

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 653 851**

51 Int. Cl.:

**C07H 21/00** (2006.01)

**C12Q 1/68** (2006.01)

**G01N 33/48** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **01.05.2008 PCT/US2008/062281**

87 Fecha y número de publicación internacional: **13.11.2008 WO08137586**

96 Fecha de presentación y número de la solicitud europea: **01.05.2008 E 08747396 (3)**

97 Fecha y número de publicación de la concesión europea: **04.10.2017 EP 2162459**

54 Título: **Biomarcadores transcriptómicos para evaluación de riesgo individual en insuficiencia cardíaca de nueva aparición**

30 Prioridad:

**01.05.2007 US 915224 P**  
**08.01.2008 US 19749 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**09.02.2018**

73 Titular/es:

**UNIVERSITY OF MIAMI (100.0%)**  
**1475 NW 12TH AVENUE, SUITE 2012**  
**MIAMI, FL 33136, US**

72 Inventor/es:

**HARE, JOSHUA, M. y**  
**HEIDECKER, BETTINA**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 653 851 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

Biomarcadores transcriptómicos para evaluación de riesgo individual en insuficiencia cardíaca de nueva aparición

### Campo de la invención

5 La presente invención se refiere a biomarcadores de dianas terapéuticas farmacológicas nuevas de enfermedad cardíaca, composiciones y procedimientos para predecir, diagnosticar y tratar enfermedades cardíacas y trastornos relacionados con las mismas. Más específicamente, la invención se refiere a procedimientos y composiciones basados en firmas moleculares únicas asociadas con diversos aspectos de enfermedades y trastornos cardíacos.

### Antecedentes

10 La evolución clínica de pacientes con insuficiencia cardíaca de nuevo diagnóstico varía drásticamente, con algunos pacientes que se recuperan y vuelven a niveles completamente normales de fracción de eyección (FE), mientras que otros desarrollan síntomas graves de descompensación cardíaca que requieren la inserción de un dispositivo de asistencia ventricular izquierda (DAVI) o un trasplante cardíaco. Una evaluación de riesgos y predicción de pronóstico en primera presentación precisas son cruciales para la asignación apropiada de supervisión de terapia y tratamiento de pacientes. Las herramientas de predicción basadas en criterios convencionales tienen una precisión limitada. La identificación de pacientes en fase temprana proporcionaría tanto a los pacientes individuales como a los gobiernos ahorros financieros sustanciales. Existe por lo tanto la necesidad de una evaluación de pronóstico precisa que permita el ajuste del tratamiento de forma apropiada y lo suficientemente pronto.

20 El documento US 2006/094038 A1 describe la identificación de genes cuyos productos génicos se expresan diferencialmente tras la sobrecarga de presión del corazón, y procedimientos de diagnóstico y terapéuticos relacionados.

Hari y col. (Cell Biochemistry and Biophysics 2006; 44 (1): 1085-9195) describen la identificación de genes implicados en enfermedad cardíaca congénita cianótica, tetralogía de Fallot (TF), que identifica 236 genes que muestran patrón de expresión alterado en pacientes de TF.

25 Archacki y col. (Physical Genomics, American Physiological Society 2004; 15: 65-74) describen el análisis de expresión de aproximadamente 12.000 genes humanos en nueve arterias coronarias humanas gravemente ateroscleróticas y seis no ateroscleróticas usando micromatrices de oligonucleótidos, que identifican 56 genes que muestran expresión diferencial en tejidos de arterias coronarias ateroscleróticas.

30 Heidecker y col. (Heart Failure Reviews 2007; 12: 1-11) revisan el uso de micromatrices para análisis transcriptómico para aplicaciones clínicas, particularmente en el área cardiovascular, y menciona, por ejemplo, la Matriz U133 PLUS 2.0 de Affymetrix.

Xiaohong y col. (BMC Bioinformatics 2005; 6 (1): 205) presentan un estudio comparativo de etiología de insuficiencia cardíaca humana diferenciadora usando perfiles de expresión génica, usando, por ejemplo, la Matriz U133 PLUS 2.0.

### Sumario

35 Los biomarcadores transcriptómicos comprenden biomoléculas y permiten la predicción del resultado de pronóstico de enfermedades cardíacas y trastornos de las mismas.

40 En un aspecto, la presente invención se refiere al uso de una firma molecular que consiste en secuencias génicas complementarias de las secuencias de ácido nucleico de la Tabla 2 para predecir un pronóstico de insuficiencia cardíaca de nueva aparición, que comprende examinar la expresión de las secuencias génicas de la firma molecular en una muestra tisular o celular.

En un aspecto, la presente invención se refiere al uso de una firma molecular que consiste en secuencias génicas complementarias de las secuencias de ácido nucleico de la Tabla 8 para predecir un pronóstico de insuficiencia cardíaca, que comprenden examinar la expresión de las secuencias génicas de la firma molecular en una muestra tisular o celular.

45 Estos y otros aspectos de la invención se exponen en las reivindicaciones adjuntas, y se describen adicionalmente posteriormente.

50 Los biomarcadores transcriptómicos, procedimientos y ensayos desvelados en el presente documento se dirigen a la examinación de la expresión de biomarcadores transcriptómicos en una muestra tisular o celular de mamífero, en la que la determinación de esa expresión de uno o más de dichos biomarcadores transcriptómicos es predictiva del resultado de pronóstico o diagnóstico de enfermedades y trastornos cardíacos y cardiovasculares, tales como por ejemplo, miocarditis, enfermedad cardíaca coronaria, angina, síndrome coronario agudo, aneurisma y disección aórtica, arritmias, cardiomiopatía, enfermedad cardíaca congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis y similares.

En una realización preferida una composición molecular comprende secuencias génicas o de ácido nucleico: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L) (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 244260\_at (clon de ADNC), 238643\_at (Neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito muy similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor del factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida, la detección en una célula o un paciente de las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas, es pronóstico de un buen pronóstico clínico en insuficiencia cardíaca, los biomarcadores son predictivos de resultado de pronóstico o diagnóstico de enfermedades y trastornos cardíacos y cardiovasculares, tales como por ejemplo, miocarditis, Enfermedad Cardíaca Coronaria, angina, Síndrome Coronario Agudo, Aneurisma y Disección Aórtica, arritmias, Cardiomiopatía, Enfermedad Cardíaca Congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis y similares.

En otra realización preferida, las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas, se sobreexpresan a niveles de al menos 1 % a 100 % o más en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal.

En otra realización preferida, las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas, se sobreexpresan en aproximadamente 50 % en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal.

En otra realización preferida, las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas, se sobreexpresan en aproximadamente 75 % en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal.

En otra realización preferida, las al menos diez secuencias génicas son pronóstico en evaluación de riesgos individual en un paciente con respecto a la aparición de insuficiencia cardíaca, enfermedades y trastornos cardíacos y cardiovasculares, tales como por ejemplo, miocarditis, Enfermedad Cardíaca Coronaria, angina, Síndrome Coronario Agudo, Aneurisma y Disección Aórtica, arritmias, Cardiomiopatía, Enfermedad Cardíaca Congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis y similares.

En otra realización preferida, una pluralidad de secuencias génicas son pronóstico en la evaluación de riesgos individual en un paciente para la aparición de insuficiencia cardíaca y predictivas de resultado de pronóstico o diagnóstico de enfermedades y trastornos cardíacos y cardiovasculares, tales como por ejemplo, miocarditis, Enfermedad Cardíaca Coronaria, angina, Síndrome Coronario Agudo, Aneurisma y Disección Aórtica, arritmias, Cardiomiopatía, Enfermedad Cardíaca Congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis y similares.

En otra realización preferida, un biomarcador (BBT) para predecir un pronóstico de insuficiencia cardíaca, comprende secuencias/biomoléculas de ácido nucleico que comprenden: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho

- GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7) (S. cerevisiae), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citotóxicos naturales), 224260\_at (clon de ADNC), 238643\_at (Neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (S. cerevisiae)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito muy similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor del factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.
- 5 El biomarcador de la reivindicación 8, en el que las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas, se sobreexpresan a niveles en al menos 1 % a 100 % o más en una célula o paciente en comparación con niveles en una célula normal o sujeto normal.
- 10 En otra realización preferida, las secuencias de ácido nucleico, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas, se sobreexpresan en aproximadamente 50 % en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal.
- 15 En otra realización preferida, las secuencias de ácido nucleico, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas se sobreexpresan en aproximadamente 75 % en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal.
- 20 En otra realización preferida, un anticuerpo o aptámero específico para cada secuencia génica que comprende:
- 30 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7 (S. cerevisiae)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citotóxicos naturales), 224260\_at (clon de ADNC), 238643\_at (Neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (S. cerevisiae)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito muy similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor del factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.
- 45
- 50
- 55 En otra realización preferida, un chip comprende las secuencias de ácido nucleico: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones
- 60

de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (Neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito muy similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor del factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En una realización, el biochip comprende al menos diez secuencias de ácido nucleico, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En una realización preferida, un procedimiento para evaluar, identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación comprende: identificar en una muestra biológica de un paciente una firma molecular que comprende un biomarcador basado en transcriptoma (BBT): 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7) (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (Neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito muy similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor del factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas; evaluar la probabilidad de identificación de cada gen componente en cada muestra; asignar cada uno a una clase; y diferenciar entre cardiomiopatía idiopática y miocarditis.

En una realización preferida, el biomarcador se identifica de un paciente aislando ácidos nucleicos obtenidos de una muestra biológica.

En otra realización preferida, los ácidos nucleicos hibridan con el biochip y los valores de intensidad sin procesar de la hibridación de micromatrices se normalizan y se identifican las diferencias específicas de fenotipo en la expresión génica. Las diferencias en la expresión génica se identifican por análisis de significación de micromatrices, en el que la significación se define con un valor q y múltiples comparaciones comprenden un valor p ajustado.

En otra realización preferida, se identifica la especificidad de fenotipos creando un clasificador en un conjunto de entrenamiento que comprende aproximadamente 66 % de datos obtenidos, con validación posterior en un conjunto de ensayo que comprende aproximadamente 33 % de datos obtenidos y definiendo un centroide reducido más próximo específico de fenotipo para clasificación. El centroide reducido más próximo específico de fenotipo comprende equilibrar una validación cruzada de aproximadamente 10 veces en un conjunto de entrenamiento.

En otra realización preferida, un procedimiento para predecir un resultado de pronóstico para enfermedad cardíaca o

recuperación de insuficiencia cardíaca comprende: identificar en una muestra biológica de un paciente una composición que comprende las secuencias génicas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citotóxicos naturales), 224260\_at (clon de ADNC), 238643\_at (Neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito muy similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor del factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas; y evaluar la probabilidad de identificación de cada gen componente en cada muestra; asignar cada uno a una clase; y predecir el resultado de pronóstico para enfermedad cardíaca o recuperación de la insuficiencia cardíaca.

En una realización, las secuencias génicas están sobreexpresadas en al menos aproximadamente 5 % en comparación con una célula normal o un sujeto normal.

En otra realización preferida, un procedimiento para identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación comprende: identificar secuencias génicas que comprenden: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citotóxicos naturales), 224260\_at (clon de ADNC), 238643\_at (Neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito muy similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor del factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas; evaluar la probabilidad de identificación de cada gen componente en cada muestra; asignar cada uno a una clase; e identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico de recuperación.

En otra realización preferida, una composición molecular comprende las secuencias génicas: 1558458\_at (LOC401320 Hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2

(perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contienen dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (Polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968 at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208 at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

Preferentemente, la detección en una célula o un paciente de las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas, es predictiva de resultado de diagnóstico clínico y pronóstico de insuficiencia cardíaca, predictivo de resultado de pronóstico de enfermedades de enfermedades y trastornos cardíacos y cardiovasculares, tales como, por ejemplo, miocarditis, Enfermedad Cardíaca Coronaria, angina, Síndrome Coronario Agudo, Aneurisma y Disección Aórtica, arritmias, Cardiomiopatía, Enfermedad Cardíaca Congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis y similares.

Preferentemente, las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas se sobreexpresan a niveles en al menos 1 % a 100 %, 200 %, 300 % o más en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal.

En una realización alternativa, al menos nueve secuencias génicas son pronóstico en la evaluación de riesgos individual en un paciente, predictivas de resultado de diagnóstico clínico y pronóstico de insuficiencia cardíaca.

En otra realización preferida, una pluralidad de secuencias génicas son pronóstico en evaluación de riesgos individual en un paciente para aparición de insuficiencia cardíaca.

En otra realización preferida, un biomarcador (BBT) predictivo de resultado de diagnóstico clínico y pronóstico de insuficiencia cardíaca que comprende secuencias/biomoléculas de ácido nucleico comprende: 1558458\_at (LOC401320 Hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contienen dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (Polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968 at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208 at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las

mismas.

5 En una realización, la detección en una célula o un paciente de las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas es predictiva de resultado de diagnóstico clínico y pronóstico de insuficiencia cardíaca, y es predictiva de resultado de pronóstico o diagnóstico de enfermedades y trastornos cardíacos y cardiovasculares, tales como por ejemplo, miocarditis, Enfermedad Cardíaca Coronaria, angina, Síndrome Coronario Agudo, Aneurisma y Disección Aórtica, arritmias, Cardiomiopatía, Enfermedad Cardíaca Congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis y similares.

10 En una realización preferida, las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas están sobreexpresados a niveles en al menos 1 % a 100 %, 200 %, 300 % o más en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal.

15 En otra, un biochip comprende las secuencias de ácido nucleico: 1558458\_at (LOC401320 Hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contienen dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (Polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968 at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208 at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En una realización, el biochip comprende al menos diez secuencias de ácido nucleico, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

40 En otra realización preferida, un anticuerpo o aptámero es específico para cada una de: 1558458\_at (LOC401320 Hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contienen dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (Polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968 at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208 at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho

GTPasa 26, ARHGAP26), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida, un procedimiento para evaluar, identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación comprende identificar en una muestra biológica de un paciente una firma molecular que comprende un biomarcador basado en transcriptoma (BBT): 1558458\_at (LOC401320 Hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citotóxicos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contienen dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (Polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNC DKFZp762N156 (del clon DKFZp762N156), 227968 at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208 at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas; evaluar la probabilidad de identificación de cada gen componente en cada muestra; asignar cada uno a una clase; y diferenciar entre cardiomiopatía idiopática y miocarditis.

En otra realización preferida, un procedimiento para predecir un resultado de pronóstico para enfermedad cardíaca o recuperación de insuficiencia cardíaca comprende identificar en una muestra biológica de un paciente una composición que comprende las secuencias génicas: 1558458\_at (LOC401320 Hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citotóxicos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contienen dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (Polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNC DKFZp762N156 (del clon DKFZp762N156), 227968 at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208 at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas; y evaluar la probabilidad de identificación de cada gen componente en cada muestra; asignar cada uno a una clase; y predecir la enfermedad cardíaca o cardiomiopatía.

En otra realización preferida, un procedimiento para identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación comprende identificar secuencias génicas que comprenden: 1558458\_at (LOC401320 Hipotética), 1560049\_at (repetición de triplete CUG, proteína de

unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contienen dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (Polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas; y evaluar la probabilidad de identificación de cada gen componente en cada muestra; asignar cada uno a una clase; e identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación.

En otra realización preferida, una célula expresa una cualquiera o más de las secuencias de ácido nucleico o productos de las mismas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (Neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito muy similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor del factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at (factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial).

En otra realización preferida, un vector que expresa uno cualquiera o más productos de ácido nucleico comprende: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión

a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (Neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito muy similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor del factor de necrosis tumoral 14), 22651\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida, una célula expresa una cualquiera o más de las secuencias de ácido nucleico o productos de las mismas que comprenden: 1558458\_at (LOC401320 Hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contienen dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (Polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNC DKFZp762N156 (del clon DKFZp762N156), 227968 at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208 at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida, un vector expresa una cualquiera o más de las secuencias de ácido nucleico que comprende: 1558458\_at (LOC401320 Hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contienen dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (Polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNC DKFZp762N156 (del clon DKFZp762N156), 227968 at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208 at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho

GTPasa 26, ARHGAP26), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

### **Breve descripción de los dibujos**

5 La invención se señala particularmente en las reivindicaciones adjuntas. Las ventajas anteriores y adicionales de la invención puede entenderse mejor haciendo referencia a la siguiente descripción tomada junto con los dibujos adjuntos, en los que:

10 La Figura 1 es una exploración de una fotografía que muestra un análisis de ARN total extraído con Bioanalizador Agilent 2100: cada muestra se ensayó con respecto a su integridad y pureza antes de hibridación de micromatrices. La fotografía representa un gel de 12 muestras con bandas uniformes de ARN 18S y 28S. El carril izquierdo contiene el marcador de referencia.

15 La Figura 2 es una representación esquemática de conjuntos de entrenamiento y ensayo como se usan para el desarrollo de un clasificador con Análisis de Predicción de Micromatrices (PAM). Todas las muestras obtenidas de pacientes con mal pronóstico (MP, n = 18) se seleccionaron de un depósito biológico (n = 180) y las que tenían un buen pronóstico (BP, n = 25) se seleccionaron de una manera con control de casos (véase el texto para las definiciones de MP y BP). El clasificador, un centroide reducido más próximo, se desarrolló en 2/3 de los datos (17 muestras con BP; 12 muestras con MP) con validación posterior en el 1/3 restante de los datos (8 muestras con BP, 6 muestras con MP). La precisión de ensayo global del BBT se calculó a partir de 50 particiones aleatorias en conjuntos de entrenamiento y ensayo.

20 La Figura 3 es un mapa de calor de muestras de todos los pacientes con cardiomiopatía dilatada idiopática (n = 43). Cada columna corresponde a una muestra de paciente y cada fila representa un gen. Las muestras clasificadas como con MP forman un grupo definido y están destacadas en un cuadrado rojo. Los genes regulados negativamente se representan con rojo, mientras que los genes regulados positivamente se marcan en azul. Las flechas amarillas indican muestras clasificadas erróneamente.

25 La Figura 4 es un diagrama circular que ilustra rutas implicadas dentro del biomarcador de pronóstico. Las rutas principales sobreexpresadas en pacientes con BP incluyeron transcripción (26 %), unión de proteínas (15 %), transporte de iones (13 %) y desarrollo neuromuscular (10 %).

30 La Figura 5 muestra dos gráficos que ilustran la mejora funcional de la fracción de eyección (FE) desde la línea basal hasta el punto final: dentro de todos los casos admitidos de cardiomiopatía idiopática (n = 43), se analizaron adicionalmente aquellos de los que estuvieron disponibles medidas ecocardiográficas en línea basal y punto final (n = 17). Las muestras se clasifican en BP o MP basándose en la predicción de BBT. Los pacientes clasificados como BP (n = 11, promedio de seguimiento: 49,9±21 m) experimentaron mejora de FE (\*P = 0,0009), mientras que los MP (n = 6, promedio de seguimiento: 6,2±2,9 m) no. La línea roja representa una muestra clasificada erróneamente. Las barras de error representan ETM.

35 La Figura 6 es una gráfica que muestra la validación cruzada del conjunto de entrenamiento: partiendo de un conjunto de 9 genes, el aumento del umbral conduce a un aumento constante del error de clasificación errónea después de un límite de 2,1.

La Figura 7 es una gráfica que muestra la validación cruzada del conjunto de entrenamiento: el aumento del umbral aumenta constantemente el error de clasificación errónea después de un límite de 2,1.

40 La Figura 8 es un mapa de calor que usa una firma de 9 genes: este mapa de calor se creó por un enfoque de agrupamiento no supervisado usando normalización por niveles de expresión medios. Cada columna representa una muestra y cada línea corresponde a un gen, para el que se enumeran las ID de affymetrix en el lado derecho. Para más detalles acerca de las anotaciones de genes véase tabla 5. Un color rojo significa baja expresión del gen, mientras que un color azul demuestra altos niveles de expresión génica. Tres muestras de pacientes con mal pronóstico se agruparon erróneamente (rodeadas).

45 La Figura 9 muestra un mapa de calor usando una firma de 43 genes: este mapa de calor se creó con el mismo algoritmo de agrupamiento jerárquico no supervisado que en la Figura 9. Solamente una muestra se clasificó erróneamente (BP-10). Resulta interesante que, realizando investigación adicional acerca del historial de este paciente, se descubrió que se le había diagnosticado en primer lugar insuficiencia cardíaca más de 5 años antes de morir, lo que sugiere que era de hecho un superviviente a largo plazo y fue reconocido correctamente por el algoritmo de los inventores. En el estudio de los inventores, se le clasificó como un paciente con mal pronóstico, porque la biopsia se obtuvo un año antes de morir.

50 La Figura 10 es una gráfica que muestra el centroide reducido más próximo del "clasificador de 9 genes": los centroides se calcularon en PAM a partir de la expresión promedio para cada gen en cada clase dividida por la desviación típica dentro de la clase para ese gen. La clasificación de centroide reducido más próximo "reduce" cada uno de los centroides de clase hacia el centroide general para todas las clases por el umbral. La clasificación del centroide más próximo toma el perfil de expresión génica de una nueva muestra, y lo compara con cada uno de estos centroides de clase. La clase de cuyo centroide está más cerca, en distancia al cuadrado, es la clase predicha para esa nueva muestra. Cada línea en la gráfica representa un gen. El centroide rojo caracteriza el grupo 1 (mal pronóstico), el centroide verde caracteriza el grupo 2 (buen pronóstico). Todos los genes, excepto uno, que se regularon positivamente en el grupo con buen pronóstico, estaban infraexpresados en el grupo con mal resultado clínico.

