

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 640 900**

51 Int. Cl.:

**G01N 33/68** (2006.01)

**G06F 19/00** (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **30.10.2009 PCT/EP2009/064410**

87 Fecha y número de publicación internacional: **06.05.2010 WO10049538**

96 Fecha de presentación y número de la solicitud europea: **30.10.2009 E 09752320 (3)**

97 Fecha y número de publicación de la concesión europea: **21.06.2017 EP 2362942**

54 Título: **Biomarcadores**

30 Prioridad:

**30.10.2008 US 109649 P**

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

**07.11.2017**

73 Titular/es:

**FIRALIS (100.0%)  
35 rue du Fort  
68330 Huningue, FR**

72 Inventor/es:

**DEVAUX, YVAN;  
WAGNER, DANIEL R.;  
AZUAJE, FRANCISCO y  
VAUSORT, MÉLANIE**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 640 900 T3

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

## DESCRIPCIÓN

Biomarcadores

### Campo de la invención

5 La presente invención se refiere a un conjunto de nuevos biomarcadores para evaluar el riesgo o la gravedad de la insuficiencia cardíaca (IC) en un paciente, particularmente después de que el paciente ha sufrido un infarto de miocardio (IM). Se describen además kits de diagnóstico para medir los niveles de los biomarcadores, no siendo los kits parte de la invención.

### Antecedentes de la invención

10 La insuficiencia cardíaca (IC) no es una enfermedad específica, sino un conjunto de signos y síntomas, todos ellos causados por una incapacidad del corazón para aumentar adecuadamente la capacidad cardíaca según sea necesario. Los pacientes típicamente presentan dificultad para respirar, edema y fatiga. La IC se ha convertido en una enfermedad de proporciones epidémicas, que afecta al 3 % de la población adulta. La mortalidad de la IC es más grave que muchas formas de cáncer con una supervivencia de cinco años menos en el 30 %. El infarto de miocardio (IM) es una de las principales causas de la IC. El 63 % de los pacientes desarrollan IC en los 6 años  
15 posteriores al IM. La remodelación ventricular izquierda contribuye en gran medida a la IC. Debido a que la HF se vuelve más común en las personas ancianas, el número de individuos afectados continuará aumentando con nuestra población envejecida. Una extensa investigación en la última década ha traído una mejor comprensión de la fisiopatología de la IC. Al mismo tiempo, se han identificado nuevas dianas terapéuticas, aunque solo unas pocas de ellas son de uso terapéutico potencial. Algunas de estas dianas se han denominado "biomarcadores pronósticos" o "biomarcadores" cuando su valor de pronóstico podría demostrarse.

20 Los biomarcadores pueden clasificarse en tres categorías. Los biomarcadores que pueden ayudar en el cuidado de individuos aparentemente sanos se denominan "biomarcadores de detección". Los biomarcadores que se observan en pacientes con sospecha de enfermedad se denominan "biomarcadores de diagnóstico" y los biomarcadores que se observan en pacientes con enfermedad manifiesta se denominan "biomarcadores pronósticos". Aunque en la práctica clínica se utilizan biomarcadores diagnósticos como la troponina I y la troponina T para el IM y el péptido natriurético cerebral (BNP) para la insuficiencia cardíaca, el uso potencial de estos biomarcadores como biomarcadores de pronósticos para adaptar o modificar el tratamiento del paciente individualmente ("medicina personalizada") no ha sido probado.

25 Debe tenerse en cuenta que cuanto más temprano se pueda identificar un paciente propenso a desarrollar IC después de IM, más eficientemente se puede ajustar el tratamiento. Sin embargo, una limitación mayor del uso de biomarcadores individuales es que no todos los pacientes pueden presentar tales factores de riesgo de forma independiente. Por lo tanto, se ha hecho evidente que la identificación temprana de los pacientes propensos a desarrollar IC después de IM se beneficiaría considerablemente de la multiplicación e integración de biomarcadores.

