

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 583 003**

51 Int. Cl.:

C12Q 1/68 (2006.01)

C40B 40/08 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.05.2008 E 08754986 (1)**

97 Fecha y número de publicación de la concesión europea: **20.04.2016 EP 2152916**

54 Título: **Un biomarcador transcriptómico de miocarditis**

30 Prioridad:

01.05.2007 US 915215 P

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

16.09.2016

73 Titular/es:

**UNIVERSITY OF MIAMI (100.0%)
1475 NW 12th Avenue, Room 2012
Miami FL 33136, US**

72 Inventor/es:

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

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 583 003 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).

DESCRIPCION

Un biomarcador transcriptómico de miocarditis

5 Campo de la invención

Esta invención se refiere a biomarcadores de enfermedad cardíaca, miocarditis, nuevos objetivos terapéuticos del fármaco, composiciones y procedimientos para predecir, diagnosticar y tratar enfermedades del corazón y trastornos relacionados con las mismas. Más específicamente, la invención se refiere a procedimientos y composiciones con base en firmas moleculares únicas asociadas con diversos aspectos de las enfermedades y trastornos cardíacos.

Antecedentes

15 El enfoque actual para el tratamiento de pacientes con insuficiencia cardíaca debida al deterioro de la función cardiaca adolece de individualización. Este problema es de importancia creciente, pues aumenta el número de clases de medicinas para la insuficiencia cardíaca. Además, existe un creciente reconocimiento de que puede haber utilidad para engendrar terapias específicas. Se requieren biomarcadores precisos para perfeccionar la precisión diagnóstica a fin de mejorar la aplicación de la medicina personalizada en el campo de la insuficiencia cardíaca.

20 Hay una necesidad en la técnica para proporcionar un diagnóstico y pronóstico temprano de una enfermedad cardíaca. La miocarditis provoca una minoría significativa de la función cardiaca deprimida y por lo tanto provoca insuficiencia cardíaca y muerte cardíaca súbita prematura e inesperada. La miocarditis afecta a los seres humanos a lo largo de la vida, incluyendo los niños. El enfoque del diagnóstico actual usando análisis histológico del tejido cardiaco obtenido por biopsia carece de sensibilidad y especificidad. Dado el alto riesgo de desarrollar complicaciones cardíacas graves por miocarditis y la disponibilidad de terapias específicas para la enfermedad, existe la necesidad de mejores biomarcadores, para ajustar el tratamiento de manera adecuada y lo suficientemente temprano.

30 NANNI L Y COLABORADORES (2006) JOURNAL OF MOLECULAR AND CELLULAR CARDIOLOGY, vol 41, 934-948 revisa perfiles de expresión génica en enfermedades cardiovasculares, incluyendo miocarditis.

Resumen

35 Se identificaron firmas moleculares que funcionan como un biomarcador de diagnóstico muy sensible para la miocarditis, enfermedades del corazón y trastornos del mismo.

En una realización preferida, una composición molecular comprende secuencias génicas: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (proteína en dedo de zinc 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína nuclear heterogénea H1 (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza a ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (dedo de zinc de la proteína 623), 212637_s_at (dominio WW que contiene la proteína ubiquitina E3 ligasa 1), 212920_at, 213317_at (canal intracelular de cloro 5), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y el dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con un dominio del tipo EGF y dos dominios del tipo de la folistatina 2), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina del tipo aminotransferasa 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína quinasa nuclear y sustrato 1 de quinasa que depende de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOC1 97135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio de tipo AN1 3), 240971_x_at (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

En otra realización preferida, la detección de las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, es diagnóstica de la miocarditis y trastornos de miocardio.

60 En otra realización preferida, la detección en una célula o paciente de al menos diez secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, es diagnóstica del diagnóstico de la miocarditis, cardiomiopatía idiopática, enfermedades del corazón y trastornos del mismo.

En otra realización preferida, las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, están sobreexpresados en niveles de al menos 1 %, 2 %, 5 %, 10 % en una célula o paciente en comparación con los niveles en una célula normal o sujeto normal.

5 Un biomarcador (TBB-I) para el diagnóstico de la miocarditis comprende: secuencias/biomoléculas de ácido nucleico que comprenden: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (proteína en dedo de zinc 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína H1 nuclear heterogénea (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza con ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (dedo de cinc de proteína 623), 212637_s_at (dominio WW que contiene la proteína ubiquitina E3 ligasa 1), 212920_at, 213317_at (canal intracelular de cloro 5), 213619_at (ribonucleoproteína H1 nuclear heterogénea (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y el dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con un dominio del tipo EGF y dos dominios del tipo de la folistatina 2), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina del tipo aminotransferasa 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína quinasa nuclear y sustrato 1 de quinasa que depende de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOC1 97135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio de tipo AN1 3), 240971_x_at (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

25 En una realización alternativa, la detección de al menos diez biomoléculas es diagnóstico de miocarditis.

En otra realización preferida, un anticuerpo o aptámero específico para cada secuencia génica que comprende secuencias de ácidos nucleicos/biomoléculas que comprenden: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (dedo de cinc de la proteína 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína H1 nuclear heterogénea (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza con ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (dedo de cinc de proteína 623), 212637_s_at (dominio WW que contiene la proteína ubiquitina E3 ligasa 1), 212920_at, 213317_at (canal intracelular de cloro 5), 213619_at (ribonucleoproteína H1 nuclear heterogénea (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y el dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con un dominio del tipo EGF y dos dominios del tipo de la folistatina 2), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina del tipo aminotransferasa 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína quinasa nuclear y sustrato 1 de quinasa que depende de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOC1 97135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio de tipo AN1 3), 240971_x_at (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

50 En otra realización preferida, un biochip comprende secuencias de ácidos nucleicos: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (dedo de cinc de la proteína 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína H1 nuclear heterogénea (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza con ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (dedo de cinc de proteína 623), 212637_s_at (dominio WW que contiene la proteína ubiquitina E3 ligasa 1), 212920_at, 213317_at (canal intracelular de cloro 5), 213619_at (ribonucleoproteína H1 nuclear heterogénea (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y el dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con un dominio del tipo EGF y dos dominios del tipo de la folistatina 2), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina del tipo aminotransferasa 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína quinasa nuclear y sustrato 1 de quinasa que depende de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOC1 97135 hipotético), 235803_at (factor 3 del

tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio de tipo AN1 3), 24097_1_x_at (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

5 En otra realización preferida, el biochip comprende al menos diez secuencias nucleicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

En una realización preferida, un biomarcador transcriptómico (TBB-II) comprende secuencias génicas:

10 1552419_s_at (familia similar a la de la tubulina tirosina ligasa, el miembro 10), 1553212_at (queratina 78), 1555124_at (proteína hipotética MGC40574), 1556192_x_at (supresor 1 de metástasis), 1556320_at (similar a la estomatina 1 (EPB72)), 1556510_at (clon de ADNc IMAGE: 4796864), 1558484_s_at (repetición rica en leucina que contiene 27), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567410_at (dedo de zinc de proteína 135), 1568513_x_at (proteasa, serina, 1 (tripsina 1)), 1570408_at (serina/treonina quinasa 24 (homólogo de STE20, levadura)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina, 204581_at (molécula CD22, glicoproteína asociada a mielina), 205586_x_at (factor de crecimiento nervioso inducible VGF), 206333_at (homólogo 1 de musashi (Drosophila)), 207004_at (CLL de células B/linfoma 2), 210059_s_at (proteína quinasa 13 activada por mitógeno), 210228_at (factor 2 estimulador de colonias 2 (macrófago-granulocito)), 210384_at (proteína arginina metiltransferasa 2), 210923_at (familia 1 portador de soluto (transportador de glutamato), miembro 7), el factor 211024_s_at (factor 1 de transcripción tiroides /// factor 1 de transcripción de tiroides), 211062_s_at (carboxipeptidasa Z /// carboxipeptidasa Z), 211096_at (factor 2 de transcripción de leucemia de pre-células B), 211181_x_at (factor 1 de transcripción relacionado con enanismo (leucemia 1 mieloide aguda; oncogén aml1)), 211710_x_at (proteína ribosomal L4///proteína ribosomal L4), 213096_at (familia 2 del dominio transmembrana y del dominio de bobina en espiral), 213121_at (polipéptido de 70 kDa de la ribonucleoproteína nuclear pequeña (antígeno RNP)), 213242_x_at (KIAA0284), 213568_at (impar omitida relacionada con 2 (Drosophila)), 213770_at supresor de la quinasa de ras 1 (214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 216116_at (proteína que interactúa con NCK con el dominio SH3), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 216820_at , 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217180_at (proteína hipotética similar al producto génico KIAA0187), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 217430_x_at (colágeno, tipo 1, alfa 1), 219070_s_at (dominio de esperma móvil que contiene 3), 219425_at (familia 4A de la sulfotransferasa, miembro 1), 221663_x_at (receptor H3 de histamina), 221684_s_at (Nictalopina), 223974_at (proteína hipotética MGC11082), 226640_at (diacilglicerol lipasa beta), 228074_at (proteína hipotética LOC162073), 229191_at (cofactor D de plegamiento de la tubulina), 229257_at (proteína KIAA1856), 229335_at (superfamilia de las inmunoglobulinas, miembro 4C), 229358_at (homólogo hedgehog indio (Drosophila)), 230341_x_at (metalopeptidasa ADAM con el motivo tipo 1 de trombospondina, 10), 230693_at (ATPasa, que transporta Ca⁺⁺, músculo cardiaco, contracción rápida 1), 230768_at (FERM, RhoGEF y dominio pleckstrina de la proteína 2), 231510_at (miembro de la familia GLI-Kruppel GLI2), 231629_x_at (peptidasa 3 relacionada con calicreína), 231998_at, 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 234637_at (proteína 4-5 relacionada con queratina), 234881_at, 235568_at (marco de lectura abierto 59 del cromosoma 19) , 235600_at (locus transcrito), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 237087_at (marco de lectura abierto 105 del cromosoma 14), 237144_at (factor de crecimiento beta transformante latente que se enlaza a la proteína 3), 237398_at (locus transcrito), 237547_at (proteína hipotética LOC728730), 237679_at (66 que contiene al motivo tripartita), 238267_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239026_x_at (centaurina, gamma 3), 239463_at (locus transcrito), 239756_at (tipo 1 deficiente en la captura mitótica de MAD1 (levadura)), 240039_at (locus transcrito /// locus transcrito), 240147_at (proteína hipotética MGC11257), 240517_at (cistationina beta sintasa), 241270_at (homólogo 2 romboide 5 (Drosophila)), 241431_at, 242365_at (dominio de bobina en espiral que contiene 32), 243297_at (proteína vacuolar para clasificar al homólogo D 13 (*S. cerevisiae*)), 243497_at (locus transcrito), 243766_s_at (miembro 2 de la familia del dominio TEA), 43934_at (receptor 137 acoplado a la proteína G), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

En otra realización preferida, la detección de las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, es diagnóstico de miocarditis y trastornos de miocardio, enfermedades y trastornos cardíacos y cardiovasculares, tales como, por ejemplo, enfermedad cardíaca coronaria, angina, síndrome coronario agudo, aneurisma aórtico y disección, arritmias, cardiomiopatía, enfermedad cardíaca congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis, y similares.

En otra realización alternativa, la detección en una célula o paciente de al menos diez secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, es diagnóstico del diagnóstico de miocarditis, cardiomiopatía idiopática, enfermedades cardíacas y trastornos del mismo.

En otra realización, las secuencias génicas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, se modulan en niveles de al menos aproximadamente 1 %, 5 %, 10 %, 100 %, 200 % o más en una célula o el paciente en comparación con los niveles en una célula normal o sujeto normal.

- En otra realización preferida, un biochip comprende secuencias de ácido nucleico: 1552419_s_at (familia similar a la tubulina tirosina, miembro 10), 1553212_at (queratina 78), 1555124_at (proteína hipotética MGC40574), 1556192_x_at (supresor 1 de metástasis), 1556320_at (similar a la estomatina 1 (EPB72)), 1556510_at (clon de ADNc IMAGE: 4796864), 1558484_s_at (repetición rica en leucina que contiene 27), 1565662_at (mucina 6, moco oligomérico/ formador de gel), 1567410_at (dedo de zinc de la proteína 135), 1568513_x_at (proteasa, serina, 1 (tripsina 1)), 1570408_at (serina / treonina quinasa 24 (homólogo de STE20, levadura)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 204581_at (molécula CD22, glicoproteína asociada a mielina), 205586_x_at (factor de crecimiento nervioso inducible VGF), 206333_at (homólogo 1 musashi (Drosophila)), 207004_at (CLL de células B/ linfoma 2), 210059_s_at (proteína quinasa 13 activada por mitógeno), 210228_at (factor 2 estimulante de colonias (granulocitos y macrófagos)), 210384_at (proteína arginina metiltransferasa 2), 210923_at (familia 1 portadora de soluto (transportador de glutamato), miembro 7), 211024_s_at (factor 1 de transcripción de tiroides /// factor 1 de transcripción de tiroides), 211062_s_at (carboxipeptidasa Z /// carboxipeptidasa Z), 211096_at (factor 2 de transcripción de leucemia de pre-células B), 211181_x_at (factor 1 de transcripción relacionada con enanismo (leucemia 1 mieloide aguda; oncogén aml1)), 211710_x_at (proteína ribosomal L4 /// proteína ribosomal L4), 213096_at (familia 2 del dominio transmembrana y de bobina en espiral), 213121_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña (antígeno RNP)), 213242_x_at (KIAA0284), 213568_at (impar omitida relacionada con 2 (Drosophila)), 213770_at supresor quinasa de ras 1 (214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 216116_at (proteína que interactúa NCK con el dominio SH3), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 216820_at , 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217180_at (proteína hipotética similar al producto génico KIAA0187), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 217430_x_at (colágeno, tipo 1, alfa 1), 219070_s_at (dominio de espermatozoides móviles que contiene 3), 219425_at (familia 4A de sulfotransferasa, miembro 1), 221663_x_at (receptor de histamina H3), 221684_s_at (Nictalopina), 223974_at (proteína hipotética MGC11082), 226640_at (diacilglicerol lipasa beta), 228074_at (proteína hipotética LOC162073), 229191_at (cofactor D de pliegue de tubulina), 229257_at (proteína KIAA1856), 229335_at (superfamilia de las inmunoglobulinas, miembro 4C), 229358_at (homólogo hedgehog indio (Drosophila)), 230341_x_at (metalopeptidasa ADAM con el motivo tipo 1 de trombospondina, 10), 230693_at (ATPasa, transporte de Ca⁺⁺, músculo cardíaco, contracción rápida 1), 230768_at (FERM, RhoGEF y dominio pleckstrina de la proteína 2), 231510_at (miembro de la familia GLI-Kruppel GLI2), 231629_x_at (peptidasa 3 relacionada con calicreína), 231998_at, 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 234637_at (proteína 4-5 asociada con queratina), 234881_at, 235568_at (marco de lectura abierto 59 del cromosoma 19), 235600_at (locus transcrito), 236496_at (homólogo 2 de espermatozoides degenerativos, lipido desaturasa (Drosophila)), 237087_at (marco de lectura abierto 105 del cromosoma 14), 237144_at (factor de crecimiento beta transformante latente que se enlaza a la proteína 3), 237398_at (locus transcrito), 237547_at (proteína hipotética LOC728730), 237679_at (66 que contiene un motivo tripartita), 238267_s_at, 238445_x_at (manosil (alfa-1,6-)glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239026_x_at (centaurina, gamma 3), 239463_at (locus transcrito), 239756_at (tipo 1 deficiente en la captura mitótica de MAD1 (levadura)), 240039_at (locus transcrito /// locus transcrito), 240147_at (proteína hipotética MGC11257), 240517_at (cistationina beta sintasa), 241270_at (homólogo 2 romboide 5 (Drosophila)), 241431_at, 242365_at (dominio de bobina en espiral que contiene 32), 243297_at (proteína vacuolar para clasificar al homólogo D 13 (S. cerevisiae)), 243497_at (locus transcrito), 243766_s_at (miembro 2 de la familia del dominio TEA), 43934_at (receptor 137 acoplado a la proteína G), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.
- En aún otra realización, un anticuerpo o aptámero específico para cada secuencia génica que comprende secuencias de ácido nucleico/biomoléculas que comprenden: 1552419_s_at (familia similar a la tirosina ligasa de tubulina, miembro 10), 1553212_at (queratina 78), 1555124_at (proteína hipotética MGC40574), 1556192_x_at (supresor 1 de metástasis), 1556320_at (similar a la estomatina 1 (EPB72)), 1556510_at (clon de ADNc IMAGE: 4796864), 1558484_s_at (repetición rica en leucina que contiene 27), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567410_at (dedo de zinc de proteína 135), 1568513_x_at (proteasa, serina, 1 (tripsina 1)), 1570408_at (serina / treonina quinasa 24 (homólogo de STE20, levadura)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido guanina), 204581_at (molécula CD22, glicoproteína asociada a mielina), 205586_x_at (factor de crecimiento nervioso inducible VGF), 206333_at (homólogo 1 musashi (Drosophila)), 207004_at (CLL de células B/ linfoma 2), 210059_s_at (proteína quinasa 13 activada por mitógeno), 210228_at (factor 2 estimulante de colonias (granulocitos y macrófagos)), 210384_at (proteína arginina metiltransferasa 2), 210923_at (familia 1 portadora de soluto (transportador de glutamato), miembro 7), 211024_s_at (factor 1 de transcripción de tiroides /// factor 1 de transcripción de tiroides), 211062_s_at (carboxipeptidasa Z /// carboxipeptidasa Z), 211096_at (factor 2 de transcripción de leucemia de pre-células B), 211181_x_at (factor 1 de transcripción relacionada con enanismo (leucemia 1 mieloide aguda; oncogén aml1)), 211710_x_at (proteína ribosomal L4 /// proteína ribosomal L4), 213096_at (familia 2 del dominio transmembrana y de bobina en espiral), 213121_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña (antígeno RNP)), 213242_x_at (KIAA0284), 213568_at (impar omitida relacionada con 2 (Drosophila)), 213770_at supresor quinasa de ras 1 (214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 216116_at (proteína que interactúa NCK con el dominio SH3), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 216820_at , 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217180_at (proteína hipotética similar al producto génico KIAA0187), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 217430_x_at (colágeno, tipo 1, alfa 1), 219070_s_at (dominio de espermatozoides móviles que

5 contiene 3), 219425_at (familia 4A de sulfotransferasa, miembro 1), 221663_x_at (receptor de histamina H3), 221684_s_at (Nictalopina), 223974_at (proteína hipotética MGC11082), 226640_at (diacilglicerol lipasa beta), 228074_at (proteína hipotética LOC162073), 229191_at (cofactor D de plegue de tubulina), 229257_at (proteína KIAA1856), 229335_at (superfamilia de las inmunoglobulinas, miembro 4C), 229358_at (homólogo hedgehog indio (Drosophila)), 230341_x_at (metalopeptidasa ADAM con el motivo tipo 1 de trombospondina, 10), 230693_at (ATPasa, transporte de Ca⁺⁺, músculo cardíaco, contracción rápida 1), 230768_at (FERM, RhoGEF y dominio pleckstrina de la proteína 2), 231510_at (miembro de la familia GLI-Kruppel GLI2), 231629_x_at (peptidasa 3 relacionada con calicreína), 231998_at, 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 234637_at (proteína 4-5 asociada con queratina), 234881_at, 235568_at (marco de lectura abierto 59 del cromosoma 19), 235600_at (locus transcrito), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 237087_at (marco de lectura abierto 105 del cromosoma 14), 237144_at (factor de crecimiento beta transformante latente que se enlaza a la proteína 3), 237398_at (locus transcrito), 237547_at (proteína hipotética LOC728730), 237679_at (66 que contiene un motivo tripartita), 238267_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239026_x_at (centaurina, gamma 3), 239463_at (locus transcrito), 239756_at (tipo 1 deficiente en la captura mitótica de MAD1 (levadura)), 240039_at (locus transcrito /// locus transcrito), 240147_at (proteína hipotética MGC11257), 240517_at (cistationina beta sintasa), 241270_at (homólogo 2 romboide 5 (Drosophila)), 241431_at, 242365_at (dominio de bobina en espiral que contiene 32), 243297_at (proteína vacuolar para clasificar al homólogo D 13 (*S. cerevisiae*)), 243497_at (locus transcrito), 243766_s_at (miembro 2 de la familia del dominio TEA), 43934_at (receptor 137 acoplado a la proteína G), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

25 En otra realización, un biomarcador (TBB-III) comprende secuencias de ácido nucleico/biomoléculas que comprenden: 1553212_at (queratina 78), 1557236_at (apolipoproteína L), 1558142_at (repetición de trinucleótidos que contiene 6B), 1558484_s_at (repetición rica en leucina que contiene 27), 1565614_at (dedo de zinc de proteína 337), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 205758_at (molécula CD8a /// molécula CD8a), 206333_at (homólogo 1 de musashi (Drosophila)), 212920_at, 213242_x_at (KIAA0284), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 219425_at (familia 4A de sulfotransferasa, miembro 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 229191_at (cofactor D de plegamiento de tubulina), 229569_at (clon de ADNc IMAGE: 5263455), 231_629_x_at (peptidasa 3 relacionada con calicreína), 233765_at (LOC197135 hipotético), 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 235568_at (marco de lectura abierto 59 del cromosoma 19), 235803_at (factor 3 del tipo receptor de citoquinas), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 236953_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239463_at (locus transcrito), 240544_at (dedo de zinc, dominio 3 tipo AN-1), 243766_s_at (miembro 2 de la familia del dominio TEA), y secuencias complementarias, fragmentos, alelos, variantes de 244042_x_at y productos génicos de los mismos.

45 En otra realización preferida, un procedimiento de diagnóstico de miocarditis, comprende identificar en una muestra biológica de un paciente una firma molecular de secuencias de ácido nucleico/biomoléculas que comprenden al biomarcador-I de base transcriptómica (TBB-I) que comprende: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (dedo de zinc de proteína 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína nuclear heterogénea H1 (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza a ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (proteína en dedo de zinc 623), 212637_s_at (dominio WW que contiene proteína ligasa 1 ubiquitina E3), 212920_at, 213317_at (canal intracelular de cloro 5), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con un dominio del tipo EGF y dos dominios del tipo de la folistatina 2), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina aminotransferasa tipo 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína quinasa nuclear y sustrato 1 de quinasa dependiente de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOC197135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio 3 de tipo AN1), 240971-x-al (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos; evaluación de la probabilidad de identificación de cada gen componente en cada muestra; asignación de cada uno a una clase; y, diferenciación entre cardiomiopatía idiopática y miocarditis.

En otra realización preferida, un procedimiento de diagnóstico de enfermedades del corazón o miocarditis comprende identificar en una muestra biológica de un paciente una firma molecular de secuencias de ácido nucleico/biomoléculas que comprende un biomarcador-II con base en transcriptómica (TBB-II) que comprende:

5 1552419_s_at (familia similar a la tirosina quinasa de tubulina, miembro 10), 1553212_at (queratina 78), 1555124_at (proteína hipotética MGC40574), 1556192_x_at (supresor 1 de metástasis), 1556320_at (similar a la estomatina 1 (EPB72)), 1556510_at (clon de ADNc IMAGE: 4796864), 1558484_s_at (repetición rica en leucina que contiene 27), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567410_at (dedo de zinc de proteína 135), 1568513_x_at (proteasa, serina, 1 (tripsina 1)), 1570408_at (serina / treonina quinasa 24 (homólogo de STE20, levadura)),

10 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 204581_at (molécula CD22, glicoproteína asociada a mielina), 205586_x_at (factor de crecimiento nervioso inducible VGF), 206333_at (homólogo 1 de musashi (Drosophila)), 207004_at (CLL de células B / linfoma 2), 210059_s_at (proteína quinasa 13 activada por mitógeno), 210228_at (factor 2 estimulante de colonias (granulocito-macrófago)), 210384_at (proteína arginina metiltransferasa 2), 210923_at (familia 1 portadora de soluto (transportador de glutamato), miembro 7), 211024_s_at (factor 1 de transcripción de la tiroides /// factor 1 de transcripción de la tiroides), 211062_s_at (carboxipeptidasa Z /// carboxipeptidasa Z), 211096_at (factor 2 de transcripción de leucemia de pre-células B), 211181_x_at (factor 1 de transcripción relacionado con enanismo (leucemia 1 mieloide aguda; oncogén aml1)), 211710_x_at (proteína ribosomal L4///proteína ribosomal L4), 213096_at (familia 2 del dominio transmembrana y del dominio de bobina en espiral), 213121_at (polipéptido de 70 kDa de la ribonucleoproteína nuclear pequeña (antígeno RNP)), 213242_x_at (KIAA0284), 213568_at (impar omitida relacionada con 2 (Drosophila)), 213770_at supresor quinasa de ras 1 (214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 216116_at (proteína que interactúa con NCK con el dominio SH3), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 216820_at, 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217180_at (proteína hipotética similar al producto génico KIAA0187), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 217430_x_at (colágeno, tipo 1, alfa 1), 219070_s_at (dominio de esperma móvil que contiene 3), 219425_at (familia 4A de sulfotransferasa, miembro 1), 221663_x_at (receptor H3 de histamina), 221684_s_at (Nictalopina), 223974_at (proteína hipotética MGC11082), 226640_at (diacilglicerol lipasa beta), 228074_at (proteína hipotética LOC162073), 229191_at (cofactor D de plegamiento de la tubulina), 229257_at (proteína KIAA1856), 229335_at (superfamilia de las inmunoglobulinas, miembro 4C), 229358_at (homólogo hedgehog indio (Drosophila)), 230341_x_at (metalopeptidasa ADAM con el motivo tipo 1 de trombospondina, 10), 230693_at (ATPasa, que transporta Ca⁺⁺, músculo cardíaco, contracción rápida 1), 230768_at (FERM, RhoGEF y dominio pleckstrina de la proteína 2), 231510_at (miembro de la familia GLI-Kruppel GLI2), 231629_x_at (peptidasa 3 relacionada con calicreína), 231998_at, 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 234637_at (proteína 4-5 relacionada con queratina), 234881_at, 235568_at (marco de lectura abierto 59 del cromosoma 19), 235600_at (locus transcrito), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 237087_at (marco de lectura abierto 105 del cromosoma 14), 237144_at (proteína 3 que se enlaza al factor de crecimiento beta transformante latente), 237398_at (locus transcrito), 237547_at (proteína hipotética LOC728730), 237679_at (66 que contiene al motivo tripartita), 238267_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239026_x_at (centaurina, gamma 3), 239463_at (locus transcrito), 239756_at (tipo 1 deficiente en la captura mitótica de MAD1 (levadura)), 240039_at (locus transcrito /// locus transcrito), 240147_at (proteína hipotética MGC11257), 240517_at (cistationina beta sintasa), 241270_at (homólogo 2 romboide 5 (Drosophila)), 241431_at, 242365_at (dominio de bobina en espiral que contiene 32), 243297_at (proteína vacuolar para clasificar al homólogo D 13 (*S. cerevisiae*)), 243497_at (locus transcrito), 243766_s_at (miembro 2 de la familia del dominio TEA), 43934_at (receptor 137 acoplado a la proteína G), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos; evaluación de la probabilidad de identificación de cada gen componente en cada muestra; asignación de cada uno a una clase; y, diagnóstico de la enfermedad cardíaca o miocarditis.