60 La Figura 11 muestra el centroide reducido más próximo del "clasificador de 43 genes": después de extender el clasificador de los inventores a 43 genes, apareció un segundo gen regulado positivamente en el grupo con mal

pronóstico.

### **Descripción detallada**

La invención se refiere a firmas moleculares que actúan como un biomarcador de pronóstico muy sensible para insuficiencia cardíaca, enfermedades cardíacas, miocarditis y otros trastornos cardíacos.

- 5 La predicción de pronóstico sigue siendo una necesidad no satisfecha importante en insuficiencia cardíaca (IC) de nueva aparición. Aunque se usan varios ensayos clínicos, ninguno distingue con precisión entre pacientes con mala supervivencia frente a supervivencia excelente. Una firma transcriptómica, generada a partir de una biopsia endomiocárdica (BEM), actúa como un biomarcador de pronóstico nuevo en IC.

#### Definiciones

- 10 De acuerdo con la presente invención y como se usa en el presente documento, los siguientes términos se definen con los siguientes significados, a no ser que se indique de forma explícita otra cosa.

Como se usa en el presente documento, “un” y “el” incluyen referencias plurales a no ser que el contexto claramente dicte otra cosa.

- 15 Como se usa en el presente documento, una “firma molecular” o “firma” o “biomarcador” o “biomarcador basado en transcriptoma” se usan indistintamente en el presente documento y se refieren a todas las biomoléculas identificadas en las Tablas 2 y 8. A medida que se descubren más biomoléculas, cada biomolécula de nueva identificación puede asignarse a uno o más biomarcadores o firmas moleculares. Cada biomolécula también puede retirarse, reasignarse o redistribuirse a una firma molecular.

- 20 El término “biomolécula” se refiere a ADN, ARN (incluyendo ARNm, ARNr, ARNt y ARNtm), nucleótidos, nucleósidos, polinucleótidos, péptidos y cualquier combinación de los mismos.

Una “posición” base como se usa en el presente documento se refiere a la localización de una base o resto de nucleótido dado dentro de un ácido nucleico.

Como se usa en el presente documento, el término “matriz” se refiere a una disposición espacial ordenada, particularmente una disposición de biomoléculas inmovilizadas.

- 25 Como se usa en el presente documento, la expresión “matriz direccionable” se refiere a una matriz en la que los elementos individuales tienen coordenadas x e y definidas con precisión, de modo que puede identificarse un elemento dado en una posición particular en la matriz.

- 30 Como se usan en el presente documento, las expresiones “sonda” y “sonda biomolecular” se refieren a una biomolécula usada para detectar una biomolécula complementaria. Los ejemplos incluyen antígenos que detectan anticuerpos, oligonucleótidos que detectan oligonucleótidos complementarios y ligandos que detectan receptores. Dichas sondas se inmovilizan preferentemente en un microelectrodo que comprende un sustrato.

- 35 Como se usa en el presente documento, las expresiones “biomatriz”, “biochip” y “matriz de biochip” se refieren a una disposición espacial ordenada de biomoléculas inmovilizadas en un microelectrodo dispuesto sobre un sustrato de soporte sólido. Las moléculas sonda preferidas incluyen aptámeros, ácidos nucleicos, oligonucleótidos, péptidos, ligandos, anticuerpos y antígenos; las especies de sonda más preferidas son péptidos y proteínas. Las microplacas biológicas, como se usan en la técnica, abarcan sustratos que contienen matrices o micromatrices, preferentemente matrices ordenadas y más preferentemente matrices ordenadas, direccionables, de moléculas biológicas que comprenden un miembro de un par de unión biológico. Típicamente, dichas matrices son matrices de oligonucleótidos que comprenden una secuencia de nucleótidos que es complementaria de al menos una secuencia  
40 que puede estar o se espera que esté presente en una muestra biológica. Como alternativa, y preferentemente, las proteínas, los péptidos u otras moléculas pequeñas pueden disponerse en dichas microplacas biológicas para realizar, entre otros, análisis inmunológicos (en los que las moléculas dispuestas son antígenos) o ensayar receptores biológicos (en los que las moléculas dispuestas son ligandos, agonistas o antagonistas de dichos receptores).

- 45 La expresión/cantidad de un gen, una biomolécula o un biomarcador en una primera muestra está a un nivel “mayor que” el nivel en una segunda muestra si el nivel de expresión/cantidad del gen o biomarcador en la primera muestra es al menos aproximadamente 1 vez, 1,2 veces, 1,5 veces, 1,75 veces, 2 veces, 3 veces, 4 veces, 5 veces, 6 veces, 7 veces, 8 veces, 9 veces, 10 veces, 20 veces, 30 veces el nivel de expresión/cantidad del gen o biomarcador en la segunda muestra o una muestra normal. Los niveles de expresión/cantidades pueden determinarse basándose en  
50 cualquier criterio adecuado conocido en la técnica, incluyendo, pero sin limitación, ARNm, ADNc, proteínas, fragmentos proteicos y/o copias génicas. Los niveles de expresión/cantidades pueden determinarse de forma cualitativa y/o cuantitativa.

Por el término “modular”, se entiende que cualquiera de las actividades mencionadas, por ejemplo, se aumentan, potencian, actúan como un agonista, promueven, reducen, disminuyen, suprimen, bloquean o antagonizan (actúan

como un agonista). La modulación puede aumentar la actividad en más de 1 vez, 2 veces, 3 veces, 5 veces, 10 veces, 100 veces, etc., con respecto a valores de línea basal. La modulación también puede reducir su actividad por debajo de valores de línea basal.

5 El término “complementario” significa que dos secuencias son complementarias cuando la secuencia de una puede unirse con la secuencia de la otra en un sentido antiparalelo en el que el extremo 3' de cada secuencia se une al extremo 5' de la otra secuencia y cada A, T (U), G y C de una secuencia se alinea después con una T (U), A, C y G, respectivamente, de la otra secuencia. Normalmente, la secuencia complementaria del oligonucleótido tiene al menos 80 % o 90 %, preferentemente 95 %, más preferentemente 100 %, de complementariedad con una secuencia definida. Preferentemente, pueden identificarse alelos o variantes de los mismos. También puede emplearse un programa BLAST para abordar dicha identidad de secuencia.

10 La expresión “secuencia complementaria” en referencia a una secuencia polinucleotídica, se refiere a la secuencia de bases en otra molécula de ácido nucleico por las normas de formación de pares de bases. Más particularmente, la expresión o una expresión similar se refiere a la hibridación o formación de pares de bases entre nucleótidos o ácidos nucleicos, tal como, por ejemplo, entre las dos cadenas de una molécula de ADN bicatenario o entre un cebador oligonucleotídico y un sitio de unión a cebadores en un ácido nucleico monocatenario para secuenciar o amplificar. Los nucleótidos complementarios son, en general, A y T (o A y U), o C y G. Se dice que dos moléculas de ARN o ADN monocatenarias son sustancialmente complementarias cuando los nucleótidos de una cadena, alineados de forma óptima y comparados y con inserciones o supresiones de nucleótidos apropiadas, se emparejan con al menos aproximadamente 95 % de los nucleótidos de la otra cadena, habitualmente al menos aproximadamente 98 % y más preferentemente de aproximadamente 99 % a aproximadamente 100 %. Pueden identificarse secuencias polinucleotídicas complementarias por una diversidad de enfoques incluyendo el uso de algoritmos informáticos bien conocidos y software, por ejemplo el programa BLAST.

20 Un “alelo” o una “variante” es una forma alternativa de un gen. Las variantes pueden resultar de al menos una mutación en la secuencia de ácido nucleico y pueden dar como resultado ARNm alterados o polipéptidos cuya estructura o función puede estar o no alterada. Cualquier gen natural o recombinante dado puede no tener ninguna forma alélica, tener una o tener muchas. Los cambios mutacionales comunes que dan lugar a variantes se atribuyen en general a supresiones, adiciones o sustituciones naturales de nucleótidos. Cada uno de estos tipos de cambios puede producirse solo, o en combinación con los otros, una o más veces en una secuencia dada.

25 Como se usa en el presente documento, la expresión “aptámero” o “especie de unión de ácido nucleico seleccionada” incluirá ARN o ADN modificado químicamente o no modificado. El procedimiento de selección puede ser, pero sin limitación, mediante cromatografía de afinidad y el procedimiento de amplificación por transcripción inversa (RT) o reacción en cadena de la polimerasa (PCR).

30 Como se usa en el presente documento, la expresión “aptámero de señalización” incluirá aptámeros con moléculas indicadoras, preferentemente un colorante fluorescente, unidas a un nucleótido de tal modo que tras cambios conformacionales resultantes de la interacción del aptámero con un ligando, las moléculas indicadoras producen una señal diferencial, preferentemente un cambio en la intensidad de fluorescencia.

35 Como se usa en el presente documento, la expresión “fragmento o segmento”, como se aplica a una secuencia de ácido nucleico, gen o polipéptido, será habitualmente de al menos aproximadamente 5 bases de ácido nucleico contiguas (para secuencia de ácido nucleico o gen) o aminoácidos (para polipéptidos), típicamente al menos aproximadamente 10 bases de ácido nucleico o aminoácidos contiguos, más típicamente al menos aproximadamente 20 bases de ácido nucleico o aminoácidos contiguos, habitualmente al menos aproximadamente 30 bases de ácido nucleico o aminoácidos contiguos, preferentemente al menos aproximadamente 40 bases de ácido nucleico o aminoácidos contiguos, más preferentemente al menos aproximadamente 50 bases de ácido nucleico o aminoácidos contiguos y aún más preferentemente al menos aproximadamente 60 a 80 o más bases de ácido nucleico o aminoácidos contiguos de longitud. Los “fragmentos solapantes” como se usa en el presente documento, se refieren a fragmentos de ácido nucleico o peptídicos contiguos que comienzan en el extremo amino terminal de un ácido nucleico o una proteína y terminan en el extremo carboxilo terminal del ácido nucleico o la proteína. Cada fragmento de ácido nucleico o péptido tiene al menos aproximadamente una posición de ácido nucleico o aminoácido contigua en común con el siguiente fragmento de ácido nucleico o péptido, más preferentemente al menos aproximadamente tres bases de ácido nucleico o posiciones de aminoácidos contiguas en común, más preferentemente al menos aproximadamente diez bases de ácido nucleico o posiciones de aminoácidos contiguas en común.

40 Las “muestras biológicas” incluyen muestras sólidas y de fluidos corporales. Preferentemente, la muestra se obtiene del corazón. Sin embargo, las muestras biológicas usadas en la presente invención pueden incluir células, proteína o extractos de membrana de células, sangre o fluidos biológicos tales como líquido ascítico o líquido cefalorraquídeo (por ejemplo, líquido cefalorraquídeo). Los ejemplos de muestras biológicas sólidas incluyen, pero sin limitación, muestras tomadas de tejidos del sistema nervioso central, hueso, mama, riñón, cuello uterino, endometrio, cabeza/cuello, vesícula biliar, glándula parótida, próstata, glándula hipófisis, músculo, esófago, estómago, intestino delgado, colon, hígado, bazo, páncreas, tiroides, corazón, pulmón, vejiga, tejido adiposo, ganglio linfático, útero, ovario, glándula adrenal, testículo, amígdala y timo. Los ejemplos de “muestras de fluidos corporales” incluyen, pero

sin limitación, sangre, suero, semen, líquido prostático, líquido seminal, orina, saliva, esputo, moco, médula ósea, linfa y lágrimas.

La “muestra” se usa en el presente documento en su sentido más amplio. Una muestra que comprende polinucleótidos, polipéptidos, péptidos, anticuerpos y similares puede comprender un líquido corporal; una fracción soluble de una preparación celular, o un medio en el que se cultivaron las células; un cromosoma, un orgánulo o membrana aislada o extraída de una célula; ADN genómico, ARN o ADNc, polipéptidos o péptidos en solución o unidos con un sustrato; una célula; un tejido; una impresión tisular; una huella dactilar, piel o pelo; y similares.

“Diagnóstico” significa identificar la presencia o naturaleza de una afección patológica. Los procedimientos de diagnóstico difieren en su sensibilidad y especificidad. La “sensibilidad” de un ensayo de diagnóstico es el porcentaje de individuos enfermos que tienen un resultado de ensayo positivo (porcentaje de “verdaderos positivos”). Los individuos enfermos no detectados por el ensayo son “falsos negativos”. Los sujetos que no están enfermos y que tienen resultado de ensayo negativo se denominan “verdaderos negativos”. La “especificidad” de un ensayo de diagnóstico es 1 menos la tasa de falsos positivos, donde la tasa de “falsos positivos” se define como la proporción de los que no tienen la enfermedad que dan resultado de ensayo positivo. Aunque un procedimiento de diagnóstico particular puede no proporcionar un diagnóstico definitivo de una afección, es suficiente si el procedimiento proporciona una indicación positiva que ayuda en el diagnóstico.

Los términos “detección” o “detectar” incluyen ensayar, cuantificar, capturar imágenes o establecer de otro modo la presencia o ausencia del biomarcador transcriptómico, o combinaciones de biomoléculas que comprenden el biomarcador, y similares, o ensayar con respecto a, capturar imágenes, averiguar, establecer o determinar de otro modo el pronóstico de insuficiencia cardíaca.

#### *Biomarcador transcriptómico/firmas moleculares*

En una realización preferida, un biomarcador para el pronóstico del resultado de insuficiencia cardíaca comprende: biomoléculas/secuencias de ácido nucleico que comprenden las secuencias génicas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición WD 33), 213773\_x\_at (familia del dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida, la detección en una célula o un paciente de un biomarcador (BBT) que comprende: secuencias de ácido nucleico/biomoléculas que comprenden: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG), proteína de unión a ARN 2), 228198\_s\_at (proteína

ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición WD 33), 213773\_x\_at (familia del dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas, es diagnóstico de trastornos miocárdicos. Preferentemente, las secuencias de ácido nucleico, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas se sobreexpresan a niveles de al menos entre aproximadamente 1 %, 1,2 %, 2 %, 5 %, 10 %, 20 % en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal; más preferentemente, las secuencias de ácido nucleico, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas se sobreexpresan en aproximadamente 50 % en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal; más preferentemente, las secuencias de ácido nucleico, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas se sobreexpresan en aproximadamente 75 % en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal.

En otra realización preferida, un biomarcador comprende las secuencias de ácido nucleico: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 // MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (Lon peptidasa 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida, un biochip comprende una firma molecular/biomarcador que comprende las secuencias de ácido nucleico: un biomarcador comprende las secuencias génicas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG), proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de

Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición WD 33), 213773\_x\_at (familia del dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida un biochip comprende las secuencias de ácido nucleico: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRR37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (Lon peptidasa 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNC DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida un procedimiento de pronóstico de insuficiencia cardíaca comprende identificar en una muestra biológica de un paciente una firma molecular que comprende un biomarcador basado en transcriptoma (BBT): un biomarcador comprende las secuencias génicas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG), proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición WD 33), 213773\_x\_at (familia del dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las

mismas; evaluar la probabilidad de identificación de cada gen componente en cada muestra; y asignar cada uno a una clase.

En otra realización preferida, un procedimiento para predecir el resultado clínico y pronóstico en insuficiencia cardíaca comprende detección de un biomarcador que comprende: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (Lon peptidasa 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En una realización preferida, la especificidad de fenotipo se identifica creando un clasificador en un conjunto de entrenamiento que comprende aproximadamente 66 % de los datos obtenidos, con validación posterior en un conjunto de ensayo que comprende aproximadamente 33 % de los datos obtenidos y que define un centroide reducido más próximo específico de fenotipo para clasificación.

En otra realización preferida, un procedimiento para identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación comprende: identificar en una muestra biológica de un paciente una firma molecular que comprende un biomarcador basado en transcriptoma (BBT): un biomarcador comprende las secuencias génicas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG), proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición WD 33), 213773\_x\_at (familia del dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas; evaluar la probabilidad de identificación de cada gen componente en cada muestra; asignar cada uno a una clase; e identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación.

En otra realización preferida, un procedimiento para identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación comprende: identificar en una muestra biológica de un paciente una firma molecular que comprende un biomarcador basado en transcriptoma (BBT): un biomarcador comprende las secuencias génicas: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH66 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (Lon peptidasa 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNC DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas; evaluar la probabilidad de identificación de cada gen componente en cada muestra; asignar cada uno a una clase; e identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación.

En otra realización preferida, la predicción de pronóstico de insuficiencia cardíaca comprende detectar al menos diez moléculas que tienen secuencias génicas que comprenden: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG), proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición WD 33), 213773\_x\_at (familia del dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida, la predicción de pronóstico de insuficiencia cardíaca comprende detectar al menos nueve moléculas que tienen secuencias génicas que comprenden: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D

(ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 // MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (Lon peptidasa 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFzP762N156 (del clon DKFzP762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDCC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida, una composición de biochip que comprende las secuencias génicas: un biomarcador comprende las secuencias génicas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG), proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición WD 33), 213773\_x\_at (familia del dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En una realización, el biochip comprende a menos diez biomoléculas seleccionadas de un biomarcador que comprende las secuencias génicas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG), proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at

- (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición WD 33), 213773\_x\_at (familia del dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.
- 5
- 10 En otra realización preferida, una composición de biochip que comprende las secuencias génicas: un biomarcador comprende las secuencias génicas: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 // MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (Lon peptidasa 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.
- 15
- 20
- 25
- 30
- 35

- En una realización preferida, un biochip comprende al menos nueve moléculas que comprenden: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (Lon peptidasa 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC 1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.
- 40
- 45
- 50
- 55
- 60

En otra realización preferida, un procedimiento para predecir un pronóstico de insuficiencia cardíaca de nueva aparición, enfermedad cardíaca o afecciones de la misma comprende: identificar en una muestra biológica de un

paciente un biomarcador que comprende las secuencias génicas: un biomarcador que comprende las secuencias génicas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG), proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición WD 33), 213773\_x\_at (familia del dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas; y evaluar la probabilidad de identificación de cada gen componente en cada muestra; y asignar cada uno a una clase. Se detalla un ejemplo de un procedimiento preferido en los ejemplos a continuación.

En otra realización preferida, un procedimiento para identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y en pacientes con un buen pronóstico para recuperación comprende: identificar un biomarcador que comprende secuencias génicas que comprenden: un biomarcador comprende las secuencias génicas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG), proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición WD 33), 213773\_x\_at (familia del dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas; y evaluar la probabilidad de identificación de cada gen componente en cada muestra; y asignar cada uno a una clase; identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación.

En otra realización preferida, un procedimiento para identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación comprende: identificar un biomarcador que comprende secuencias génicas que comprenden: un biomarcador comprende las secuencias génicas: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales,

NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, PATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (Lon peptidasa 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

#### *Procedimientos y materiales alternativos para identificar firmas moleculares o biomarcadores transcriptómicos*

*Detección de ácidos nucleicos y proteínas como marcadores:* en realizaciones preferidas, cada biomarcador se detecta en procedimientos basados en microplacas tales como los descritos en detalle en los ejemplos a continuación, para proporcionar un pronóstico y diagnóstico precisos de pacientes con insuficiencia cardíaca con buen resultado de pronóstico, trastornos y enfermedades cardíacos, por ejemplo, insuficiencia cardíaca, miocarditis, cardiomiopatía idiopática y similares. También se conocen en la técnica otros procedimientos y pueden utilizarse uno o más procedimientos.

Los procedimientos y ensayos desvelados en el presente documento se dirigen al examen de la expresión de biomarcadores transcriptómicos en una muestra tisular o celular de mamífero, en el que la determinación de esa expresión de uno o más de dichos biomarcadores transcriptómicos es predictiva del resultado de pronóstico o diagnóstico de enfermedades y trastornos cardíacos y cardiovasculares, tales como por ejemplo, miocarditis, Enfermedad Cardíaca Coronaria, angina, Síndrome Coronario Agudo, Aneurisma y Disección Aórtica, arritmias, Cardiomiopatía, Enfermedad Cardíaca Congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis y similares. Las firmas moleculares o biomarcadores transcriptómicos comprenden las biomoléculas identificadas en la Tabla 2. A medida que se descubren más biomoléculas, cada biomolécula de nueva identificación puede asignarse a uno cualquiera o más biomarcadores o firmas moleculares. Cada biomolécula también puede retirarse, reasignarse o redistribuirse a una firma molecular.

Se describen realizaciones preferidas en la identificación de biomoléculas, procedimientos analíticos, etc., en detalle en los Ejemplos a continuación.