30 El factor de crecimiento endotelial vascular (VEGF), una sub-familia de factores de crecimiento, es un término amplio que cubre un número de proteínas de un número de familias. Estos factores de crecimiento se han estudiado principalmente por sus propiedades angiogénicas. Se han identificado hasta la fecha diversos factores de crecimiento, el mejor conocido de los cuales es el VEGFA. El VEGFA normalmente se denomina simplemente "VEGF" Ferrara y col. "The biology of VEGF and its receptors" Nat Med. 2003 jun; 9(6):669-676). Uno de los biomarcadores propuestos de la presente invención, VEGFB, es parte de la familia de factores de crecimiento endotelial vascular pero es distinto del VEGFA.

### Objetos de la invención

Los objetos de la presente invención son:

#### 1. Proporcionar una herramienta pronóstica

35 Es un objeto de la presente invención proporcionar una herramienta para el pronóstico temprano de la aparición de IC con el fin de mejorar la tasa supervivencia y disminuir el desarrollo de la IC.

Es otro objeto de la presente invención utilizar esta herramienta pronóstica para identificar pacientes en riesgo de desarrollar insuficiencia cardíaca.

Es otro objeto que no es parte la presente invención utilizar esta herramienta pronóstica para ajustar los tratamientos para prevenir de mejor manera el desarrollo de la remodelación ventricular y la IC después de IM.

#### 2. Proporcionar un nuevo kit de diagnóstico

50 Es un objeto, que no es parte la invención, proporcionar un kit de diagnóstico capaz de medir las concentraciones de marcadores en fluidos biológicos con el fin de ayudar en el pronóstico temprano de la aparición de IC para mejorar la supervivencia y disminuir el desarrollo de IC.

Es otro objeto, que no es parte la invención, proporcionar un nuevo kit de diagnóstico para identificar pacientes en riesgo de desarrollar remodelado ventricular y HF.

5 Es otro objeto, que no es parte la invención, proporcionar un nuevo kit de diagnóstico para ajustar los tratamientos para mejor prevenir el desarrollo de la remodelación ventricular y la IC después del IM. Además, es también un objeto de la presente invención proporcionar biomarcadores que puedan usarse en el proceso de diagnóstico de pacientes, post-IM, para la susceptibilidad de un paciente a desarrollar HF o Remodelación Ventricular.

Sorprendentemente, los presentes inventores han descubierto que el ARNm y los niveles plasmáticos de las siguientes proteínas VEGFB, THBS1 y PGF varían después de IM y son excelentes indicadores de la probabilidad de que el paciente continúe a desarrollar IC y/o Ventricular Remodelación.

10 Por lo tanto, estos biomarcadores pueden utilizarse para detectar pacientes con IM y, en particular, proporcionar una herramienta de pronóstico temprano para identificar a aquellos pacientes que, habiendo sufrido de IM, corren un mayor riesgo de desarrollar después un IC y/o la Remodelación Ventricular. También se proporcionan kits de diagnóstico para medir los niveles de estos tres biomarcadores y son útiles en el contexto del IM para predecir la ocurrencia de IC y/o Remodelación Ventricular.

15 **Sumario de la invención**

Se desvela en el presente documento un procedimiento para identificar los pacientes con infarto de miocardio que tienen un mayor riesgo de desarrollar una afección cardíaca, que comprende:

- 20
- analizar, después del infarto, una muestra de fluido corporal del paciente para medir los niveles del factor de crecimiento endotelial vascular B (VEGFB), trombospondina 1 (THBS1) y/o factor de crecimiento placentario (PGF);
  - comparar los niveles de VEGFB, THBS1 y/o PGF con los niveles correspondientes de VEGFB, THBS1 y/o PGF de una muestra de referencia, teniendo dicha muestra de referencia un resultado clínico conocido; y
  - determinar si el paciente tiene un mayor riesgo de desarrollar una afección cardíaca, sobre la base de dicha comparación.

