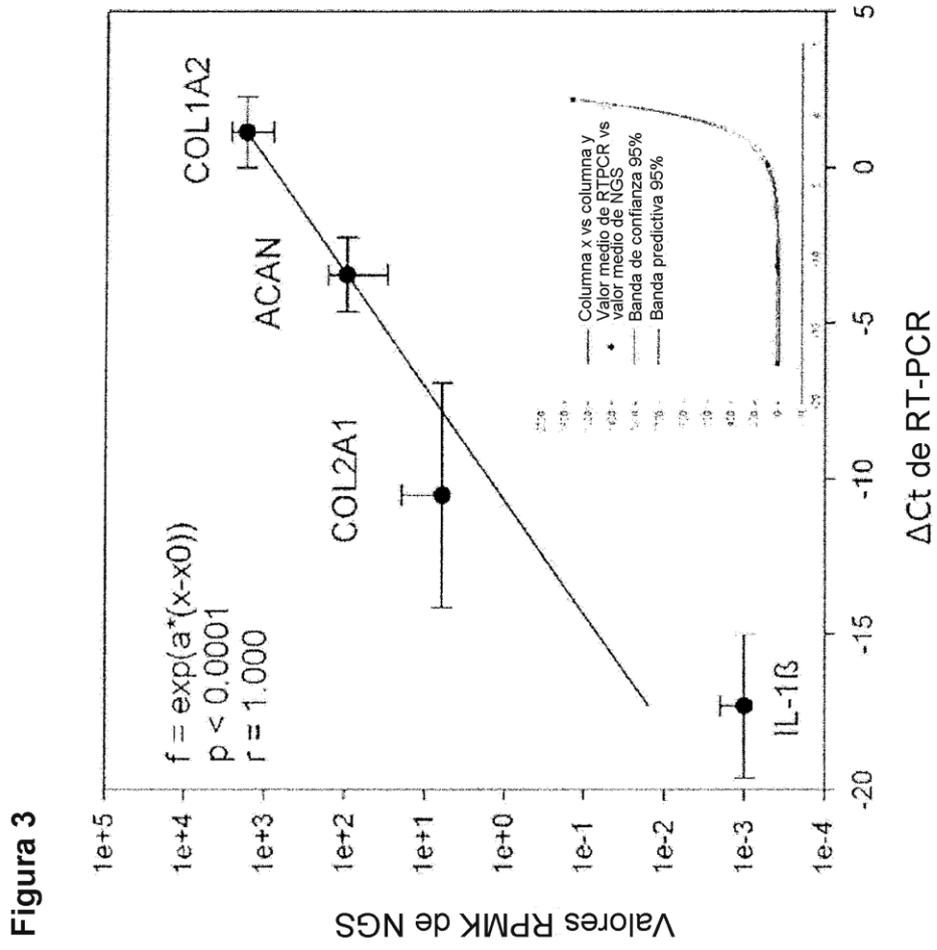
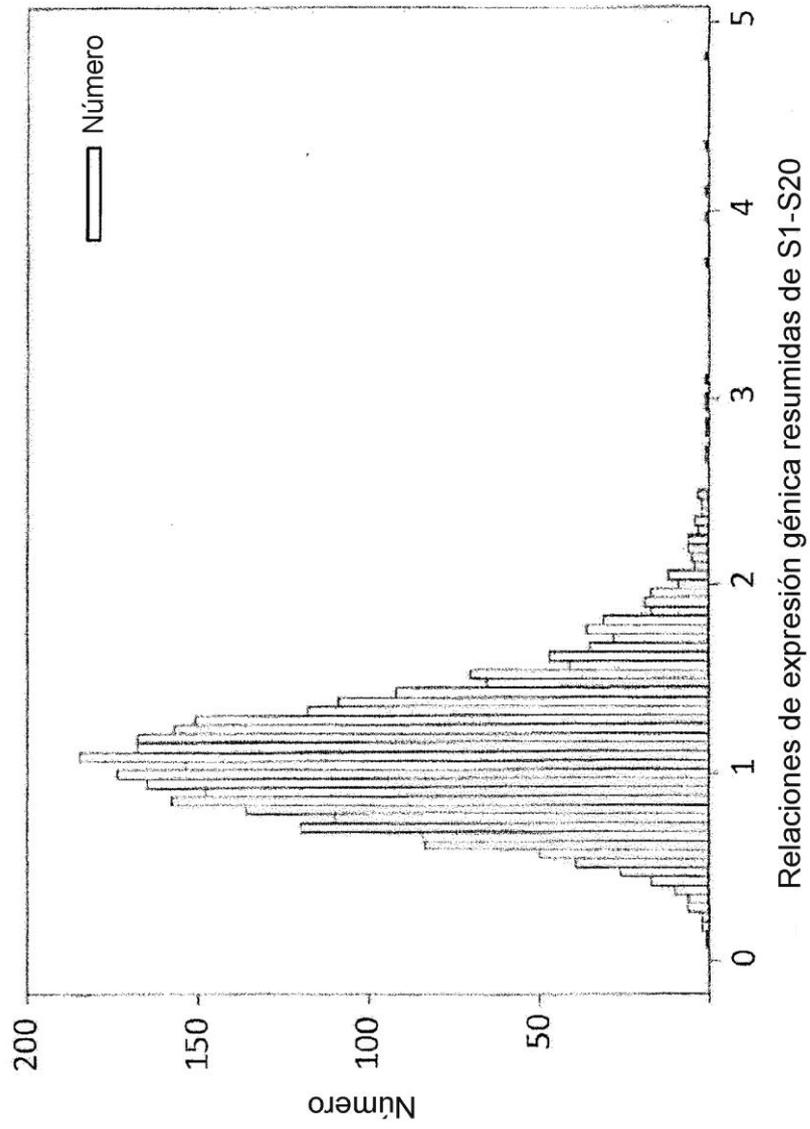


Figura 4



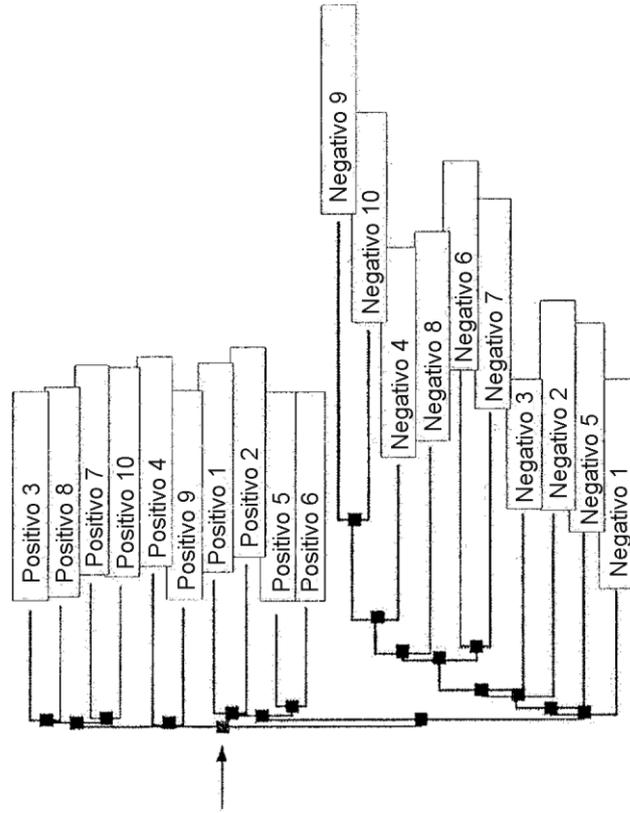


Figura 5

Figura 6