*Micromatrices:* en general, usando micromatrices de ácido nucleico, se describen de forma inversa muestras de ARNm de ensayo y de control de muestras tisulares de ensayo y de control y se marcan para generar sondas de ADNc. Las sondas se hibridan después con una matriz de ácidos nucleicos inmovilizada en un soporte sólido. La matriz se configura de modo que se conozca la secuencia y posición de cada miembro de la matriz. Por ejemplo, una selección de genes que tienen potencial para expresarse en ciertas patologías puede disponerse en un soporte sólido. La hibridación de una sonda marcada con un miembro de matriz particular indica que la muestra de la que se obtuvo la sonda expresa ese gen. El análisis de expresión génica diferencial de tejido enfermo puede proporcionar información valiosa. La tecnología de micromatrices utiliza técnicas de hibridación de ácido nucleico y tecnología de computación para evaluar el perfil de expresión de ARNm de miles de genes en un único experimento (véase, por ejemplo, documento WO 01/75166 publicado el 11 de octubre de 2001; (véase, por ejemplo, Patente de Estados Unidos n.º 5.700.637, Patente de Estados Unidos n.º 5.445.934 y Patente de Estados Unidos n.º 5.807.522, Lockart, Nature Biotechnology, 14: 1675-1680 (1996); Cheung, VG y col., Nature Genetics 21 (Supl): 15-19 (1999) para un análisis de la fabricación de matrices). Las micromatrices de ADN son matrices en miniatura que contienen fragmentos génicos que se sintetizan directamente en o se aplican puntualmente sobre vidrio u otros sustratos. En una única matriz se representan habitualmente miles de genes. Un experimento de micromatriz típico implica las siguientes etapas: 1) preparación de diana marcada con fluorescencia a partir de ARN aislado de la muestra, 2) hibridación de la diana marcada con la micromatriz, 3) lavado, tinción y exploración de la matriz, 4) análisis de la imagen explorada y 5) generación de perfiles de expresión génica. En la actualidad se usan dos tipos principales de micromatrices de ADN: matrices de oligonucleótidos (habitualmente 25 a 70 unidades) y matrices de expresión génica que contienen productos de PCR preparados a partir de ADNc. En la formación de una matriz, los oligonucleótidos pueden prefabricarse y aplicarse puntualmente a la superficie o sintetizarse directamente en la

superficie (*in situ*). El sistema Affymetrix GeneChip™ es un sistema de micromatrices disponible en el mercado que comprende matrices fabricadas por síntesis directa de oligonucleótidos en una superficie de vidrio.

*Matrices de sondas/genes:* se sintetizan oligonucleótidos, habitualmente de 25 unidades, directamente sobre una oblea de vidrio por una combinación de fotolitografía basada en semiconductor y tecnologías de síntesis química de fase sólida. Cada matriz contiene hasta 400.000 oligonucleótidos diferentes y cada oligonucleótido está presente en millones de copias. Ya que las sondas oligonucleotídicas se sintetizan en localizaciones conocidas en la matriz, los patrones de hibridación y las intensidades de señal pueden interpretarse con respecto a identidad génica y niveles de expresión relativos por el software del Grupo de Micromatrices Affymetrix. Cada gen está representado en la matriz por una serie de sondas oligonucleotídicas diferentes. Cada par de sondas consiste en un oligonucleótido perfectamente coincidente y un oligonucleótido desapareado. La sonda perfectamente coincidente tiene una secuencia exactamente complementaria del gen particular y por lo tanto mide la expresión del gen. La sonda desapareada difiere de la sonda perfectamente coincidente en una sustitución de una única base en la posición de base central, alterando la unión del transcrito génico diana. Esto ayuda a determinar la hibridación de fondo inespecífica que contribuye a la señal medida para el oligonucleótido perfectamente coincidente. El software de Conjunto de Micromatrices resta las intensidades de hibridación de las sondas desapareadas de las de las sondas perfectamente coincidentes para determinar el valor de intensidad absoluto o específico para cada conjunto de sondas. Las sondas se eligen basándose en información actual de GenBank y otros depósitos de nucleótidos. Se cree que las secuencias reconocen regiones únicas del extremo 3' del gen. Se usa un horno de hibridación GeneChip (horno "asador") para llevar a cabo la hibridación de hasta 64 matrices a la vez. La estación fluidica realiza lavado y tinción de las matrices de sondas. Está completamente automatizada y contiene cuatro módulos, conteniendo cada módulo una matriz de sonda. Cada módulo se controla de forma independiente mediante software de Conjunto de Micromatrices usando protocolos de fluidos preprogramados. El explorador es un explorador de fluorescencia de láser confocal que mide la intensidad de fluorescencia emitida por el ARNc marcado unido a las matrices de sondas. La estación de trabajo informática con el software de Conjunto de Micromatrices controla la estación fluidica y el explorador. El software de Conjunto de Micromatrices puede controlar hasta ocho estaciones fluidicas usando protocolos de hibridación, lavado y tinción preprogramados para la matriz de sondas. El software también adquiere y convierte datos de intensidad de hibridación en una determinación de presencia/ausencia para cada gen usando algoritmos apropiados. Finalmente, el software detecta cambios en la expresión génica entre experimentos por análisis de comparación y formatea el resultado en archivos.txt, que pueden usarse con otros programas informáticos para análisis de datos adicionales.

La expresión de un biomarcador seleccionado también puede evaluarse examinando la supresión de genes o amplificación de genes. La supresión o amplificación de genes puede medirse por uno cualquiera de una amplia diversidad de protocolos conocidos en la técnica, por ejemplo, mediante transferencia de Southern convencional, transferencia de Northern para cuantificar la transcripción de ARNm (Thomas, Proc. Natl. Acad. Sci. USA, 77: 5201-5205 (1980)), transferencia puntual (análisis de ADN) o hibridación *in situ* (por ejemplo, FISH), usando una sonda marcada de forma apropiada, procedimientos citogenéticos o hibridación genómica comparativa (CGH) usando una sonda marcada de forma apropiada.

*Detección de polipéptidos:* en otra realización, se detecta un polipéptido correspondiente a un marcador. Un agente preferido para detectar un polipéptido es un anticuerpo o aptámero capaz de unirse con un polipéptido correspondiente a un marcador de la invención, preferentemente un anticuerpo con un marcador detectable. Los anticuerpos pueden ser policlonales, o más preferentemente, monoclonales. Puede usarse un anticuerpo intacto, o un fragmento del mismo, por ejemplo, Fab o F(ab')<sub>2</sub>. Se pretende que el término "marcado", con respecto a la sonda o el anticuerpo, abarque marcaje directo de la sonda o el anticuerpo por acoplamiento, es decir, uniendo físicamente, una sustancia detectable con la sonda o el anticuerpo, así como un marcaje indirecto de la sonda o el anticuerpo por reactividad con otro reactivo que está marcado directamente. Los ejemplos de marcaje indirecto incluyen detección de un anticuerpo primario usando un anticuerpo secundario marcado con fluorescencia y marcaje de extremos de una sonda de ADN con biotina de modo que pueda detectarse con estreptavidina marcada con fluorescencia.

Pueden aislarse proteínas de individuos usando técnicas que son bien conocidas por los expertos en la materia. Los procedimientos de aislamiento de proteínas empleados pueden ser, por ejemplo, tales como los descritos en Harlow y Lane (1988), mencionado anteriormente. Puede emplearse una diversidad de formatos para determinar si una muestra contiene una proteína que se une con un anticuerpo dado. La expresión de diversos biomarcadores en una muestra puede analizarse mediante varias metodologías, muchas de las cuales se conocen en la técnica y son entendidas por los expertos en la materia, incluyendo pero sin limitación, inmunohistoquímica y/o análisis de Western, ensayos basados en sangre cuantitativos (como por ejemplo ELISA de suero) (para examinar, por ejemplo, los niveles de expresión de proteínas), ensayos de actividad enzimática bioquímica, hibridación *in situ*, análisis de Northern y/o análisis por PCR de ARNm, así como uno cualquiera de una amplia diversidad de ensayos que pueden realizarse por análisis de matrices génicas y/o tisulares. Se encuentran protocolos típicos para evaluar el estado de genes y productos génicos, por ejemplo en Ausubel y col. eds., 1995, Current Protocols In Molecular Biology, Unidades 2 (Transferencia de Northern), 4 (Transferencia de Southern), 15 (Inmunotransferencia) y 18 (Análisis por PCR). Un experto en la materia puede adaptar fácilmente procedimientos de detección de anticuerpos/proteínas conocidos para su uso en la determinación de si las células expresan un marcador de la presente invención y la concentración relativa de ese producto de expresión polipeptídico específico en sangre u otros tejidos corporales.

En dichos procedimientos alternativos, una muestra puede ponerse en contacto con un anticuerpo específico para dicho biomarcador en condiciones suficientes para que se forme un complejo de anticuerpo-biomarcador, y después detectar dicho complejo. La presencia del biomarcador puede detectarse de varias maneras, tales como por transferencia de Western y procedimientos de ELISA para ensayar una amplia diversidad de tejidos y muestras, incluyendo plasma o suero. Está disponible una amplia serie de técnicas de inmunoensayo usando dicho formato de ensayo, véase, por ejemplo, Patentes de Estados Unidos n.º 4.016.043, 4.424.279 y 4.018.653. Estas incluyen ensayos tanto en un único sitio como en dos sitios o “de tipo sándwich” de los tipos no competitivos, así como en los ensayos de unión competitiva tradicionales. Estos ensayos también incluyen unión directa de un anticuerpo marcado con un biomarcador diana.

Los ensayos de tipo sándwich están entre los ensayos más útiles y más habitualmente usados. Existen varias variaciones de la técnica de ensayo de tipo sándwich. Brevemente, en un ensayo directo típico, un anticuerpo no marcado se inmoviliza en un sustrato sólido y la muestra a ensayar se pone en contacto con la molécula unida. Después de un periodo adecuado de incubación, durante un periodo de tiempo suficiente para permitir la formación de un complejo de anticuerpo-antígeno, se añade después un segundo anticuerpo específico del antígeno, marcado con una molécula indicadora capaz de producir una señal detectable y se incuba, permitiendo un tiempo suficiente para la formación de otro complejo de anticuerpo-antígeno-anticuerpo marcado. Cualquier material que no haya reaccionado se retira por lavado, y se determina la presencia del antígeno mediante observación de una señal producida por la molécula indicadora. Los resultados pueden ser cualitativos, mediante observación sencilla de la señal visible, o pueden cuantificarse comparando con una muestra de control que contiene cantidades conocidas de biomarcador.

Las variaciones del ensayo directo incluyen un ensayo simultáneo, en el que tanto la muestra como el anticuerpo marcado se añaden simultáneamente al anticuerpo unido. Los expertos en la materia conocen bien estas técnicas, incluyendo cualquier variación menor que sea fácilmente evidente. En un ensayo de tipo sándwich directo típico, un primer anticuerpo que tiene especificidad por el biomarcador se une de forma covalente o pasiva a una superficie sólida. La superficie sólida es típicamente vidrio o un polímero, siendo los polímeros más habitualmente usados celulosa, poliácridamida, nylon, poliestireno, cloruro de polivinilo o polipropileno. Los soportes sólidos pueden estar en forma de tubos, perlas, discos de microplacas o cualquier otra superficie adecuada para realizar un inmunoensayo. Los procesos de unión se conocen bien en la técnica y consisten en general en reticulación, unión covalente o adsorción física. El complejo de polímero-anticuerpo que se lava para preparar la muestra de ensayo. Se añade después una alícuota de la muestra para ensayar al complejo de fase sólida y se incuba durante un periodo de tiempo suficiente (por ejemplo 2-40 minutos o durante una noche si es más conveniente) y en condiciones adecuadas (por ejemplo de temperatura ambiente a 40 °C, tal como entre 25 °C y 32 °C, inclusive) para permitir la unión de cualquier subunidad presente en el anticuerpo. Después del periodo de incubación, la fase sólida de subunidad de anticuerpo se lava y se seca y se incuba con un segundo anticuerpo específico para una parte del biomarcador. El segundo anticuerpo se une con una molécula indicadora que se usa para indicar la unión del segundo anticuerpo con el marcador molecular.

Un procedimiento alternativo implica inmovilizar los biomarcadores diana en la muestra y disponer después la diana inmovilizada a anticuerpo específico que puede estar o no marcado con una molécula indicadora. Dependiendo de la cantidad de diana y la fuerza de la señal de molécula indicadora, una diana unida puede ser detectable por marcaje directo con el anticuerpo. Como alternativa, un segundo anticuerpo marcado, específico del primer anticuerpo se expone al complejo de diana-primer anticuerpo para formar un complejo terciario de diana-primer anticuerpo-segundo anticuerpo. El complejo se detecta por la señal emitida por la molécula indicadora. Por “molécula indicadora”, como se usa en la presente memoria descriptiva, se entiende una molécula que, por su naturaleza química, proporciona una señal identificable de forma analítica que permite la detección de anticuerpo unido a antígeno. Las moléculas indicadoras más habitualmente usadas en este tipo de ensayo son enzimas, fluoróforos o moléculas que contienen radionúclidos (es decir radioisótopos) y moléculas quimioluminiscentes.

En el caso de un inmunoensayo enzimático, una enzima se conjuga con el segundo anticuerpo, generalmente por medio de glutaraldehído o peryodato. Como se reconocerá fácilmente, sin embargo, existe una amplia diversidad de diferentes técnicas de conjugación, que están fácilmente disponibles para los expertos en la materia. Las enzimas habitualmente usadas incluyen peroxidasa de rábano rústico, glucosa oxidasa, galactosidasa y fosfatasa alcalina, entre otras. Los sustratos para usar con las enzimas específicas se eligen en general para la producción, tras hidrólisis por la enzima correspondiente, de un cambio de color detectable. Los ejemplos de enzimas adecuadas incluyen fosfatasa alcalina y peroxidasa. También es posible emplear sustratos fluorogénicos, que producen un producto fluorescente en lugar de los sustratos cromogénicos indicados anteriormente. En todos los casos, el anticuerpo marcado con enzima se añade al complejo de primer anticuerpo-marcador molecular, se permite que se una y después se retira por lavado el reactivo en exceso. Después se añade una solución que contiene el sustrato apropiado al complejo de anticuerpo-antígeno-anticuerpo. El sustrato reaccionará con la enzima unida al segundo anticuerpo, proporcionando una señal visual cualitativa, que puede cuantificarse adicionalmente, habitualmente de forma espectrofotométrica, para proporcionar un indicio de la cantidad de biomarcador que estaba presente en la muestra. Como alternativa, pueden acoplarse químicamente compuestos fluorescentes, tales como fluoresceína y rodamina, con anticuerpos sin alterar su capacidad de unión. Cuando se activan por iluminación con luz de una longitud de onda particular, el anticuerpo marcado con fluorocromo adsorbe la energía de luz, induciendo un estado a excitabilidad en la molécula, seguido de emisión de la luz a un color característico visualmente detectable con un

microscopio óptico. Como en el EIA, se permite que el anticuerpo marcado con fluorescencia se una con el complejo del primer anticuerpo-marcador molecular. Después de retirar por lavado el reactivo no unido, el complejo terciario restante se expone después a la luz de la longitud de onda apropiada, la fluorescencia observada indica la presencia del marcador molecular de interés. Las técnicas de inmunofluorescencia y EIA están ambas muy bien establecidas en este campo. Sin embargo, también pueden emplearse otras moléculas indicadoras, tales como radioisótopo, moléculas quimioluminiscentes o bioluminiscentes.

Los procedimientos de la invención incluyen además protocolos que examinan la presencia y/o expresión de ARNm, en una muestra tisular o celular. Se conocen bien procedimientos para la evaluación de ARNm en células e incluyen, por ejemplo, ensayos de hibridación usando sondas de ADN complementarias (tales como hibridación *in situ* usando ribosondas marcadas, transferencia de Northern y técnicas relacionadas) y diversos ensayos de amplificación de ácido nucleico (tales como RT-PCR y otros procedimientos de detección de tipo amplificación, tales como, por ejemplo, ADN ramificado, SISBA, TMA y similares).

En una realización, el nivel de ARNm correspondiente al marcador puede determinarse por formatos tanto *in situ* como *in vitro* en una muestra biológica usando procedimientos conocidos en la técnica. Muchos procedimientos de detección de expresión usan ARN aislado. Para procedimientos *in vitro*, puede utilizarse cualquier técnica de aislamiento de ARN que no seleccione contra el aislamiento de ARNm para la purificación de ARN de células. Véase, por ejemplo, Ausubel y col., Ed., Curr. Prot. Mol. Biol., John Wiley & Sons, NY (1987-1999). Adicionalmente, pueden procesarse fácilmente grandes números de muestras tisulares usando técnicas bien conocidas por los expertos en la materia, tales como, por ejemplo, el proceso de aislamiento de ARN de una única etapa de la Patente de Estados Unidos n.º 4.843.155. El ARNm aislado puede usarse en ensayos de hibridación o amplificación que incluyen, pero sin limitación, análisis de Southern o Northern, análisis de PCR y matrices de sondas. Un procedimiento de diagnóstico preferido para la detección de niveles de ARNm implica poner en contacto el ARNm aislado con una molécula de ácido nucleico (sonda) que puede hibridar con el ARNm codificado por el gen que se detecta. La sonda de ácido nucleico puede ser, por ejemplo, un ADNc de longitud completa, o una parte del mismo, tal como un oligonucleótido de al menos 7, 15, 30, 50, 100, 250 o 500 nucleótidos de longitud y suficientes para hibridar específicamente en condiciones rigurosas con un ARN o ADN genómico que codifica un marcador de la presente invención. Otras sondas adecuadas para uso en los ensayos de diagnóstico de la invención se describen en el presente documento. La hibridación de un ARNm con la sonda indica que el marcador en cuestión se está expresando.

En un formato, el ARNm se inmoviliza en una superficie sólida y se pone en contacto con una sonda, por ejemplo, procesando el ARNm aislado en un gel de agarosa y transfiriendo el ARNm del gel a una membrana, tal como nitrocelulosa. En un formato alternativo, la sonda o las sondas se inmovilizan en una superficie sólida y el ARNm se pone en contacto con la sonda o las sondas, por ejemplo, en una matriz de microplacas de genes Affymetrix. Un experto en la materia puede adaptar fácilmente procedimientos de detección de ARNm conocidos para su uso en la detección del nivel de ARNm codificado por los marcadores de la presente invención.

Aunque la amplificación de moléculas no se requiere en la presente invención como se analiza en la sección de ejemplos, un experto en la materia podría usar procedimientos de amplificación. Un procedimiento alternativo para determinar el nivel de ARNm correspondiente a un marcador de la presente invención en una muestra implica el proceso de amplificación de ácido nucleico, por ejemplo, mediante RT-PCR (la realización experimental expuesta en Mullis, Patente de Estados Unidos n.º 4.683.202 (1987), reacción en cadena de la ligasa, replicación de secuencia autosostenida, Guatelli y col., Proc. Natl. Acad. Sci. USA, Vol. 87, págs. 1874-1878 (1990), sistema de amplificación transcripcional, Kwoh y col., Proc. Natl. Acad. Sci. USA, Vol. 86, pp. 1173-1177 (1989); Q-Beta Replicasa, Lizardi y col., Biol. Technology, Vol. 6, página 1197 (1988); replicación por círculo rodante, Patente de Estados Unidos n.º 5.854.033 (1988); o cualquier otro procedimiento de amplificación de ácido nucleico seguido de la detección de las moléculas amplificadas usando técnicas bien conocidas por los expertos en la materia. Estos esquemas de detección son especialmente útiles para la detección de las moléculas de ácido nucleico si dichas moléculas están presentes en números muy bajos. Como se usa en el presente documento, los cebadores de amplificación se definen como un par de moléculas de ácido nucleico que pueden hibridar con regiones 5' o 3' de un gen (cadenas más y menos, respectivamente, o viceversa) y contienen una región corta entre medias. En general, los cebadores de amplificación son de aproximadamente 10-30 nucleótidos de longitud y flanquean una región de aproximadamente 50-200 nucleótidos de longitud. En condiciones apropiadas y con reactivos apropiados, dichos cebadores permiten la amplificación de una molécula de ácido nucleico que comprende la secuencia de nucleótidos flanqueada por los cebadores.

Para procedimientos *in situ*, no es necesario que el ARNm se aisle de las células antes de su detección. En dichos procedimientos una muestra celular o tisular se prepara/procesa usando procedimientos histológicos conocidos. La muestra se inmoviliza después en un soporte, típicamente un portaobjetos de vidrio, y después se pone en contacto con una sonda que puede hibridar con ARNm que codifica el marcador.

- 5 Como alternativa para realizar determinaciones basadas en el nivel de expresión absoluto del marcador, las determinaciones pueden basarse en el nivel de expresión normalizado del marcador. Los niveles de expresión se normalizan corrigiendo el nivel de expresión absoluto de un marcador por comparación de su expresión con la expresión de un gen que no es un marcador, por ejemplo, un gen constitutivo que se expresa de forma constitutiva. Los genes adecuados para normalización incluyen genes constitutivos, tales como el gen de actina o genes  
10 específicos de células epiteliales. Esta normalización permite la comparación del nivel de expresión en una muestra, por ejemplo, una muestra de paciente, con otra muestra o entre muestras de diferentes fuentes.

- 15 Como alternativa, el nivel de expresión puede proporcionarse como un nivel de expresión relativo. Para determinar un nivel de expresión relativo de un marcador, el nivel de expresión del marcador se determina para 10 o más muestras de muestras biológicas normales frente a enfermedades, preferentemente 50 o más muestras, antes de la determinación del nivel de expresión para la muestra en cuestión. Se determina el nivel de expresión medio de cada uno de los genes ensayados en el número de muestras mayor y esto se usa como un nivel de expresión de línea basal para el marcador. El nivel de expresión del marcador determinado para la muestra de ensayo (nivel absoluto de expresión) se divide después por el valor de expresión medio obtenido para ese marcador. Esto proporciona un nivel de expresión relativo.

- 20 Preferentemente, las muestras usadas en la determinación de línea basal serán de pacientes que no tienen el polimorfismo. La elección de fuente de celular depende del uso del nivel de expresión relativo. El uso de expresión hallada en tejidos normales como una puntuación de expresión media ayuda a validar si el marcador ensayado es específico (frente a células normales). Además, a medida que se acumulan más datos, el valor de expresión medio puede revisarse, proporcionando valores de expresión relativos mejorados basándose en datos acumulados.

#### 25 *Anticuerpos y aptámeros*

En una realización preferida, los anticuerpos y aptámeros se unen específicamente con cada componente de los biomarcadores descritos en el presente documento. Los componentes incluyen las secuencias de ácido nucleico, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas de cada componente en cada biomarcador.