25 De acuerdo con un primer aspecto, la presente invención proporciona un procedimiento para identificar pacientes con infarto de miocardio que tienen un mayor riesgo de desarrollar insuficiencia cardíaca, que comprende:

- 30
- analizar, después del infarto, una muestra de fluido corporal del paciente para medir los niveles del factor de crecimiento endotelial vascular B (VEGFB), trombospondina-1 (THBS1) y/o factor de crecimiento placentario (PGF);
  - comparar los niveles de VEGFB, THBS1 y/o PGF con los niveles correspondientes de VEGFB, THBS1 y/o PGF con un control; y
  - determinar si el paciente tiene un mayor riesgo de desarrollar una afección cardíaca, sobre la base de dicha comparación.

35 En algunas realizaciones, al menos uno de los siguientes indica una mayor probabilidad de que dicho paciente sufra de dicha afección cardíaca:

- 40
- niveles más bajos de VEGFB en la muestra del paciente analizado en comparación con el nivel de control de VEGFB;
  - niveles más altos de THBS1 en la muestra del paciente analizado en comparación con el nivel de control de THBS1; Y/o
  - niveles más altos de PGF en la muestra de paciente analizado en comparación con el nivel de control de PGF.

45 Como se desvela en el presente documento, la afección cardíaca puede ser infarto de miocardio, síndrome coronario agudo, cardiomiopatía isquémica o cardiomiopatía no isquémica. Más preferentemente, el paciente puede seguir desarrollando o sufrir de insuficiencia cardíaca. Preferentemente, el paciente puede desarrollar remodelación ventricular. Se notará que muchos pacientes con infarto de miocardio experimentan remodelación ventricular y posteriormente, o al mismo tiempo, desarrollan la condición conocida como insuficiencia cardíaca. Por lo tanto, existe una clara correlación entre la remodelación ventricular y la insuficiencia cardíaca y, preferentemente, los pacientes con infarto de miocardio sometidos a remodelado ventricular también desarrollará insuficiencia cardíaca.

50 Como se desvela en el presente documento la muestra de fluido corporal tomada del paciente puede ser una muestra de sangre, una muestra de plasma, o una muestra de suero. Preferentemente, los niveles de VEGFB, THBS1 y PGF ensayados son niveles de ARNm. Estos pueden determinarse mediante el análisis de ARNm en glóbulos rojos y/o blancos. Preferentemente, las células sanguíneas son leucocitos, neutrófilos, basófilos, eosinófilos, linfocitos, monocitos, plaquetas o eritrocitos.

55 Preferentemente, el VEGFB, THBS1 y PGF pueden medirse a nivel de ARNm a partir de células sanguíneas mediante cualquier técnica capaz de cuantificar ARNm, lo más recomendado es PCR cuantitativa, preferentemente micromatrices. También se prefiere que el VEGFB, THBS1 y PGF se puedan medir al nivel de las proteínas en el

plasma mediante cualquier técnica capaz de cuantificar proteínas, preferentemente ELISA. Los resultados del ensayo de ARNm pueden usarse en combinación con los resultados de los ensayos de proteínas plasmáticas para una evaluación más precisa.

5 En algunos casos, que no son parte de la invención, solo se ensaya el nivel de PGF, que puede ser preferentemente cualquiera de las secuencias dadas para PGF en SEQ ID NO 7-9, o fragmentos de las mismas. En otros casos, que no son parte de la invención, solo se ensaya el nivel de THBS1, que puede ser preferentemente una cualquiera de las secuencias de THBS1 proporcionadas en SEQ ID NO 4-6, o fragmentos de las mismas. En otros casos, que no son parte de la invención, se ensayan tanto los niveles de THBS1 como de PGF.