En otra realización preferida, un procedimiento de diagnóstico de enfermedades del corazón o miocarditis comprende identificar en una muestra biológica de un paciente una firma molecular de secuencias de ácido nucleico/biomoléculas que comprende un biomarcador-III con base en transcriptómica (TBB-III) que comprende:

50 1553212_at (queratina 78), 1557236_at (apolipoproteína L), 1558142_at (repetición de trinucleótidos que contiene 6B), 1558484_s_at (una repetición rica en leucina que contiene 27), 1565614_at (dedo de zinc de proteína 337), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567100_at (homólogo de Dachshund 1 (Drosophila)),

55 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 205758_at (molécula CD8a /// molécula CD8a), 206333_at (homólogo 1 de musashi (Drosophila)), 212920_at, 213242_x_at (KIAA0284), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 219425_at (familia 4A de sulfotransferasa, miembro 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 229191_at (cofactor D de plegamiento de tubulina), 229569_at (clon de ADNc IMAGE: 5263455), 231629_x_at (peptidasa 3 relacionada con calicreína), 233765_at (LOC197135 hipotético), 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencias 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 235568_at (marco de lectura abierto 59 del cromosoma 19), 235803_at (factor 3 del tipo receptor de citoquinas), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa

(Drosophila)), 236953_s_at, 238445_x_at (manosil (alfa-1,6)- glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isoforma B), 239463_at (locus transcrito), 240544_at (dedo de cinc, dominio 3 de tipo AN1), 243766_s_at (miembro 2 de la familia del dominio TEA), y secuencias complementarias, fragmentos, alelos, variantes de 244042_x_at y productos génicos de los mismos; evaluación de la probabilidad de identificación de cada gen componente en cada muestra; asignación de cada uno a una clase; y, diagnóstico de la enfermedad cardíaca o miocarditis.

En otra realización preferida, un kit comprende un biomarcador-I con base en transcriptómica (TBB-I): 1553212_at (queratina 78), 1557236_at (apolipoproteína L), 1558142_at (repetición de trinucleótidos que contiene 6B), 1558484_s_at (una repetición rica en leucina que contiene 27), 1565614_at (dedo de zinc de proteína 337), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 205758_at (molécula CD8a /// molécula CD8a), 206333_at (homólogo 1 de musashi (Drosophila)), 212920_at, 213242_x_at (KIAA0284), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 216427_at (ADNc: FLJ22786 fis, clon KAlA2150), 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 219425_at (familia 4A de sulfotransferasa, miembro 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 229191_at (cofactor D de plegamiento de tubulina), 229569_at (clon de ADNc IMAGE: 5263455), 231629_x_at (peptidasa 3 relacionada con calicreína), 233765_at (LOC197135 hipotético), 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 235568_at (marco de lectura abierto 59 del cromosoma 19), 235803_at (factor 3 del tipo receptor de citoquinas), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 236953_s_at, 238445_x_at (manosil (alfa-1,6)- glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isoforma B), 239463_at (locus transcrito), 240544_at (dedo de cinc, dominio 3 de tipo AN1), 43766_s_at (miembro 2 de la familia de dominio TEA), y secuencias complementarias, fragmentos, alelos, variantes de 244042_x_at y productos génicos de los mismos.

En otra realización preferida, un kit comprende un biomarcador-II con base en transcriptómica (TBB-II): 1552419_s_at (familia similar a la tirosina ligasa de tubulina, miembro 10), 1553212_at (queratina 78), 1555124_at (proteína hipotética MGC40574), 1556192_x_at (supresor 1 de metástasis), 1556320_at (similar a la estomatina 1 (EPB72)), 1556510_at (clon de ADNc IMAGE: 4796864), 1558484_s_at (repetición rica en leucina que contiene 27), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567410_at (dedo de zinc de proteína 135), 1568513_x_at (proteasa, serina, 1 (tripsina 1)), 1570408_at (serina/treonina quinasa 24 (homólogo de STE20, levadura)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 204581_at (molécula CD22, glicoproteína asociada a la mielina), 20S586_x_at (factor de crecimiento nervioso inducible VGF), 206333_at (homólogo 1 de musashi (Drosophila)), 207004_at (CLL de células B / linfoma 2), 210059_s_at (proteína quinasa 13 activada por mitógeno), 210228_at (factor 2 estimulante de colonias (granulocitos y macrófagos)), 210384_at (proteína arginina metiltransferasa 2), 210923_at (familia 1 portadora de soluto 1 (transportador de glutamato), miembro 7), 211024_s_at (factor 1 de transcripción de la tiroides /// factor 1 de transcripción de la tiroides), 211062_s_at (carboxipeptidasa Z /// carboxipeptidasa Z), 211096_at (factor 2 de transcripción de leucemia de pre-células B), 211181_x_at (factor 1 de transcripción relacionado con enanismo (leucemia mieloide aguda 1; oncogén aml1)), 211710_x_at (proteína ribosomal L4 /// proteína ribosomal L4), 213096_at (familia 2 del dominio transmembrana y de bobina en espiral), 213121_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña (antígeno RNP)), 213242_x_at (KIAA0284), 213568_at (impar omitida relacionada con 2 (Drosophila)), 213770_at supresor quinasa de ras 1 (214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 216116_at (proteínas que interactúan NCK con el dominio SH3), 216427_at (ADNc: FLJ22786 fis, clon KAlA2150), 216820_at, 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217180_at (hipotética proteína similar a KIAA0187 producto génico), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 217430_x_at (colágeno tipo I, alfa 1), 219070_s_at (dominio motilidad de esperma que contiene 3), 219425_at (familia 4A de sulfotransferasa, miembro 1), 221663_x_at (histamina H3 receptor), 221684_s_at (Nictalopina), 223974_at (proteína hipotética MGC11082), 226640_at (diacilglicerol lipasa beta), 228074_at (proteína hipotética LOC162073), 229191_at (cofactor D de plegamiento de tubulina), 229257_at (proteína KIAA1856), 229335_at (superfamilia de las inmunoglobulinas, 4C miembros), 229358_at (India homólogo hedgehog (Drosophila)), 230341_x_at (metalopeptidasa ADAM con el tipo de trombospondina 1 adorno, 10), 230693_at (ATPasa, Ca <+> transporte, músculo cardíaco, rápido tic 1), 230768_at (FERM, RhoGEF y pleckstrina proteína de dominio 2), 231510_at (miembro de la familia GLI-Kruppel GLI2), 231629_x_at (peptidasa 3 relacionada con calicreína), 231998_at, 233794_at (proteína de ADN de cadena sencilla de unión 3), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa-calicreína relacionada 15), 234637_at (proteína queratina asociado 4-5), 234881_at, 235568_at (marco de lectura abierto 59 del cromosoma 19), 235600_at (locus transcrito), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 237087_at (cromosoma 14 marco de lectura abierto 105), 237144_at (factor de crecimiento transformante latente en la proteína de unión 3), 237398_at (locus transcrito), 237547_at (proteína hipotética LOC728730), 237679_at (tripartito 66 que contiene motivo de), 238267_s_at, 238445_x_at (manosil (alfa-1,6 -) - glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isoforma B), 239026_x_at (centaurina, gamma 3), 239463_at (locus transcrito), 239756_at (detención de la mitosis MAD1 deficientes tipo 1 (levadura)), 240039_at (locus transcrito /// locus transcrito), 240147_at (hipotética proteína MGC11257), 240517_at (cistationina beta

sintasa), 241270_at (homólogo 2 romboide 5 (*Drosophila*)), 241431_at, 242365_at (dominio de bobina en espiral que contiene 32), 243297_at (proteína vacuolar para clasificar al homólogo D 13 (*S. cerevisiae*)), 243497_at (locus transcrito), 243766_s_at (familia del dominio TEA miembro 2), 43934_at (receptor 137 acoplado a la proteína G), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

5 En otra realización preferida, un kit comprende un biomarcador-III de base transcriptómica (TBB-III): 1553212_at (queratina 78), 1557236_at (apolipoproteína L), 1558142_at (repetición de trinucleótidos que contiene 6B), 1558484_s_at (repetición rica en leucina que contiene 27), 1565614_at (dedo de zinc de proteína 337), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567100_at (homólogo de Dachshund 1 (*Drosophila*)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 205758_at (molécula CD8a /// molécula CD8a), 10 206333_at (homólogo 1 de musashi (*Drosophila*)), 212920_at, 213242_x_at (KIAA0284), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 216427_at (ADNc: FLJ22786 fis, el clon KAI2150), 15 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 219425_at (familia 4A de sulfotransferasa, miembro 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 229191_at (cofactor D de plegamiento de tubulina), 229569_at (clon de ADNc IMAGE: 5263455), 231629_x_at (peptidasa 3 relacionada con calicreína), 233765_at (LOC197135 hipotético), 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 235568_at (marco de lectura abierto 59 del cromosoma 19), 235803_at (factor 3 del tipo receptor de citoquinas), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (*Drosophila*)), 236953_s_at, 238445_x_at (manosil (alfa-1,6)- glicoproteína beta-1,6-N-acetilglucosaminiltransferasa, isozima B), 239463_at (locus transcrito), 240544_at (dedo de cinc, dominio 3 de tipo AN1), 243766_s_at (miembro 2 de la familia del dominio TEA), y secuencias complementarias, fragmentos, alelos, 20 25 variantes de 244042_x_at y productos génicos de los mismos.

En otra realización, una célula expresa una o más biomoléculas que comprenden el biomarcador TBB-I.

En otra realización, una célula expresa una o más biomoléculas que comprenden el biomarcador TBB-II.

En otra realización, una célula expresa una o más biomoléculas que comprenden el biomarcador TBB-III

En otra realización preferida, un vector codifica una o más biomoléculas que comprenden TBB-I.

En otra realización preferida, un vector codifica una o más biomoléculas que comprenden TBB-II.

En otra realización preferida, un vector codifica una o más biomoléculas que comprenden TBB-III.

Los aspectos de la invención se describen más adelante.

Breve descripción de los dibujos

La invención se señala en forma particular en las reivindicaciones adjuntas. Las ventajas anteriores y adicionales de esta invención se pueden entender mejor haciendo referencia a la siguiente descripción tomada junto con los dibujos adjuntos, en los que:

La Figura 1 muestra un gráfico SAM de muestras de pacientes con cardiomiopatía idiopática frente a los pacientes con miocarditis: el color rojo indica los genes que fueron significativamente sobrerregulados en pacientes con miocarditis, el color verde indica los genes que fueron significativamente subregulados. Las líneas azules exteriores representan los puntos de corte positivos y negativos que fueron elegidos con el valor delta. 134 genes eran significativamente diferentes entre los dos grupos, cuando se eligió un valor delta de 1,41 (FDR = 0,49). Se redujo este subconjunto de genes a 122 genes candidatos con un valor q de 0 % que se usó para el agrupamiento jerárquico y análisis PAM.

La Figura 2 muestra un mapa térmico creado por la agrupación jerárquica usando 122 genes candidatos: las muestras de pacientes con miocarditis se etiquetan como "mio-", las muestras de pacientes con cardiomiopatía idiopática se etiquetan ya sea con "GP-" o "BP-". Cada columna representa una muestra y cada línea corresponde a un gen, para el que se enumeran los ID de Affymetrix en el lado derecho. Para más detalles acerca de las anotaciones de genes véase la Tabla 1. Un color rojo significa una baja expresión del gen, mientras que un color azul demuestra altos niveles de expresión génica. Seis muestras se agruparon mal.

La Figura 3 es un gráfico que muestra el error de clasificación del conjunto de entrenamiento. El clasificador fue entrenado en 8 muestras de pacientes con miocarditis y 25 muestras de pacientes con cardiomiopatía idiopática. Después de aumentar el umbral a 3,2 y la reducción de los genes del clasificador a menos de 22, el error de clasificación aumentó dramáticamente.

La figura 4 es un gráfico que muestra los resultados del "clasificador de 39 genes" para miocarditis: Este gráfico visualiza las probabilidades calculadas para cada clase después de que se aplicó la "firma molecular de 39 genes" con un umbral de 2.6. La probabilidad se puede leer en el eje y. El grupo 1 representa las muestras de pacientes con cardiomiopatía idiopática, el grupo 2 representa las muestras de pacientes con miocarditis.

La Figura 5 muestra el centroide más cercano reducido del "clasificador de 39 genes": Los centroides se calcularon en PAM a partir de la expresión promedio de cada gen en cada clase dividido por la desviación estándar dentro de una clase para ese gen. La clasificación más cercana del centroide reducido "reduce" cada uno de los centroides de clase hacia el centroide global para todas las clases por el umbral. La clasificación del centroide más cercano toma el perfil de expresión génica de una nueva muestra, y lo compara con cada uno de estos centroides de clase. La clase cuyo centroide que es más cercana a, en distancia al cuadrado, es la clase predicha para esa nueva muestra. Cada línea en el gráfico representa un gen. El centroide rojo caracteriza al grupo 1 (cardiomiopatía idiopática), el centroide verde caracteriza al grupo 2 (miocarditis). La sobreexpresión se ilustra como un vector a la derecha, la subregulación como vector a la izquierda en el gráfico.

La Figura 6 muestra un mapa térmico del "biomarcador transcriptómico de 39 genes": Este mapa térmico fue creado por el mismo procedimiento de agrupamiento no supervisado que en la figura 2.

La Figura 7 muestra el análisis de componentes principales (PCA) y diversos algoritmos de agrupación en muestras de miocarditis frente a otros tipos de cardiomiopatía: Para ilustrar la contribución de cada uno de los 122 genes ($FC > 1,2$; $q < 0,1$ %) para cada fenotipo, se realizó un PCA ($n = 61$). Se usó una matriz de correlación con los genes como variables. Los genes menos significativos se indican con los vectores cerca del centro y corresponden a los genes que fueron excluidos usando el análisis PAM. Las muestras de miocarditis fueron etiquetadas con "M", las muestras de pacientes con otras formas de cardiomiopatía se marcaron con "O". Los genes que se sobreexpresaron están etiquetados con números seriales y se agrupan con la clase correspondiente. Todas las muestras de miocarditis, salvo dos, fueron claramente agrupadas juntas. Además, se analizó la robustez del conjunto reducido de 39 genes con 3 tipos diferentes de algoritmos de agrupamiento (el procedimiento de Ward, enlace completo y enlace promedio). Las muestras de miocarditis se destacan en rojo - muestras incorrectamente clasificadas están en círculos en azul. Todos los procedimientos consiguen la misma precisión que PAM ($n = 33$, 97 % de precisión).

Las Figuras 8A a 8J muestran las ID de Affymetrix y las correspondientes secuencias de las sondas.

Descripción detallada

La invención comprende firmas moleculares que funcionan como un biomarcador de diagnóstico muy sensible para la miocarditis. La miocarditis es una enfermedad común que se estima que causa hasta 30 % de la cardiomiopatía dilatada, incluso en pacientes inicialmente asintomáticos. La miocarditis también puede presentar una muerte cardíaca súbita y afecta a individuos de todas las edades. En la infancia, la miocarditis causa un mayor porcentaje de insuficiencia cardíaca que en la edad adulta. El hecho de que la mayoría de los casos virales inducidos pasen en un curso clínicamente no perceptible, señala la importancia de hallar biomarcadores más confiables que las herramientas de diagnóstico estándar que están disponibles actualmente, por ejemplo, ECG, enzimas cardíacas e inmunohistoquímica.

Las herramientas de diagnóstico estándar actuales para la miocarditis (ECG, enzimas cardíacas, inmunohistoquímica) no siempre son lo suficientemente confiables y muchos pacientes se someten a cursos clínicamente no perceptibles sin ser tratados. Especialmente en pediatría es crucial detectar miocarditis en una etapa temprana ya que se ha observado muchas veces un curso fatal en los niños.

Definiciones

De acuerdo con la presente invención y como se usa en este documento, los siguientes términos se definen con los siguientes significados, a menos que se indique explícitamente lo contrario.

Tal como se usa en este documento, "un", "uno, una" y "el, la" incluyen referencias plurales a menos que el contexto indique claramente lo contrario.

Tal como se usa en este documento, una "firma molecular" o "firma" o "biomarcador" o "biomarcador con base transcriptómica" se usan indistintamente en la presente memoria y se refieren a todas las biomoléculas identificadas en las Tablas 1, 2 y 4. Por lo tanto, la Tabla 1 que comprende las biomoléculas enumeradas en la misma, representa un biomarcador o firma molecular; la Tabla 2 que comprende las biomoléculas enumeradas en la misma, representa otro biomarcador o firma molecular; etcétera. A medida que se descubren más biomoléculas, cada una de las biomoléculas recientemente identificadas se pueden asignar a uno cualquier o más biomarcadores o firma molecular. Cada biomolécula también se puede quitar, reasignar o relocalizar a una firma molecular. Por lo tanto, en algunas realizaciones la firma molecular comprende al menos diez biomoléculas. Las diez biomoléculas se seleccionan de los genes identificados en la presente memoria, o de las biomoléculas recientemente identificadas. Los biomarcadores de la Tabla 1 que comprenden los 38 genes sobreexpresados se denomina TBB-I por razones de brevedad. El biomarcador que comprende las biomoléculas en la Tabla 2 se denomina TBB-II por razones de

- 5 brevedad. El biomarcador que comprende las biomoléculas en la Tabla 4 se denomina TBB-III por razones de brevedad. Al realizar un diagnóstico es deseable detectar al menos 10 o más biomoléculas. Cualquiera de TBB-I, TBB-II y TBB-III o combinaciones de los mismos se pueden usar en el diagnóstico de miocarditis. Cualquiera de TBB-I, TBB-II y TBB-III o combinaciones de los mismos se puede usar en el diagnóstico de la miocarditis y la cardiomiopatía idiopática y diferenciar entre las dos condiciones.
- El término "biomolécula" se refiere a ADN, ARN (incluyendo ARNm, ARNr, ARNt y ARNmt), nucleótidos, nucleósidos, análogos, polinucleótidos, péptidos y cualquiera de sus combinaciones.
- 10 Una "posición" base tal como se usa aquí se refiere a la localización de una base dada o residuo de nucleótido dentro de un ácido nucleico.
- Tal como se usa en el presente documento, el término "matriz" se refiere a una disposición espacial ordenada, en particular una disposición de biomoléculas inmovilizadas.
- 15 Tal como se usa en el presente documento, el término "matriz direccionable" se refiere a una matriz en el que los elementos individuales tienen coordenadas x e y definidas en forma precisa, de manera que se puede identificar un elemento dado en una posición particular en la matriz.
- 20 Tal como se usa en el presente documento, los términos "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. Tales sondas se inmovilizan preferiblemente sobre un microelectrodo que comprende un sustrato.
- 25 Tal como se usa en el presente documento, los términos "biomatriz", "biochip" y "matriz de biochip" se refieren a una disposición espacial ordenada de biomoléculas inmovilizadas sobre un microelectrodo dispuesto sobre un sustrato de soporte sólido. Las moléculas sonda preferidas incluyen aptámeros, ácidos nucleicos, oligonucleótidos, péptidos, ligandos, anticuerpos y antígenos; los péptidos y proteínas son las especies de sondas más preferidas. Biochips, tal como se usa en la técnica, abarcan sustratos que contienen matrices o micromatrices, preferiblemente matrices ordenadas y lo más preferiblemente matrices ordenadas, direccionables, de moléculas biológicas que comprenden un miembro de un par biológico que se enlaza. Normalmente, tales matrices son matrices de oligonucleótidos que comprenden una secuencia de nucleótidos que es complementaria al menos a una secuencia que puede ser o se espera que esté presente en una muestra biológica. Alternativamente, y preferiblemente, las proteínas, péptidos u otras moléculas pequeñas pueden ser dispuestas en dichos biochips para realizar, entre otros, los análisis inmunológicos (en los que las moléculas dispuestas son antígenos) o el ensayo de receptores biológicos (en el que las moléculas dispuestas son ligandos, agonistas o antagonistas de dichos receptores).
- 30
- 35 Expresión/cantidad de un gen, biomolécula, o biomarcador en una primera muestra está a un nivel "mayor que" el nivel en una segunda muestra si el nivel/cantidad de expresión 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/cantidad de expresión del gen o biomarcador en la segunda muestra o una muestra normal. Los niveles/cantidades de expresión se pueden determinar con base en cualquier criterio adecuado conocido en la técnica, incluyendo pero no limitado a ARNm, ADNc, proteínas, fragmentos de proteínas y/o copias de genes. Los niveles/cantidades de expresión se pueden determinar cualitativa y/o cuantitativamente.
- 40
- 45 Por el término "modular", se entiende que cualquiera de las actividades mencionadas, son, por ejemplo, aumentadas, mejoradas, agonizadas (actúan como un agonista), promovidas, disminuidas, reducidas, suprimidas, bloqueadas, o antagonizadas (actúan como un antagonista). La modulación puede aumentar la actividad más de 1 vez, 2 veces, 3 veces, 5 veces, 10 veces, 100 veces, etc., sobre los valores basales. La modulación también puede disminuir su actividad por debajo de los valores basales.
- 50 Un "alelo" o "variante" es una forma alternativa de un gen. Las variantes pueden ser el resultado de al menos una mutación en la secuencia de ácido nucleico y puede dar como resultado ARNm alterados o polipéptidos cuya estructura o función puede o no ser alterada. Cualquier gen natural o recombinante dado puede tener ninguna, una o muchas formas alélicas. Los cambios mutacionales comunes que dan lugar a variantes se atribuyen generalmente a supresiones, adiciones o sustituciones naturales de nucleótidos. Cada uno de estos tipos de cambios pueden ocurrir solos o en combinación con los otros, una o más veces en una secuencia dada.
- 55 El término "complementario" significa que dos secuencias son complementarias cuando la secuencia de una se puede unir a 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 entonces 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 %, preferiblemente 95 %, lo más preferiblemente 100 %, de complementariedad con una secuencia definida. Preferiblemente, se pueden identificar los alelos o variantes de los mismos. También se puede emplear un programa BLAST para evaluar dicha identidad de secuencia.
- 60
- 65

- El término "secuencia complementaria" que se refiere a una secuencia de polinucleótidos, se refiere a la secuencia de bases en otra molécula de ácido nucleico por las reglas de apareamiento de bases. Más particularmente, el término o un término similar se refiere a la hibridación o emparejamiento de bases entre los nucleótidos o ácidos nucleicos, tales como, por ejemplo, entre las dos hebras de una molécula de ADN de doble hebra o entre un cebador oligonucleótido y un sitio de unión del cebador en un ácido nucleico de una sola cadena que se secuencia o amplifica. Los nucleótidos complementarios son, en general, A y T (o A y U), o C y G. Dos moléculas de ARN o ADN monocatenarias se dice que son sustancialmente complementarias cuando los nucleótidos de una hebra, óptimamente alineados y comparados con las inserciones o supresiones de nucleótidos apropiados, se emparejan con al menos aproximadamente 95 % de los nucleótidos de la otra hebra, por lo general al menos aproximadamente 98 %, y más preferiblemente desde aproximadamente 99 % hasta aproximadamente 100 %. Las secuencias de polinucleótidos complementarios se pueden identificar por varios enfoques incluyendo el uso de algoritmos informáticos y software bien conocidos, por ejemplo el programa BLAST.
- 15 Tal como se usa en el presente documento, el término "aptámero" o "especie que se enlaza a un ácido nucleico seleccionado" incluirá ARN o ADN no modificado o modificado químicamente. El procedimiento de selección puede ser por, pero no se limitan a, cromatografía de afinidad y el procedimiento de amplificación por transcripción inversa (RT) o reacción en cadena de la polimerasa (PCR).
- 20 Tal como se usa en el presente documento, el término "aptámero de señalización" incluirá aptámeros con moléculas informadoras, preferiblemente un colorante fluorescente, unidas a un nucleótido de tal manera que tras los cambios conformacionales que resultan de la interacción del aptámero con un ligando, las moléculas informadoras producen una señal diferencial, preferiblemente un cambio en la intensidad de fluorescencia.
- 25 Tal como se usa en el presente documento, el término "fragmento o segmento", como se aplica a una secuencia de ácido nucleico, gen o polipéptido, será ordinariamente de al menos aproximadamente 5 bases de ácido nucleico contiguas (para una secuencia de ácido nucleico o gen) o aminoácidos (para los polipéptidos), normalmente al menos aproximadamente 10 bases de ácido nucleico o aminoácidos contiguos, más normalmente al menos aproximadamente 20 bases de ácido nucleico o aminoácidos contiguos, usualmente al menos aproximadamente 30 bases de ácido nucleico o aminoácidos contiguos, preferiblemente al menos aproximadamente 40 bases de ácido nucleico o aminoácidos contiguos, más preferiblemente al menos aproximadamente 50 bases de ácido nucleico o aminoácidos contiguos, e incluso más preferiblemente al menos aproximadamente 60 a 80 o más bases de ácido nucleico o aminoácidos contiguos de longitud. "Fragmentos que se superponen", como se usa en el presente documento, se refiere a fragmentos de ácido nucleico o de péptidos contiguos que comienzan en el extremo del terminal amino de un ácido nucleico o proteína y que termina en el extremo del terminal carboxilo del ácido nucleico o proteína. Cada fragmento de ácido nucleico o de péptido tiene al menos aproximadamente una posición contigua de ácido nucleico o aminoácido en común con el fragmento siguiente de ácido nucleico o péptido, más preferiblemente al menos aproximadamente tres posiciones contiguas de bases de ácido nucleico o aminoácidos en común, lo más preferiblemente al menos aproximadamente diez posiciones contiguas de bases de ácido nucleico o aminoácidos en común.
- 40 "Muestras biológicas" incluyen muestras sólidas y de fluidos corporales. Preferiblemente, la muestra se obtiene del corazón. Sin embargo, las muestras biológicas usadas en la presente invención pueden incluir células, proteínas o extractos de células de membrana, sangre o fluidos biológicos tales como fluido ascítico o fluido cerebral (por ejemplo, fluido cefalorraquídeo). Ejemplos de muestras biológicas incluyen, pero no se limitan a, muestras tomadas de tejidos del sistema nervioso central, hueso, mama, riñón, cuello uterino, endometrio, cabeza/cuello, vesícula biliar, glándula parótida, próstata, glándula pituitaria, músculo, esófago, estómago, intestino delgado, colon, hígado, bazo, páncreas, tiroides, corazón, pulmón, vejiga, tejido adiposo, ganglios linfáticos, útero, ovarios, glándulas suprarrenales, testículos, amígdalas y el timo. Ejemplos de "muestras de fluidos corporales" incluyen, pero no se limitan a sangre, suero, semen, fluido prostático, fluido seminal, orina, saliva, esputo, mucosidad, médula ósea, linfa, y lágrimas.
- 45 "Muestra" se usa aquí en su sentido más amplio. Una muestra que comprende polinucleótidos, polipéptidos, péptidos, anticuerpos y similares pueden comprender un fluido corporal; una fracción soluble de una preparación de células, o medios en los que se cultivaron las células; un cromosoma, un orgánulo, o membrana aislados o extraídos de una célula; ADN, ARN o ADNc genómico, polipéptidos, o péptidos en solución o unidos a un sustrato; una célula; un tejido; una impresión de tejido; una huella digital, piel o cabello; y similares.
- 50 "Diagnóstico" significa identificar la presencia o naturaleza de una condición 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 dan positivo (porcentaje de "verdaderos positivos"). Los individuos enfermos no detectados por el ensayo son "falsos negativos". Los sujetos que no están enfermos y que obtienen un resultado negativo en el ensayo, 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 aquellos sin la enfermedad que dan positivo. Aunque un procedimiento de diagnóstico en particular puede no proporcionar un

diagnóstico definitivo de una condición, es suficiente si el procedimiento proporciona una indicación positiva que ayuda en el diagnóstico.