- 30 Los polinucleótidos aptámeros son típicamente ADN fosfodiéster convencional monocatenario (ADNmc). También pueden incorporarse análogos de ADN cercanos al aptámero como se describe posteriormente.

Un procedimiento de descubrimiento de aptámero típico se describe a continuación:

- 35 se sintetiza un polinucleótido que comprende una secuencia aleatoria entre "ramas" que tienen secuencia constante. Las ramas pueden incluir sitios de restricción para clonación conveniente y también pueden actuar como sitios de cebado para cebadores de PCR. La síntesis puede realizarse fácilmente en instrumentos comerciales.

- 40 Las proteínas diana se tratan con el polinucleótido aleatorio. La proteína diana puede estar en solución y después los complejos pueden inmovilizarse y separarse de ácidos nucleicos no unidos mediante el uso de una columna de afinidad de anticuerpos. Como alternativa, la proteína diana podría inmovilizarse antes del tratamiento con el polinucleótido aleatorio.

- 45 Los complejos de proteína diana-polinucleótido se separan del material que no está en complejo y después se separan los polinucleótidos unidos de la proteína diana. El ácido nucleico unido puede después caracterizarse, pero más habitualmente se amplifica, por ejemplo mediante PCR y se repiten las etapas de unión, separación y amplificación. En muchos casos, el uso de condiciones que promueven cada vez más la separación del ácido nucleico de la proteína diana, por ejemplo mayor concentración salina, en el tampón de unión usado en la etapa 2) en iteraciones posteriores, da como resultado la identificación de polinucleótidos que tienen afinidad crecientemente alta por la proteína diana.

- 50 Los ácidos nucleicos que muestran alta afinidad por las proteínas diana se aíslan y caracterizan. Esto se consigue típicamente clonando los ácidos nucleicos usando sitios de restricción incorporados en las ramas, y después secuenciando el ácido nucleico clonado.

La afinidad de los aptámeros por sus proteínas diana está típicamente en el intervalo nanomolar, pero puede ser tan baja como el intervalo picomolar. Es decir  $K_D$  es típicamente 1 pM a 500 nM, más típicamente de 1 pM a 100 nM. Los aptámeros que tienen una afinidad de  $K_D$  en el intervalo de 1 pM a 10 nM también son útiles.

- 55 Pueden sintetizarse polinucleótidos de aptámeros en un sintetizador de ácidos nucleicos disponible en el mercado por procedimientos conocidos en la técnica. El producto puede purificarse por procedimientos de selección de

tamaño o cromatográficos.

Los polinucleótidos de aptámeros son típicamente de aproximadamente 10 a 200 nucleótidos de longitud, más típicamente de aproximadamente 10 a 100 nucleótidos de longitud, aún más típicamente de aproximadamente 10 a 50 nucleótidos de longitud y aún más típicamente de aproximadamente 10 a 25 nucleótidos de longitud. Un intervalo de longitud preferido es de aproximadamente 10 a 50 nucleótidos.

Las secuencias de aptámeros pueden seleccionarse como una secuencia deseada, o poblaciones aleatorias o parcialmente aleatorias de secuencias pueden prepararse y después seleccionarse con respecto a unión específica con una proteína diana deseada mediante ensayo *in vitro*. Puede usarse cualquiera de los ensayos de unión de ácido nucleico-proteína típicos conocidos en la técnica, por ejemplo, transferencia de "Southwestern" usando oligonucleótido marcado o proteína marcada como la sonda. Véase también Patente de Estados Unidos n.º 5.445.935 para un ensayo de polarización de fluorescencia de interacción de proteína-ácido nucleico.

Se consideran conocidos en la técnica nucleótidos apropiados para síntesis de aptámeros y su uso, y reactivos para enlace covalente de proteínas con ácidos nucleicos y su uso.

Un complejo de aptámero-proteína deseado, por ejemplo, complejo de aptámero-trombina puede marcarse y usarse como un agente de diagnóstico *in vitro* de una manera muy similar a cualquier agente de unión a proteína específico, por ejemplo un anticuerpo monoclonal. Por lo tanto, un complejo de aptámero-proteína puede usarse para detectar y cuantificar la cantidad de su proteína diana en una muestra, por ejemplo una muestra de sangre, para proporcionar un diagnóstico de una patología correlacionado con la cantidad de la proteína en la muestra.

Un complejo molecular de aptámero-diana/cebo deseado también puede usarse para captura de imágenes de diagnóstico. En usos de captura de imágenes, los complejos se marcan de modo que puedan detectarse fuera del cuerpo. Son marcadores típicos radioisótopos, habitualmente radioisótopos con semividas cortas. Pueden usarse los radioisótopos de captura de imágenes habituales, tales como <sup>123</sup>I, <sup>124</sup>I, <sup>125</sup>I, <sup>131</sup>I, <sup>99m</sup>Tc, <sup>186</sup>Re, <sup>188</sup>Re, <sup>64</sup>Cu, <sup>67</sup>Cu, <sup>212</sup>Bi, <sup>213</sup>Bi, <sup>67</sup>Ga, <sup>90</sup>Y, <sup>111</sup>In, <sup>18</sup>F, <sup>3</sup>H, <sup>14</sup>C, <sup>35</sup>S o <sup>32</sup>P. También pueden usarse potenciadores de captura de imágenes por resonancia magnética nuclear (RMN), tales como gadolinio-153, para marcar el complejo para detección por RMN. Se consideran conocidos en la técnica procedimientos y reactivos para realizar el marcaje, bien en el polinucleótido o bien en el resto proteico.

En una realización preferida, un anticuerpo o aptámero es específico para cada secuencia génica del biomarcador que comprende un biomarcador que comprende las secuencias génicas: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA020 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura de cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_art (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En una realización preferida, un anticuerpo o aptámero es específico para cada secuencia génica del biomarcador que comprende un biomarcador que comprende las secuencias génicas: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglucono 2 de heparán sulfato (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP,

subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de solutos 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (S. cerevisiae), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

## 20 *Descubrimiento de fármacos*

En otras realizaciones preferidas, las firmas moleculares son útiles para la identificación de nuevos fármacos en el tratamiento de enfermedades y trastornos cardiovasculares.

*Moléculas pequeñas:* los compuestos de ensayo de moléculas pequeñas o compuestos terapéuticos candidatos pueden ser inicialmente miembros de una biblioteca química orgánica o inorgánica. Como se usa en el presente documento, "moléculas pequeñas" se refiere a moléculas orgánicas o inorgánicas pequeñas de peso molecular por debajo de aproximadamente 3.000 Dalton. Las moléculas pequeñas pueden ser productos naturales o miembros de una biblioteca química combinatoria. Debería usarse un conjunto de moléculas diversas para abarcar una diversidad de funciones tales como carga, aromaticidad, enlaces de hidrógeno, flexibilidad, tamaño, longitud de cadena lateral, hidrofobicidad y rigidez. Se conocen en este campo técnicas combinatorias adecuadas para sintetizar moléculas pequeñas, por ejemplo, como se ejemplifica en Obrecht y Villalgordo, Solid-Supported Combinatorial and Parallel Synthesis of Small-Molecular-Weight Compound Libraries, Pergamon-Elsevier Science Limited (1998), e incluyen aquellas tales como las técnicas de síntesis "en paralelo" o "de división y agrupamiento", técnicas de fase sólida y fase en solución y técnicas de codificación (véase, por ejemplo, Czarnik, Curr. Opin. Chem. Bio., 1: 60 (1997)). Además, están disponibles en el mercado varias bibliotecas de moléculas pequeñas.

Las aplicaciones de exploración particulares de la presente invención se refieren al ensayo de compuestos farmacéuticos en investigación farmacológica. El lector se refiere en general al libro de texto convencional "In vitro Methods in Pharmaceutical Research", Academic Press, 1997, y Patente de Estados Unidos n.º 5.030.015). La evaluación de la actividad de los compuestos farmacéuticos candidatos generalmente implica administrar un compuesto candidato, determinar cualquier cambio en la morfología, fenotipo y expresión del marcador, o actividad metabólica de las células y función de las células que sea atribuible al compuesto (en comparación con células no tratadas o células tratadas con un compuesto inerte), y después correlacionar el efecto del compuesto con el cambio observado.

La exploración puede realizarse, por ejemplo, debido a que el compuesto se diseña para tener un efecto farmacológico en ciertos tipos celulares, o debido a que un compuesto diseñado para tener efectos en otra parte puede tener efectos secundarios no pretendidos. Pueden ensayarse dos o más fármacos en combinación (combinando con las células bien de forma simultánea o bien secuencial), para detectar posibles efectos de interacción fármaco-fármaco. En algunas aplicaciones, los compuestos se exploran inicialmente con respecto a toxicidad potencial (Castell y col., págs. 375-410 en "In vitro Methods in Pharmaceutical Research", Academic Press, 1997). La citotoxicidad puede determinarse en primer lugar por el efecto en la viabilidad celular, supervivencia, morfología y expresión o liberación de ciertos marcadores, receptores o enzimas. Los efectos de un fármaco en el ADN cromosómico pueden determinarse midiendo la síntesis o reparación de ADN. La incorporación de [<sup>3</sup>H] timidina o BrdU, especialmente en momentos no programados en el ciclo celular, o por encima del nivel requerido para replicación celular, es coherente con un efecto farmacológico. Los efectos no deseados también pueden incluir tasas poco habituales de intercambio de cromátidas hermanas, determinadas por propagación de metafase. El lector se refiere a A. Vickers (PP 375-410 en "In Vitro Methods in Pharmaceutical Research", Academic Press, 1997) para desarrollo adicional.

En una realización, se proporciona un procedimiento para identificar un agente candidato comprendiendo dicho procedimiento: (a) poner en contacto una muestra biológica de un paciente con el agente candidato y determinar el nivel de expresión de uno o más biomarcadores descritos en el presente documento; (b) determinar el nivel de expresión de un biomarcador o biomarcadores correspondientes en una alícuota de la muestra biológica que no está en contacto con el agente candidato; (c) observar el efecto del agente candidato comparando el nivel de expresión

del biomarcador o los biomarcadores en la alícuota de la muestra biológica en contacto con el agente candidato y el nivel de expresión del biomarcador o los biomarcadores correspondientes en la alícuota de la muestra biológica que no está en contacto con el agente candidato; y (d) identificar dicho agente a partir de dicho efecto observado, en el que una diferencia de al menos 10 % entre el nivel de expresión del gen biomarcador o combinación de genes biomarcadores en la alícuota de la muestra biológica en contacto con el agente candidato y el nivel de expresión del gen biomarcador correspondiente o combinación de genes biomarcadores en la alícuota de la muestra biológica que no está en contacto con el agente candidato es un indicio de un efecto del agente candidato.

En realizaciones preferidas, los efectos del fármaco se correlacionan con la expresión de las firmas moleculares asociadas con un buen pronóstico como se describe en detalle en los ejemplos a continuación.

10 En otra realización, se proporciona un agente candidato derivado por el procedimiento.

En otra realización, se proporciona una preparación farmacéutica que comprende un agente.

En otra realización preferida, un procedimiento para producir un fármaco comprende las etapas del procedimiento (i) sintetizar el agente candidato identificado en la etapa (c) anterior o un análogo o derivado del mismo en una cantidad suficiente para proporcionar dicho fármaco en una cantidad terapéuticamente eficaz a un sujeto; y/o (ii) combinar el candidato farmacológico, el agente candidato identificado en la etapa (c) anterior o un análogo o derivado del mismo con un vehículo farmacéuticamente aceptable.

20 *Vectores, células:* en algunas realizaciones es deseable expresar las biomoléculas que comprenden un biomarcador, en un vector y en células. Las aplicaciones de dichas combinaciones son ilimitadas. Los vectores y células que expresan la o las biomoléculas pueden usarse en ensayos, kits, descubrimiento de fármacos, diagnóstico, pronóstico y similares. Las células pueden ser células madre aisladas de la médula ósea como una célula progenitora, o células obtenidas de cualquier otra fuente, tal como por ejemplo ATCC.

“Célula progenitora derivada de médula ósea” (BMDC) o “célula madre derivada de médula ósea” se refiere a una célula madre primitiva con la maquinaria para autorrenovación activa de forma constitutiva. Se incluyen en esta definición células madre que son totipotentes, pluripotentes y precursoras. Una “célula precursora” puede ser cualquier célula en una ruta de diferenciación celular que sea capaz de diferenciarse en una célula más madura. Como tal, la expresión “población de células precursoras” se refiere a un grupo de células capaz de desarrollarse a una célula más madura. Una población de células precursoras puede comprender células que son totipotentes, células que son pluripotentes y células que están restringidas a linaje de células madre (es decir células capaces de desarrollarse a menos linajes que todos los hematopoyéticos, o a, por ejemplo, solamente células de linaje eritroide). Como se usa en el presente documento, la expresión “célula totipotente” se refiere a una célula capaz de desarrollarse a todos los linajes de células. De forma similar, la expresión “población totipotente de células” se refiere a una composición de células capaz de desarrollarse a todos los linajes de células. También como se usa en el presente documento, la expresión “célula pluripotente” se refiere a una célula capaz de desarrollarse a una diversidad (aunque no todos) de linajes y son al menos capaces de desarrollarse a todos los linajes hematopoyéticos (por ejemplo, linajes linfoide, eritroide y trombocítico). Las células madre derivadas de médula ósea contienen dos tipos bien caracterizados de células madre. Las células madre mesenquimatosas (MSC) normalmente forman condrocitos y osteoblastos. Las células madre hematopoyéticas (HSC) son de origen mesodérmico lo que normalmente da lugar a células de la sangre y el sistema inmunitario (por ejemplo, linajes eritroide, de granulocitos/macrófagos, megacariocitos y linfoide). Además, también se ha mostrado que las células madre hematopoyéticas tienen el potencial de diferenciarse en las células del hígado (incluyendo hepatocitos, células del conducto biliar), pulmón, riñón (por ejemplo, células epiteliales tubulares renales y parénquima renal), tracto gastrointestinal, fibras del músculo esquelético, astrocitos del SNC, neuronas de Purkinje, músculo cardíaco (por ejemplo, cardiomiocitos), endotelio y piel.

En una realización preferida, un procedimiento para identificar compuestos terapéuticos candidatos comprende cultivar células que expresan al menos una biomolécula seleccionada de: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA020 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura de cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citotóxicos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito fuertemente similar a Up\_0010813421),

233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_art (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas, con un agente terapéutico candidato; identificar agentes terapéuticos candidatos que modulan la expresión de las biomoléculas e identificar un agente terapéutico candidato. Preferentemente, un agente terapéutico candidato comprende moléculas orgánicas, moléculas inorgánicas, vacunas, anticuerpos, moléculas de ácido nucleico, proteínas, péptidos y vectores que expresan moléculas de ácido nucleico.

En una realización preferida, un procedimiento para identificar compuestos terapéuticos candidatos comprende cultivar células que expresan al menos una biomolécula seleccionada de: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano 2 de heparán sulfato (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 // MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de solutos 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNC DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas, con un agente terapéutico candidato; identificar agentes terapéuticos candidatos que modulan la expresión de las biomoléculas e identificar un agente terapéutico candidato. Preferentemente, un agente terapéutico candidato comprende moléculas orgánicas, moléculas inorgánicas, vacunas, anticuerpos, moléculas de ácido nucleico, proteínas, péptidos y vectores que expresan moléculas de ácido nucleico.

Dichos compuestos son útiles, por ejemplo, como compuestos terapéuticos candidatos para el tratamiento de enfermedad cardíaca, trastornos cardíacos y afecciones de los mismos. Por lo tanto, se incluyen en el presente documento procedimientos para explorar con respecto a compuestos terapéuticos candidatos para el tratamiento de, por ejemplo, miocarditis, Enfermedad Cardíaca Coronaria, angina, Síndrome Coronario Agudo, Aneurisma y Disección Aórtica, arritmias, Cardiomiopatía, Enfermedad Cardíaca Congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis y similares. Los procedimientos incluyen administrar el compuesto a un modelo de la afección, por ejemplo, poner en contacto un modelo celular (*in vitro*) con el compuesto, o administrar el compuesto a un modelo animal de la afección, por ejemplo, un modelo animal de una afección asociada con enfermedad cardíaca. El modelo se evalúa después con respecto a un efecto del compuesto candidato en el resultado clínico en el modelo y puede considerarse un compuesto terapéutico candidato para el tratamiento de la afección. Dichos efectos pueden incluir efectos clínicamente relevantes, dolor reducido; aumento de la esperanza de vida; y así sucesivamente. Dichos efectos pueden determinarse en una escala macroscópica o microscópica. Los compuestos terapéuticos candidatos identificados por estos procedimientos pueden verificarse adicionalmente, por ejemplo, mediante administración a sujetos humanos en un ensayo clínico.

Las biomoléculas pueden expresarse a partir de uno o más vectores. Un "vector" (denominado en ocasiones "vehículo" de suministro génico o transferencia génica) se refiere a una macromolécula o complejo de moléculas que comprende un polinucleótido para suministrar a una célula hospedadora, bien *in vitro* o bien *in vivo*. El polinucleótido para suministrar puede comprender una secuencia codificante de interés en la terapia génica. Los vectores incluyen, por ejemplo, vectores víricos (tales como adenovirus ("Ad"), virus adenoasociados (VAA) y retrovirus), liposomas y otros complejos que contienen lípidos, y otros complejos macromoleculares capaces de mediar en el suministro de

un polinucleótido a una célula hospedadora. Los vectores también pueden comprender otros componentes o funcionalidades que modulan además el suministro génico y/o la expresión génica, o que proporcionan de otro modo propiedades beneficiosas a las células diana. Como se describe y se ilustra en más detalle posteriormente, dichos otros componentes incluyen, por ejemplo, componentes que influyen en la unión o la dirección a células (incluyendo componentes que median en la unión específica de tipo celular o tejido); componentes que influyen en la captación del ácido nucleico del vector por la célula; componentes que influyen en la localización del polinucleótido dentro de la célula después de la captación (tales como agentes que median en la localización nuclear); y componentes que influyen en la expresión del polinucleótido. Dichos componentes también podrían incluir marcadores, tales como marcadores detectables y/o seleccionables que pueden usarse para detectar o seleccionar células que han captado y expresan el ácido nucleico suministrado por el vector. Dichos componentes pueden proporcionarse como un elemento natural del vector (tal como el uso de ciertos vectores víricos que tienen componentes o funcionalidades que median en la unión y captación) o pueden modificarse vectores para proporcionar dichas funcionalidades. Otros vectores incluyen los descritos en Chen y col; *BioTechniques*, 34: 167-171 (2003). Se conocen en la técnica y están disponibles en general una amplia diversidad de dichos vectores.

### 15 Kits

En otra realización preferida, se proporciona un kit que comprende una cualquiera o más de las biomoléculas que comprenden: 232669\_at (factor inducible por hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA020 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura de cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia del receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_art (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE: 4821815), 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

En otra realización preferida, se proporciona un kit que comprende una cualquiera o más de las biomoléculas que comprenden: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano 2 de heparán sulfato (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 // MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14 TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA020, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de solutos 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNC IMAGE: 4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDCC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE),

RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de las mismas.

5 Para su uso en las aplicaciones descritas o sugeridas anteriormente, los kits o artículos de fabricación también son útiles en la invención. Dichos kits pueden comprender un medio de vehículo que se compartimentaliza para recibir en confinamiento estrecho uno o más medios recipientes tales como viales, tubos y similares, comprendiendo cada uno de los medios recipientes uno de los elementos separados para usar en el procedimiento. Por ejemplo, uno de los medios recipientes puede comprender una sonda que está o puede estar marcada de forma detectable. Cuando el kit utiliza hibridación de ácido nucleico para detectar el ácido nucleico diana, el kit también puede tener recipientes que contienen nucleótido o nucleótidos para amplificación de la secuencia de ácido nucleico diana y/o un recipiente que comprende un medio indicador, tal como una proteína de unión a biotina, tal como avidina o estreptavidina, unida a una molécula indicadora, tal como un marcador enzimático, fluorescente o radioisotópico.

15 El kit comprenderá típicamente el recipiente descrito anteriormente y uno o más recipientes adicionales que comprenden materiales deseables desde un punto de vista comercial y de usuario, incluyendo tampones, diluyentes, filtros, agujas, jeringas y prospectos con instrucciones para su uso. Un marcador puede estar presente en el recipiente para indicar que la composición se usa para una terapia específica o aplicación no terapéutica, y también puede indicar instrucciones para uso *in vivo* o *in vitro*, tales como las descritas anteriormente.

20 Los kits tienen varias realizaciones. Una realización típica es un kit que comprende un recipiente, un marcador en dicho recipiente y una composición contenida dentro de dicho recipiente; en el que la composición incluye un anticuerpo primario que se une con las biomoléculas de cada firma molecular e instrucciones para usar el anticuerpo para evaluar la presencia de biomoléculas en al menos un tipo de célula de mamífero. El kit puede comprender además un conjunto de instrucciones y materiales para preparar una muestra tisular y aplicar un anticuerpo y sonda a la misma sección de una muestra tisular. El kit puede incluir un anticuerpo tanto primario como secundario, en el que el anticuerpo secundario se conjuga con un marcador, por ejemplo, un marcador enzimático.

25 Otra realización es un kit que comprende un recipiente, una etiqueta en dicho recipiente y una composición contenida dentro de dicho recipiente; en el que dicha composición incluye un polinucleótido que hibrida con un complemento de los polinucleótidos en condiciones rigurosas, la etiqueta en dicho recipiente indica que la composición puede usarse para evaluar la presencia de una firma molecular en al menos un tipo de célula de mamífero, e instrucciones para usar el polinucleótido para evaluar la presencia de biomolécula de ARN o ADN en al menos un tipo de célula de mamífero.