10 En casos todavía adicionales, se ensaya el nivel de VEGFB, que puede ser preferentemente cualquiera de las secuencias de VEGFB proporcionadas en SEQ ID NO 1-3, o fragmentos de las mismas. Esto es en combinación con THBS1 y PGF. Se prefiere que el VEGFB pueda ser la variante de empalme VEGFB 186, la variante de empalme VEGFB 167 o ambas. Como se desvela en el presente documento, el nivel de ARNm puede ensayarse el día del infarto de miocardio. Los niveles de proteína plasmática para el análisis pueden obtenerse del paciente el día 1, siendo el día siguiente al infarto de miocardio. También puede ensayarse una muestra posterior, obtenida del paciente más tarde en el día 1 o en el día 2 o 3 o 4 o 5 o 6 o 7 o en cualquier momento hasta 1 mes después del infarto.

20 Como se desvela en el presente documento, la muestra posterior se obtiene del paciente el día 1 y el valor ensayado del día a partir del día 1 se compara con la muestra de referencia. Sin embargo, en otros casos, la muestra posterior se obtiene del paciente el día 1 y su valor del día 1 se compara con los niveles ensayados en la muestra del día del infarto. Esto es lo más preferible con respecto a VEGFB. En otras palabras, el control, la primera muestra post-IM y las muestras siguientes pueden ensayarse para los niveles de VEGFB y estos niveles se comparan. Un aumento en los niveles de VEGFB desde la muestra de referencia hasta el día 1 o 2 es particularmente útil en la identificación del paciente como de menor riesgo de desarrollar dicha afección. De manera similar, una disminución en los niveles de VEGFB desde la muestra control hasta el día 1 o 2 es particularmente útil en la identificación del paciente con un riesgo mayor de desarrollar dicha afección.

25 Sin embargo, también se prefiere que se comparen los niveles de VEGFB de la primera muestra post-IM al día 0 o 1 con la muestra posterior (post IM) al día 1 o 2, el aumento en VEGFB desde el día 0 o 1 al día 1 o 2 es altamente indicativo de un paciente con menor riesgo de desarrollar las presentes afecciones. Los cambios en VEGFB medidos entre el día 0 y el día 2 o el día 1 y el día 3 se prefieren particularmente.

30 De hecho, los presentes inventores descubrieron que los niveles en plasma de VEGFB fueron similares entre grupos EF altos y bajos el día 0 y el día 1 después de IM. El día 2 sin embargo, los niveles de VEGFB aumentaron en pacientes de EF alto (2 veces en comparación con el día 0) mientras que cayeron en pacientes de EF bajo (2,5 veces en comparación con el día 0) (*Figura 9*). Estos datos están en concordancia con la regulación positiva del ARNm de VEGFB en pacientes con EF alto (*Figura 8*). Estos resultados sugieren que los pacientes capaces de aumentar su producción de VEGFB después de IM son más propensos a tener un resultado favorable.

35 Una disminución en VEGFB desde el día 0 o 1 hasta el día 1 o 2 y particularmente del día 0 al día 2, es altamente indicativa de un paciente con un elevado riesgo de desarrollar las presentes afecciones.

40 La presente invención mide los cambios en los niveles de expresión o prevalencia de ciertos biomarcadores, en lugar de solo la presencia o ausencia de los mismos, como es a veces en el caso de la técnica. Para medir un cambio en dichos niveles, necesita establecerse un punto de referencia. Este puede ser preferentemente los niveles en una muestra adicional, en particular una muestra temprana, preferentemente tomada el día del infarto (día cero) o el primer o el segundo día después del infarto. Preferentemente, la muestra control es el nivel basal de VEGFB, THBS1 o PGF el día del infarto, respectivamente.