Biomarcador transcriptómico/Firmas moleculares

5 En una realización preferida, un biomarcador (TBB-I) comprende secuencias de ácido nucleico/biomoléculas que comprenden: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (proteína en dedo de zinc 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína nuclear heterogénea H1 (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza a ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (proteína en dedo de zinc 623), 212637_s_at (dominio WW que contiene la proteína ubiquitina E3 ligasa 1), 212920_at, 213317_at (canal 5 intracelular de cloro), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y el dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con dominios 2 similares a EGF y dos folistatinas), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina del tipo aminotransferasa 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína quinasa nuclear y sustrato 1 de quinasa que depende de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOC197135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio 3 de tipo AN-1), 240971_x_at (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

En otra realización preferida, las biomoléculas comprenden: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (dedo de zinc de proteína 337), 1567100_at (homólogo Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína nuclear heterogénea H1 (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza a ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (proteína en dedo de zinc 623), 212637_s_at (dominio WW que contiene la proteína ligasa ubiquitina E3), 212920_at, 213317_at (canal 5 intracelular de cloro), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con dominios 2 similares a EGF y similares a dos folistatinas), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina aminotransferasa tipo 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína nuclear quinasa y sustrato 1 de quinasa que depende de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOC197135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio 3 de tipo AN1), 240971_x_at (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2) se sobreexpresan en pacientes con miocarditis en comparación con sujetos normales. En algunas formas de realización se sobreexpresan al menos diez biomoléculas.

En otra realización preferida, un biomarcador transcriptómico (TBB- II) comprende biomoléculas que comprenden: 1552419_s_at (la familia similar a la tirosina ligasa de tubulina, miembro 10), 1553212_at (queratina 78), 1555124_at (proteína hipotética MGC40574), 1556192_x_at (supresor 1 de metástasis), 1556320_at (similar a la estomatina 1 (EPB72)), 1556510_at (clon de ADNc IMAGE: 4796864), 1558484_s_at (repetición rica en leucina que contiene 27), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567410_at (dedo de zinc de proteína 135), 1568513_x_at (proteasa, serina, 1 (tripsina 1)), 1570408_at (serina/treonina quinasa 24 (homólogo de STE20, levadura)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido guanina), 204581_at (molécula CD22, glicoproteína asociada a mielina), 205586_x_at (factor de crecimiento nervioso inducible VGF), 206333_at (homólogo 1 de musashi (Drosophila)), 207004_at (CLL de células B / linfoma 2), 210059_s_at (proteína quinasa 13 activada por mitógeno), 210228_at (factor estimulante de colonias 2 (granulocito -macrófago)), 210384_at (proteína arginina metiltransferasa 2), 210923_at (familia 1 portadora de soluto (transportador de glutamato), miembro 7), 211024_s_at (factor 1 de transcripción de la tiroides /// factor 1 de transcripción de la tiroides), 211062_s_at (carboxipeptidasa Z /// carboxipeptidasa Z), 211096_at (factor 2 de transcripción de leucemia de pre-células B), 211181_x_at (factor 1 de transcripción relacionado con enanismo (leucemia mieloide aguda 1; oncogén aml1)), 211710_x_at (proteína ribosomal L4 /// proteína ribosomal L4), 213096_at (familia 2 del dominio transmembrana y de bobina en espiral), 213121_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña (antígeno RNP)), 213242_x_at (KIAA0284), 213568_at (impar omitida relacionada con 2 (Drosophila)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 216116_at (proteína que interactúa con NCK con el dominio SH3), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 216820_at , 217054_at (ADNc FLJ39484 fis, clon

PROST2014925), 217180_at (proteína hipotética similar al producto génico KIAA0187), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 217430_x_at (colágeno, tipo 1, alfa 1), 219070_s_at (dominio de esperma móvil que contiene 3), 219425_at (familia 4A de la sulfotransferasa, miembro 1), 221663_x_at (receptor H3 de histamina), 221684_s_at (Nictalopina), 223974_at (proteína hipotética MGC11082), 226640_at (diacilglicerol lipasa beta), 228074_at (proteína hipotética LOC162073), 229191_at (cofactor D de plegamiento de la tubulina), 229257_at (proteína KIAA1856), 229335_at (superfamilia de las inmunoglobulinas, miembro 4C), 229358_at (homólogo hedgehog indio (*Drosophila*)), 230341_x_at (metalopeptidasa ADAM con el motivo tipo 1 de trombospondina, 10), 230693_at (ATPasa, que transporta Ca⁺⁺, músculo cardíaco, contracción rápida 1), 230768_at (FERM, RhoGEF y dominio pleckstrina de la proteína 2), 231510_at (miembro de la familia GLI-Kruppel GLI2), 231629_x_at (peptidasa 3 relacionada con calicreína), 231998_at, 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 234637_at (proteína 4-5 relacionada con queratina), 234881_at, 235568_at (marco de lectura abierto 59 del cromosoma 19), 235600_at (locus transcrito), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (*Drosophila*)), 237087_at (marco de lectura abierto 105 del cromosoma 14), 237144_at (factor de crecimiento beta transformante latente que se enlaza a la proteína 3), 237398_at (locus transcrito), 237547_at (proteína hipotética LOC728730), 237679_at (66 que contiene al motivo tripartita), 238267_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239026_x_at (centaurina, gamma 3), 239463_at (locus transcrito), 239756_at (tipo 1 deficiente en la captura mitótica de MAD1 (levadura)), 240039_at (locus transcrito /// locus transcrito), 240147_at (proteína hipotética MGC11257), 240517_at (cistationina beta sintasa), 241270_at (homólogo 2 romboide 5 (*Drosophila*)), 241431_at, 242365_at (dominio de bobina en espiral que contiene 32), 243297_at (proteína vacuolar para clasificar al homólogo D 13 (*S. cerevisiae*)), 243497_at (locus transcrito), 243766_s_at (miembro 2 de la familia del dominio TEA), 43934_at (receptor 137 acoplado a la proteína G), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

En otra realización preferida, las biomoléculas que comprenden el biomarcador transcriptómico (TBB-II) secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos se subregulan en un paciente que sufre miocarditis en comparación con un sujeto normal.

En otra realización preferida, un biomarcador (TBB-III) comprende secuencias de ácido nucleico/biomoléculas que comprenden: 1553212_at (queratina 78), 1557236_at (apolipoproteína L), 1558142_at (repetición de trinucleótidos que contiene 6B), 1558484_s_at (una repetición rica en leucina que contiene 27), 1565614_at (dedo de zinc de proteína 337), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567100_at (homólogo de Dachshund 1 (*Drosophila*)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 205758_at (molécula CD8a /// molécula CD8a), 206333_at (homólogo 1 de musashi (*Drosophila*)), 212920_at, 213242_x_at (KIAA0284), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 216427_at (ADNc: FLJ22786 fis, el clon KAIA2150), 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 219425_at (familia 4A de sulfotransferasa, miembro 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 229191_at (cofactor D de plegamiento de tubulina), 229569_at (clon de ADNc IMAGE: 5263455), 231629_x_at (peptidasa 3 relacionada con la calicreína), 233765_at (LOC197135 hipotético), 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 235568_at (marco de lectura abierto 59 del cromosoma 19), 235803_at (factor 3 del tipo receptor de citoquinas), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (*Drosophila*)), 236953_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239463_at (locus transcrito), 240544_at (dedo de cinc, dominio 3 de tipo AN1), 243766_s_at (miembro 2 de la familia del dominio TEA), y secuencias complementarias, fragmentos, alelos, variantes de 244042_x_at y productos génicos de los mismos.

En otra realización preferida, la detección de cualquiera de los biomarcadores, TBB-I, TBB-II y TBB-III o combinaciones de los mismos, son para diagnóstico de la miocarditis, cardiomiopatía idiopática, enfermedades del corazón y trastornos del mismo. En otra realización preferida, la detección de al menos diez biomoléculas en TBB-I es diagnóstica de la miocarditis, cardiomiopatía idiopática, enfermedades del corazón y trastornos del mismo.

En otra realización preferida, la detección de al menos diez moléculas en TBB-II es diagnóstica de la miocarditis, cardiomiopatía idiopática, enfermedades del corazón y trastornos del mismo. Puesto que las biomoléculas en TBB-II son subreguladas en miocarditis, cuando se usa el término "detección" con respecto a las biomoléculas de TBB-II, se referirá a la subregulación de las biomoléculas, en comparación con la expresión en células o sujetos normales o sanos.

En otra realización preferida, la detección de al menos diez moléculas en TBB-III es diagnóstica de la miocarditis, cardiomiopatía idiopática, enfermedades del corazón y trastornos del mismo.

En otra realización preferida, la detección en una célula o paciente de las biomoléculas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, es diagnóstico de la miocarditis,

cardiomiopatía idiopática, enfermedades del corazón y trastornos del mismo. Preferiblemente, las secuencias de biomoléculas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, son modulados en niveles de al menos entre 1 %, 2 %, 5 %, 10 % en una célula o paciente en comparación con los niveles en una célula normal o sujeto normal; más preferiblemente, las secuencias de biomarcadores de genes, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, son modulados en aproximadamente 50 % en una célula o un paciente, en comparación con los niveles en una célula normal o sujeto normal; más preferiblemente, las secuencias de biomarcadores de genes, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, son modulados en aproximadamente 75 % en una célula o un paciente, en comparación con los niveles en una célula normal o sujeto normal. El término "modulado" se refiere a un aumento o disminución en el nivel, concentración, cantidad, etc., en comparación con una célula normal o sujeto sano normal.

En otra realización preferida, un biochip comprende una firma molecular que comprende TBB-I, TBB-II, o TBB-III, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

En otra realización preferida, un biochip comprende al menos diez biomoléculas seleccionadas de las biomoléculas que comprenden TBB-I, TBB-II, o TBB-III, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos. Se pueden seleccionar cualquiera entre diez o más de los biomarcadores TBB-I, TBB-II, o TBB-III o combinaciones seleccionadas de los mismos. No hay nada que impida la adición de cualquier biomarcador identificado recientemente o cualesquiera otros biomarcadores conocidos o biomoléculas de los mismos.

En otra realización preferida, un procedimiento para diferenciar entre cardiomiopatía idiopática y miocarditis, comprende: la identificación en una muestra biológica de un paciente de una firma molecular que comprende un biomarcador TBB-I, TBB-II y TBB-III con base en transcriptómica.

En otra realización preferida, un procedimiento para diferenciar entre cardiomiopatía idiopática y miocarditis, comprende la identificación en una muestra biológica de un paciente de una firma molecular que comprende las biomoléculas del biomarcador basado en transcriptómica (TBB), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos; la evaluación de la probabilidad de identificación de cada gen componente en cada muestra; y la asignación de cada uno a una clase.

En otra realización preferida, las biomoléculas se seleccionan de TBB-I, TBB-II, y TBB-III, o combinaciones de los mismos.

En una realización preferida, la especificidad del fenotipo se identifica mediante la creación de un clasificador en un conjunto de entrenamiento que comprende aproximadamente 66 % de los datos obtenidos, con la validación posterior en un conjunto de prueba que comprende aproximadamente 33 % de los datos obtenidos y la definición de un centroide encogido más cercano al fenotipo específico para la clasificación.

En otra realización preferida, un procedimiento de diagnóstico de enfermedades del corazón o miocarditis comprende la identificación en una muestra biológica de un paciente de una firma molecular que comprende un biomarcador con base transcriptómica (TBB), TBB-I, TBB-II y TBB-III.

En otra realización preferida, un procedimiento de diagnóstico de enfermedades del corazón o miocarditis comprende la detección de al menos diez biomoléculas, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos en una muestra; evaluación de la probabilidad de identificación de cada gen componente en cada muestra; asignación de cada uno a una clase; y, el diagnóstico de enfermedades del corazón o miocarditis.

Procedimientos alternativos y materiales para la identificación de firmas moleculares o biomarcadores transcriptómicos

Detección de ácidos nucleicos y proteínas como marcadores: En realizaciones preferidas, se detecta cada biomarcador en procedimientos basados en chips, tal como los descritos en detalle en los ejemplos que se presentan más adelante. Con el fin de proporcionar un diagnóstico preciso de trastornos y enfermedades cardíacas, por ejemplo, insuficiencia cardíaca, miocarditis, cardiomiopatía idiopática y similares. También se conocen otros procedimientos en la técnica, y se pueden usar uno o más procedimientos.

Los procedimientos y ensayos descritos en este documento están dirigidos al examen de la expresión de biomarcadores transcriptómicos en un tejido de mamífero o muestra de células, en el que la determinación de que la expresión de uno o más de tales biomarcadores transcriptómicos es predictiva del resultado de pronóstico o diagnóstico de enfermedades y trastornos cardíacos y cardiovasculares, tales como por ejemplo, miocarditis, enfermedad coronaria, angina, síndrome coronario agudo, aneurisma de la aorta y disección, arritmias, cardiomiopatía, enfermedad cardíaca congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis, y similares. Las firmas moleculares o el biomarcador transcriptómico comprenden las biomoléculas identificadas en las Tablas 1, 2 y 4. Al hacer un diagnóstico es deseable detectar al menos 10 o más biomoléculas.

Cualquiera de TBB-I (Tabla 1), TBB-II (Tabla 2), TBB-III (Tabla 4) o combinaciones de los mismos se pueden usar en el diagnóstico de miocarditis. Cualquiera de TBB-I, TBB-II y TBB-III o combinaciones de los mismos se pueden usar en el diagnóstico de la miocarditis y la cardiomiopatía idiopática y la diferenciación entre las dos condiciones.

- 5 Las realizaciones preferidas en la identificación de biomoléculas, procedimientos analíticos, etc., se describen en detalle en los Ejemplos que se presentan más adelante.

Micromatrices: En general, el uso de micromatrices de ácido nucleico, muestras de prueba y de control de ARNm a partir de muestras de tejido de prueba y control se someten a transcripción inversa y se marcan para generar sondas de ADNc. Las sondas se hibridan luego con una matriz de ácidos nucleicos inmovilizados en un soporte sólido. La matriz está configurada de tal manera que la secuencia y la posición de cada miembro de la matriz se conoce. Por ejemplo, se puede disponer una selección de genes que tienen potencial para ser expresados en ciertos estados de enfermedad sobre un soporte sólido. La hibridación de una sonda marcada con un miembro particular de la matriz indica que la muestra de la que se derivó la sonda expresa ese gen. El análisis de la expresión génica diferencial de tejido enfermo puede proporcionar información valiosa. La tecnología de micromatrices usa técnicas de hibridación de ácido nucleico y tecnología informática para evaluar el perfil de expresión de ARNm de miles de genes dentro de un único experimento. (véase, por ejemplo, el documento WO 01/75166 publicado el 11 de octubre de 2001; (véase, por ejemplo, la patente estadounidense No. 5.700.637, la patente estadounidense No. 5.445.934, y la patente estadounidense No. 5.807.522, Lockart, Nature Biotechnology, 14: 1675-1680 (1996); Cheung, VG y colaboradores, Nature Genetics 21 (suplemento): 15-19 (1999) para una discusión de la fabricación de la matriz). Las micromatrices de ADN son matrices en miniatura que contienen fragmentos de genes que o bien son sintetizados directamente sobre, o colocados sobre vidrio u otros sustratos. Miles de genes se representan generalmente en una única matriz. Un experimento típico de micromatrices implica los siguientes pasos: 1) preparación del objetivo marcado con fluorescencia a partir de ARN aislado de la muestra, 2) hibridación del objetivo marcado con la micromatriz, 3) lavado, tinción y barrido de la matriz, 4) análisis de la imagen escaneada y 5) generación de perfiles de expresión génica. En la actualidad se están usando dos tipos principales de micromatrices de ADN: matrices de oligonucleótidos (por lo general de 25 a 70 mers) y matrices de expresión génica que contienen productos de PCR preparados a partir de ADNc. En la formación de una matriz, los oligonucleótidos pueden ser o bien prefabricados y colocados en la superficie o sintetizados directamente sobre la superficie (*in situ*). El sistema Affymetrix GENECHIP^{MR} es un sistema de micromatrices comercialmente disponible que comprende matrices fabricadas mediante síntesis directa de oligonucleótidos sobre una superficie de vidrio.

Matrices de sondas/genes: Oligonucleótidos, por lo general 25 mers, se sintetizan directamente sobre una oblea de vidrio mediante una combinación de fotolitografía con base en semiconductores y tecnologías de síntesis química en fase sólida. Cada matriz contiene hasta 400.000 oligonucleótidos diferentes y cada oligonucleótido está presente en millones de copias. Dado que las sondas de oligonucleótidos se sintetizan en sitios conocidos sobre la matriz, se pueden interpretar los patrones de hibridación y las intensidades de señal en términos de identidad génica y de niveles relativos de expresión mediante el software Affymetrix Microarray Suite. Cada gen está representado sobre la matriz por una serie de diferentes sondas de oligonucleótidos. Cada par de sondas consiste en un oligonucleótido que coincide perfectamente y un oligonucleótido que no lo hace. La sonda que coincide perfectamente tiene una secuencia exactamente complementaria al gen particular y, por tanto mide la expresión del gen. La sonda que no coincide difiere de la sonda que coincide perfectamente por una sola sustitución de bases en la posición de una base central, perturbando el enlazamiento del transcrito génico objetivo. Esto ayuda a determinar la hibridación de fondo y no específica que contribuye a la señal medida para el oligonucleótido que coincide perfectamente. El software Microarray Suite resta las intensidades de hibridación de las sondas que no coinciden de las de las sondas que coinciden perfectamente para determinar el valor de la intensidad absoluta o específica para cada conjunto de sondas. Las sondas se eligen con base en la información actual del GenBank y otros repositorios 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 de fluidos realiza el lavado y la tinción de las matrices de sondas. Está completamente automatizada y contiene cuatro módulos, conteniendo cada módulo una matriz de sondas. Cada módulo está controlado independientemente por el software Microarray Suite usando protocolos preprogramados de fluidos. El escáner es un escáner de fluorescencia láser confocal que mide la intensidad de la fluorescencia emitida por el ARNc marcado unido a las matrices de sondas. La estación de trabajo del ordenador con el software Microarray Suite controla la estación de fluidos y el escáner. El software Microarray Suite puede controlar hasta ocho estaciones de fluidos mediante hibridación previamente programada, lavado, y protocolos de tinción para la matriz de la sonda. El software también adquiere y convierte los datos de intensidad de hibridación en una llamada de presencia/ausencia para cada gen usando algoritmos apropiados. Finalmente, el software detecta cambios en la expresión génica entre experimentos mediante análisis de comparación y formatea la salida en archivos .txt que pueden usarse con otros programas de software para el análisis de datos.

La expresión de un biomarcador seleccionado también puede evaluarse examinando la supresión del gen o la amplificación del gen. La supresión o amplificación del gen se pueden medir mediante uno cualquiera de una amplia variedad de protocolos conocidos en la técnica, por ejemplo, mediante transferencia Southern convencional, transferencia Northern para cuantificar la transcripción de ARNm (Thomas, Proc. Natl. Acad. Sci. EE.UU., 77: 5201 - 5205 (1980)), transferencia de manchas (análisis de ADN), o hibridación *in situ* (por ejemplo, FISH), usando una

sonda marcada apropiadamente, procedimientos citogenéticos o hibridación genómica comparativa (CGH) usando una sonda marcada apropiadamente.

5 Detección de polipéptidos: En otra realización, se detecta un polipéptido correspondiente a un marcador. Un agente preferido para la detección de un polipéptido es un anticuerpo o aptámero capaz de unirse a un polipéptido correspondiente a un marcador de la invención, preferiblemente un anticuerpo con una etiqueta detectable. Los anticuerpos pueden ser policlonales, o más preferiblemente, monoclonales. Se pueden usar un anticuerpo intacto, o un fragmento del mismo, por ejemplo, Fab o F(ab')₂. El término "marcado", con respecto a la sonda o anticuerpo, pretende abarcar la marcación directa de la sonda o anticuerpo por acoplamiento, es decir, enlazando físicamente, 10 una sustancia detectable a la sonda o anticuerpo, así como marcación indirecta de la sonda o anticuerpo por reactividad con otro reactivo que está directamente marcado. Ejemplos de marcación indirecta incluyen la detección de un anticuerpo primario usando un anticuerpo secundario marcado en forma fluorescente y marcación final de una sonda de ADN con biotina de manera que puede ser detectada con estreptavidina marcada en forma fluorescente.

15 Las proteínas de los individuos se pueden aislar usando técnicas que son bien conocidas por los expertos en la técnica. Los procedimientos de aislamiento de proteínas empleados pueden, por ejemplo, ser tales como los descritos en Harlow y Lane (1988), ver más arriba. Se pueden emplear varios formatos para determinar si una muestra contiene una proteína que se une a un anticuerpo dado. La expresión de diversos biomarcadores en una muestra se puede analizar por una cantidad de metodologías, muchas de las cuales son conocidas en la técnica y comprendidas por el experto en la técnica, incluyendo, pero no limitado a, inmunohistoquímica y/o análisis Western, 20 ensayos cuantitativos con base en sangre (como por ejemplo ELISA en suero) (para examinar, por ejemplo, los niveles de expresión de proteína), ensayos de actividad enzimática bioquímica, hibridación *in situ*, análisis Northern y/o análisis por PCR de los ARNm, así como cualquiera de la amplia variedad de ensayos que pueden ser realizados mediante el análisis de una matriz génica y/o de tejido. Los protocolos típicos para evaluar el estado de genes y productos génicos se encuentran, por ejemplo, en Ausubel y colaboradores eds., 1995, Current Protocols in Molecular Biology, Unidades 2 (transferencia de Northern), 4 (transferencia de Southern), 15 (inmunotransferencia) y 25 18 (análisis por PCR). Un experto puede adaptar fácilmente los procedimientos de detección conocidos de proteína/anticuerpo para uso en la determinación de si las células expresan un marcador de la presente invención y la concentración relativa de ese producto específico de expresión del polipéptido en la sangre u otros tejidos del cuerpo. 30

En tales procedimientos alternativos, puede ponerse en contacto una muestra con un anticuerpo específico para dicho biomarcador en condiciones suficientes para que se forme un complejo anticuerpo-biomarcador, y luego detectar dicho complejo. La presencia del biomarcador puede ser detectada en una cantidad de formas, tales como 35 mediante transferencia Western y procedimientos de ELISA para el ensayo de una amplia variedad de tejidos y muestras, incluyendo plasma o suero. Se encuentra disponible una amplia variedad de técnicas de inmunoensayo usando un formato de ensayo de este tipo, véase, por ejemplo, las patentes estadounidenses Nos. 4.016.043, 4.424.279 y 4.018.653. Estas incluyen tanto ensayos de sitio único y de dos sitios o tipo "sándwich" de los tipos no competitivos, así como en los ensayos tradicionales de enlazamiento competitivo. Estos ensayos también incluyen el enlazamiento directo de un anticuerpo marcado a un biomarcador objetivo. 40

Los ensayos tipo sándwich se encuentran entre los más útiles y más empleados. Existe una cantidad de variaciones de la técnica de ensayo tipo sándwich. Brevemente, en un ensayo directo típico, se inmoviliza un anticuerpo no marcado sobre un sustrato sólido, y se pone en contacto la muestra a ensayar con la molécula unida. Después de 45 un período de incubación adecuado, durante un período de tiempo suficiente para permitir la formación de un complejo anticuerpo-antígeno, se añade luego un segundo anticuerpo específico para el antígeno, marcado con una molécula informadora capaz de producir una señal detectable y se incuba, permitiendo el tiempo suficiente para la formación de otro complejo de anticuerpo marcado con anticuerpo-antígeno. Cualquier material sin reaccionar se elimina por lavado, y se determina la presencia del antígeno mediante la observación de una señal producida por la molécula informadora. Los resultados pueden ser cualitativos, por simple observación de la señal visible, o pueden 50 cuantificarse por comparación con una muestra de control que contiene cantidades conocidas de biomarcador.