30 Otros componentes opcionales en el kit incluyen micromatrices, uno o más tampones (por ejemplo, tampón de bloqueo, tampón de lavado, tampón de sustrato, etc.), otros reactivos tales como sustrato (por ejemplo, cromógeno) que se altera químicamente por un marcador enzimático, solución de recuperación de epítomos, muestras de control (controles positivos y/o negativos), portaobjeto o portaobjetos de control, etc.

35 Se ilustran realizaciones de composiciones y procedimientos de la invención en los siguientes ejemplos. Estos ejemplos se proporcionan para fines ilustrativos y no se consideran limitaciones del alcance de composiciones y procedimientos de la invención.

### Ejemplos

#### 40 **Ejemplo 1: biomarcadores transcriptómicos para evaluación de riesgos individual en insuficiencia cardíaca de nueva aparición**

La evaluación de riesgos y predicción de pronóstico en primera presentación precisos son cruciales para la asignación apropiada de terapia, supervisión y tratamientos de pacientes. En este estudio, se presenta un biomarcador basado en transcriptoma (BBT) que se ha obtenido de una única biopsia endomiocárdica (BEM) y que predice el resultado clínico a largo plazo de pacientes con cardiomiopatía dilatada idiopática (CMDI) con precisión muy alta.

#### *Materiales y procedimientos*

45 **Abreviaturas:** NYHA, clasificación de la Asociación del Corazón de Nueva York; FE VI, fracción de eyección ventricular izquierda, DDIVL, dimensión diastólica interna ventricular izquierda; PAP, presión arterial pulmonar; PECP, presión de enclavamiento capilar pulmonar.

55 **Pacientes:** se recogieron BEM de pacientes que se refirieron al Hospital Johns Hopkins entre 1997 y 2006 para evaluación de cardiomiopatía (n = 35012). Se mantuvo una base de datos clínica de resultado de pacientes simultáneamente durante un periodo de 10 años comenzando en 1997. Todos los pacientes proporcionaron consentimiento informado por escrito para recogida de muestras y abstracción de diagrama médico. Se obtuvieron BEM transvenosas del septo ventricular derecho para análisis de micromatrices como se ha descrito previamente (Felker GM, Thompson RE, Hare JM y col., N Engl J Med 2000; 342: 1077-8). Para evitar posibles factores de

confusión específicos de enfermedad, se seleccionaron muestras de un único tipo frecuente de cardiomiopatía (CMDI). CMDI fue un diagnóstico de exclusión después de trabajo histológico exhaustivo sin ninguna señal patológica detectable. Dentro de un depósito de 180 muestras de CMDI, se seleccionaron biopsias de una manera con control de casos basándose en los extremos fenotípicos en la supervivencia de la cohorte. Un grupo con buen pronóstico (BP, n = 25) se definió como con supervivencia sin acontecimientos durante al menos 5 años después de la presentación inicial con síntomas de insuficiencia cardíaca; un grupo con mal pronóstico (MP, n = 18) experimentó un acontecimiento en los primeros 2 años. Los acontecimientos incluyeron muerte (n = 14), requisito de dispositivo de asistencia ventricular izquierdo (n = 2) o trasplante cardíaco (n = 2).

*Procesamiento de biopsias:* las BEM se congelaron inmediatamente de forma instantánea en nitrógeno líquido para almacenamiento en un biodeposito. Todas las etapas de aislamiento y procesamiento de ARN se realizaron de acuerdo con las directrices de MIAME (información mínima acerca de un experimento de micromatrices). Las muestras tisulares (diámetro promedio -2 mm) se homogeneizaron con el molino mezclador MM 301 (Retsch, Cat. n.º 85120). Se usó reactivo de Trizol junto con el sistema de purificación de ARN total Micro-a-Midi (Invitrogen, Cat. n.º 12183-018) para extracción de ARN total (tasa de éxito: 97 % de muestras). Se midió la concentración e integridad de ARN total con el Bioanalizador Agilent 2100. Todas las muestras de ARN mostraron ARN ribosómico 28S y 18S intacto en electroforesis en gel de agarosa desnaturizante y las lecturas de absorbancia a 260/280 nm quedaron en el intervalo aceptable de 1,8-2,1. Se aisló una cantidad promedio de 586 ng de ARN total y se preprocesó con el Sistema de Amplificación y Marcaje de ARN de Biotina Ovation (NuGen, Cat. n.º 2300-12).

*Hibridación de micromatrices:* se hibridaron muestras con la Matriz de Genoma Humano U133 Plus 2.0 de Affymetrix sin etapa de amplificación adicional. Los experimentos de micromatrices se valoraron como exitosos cuando el aislamiento de ARN y la hibridación de micromatrices cumplieron todos los índices de control de calidad como se especifica en las Directrices de Affymetrix para Evaluar la Calidad de Muestras y Matrices. El fondo promedio y ruido de todas las microplacas se registraron dentro de intervalos aceptables y las eficacias de hibridación fueron similares para todas las muestras.

*Análisis estadístico:* los valores de intensidad sin procesar de hibridación de micromatrices se normalizaron con Promedio de Multimatrices Robusto (RMA) implementado en el paquete R para cálculo estadístico (disponible en [www.R-project.com](http://www.R-project.com)). En la siguiente etapa, se usó Análisis de Significación de Micromatrices (SAM) para identificar diferencias específicas de fenotipo en la expresión génica. SAM define la significación con el valor q, un valor p ajustado para múltiples comparaciones (Storey J. Journal of the Royal Statistical Society 2007: 64: 479-498). Para el desarrollo de un BBT, se usó Análisis de Predicción de Micromatrices (PAM) para crear un clasificador en un conjunto de entrenamiento (que contiene 2/3 de los datos, n = 29), con validación posterior en un conjunto de ensayo (que contiene 1/3 de los datos, n = 14). La precisión general se evaluó después de 50 particiones aleatorias. Para ensayar con respecto a condiciones de línea basal equilibradas de subgrupos en la cohorte (conjunto de entrenamiento y ensayo), se usó un ensayo de t y ANOVA de una vía, o ensayo de suma de rangos de Mann-Whitney y ANOVA de una vía de Kruskal-Wallis en rangos si se requiere.

#### Resultados

La capacidad de distinguir pacientes que mejorarán su estado funcional de los que continuarán a colapso circulatorio y requieren trasplante cardíaco o colocación de DAVI, sigue siendo un reto clínico importante.

*Características de pacientes:* se analizaron un total de 43 BEM con tecnología de micromatrices para identificar diferencias específicas de fenotipo en la expresión génica y desarrollar un BBT de pronóstico. La Tabla 1 contiene las condiciones de línea basal de todos los pacientes. La tabla se dividió en 4 subgrupos de acuerdo con el resultado clínico y las subdivisiones usadas para análisis de PAM, para excluir posible sesgo que pueda estar provocado por parámetros de riesgo desequilibrados en las cohortes de entrenamiento y ensayo. No hubo ninguna diferencia significativa en la edad, el sexo, la función ventricular, la hemodinámica o la terapia farmacológica entre los subgrupos. La población general con CMDI se presentó a una edad promedio de  $46 \pm 15$  años, con sobrerrepresentación ligera del sexo masculino (67 %). Todos los subgrupos estuvieron en un estadio de NYHA avanzado de  $2,6 \pm 0,7$  con fracción de eyección (FE) gravemente comprometida de  $23 \pm 13$  %, DDIIV promedio de  $6,1 \pm 1,5$  cm y presión de enclavamiento capilar pulmonar (PECP) de  $2 \pm 1,20$  kPa.

Tabla 1

|                | Buen pronóstico<br>(n=25) | Mal pronóstico<br>(n=18) |
|----------------|---------------------------|--------------------------|
| Edad           | 46±15                     | 48±17                    |
| Hombres, n (%) | 17(68 %)                  | 12(67 %)                 |

(continuación)

|  | Buen pronóstico<br>(n=25) | Mal pronóstico<br>(n=18) |
|--|---------------------------|--------------------------|
| NYHA, n (%)                                    |                           |                          |
| I  | 1(4 %)                    | 1(5 %)                   |
| II   | 13(52 %)                  | 6(33 %)                  |
| III  | 10(40 %)                  | 8(44 %)                  |
| IV   | 1(4 %)                    | 4(22 %)                  |
| FE VI, %                                       | 24±13                     | 23±13                    |
| DDIVI, cm                                      | 6,4±1                     | 6,3±2                    |
| PAP, kPa                                       |                           |                          |
| Sistólica                                      | 4,79±1,73                 | 5,45±1,73                |
| Diastólica                                     | 2,13±0,80                 | 2,66±1,46                |
| Presión de enclavamiento capilar pulmonar, kPa | 1,73±0,93                 | 2,39±1,33                |
| Medicaciones, n (%)                            |                           |                          |
| β-Antagonista                                  | 17(68 %)                  | 13(72 %)                 |
| Inhibidor de ACE                               | 17(68 %)                  | 12(67 %)                 |
| Antagonista de aldosterona                     | 4(16 %)                   | 4(22 %)                  |
| Diurético                                      | 17(68 %)                  | 15(83 %)                 |
| Terapia inotrópica intravenosa                 | 0                         | 0                        |

5 *Análisis de micromatrices:* se aisló un promedio de  $568 \pm 92$  ng de ARN total de todas las BEM y se ensayaron con el Bioanalizador Agilent 2100, revelando alta integridad y pureza de ARN de todas las muestras con bandas  
10 uniformes de ARN 18S y 28S (Figura 1). Hubo 46 genes sobreexpresados de forma significativa en pacientes que se recuperaron de insuficiencia cardíaca ( $q < 5 \%$ ,  $FC > 1,2$ , Tabla 1) como se determinó con SAM. Se usó PAM para evaluar el valor predictivo de este conjunto de genes. Para conseguir esto y ensayar con respecto a validez, los pacientes se asignaron a conjuntos de entrenamiento que consistían en 2/3 de las muestras ( $n = 29$ ) y conjuntos de ensayo que contenían 1/3 de las muestras ( $n = 14$ ; Figura 2). Este enfoque dio como resultado un "centroide reducido más próximo" de 45 genes, que distinguió con alta precisión entre pacientes de alto riesgo y los que tenían un excelente pronóstico.

Tabla 2: 46 genes sobreexpresados de forma significativa en pacientes con insuficiencia cardíaca con buen resultado de pronóstico ( $q < 5 \%$ ,  $FC > 1,2$ ).

| ID de affy  | Título del gen  | Factor de Cambio |
|-------------|---|------------------|
| 232669_at   | Factor inducible por hipoxia 3, subunidad alfa                              | 1,8              |
| 214451_at   | familia de vehículo de soluto 26, miembro 10                                | 1,8              |
| 243482_at   | Tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1 | 1,8              |
| 226210_s_at | expresado por vía materna 3   | 1,7              |
| 232159_at   | Tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1 | 1,7              |
| 233026_s_at | Que contiene dominio PDZ 2  | 1,6              |

ES 2 653 851 T3

(continuación)

| <i>ID de affy</i> | <i>Título del gen</i>   | <i>Factor de Cambio</i> |
|-------------------|---|-------------------------|
| 21199b_s_at       | Gen hipotético de proteína de tipo KIAA0220 LOC 283846  | 1,6                     |
| 243774_at         | mucina 20, asociada a superficie celular  | 1,6                     |
| 242551_at         | Fase abierta de lectura 1 de cromosoma 18   | 1,6                     |
| 244548_at         | Proteína activadora de Rho GTPasa 26  | 1,6                     |
| 244208_at         | Supresor de punto de control 1  | 1,6                     |
| 239984_at         | Canal de sodio, abierto por tensión, tipo VII, alfa   | 1,6                     |
| 230683_at         | ADNC:FLJ20892 fis, clon ADKA03430   | 1,5                     |
| 241869_at         | apolipoproteína L, 6  | 1,5                     |
| 241597_at         | Repeticiones de dipéptido de arginina-ácido glutámico (RE)  | 1,5                     |
| 235887_at         | Homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora ( <i>C. elegans</i> ) | 1,5                     |
| 229957_at         | proteína transmembrana 91   | 1,5                     |
| 223546_x_at       | Tipo LUC7L ( <i>S. cerevisiae</i> )   | 1,                      |
| 239567_at         | Proteína activadora de Rho GTPasa 10  | 1,5                     |
| 242194_at         | Culina 4A   | 1,5                     |
| 1558525_at        | proteína hipotética LOC283901   | 1,4                     |
| 227178_at         | repetición de triplete CUG, proteína de unión a ARN 2   | 1,4                     |
| 228198_s_at       | Proteína ribosómica mitocondrial S9   | 1,4                     |
| 202379_s_at       | secuencia de reconocimiento de tumor por linfocitos citolíticos naturales                               | 1,4                     |
| 224260_at         | clon de ADNC  | 1,4                     |
| 238643_at         | Neuroblastoma, supresión de tumorigenicidad 1   | 1,4                     |
| 232253_at         | homólogo de RAD50 ( <i>S. cerevisiae</i> )  | 1,4                     |
| 227968_at         | Que contiene dominio 7 de enfermedad de Parkinson 1   | 1,4                     |
| 233197_at         | tipo kelch 9 ( <i>Drosophila</i> )  | 1,4                     |
| 244512_at         | Locus transcrito, fuertemente similar a XP_001081342.1  | 1,4                     |
| 233443_at         | proteína hipotética LOC389362   | 1,4                     |
| 231275_at         | proteína FLJ42875   | 1,4                     |
| 226419_s_at       | proteína hipotética LOC645460   | 1,4                     |
| 201221_s_at       | polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña   | 1,4                     |
| 209354_at         | miembro de la familia del receptor de factor de necrosis tumoral 14                                     | 1,4                     |
| 226571_s_at       | Tipo receptor de proteína tirosina fosfatasa S  | 1,4                     |
| 220728_at         | EST   | 1,3                     |
| 203071_at         | dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto                                     | 1,3                     |
| 213946_s_at       | tipo obscurina 1, similar a isoforma de titina N2-B   | 1,3                     |

(continuación)

| <i>ID de affy</i> | <i>Título del gen</i>   | <i>Factor de Cambio</i> |
|-------------------|---|-------------------------|
| 201394_s_at       | proteína de motivo de unión a ARN 5   | 1,3                     |
| 203748_x_at       | motivo de unión a ARN, proteína de interacción monocatenaria 1                      | 1,3                     |
| 223147_s_at       | dominio de repetición de WD 33  | 1,3                     |
| 2I3773T_x_at      | familia de dominio NOL1/NOP2/Sun, miembro 5   | 1,3                     |
| 1560049_at        | repetición de triplete CUG, proteína de unión a ARN 2                               | 1,3                     |
| 243974_at         | clon de ADNC IMAGE 4821815  | 1,2                     |
| 201510_at         | factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial) | 1,2                     |

5 *Validación:* para obtener el rendimiento general del biomarcador de los inventores, se realizaron 50 particiones aleatorias en conjuntos de entrenamiento y ensayo, revelando una sensibilidad general del 74 % (IC al 95 %: 69 %-79 %) y una especificidad general de 90 % (IC al 95 %: 87 %-93 %). El valor predictivo positivo fue de 85 % (IC al 95 %: 80 %-89 %), mientras que el valor predictivo negativo fue de 82 % (IC al 95 %: 78 %-86 %). La relación de probabilidad logarítmica fue de 3,3. Para evaluar la influencia de diversos factores hospedadores, los 43 sujetos de CMDI se dividieron en múltiples estratos de dos vías basándose en la edad ( $\geq$  o  $<$ 50 años), fracción de eyección ( $\geq$  o  $<$ 15 %) y uso de inótrópos para ensayar si la precisión del marcador de pronóstico se vio afectada por parámetros de línea basal. La sensibilidad y especificidad del BBT se evaluó después, determinando la proporción de sujetos clasificados de forma correcta dentro de cada estrato. Notablemente, la precisión predictiva fue perfecta (sensibilidad y especificidad: 100 %) tanto en pacientes de más de 50 años como en pacientes con fracción de eyección por encima del 15 %. Ninguno de los pacientes recibió medicación intróptica.

10  
15 La firma molecular se ilustra en un mapa de calor, que se creó por distancia euclídea (Figura 3). Este enfoque independiente de agrupamiento no supervisado confirmó adicionalmente la robustez del conjunto descubierto de genes con muy clara distinción de muestras con MP de muestras con BP. Las rutas con implicación importante fueron mecanismos de transporte iónico (13 %), desarrollo neuromuscular (10 %), unión a proteínas (15 %) y transcripción (26 %) ([www.gene-ontology.org](http://www.gene-ontology.org), Figura 4).

20 *Predicción de recuperación en la función ventricular izquierda:* el resultado clínico mejorado se asocia con frecuencia con la recuperación de la función VI. En consecuencia, se ensayó la hipótesis de que la firma de BP se asociaría con fracción de eyección mejorada. Entre la muestra de estudio ( $n = 43$ ), todos los sujetos en los que estuvieron disponibles datos de ecocardiografía emparejados de línea basal y seguimiento ( $n = 17$ ) se seleccionaron y caracterizaron como BP o MP según su BBT (Figura 5). Los pacientes clasificados como BP experimentaron una mejora sustancial de FE de  $23 \pm 3$  % a  $42 \pm 5$  % ( $p = 0,0009$ ), mientras que los pacientes con una firma molecular de MP continuaron teniendo una FE reducida a  $20 \pm 3$  % y  $21 \pm 3$  % en línea basal y seguimiento, respectivamente.

#### *Análisis:*

30 Una de las aplicaciones más valiosas de la información genómica ha demostrado ser la predicción clínica. Mientras que el patrón de genes expresados diferencialmente proporciona información acerca de la etiología de enfermedad, se ha usado también para el desarrollo de biomarcadores. Este enfoque ha sido altamente exitoso en enfermedad neoplásica y está apareciendo en una diversidad de otros procesos de enfermedad. Aquí, se ha buscado identificar un BBT que predijera el resultado clínico de pacientes con insuficiencia cardíaca de nueva aparición. Los inventores han desarrollado un biomarcador de pronóstico altamente preciso, que es capaz de diferenciar pacientes con trayectoria muy mala de los que tienen pronóstico a largo plazo excelente. Resulta importante que este estudio ha usado biopsias de pacientes que presentan un estadio de enfermedad temprano, sin amplificación adicional de ARN.

35 Por dos razones principales, este enfoque proporciona mejor precisión y resultados más representativos en comparación con estudios previos: en primer lugar, se están presentando diferencias transcriptómicas iniciales sin "ruido" de cambios compensatorios inespecíficos que pueden activarse durante el progreso de enfermedad. En segundo lugar, se evitó el sesgo de amplificación que puede producirse debido a diferentes preferencias de unión de cebadores.

40 Además de proporcionar un biomarcador clínico nuevo y útil, este estudio ha ofrecido información acerca de diferencias individuales en la expresión génica que pueden estar implicadas de forma mecánica en un mal resultado de insuficiencia cardíaca. Más notablemente, se ha abordado una necesidad no satisfecha importante en el tratamiento de la insuficiencia cardíaca, la capacidad de evaluar con precisión el pronóstico del paciente. Aunque están apareciendo biomarcadores de valor pronóstico, estos marcadores no se dicotomizan sustancialmente entre

45 pacientes con resultados muy variables. Los pacientes con características casi idénticas en el momento de

presentación, que recibieron terapia idéntica, pueden tener pronósticos drásticamente diferentes. Aunque algunos pacientes experimentan recuperación completa de su función cardíaca en un periodo de 5 años de promedio, otros progresan al colapso circulatorio en los primeros 2 años de presentación y requieren intervenciones terapéuticas agresivas, tales como asistencia circulatoria mecánica o trasplante cardíaco.

5 El BBT desarrollado, que contiene un centrote reducido más próximo de 45 genes, predijo el fenotipo de muestras con muy alta especificidad (90 %) lo que sugiere su capacidad en particular como un ensayo de descarte y puede usarse además de la evaluación de riesgos convencional actual. Además, pueden incluirse resultados del BBT de los inventores en la evaluación de una puntuación de prioridad en listas de trasplante. En combinación con datos clínicos y valores de laboratorio, el BBT de los inventores es prometedor para proporcionar una mejor orientación en el tratamiento de pacientes.

10 Es importante que el BBT de pronóstico de los inventores rindió perfectamente en pacientes que tenían más de 50 años de edad o que tenían una FE mayor del 15 %. A este respecto, los marcadores específicos para edad y función podrían considerarse en el futuro. Se requerirán estudios futuros para validar el valor clínico preciso del enfoque de BBT para IC de nueva aparición. Además, los datos de los inventores indican fuertemente un componente molecular que influye en la recuperación de pacientes con insuficiencia cardíaca. De los 46 genes expresados diferencialmente (Tabla 2), todos se sobreexpresaron en pacientes que se recuperaron de insuficiencia cardíaca. Muchos genes sobreexpresados en estos pacientes tienen funciones de unión de ARN o ADN, por ejemplo RBMS1 y WDR33, que desempeñan un papel crucial en la replicación de ADN, transcripción génica, progresión del ciclo celular y apoptosis. Se han identificado diversos genes con funciones reguladoras críticas, es decir: SNRP70; los factores de transcripción nuclear ELF3, CHES1 y RERE; NSUN4, un gen con actividad metiltransferasa y el factor de transcripción HIF3A. HIF3A se ha analizado como un inductor de transportadores de glucosa, factor de crecimiento endotelial vascular y eritropoyetina similar a HIF1A y HIF2A, mientras que otros han postulado su neutralización con las otras dos subunidades.