45 Alternativamente, el control puede ser un valor de referencia obtenible de una población de pacientes infartados con un intervalo conocido de resultados clínicos. Puede construirse una base de datos de pacientes infartados y una vez calibrada para edad, sexo, etc., puede cerciorarse un valor promedio (modo, media o mediana como se considere apropiado) o un intervalo de valores para pacientes que tienen ciertos criterios (por ejemplo sexo, peso y edad), medidos en un momento particular después del infarto y con resultados clínicos conocidos (es decir insuficiencia cardíaca o no). Los datos del paciente ensayado pueden compararse entonces contra su valor de referencia o intervalo para determinar la probabilidad de que el paciente ensayado tenga uno u otro de los resultados clínicos. Por ejemplo, si se establece que los niveles de VEGFB menores de alrededor de "X" microorganismos/ml en plasma de pacientes masculinos infartados con edad de 55-60 medidos el día 1 después del infarto da un 50 % de probabilidad de insuficiencia cardíaca, entonces puede tomarse una muestra ensayada de un hombre correspondiente el día 1 y compararse con su valor para determinar si el 50 % de probabilidad se aplica a este paciente.

55 La etapa de determinación puede ser por medio de un análisis estadístico adecuado, por ejemplo técnicas de comparación de "vecino más cercano", tales como los programas Kstar y SVM. Un ejemplo particularmente preferido es el uso de una plataforma de extracción de datos, tal como Weka. Esto puede ser seguido por agrupamiento

jerárquico, preferentemente implementado utilizando el procedimiento de grupo de pares no ponderado con promedios aritméticos y coeficientes de correlación. La visualización en clústeres puede realizarse con GEPAS. Opcionalmente, las pruebas de significación estadística, los valores de correlación de Pearson y las gráficas se pueden generar con el paquete Statistica (v. 6.0).

5 El procedimiento puede comprender además la recogida de datos sobre uno o más pacientes IM, preferentemente para incluir el análisis de los niveles o valores de al menos uno de, y preferentemente cada uno de, los tres ARNm y/o proteínas plasmáticas (VEGFB, THBS1 y/o PGF) y el resultado clínico asociado para ese paciente. Esto se utiliza para crear datos característicos/valores para VEGFB, THBS1 y/o PGF asociados con un resultado clínico particular.

10 En lo sucesivo, una base de datos puede crearse con los datos de cada individuo. Estos valores conocidos podrían ser referidos como muestras de referencia (o "vistas").

Una muestra de "consulta" o "no vista", es decir, de un nuevo paciente recién infartado, se introduce y se comparado con los datos de referencia en la base de datos. Los datos de consulta se recogen de un paciente que acaba de tener un IM y para los que se desea establecer la probabilidad de desarrollar insuficiencia cardíaca. En otras palabras, el resultado clínico para este paciente es desconocido y el operador está pidiendo al programa predecir el resultado de este paciente (aumento o reducción de riesgo de IC después del infarto de miocardio).

15 El programa compara la muestra no vista/consulta con el conjunto de datos de referencia y hace su predicción sobre el resultado clínico del paciente.

20 Cuando se hace referencia al ARNm anterior, se apreciará que lo mismo es válido para los niveles de proteína plasmática.

Se prefiere un clasificador para determinar el pronóstico. El clasificador puede incluir programas como PAM, Kstar y SVM que son bien conocidos por hacer la predicción de diferentes maneras. Sin embargo, siempre existe una comparación de los datos de ARNm (o proteína) en la muestra "no vista" con uno o más de los puntos de datos "vistos", por ejemplo cuando se busca el "vecino más cercano". Kstar y SVM, por ejemplo, usan algoritmos diferentes, pero básicamente trabajan de manera similar comparando los datos o valores de la consulta con el conjunto de valores de referencia más cercano.

25 Por lo tanto, se prefiere que el clasificador busque en la base de datos y compare el conjunto de datos de la muestra de consulta/no vista con los conjuntos de datos conocidos/vistos. Habiendo encontrado el emparejamiento más cercano, (por ejemplo en un sentido 3D cuando se analizan los tres ARNm o niveles de proteína plasmática) en la base de datos, el programa basa su predicción en el resultado clínico del paciente de la consulta sobre la base del resultado clínico del emparejamiento más cercano.