Las variaciones sobre el ensayo directo incluyen un ensayo simultáneo, en el que se añaden tanto la muestra como el anticuerpo marcado de forma simultánea al anticuerpo enlazado. Estas técnicas son bien conocidas por los 55 expertos en la técnica, incluyendo cualesquiera pequeñas variaciones como será fácilmente evidente. En un ensayo típico tipo sándwich directo, se enlaza un primer anticuerpo que tiene especificidad por el biomarcador ya sea en forma covalentemente o pasiva a una superficie sólida. La superficie sólida es normalmente vidrio o un polímero, siendo los polímeros usados más comúnmente celulosa, poli(acrilamida), nailon, poliestireno, cloruro de polivinilo o polipropileno. Los soportes sólidos pueden estar en forma de tubos, perlas, discos de microplacas, o cualquier otra 60 superficie adecuada para la realización de un inmunoensayo. Los procesos de enlazamiento son bien conocidos en la técnica y generalmente consisten en entrecruzamiento mediante enlazamiento covalente o adsorción física, se lava el complejo polímero-anticuerpo en preparación para la muestra de ensayo. A continuación se añade una alícuota de la muestra a ensayar al complejo en fase sólida y se incuba durante un período de tiempo suficiente (por ejemplo, 2-40 minutos o durante la noche si es más conveniente) y bajo condiciones adecuadas (por ejemplo, de temperatura ambiente hasta 40° C, tal como entre 25° C y 32° C, inclusive) para permitir el enlazamiento de cualquier subunidad presente en el anticuerpo. Tras el período de incubación, se lava la fase sólida de la subunidad 65

del anticuerpo y se seca y se incuba con un segundo anticuerpo específico para una porción del biomarcador. El segundo anticuerpo se enlaza a una molécula informadora que se usa para indicar el enlazamiento del segundo anticuerpo al marcador molecular.

5 Un procedimiento alternativo implica inmovilizar los biomarcadores objetivo en la muestra y luego exponer el objetivo inmovilizado a un anticuerpo específico que puede o no estar marcado con una molécula informadora. Dependiendo de la cantidad de objetivo y de la fuerza de la señal de la molécula informadora, se puede detectar un objetivo enlazado mediante marcación directa con el anticuerpo. Alternativamente, se expone un segundo anticuerpo marcado, específico para el primer anticuerpo al complejo primer anticuerpo-objetivo para formar un complejo terciario de objetivo-primer anticuerpo-segundo anticuerpo. El complejo se detecta por la señal emitida por la molécula informadora. Por "molécula informadora", como se usa en la presente memoria, se entiende una molécula que, por su naturaleza química, proporciona una señal analíticamente identificable que permite la detección del anticuerpo enlazado al antígeno. Las moléculas informadoras más comúnmente usadas en este tipo de ensayo son o bien 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 variedad de diferentes técnicas de conjugación, que están fácilmente disponibles para el experto en la materia. Enzimas usadas comúnmente incluyen peroxidasa de rábano picante, glucosa oxidasa, galactosidasa y fosfatasa alcalina, entre otras. Los sustratos a usar con las enzimas específicas se eligen generalmente para la producción, tras la hidrólisis por la enzima correspondiente, de un cambio de color detectable. 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 mencionados anteriormente. En todos los casos, se añade el anticuerpo marcado con enzima al complejo primer anticuerpo-marcador molecular, se permite que se enlace, y luego se lava el exceso de reactivo. Se añade luego una solución que contiene el sustrato apropiado al complejo de anticuerpo-antígeno-anticuerpo. El sustrato reaccionará con la enzima ligada al segundo anticuerpo, dando una señal visual cualitativa, que puede cuantificarse adicionalmente, usualmente espectrofotométricamente, para dar una indicación de la cantidad de biomarcador que estaba presente en la muestra. Alternativamente, compuestos fluorescentes, tales como fluoresceína y rodamina, pueden acoplarse químicamente a anticuerpos sin alterar su capacidad de enlazamiento. Cuando se activa por iluminación con luz de una longitud de onda particular, el anticuerpo marcado con fluorocromo adsorbe la energía de la luz, induciendo un estado de excitabilidad en la molécula, seguido de emisión de la luz a un color característico detectable visualmente con un microscopio de luz. Como en el EIA, se permite que el anticuerpo marcado en forma fluorescente se enlace al complejo primer anticuerpo-marcador molecular. Después de lavar el reactivo no enlazado, se expone luego el complejo terciario restante a la luz de la longitud de onda apropiada, la fluorescencia observada indica la presencia del marcador molecular de interés. Técnicas de inmunofluorescencia y EIA están ambas muy bien establecidas en la técnica. Sin embargo, otras moléculas informadoras, tales como moléculas de radioisótopos, quimioluminiscentes o bioluminiscentes, pueden también ser empleadas. Los procedimientos de la invención incluyen además los protocolos que examinan la presencia y/o expresión de los ARNm, en una muestra de tejido o célula. Los procedimientos para la evaluación de los ARNm en células son bien conocidos 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 ácidos nucleicos (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, se puede determinar el nivel de ARNm correspondiente al marcador tanto mediante formatos *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 los procedimientos *in vitro*, se puede usar cualquier técnica de aislamiento de ARN que no seleccione contra el aislamiento de ARNm para la purificación de ARN de las células. Véase, por ejemplo, Ausubel y colaboradores, Ed., Curr. Prot. Mol. Biol., John Wiley & Sons, Nueva York (1987-1999). Adicionalmente, se pueden procesar fácilmente grandes cantidades de muestras de tejido usando técnicas bien conocidas por los expertos en la técnica, tales como, por ejemplo, el proceso de aislamiento de ARN de un solo paso de la patente estadounidense No. 4.843.155. El ARNm aislado puede usarse en ensayos de hibridación o amplificación que incluyen, pero no se limitan a, análisis 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 está siendo detectado. La sonda de ácido nucleico puede ser, por ejemplo, un ADNc de longitud completa, o una porción del mismo, tal como un oligonucleótido de al menos 7, 15, 30, 50, 100, 250 o 500 nucleótidos de longitud y suficiente para hibridar específicamente bajo condiciones rigurosas con un ARNm o ADN genómico que codifica un marcador de la presente invención. Moléculas de ácido nucleico para uso en los ensayos de diagnóstico de la invención se exponen en las reivindicaciones. La hibridación de un ARNm con la sonda indica que el marcador en cuestión está siendo expresado.

65 En un formato, el ARNm se inmoviliza sobre una superficie sólida y en contacto con una sonda, por ejemplo, mediante la corrida del ARNm aislado en un gel de agarosa y transfiriendo el ARNm del gel a una membrana, tal

como nitrocelulosa. En un formato alternativo, la(s) sonda(s) se inmovilizan sobre una superficie sólida y el ARNm se pone en contacto con la(s) sonda(s), por ejemplo, en una matriz de chips de genes Affymetrix. Un experto puede adaptar fácilmente los procedimientos de detección de ARNm conocidos para uso en la detección del nivel de ARNm codificado por los marcadores de la presente invención.

5 Aunque no se requiere la amplificación de las moléculas en la presente invención como se discute en la sección de ejemplos, un experto en la técnica 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 involucra el proceso de amplificación de ácido nucleico, por ejemplo, por RT-PCR (la forma de realización experimental expuesta en Mullis, patente estadounidense No. 4.683.202 (1987); reacción en cadena de la ligasa, replicación de secuencia autosostenida, Guatelli y colaboradores, Proc. Natl. Acad. Sci. EE.UU., Vol 87, páginas 1874-1878 (1990); sistema de amplificación transcripcional, Kwoh y colaboradores, Proc. Natl. Acad. Sci. EE.UU., Vol 86, páginas 1173-1177 (1989); Q-Beta replicasa, Lizardi y colaboradores, Biol. Technology, Vol 6, página 1197 (1988); replicación de círculo rodante, patente estadounidense No. 5.854.033 (1988); o cualquier otro procedimiento de amplificación de ácido nucleico, seguido por la detección de las moléculas amplificadas usando técnicas bien conocidas por los expertos en la técnica. Estos esquemas de detección son especialmente útiles para la detección de las moléculas de ácido nucleico si tales moléculas están presentes en cantidades muy pequeñas. Como se usa aquí, los cebadores de amplificación se definen como un par de moléculas de ácido nucleico que pueden hibridarse con regiones 5' o 3' de un gen (más y menos hebras, respectivamente, o viceversa) y contienen una región corta en el medio. 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, tales cebadores permiten la amplificación de una molécula de ácido nucleico que comprende la secuencia de nucleótidos flanqueada por los cebadores.

25 Para procedimientos *in situ*, el ARNm no necesita ser aislado de las células antes de la detección. En tales procedimientos, se prepara/procesa una muestra de células o tejido usando procedimientos histológicos conocidos. La muestra se inmoviliza luego sobre un soporte, normalmente un portaobjetos de vidrio, y luego se pone en contacto con una sonda que puede hibridar con ARNm que codifica el marcador.

30 Como una alternativa a las 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 mediante la comparación de su expresión con la expresión de un gen que no es un marcador, por ejemplo, un gen de mantenimiento que se expresa constitutivamente. Los genes adecuados para la normalización incluyen genes de mantenimiento, tales como el gen de actina o genes específicos de células epiteliales. Esta normalización permite la comparación del nivel de expresión en una muestra, por ejemplo, una muestra del paciente, con otra muestra o entre muestras de diferentes fuentes.

40 Alternativamente, el nivel de expresión se puede proporcionar 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 versus enfermas, preferiblemente 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 mayor número de muestras y se usa como un nivel de expresión base para el marcador. El nivel de expresión del marcador determinado para la muestra de ensayo (nivel absoluto de expresión) se divide luego por el valor de expresión medio obtenido para ese marcador. Esto proporciona un nivel de expresión relativo.

50 Preferentemente, las muestras usadas en la determinación de línea base serán de pacientes que no tienen el polimorfismo. La elección de la fuente de células depende del uso del nivel de expresión relativo. El uso de la expresión que se encuentra en tejidos normales como una puntuación de expresión media ayuda a la validación si el marcador ensayado es específico (frente a células normales). Además, entre más datos se acumulen, se puede revisar el valor de expresión medio, proporcionando valores de expresión relativos mejorados con base en los datos acumulados.

55 Anticuerpos y aptámeros

En una realización preferida, los anticuerpos y aptámeros enlazan específicamente 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 los mismos de cada componente en cada biomarcador.

Los polinucleótidos aptámeros son normalmente ADN fosfodiéster estándar monocatenario (ADNmc). Los análogos cercanos de ADN también se pueden incorporar en el aptámero, como se describe a continuación. Un procedimiento de descubrimiento típico de aptámero se describe a continuación:

65

Se sintetiza un polinucleótido que comprende una secuencia al azar entre "brazos" que tienen secuencia constante. Los brazos pueden incluir sitios de restricción para la clonación conveniente y pueden también funcionar como sitios de cebado para cebadores de PCR. La síntesis se puede realizar fácilmente en instrumentos comerciales.

5 La proteína objetivo se trata con el polinucleótido aleatorizado. La proteína objetivo puede estar en solución y luego se inmovilizan y separan los complejos de los ácidos nucleicos no enlazados mediante el uso de una columna de afinidad de anticuerpos. Alternativamente, la proteína objetivo se puede inmovilizar antes del tratamiento con el polinucleótido aleatorizado.

10 Los complejos de proteína objetivo-polinucleótido se separan del material no complejado y luego se separan los polinucleótidos enlazados de la proteína objetivo. El ácido nucleico enlazado, se puede caracterizar luego, pero es más comúnmente amplificado, por ejemplo, mediante PCR y se repiten las etapas de enlazamiento, separación y amplificación. En muchos casos, el uso de condiciones que promueven en forma creciente la separación del ácido nucleico de la proteína objetivo, por ejemplo, mayor concentración de sal, en el regulador de enlazamiento usado en la etapa 2) en iteraciones subsiguientes, resulta en la identificación de polinucleótidos que tienen cada vez más alta afinidad por la proteína objetivo.

15 Los ácidos nucleicos que muestran alta afinidad por las proteínas objetivo se aíslan y se caracterizan. Esto se realiza normalmente clonando los ácidos nucleicos usando sitios de restricción incorporados en los brazos, y luego secuenciando el ácido nucleico clonado.

20 La afinidad de los aptámeros por sus proteínas objetivo está normalmente en el intervalo nanomolar, pero puede ser tan bajo como el intervalo picomolar. Es decir K_D es normalmente de 1 pM a 500 nM, más normalmente 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.

25 Los polinucleótidos aptámeros pueden ser sintetizados en un sintetizador de ácidos nucleicos comercialmente disponible mediante procedimientos conocidos en la técnica. El producto se puede purificar mediante procedimientos de selección por tamaño o cromatográficos.

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

35 Las secuencias del aptámero pueden elegirse como una secuencia deseada, o se pueden elaborar poblaciones al azar o parcialmente al azar de secuencias y luego se seleccionan por el enlazamiento específico con una proteína objetivo deseada mediante un ensayo *in vitro*. Se pueden usar cualquiera de los ensayos típicos de enlazamiento ácido nucleico-proteína típicos conocidos en la técnica, por ejemplo, transferencia tipo "Southwestern" usando ya sea un oligonucleótido marcado o una proteína marcada como sonda. Véase también la patente estadounidense No. 5.445.935 para un ensayo de polarización por fluorescencia de interacción proteína-ácido nucleico.

40 Los nucleótidos apropiados para la síntesis de aptámeros y su uso, y los reactivos para el enlace covalente de las proteínas con ácidos nucleicos y su uso, se consideran conocidos en la técnica. Un complejo deseado de aptámero-proteína, por ejemplo, complejo aptámero-trombina se puede marcar y usar como agente de diagnóstico, *in vitro* casi de la misma forma que cualquier agente de enlazamiento de proteínas específicas, por ejemplo, un anticuerpo monoclonal. Por lo tanto, un complejo de aptámero-proteína se puede usar para detectar y cuantificar la cantidad de su proteína objetivo en una muestra, por ejemplo, una muestra de sangre, para proporcionar un diagnóstico de un estado de enfermedad en correlación con la cantidad de la proteína en la muestra.

45 También se puede usar un complejo molecular deseado de aptámero-objetivo/cebo para formación de imágenes de diagnóstico. En usos de formación de imágenes, los complejos se marcan de forma que puedan ser detectados fuera del cuerpo. Etiquetas típicas son radioisótopos, por lo general los que tienen una vida media corta. Se pueden usar radioisótopos usuales para formación de imágenes, tales como ^{123}I , ^{124}I , ^{125}I , ^{131}I , $^{99\text{m}}\text{Tc}$, ^{186}Re , ^{188}Re , ^{64}Cu , ^{67}Cu , ^{212}Bi , ^{213}Bi , ^{67}Ga , ^{90}Y , ^{111}In , ^{18}F , ^3H , ^{14}C , ^{35}S o ^{32}P . Se pueden usar también mejoradores de la formación de imágenes de resonancia magnética nuclear (RMN), tales como gadolinio 153, para marcar el complejo para detección por RMN. Los procedimientos y reactivos para realizar el etiquetado, ya sea en el polinucleótido o en la fracción de la proteína, se consideran conocidos en la técnica.

50 En una realización preferida, un anticuerpo o aptámero es específico para cada biomolécula de biomarcador (TBB-I) que comprende: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (proteína en dedo de zinc 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031Ts_at (Ribonucleoproteína nuclear heterogénea H1 (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza a ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (proteína en dedo de zinc 623), 212637_s_at (dominio WW que contiene proteína ligasa 1 ubiquitina E3), 212920_at, 213317_at (canal 5 intracelular de cloro), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 215443_at

(receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbinda y el dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, LNG12403 clon), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con dominios 2 similares a EGF y similares a dos folistatinas), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina aminotransferasa tipo 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína quinasa nuclear y sustrato 1 de quinasa dependiente de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOC1 97135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio 3 de tipo AN-1), 240971_x_at (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

En una realización preferida, un anticuerpo o aptámero es específico para cada biomolécula de biomarcador (TBB-II) que comprende: 1552419_s_at (familia similar a la tirosina ligasa de tubulina, miembro 10), 1553212_at (queratina 78), 1555124_at (proteína hipotética MGC40574s), 1556192_x_at (supresor 1 de metástasis), 1556320_at (similar a la estomatina 1 (EPB72)), 1556510_at (clon de ADNc IMAGE: 4796864), 1558484_s_at (repetición rica en leucina que contiene 27), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567410_at (dedo de zinc de proteína 135), 1568513_x_at (proteasa, serina, 1 (tripsina 1)), 1570408_at (serina/treonina quinasa 24 (homólogo de STE20, levadura)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido guanina), 204581_at (molécula CD22, glicoproteína asociada a mielina), 205586_x_at (factor de crecimiento nervioso inducible VGF), 206333_at (homólogo 1 de musashi (Drosophila)), 207004_at (CLL de células B/linfoma 2), 210059_s_at (proteína quinasa 13 activada por mitógeno), 210228_at (factor 2 estimulador de colonias (granulocito/macrófago)), 210384_at (proteína arginina metiltransferasa 2), 210923_at (familia 1 portadora de soluto (transportador de glutamato), miembro 7), 211024_s_at (factor 1 de transcripción de la tiroides /// factor 1 de transcripción de la tiroides), 211062_s_at (carboxipeptidasa Z /// carboxipeptidasa Z), 211096_at (factor 2 de transcripción de leucemia de pre-células B), 211181_x_at (factor 1 de transcripción relacionado con enanismo (leucemia mieloide aguda 1; oncogén aml1)), 211710_x_at (proteína ribosomal L4 /// proteína ribosomal L4), 213096_at (familia 2 del dominio transmembrana y de bobina en espiral), 213121_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña (antígeno RNP)), 213242_x_at (KIAA0284), 213568_at (impar omitida relacionada con 2 (Drosophila)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 216116_at (proteína que interactúa con NCK con el dominio SH3), 216427_at (ADNc: FLJ22786 fis, clon KAlA2150), 216820_at, 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217180_at (proteína hipotética similar al producto génico KIAA0187), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 217430_x_at (colágeno, tipo I, alfa 1), 219070_s_at (dominio de esperma móvil que contiene 3), 219425_at (familia 4A de la sulfotransferasa, miembro 1), 221663_x_at (receptor H3 de histamina), 221684_s_at (Nictalopina), 223974_at (proteína hipotética MGC11082), 226640_at (diacilglicerol lipasa beta), 228074_at (proteína hipotética LOC162073), 229191_at (cofactor D de plegamiento de la tubulina), 229257_at (proteína KIAA1856), 229335_at (superfamilia de las inmunoglobulinas, miembro 4C), 229358_at (homólogo hedgehog indio (Drosophila)), 230341_x_at (metalopeptidasa ADAM con el motivo tipo 1 de trombospondina, 10), 230693_at (ATPasa, que transporta Ca⁺⁺, músculo cardíaco, contracción rápida 1), 230768_at (FERM, RhoGEF y dominio pleckstrina de la proteína 2), 231510_at (miembro de la familia GLI-Kruppel GLI2), 231629_x_at (peptidasa 3 relacionada con calicreína), 231998_at, 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 234637_at (proteína 4-5 relacionada con queratina), 234881_at, 235568_at (marco de lectura abierto 59 del cromosoma 19), 235600_at (locus transcrito), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 237087_at (marco de lectura abierto 105 del cromosoma 14), 237144_at (factor de crecimiento beta transformante latente que se enlaza a la proteína 3), 237398_at (locus transcrito), 237547_at (proteína hipotética LOC728730), 237679_at (66 que contiene al motivo tripartita), 238267_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239026_x_at (centaurina, gamma 3), 239463_at (locus transcrito), 239756_at (tipo 1 deficiente en la captura mitótica de MAD1 (levadura)), 240039_at (locus transcrito /// locus transcrito), 240147_at (proteína hipotética MGC11257), 240517_at (cistationina beta sintasa), 241270_at (homólogo 2 romboide 5 (Drosophila)), 241431_at, 242365_at (dominio de bobina en espiral que contiene 32), 243297_at (proteína vacuolar para clasificar al homólogo D 13 (S. cerevisiae)), 243497_at (locus transcrito), 243766_s_at (miembro 2 de la familia del dominio TEA), 43934_at (receptor 137 acoplado a la proteína G), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

En una realización preferida, un anticuerpo o aptámero es específico para cada biomolécula de biomarcador (TBB-III) que comprende: 1553212_at (queratina 78), 1557236_at (apolipoproteína L), 1558142_at (repetición de trinucleótidos que contiene 6B), 1558484_s_at (repetición rica en leucina que contiene 27), 1565614_at (dedo de zinc de proteína 337), 1565662_at (mucina 6, formadora de moco/gel oligomérico), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 205758_at (molécula CD8a /// molécula CD8a), 206333_at (homólogo 1 de musashi (Drosophila)), 212920_at, 213242_x_at (KIAA0284), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 213770_at (supresor quinasa de ras 1), 214171_s_at (Factor 2 auxiliar de ARN nuclear pequeño U2), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 216427_at (ADNc: FLJ22786

5 fis, clon KAIA2150), 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217182_at (mucina 5AC, formadora de moco/gel oligomérico), 217322_x_at, 219425_at (familia 4A de sulfotransferasa, miembro 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 229191_at (cofactor D de plegamiento de tubulina), 229569_at (clon de ADNc IMAGE: 5263455), 231629_x_at (peptidasa 3 relacionada con calicreína), 233765_at (LOC197135 hipotético), 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 235568_at (marco de lectura abierto 59 del cromosoma 19), 235803_at (factor 3 del tipo receptor de citoquinas), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (*Drosophila*)), 236953_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isoenzima B), 239463_at (locus transcrito), 240544_at (dedo de zinc, dominio 3 de tipo AN-1), 243766_s_at (miembro de la familia de dominio TEA 2), y 244042_x_at secuencias, fragmentos, alelos, variantes y productos génicos complementarias de los mismos.

Descubrimiento del fármaco

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

20 Moléculas pequeñas: los compuestos de ensayo de molécula pequeña de moléculas o los compuestos terapéuticos candidatos pueden ser inicialmente miembros de una biblioteca química orgánica o inorgánica. En la presente memoria, "moléculas pequeñas" se refiere a moléculas orgánicas o inorgánicas pequeñas de peso molecular inferior a aproximadamente 3.000 Daltons. Las moléculas pequeñas pueden ser productos naturales o miembros de una biblioteca química combinatoria. Un conjunto de diversas moléculas se debe usar para cubrir varias funciones tales como la carga, aromaticidad, enlaces de hidrógeno, flexibilidad, tamaño, longitud de cadena lateral, hidrofobicidad, y rigidez. Las técnicas combinatorias adecuadas para síntesis de moléculas pequeñas son conocidos en la técnica, por ejemplo, como se ejemplifica por Obrecht y Villalgorido, 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 de "división y combinación", o en "paralelo", técnicas en fase sólida y en fase de solución, y técnicas de codificación (véase, por ejemplo, Czarnik, Curr. Opin. Chem. Bio., 1:60 (1997). Además, se encuentran comercialmente disponibles una cantidad de bibliotecas de moléculas pequeñas.

30 Las aplicaciones de cribado particular, se relacionan con la prueba de compuestos farmacéuticos en investigación de fármacos. Se remite al lector generalmente al libro de texto estándar "In vitro Methods in Pharmaceutical Research", Academic Press, 1997, y la patente estadounidense No. 5.030.015). La evaluación de la actividad de los compuestos farmacéuticos candidatos generalmente implica la administración de un compuesto candidato, la determinación de cualquier cambio en la morfología, fenotipo marcador y de expresión, o actividad metabólica de las células y función de las células que sea atribuible al compuesto (comparado con células no tratadas o con células tratadas con un compuesto inerte), y luego correlacionar el efecto del compuesto con el cambio observado.

40 El cribado puede realizarse, por ejemplo, ya sea porque el compuesto está diseñado para tener un efecto farmacológico sobre determinados tipos de células, o porque un compuesto diseñado para tener efectos en otros lugares puede tener efectos secundarios no deseados. Dos o más fármacos pueden ser probados en combinación (mediante la combinación con las células ya sea de forma simultánea o secuencialmente), para detectar posibles efectos de interacción fármaco-fármaco. En algunas aplicaciones, los compuestos se criban inicialmente para toxicidad potencial (Castell y colaboradores, páginas 375-410 en "In vitro Methods in Pharmaceutical Research", Academic Press, 1997). Se puede determinar la citotoxicidad en primer lugar por el efecto sobre la viabilidad, supervivencia, morfología, y la expresión celular o la liberación de ciertos marcadores, receptores o enzimas. Los efectos de un fármaco sobre el ADN cromosómico se pueden determinar mediante la medición de la síntesis o reparación del ADN. La incorporación de [³H] timidina o BrdU, especialmente en tiempos no programados en el ciclo celular, o por encima del nivel requerido para la replicación celular, es consistente con un efecto del fármaco. Los efectos no deseados también pueden incluir tasas inusuales de intercambio de cromátidas hermanas, determinados por la propagación de la metafase. Se remite al lector a A. Vickers (páginas 375-410 en "In vitro Methods in Pharmaceutical Research", Academic Press, 1997) para una elaboración adicional.

55 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 no en contacto con el agente candidato; (c) observar el efecto del agente candidato comparando el nivel de expresión del biomarcador o biomarcadores en la alícuota de la muestra biológica en contacto con el agente candidato y el nivel de expresión del biomarcador o 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 donde al menos una diferencia de 1 %, 2 %, 5 %, 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 una indicación 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 que se presenten más adelante.

En otra realización, se deriva un agente candidato mediante el procedimiento.