15 Los genes CUGBP2, de tipo LUC7 y SEMA3B están implicados en el desarrollo neuromuscular, la orientación de axones y la regulación de la contractilidad cardíaca ([www.geneontology.org/](http://www.geneontology.org/)). También está estrechamente relacionado en su función el gen de tipo obscurina 1 sobreexpresado, un enlazador que estabiliza contactos celulares y orgánulos dentro del citoesqueleto y que se localiza en los discos intercalados en el cardiomiocito adulto. Se supone que el gen de tipo obscurina 1 tiene funciones similares a la obscurina, una proteína multifuncional responsable del ensamblaje de miofibrillas y organización celular de miocitos. Por su interacción con titina y anquirina, así como unión con sarcómero y retículo sarcoplásmico, proporciona los alineamientos requeridos para contracción.

20 Además los inventores han descubierto dos genes regenerativos, Rad50 y SMG6, reguladores clave importantes en la actividad de telomerasa y reparación de ADN. Rad50 es parte del complejo Mrel I/Rad50/Nbs1 (MRN), una unidad funcional que genera bucles t insertando salientes G 3' en extremos teloméricos humanos. Estos bucles t evitan que los extremos de los cromosomas sean reconocidos como ADN dañado y proporcionan un molde para telomerasa y conservación de estabilidad genómica. El aumento de la actividad telomerasa podría explicar un efecto protector contra procesos degenerativos y envejecimiento en el corazón de pacientes con buen pronóstico. Es atractivo especular que la razón de la mejor recuperación puede deberse a aumento de la capacidad regenerativa celular ya que Rad50 puede ser importante para la viabilidad de células madre. Varios genes dentro del BBT de pronóstico de los inventores (SNRP de 70 kDa, tipo Obscurina 1 y motivo de unión a ARN) fueron coherentes con hallazgos de un estudio de 199 muestras humanas de DAVI que indicaban que estos genes están implicados en la recuperación. No hubo ningún solapamiento con los resultados publicados por Lowes y col. (N Engl J Med 2 de mayo de 2002; 346 (18): 1357-65; Lowes BD, y col. J Heart Lung Transplant mayo de 2006; 25 (5): 579-88), que investigaron genes sobreexpresados durante la recuperación de IC. Esto puede indicar que hay importantes diferencias entre los cambios de expresión génica provocados por terapia y características de la expresión génica que pueden predecir resultado clínico. Merece la pena mencionar varias características técnicas de este estudio. Aunque la mayoría de estudios previos han usado tejidos cardíacos obtenidos en estadios avanzados de cardiomiopatía (con frecuencia en el momento de trasplante), los inventores han investigado biopsias de pacientes obtenidas en la primera presentación con insuficiencia cardíaca. Por lo tanto, los biomarcadores obtenidos son muy relevantes para una posible aplicación clínica y sugieren cambios patológicos iniciales en el transcriptoma, que pueden ser causantes de la enfermedad y posibles dianas terapéuticas. Adicionalmente, se ha usado una técnica de procesamiento de ARN altamente eficaz, que permite el análisis de micromatrices sin amplificación adicional, y por lo tanto evitando posible sesgo debido a preferencias de unión diferentes de cebadores y reducción de tiempo y costes con respecto a una aplicación clínica posterior.

25 Aunque la biopsia endomiocárdica es un procedimiento de bajo riesgo, pueden desarrollarse procedimientos alternativos para obtener biomarcadores transcriptómicos. Ha habido pruebas de que el tejido afectado y células sanguíneas en circulación comparten un alto porcentaje de genes comunes. La fácil accesibilidad de PBMC por una sencilla punción venosa hace a esas células muy atractivas para una aplicación clínica.

30 En resumen, los inventores han desarrollado un nuevo biomarcador transcriptómico para pronóstico en insuficiencia cardíaca que puede obtenerse de una única biopsia endomiocárdica con potencial para una aplicación clínica directa. Este enfoque debería mejorar la individualización del cuidado cardíaco y ayudar a identificar pacientes con mayor riesgo de colapso circulatorio en los primeros años de presentación con insuficiencia cardíaca. La implicación de genes en el desarrollo neuromuscular, angiogénesis y mecanismo de reparación de ADN incluyendo actividad

telomerasa dentro del BBT para pacientes con excelente resultado clínico ofrece la posibilidad biológica con respecto a una ventaja de recuperación.

**Ejemplo 2: biomarcador transcriptómico para predecir el resultado clínico y pronóstico de insuficiencia cardíaca usando un biomarcador transcriptómico**

5 Para identificar este biomarcador (firma génica) se recogieron muestras de corazón de pacientes que se sometieron a biopsia cardíaca pronto en su ciclo clínico. Después de esto las muestras se almacenaron en un depósito biológico durante 5-10 años, tiempo durante el cual se determinó el resultado de los pacientes.

Se han recogido muestras de biopsia endomiocárdica de pacientes con cardiomiopatía idiopática en el Hospital Johns Hopkins entre 1997 y 2004 y se almacenaron en nitrógeno líquido. Se eligieron muestras de biopsia de 10 pacientes con cardiomiopatía idiopática de este depósito biológico, 18 pacientes con buen pronóstico y 12 pacientes con mal pronóstico. El mal pronóstico se definió como aparición de muerte o una intervención, concretamente colocación de dispositivo de asistencia ventricular izquierda o trasplante cardíaco, en los primeros 2 años después del diagnóstico, mientras que el buen pronóstico se definió como una supervivencia sin acontecimientos durante más de 5 años después del diagnóstico. Se siguieron para todos los experimentos las directrices de MIAME 15 (Información mínima acerca de un experimento de micromatrices). El tejido se homogeneizó con el instrumento MM301 de Retsch y se aisló ARN total con el sistema de purificación de ARN total Micro a Midi de Invitrogen. Se realizó análisis de micromatrices de ARN total con la Matriz de Genoma Humano U133 Plus 2.0 de Affymetrix. En todas las muestras, tanto el aislamiento de ARN como la hibridación de micromatrices cumplieron todos los índices de control de calidad especificados en las Directrices de Affymetrix para Evaluación de Calidad de Muestras y 20 Matrices.

Los valores de expresión sin tratamiento previo de todas las micropelículas de micromatrices se preprocesaron con Promedio Multimatriz Robusto (RMA) en R. Para hallar un biomarcador transcriptómico para predicción de clase de pronóstico bueno o malo, se usó Análisis de Predicción de Micromatrices (PAM) implementado en R. PAM usa el procedimiento de "centroide reducido más próximo" para predicción de clases. Después de entrenamiento del clasificador (agrupamiento supervisado) en un subconjunto de 13 muestras de pacientes con buen pronóstico y 7 25 muestras de pacientes con mal pronóstico, se encontró un grupo de 9 genes que predecía el resultado clínico con 60 % de sensibilidad y 100 % de especificidad (Figura 6, Figura 10, Tabla 3, Tabla 5). Los mismos 9 genes se usaron para producir un mapa de calor (agrupamiento no supervisado), normalizado por niveles medios de expresión (Figura 8). Cada fila representaba uno de los 9 genes, y cada columna fue una muestra de paciente. Una celda roja representa un gen infraexpresado en un paciente dado en relación con la expresión de gen promedio en 30 todos los pacientes, mientras que una celda azul indica un gen sobreexpresado. El dendrograma en la parte superior es un algoritmo de agrupamiento jerárquico no supervisado que divide muestras en grupos basándose en la similitud de los perfiles de expresión génica. Las "muestras de buen pronóstico" y las "muestras de mal pronóstico" se segregan en dos grupos dominantes. Sin embargo, 3 muestras se clasificaron erróneamente con este enfoque 35 (sensibilidad 75 %, especificidad 100 %). En todos los casos, las clasificaciones erróneas fueron pacientes de mal pronóstico que se clasificaron erróneamente como buen pronóstico. El número de genes se aumentó a 43, para minimizar el error de clasificación errónea (Figura 6, Figura 7) y se desarrolló un clasificador en PAM en el mismo conjunto de entrenamiento (n = 20), que predijo de nuevo con un 60 % de sensibilidad y 100 % de especificidad. Sin embargo, las probabilidades individuales para predicción de clase correcta mejoraron ligeramente, como se indica 40 en la Tabla 5. Con un mapa de calor que incluía esos 43 genes, se consiguió una mayor precisión que con PAM, clasificando erróneamente solamente una muestra de 30 totales (92 % de sensibilidad y 100 % de especificidad) Una muestra (MP-10) se clasificó erróneamente en todos los enfoques, supuestamente un valor atípico (Figura 9). El hecho de que el conjunto descubierto de genes rindió con mucha precisión en todas las muestras, refuerza los resultados de la identificación de biomarcador. De todos los genes que contribuían al biomarcador transcriptómico 45 de 9 genes, solamente uno estaba regulado positivamente de forma significativa en el grupo con mal pronóstico. El conjunto de 43 genes contenía un transcrito no identificado regulado positivamente adicional en el grupo con mal pronóstico (Figura 11). Las Tablas 6 y 7 enumeran todos los genes comprendidos dentro de la firma de 9 y 43 genes y que se sobreexpresaron en el grupo de pacientes con supervivencia a largo plazo. El biomarcador de 43 genes contenía genes que son relevantes para pronóstico, como por ejemplo el polipéptido pesado de miosina 7. Se ha 50 mostrado que sus mutaciones están asociadas con gravedad de insuficiencia cardíaca, en particular en cardiomiopatía dilatada familiar. Los valores atípicos pueden reducirse por entrenamiento del clasificador en un mayor número de muestras investigadas. El biomarcador presentado consiguió precisión de pronóstico significativamente mayor que los factores de riesgo clásicos y esquemas que han sido sugeridos por otros grupos, por ejemplo Índice de Insuficiencia Cardíaca de Seattle.

Tabla 3

| Predicción para umbral = 0                                      |             |                       |      |          |          |          |          |        |          |         |   |
|---|-------------|-----------------------|------|----------|----------|----------|----------|--------|----------|---------|---|
|   |             |                       |      |          |          |          |          |        |          |         |   |
| Nombre de Ajuste: Ajustes 4                                     |             |                       |      |          |          |          |          |        |          |         |   |
| Cuantil de Compensación   | 50          | Valor de Compensación | 0,3  |          |          |          |          |        |          |         |   |
|   | ambos       | Semilla de RNG        | 420  |          |          |          |          |        |          |         |   |
| Distribución Previa (Muestra Previa)                            |             |                       |      |          |          |          |          |        |          |         |   |
| Clase   | 1           | 2                     |      |          |          |          |          |        |          |         |   |
| Probabilidad  | 0,35        | 0,65                  |      |          |          |          |          |        |          |         |   |
| Matriz de Confusión de Predicción de Conjunto de Ensayo         |             |                       |      |          |          |          |          |        |          |         |   |
| Verdadero\Predicho  | 1           | 2                     |      |          |          |          |          |        |          |         |   |
|   | 1           | 3                     |      |          |          |          |          |        |          |         |   |
|   | 2           | 0                     | 5    |          |          |          |          |        |          |         |   |
|   |             |                       |      |          |          |          |          |        |          |         |   |
|   |             |                       |      |          |          |          |          |        |          |         |   |
|   |             |                       |      |          |          |          |          |        |          |         |   |
| Clases Reales, Predichas y Probabilidades Posteriores Predichas |             |                       |      |          |          |          |          |        |          |         |   |
| Etiquetas de Muestra  | BP-10       | BP-18                 | BP-6 | BP-7     | BP-9     | GP-20    | GP-22    | GP-23  | GP-24    | GP-25   |   |
| Etiquetas Cerradas  | 1           | 1                     | 1    | 1        | 1        | 2        | 2        | 2      | 2        | 2       | 2 |
| Predichas   | 2           | 2                     | 1    | 1        | 1        | 2        | 2        | 2      | 2        | 2       | 2 |
| Mal pronóstico  | 0,000842577 | 0,008808878           |      | 0,806138 | 0,999852 | 0,000469 | 0,098695 | 0,0002 | 0,000244 | 0,00194 |   |
| Buen pronóstico   | 0,99157422  | 0,91191122            |      | 0,193861 | 0,000348 | 0,999531 | 0,901305 | 0,9998 | 0,999756 | 0,99805 |   |

## ES 2 653 851 T3

5 Resultado de predicción del conjunto de ensayo ( $n = 10$ ): se usó una firma de 9 genes (Tabla 3). Las etiquetas de muestra se enumeran en la quinta fila desde abajo, seguido de las etiquetas de clase reales y las clases predichas. Se asignaron muestras de mal pronóstico (MP) a la clase 1 - se asignaron muestras de buen pronóstico (BP) a la clase 2. Ocho muestras se clasificaron correctamente (probabilidades entre 73 y 99 %). Solamente dos muestras se clasificaron erróneamente (con una probabilidad mayor del 90 %). Las probabilidades predichas se enumeran para cada clase en las últimas dos líneas.

Tabla 4  
Predicción para umbral = 0

| Predicción para umbral = 0   |             |                       |             |          |          |         |          |          |          |          |   |   |
|--|-------------|-----------------------|-------------|----------|----------|---------|----------|----------|----------|----------|---|---|
| Nombre de ajustes: Ajustes 6                                       |             |                       |             |          |          |         |          |          |          |          |   |   |
| Cuantil de Compensación  | 50          | Valor de Compensación | 0,350 764   |          |          |         |          |          |          |          |   |   |
|  | ambos       | Semilla de RNG        | 42047 3     |          |          |         |          |          |          |          |   |   |
| Distribución Previa (Muestra Previa)                               |             |                       |             |          |          |         |          |          |          |          |   |   |
| Clase  | 1           | 2                     |             |          |          |         |          |          |          |          |   |   |
| Probabilidad   | 0,35        | 0,65                  |             |          |          |         |          |          |          |          |   |   |
| Matriz de Confusión de Predicción de Conjunto de Ensayo(Umbral -0) |             |                       |             |          |          |         |          |          |          |          |   |   |
| Verdadero Predicho   | 1           | 2                     |             |          |          |         |          |          |          |          |   |   |
|  | 1           | 3                     | 2           |          |          |         |          |          |          |          |   |   |
|  | 2           | 0                     | 5           |          |          |         |          |          |          |          |   |   |
| Clases Predichas, Reales y Probabilidades Posteriores Predichas    |             |                       |             |          |          |         |          |          |          |          |   |   |
| Etiquetas de Muestras  | MP-18       | MP-10                 | MP-6        | MP-7     | MP-9     | BP-20   | BP-22    | BP-23    | BP-24    | BP-25    |   |   |
| Etiquetas de Clase   | 1           | 1                     | 1           | 1        | 1        | 2       | 2        | 2        | 2        | 2        | 2 | 2 |
| Etiquetas de Clase predicha  | 2           | 2                     | 1           | 1        | 1        | 2       | 2        | 2        | 2        | 2        | 2 | 2 |
| Mal pronóstico   | 0,011971912 | 1,92842E-06           | 0,995824315 | 0,896715 | 1        | 2,3E-13 | 1,47F-05 | 2,74E-13 | 1,48E-12 | 3,89E-08 |   |   |
| Buen pronóstico  | 0,988028088 | 0,999998072           | 0,004175685 | 0,103285 | 4,51E-08 | 1       | 0,999985 | 1        | 1        | 1        | 1 | 1 |

*Resultado de predicción del conjunto de ensayo (n = 10):* firma de 43 genes (Tabla 4). La descripción de la tabla es la misma que en la tabla 6. Las probabilidades para la clasificación correcta aumentaron ligeramente.

Tabla 5

| Etiquetas de muestra | MP-10       | MP-18       | MP-6        | MP-7     | MP-9      | BP-20    | BP-22    | BP-23    | BP-24    | BP-25    |
|----------------------|-------------|-------------|-------------|----------|-----------|----------|----------|----------|----------|----------|
| Firma de 9 genes     |             |             |             |          |           |          |          |          |          |          |
| Mal pronóstico       | 0,008425774 | 0,08808878  | 0,806138017 | 0,73512  | 0,999652  | 0,000469 | 0,098695 | 0,0002   | 0,000244 | 0,001943 |
| Buen pronóstico      | 0,991574226 | 0,91191122  | 0,193861983 | 0,26488  | 0,000348  | 0,999531 | 0,901305 | 0,9998   | 0,999756 | 0,998057 |
| Firma de 43 genes    |             |             |             |          |           |          |          |          |          |          |
| Mal pronóstico       | 0,011971912 | 1,92842E-06 | 0,995824315 | 0,896715 | 1         | 2,3E-13  | 1,47E-05 | 2,74E-13 | 1,48E-12 | 3,69E-08 |
| Buen pronóstico      | 0,988028088 | 0,999998072 | 0,004175685 | 0,103285 | 4,51 E-08 | 1        | 0,999985 | 1        | 1        | 1        |

La Tabla 5 muestra los resultados de PAM: rendimiento tanto de firma de 9 genes como de firma de 43 genes. Cada muestra marcada con MP se obtuvo de un paciente con mal pronóstico, cada muestra marcada con BP de un paciente con buen pronóstico. La probabilidad predicha para cada clase usando el clasificador desarrollado se enumera en cada fila. Las probabilidades para clasificación correcta aumentaron ligeramente al aumentar el número de genes.

5

**Tabla 6**

| Procedimiento usado     | Sensibilidad | Especificidad |
|-------------------------|--------------|---------------|
| PAM: 9 genes            | 60 %         | 100 %         |
| PAM: 43 genes           | 60 %         | 100 %         |
| Mapa de calor: 9 genes  | 75 %         | 100 %         |
| Mapa de calor: 43 genes | 92 %         | 100 %         |

La Tabla 6 es una visión del conjunto del rendimiento de todos los algoritmos aplicados. Cada procedimiento usado fue capaz de predecir un resultado de mal pronóstico con 100 % de especificidad. Sin embargo, el agrupamiento no supervisado (mapa de calor) predijo con mayor sensibilidad en general y rindió mejor, especialmente cuando se aumentó el número de genes en la firma molecular de 9 a 43 genes.

10

**Tabla 7**

| <b>Sonda</b> | <b>Título del Gen</b>   | <b>Símbolo del Gen</b> | <b>Descripción de Proceso biológico GO</b>  |
|--------------|---|------------------------|---|
| 201655       | Proteoglucono de heparán sulfato 2 (perlecano)                      | HSPG2                  | adhesión celular, localización de proteína  |
| 208976       | Proteína rica en cisteína 2   | CRIP2                  | ---   |
| 209354       | Superfamilia del receptor de factor de necrosis tumoral, miembro 14 | TNFRSF14               | apoptosis, respuesta inmunitaria, señal unida a receptor de superficie celular    |
| 211996       | Proteína de tipo KIAA0220, complejo de poro nuclear                 | LOC23117               | proceso biológico   |
| 213946       | Tipo obscurina 1, similar a isoforma de titina N2-B                 | OBSL1                  | ---   |
| 214951       | Familia de vehículo de soluto 26, miembro 10                        | SLC26A10               | transporte, regulación de alta resolución de señal de proteína Rho                |
| 227968       | Que contiene dominio 7 de enfermedad de Parkinson 1                 | PDDC1                  | ---   |
| 241597       | Repeticiones de dipéptido de arginina-ácido glutámico (RE)          | RERE                   | remodelación de cromatina, transcripción. Importación de sustrato portador de NLS |
| 242551       | Fase abierta de lectura 1 del Cromosoma 18                          | C18orf1                | proceso biológico   |

La Tabla 7 muestra 8 genes regulados positivamente en pacientes con buen pronóstico: la primera columna contiene los ID de Affymetrix para todos los transcritos regulados positivamente del biomarcador de 9 genes. Solamente 1 gen estaba regulado negativamente, concretamente el receptor muscarínico de acetilcolina M3. Se obtuvieron anotaciones para la función biológica del sitio web de Gene Ontology (GO).