También se proporciona un procedimiento que comprende:

- 35 - analizar una muestra de fluido corporal de un paciente post-IM para los niveles de ARNm o proteína plasmática de VEGFB, THBS1 y/o PGF para determinar los datos característicos del paciente, asociando estos datos característicos con un resultado clínico particular relacionado con la incidencia de insuficiencia cardíaca en dicho paciente, e introducir estos datos de entidad en una base de datos;
- repetir el análisis para una pluralidad de pacientes post-IM, para rellenar la base de datos para contener información de referencia sobre la relación entre HF y niveles de VEGFB, THBS1 y/o PGF;
- 40 - determinar los datos característicos de un paciente post-IM con pronóstico desconocido; y
- procesar los datos característicos para los niveles de VEGFB, THBS1 y/o PGF del paciente para compararlos con los datos característicos en la base de datos; y
- emitir un pronóstico aceptable para el paciente desconocido en función del resultado de la comparación.

Estas etapas también pueden incluirse en el procedimiento de acuerdo con el primer aspecto de la invención.

45 El pronóstico de probabilidad puede ser un aumento o disminución del riesgo de IC, lo que puede afectar el tratamiento clínico proscrito para el paciente. Los términos pronóstico y diagnóstico se pueden usar indistintamente, a menos que se evidencie otra cosa.

También se proporciona una base de datos que comprende datos característicos de pacientes con IM, incluyendo los datos el resultado clínico del paciente coincidiendo con al menos uno de, y preferentemente los tres, niveles de VEGFB, THBS1 y PGF analizados después del IM, como se describe en el presente documento. La invención también proporciona un dispositivo de comparación tal como un ordenador para acceder a la base de datos y/o procesar la consulta. La invención también proporciona un sistema que comprende una base de datos y al menos un ordenador para acceder y/o operar la base de datos. Un o el ordenador también puede usarse para procesar o administrar la consulta de la base de datos con los datos característicos del paciente a ensayar. La base de datos puede almacenarse centralmente, por ejemplo en un servidor, o puede mantenerse en el equipo de laboratorio o de campo usado para ensayar los niveles de VEGFB, THBS1 y PGF, o en un ordenador asociado a dicho equipo, como se analiza a continuación. El ordenador también puede estar situado centralmente con el servidor o de forma

remota, por ejemplo en un laboratorio intermedio, o situado en el equipo de campo.

5 Un kit, que puede incluir equipo de laboratorio o de campo para ensayar la muestra del paciente también se desvela en el presente documento. El equipo puede comprender la base de datos o puede simplemente comprender una visualización o lectura de los resultados del análisis. El equipo puede comprender la capacidad de contactar la base de datos de forma remota, por ejemplo a través de una red de Internet o de Internet, ya sea por cable o por transmisión inalámbrica.

También se desvela en el presente documento un procedimiento para obtener los datos característicos del paciente, mediante el análisis de los niveles de VEGFB, THBS1 y PGF y preferentemente asociando los datos del nivel con un paciente, por ejemplo utilizando un identificador de paciente, tal como un código.

10 También se desvela en el presente documento un procedimiento para recibir datos característicos para los niveles de VEGFB, THBS1 y/o PGF del paciente para compararlo con los datos en la base de datos. Los datos pueden procesarse por el receptor o transmitirse a un procesador.

15 Por lo tanto, se desvela en el presente documento un procedimiento para procesar datos característicos para los niveles de VEGFB, THBS1 y/o PGF del paciente y compararlos con los datos de entidad en la base de datos. El pronóstico probable para el paciente desconocido puede ser emitido por el procesador en función del resultado de la comparación y puede, opcionalmente, ser transmitido a un ordenador separado y/o al equipo o equipo discutido anteriormente, a través de una red, Internet y transmisión por cable o inalámbrica.

20 La base de datos también puede almacenarse en un medio portador, tal como un disco o dispositivo de memoria. Por lo tanto, la invención también proporciona un medio portador que comprende una base de datos dispuesta para que un computador determine la probabilidad de que un paciente desarrolle HF post-IM cuando se le consulte con datos característicos de los niveles de VEGFB, THBS1 y PGF de un paciente.