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

En otra realización preferida, un procedimiento de producción de un fármaco que comprende las etapas del procedimiento (i) sintetizar el agente candidato identificado en el paso (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 a fármaco, el agente candidato identificado en el paso (c) anterior, o un análogo o derivado del mismo con un vehículo farmacéuticamente aceptable.

Vectores, células: En algunas realizaciones es deseable expresar las biomoléculas que comprenden un biomarcador, en un vector y en células. Las aplicaciones de tales combinaciones son ilimitadas. Los vectores y células que expresan las una o más biomoléculas se pueden usar en ensayos, kits, el descubrimiento de fármacos, diagnósticos, pronósticos y similares. Las células pueden ser células madre aisladas de la médula ósea como una células progenitora, o células obtenidas de cualquier otra fuente, como por ejemplo, ATCC.

"Célula progenitora derivada de la médula ósea" (BMDC) o "célula madre derivada de médula ósea" se refiere a una célula madre primitiva, con la maquinaria para autorrenovación constitutivamente activa. En esta definición se incluyen células madre que son totipotentes, pluripotentes y precursores. Una "célula precursora" puede ser cualquier célula en una ruta de diferenciación celular que es capaz de diferenciarse en una célula más madura. Como tal, el término "población de células precursoras" se refiere a un grupo de células capaces de desarrollarse en células más maduras. Una población de células precursoras puede comprender células que son totipotentes, células que son pluripotentes y células que son un linaje restringido de células madre (es decir, células capaces de desarrollarse en menos de todos los linajes hematopoyéticos, o, por ejemplo, únicamente en células de linaje eritroide). Tal como se usa en el presente documento, el término "célula totipotente" se refiere a una célula capaz de desarrollarse en todos los linajes de células. Del mismo modo, el término "población totipotente de células" se refiere a una composición de células capaces de desarrollarse en todos los linajes de células. También, como se usa aquí, el término "células pluripotentes" se refiere a una célula capaz de desarrollarse en una variedad (aunque no todos) los linajes y son al menos capaces de desarrollarse en todos los linajes hematopoyéticos (por ejemplo, linajes linfoides, eritroide, y trombocítico). Las células madre derivadas de la médula ósea contienen dos tipos bien caracterizados de células madre. Las células madre mesenquimales (MSC) normalmente forman condrocitos y osteoblastos. Las células madre hematopoyéticas (HSC) son de origen mesodérmico que normalmente dan lugar a células de la sangre y del sistema inmune (por ejemplo, linajes eritroides, granulocitos/macrófagos, megacariocitos y linajes linfoides). Además, las células madre hematopoyéticas también han demostrado tener el potencial para diferenciarse en células del hígado (incluyendo hepatocitos, células del conducto biliar), pulmón, riñón (por ejemplo, células epiteliales tubulares renales y el parénquima renal), el tracto gastrointestinal, fibras de músculo esquelético, astrocitos del SNC, neuronas de Purkinje, músculo cardíaco (por ejemplo, cardiomiocitos), endotelio y de la piel.

En una realización preferida, un procedimiento de identificación de compuestos terapéuticos candidatos comprende el cultivo de células que expresan al menos una biomolécula seleccionada del biomarcador (TBB-1) que tienen secuencias de ácido nucleico/biomoléculas que comprenden: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (dedo de zinc de proteína 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína nuclear heterogénea H1 (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza a ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (proteína en dedo de zinc 623), 212637_s_at (dominio WW que contiene proteína ligasa 1 ubiquitina E3), 212920_at, 213317_at (canal intracelular de cloro 5), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con un dominio del tipo EGF y dos dominios del tipo de la folistatina 2), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina aminotransferasa tipo 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína nuclear quinasa y sustrato de quinasa dependiente de ciclina 1), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOC197135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio 3 de tipo AN1), 240971_x_at (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (tazaroteno inducida) 2), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, con un agente terapéutico candidato; la identificación de agentes terapéuticos candidatos que modulan la expresión de las biomoléculas y la identificación de un agente terapéutico candidato. Preferiblemente, 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.

5 En una realización preferida, un procedimiento de identificación de compuestos terapéuticos candidatos comprende el cultivo de células que expresan al menos una biomolécula seleccionada de: un biomarcador (TBB-I) que tienen
 10 secuencias de ácido nucleico/biomoléculas que comprenden: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (dedo de zinc de proteína 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína nuclear heterogénea H1 (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza a ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (proteína en dedo de zinc 623), 212637_s_at (dominio WW que contiene proteína ligasa 1 ubiquitina E3), 212920_at, 213317_at (canal intracelular de cloro 5), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con un dominio del tipo EGF y dos dominios del tipo de la folistatina 2), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina aminotransferasa tipo 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína nuclear quinasa y el sustrato 1 de quinasa que depende de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOC197135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio 3 de tipo AN 1), 240971_x_at (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (tazaroteno inducida) 2), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos, secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

En una realización preferida, un procedimiento de identificación de compuestos terapéuticos candidatos comprende el cultivo de células que expresan al menos una biomolécula seleccionada de: un biomarcador (TBB-II) que tiene
 30 secuencias de ácido nucleico/biomoléculas que comprenden: 1552419_s_at (familia del tipo tirosina ligasa de tubulina, miembro 10), 1553212_at (queratina 78), 1555124_at (proteína hipotética MGC40574), 1556192_x_at (supresor 1 de metástasis), 1556320_at (similar a la estomatina 1 (EPB72)), 1556510_at (clon de ADNc IMAGE: 4796864), 1558484_s_at (repetición rica en leucina que contiene 27), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567410_at (dedo de zinc de proteína 135), 1568513_x_at (proteasa, serina, 1 (tripsina 1)), 1570408_at (serina/treonina quinasa 24 (homólogo de STE20, levadura)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 204581_at (molécula CD22, glicoproteína asociada a mielina), 205586_x_at (factor de crecimiento nervioso inducible VGF), 206333_at (homólogo 1 de musashi (Drosophila)), 207004_at (CLL de células B/linfoma 2), 210059_s_at (proteína quinasa 13 activada por mitógeno), 210228_at (factor 2 estimulante de colonias (granulocito-macrófago)), 210384_at (proteína arginina metiltransferasa 2), 210923_at (familia 1 portadora de soluto (transportador de glutamato), miembro 7), 211024_s_at (factor 1 de transcripción de la tiroides /// factor 1 de transcripción de la tiroides), 211062_s_at (carboxipeptidasa Z /// carboxipeptidasa Z), 211096_at (factor 2 de transcripción de leucemia de pre-células B), 211181_x_at (factor 1 de transcripción relacionado con enanismo (leucemia mielode aguda 1; oncogén aml1)), 211710_x_at (proteína ribosomal L4 /// proteína ribosomal L4), 213096_at (familia 2 del dominio transmembrana y de bobina en espiral), 213121_at (polipéptido de 70 kDa de ribonucleoproteína nuclear pequeña (antígeno RNP)), 213242_x_at (KIAA0284), 213568_at (impar omitida relacionada con 2 (Drosophila)), 213770_at supresor quinasa de ras 1 (214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 216116_at (proteínas que interactúan con NCK con el dominio SH3), 216427_at (ADNc: FLJ22786 fis, clon KIAA2150), 216820_at, 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217180_at (proteína hipotética similar al producto génico KIAA0187), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 217430_x_at (colágeno, tipo I, alfa 1), 219070_s_at (dominio de esperma móvil que contiene 3), 219425_at (familia 4A de sulfotransferasa, miembro 1), 221663_x_at (receptor de histamina H3), 221684_s_at (Nictalopina), 223974_at (proteína hipotética MGC11082), 226640_at (diacilglicerol lipasa beta), 228074_at (proteína hipotética LOC162073), 229191_at (cofactor D de plegamiento de tubulina), 229257_at (proteína KIAA1856), 229335_at (superfamilia de inmunoglobulina, miembro 4C), 229358_at (homólogo hedgehog indio (Drosophila)), 2303411_x_at (metalopeptidasa ADAM con el motivo tipo 1 de trombospondina, 10), 230693_at (ATPasa, que transporta Ca⁺⁺, músculo cardíaco, contracción rápida 1), 230768_at (FERM, RhoGEF y dominio pleckstrina de la proteína 2), 231510_at (miembro de la familia GLI-Kruppel GLI2), 231629_x_at (peptidasa 3 relacionada con caliceína), 231998_at, 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con caliceína), 234637_at (proteína 4-5 relacionada con queratina), 234881_at, 235568_at (marco de lectura abierto 59 del cromosoma 19), 235600_at (locus transcrito), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 237087_at (marco de lectura abierto 105 del cromosoma 14), 237144_at (factor de crecimiento beta transformante latente que se enlaza a la proteína 3), 237398_at (locus transcrito), 237547_at (proteína hipotética LOC728730), 237679_at (66 que contiene al motivo tripartita), 238267_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239026_x_at (centaurina, gamma 3), 239463_at (locus transcrito), 239756_at (tipo 1 deficiente en la captura mitótica de MAD1 (levadura)), 240039_at (locus transcrito /// locus

transcrito), 240147_at (proteína hipotética MGC11257), 240517_at (cistationina beta sintasa), 241270_at (homólogo 2 romboide 5 (*Drosophila*)), 241431_at, 242365_at (dominio de bobina en espiral que contiene 32), 243297_at (proteína vacuolar para clasificar al homólogo D 13 (*S. cerevisiae*)), 243497_at (locus transcrito), 243766_s_at (miembro 2 de la familia del dominio TEA), 43934_at (receptor 137 acoplado a la proteína G), secuencias complementarias, fragmentos, alelos, variantes y productos génicos de los mismos.

En una realización preferida, un procedimiento de identificación de compuestos terapéuticos candidatos comprende cultivar células que expresan al menos una biomolécula seleccionada de: un biomarcador (TBB-III) que tiene secuencias de ácido nucleico/biomoléculas que comprenden: 1553212_at (queratina 78), 1553212_at (queratina 78), 1557236_at (apolipoproteína L), 1558142_at (repetición de trinucleótidos que contiene 6B), 1558484_s_at (repetición rica en leucina que contiene 27), 1565614_at (dedo de zinc de proteína 337), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567100_at (homólogo de Dachshund 1 (*Drosophila*)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 205758_at (molécula CD8a /// molécula CD8a), 206333_at (homólogo 1 de musashi (*Drosophila*)), 212920_at, 213242_x_at (KIAA0284), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 219425_at (familia 4A de sulfotransferasa, miembro 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 229191_at (cofactor D de plegamiento de tubulina), 229569_at (clon de ADNc IMAGE: 5263455), 231_629_x_at (peptidasa 3 relacionada con calicreína), 233765_at (LOC197135 hipotético), 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 235568_at (marco de lectura abierto 59 del cromosoma 19), 235803_at (factor 3 del tipo receptor de citoquinas), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (*Drosophila*)), 236953_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239463_at (locus transcrito), 240544_at (dedo de zinc, dominio 3 tipo AN-1), 243766_s_at (miembro 2 de la familia del dominio TEA), y productos génicos de los mismos.

Tales compuestos son útiles, por ejemplo, como compuestos terapéuticos candidatos para el tratamiento de una enfermedad del corazón, trastornos cardíacos y de las condiciones de los mismos. Por lo tanto, se incluyen en este documento procedimientos para el cribado de compuestos terapéuticos candidatos para el tratamiento, por ejemplo, de miocarditis, enfermedad coronaria, angina, síndrome coronario agudo, aneurisma de la aorta y disección, arritmias, cardiomiopatía, enfermedad cardíaca congénita, insuficiencia cardíaca congestiva o insuficiencia cardíaca crónica, pericarditis, y similares. Los procedimientos incluyen la administración del compuesto a un modelo de la condición, por ejemplo, poner en contacto un modelo de célula (*in vitro*) con el compuesto, o la administración del compuesto a un modelo animal de la condición, por ejemplo, un modelo animal de una condición asociada con la enfermedad cardíaca. Se evalúa luego el modelo para un efecto del compuesto candidato sobre el resultado clínico en el modelo y se puede considerar un compuesto terapéutico candidato para el tratamiento de la condición. Tales efectos pueden incluir efectos clínicamente relevantes como la disminución del dolor; el aumento de la expectativa de vida; etcétera. Tales efectos pueden ser determinados en una escala macroscópica o microscópica. Los compuestos terapéuticos candidatos identificados por estos procedimientos se pueden verificar adicionalmente, por ejemplo, mediante la administración a sujetos humanos en un ensayo clínico.

Las biomoléculas se pueden expresar a partir de uno o más vectores. Un "vector" (a veces denominado como un "vehículo" para el suministro de un gen o la transferencia de un gen) se refiere a una macromolécula o complejo de moléculas que comprenden un polinucleótido que va a ser suministrado a una célula huésped, ya sea *in vitro* o *in vivo*. El polinucleótido que va a ser suministrado puede comprender una secuencia de codificación de interés en terapia génica. Los vectores incluyen, por ejemplo, vectores virales (tales como adenovirus ("Ad"), virus adeno-asociados (AAV), y retrovirus), liposomas y otros complejos que contienen lípidos, y otros complejos macromoleculares capaces de mediar el suministro de un polinucleótido a una célula huésped. Los vectores también pueden comprender otros componentes o funcionalidades que modulan el suministro de genes y/o la expresión génica, o bien proporcionan propiedades beneficiosas a las células objetivo. Como se describe y se ilustra en más detalle más adelante, tales otros componentes incluyen, por ejemplo, componentes que influyen en el enlazamiento o el direccionamiento a las células (incluyendo los componentes que median el tipo de célula o el enlazamiento específico del tejido); los componentes que influyen en la absorción del ácido nucleico del vector por la célula; los componentes que influyen en la localización del polinucleótido dentro de la célula después de la absorción (tales como agentes mediadores de localización nuclear); y los componentes que influyen en la expresión del polinucleótido. Tales componentes también pueden incluir marcadores, tales como marcadores detectables y/o seleccionables que se pueden usar para detectar o seleccionar las células que han absorbido y expresan el ácido nucleico suministrado por el vector. Tales componentes pueden ser proporcionados como una característica natural del vector (tal como el uso de ciertos vectores virales que tienen componentes o funcionalidades que median el enlazamiento y la absorción), o se pueden modificar vectores para proporcionar tales funcionalidades. Otros vectores incluyen los descritos por Chen y colaboradores; *BioTechniques*, 34: 167-171 (2003). Una gran variedad de tales vectores es conocida en la técnica y está generalmente disponible.

En otra realización preferida, un vector expresa una o más biomoléculas que comprenden o forman TBB-I.

En otra realización preferida, un vector expresa una o más biomoléculas que comprenden o forman TBB-II.

En otra realización preferida, un vector expresa una o más biomoléculas que comprenden o forman TBB-III.

5

Kits

En otra realización preferida, se proporciona un kit que comprende uno cualquiera o más de los biomarcadores o firmas moleculares que comprenden las Tablas 1, 2 y 4.

10

Para su uso en las aplicaciones descritas o sugeridas anteriormente, también se proporcionan kits para uso de acuerdo con la invención. Tales kits pueden comprender un medio portador que está compartimentalizado para recibir en estrecho confinamiento uno o más medios contenedores tales como viales, tubos y similares, cada uno de los medios contenedores comprende uno de los elementos separados para ser usados en el procedimiento. Por ejemplo, uno de los medios que sirven como contenedor puede comprender una sonda que es o puede ser marcada de forma detectable. Cuando el kit usa hibridación de ácidos nucleicos para detectar el ácido nucleico objetivo, el kit también puede tener contenedores que contienen nucleótido(s) para la amplificación de la secuencia de ácido nucleico objetivo y/o un contenedor que comprende un medio informador, tal como una proteína que se enlaza a biotina, tal como avidina o estreptavidina, enlazada a una molécula informadora, tal como un marcador enzimático, fluorescente, o un radioisótopo.

15

20

El kit para uso de acuerdo con la invención comprenderá normalmente el contenedor descrito anteriormente y uno o más de otros contenedores que comprenden materiales deseables desde un punto de vista comercial y del usuario, incluyendo reguladores, diluyentes, filtros, agujas, jeringas y prospectos con instrucciones de uso. Una etiqueta puede estar presente en el contenedor para indicar que la composición se usa para una terapia específica o aplicación no terapéutica, y también puede indicar instrucciones ya sea para uso *in vivo* o *in vitro*, tales como las descritas anteriormente.

25

Los kits para uso de acuerdo con la invención tienen una serie de formas de realización. Una realización típica es un kit que comprende un contenedor, una etiqueta en dicho contenedor y una composición contenida dentro de dicho contenedor; en el que la composición incluye un anticuerpo primario que se enlaza a las biomoléculas de cada firma molecular e instrucciones para usar el anticuerpo para la evaluación de 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 de tejido y aplicar anticuerpo y sonda a la misma sección de una muestra de tejido. El kit puede incluir tanto un anticuerpo primario como secundario, en el que el anticuerpo secundario se conjuga con un marcador, por ejemplo, un marcador enzimático.

30

35

Otra realización es un kit que comprende un contenedor, una etiqueta en dicho contenedor y una composición contenida dentro de dicho contenedor; en el que la composición incluye un polinucleótido que se hibrida con un complemento de los polinucleótidos bajo condiciones rigurosas, la etiqueta en dicho contenedor indica que la composición se puede usar 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 una biomolécula de ARN o ADN en al menos un tipo de célula de mamífero.

40

Otros componentes opcionales en el kit incluyen, micromatrices, uno o más reguladores (por ejemplo, regulador de bloqueo, regulador de lavado, regulador de sustrato, etc.), otros reactivos, tales como sustrato (por ejemplo, un cromógeno) que está químicamente alterado por un marcador enzimático, solución de recuperación del epítipo, muestras de control (controles positivos y/o negativos), portaobjeto(s) de control, etc.

45

Las realizaciones de composiciones y procedimientos de la invención se ilustran en los siguientes ejemplos. Estos ejemplos se proporcionan con fines ilustrativos y no se consideran limitaciones sobre el alcance de las composiciones y procedimientos de la invención.

50

Ejemplos

55

Materiales y procedimientos:

Pacientes: los EMB fueron recolectados de pacientes que fueron remitidos al hospital Johns Hopkins entre 1997 y 2006 para evaluación de cardiomiopatía (n = 350). Se mantuvo simultáneamente una base de datos clínica de los resultados del paciente durante un período de 10 años a partir de 1997. Todos los pacientes dieron su consentimiento informado por escrito para la recolección de muestras y la extracción del cuadro médico. Los EMB transvenosos del tabique ventricular derecho se obtuvieron como se ha descrito anteriormente (Felker GM, Thompson RE, Hare JM, Hruban RH, Clemetson DE, Howard DL, Baughman KL, Kasper EK. N Engl J Med, 2000, 13 de abril; 342 (15): 1077-1084) para el análisis posterior de micromatrices. Con el fin de evitar posibles factores de confusión específicos de la enfermedad, sólo se seleccionaron muestras de pacientes con IDCM. IDCM era un diagnóstico de exclusión después de una extensa elaboración histológica sin signos patológicos detectables. Dentro

60

65

de un depósito de 180 muestras de IDCM, se seleccionaron las biopsias en la forma de un control de caso con base en los extremos fenotípicos en la supervivencia de la cohorte. Un grupo de buen pronóstico (GP, n = 25), se definió como la supervivencia libre de eventos durante al menos 5 años después de la presentación inicial con síntomas de insuficiencia cardíaca; un grupo con mal pronóstico (PP, n = 18), experimentó un acontecimiento dentro de los primeros 2 años. Los eventos incluyeron la muerte (n = 14), la necesidad de dispositivo de asistencia ventricular (n = 2) o trasplante cardíaco (n = 2).

Preparación de biopsias: Los EMB fueron congelados inmediatamente en nitrógeno líquido para su almacenamiento en un biorrepositorio. Todas las etapas de aislamiento y procesamiento del ARN se realizaron de acuerdo con las directrices MIAME (información mínima acerca de un experimento de micromatrices). Las muestras de tejido (~2mm de diámetro promedio) se homogeneizaron con el molino mezclador MM 301 (Retsch, Cat. No. 85120). Se usó el reactivo Trizol junto con el sistema de purificación de ARN total Micro-to-Midi (Invitrogen, Cat. No. 12183-018) para la extracción del ARN total (tasa de éxito: 97 % de las muestras). Se midió la concentración y la integridad del ARN total con el bioanálizador Agilent 2100. Todas las muestras de ARN exhiben ARN ribosomal 28S y 18S intacto en electroforesis desnaturante en gel de agarosa y las lecturas de absorbancia a 260/280 nm cayeron en el intervalo aceptable de 1,8 a 2,1. Se aisló y procesó una cantidad promedio de 568 ± 92 ng (SEM) de ARN total con el sistema de amplificación y marcación de ARN con biotina Ovation (NuGen, Cat. No. 2300-12).

Hibridación de micromatrices: Se hibridaron las muestras con el Genoma Humano U133 Plus 2.0 Array de Affymetrix y sin etapa adicional de amplificación. Los experimentos de micromatrices fueron juzgados exitosos cuando el aislamiento de ARN y la hibridación de micromatrices cumplieron con todos los índices de control de calidad que se especifican en la guía de Affymetrix para la evaluación de la muestra y calidad de la matriz (Kittleson MM, Irizarry RA, Heidecker B, Hare JM Transcriptomics: Translation of Global Expression Analysis to Genomic Medicine. En: Willard H.F., Ginsburg G.S., eds. Handbook of Genomic Medicine. 1era ed. Elsevier; 2008). El ruido y la señal promedio de fondo de todos los chips registrados dentro de los intervalos aceptables y las eficiencias de hibridación fueron similares para todas las muestras.

Análisis estadístico: los valores de intensidad en bruto de la hibridación de micromatrices se normalizaron con el Robust Multiarray Average (RMA) implementado en el paquete R para cálculo estadístico (disponible en www.R-project.org). En la siguiente etapa, análisis de significancia de micromatrices (SAM) (Tusher VG, y colaboradores Proc Natl Acad Sci EE.UU, 2001 24 de abril; 98 (9): 5116-21) para identificar las diferencias específicas de fenotipo en la expresión génica. SAM define la significancia con el valor q, un valor p ajustado para comparaciones múltiples. Para el desarrollo de un análisis TBB, se usó un análisis de predicción de micromatrices (PAM) (Tibshirani R, y colaboradores Proc Natl Acad Sci EE.UU. 2002 14 de Mayo; 99 (10): 6567-72) para crear un clasificador en un conjunto de entrenamiento (que contiene 2/3 de los datos, n = 29), con la posterior validación en un conjunto de prueba (que contiene un tercio de los datos, n = 14) (Kittleson MM, y colaboradores Circulation 2004 30 de noviembre; 110 (22): 3444-51). PAM es un software que permite encontrar el número mínimo de genes necesarios para crear un fenotipo específico del "centroide encogido más cercano" para clasificación (Tibshirani R, y colaboradores, 2002, ver más arriba). Este se hizo mediante una validación cruzada equilibrada 10 veces en un conjunto de entrenamiento, que permite escoger un umbral que minimiza los errores de la clasificación. Se evaluó la precisión total del biomarcador transcriptómica descubierto después de 50 divisiones aleatorias. Para probar las condiciones de base equilibradas de nuestra cohorte con pronóstico versus pobre, se usó una prueba t de Student para el análisis numérico y exacto de Fisher para variables categóricas.

45 Ejemplo 1: Biomarcador transcriptómico

Para identificar este biomarcador (firma genética) se recogieron muestras de corazón de pacientes sometidos a biopsia de corazón en una etapa temprana de la insuficiencia cardíaca, ya sea procedente de la miocarditis o de la cardiomiopatía idiopática, y se almacenaron en biorrepositorio durante 10 años. Las muestras de biopsias de endomiocardio de pacientes con miocarditis o cardiomiopatía idiopática han sido recolectadas en el Hospital Johns Hopkins entre 1997 - 2004 y almacenadas en nitrógeno líquido. El biorrepositorio fue revisado y se seleccionaron las muestras de biopsias de 19 pacientes con miocarditis y de 42 pacientes con cardiomiopatía idiopática. La miocarditis se diagnostica a partir de muestras de biopsia de endomiocardio mediante técnicas de inmunohistoquímica. Las directrices de la MIAME (Información mínima acerca de un experimento de micromatrices) fueron seguidas para todos los experimentos. Se homogeneizó el tejido con el instrumento MM301 de Retsch y se aisló el ARN total con el sistema de purificación de ARN total Microto-Midi de Invitrogen. El análisis de micromatrices del ARN total se realizó con el Genoma Humano U 133 Plus 2,0 Array de Affymetrix. En todas las muestras, tanto el aislamiento de ARN como la hibridación de la micromatriz cumplieron todos los índices de control de calidad como se especifica en la directriz de Affymetrix para evaluación de la muestra y calidad de la matriz. Los valores de expresión en bruto de todos los chips de micromatrices fueron procesados con Robust Multiarray Average (RMA) en R. Se realizó el análisis de significancia de micromatrices (SAM) en un conjunto de entrenamiento de 28 muestras (11 muestras de la miocarditis, 17 muestras de cardiomiopatía idiopática) para seleccionar un subconjunto de 122 genes expresados en forma diferencial (FC > 1,2, q = 0 %). Estos 122 genes contenían 38 genes que fueron sobreexpresados y 84 genes fueron subregulados en pacientes con miocarditis (Figura 1).