15

**Tabla 8**

| <b>Sonda</b> | <b>Título del Gen</b>                                 | <b>Símbolo del Gen</b> | <b>Descripción de Proceso Biológico de GO</b>                             |
|--------------|---|------------------------|---|
| 155845       | LOC401320 hipotética                                  | LOC401320              | ---   |
| 156004       | repetición de triplete CUG, proteína de unión a ARN 2 | CUGBP2                 | procesamiento de ARN, desarrollo de unión neuromuscular, regulación de    |
| 201394       | proteína de motivo de unión a ARN 5                   | RBM5                   | procesamiento de ARN, ciclo celular, regulación negativa de la progresión |
| 201655       | proteoglucono de heparán sulfato 2 (perlecano)        | HSPG2                  | adhesión celular, localización de proteínas                               |

20

(continuación)

| <b>Sonda</b> | <b>Título del Gen</b>   | <b>Símbolo del Gen</b> | <b>Descripción de Proceso Biológico de GO</b>  |
|--------------|---|------------------------|--|
| 202379       | secuencia de reconocimiento de tumor por linfocitos citotóxicos naturales | NKTR                   | plegamiento de proteínas   |
| 202808       | ---   | ---                    | ---  |
| 203071       | dominio sema, dominio de inmunoglobulina (Ig), básico corto               | SEMA3B                 | señalización célula-célula, desarrollo, orientación axónica                              |
| 203748       | motivo de unión a ARN, proteína de interacción monocatenaria 1            | RBMS1                  | replicación de ADN, procesamiento de ARN, regulación de la traducción                    |
| 203981       | casete de unión a ATP, subfamilia D (ALD), miembro 4                      | ABCD4                  | transporte   |
| 204737       | miosina, cadena pesada 6, miosina, cadena pesada 7                        | MYH6 /// MYH7          | contracción de músculo estriado, contracción de músculo                                  |
| 204978       | factor de corte y empalme, rico en arginina/serina 16                     | SFRS16                 | procesamiento de ARNm, corte y empalme de ARN  |
| 206209       | anhidrasa carbónica IV  | CA4                    | metabolismo de compuesto de un carbono, percepción visual, respuesta a                   |
| 207541       | componente de exosoma 10  | EXOSC10                | catabolismo de ARNm, degradación mediada por mutación terminadora, procesamiento de ARNr |
| 207798       | tipo ataxina 2  | ATXN2L                 | proceso biológico  |
| 208978       | proteína rica en cisteína 2   | CRIP2                  | ---  |
| 209354       | superfamilia del receptor de factor de necrosis tumoral, miembro 14       | TNFRSF14               | apoptosis, respuesta inmunitaria, señal unida a receptor de superficie celular           |
| 210628       | proteína de unión a factor de crecimiento transformante beta latente 4    | LTBP4                  | regulación de crecimiento celular, plegamiento de proteínas, desarrollo, regulación de   |
| 211909       | receptor de prostaglandina E 3 (subtipo EP3)                              | PTGER3                 | transcripción, dependiente de ADN, transducción de señales, acoplado a proteína-G        |
| 211996       | proteína de tipo KIAA0220, complejo de poro nuclear                       | LOC23117               | proceso biológico  |
| 212487       | que contiene dominio de parche G 8  | GPATCH8                | proceso biológico  |
| 213946       | tipo obscurina 1, isoforma de titina similar N2-B                         | OBSL1                  | ---  |
| 214951       | familia de vehículo de soluto 26, miembro 10                              | SLC26A10               | transporte, regulación de transducción de señal de proteína Rho                          |
| 220219       | que contiene repetición rica en leucina 37A                               | LRRC37A                | ---  |
| 221071       | ---   | ---                    | ---  |
| 221780       | polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27                             | DDX27                  | ---  |
| 221806       | que contiene dominio SET 5  | SETD5                  | ---  |
| 221833       | peptidasa LON 2, peroxisómica   | LONP2                  | proteólisis, proteólisis dependiente de ATP  |
| 223546       | tipo LUC7 ( <i>S. cerevisiae</i> )  | LUC7L                  | regulación negativa del desarrollo de músculo estriado                                   |
| 224260       | clon de ADNc IMAGE:4478733  | ---                    | ---  |

(continuación)

| <b>Sonda</b> | <b>Título del Gen</b>  | <b>Símbolo del Gen</b> | <b>Descripción de Proceso Biológico de GO</b>  |
|--------------|--|------------------------|--|
| 225562       | activador de proteína RAS p21 3                                | RASA3                  | casca de señalización intracelular, mediada por regulación de GTPasa pequeña         |
| 226040       | ADNc de ARNm DKFZp762N156 (de clon)                            | ---                    | ---  |
| 227968       | que contiene dominio 7 de enfermedad de Parkinson 1            | PDDC1                  | ---  |
| 228198       | proteína ribosómica mitocondrial S9                            | MRPS9                  | biosíntesis de proteínas, respuesta al daño de ADN, detección de ADN                 |
| 229830       | Locus transcrito   | ---                    | ---  |
| 230683       | ADNC: FLJ20892 fis, clon ADKA03430                             | ---                    | ---  |
| 230185       | motivo de unión a ARN, proteína de interacción monocatenaria 1 | RBMS1                  | replicación de ADN, procesamiento de ARN /// regulación de la traducción             |
| 241597       | repeticiones de dipéptido de arginina-ácido glutámico (RE)     | RERE                   | remodelación de la cromatina, transcripción, importación de sustrato portador de NLS |
| 242551       | fase abierta de lectura 1 del Cromosoma 18                     | C18orf1                | proceso biológico  |
| 244208       | supresor de punto de control 1                                 | CHES1                  | punto de control de daño de ADN, fase G2 del ciclo celular mitótico                  |
| 244494       | dedo de cinc, que contiene tipo DHHC 1                         | ZDHHC1                 | proceso biológico, palmitoilación de proteínas                                       |
| 244548       | proteína activadora de Rho GTPasa 26                           | ARHGAP26               | transducción de señal, desarrollo del sistema nervioso, citoesqueleto de actina      |

5 Tabla 8: 41 genes regulados positivamente en pacientes con buen pronóstico: 41 transcritos del biomarcador de 43 genes se sobreexpresaron en el grupo con buen resultado clínico. Solamente 2 genes se regularon negativamente, de nuevo el receptor muscarínico de acetilcolina M3, como en la firma molecular de 9 genes, y un transcrito con función desconocida.

## REIVINDICACIONES

1. Uso de una firma molecular que consiste en secuencias génicas complementarias de las secuencias de ácido nucleico: 232669\_at (factor inducible por Hipoxia 3, subunidad alfa), 214951\_at (familia de vehículo de soluto 26, miembro 10), 243482\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 226210\_s\_at (expresado por vía materna 3), 232159\_at (tipo sustrato 15 de ruta del receptor de factor de crecimiento epidérmico 1), 233026\_s\_at (que contiene dominio PDZ), 211996\_s\_at (gen hipotético de proteína de tipo KIAA0220 LOC 283846), 243774\_at (mucina 20, asociado a superficie celular), 242551\_at (fase abierta de lectura del cromosoma 18), 244548\_at (proteína activadora de Rho GTPasa 26), 244208\_at (supresor de punto de control 1), 239984\_at (canal de sodio, abierto por tensión, tipo VII, alfa), 230683\_at (ADNC:FLJ20892 fis, clon ADKA03430), 214869\_at (apolipoproteína L, 6), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE)), 235887\_at (homólogo de Smg-6, factor de degradación de ARNm mediada por mutación terminadora (*C. elegans*)), 229957\_at (proteína transmembrana 91), 223546\_x\_at (tipo LUC7L (*S. cerevisiae*)), 239567\_at (proteína activadora de Rho GTPasa 10), 242194\_at (Culina 4A), 1558525\_at (proteína hipotética LOC283901), 227178\_at (repetición de triplete CUG, proteína de unión a ARN 2), 228198\_s\_at (proteína ribosómica mitocondrial S9), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales), 224260\_at (clon de ADNC), 238643\_at (Neuroblastoma, supresión de tumorigenicidad 1), 232253\_at (homólogo de RAD50 (*S. cerevisiae*)), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1), 233197\_at (tipo kelch 9 (*Drosophila*)), 244512\_at (locus transcrito fuertemente similar a XP\_0010813421), 233443\_at (proteína hipotética LOC389362), 231275\_at (proteína FLJ42875), 226419\_s\_at (proteína hipotética LOC64546), 201221\_s\_at (polipéptido de 70kDa de ribonucleoproteína nuclear pequeña), 209354\_at (miembro de la familia de receptor de factor de necrosis tumoral 14), 226571\_s\_at (tipo receptor de proteína tirosina fosfatasa, S), 220728\_at (EST), 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto), 213946\_s\_at (tipo obscurina 1, similar a isoforma de titina N2-B), 201394\_s\_at (proteína de motivo de unión a ARN 5), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1), 223147\_s\_at (dominio de repetición de WD 33), 213773\_x\_at (familia de dominio NOL/NOP2/Sun, miembro 5), 1560049\_at (repetición de triplete CUG, proteína de unión a ARN 2), 243974\_at (clon de ADNC IMAGE:4821815) y 201510\_at factor de tipo E74 3 (factor de transcripción de dominio ets, específico epitelial) para predecir un pronóstico de insuficiencia cardíaca de nueva aparición, que comprende examinar la expresión de las secuencias génicas de la firma molecular en una muestra tisular o celular.
2. El uso de la reivindicación 1, en el que la sobreexpresión de las secuencias génicas es pronóstico de un buen resultado clínico.
3. El uso de la reivindicación 1, en el que la detección en una célula o en una muestra biológica obtenida de un paciente de las secuencias génicas es pronóstico de un pronóstico clínico deseable en insuficiencia cardíaca, y en el que las secuencias génicas están sobreexpresadas a niveles en al menos 1 % en una célula o en una muestra biológica de un paciente en comparación con niveles en una célula normal o un sujeto normal, o en el que las secuencias génicas están sobreexpresadas en aproximadamente 50 % en una célula o un paciente en comparación con niveles en una célula normal o un sujeto normal, o en el que las secuencias génicas están sobreexpresadas en aproximadamente 75 % en una célula o en una muestra biológica obtenida de un paciente en comparación con niveles en una célula normal o un sujeto normal.
4. El uso de una cualquiera de las reivindicaciones 1 a 3, en el que un biochip comprende las secuencias de ácido nucleico.
5. Un procedimiento para identificar y distinguir entre pacientes con insuficiencia cardíaca de nueva aparición con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para la recuperación de enfermedad cardíaca que comprende identificar en una muestra biológica de un paciente la firma molecular de la reivindicación 1, en el que las secuencias génicas de la firma molecular están sobreexpresadas en una muestra biológica obtenida de un paciente con un buen resultado de pronóstico.
6. El procedimiento de la reivindicación 5, en el que la firma molecular se identifica de un paciente aislando ácidos nucleicos obtenidos de la muestra biológica, y en el que los ácidos nucleicos se hibridan preferentemente con un biochip y los valores de intensidad sin procesar de hibridación de micromatrices se normalizan y se identifican diferencias específicas de fenotipo en la expresión génica, y en el que las diferencias en la expresión génica se identifican preferentemente por análisis de significación de micromatrices, en el que la significación se define con un valor q y múltiples comparaciones comprenden un valor p ajustado, y en el que la especificidad de fenotipo se identifica preferentemente creando un clasificador en un conjunto de entrenamiento que comprende aproximadamente 66 % de los datos obtenidos, con validación posterior en un conjunto de ensayo que comprende aproximadamente 33 % de los datos obtenidos y que define un centroide reducido más próximo específico de fenotipo para clasificación, y en el que el centroide reducido más próximo específico de fenotipo comprende preferentemente equilibrar una validación cruzada de aproximadamente 10 veces en un conjunto de entrenamiento.

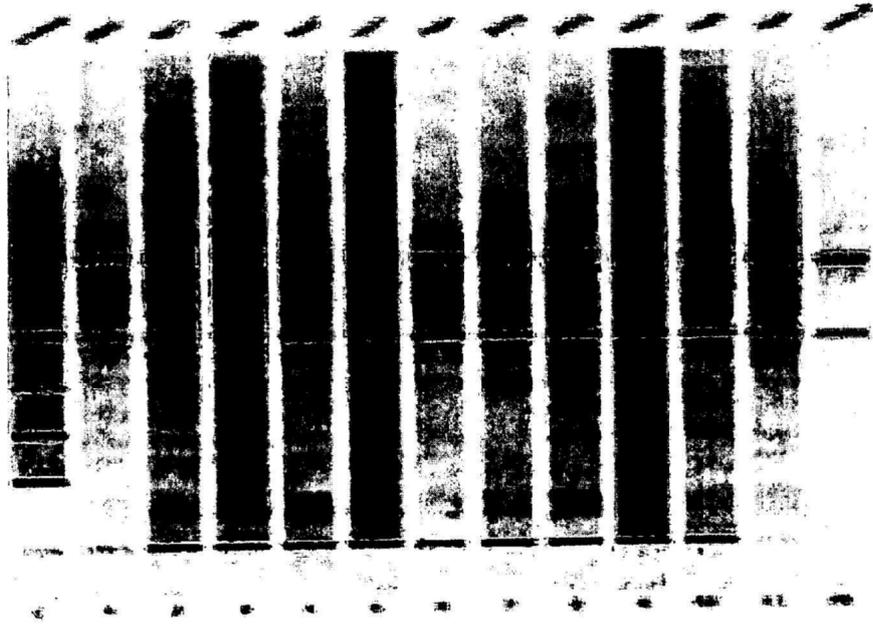
7. Un procedimiento para predecir un resultado de pronóstico para enfermedad cardíaca o recuperación de insuficiencia cardíaca de nueva aparición que comprende identificar en una muestra biológica obtenida de un paciente las secuencias génicas de la firma molecular de la reivindicación 1, en el que las secuencias génicas de la firma molecular están sobreexpresadas en pacientes con un buen resultado de pronóstico, prediciendo de este modo el resultado de pronóstico para enfermedad cardíaca o recuperación de insuficiencia cardíaca.
8. El procedimiento de la reivindicación 7, en el que las secuencias génicas están sobreexpresadas en al menos aproximadamente 20 % en comparación con una célula normal o un sujeto normal.
9. Uso de una firma molecular que consiste en secuencias génicas complementarias de las secuencias de ácido nucleico: 1558458\_at (LOC401320 hipotética), 1560049\_at (repetición de triplete CUG, proteína de unión de ARN 2, CUGBP2), 201394\_s\_at (proteína de motivo de unión a ARN 5, RBM5), 201655\_s\_at (proteoglicano de heparán sulfato 2 (perlecano), HSPG2), 202379\_s\_at (secuencia de reconocimiento de tumor por linfocitos citolíticos naturales, NKTR), 202808\_at, 203071\_at (dominio sema, dominio de inmunoglobulina (Ig), dominio básico corto, secretado, (semaforina) 3B, SEMA3B), 203748\_x\_at (motivo de unión a ARN, proteína de interacción monocatenaria 1, RBMS1), 203981\_s\_at (casete de unión a ATP, subfamilia D (ALD), miembro 4, ABCD4), 204737\_s\_at (miosina, cadena pesada 6, miosina, cadena pesada 7, MYH6 /// MYH7), 204978\_at (factor de corte y empalme, rico en arginina/serina 16, SFRS16), 206209\_s\_at (anhidrasa carbónica IV, CA4), 207541\_s\_at (componente de exosoma 10, EXOSC10), 207798\_s\_at (tipo ataxina 2, ATXN2L), 208978\_at (proteína rica en cisteína 2, CRIP2), 209354\_at (superfamilia del receptor de factor de necrosis tumoral, miembro 14, TNFRSF14), 210628\_x\_at (proteína de unión a factor de crecimiento transformante beta latente 4, LTBP4), 211909\_x\_at (receptor de prostaglandina E 3 (subtipo EP3), PTGER3), 211996\_at (proteína de tipo KIAA0220, complejo de poro nuclear (LOC23117)), 212487\_at (que contiene dominio de parche G 8, GPATCH8), 213946\_s\_at (tipo obscurina 1, OBSL1), 214951\_at (familia de vehículo de soluto 26, miembro 10, SLC26A10), 220219\_s\_at (que contiene repetición rica en leucina 37A, LRRC37A), 221071\_at, 221780\_s\_at (polipéptido de caja DEAD (Asp-Glu-Ala-Asp) 27DDX27), 221806\_s\_at (que contiene dominio SET 5, SETD5), 221833\_at (peptidasa Lon 2, peroxisómica, LONP2), 223546\_x\_at (tipo LUC7 (*S. cerevisiae*), LUC7L), 224260\_at (clon de ADNc IMAGE:4478733), 225562\_at (activador de proteína AS p21 3, RASA3), 226040\_at (ARNM; ADNc DKFZp762N156 (del clon DKFZp762N156), 227968\_at (que contiene dominio 7 de enfermedad de Parkinson 1, PDDC1), 228198\_s\_at (proteína ribosómica mitocondrial S9, MRPS9), 229830\_at (locus transcrito), 230683\_at (ADNC: FLJ20892 fis, clon ADKA03430), 238185\_at (motivo de unión a ARN, proteína de interacción monocatenaria I, RBMS1), 241597\_at (repeticiones de dipéptido de arginina-ácido glutámico (RE), RERE), 242551\_at (fase abierta de lectura 1 del cromosoma 18, C18orf1), 244208\_at (supresor de punto de control 1, CHES1), 244494\_at (dedo de cinc, que contiene tipo DHHC 1, ZDHHC1) y 244548\_at (proteína activadora de Rho GTPasa 26, ARHGAP26) para predecir un pronóstico de insuficiencia cardíaca, que comprende examinar la expresión de las secuencias génicas de la firma molecular en una muestra tisular o celular.
10. El uso de la reivindicación 9, en el que la sobreexpresión de las secuencias génicas es pronóstico de un buen resultado clínico.
11. El uso de la reivindicación 9, en el que la detección en una célula o una muestra biológica obtenida de un paciente de las secuencias génicas es predictiva de resultado de diagnóstico clínico y pronóstico de insuficiencia cardíaca, y en el que las secuencias génicas están sobreexpresadas a niveles en al menos 1 % en una célula o una muestra biológica de un paciente en comparación con niveles en una célula normal o un sujeto normal, o en el que las secuencias génicas están sobreexpresadas en aproximadamente 50 % en una célula o una muestra biológica obtenida de un paciente en comparación con niveles en una célula normal o un sujeto normal, o en el que las secuencias génicas están sobreexpresadas en aproximadamente 75 % en una célula o una muestra biológica obtenida de un paciente en comparación con niveles en una célula normal o un sujeto normal.
12. El uso de la reivindicación 9, en el que un biochip comprende las secuencias de ácido nucleico.
13. Procedimiento para identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación de enfermedad cardíaca que comprende identificar en una muestra biológica de un paciente la firma molecular de la reivindicación 9; en el que las secuencias génicas de la firma molecular están reguladas positivamente en pacientes con un buen resultado de pronóstico, e identificar y distinguir entre pacientes con un alto riesgo de enfermedad cardíaca y pacientes con un buen pronóstico para recuperación de enfermedad cardíaca.
14. El procedimiento de la reivindicación 13, en el que la firma molecular se identifica de un paciente aislando ácidos nucleicos obtenidos de la muestra biológica, y en el que los ácidos nucleicos se hibridan preferentemente con un biochip y los valores de intensidad sin procesar de hibridación de micromatrices se normalizan y se identifican diferencias específicas de fenotipo en la expresión génica, y en el que las diferencias en la expresión génica se identifican preferentemente por análisis de significación de micromatrices, en el que la significación se define con un valor q y múltiples comparaciones comprenden un valor p ajustado, y en el que la especificidad de fenotipo se identifica preferentemente creando un clasificador en un conjunto de

entrenamiento que comprende aproximadamente 66 % de datos obtenidos, con validación posterior en un conjunto de ensayo que comprende aproximadamente 33 % de los datos obtenidos y que define un centroide reducido más próximo específico de fenotipo para clasificación, y

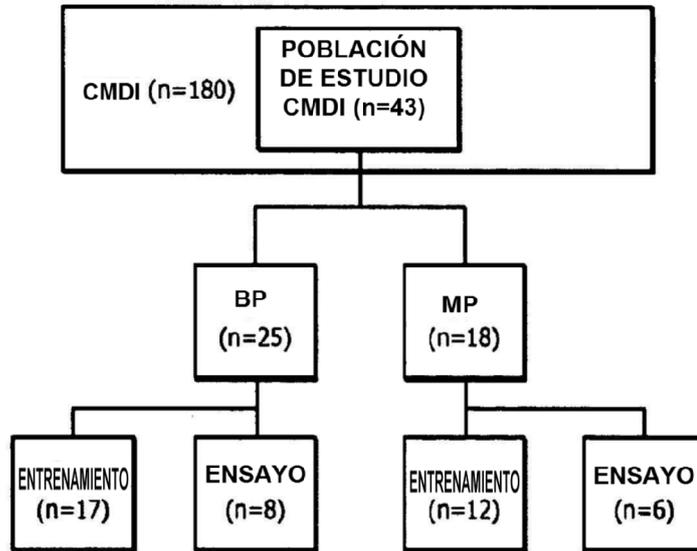
5 en el que el centroide reducido más próximo específico de fenotipo comprende preferentemente equilibrar una validación cruzada de aproximadamente 10 veces en un conjunto de entrenamiento.

15. Un procedimiento para predecir un resultado de pronóstico para enfermedad cardíaca o recuperación de insuficiencia cardíaca que comprende identificar en una muestra biológica de un paciente las secuencias génicas de la firma molecular de la reivindicación 9, en el que las secuencias génicas de la firma molecular están sobreexpresadas en pacientes con un buen resultado de pronóstico

10 16. El procedimiento de la reivindicación 15, en el que las secuencias génicas están sobreexpresadas en al menos aproximadamente 20 % en comparación con una célula normal o un sujeto normal.