Con este subconjunto predefinido de 122 genes, se creó un mapa térmico que incluye todas las muestras del estudio (n = 61) estandarizadas por los niveles medios de expresión (Figura 2, Tabla 1, Tabla 2). Un mapa térmico es un procedimiento de clasificación de aprendizaje de máquina no supervisado. Cada fila representa uno de los 122 genes, y cada columna es una muestra del paciente. Una celda roja representa un gen subexpresado en un paciente dado con respecto a la expresión génica promedio en todos los pacientes, mientras que una celda azul representa un gen sobreexpresado. El dendrograma en la parte superior es un algoritmo de agrupamiento jerárquico no supervisado que divide las muestras en grupos con base en la similitud de los perfiles de expresión génica. Este algoritmo se desempeñó con 77 % de sensibilidad y 97 % de especificidad. Para aumentar el poder predictivo, se usó el mismo subconjunto de 122 genes para un enfoque de clasificación del "centroide encogido más cercano" con el análisis de predicción de micromatrices (PAM) implementado en R. Después del entrenamiento del clasificador (agrupamiento supervisado) en un subconjunto de 11 muestras de pacientes con miocarditis y 17 muestras de pacientes con cardiomiopatía idiopática, se usó un grupo de 39 genes que identificó miocarditis con 90 % de sensibilidad y 88 % de especificidad (Figura 3, Figura 4, Figura 5, Tabla 3, Tabla 4). Para analizar si se podría aumentar el poder predictivo del biomarcador transcriptómico descubierto, se creó un mapa térmico con el nuevo subconjunto de 39 genes candidatos (Figura 6). Aunque el primer mapa térmico de 122 genes produjo 6 muestras mal clasificadas, únicamente 4 muestras fueron clasificadas en forma incorrectamente por este algoritmo (sensibilidad: 89 %, especificidad: 95 %). Con este procedimiento de diagnóstico extremadamente poderoso de perfilado de expresión génica, incluso los pacientes que no presentan signos clínicos característicos de miocarditis, serán detectados en una etapa muy temprana de la enfermedad y tratados adecuadamente. La aplicación de este biomarcador altamente sensible reducirá significativamente el número de casos que conducen a un desenlace fatal de miocarditis en el futuro. Además, este biomarcador permitirá el uso apropiado de las terapias específicamente para miocarditis.

La Tabla 1 muestra los 38 genes que fueron significativamente sobreexpresados en pacientes con miocarditis versus cardiomiopatía idiopática. La primera columna contiene la identificación (ID) de Affymetrix para cada transcrito.

Tabla 1: 38 genes sobreexpresados en forma significativa en pacientes con miocarditis (q = 0,1 %; FC > 1,2)

ID del conjunto de sondas	Título del gen	Descripción del proceso biológico GO
1553145_at	proteína hipotética FLJ39653	---
1553575_at	---	---
1557236_at	apolipoproteína L, 6	regulación de la transcripción, dependiente de ADN /// transporte de lípidos
1558142_at	que contiene repetición de trinucleótidos	---
1560752_at	proteína 2 de los dominios caja F y WD-40	ciclo de ubiquitina /// modificación de proteína /// proteólisis
1565614_at	proteína en dedo de zinc 337	transcripción /// regulación de la transcripción, dependiente de ADN
1567100_at	homólogo de Dachshund 1 (Drosophila)	transcripción /// regulación de la transcripción, dependiente de ADN
200068_s_at	calnexina /// calnexina	angiogénesis /// plegamiento de proteína /// secreción de proteína
201031_s_at	ribonucleoproteína nuclear heterogénea H1 (H)	procesamiento de ARNm /// empalme de ARN /// procesamiento de ARN
202646_s_at	dominio de choque frío que contiene E1, enlazado al ARN	regulación de la transcripción, dependiente de ADN /// desarrollo de gónadas masculinas
205758_at	molécula CD8a /// molécula CD8a	respuesta inmune /// procesamiento y presentación de antígeno /// activación de células T /// diferenciación de células T citotóxicas /// respuesta inmune
206183_at	proteína en dedo de zinc 623	transcripción /// regulación de la transcripción, dependiente de ADN
212637_s_at	dominio WW que contiene la proteína ubiquitina E3 ligasa 1	transducción de señal /// diferenciación de células T /// entrada de virus en célula huésped
212920_at	---	---
213317_at	canal intracelular de cloro 5	transporte de iones /// transporte de cloro /// transporte de cloro
213619_at	ribonucleoproteína nuclear heterogénea H1 (H)	/// embarazo /// transporte /// transporte empalme de ARN
215443_at	receptor de la hormona estimuladora de la tiroides	transducción de señal /// señalización de la proteína G
216198_at	proteína de interacción con el factor de transcripción activador 7	metilación de ADN
217870_s_at	citidilato quinasa	biosíntesis del ribonucleótido pirimidina
218087_s_at	sorbina y dominio SH3 que contiene 1	transporte /// ruta de señalización del receptor de

222145_at 223577_x_at 224321_at	ADNc: FLJ23572 fis, clon LNG 12403 proteína PRO1073 proteína transmembrana con un dominio del tipo EGF y dos dominios del tipo de la folistatina 2	insulina --- ---
224373_s_at	motivo IQ y repeticiones WD 1	apoptosis /// transducción de señal /// desarrollo /// síntesis de ATP acoplada a transporte de electrones
224644_at 226173_at 226773_at	clon de ADNc IMAGE:5278517 ornitina del tipo aminotransferasa I ADNc FLJ35131 fis, clon PLACE6008824	--- --- ---
226880_at	caseína quinasa nuclear y sustrato 1 de quinasa que depende de ciclina	---
228980_at	dominio dedo anular y dominio del tipo FYVE que contiene 1	ciclo de ubiquitina /// apoptosis
229569_at 231735_s_at 233765_at 235803_at 236131_at 236953_s_at 240544_at 240971_x_at	clon de ADNc IMAGE:5263455 proteína PRO1073 LOC197135 hipotético factor 3 del tipo receptor de citoquinas clon de ADNc IMAGE:6622963 similar al ADNc RIKEN 8030451K01 dedo de zinc, dominio de tipo AN1 3 Cullin 4A	--- --- --- --- --- --- --- ---
244042_x_at	similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2	transición G1/S del ciclo celular mitótico /// reparación de ADN /// ciclo de ubiquitina /// ciclo celular /// detención del ciclo celular ---

La Tabla 2 muestra 84 genes que fueron significativamente subregulados en pacientes con miocarditis versus pacientes con cardiomiopatía idiopática (FC> 1.2, q = 0 %).

5 Tabla 2: 84 genes significativamente subregulados en miocarditis (FC> 1.2, q = 0,1 %)

ID del conjunto de sondas	Título del gen	Descripción del proceso Biológico GO
1552419_s_at	Familia tipo tirosina quinasa de tubulina, miembro 10	modificación de proteína
1553212_at	queratina 78	---
1555124_at	proteína hipotética MGC40574	---
1556192_x_at	supresor de metástasis 1	movilidad celular
1556320_at	estomatina (EPB72) de tipo 1	procesos biológicos
1556510_at	clon de ADNc IMAGE:4796864	----
1558484_s_at	repetición rica en leucina que contiene 27	---
1565662_at	Mucina 6, moco oligomérico /que forma gel	mantenimiento del epitelio gastrointestinal
1567410_at	proteína en dedo de zinc 135	regulación de la transcripción. dependiente de ADN
1568513_x_at	Proteasa, serina, I (tripsina 1)	catabolismo de colágeno, proteólisis, catabolismo de colágeno, respuesta inmune
1570408_at	Serina/treonina quinasa 24 (homólogo de STE20, levadura)	fosforilación del aminoácido de la proteína, transducción de señal
203307_al	similar a la proteína 1 que se enlaza al nucleótido de guanina	transducción de señal
204581_at	molécula CD22, glicoproteína asociada a mielina	respuesta inmune, adhesión celular, respuesta humoral antimicrobiana (sensu Vertebrata)
205586_x_at	factor de crecimiento nervioso inducible VGF	proceso biológico
206333_at	homólogo I de musashi (Drosophila)	desarrollo del sistema nervioso
207004_at	CLL de células B/linfoma 2	regulación del progreso a través del ciclo celular, respuesta inmune humoral
210059_s_at	Proteína quinasa 13 activado por mitógeno	regulación de la traducción, fosforilación del aminoácido de la proteína
210228_at	factor 2 estimulador de colonias (granulocito-macrófago)	respuesta de defensa celular
210384_at	proteína arginina metiltransferasa 2	metilación del aminoácido de la proteína
210923_at	familia 1 portadora de soluto (transportador de	transporte de ácido dicarboxílico

211024_s_at	glutamato), miembro 7 factor 1 de transcripción de la tiroides /// factor 1 de transcripción de la tiroides	migración neuronal
211062_s_at 211096_at	carboxipeptidasa Z /// carboxipeptidasa Z factor 2 de transcripción de leucemia de pre-células B	proteólisis, ruta de señalización del receptor Wnt regulación de la transcripción, dependiente de ADN
211181_x_at	factor 1 de transcripción relacionado con enanismo (leucemia 1 mielóide aguda; oncogén aml1)	desarrollo esquelético, hemopoyesis
211710_x_at 213096_at	proteína ribosomal L4 /// proteína ribosomal L4 familia 2 del dominio transmembrana y de bobina en espiral	biosíntesis de proteína, metabolismo ---
213121_at	polipéptido de 70 kDa de la ribonucleoproteína nuclear pequeña (antígeno RNP)	procesamiento de ARNm
213242_x_at 213569_at 213770_at	KIAA0284 impar omitida relacionada con 2 (Drosophila) supresor quinasa de ras I	---
214171_s_at 216116_at	factor 2 auxiliar de ARN nuclear pequeño U2 proteína que interactúa con NCK con el dominio SH3	fosforilación del aminoácido de la proteína procesamiento de ARNm importación del sustrato que soporta NLS dentro del núcleo
216427_at 216820_at 217054_at 217180_at	ADNc: FLJ22786 fis, clon KAIA2150 ADNc FLJ39484 fis, clon PROST2014925 proteína hipotética similar al producto génico KIAA0187	---
217182_at 217322_x_at 217430_x_at 219070_s_at 219425_at 221663_x_at	mucina 5AC. Moco oligomérico/que forma gel --- colágeno, tipo 1, alfa I dominio de esperma móvil que contiene 3 familia 4A de sulfotransferasa, miembro I receptor H3 de histamina	respuesta inmune adhesión celular ---
221684_s_at 223974_at 226640_at 228074_at 129191_at 229257_at 229335_at 229358_at	Nictalopina proteína hipotética MGC11082 diacilglicerol lipasa beta proteína hipotética LOC162073 cofactor D de plegamiento de tubulina proteína KIAA1856 superfamilia de inmunoglobulina, miembro 4C homólogo hedgehog indio (Drosophila)	desarrollo esquelético desarrollo del corazón metabolismo de lípidos ruta de señalización de la proteína del receptor acoplado a proteína G respuesta a estímulos ---
230341_x_at 230693_at 230768_at	metalopectidasa ADAM con motivo tipo I de trombospondina, 10 ATPasa, que transporta Ca ⁺⁺ , músculo cardiaco, contracción rápida 1 FERM, RhoGEF y dominio pleckstrina de la proteína 2	catabolismo de lípidos, metabolismo de lípidos ---
231510_at 231629_x_at 231998_at 233794_at	miembro de la familia GLI-Kruppel GLI2 peptidasa 3 relacionada con calicreína --- proteína 3 que se enlaza a ADN monocatenario familia con similitud de secuencia 129, miembro B	plegamiento de beta tubulina ---
233974_s_at 234495_at 234637_at 234881_at 235568_at	relacionada con calicreína (peptidasa 15) proteína 4-5 asociada con queratina --- marco de lectura abierto 59 del cromosoma 19 locus transcrito	desarrollo esquelético, modelación de vasos sanguíneos proteólisis, ruta de señalización mediada por integrina transporte de iones de calcio, regulación de la contracción de músculo estriado remodelación de neuronas transcripción del promotor de ARN polimerasa II proteólisis, regulación negativa de angiogénesis ---
235600_at 236496_at 237087_at	proteína 3 que se enlaza a ADN monocatenario familia con similitud de secuencia 129, miembro B relacionada con calicreína (peptidasa 15) proteína 4-5 asociada con queratina --- marco de lectura abierto 59 del cromosoma 19 locus transcrito homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila) marco de lectura abierto 105 del cromosoma 14 factor de crecimiento beta transformante	regulación de la transcripción, dependiente de ADN ---
		proteólisis, ---

		metabolismo de lípidos ---
		desarrollo esquelético, transformación de la ruta

237144_at	latente que se enlaza a la proteína 3	de señalización del receptor beta del factor de crecimiento
237398_at	locus transcrito	---
237547_at	proteína hipotética LOC728730	---
237679_at	66 que contiene al motivo tripartita	regulación negativa de la transcripción, dependiente de ADN
238267_s_at	---	---
238445_x_at	manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B	---
239026_x_at	centaurina, gamma 3	regulación de la transcripción, dependiente de ADN
239463_at	locus transcrito	---
239756_at	tipo 1 deficiente en la captura mitótica de MAD1 (levadura)	punto de control de la mitosis
240039_at	locus transcrito /// locus transcrito	---
240147_at	proteína hipotética MGC11257	---
240517_at	Cistationina beta sintasa	biosíntesis de cisteína a partir de serina
241270_at	homólogo 2 romboide 5 (Drosophila)	---
241431_at	---	---
242365_at	dominio de bobina en espiral que contiene 32	---
243297_at	proteína vacuolar para clasificar al homólogo D 13 (<i>S. cerevisiae</i>)	localización de la proteína, ciclo celular, mitosis
243497_at	locus transcrito	---
243766_s_at	miembro 2 de la familia del dominio TEA	transcripción, regulación de la transcripción, dependiente de ADN
43934_at	receptor 137 acoplado a la proteína G	proceso biológico

La Tabla 3 muestra los resultados de la predicción del grupo de prueba (n = 28): se usaron 39 firmas de genes. Las etiquetas de la muestra se enumeran por encima de la fila resaltada en rojo, seguido por las clases predichas. Las muestras de miocarditis (Myo) fueron asignadas a la clase 2 - muestras de cardiomiopatía idiopática (GP, BP). Las muestras fueron asignadas a la clase 1. 25 muestras fueron correctamente clasificadas (probabilidades entre 75 % y 99 %). Sólo tres muestras fueron mal clasificadas (con probabilidades entre 48 % y 99 %). Las probabilidades predichas se enumeran para cada clase en las dos últimas líneas.

5

Tabla 3

Predicción para el umbral = 2,6

Nombre ajustes: Ajustes4

Quantil compensado: 50 Valor compensado: 0.257958

Preselección de RNG: 420473

Distribución anterior (Muestra anterior):

Clase	1	2
Prob.	0.757576	0.242424242

Clases predichas y probabilidades posteriores predichas, reales

Etiquetas de clase	Myo-5	Myo-8	Myo-10	Myo-1	Myo-4a	Myo-6	Myo-12	Myo-14	Myo-2	Myo-7	Myo-9	GP-13	
Etiquetas de la muestra	2	2	2	2	2	2	1	2	2	2	2	1	
Etiquetas de clase predichas	1	0.044735755	0.00317	0.000372279	0.173628	0.00064	0.000536	0.51331	0.000157	0.00408	0.021888	0.002194	0.999853
	2	0.955264245	0.99683	0.999627721	0.826372	0.99936	0.999464	0.48669	0.999843	0.99592	0.978102	0.997806	0.000147

BP-14	GP-25	BP-12	BP-13	BP-18	BP-16	BP-19	BP-20	GP-8	BP-3	BP-8	BP-15	GP-5	GP-10	GP-12	GP-26
1	1	1	1	2	1	2	1	1	1	1	1	1	1	1	1
0.961375	0.96379	0.753742	0.06179	0.990899	0.017853	0.990345	0.998788	0.996078	0.999488	0.821228	0.850091	0.867818	0.998585	0.787331	0.814156
0.038625	0.03621	0.246258	0.93821	0.019101	0.982947	0.009655	0.001232	0.003922	0.000512	0.178772	0.149309	0.132384	0.001435	0.202888	0.185844

10

Tabla 4: biomarcador transcriptómico de 39 genes para la detección de pacientes con miocarditis: la primera columna contiene las ID de Affymetrix para cada gen. Las anotaciones para la función biológica se derivaron de la base de datos de ontología de genes.

15

Tabla 4: biomarcador de 39 genes para detectar miocarditis

ID del conjunto de sondas	Título del gen	Descripción del proceso biológico GO
1553212-at 1557236_at	Queratina 78 apolipoproteína L. 6	-- regulación de la transcripción. Dependiente DNA /// transporte de lípidos /// metabolismo de lipoproteína
1558142_at 1558484_s_at 1565614_at	que contiene repetición de trinucleótidos 6B repetición rica en leucina que contiene 27 dependiente de la proteína en dedo de zinc 337	--- --- regulación de la transcripción, dependiente de ADN
1565662_at 1567100_at	Mucina 6, moco oligomérico/que forma gel Homólogo Dachshund 1(Drosophila)	mantenimiento del epitelio gastrointestinal regulación de la transcripción, dependiente de ADN transducción de señal
203307_at 205758_at	similar a la proteína 1 que se enlaza al nucleótido de guanina molécula CD8a /// molécula CD8a	procesamiento y presentación de antígeno, respuesta inmune a la diferenciación de células T citotóxicas
206333_at 212920_at 213242_x_at 213619_at 213770_at 214171_s_at	homólogo de musashi (Drosophila) --- KIAA0284 Ribonucleoproteína nuclear heterogénea H1 (H) supresor quinase de ras 1 factor 2 auxiliar de ARN nuclear pequeño U2	desarrollo del sistema nervioso --- --- procesamiento de ARNm fosforilación del aminoácido de la proteína empalme de ARN nuclear, a través del espliceosoma
215443_at	receptor de la hormona estimuladora de la tiroides	transducción de señal, señalización de la proteína G, acoplado al segundo mensajero del nucleótido cíclico
216198_at	proteína de interacción con el factor de transcripción activador 7	metilación de ADN
216427_at 217054_at 217182_at 217322_x_at 219425_at 222145_at 229191_at 229569_at 231624_x_at	ADNc: FLJ22786 fis, clon KAIA2150 ADNc FLJ39484 fis, clon PROS12014925 mucina 5AC. Moco oligomérico/que forma gel familia 4A de sulfotransferasa, miembro 1 ADNc: FLJ23572 fis, clon LNG12403 cofactor D de plegamiento de tubulina ADNc clon IMAGE:5263455 peptidasa 3 relacionada con calicreína	--- --- adhesión celular --- metabolismo de lípidos --- plegamiento de la proteína --- proteólisis, regulación negativa de angiogénesis
233765_at 233794_at 233974_s_at 234495_at 235568_at 235803_at 236496_at	LOC197135 hipotético proteína 3 que se enlaza a ADN monocatenario familia con similitud de secuencia 129. miembro B peptidasa 15 relacionada con calicreína marco de lectura abierto 59 del cromosoma 19 factor 3 del tipo receptor de citoquinas homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)	--- regulación de la transcripción, ADN --- proteólisis, proteólisis --- metabolismo de lípidos
236953_s_at 238445_x_at	similar al ADNc RIKEN 8030451K01 manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetilglucosaminiltransferasa, isozima B	--- ---
239463_at 240544_at 243766_s_at	locus transcrito Dedo de zinc, dominio 3 de tipo AN1 miembro 2 de la familia del dominio TEA	--- --- regulación de la transcripción, dependiente de ADN
244042_x_at	Similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2	---

Otras realizaciones

5 Debe entenderse que aunque la invención se ha descrito junto con la descripción detallada de la misma, la descripción anterior pretende ilustrar y no limitar el alcance de la invención. Otros aspectos, ventajas y

modificaciones están dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

1. Uso de una composición molecular que comprende moléculas de ácido nucleico que consisten en las secuencias de ácido nucleico de: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (proteína en dedo de zinc 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína nuclear heterogénea H1 (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza a ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (dedo de zinc de la proteína 623), 212637_s_at (dominio WW que contiene la proteína ubiquitina E3 ligasa 1), 212920_at, 213317_at (canal intracelular de cloro 5), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y el dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con un dominio del tipo EGF y dos dominios del tipo de la folistatina 2), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina del tipo aminotransferasa 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína quinasa nuclear y sustrato 1 de quinasa que depende de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOCI 97135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio de tipo AN1 3), 240971_x_at (Cullin 4A) y 244042_x_at (similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2) para el diagnóstico *in vitro* de la miocarditis.
2. El uso de la reivindicación 1, en el que la detección de la sobreexpresión de las moléculas de ácido nucleico, las secuencias complementarias, los alelos o los productos génicos de los mismos en una muestra biológica es diagnóstico de miocarditis.
3. El uso de la reivindicación 1, en el que las moléculas de ácido nucleico, o los productos génicos de las mismas, se sobreexpresan a niveles por al menos del 10 % en una célula o un paciente en comparación con los niveles en una célula normal o un sujeto normal.
4. El uso de la reivindicación 1, en el que las moléculas de ácido nucleico, las secuencias complementarias, los alelos o los productos génicos de los mismos se sobreexpresan en aproximadamente un 50 % en una célula o un paciente, en comparación con los niveles en una célula normal o un sujeto normal.
5. El uso de la reivindicación 1, en el que las moléculas de ácido nucleico, o los productos génicos de las mismas, se sobreexpresan en aproximadamente un 75 % en una célula o un paciente, en comparación con los niveles en una célula normal o un sujeto normal.
6. El uso de la reivindicación 1, en el que un biochip comprende las moléculas de ácido nucleico hibridadas con el biochip.
7. Uso de un biomarcador transcriptómico que consiste en moléculas de ácido nucleico complementarias a: 1553212_at (queratina 78), 1557236_at (apolipoproteína L), 1558142_at (repetición de trinucleótidos que contiene 6B), 1558484_s_at (repetición rica en leucina que contiene 27), 1565614_at (dedo de zinc de proteína 337), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 205758_at (molécula CD8a /// molécula CD8a), 206333_at (homólogo 1 de musashi (Drosophila)), 212920_at, 213242_x_at (KIAA0284), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 219425_at (familia 4A de sulfotransferasa, miembro 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 229191_at (cofactor D de plegamiento de tubulina), 229569_at (clon de ADNc IMAGE: 5263455), 231629_x_at (peptidasa 3 relacionada con calicreína), 233765_at (LOCI 97135 hipotético), 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 235568_at (marco de lectura abierto 59 del cromosoma 19), 235803_at (factor 3 del tipo receptor de citoquinas), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 236953_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetilglucosaminiltransferasa, isozima B), 239463_at (locus transcrito), 240544_at (dedo de zinc, dominio 3 tipo AN-1) y 243766_s_at (miembro 2 de la familia del dominio TEA), para el diagnóstico *in vitro* de miocarditis.

8. El uso de la reivindicación 7, en el que la detección de la sobreexpresión de las moléculas de ácido nucleico, las secuencias complementarias, los alelos o los productos génicos de los mismos en una muestra biológica es diagnóstico de miocarditis.
- 5 9. El uso de la reivindicación 7, en el que las moléculas de ácido nucleico, las secuencias complementarias, los alelos o los productos génicos de los mismos, se sobreexpresan en niveles de al menos un 10 % en una célula o un paciente en comparación con los niveles en una célula normal o un sujeto normal.
- 10 10. El uso de la reivindicación 7, en el que las moléculas de ácido nucleico, las secuencias complementarias, los alelos o los productos génicos de los mismos se sobreexpresan en aproximadamente un 50 % en una célula o un paciente, en comparación con los niveles en una célula normal o un sujeto normal.
- 15 11. El uso de la reivindicación 7, en el que las moléculas de ácido nucleico, las secuencias complementarias, los alelos o los productos génicos de los mismos se sobreexpresan en aproximadamente un 75 % en una célula o un paciente, en comparación con los niveles en una célula normal o un sujeto normal.
- 20 12. Un procedimiento in vitro de diagnóstico de miocarditis, que comprende: la generación por parte de un paciente de una firma molecular, en donde la generación comprende la medición de la expresión de moléculas de ácido nucleico que comprenden secuencias de ácido nucleico: 1553145_at (proteína hipotética FLJ39653), 1553575_at, 1557236_at (apolipoproteína L, 6), 1558142_at (repetición de trinucleótidos que contiene 6B), 1560752_at (proteína 2 de los dominios caja F y WD-40), 1565614_at (proteína en dedo de zinc 337), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 200068_s_at (calnexina /// calnexina), 201031_s_at (ribonucleoproteína H1 nuclear heterogénea (H)), 202646_s_at (dominio de choque frío que contiene E1, que se enlaza con ARN), 205758_at (molécula CD8a /// molécula CD8a), 206188_at (dedo de cinc de proteína 623), 212637_s_at (dominio WW que contiene la proteína ubiquitina E3 ligasa 1), 212920_at, 213317_at (canal intracelular de cloro 5), 213619_at (ribonucleoproteína H1 nuclear heterogénea (H)), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 217870_s_at (citidilato quinasa), 218087_s_at (sorbina y el dominio SH3 que contiene 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 223577_x_at (proteína PRO1073), 224321_at (proteína transmembrana con un dominio del tipo EGF y dos dominios del tipo de la folistatina 2), 224373_s_at (motivo IQ y repeticiones WD 1), 224644_at (clon de ADNc IMAGE: 5278517), 226173_at (ornitina del tipo aminotransferasa 1), 226773_at (ADNc FLJ35131 fis, clon PLACE6008824), 226880_at (caseína quinasa nuclear y sustrato 1 de quinasa que depende de ciclina), 228980_at (dominio dedo anular y dominio del tipo FYVE que contiene 1), 229569_at (clon de ADNc IMAGE: 5263455), 231735_s_at (proteína PRO1073), 233765_at (LOCI 97135 hipotético), 235803_at (factor 3 del tipo receptor de citoquinas), 236131_at (clon de ADNc IMAGE: 6622963), 236953_s_at (similar al ADNc RIKEN 8030451K01), 240544_at (dedo de zinc, dominio de tipo AN1 3), 240971_x_at (Cullin 4A), 244042_x_at (similar al respondedor del receptor de ácido retinoico (inducido por tazaroteno) 2); el análisis de la firma molecular generada;
- 30 y el diagnóstico de si el paciente tiene o no miocarditis tras el análisis de la firma molecular generada.
- 40 13. El procedimiento de la reivindicación 12, en el que al paciente se le diagnostica que tiene miocarditis si: 1553212_at (queratina 78), 1557236_at (apolipoproteína L), 1558142_at (repetición de trinucleótidos que contiene 6B), 1558484_s_at (repetición rica en leucina que contiene 27), 1565614_at (dedo de zinc de proteína 337), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 205758_at (molécula CD8a /// molécula CD8a), 206333_at (homólogo 1 de musashi (Drosophila)), 212920_at, 213242_x_at (KIAA0284), 213619_at (ribonucleoproteína nuclear heterogénea H1 (H)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 219425_at (familia 4A de sulfotransferasa, miembro 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 229191_at (cofactor D de plegamiento de tubulina), 229569_at (clon de ADNc IMAGE: 5263455), 231629_x_at (peptidasa 3 relacionada con calicreína), 233765_at (LOCI97135 hipotético), 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 235568_at (marco de lectura abierto 59 del cromosoma 19), 235803_at (factor 3 del tipo receptor de citoquinas), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 236953_s_at, 238445_x_at (manosil (alfa-1,6)- glicoproteína beta-1,6-N-acetilglucosaminiltransferasa, isozima B), 239463_at (locus transcrito), 240544_at (dedo de zinc, dominio 3 tipo AN-1), 243766_s_at (miembro 2 de la familia del dominio TEA) y 244042_x_at están sobrerregulados.
- 55 60 14. El procedimiento de las reivindicaciones 12 o 13, en el que la generación comprende además aislar moléculas de ácido nucleico obtenidas a partir de una muestra biológica, y las moléculas de ácido nucleico aisladas se hibridan preferiblemente con un biochip que comprende moléculas complementarias de ácido nucleico y los valores de intensidad en bruto de la hibridación se normalizan preferiblemente y se identifican preferiblemente las diferencias específicas del fenotipo en la expresión génica, y
- 65

en el que las diferencias en la expresión génica se identifican preferiblemente mediante el análisis de la significancia de las micromatrices, en donde significación se define con un valor q y comparaciones múltiples comprenden un valor p ajustado, y

5 en el que la especificidad del fenotipo se identifica preferiblemente mediante la creación de un clasificador en un conjunto de entrenamiento que comprende aproximadamente un 66 % de los datos obtenidos, con la validación posterior en un conjunto de prueba que comprende aproximadamente un 33 % de los datos obtenidos y que definen un centroide encogido más cercano específico del fenotipo para la clasificación, y

en el que el centroide encogido más cercano específico del fenotipo preferiblemente comprende equilibrar aproximadamente una validación cruzada 10 veces en un conjunto de entrenamiento.

10

15. El uso de la reivindicación 1, en el que un kit comprende las moléculas de ácido nucleico.

16. El uso de la reivindicación 15, en el que un biochip comprende las moléculas de ácido nucleico.

15

17. Uso de un biochip que comprende moléculas de ácido nucleico hibridadas con el biochip, en donde las moléculas de ácido nucleico consisten en: 1553212_at (queratina 78), 1557236_at (apolipoproteína L), 1558142_at (repetición de trinucleótidos que contiene 6B), 1558484_s_at (repetición rica en leucina que contiene 27), 1565614_at (dedo de zinc de proteína 337), 1565662_at (mucina 6, moco oligomérico/que forma gel), 1567100_at (homólogo de Dachshund 1 (Drosophila)), 203307_at (similar a la proteína 1 que se enlaza al nucleótido de guanina), 205758_at (molécula CD8a /// molécula CD8a), 206333_at (homólogo 1 de musashi (Drosophila)), 212920_at, 213242_x_at (KIAA0284), 213619_at (ribonucleoproteína nuclear heterogénea HI (H)), 213770_at (supresor quinasa de ras 1), 214171_s_at (factor 2 auxiliar de ARN nuclear pequeño U2), 215443_at (receptor de la hormona estimuladora de la tiroides), 216198_at (proteína de interacción con el factor de transcripción activador 7), 216427_at (ADNc: FLJ22786 fis, clon KAIA2150), 217054_at (ADNc FLJ39484 fis, clon PROST2014925), 217182_at (mucina 5AC, moco oligomérico/que forma gel), 217322_x_at, 219425_at (familia 4A de sulfotransferasa, miembro 1), 222145_at (ADNc: FLJ23572 fis, clon LNG12403), 229191_at (cofactor D de plegamiento de tubulina), 229569_at (clon de ADNc IMAGE: 5263455), 231629_x_at (peptidasa 3 relacionada con calicreína), 233765_at (LOC197135 hipotético), 233794_at (proteína 3 que se enlaza a ADN monocatenario), 233974_s_at (familia con similitud de secuencia 129, miembro B), 234495_at (peptidasa 15 relacionada con calicreína), 235568_at (marco de lectura abierto 59 del cromosoma 19), 235803_at (factor 3 del tipo receptor de citoquinas), 236496_at (homólogo 2 de espermatozoides degenerativos, lípido desaturasa (Drosophila)), 236953_s_at, 238445_x_at (manosil (alfa-1,6-)-glicoproteína beta-1,6-N-acetil-glucosaminiltransferasa, isozima B), 239463_at (locus transcrito), 240544_at (dedo de zinc, dominio 3 tipo AN-1), 243766_s_at (miembro 2 de la familia del dominio TEA), para el diagnóstico *in vitro* de miocarditis.

35

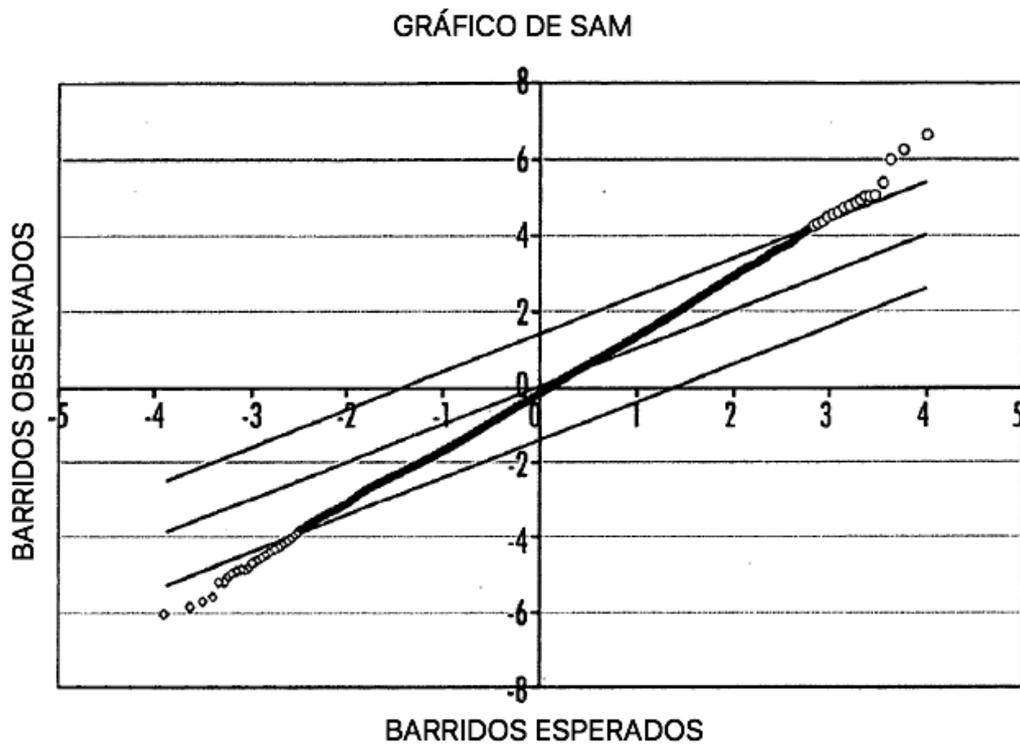


FIG. 1

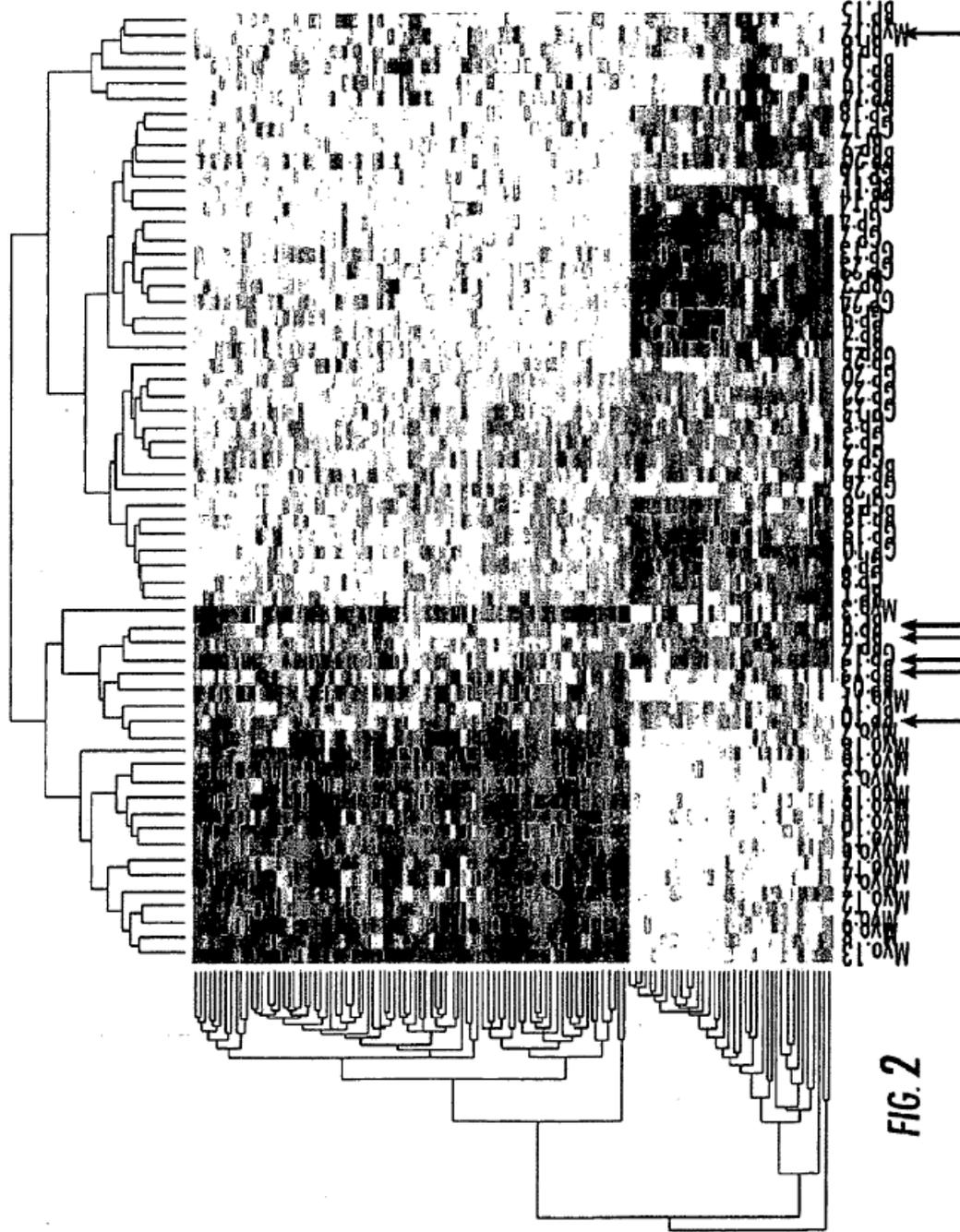


FIG. 2

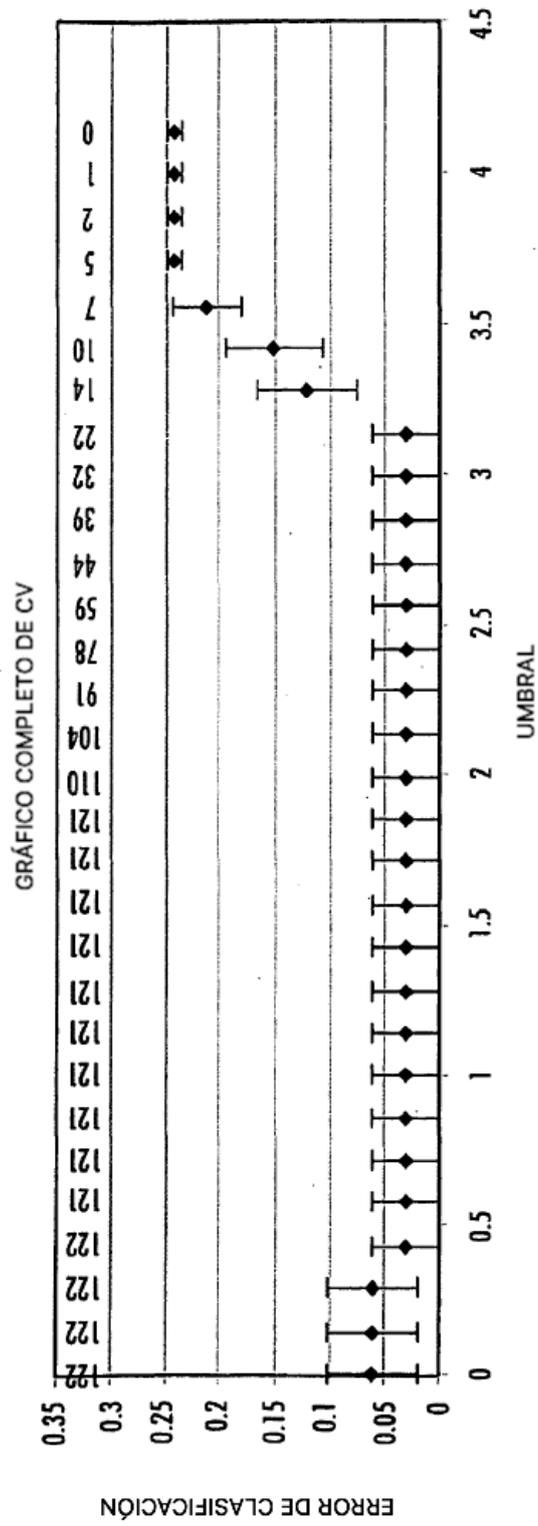


FIG. 3

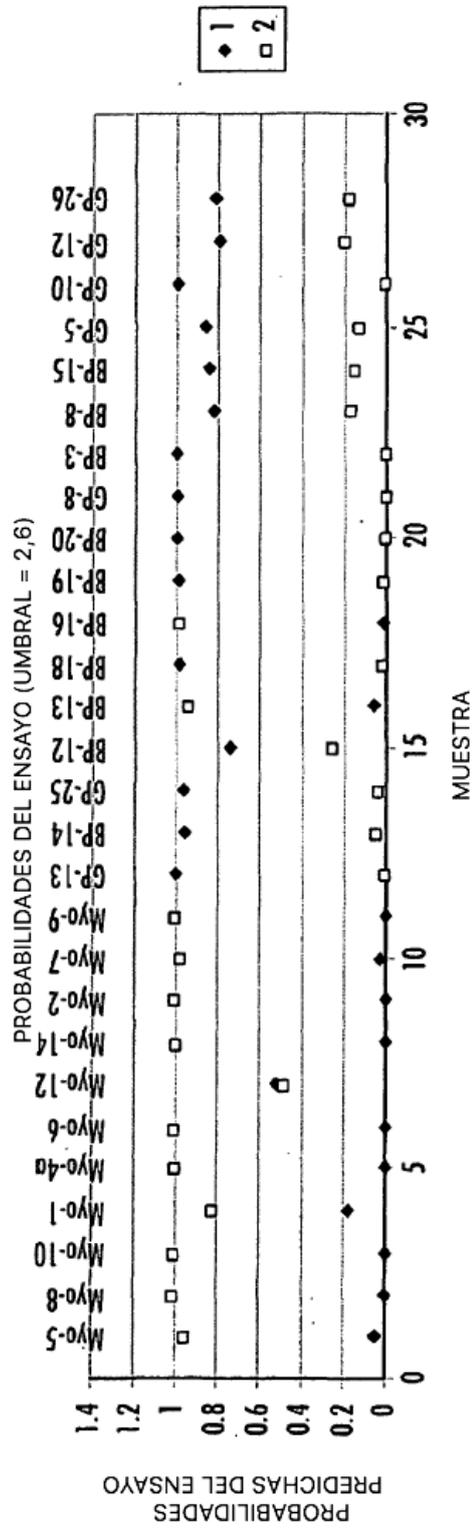


FIG. 4

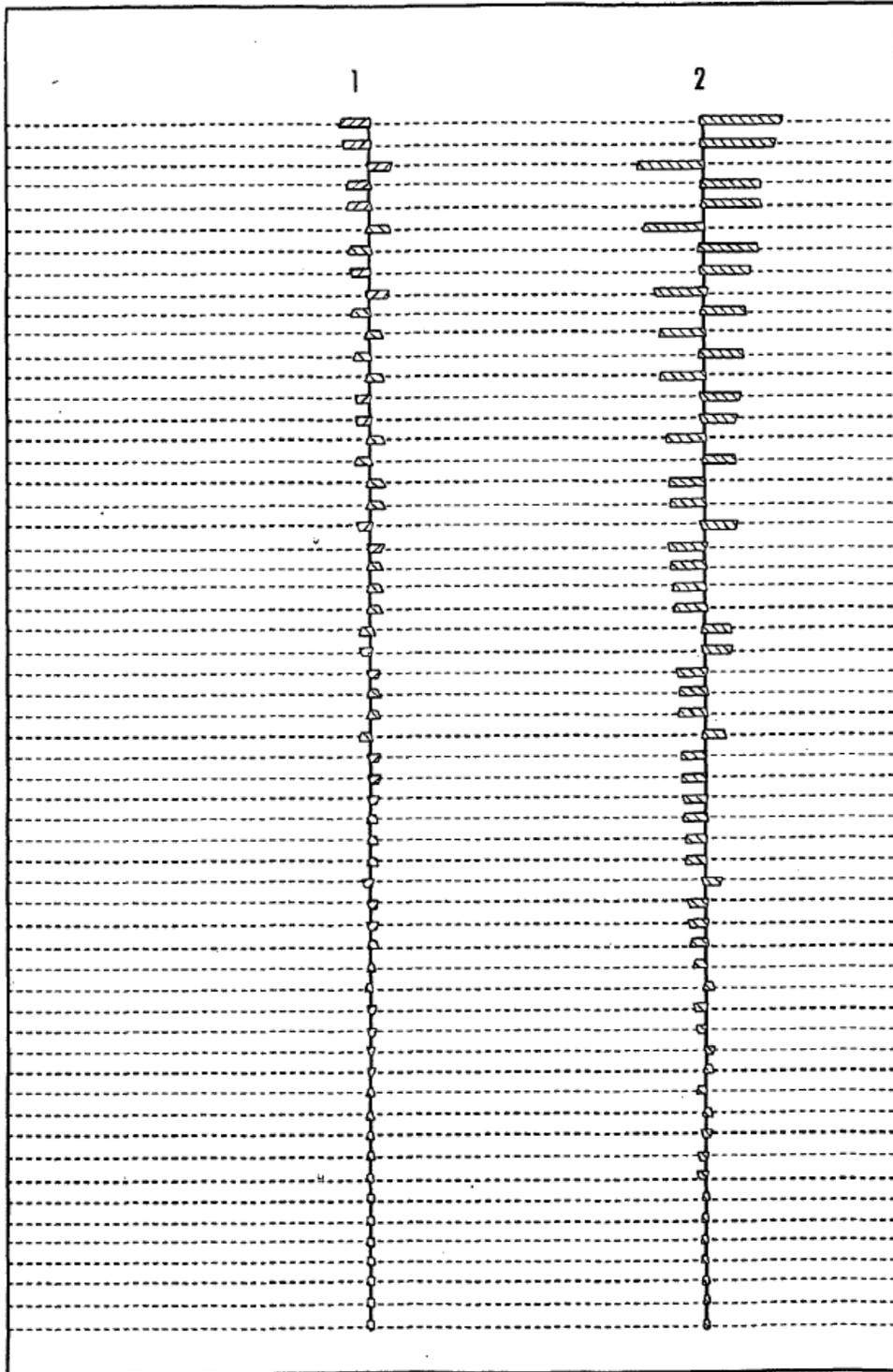


FIG. 5

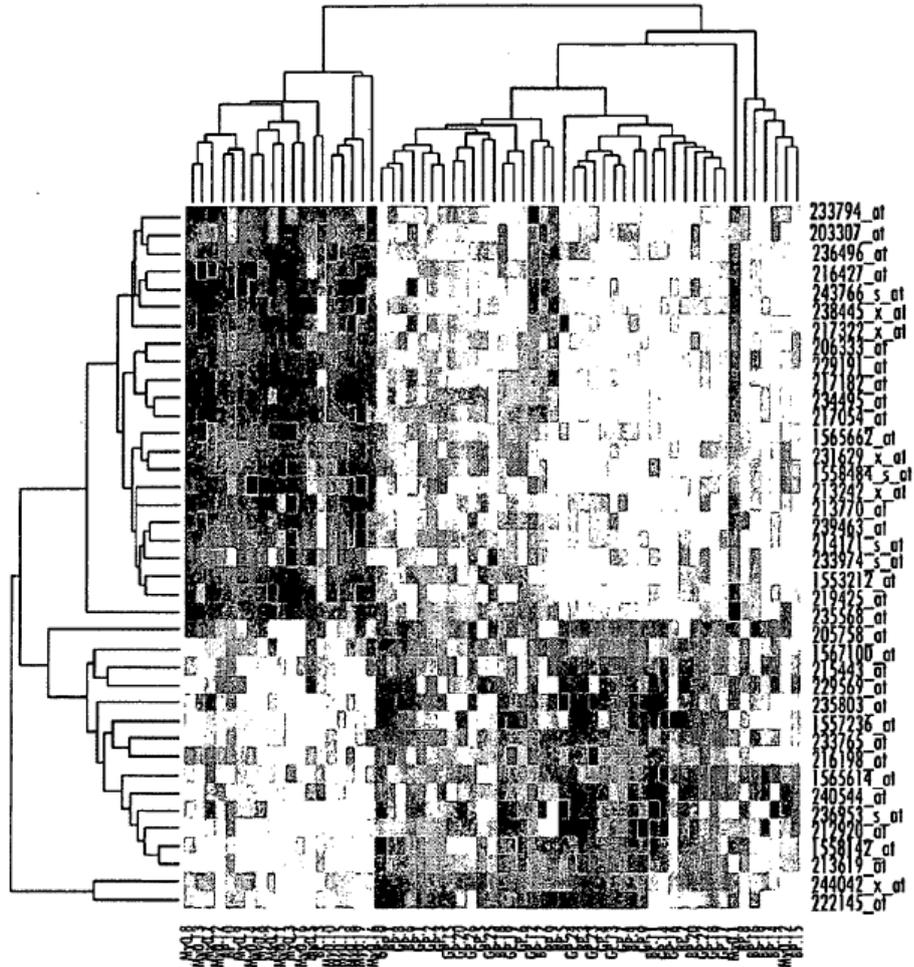


FIG. 6

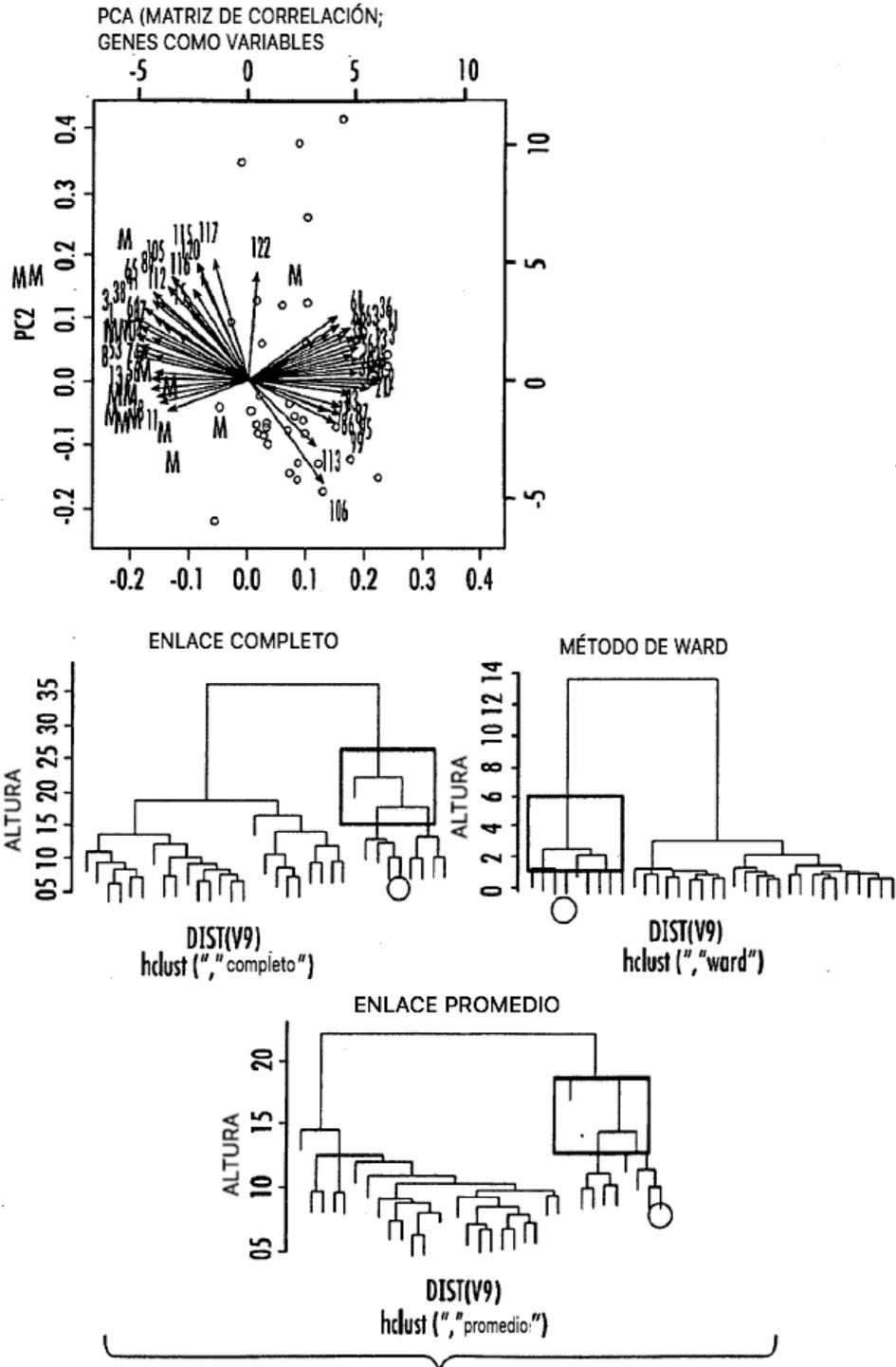


FIG. 7

Affymetrix ID	Secuencias de la sonda (5' -> 3')
1553145_at	<ul style="list-style-type: none"> • ATGTCAGTTTCCATTTACTGTCTCA • AGTTTCCATTTACTGTCTCACCACA • TACTGTCTCACCACATTTATGGGT • CACATTTATGGGTTATTGCCTTTG • GAAGGGTAGGAATCCCAAATCTGT • GTTGCCTAAGGCAGCCAGTTGTCAG • GGCAGCCAGTTGTGAGAATGGACTT • GTCAGAATGGACTTTCCTCATTTTC • ATAGAAGAGAGAGGGCCCCGGCACGG • CAGTGAACCGAACCGAGATCGCGTC • ACCGAACCGAGATCGCGTCGCTGGA
1553575_at	<ul style="list-style-type: none"> • AGTATTGATTGTTAGCGGTGTGGTC • GTGTGGTCGGGTGTGTTATTATTCT • GAATGATGGTTGTCTTTGGATATAC • TTGGATATACTACAGCGATGGCTAT • AGCGATGGCTATTGAGGAGTATCCT • TTGAGGTCTTGGTGAGTGTTTAGT • TGATTCCGGGAGGATCCTATTGGTGC • GCTTTGTATGATTATGGGCGTTGAT • ATGGGCGTTGATTAGTAGTAGTTAC • AGTTACTGGTTGAACATTGTTTGTT • ATATATTGTAATTGAGATTGCTCGG
1557236_at	<ul style="list-style-type: none"> • TTTCTTTTACTTCTCAGTTGACTG • GGAAGAATGTAGTTTCGTTTATGT • CGTTTTATGTTAATATCGCTCAAA • TTTTCTCTCTGTACCTTATCATGT • GTACCTTATCATGTAAATTTTGCTA • GCAAAAATTTAAGGCTATTTAGCTG • CAACTGCCTAGGGTTGTA AACAGG • ATGCCTAATACGTCTTTATATGTAT • GTTGTGTACACGATGTTT TAGTGCT • GTACACGATGTTT TAGTGCTAAAAA • AAAATATGTAAAAGAGCTCTACTTG
1558142_at	<ul style="list-style-type: none"> • ATTGCTCCTCATATTACTGGTTTTA • ATATTACTGGTTTTACATGGACACA • CATGGACACAGAACTAGGCACTTT • AGGCACTTTAGAGGTGCACTTGCAT • TAGAGGTGCACTTGCATGGCAGGCT • GGTACTTAAAATCACTGGTTCACTT • TTTTTTGCCATAGCAGGTACTGTAT • CATAGCAGGTACTGTATTCTCATG • GAATTAAGTCGTCGTCATTT CATT A • GATTGTATTATGTGCAACTCAGTTG • ACTCAGTTGCTTACATTATAACTAC