**FIG. 1**



*FIG. 2*

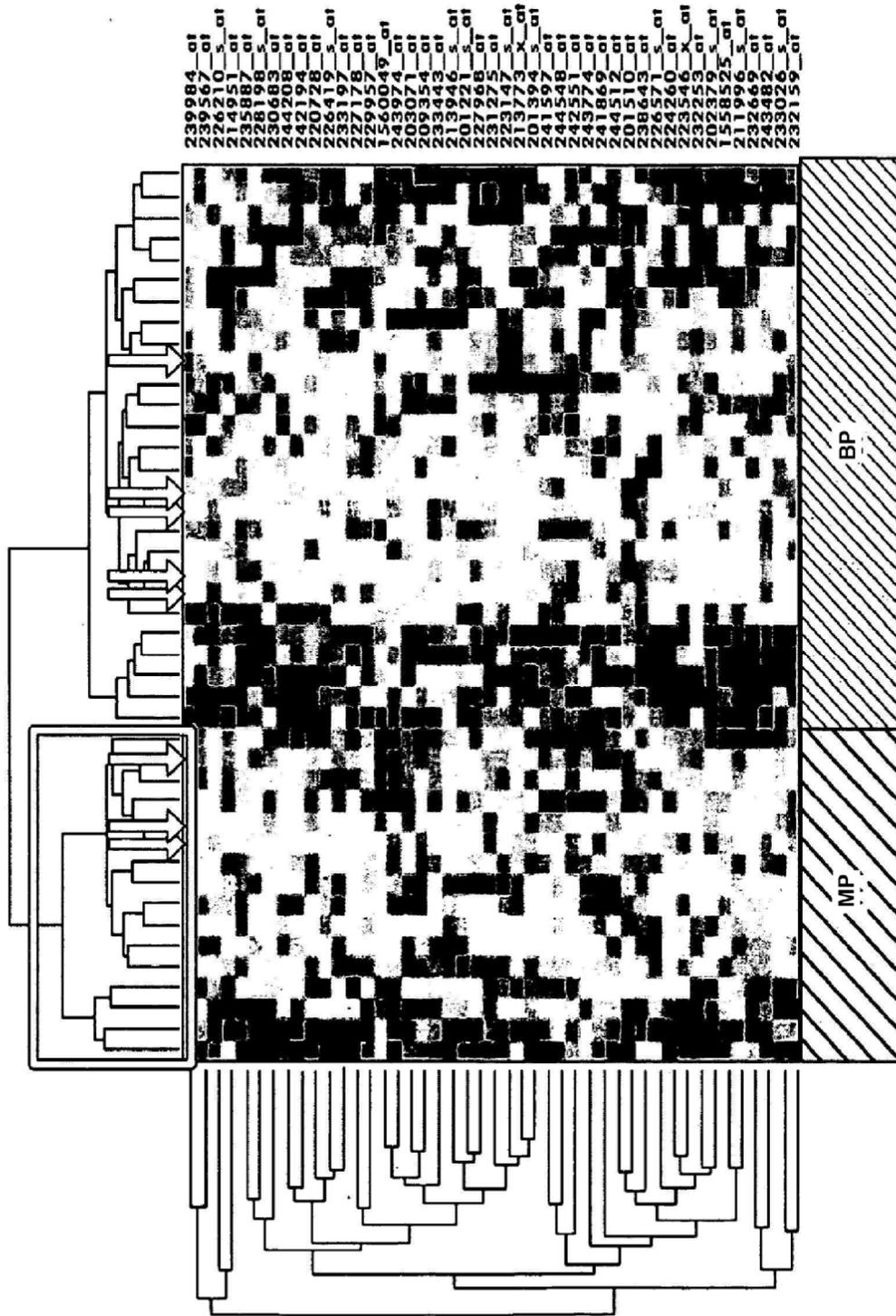


FIG. 3

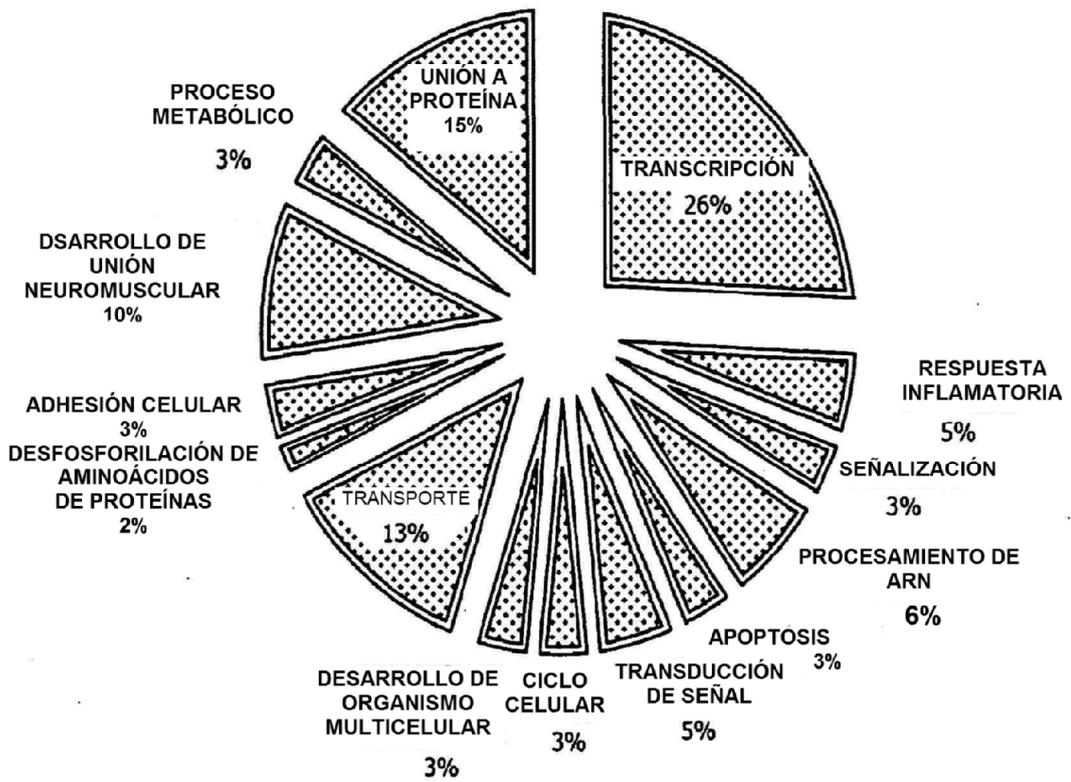


FIG. 4

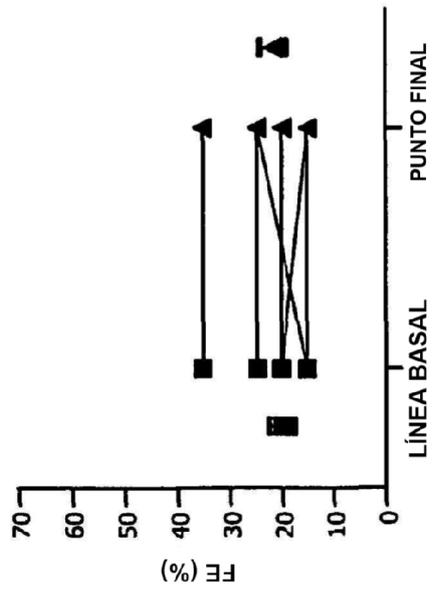
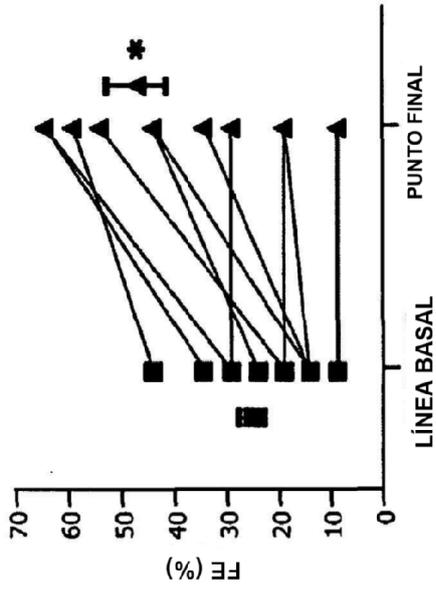


FIG. 5

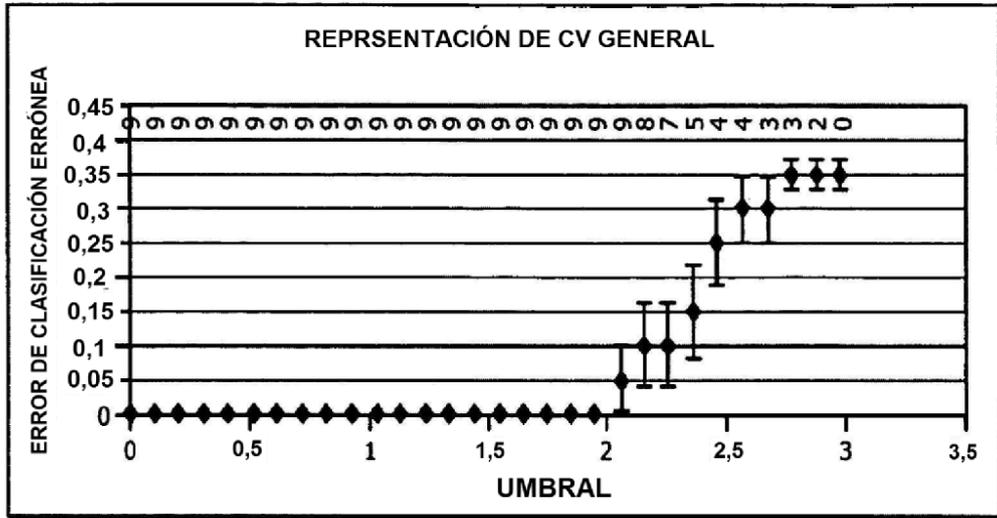


FIG. 6

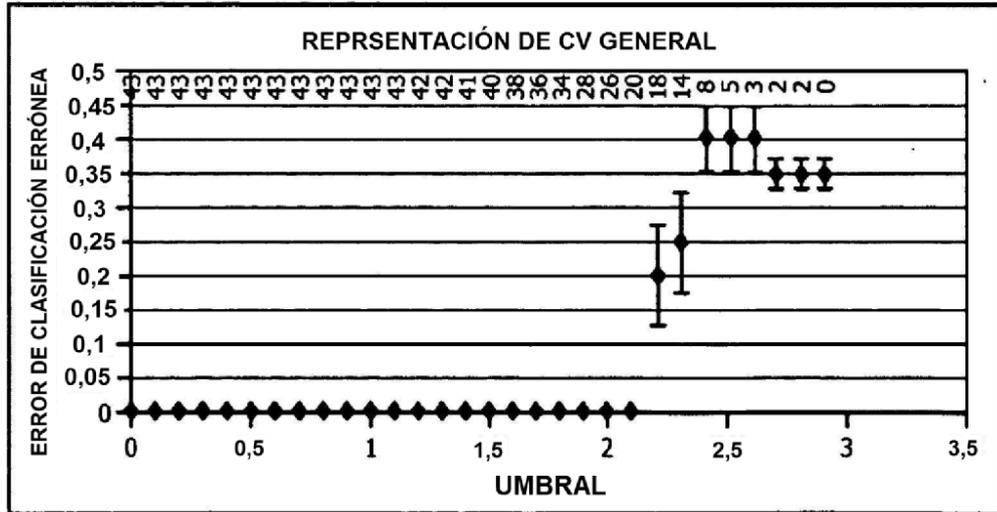


FIG. 7

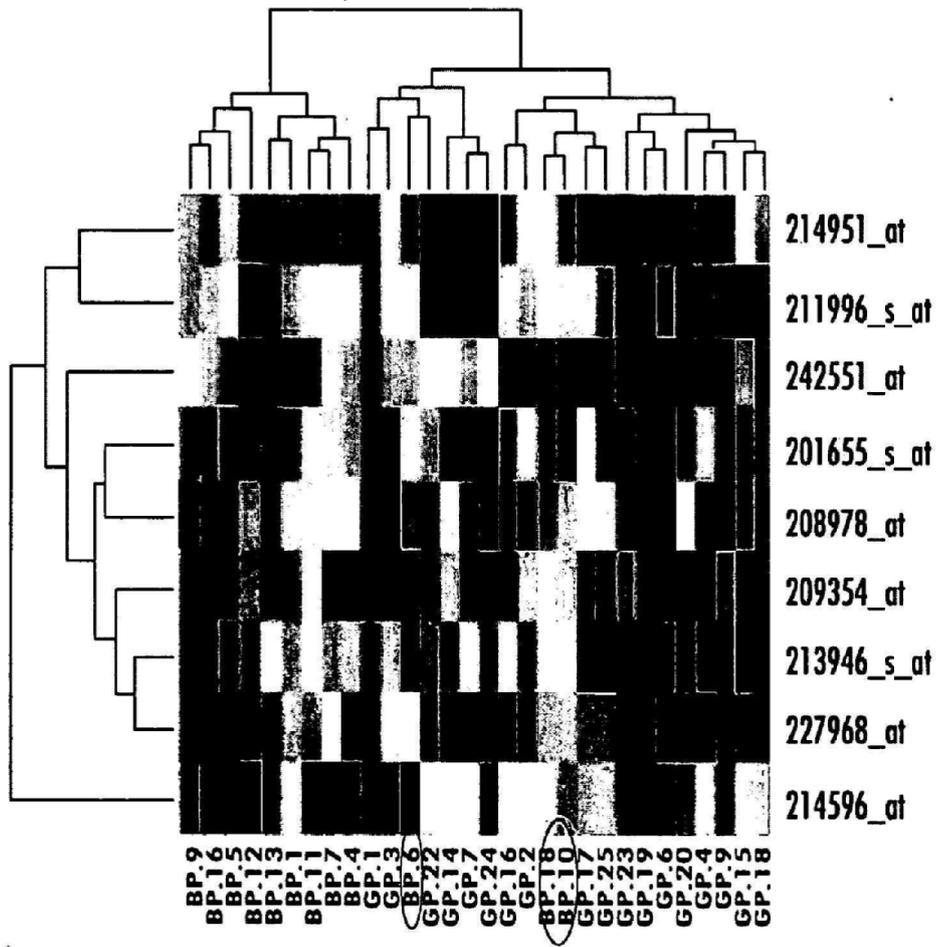


FIG. 8

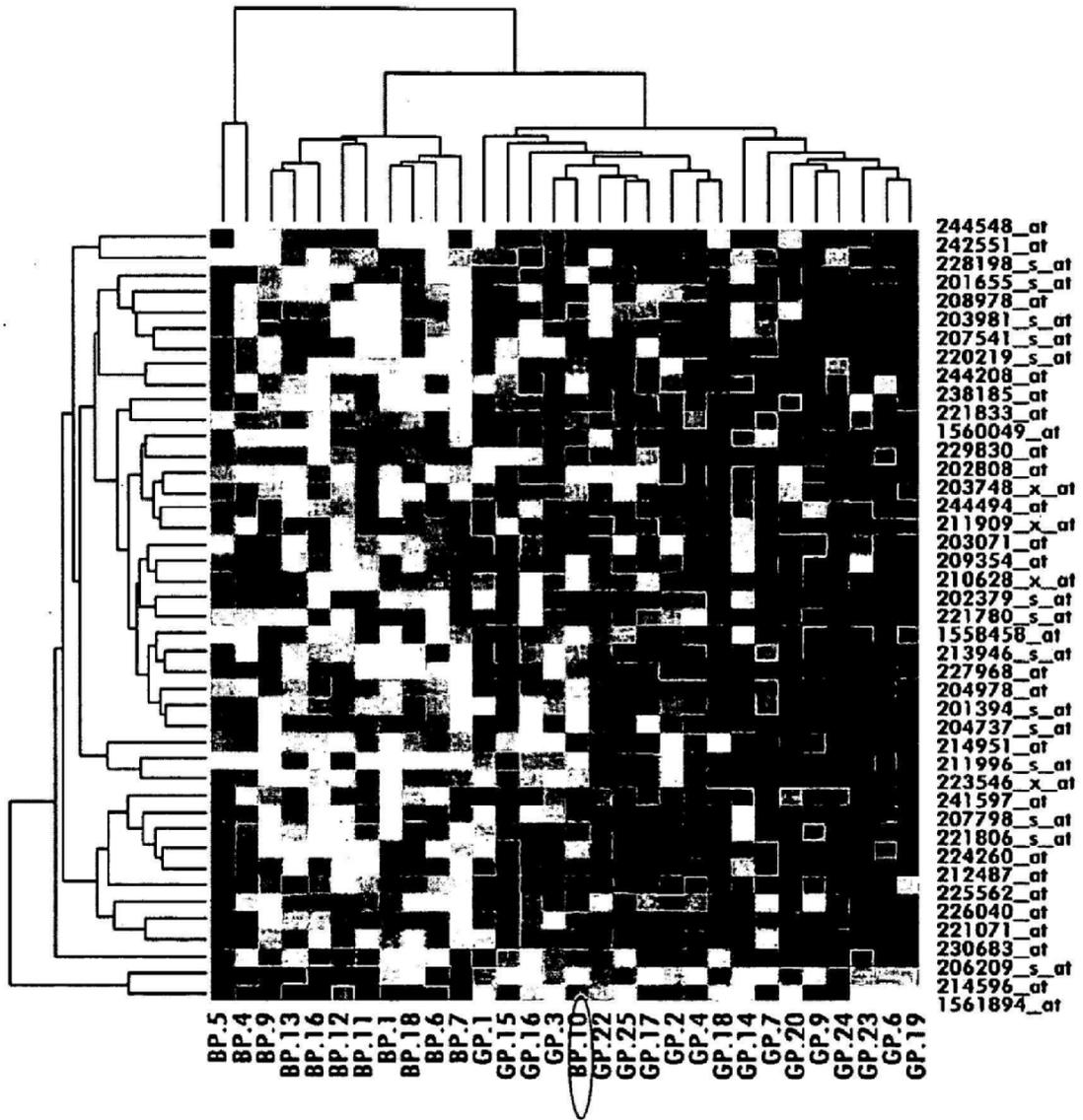


FIG. 9

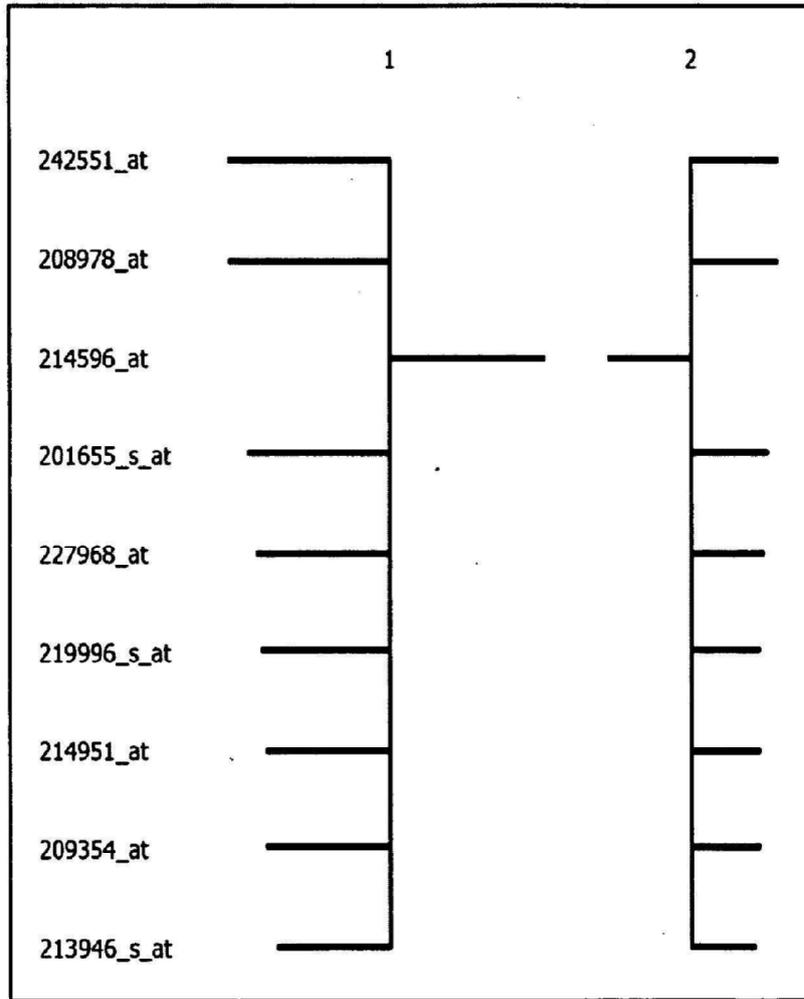


FIG. 10

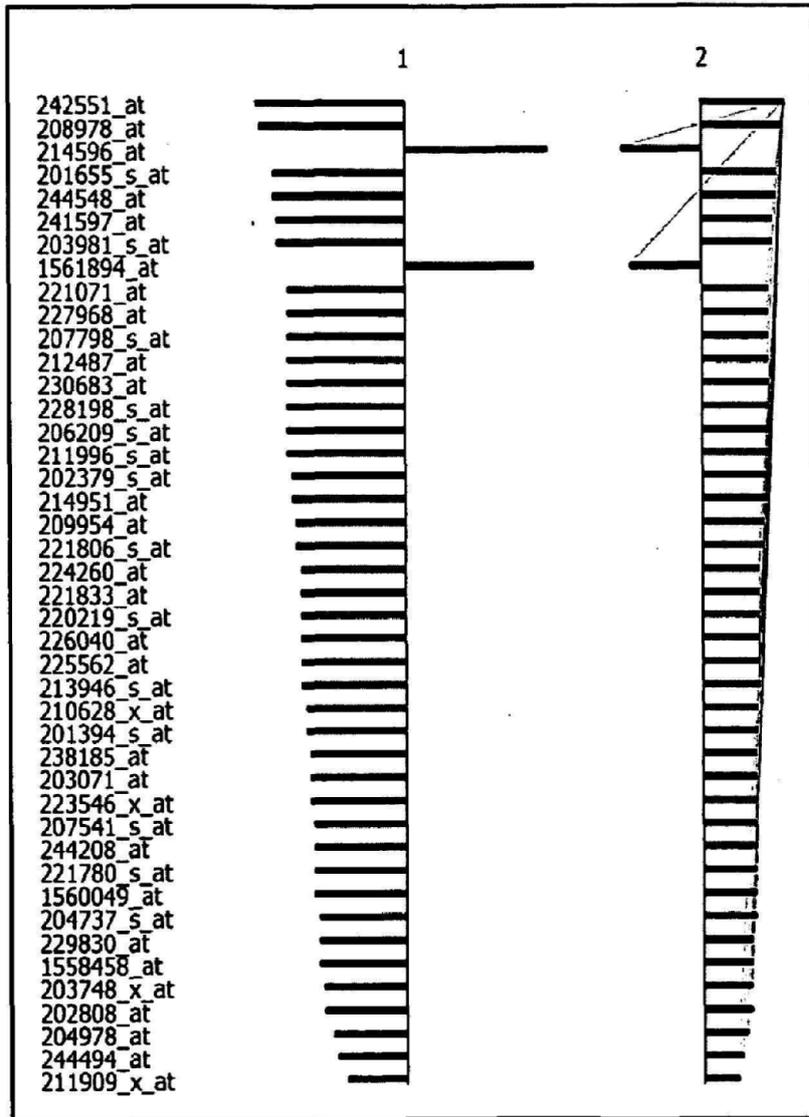


FIG. 11