FIG. 8A

Affymetrix ID	Secuencias de la sonda (5' -> 3')
1560752_at	<ul style="list-style-type: none"> • CTGGAGCACTGTAGCCCTTATAAAC • AAACAGTTGAGGCTGATCGGCTCTC • GCTGATCGGCTCTCTTTTAGGACAA • GGCTCTCTTTTAGGACAAGGCCCA • AGATGGCCCAGACTCAAGAAAGGAT • GCAACTTGCAATATGTTGAATCCA • AGGGACTGCTCTGAGAAGGTTGAAA • GAAGGTTGAAAAGCTGCCTGAACA • AGCTTGCCTGAACATACTCATTTC • GCAGATTAACATCCTTACAGACAAT • AAACACACCAGGGAGACACCATAAA
1565614_at	<ul style="list-style-type: none"> • AAAACCTACATTACTATCTTAATGC • GATAACATTTACACCCATTTCAGGG • GGCTGCAATCATCATATTTAATGGT • TCACTCTTATTTAGCATTATCCTAG • GAATAACTAGCTAACCTGTTAAGGT • ACTAGATACTAGTAACACACAGTTT • GTAACACACAGTTTAGGAACCTCAG • AAAGCAC TTGTATTTCTGTAACCTA • GCACTTGTATTTCTGTAACCTAGCA • GTAACCTAGCAGCAAACAGTTAAAA • AAAGACCGTCGAGATGGGACATGAA
1567100_at	<ul style="list-style-type: none"> • ACAATAGAAGTGAACTTCCCTGAA • AAGCCCTGGGTCCACGTTGACCTAG • TCCACGTTGACCTAGCAGGGCAGGC • AGGGCAGGCTGGAACCTGATACTAT • CTGATACTATCAACCCAAGGGCTGC • GCTAAAGCAATCAGGTTCAACAACA • TATATGTTCCCTAATGTCGCTTTGTC • TCCCTAATGTCGCTTTGCCACAAAG • GGCAAAGCAAGCTATTTCTAGTAA • ATTAATCAATGGTCTAAACGTGCA • TGCAGATAAGGCCTAATCCCACACA
200068_s_at	<ul style="list-style-type: none"> • CCTCTCCCACACCACATATAGTGTA • TTCTTGCTGTTGAGTCTTTTTCCT • TTTTCTTTTCATCTGTTGTTCTGT • GTCACAGTGACCTTAGCTACATAGC • AGCTACATAGCAGACTTTCCCAAAT • ATGTCTGCAGGTTTCTCCTTGAAGC • TTGCATTTCCAGAAATCTGCCTCCT • CTTACCCTCCGTTGACAGTATATG • CAGTATATGTCATGCCTCACTTTCT • CTCACTTTCTTCTAGCTGAGCTTTA • ATTGTGATGATCGTTTCATTGCCTTC

FIG. 8B

Affymetrix ID	Secuencias de la sonda (5' -> 3')
201031_s_at	<ul style="list-style-type: none"> • TTAACAGTTTCTGCAATACAAGCTT • GCTTGTGATTTATGCTTACTCTAAG • GACTTAAGGCCAGTATTTTGAAT • AAGTTCAGCTTTTCTCAAGTTAGT • GCCCTTGCCACGTTAAATGAACA • ATTGAACACTGTTTGGATGCATGT • AGGAGCTGTGTCTACTATTAAGT • GACGCAATACCAACTTAGGATTT • GGATTTGGTCTTGGTGTGTATG • TGTATGAAATCTGAGGCCTGATT • GATTCCTTTTAGGTATATTGCGCT
202646_s_at	<ul style="list-style-type: none"> • ATAAGTAGCCGCCTGGTTACTGTGT • CGCCTGGTTACTGTGTCTGTAAAA • AAAATACAGACACTTGACCCCTGGT • CCTTGGTGTAGCTTCTGTTCAACTT • TGGATGGGTCTGATTTCTTGGCCCT • TTCTTGGCCCTCTTCTGAATTGGC • GAATTGGCCATATACAGGGTCCCTG • CCAGTGGACTGAAGGCTTTGTCTAA • GATGTGGGGGAGGGCGGTTTATCT • TTGAGGTTTTGATCTCTGGGTAAAG • GAGGCCGTTTATCTTTGTAAACACG
205758_at	<ul style="list-style-type: none"> • CAGCCCTTGCAATGACAGAGGGGCC • TGAAGAGGACAGGCTACCCCTTTA • AAACCTAAGGCCCTCTTGAATCTCTG • GATACAAACATGTTCCTGGGATCAC • AGACAATTGTTGGAGAGCCCTCAC • ATGAGGCAGACCTGACTCTCTTAAG • AACTGCTGTCCCAAACATGCACTTC • AAACATGCACTTCTTGTCTTAAGGT • ATGGTACAAGCAATGCCTGCCATT • TATGGTGTACATTCTTCTGATTAT • CTTCTGATTATTTCTACACATAC
206188_at	<ul style="list-style-type: none"> • AATGCGGAAGAGTTAGCCAGCGGTG • AAAACAACCTTTTATCAATGTCTGCA • GAAAGTCTTCTGGGATTTATAGATC • ATTCTGAGTTTTCCAATTGTTGAAC • ATTGTTGAACATGGTGTACTGCCCC • GTGTACTGCCCATTTATTTAGATC • CAGTTTGCAGCTCTCACATTTGTT • TGTATTTAATATTTCTGCATGCTAT • GCATTAGACTGTACATTGATCTTG • GATCAGTAACTTATTCTAGTAGCT • TCTAGATTCTTTAGCATTTTCTATG

FIG. 8C

Affymetrix ID	Secuencias de la sonda (5' -> 3')
212637_s_at	<ul style="list-style-type: none"> • TGGATAGAACCATAACTTACACATG • AAGTCATATACTAGATCCAATACTA • GGAAGGATTCATTGAGCAGCATAGA • GTTTGTTTACATGTTACTTTGAGAT • CTTTGAGATGCTAGGTATTTGTGGA • AAGAATCAGGCTCTTTTGTACTTTG • GTTTTTAAATCTGTGATGCTTTTCA • AATTGATGCAATTCATACTTAGGA • ATGTAAACTCTGCCACTTTTTGTG • GGTTTTTATGAAGCCAGATGGATTG • AATATAAGGCTAATGATTTTCTGTT
212920_at	<ul style="list-style-type: none"> • GGTTGACCTCAGCAGATGAAGTGAA • AGAAATTATCTGAATCTTGGTTTGT • AGATTTACAATCTACATGCAATATT • GCTTTTACAGTTTCACATGTGTACA • ATTGAACTTCTAAACTGGCATTGA • TGTTTTGTGCACCAATTTATAAA • GTGCAATACGTGTTACTTTCTGAG • GCCTTCTTTAGCAGCATTTGATGGA • TAAACCAGATTGCAATCCTTTTTT • TCCTAAACCATGTACCAAGTTTTTG • GAGTGTATTTCTGTAAGCATAGTTA
213317_at	<ul style="list-style-type: none"> • AAGTGTGCATAATTCATTTAACGT • TAGATCCAATTCCTTTCTTGCAACC • AAGGTTTGGGCTATTCTGTGTTTCT • GTACTTTTAAGGATCCAGTCATCTG • TTGAATTTGTCTTACTTTCTGGCC • CCACCTCGTTGTCCTATTTATAAG • AGTACTTGATTTTCTCGAGACCTT • GATTTTCTCGAGACCTTAACTATG • ATGAAGATACAGCTGAGTGTTTTCC • CTCATGAATCTGAACCAATTACCAA • ACCAATTTGTGTTCCAGTCTTGATT
213619_at	<ul style="list-style-type: none"> • AATCAAGGCCTCAGAATTCATACA • ATCCTAAGTATTGGTATTGCGTCTC • ATTGCGTCTCAAATTTTCCCATTA • GCTTAAACTTACGTGCCTTACAGGT • GTAGACACAGCTCCTATATTGTTTT • AACATAATTCAACTGCTTTTACCTA • GCTTTTACCTAAATACGCTTACTGC • GTACATCCTATAACTAACTTGAGAA • GTTATAGTTTACTCAGCTTCACTGT • TCAGCTTCACTGTTACATCCTAGAT • AACAATCCTGATTTCCACTTAGAGT

FIG. 8D

Affymetrix ID	Secuencias de la sonda (5' -> 3')
215443_at	<ul style="list-style-type: none"> • GGAAGAAAGTCCTTGCCTTTGAGA • GCTCCAGTATGCCATCATGATGCCT • ATCATGATGCCTGCTAAGGCAGCCA • GGCAGCCACCTTGGTGTACATGCCT • ACCTTGGTGTACATGCACAGAGG • GAAATGCAAGATCCACAAGTATG • GATGGAAGGCACTCTAGTCTTTGCA • AAGGCACTCTAGTCTTTGCAGAAA • ACCTGAATGTACATTGCACAATGCC • AATGATAGTTCGACTCGTCTGTGGA • GACTCGTCTGTGGAAGAACTTACAA
216198_at	<ul style="list-style-type: none"> • CCACACCAGGCAAATTTTGTCTTT • GCAATGATGAAAACAGTCTGTATCT • GTGTGGTCCAATATGCTAGCCACTA • TGCTAGCCACTAGCCATTTGTAGCT • AGCCACTAGCCATTTGTAGTACTG • GTAGCTACTGAACCTTGAAATGGGG • ATGGGGTTATTTTACTGAAGAACT • ATAGCACAGCTCTAGAAGAACCACC • GCTCTAGAAGAACCACCATTTACAA • AGGACAAAAGACAGCCATCATAATG • CATCATAATGTCCATTTTCCTTTCC
217870_s_at	<ul style="list-style-type: none"> • GGAATGAGTTCCTTATCTAGTGTTC • TCTTATCTAGTGTTCAGGCCAGCA • GGTTCGAGAGCATTCCCTACTCACA • GAGCATTCCCTACTCACATAAGTGAA • TGAGACTAGGTGCTTTGCTTCCTTT • GCTTCCTTTCATCAGGTATCTTTCT • AGGTATCTTTCTGTGGCATTGAGA • TTAATAAATTATGAGGCTTTGCTTT • TAACTAGTTTGTCAATCCATTTGT • GATACAGTCACCAAGAATGTTTGA • AGACCCCAATTTAAGCCTTGCTTAT
218087_s_at	<ul style="list-style-type: none"> • AACTTGTATAGCCCATGTACCTAC • GCCCATGTACCTACCTTGTATAGAA • GGAATTCCAAATAACTCACCTACAG • AACTCACCTACAGATTTTGTAGACA • GAGGAAAGGTAAGTCTGATCTCAGT • AGTCCATTTATCTGCAGTCGTATT • ATCTGCAGTCGTATTAAGTCCATTT • ATTGCCATTTAAATTACACACTGAG • CACACTGAGAGCATGTCCATATGCAG • GCAGACATAGATTTTCTGTTCATT • GAATTACTACATTTGCTGTACATAT

FIG. 8E

Affymetrix ID	Secuencias de la sonda (5' -> 3')
222145_at	<ul style="list-style-type: none"> • CTGGCTCTGGTTGAGACCCAGAAGA • GAGACCCAGAAGAGTCATGCATCAG • CAGGAAATACCCTGGCCTTTGTAGG • TTTGTAGGGACTGAGCCTGCACCGA • CCTGCACCGACGACTTCAATTGCAG • CATCTCTTCTGCAGCAAGATAACAT • GATAACATGCTACTAGGCCTCAATT • GCTACTAGGCCTCAATTCATTGCTA • AGGCCTCAATTCATTGCTAAACATT • AACATTTTTTAACAAGTATCTCACA • AAATGAATCTGGAGGTAATCCCAGC
223577_x_at	<ul style="list-style-type: none"> • TGCTGTTTTCTGCTCCCAGATGAAA • AACATTGGCACACAGCACAGCCTC • CCAACTCATCTCTCATTTATTTCCGG • ATTTCCGGCTTCTTTTATTCCAGGAT • GTAACATTTTCATTCTTTTCGCTT • CTTTTCGCTTTTATCTGCTTTTGT • CATTGCCCTTCATTGTATTTCTCA • GACATATAATCTTTCTGCCTTTAC • CTGCCTTACTTATCAATTCACCAA • CAAGGAGCTGTTTTCTCTGCATCTA • TCTAGGCCATCATACTGCCAGGCTG
224321_at	<ul style="list-style-type: none"> • TACAGGGCACTGTGGGTATAAATACT • GAAGTAGATAGAAACCGACCTGGAT • ACCCTATTGTTGATATGGACTCTAG • GTTCCGTTGGTCAAGTTATTGGATC • GGATCAATTGAGTATAGTAGTTCCG • GTATAGTAGTTCCGCTTTGACTGGTG • GTTCGCTTTGACTGGTGAAGTCTTA • AGCATGTAAGTCTCGGAGGTTGGGT • TACTGCTCGGAGGTTGGGTTCTGCT • CGGAGGTTGGGTTCTGCTCCGAGGT • TAATGCAGGTTGGTAGTTTAGGAC
224373_s_at	<ul style="list-style-type: none"> • GGCATAATTATAACAAGCTCCATCT • TGCCTACGACAAACAGACCTAAAAT • GACAAACAGACCTAAAATCGCTCAT • ACAGACCTAAAATCGCTCATTGCAT • GACCTAAAATCGCTCATTGCATACT • TAAAATCGCTCATTGCATACTCTTC • AAATCGCTCATTGCATACTCTTCAA • ATCGCTCATTGCATACTCTTCAATC • GCTCATTGCATACTCTTCAATCAGC • CCTCGTAGTAACAGCCATTCTCATC • TCGTAGTAACAGCCATTCTCATCCA

FIG. 8F

Affymetrix ID	Secuencias de la sonda (5' -> 3')
224644_at	<ul style="list-style-type: none"> • TCAAGTGCCCAAGTGTGAGACCCA • ACACTCACCAGTCTGTTCTGTAGTC • GATTGACTGTCTGTGAACTTGCAG • GAACTTGCAGGAACTGTTTCATAGT • GTTTCATAGTTTCATTAGCACAGAG • GCAAGGTTATTTTGCATCTGCATTT • AGGGTATTGCTTTTGTGATAGCATA • AATTGCTGGCAACTTTTGTATTCCC • CAACTTTTGTATTCCCATAGACTGG • GACTGGGGAAGCTTAATTGCCTTTA • ACTTCTGTTTATATCATTTGTACAA
226173_at	<ul style="list-style-type: none"> • TAGGATCCCCCTTGATGCAAGAGGT • TGCAAGAGGTAGGCTCCCCGAAAGA • CGACCACATCATGCGCAATGGGCTG • GCACTTTGACCGCCTTGTCGCAAAA • CTCTCTTTGCTGATTACCTGCAGTC • GCCACAGCCGCATCTTGATCAGGCT • TTGATCAGGCTTTCTCAAGCCCTCC • CAGATCTCCATTCTTTGCCATGAGG • CTGCCAGCCCTAGACTTGTGGAGC • TCCTCCCCAGGCCTAAGTATGTGGA • AGAGGCCACGCCTATTTATTCTGTT
226773_at	<ul style="list-style-type: none"> • CTTTAGCACCTCCATGGTTCTCATA • GGTTCTCATATACCCATGTCTGTAA • CAACTTGTGATGGCCTGCCTTTTGC • GCCTTTTGCTTTTGGCAGTTGGGAT • GAACAATTGACTCTCCATTGGTTG • GGTGGTCATACTTAGTGTCTAGGC • ATCATGGAGTTCTTCCACTTCCAAG • ATGACTCATTGCTGTTGGATTCTA • CTGATGTGTCCCCAAACTGATTTAA • TTCTACTGAAGTGCCCTTGTGTACA • GTAATTGCTGCTTTGACCATTGTTT
226880_at	<ul style="list-style-type: none"> • AGCTCCCTAAAGACTGTAGCAGGAT • TAAAAGGATCACTGGCTCCGAGTCT • ACTGGCTCCGAGTCTCTTTGAGATA • CAGAATTCATCTTTATCTCCTACCT • GATTACTATTCAACTTGTATTTTTT • GCGAAGTTTCAATAGTGTTCATCTC • GTGTTTCACTCAAATCTTATTGCTT • AATCTTATTGCTTTACAACCGTGGT • TACAACCGTGGTACACCTTTCATTA • TAATTTGACATGGCTTTCATTGGGA • GGGGAGGGCTGGCAAAGAGACCAAT

FIG. 8G

Affymetrix ID	Secuencias de la sonda (5' -> 3')
228980_at	<ul style="list-style-type: none"> • TGAGGTTGTTTACTCATTGCCCCTT • GTTCCAAGATCCCATTCAGTTAAT • CCCCTCATGGATCCTTGGTTTTAAT • TAGAATTATCCGTAGCTTTTGCTG • ATCCGTAGCTTTTGCTGATTACTC • AAGACAGATAGCTGACTGGTTCATA • AGGAAGTAGTGCTCAATATACATTG • GGAAATCTCTGGATGTGAATTGTT • GAATTGTTACTTCAAGTGGCTTTTA • GTGGCTTTTATTAAGATTTTCAG • GATTTTCTCAGACTTACTTGGAGGT
229569_at	<ul style="list-style-type: none"> • TCCCTCCTTTCTAATCCAAGATCAT • AACCATGTGGTTGTTCCACGGTTCA • TTCCACGGTTCATCTGGCTACCGTT • TAAATCAAGCTTGACCTCCTCTTTT • TTCACGTTGACTTCTATCTCAGGA • GATATTCTTCAGTTCATACTGCTG • GCTGCAGACACTGTAAGTGGTCTCC • GCGTGAAAACATCACACCGTGTGT • CCACCTTTTTCCGTCAAAGTGCTTG • GTCAAAGTGCTTGTATGGCTTTCA • ATGGCTTTCATAGCTGGGACAAGTA
231735_s_at	<ul style="list-style-type: none"> • TGAGATGGACATTGCCTCTTCATTG • TGGTAGATTCCGTAACCTTTAAATTG • GATAATGTTCTCATCAGTAGTAAGA • ATCTCAGGGTTATGCTTATTCCCCA • GCTTATTCCCAATGGAGGTATGAC • GTATGACATATAATCTTTTCTGCCT • TTTACTTATCAATTCACCAAGGAGC • ACCAAGGAGCTGTTTTCTCTGCATC • TTTTCTCTGCATCTAGGCCATCATA • TCTAGGCCATCATACTGCCAGGCTG • ACTGCCAGGCTGGTTATGACTCAGA
233765_at	<ul style="list-style-type: none"> • CAGACAGAGCACCTGGTAGAAGGGG • ATCTCCCAGCATAGTGCTCAAGCTC • CTCCCAGCATAGTGCTCAAGCTCTG • CCCAGCATAGTGCTCAAGCTCTGCT • AGCATAGTGCTCAAGCTCTGCTGAG • ATAGTGCTCAAGCTCTGCTGAGGGA • CAGCAGGGGTCGACAGACACTTCAT • CAGGGGTCGACAGACACTTCATACA • GGAACAAGCAGCAATCTTTGCCGTT • CAAGCAGCAATCTTTGCCGTTCTGC • AAAGGAATAGCATCAACATAAAAAA

FIG. 8H

Affymetrix ID	Secuencias de la sonda (5' -> 3')
235803_at	<ul style="list-style-type: none"> • GAGACTGGAGACATGATTGTGCCAC • CCACAACAAGGTACATTTAGGCAAA • AGTCTCACAAATGCCAATTATTGGTC • GAAACTCATATACTCCTTGGAAGGC • AGAATCATCTTGTGGAAAGTTGT • GAAAGTTGTTATTACCTATCAGCG • TACCTATCAGCGTATGGATCCAATA • GGATCCAATACTCTAAGAAGACCTC • AGACCTCATAATTTTACACCTAGGT • TTTACAACTTCACATATGTCCATA • GGAAATAACTAGGTCCACTGATAGG
236131_at	<ul style="list-style-type: none"> • TGGGAAGACTCACCACATACTCATC • AGTGTTAGGGTCACAGTCCCGGACT • GAGGCAAACCTGCATCAGTGGGAG • AAGGGAGACGTTTCCATCATTCTT • TTAGGAAGTGCCATCTGTACTCAGC • TGTACTCAGCACTGTGTTCCACATT • ACACACATGCCCTGTTGTCTGGGT • CAAAAGCAGCCTACAAGCGGGTCAG • GGGTCAGGAAGGCAGTCTCACAAAG • GGCAGACTGTTTCTCTATTGGTATG • GGCTTTTGCCATGCCTATCAAAGG
236953_s_at	<ul style="list-style-type: none"> • TTAGTGTTCACAGAGGATGGATATT • CGTCTGGATCACGGATGTAGGAAGT • GGATGTAGGAAGTGGTATGTATAGT • ATATCTATTAATTATCTTACTGGA • ATCTTACTGGAAATCACATCTTGC • ATCACATCTTGCACATGTCCTTGT • TTTGCACATGTCCTGTTTGTATTG • ATCAGAGTTGCTGAATCTAATTGTA • ATTGTAATTTCTTTAACGATTCATG • GAAATCACATGTTTTAACAACTT • ACTTTATTTGTACTTCTGTGGAAT
240544_at	<ul style="list-style-type: none"> • GCATGGCTGTATCTAAAGGCTTGA • ATATGTAACAATTTTGGGCAAAT • AAAGTTGGAATCTGTCTTAGAAGTA • CTTAGAAGTAATCTACTCCACCCT • GTTTGAGTGCTCTATTTAAAGCTAG • GTTCCCTATCCTTACGCAGGATTGT • TTGTTTTACTGTAGCCAGGTTAAG • GTTAAGTTAAGCTTCTTGAGGACT • GCTTCTTGAGGACTTTAGTACTGAA • AGTACTGAAAGACAATGGCTCACTT • ATGGCTCACTTGTAAGACCTAGTTG

FIG. 81

Affymetrix ID	Secuencias de la sonda (5' -> 3')
240971_x_at	<ul style="list-style-type: none"> • TTAGTATTTTTAGAGACAGGGTCTC • GGAGTGCAGTGGCACAAATCATAACT • AAAGATATCCTTTAGGCCTTGGCTG • GATATCCTTTAGGCCTTGGCTGAAC • ATCCTTTAGGCCTTGGCTGAACACA • GGCCTTGGCTGAACACAGGATCTCT • GGCTGAACACAGGATCTCTTTAAAA • AAAGATAACATTTTAGGCTAAGTGC • GGCTAAGTGCCTTAAAATTCCTACA • AGTGCGTTAAAATTCCTACAGTCAG • AATTCCTACAGTCAGATTTACAATA
244042_x_at	<ul style="list-style-type: none"> • GGGAAGATTGAATTCCATTGAATT • ATTCCATTGAATTCATGGAAGCAGG • GGAAGCAGGAATCCATCTGGTCATA • GAAGCAGGAATCCATCTGGTCATAT • AAGCAGGAATCCATCTGGTCATATT • AGCAGGAATCCATCTGGTCATATTT • GCAGGAATCCATCTGGTCATATTTT • CAGGAATCCATCTGGTCATATTTTA • AGGAATCCATCTGGTCATATTTTAA • GGAATCCATCTGGTCATATTTTAAA • ATCCATCTGGTCATATTTTAAATAA

FIG. 8J