Los niveles de VEGFB, THBS1 y PGF pueden considerarse valores o relaciones y no necesariamente tienen que ser volúmenes o masa por unidad de volumen y así sucesivamente.

25 Preferentemente, la muestra de referencia es de un paciente en un grupo demográfico, genotípico o fenotípico similar al paciente. Se puede considerar que la muestra de referencia está en el mismo grupo demográfico que el paciente infartado si se cumplen cualquiera de los siguientes criterios: sexo, edad, raza u origen étnico e historial médico. Las muestras de control genotípicas o fenotípicas adecuadas pueden seleccionarse sobre la base de cualquier número de criterios de selección adecuados, tales como la determinación del genotipo de un paciente en uno o más loci, en particular aquellos conocidos por estar asociados con pacientes infartados, insuficiencia cardíaca y/o remodelación ventricular .

35 Determinar el genotipo puede comprender la detección de la presencia de un cambio de aminoácido en la secuencia del dominio de la hemopexina de MMP-9 (Metaloproteinasa 9 de matriz), la presencia de un cambio de aminoácido en dicho dominio indica la susceptibilidad a dicha afección cardíaca, infarto post-IM. Preferentemente, la secuencia que se detecta comprende o codifica un residuo de aminoácido Glutamina (Gln) o Arginina (Arg) en una posición correspondiente a la posición 148 del dominio de hemopexina de MMP-9. Preferentemente, la secuencia detectada es SEQ ID NO. 13, que es la secuencia de aminoácidos del dominio de hemopexina de MMP-9 del grupo de riesgo (susceptible) (que muestra Arg en la posición de aminoácido 148), o una secuencia polinucleotídica que la codifica. El polimorfismo de nucleótidos único (SNP), presente en la secuencia codificante del gen MMP-9, se encuentra en diferentes frecuencias en pacientes con pronósticos buenos o malos para la insuficiencia cardíaca después del infarto de miocardio y puede conducir a un único cambio de aminoácido en la hemopexina de la proteína MMP-9 transcrita y activa, dando como resultado un cambio electrostático en el sitio en MMP-9 que se une al inhibidor de tejido de la metaloproteinasa-1 (TIMP-1). Esto es significativo ya que TIMP-1 es el principal inhibidor de la actividad de MMP-9. Por lo tanto, puede ser útil determinar si el presente paciente tiene el SNP anterior y, si lo hace, comparar los niveles de VEGFB, THBS1 y PGF (en las muestras ensayadas de la presente invención) con los niveles en muestras de pacientes también conocidos por tener el SNP anterior, o viceversa.

50 Preferentemente, la determinación de un riesgo disminuido de una afección cardíaca es relativa a aquellos pacientes de referencia infartados que tienen niveles relativamente altos de ARNm de VEGFB (> -1,4), niveles relativamente bajos de THBS1 (<0) y niveles relativamente bajos PGF (< -0,1), respectivamente. Estos valores se expresan como relación logarítmica (ARN del paciente / ARN de referencia). Lo contrario es válido para un mayor riesgo de desarrollar una afección cardíaca.

Los valores anteriores derivan de la Tabla 4, aunque se prefiere que éstos puedan variar al menos 1 o 2 %, más preferentemente al menos 5 %, más preferentemente al menos 7 %, más preferentemente al menos 10 %, más preferentemente al menos 15 %, más preferentemente al menos 20 %, más preferentemente al menos 25 %, más preferentemente al menos 30 %, más preferentemente al menos 40 %, e incluso hasta 50 %.

55 Dado que el único biomarcador establecido de HF (insuficiencia cardíaca) es pro-BNP (péptido natriurético cerebral), el desempeño pronóstico del conjunto de 3 biomarcadores revelado en el presente documento se comparó con el desempeño pronóstico de NT-pro-BNP. El desempeño pronóstico del nivel plasmático de NT-pro-BNP, medido 1 día

después del IM, fue moderado (AUC = 0,63, Tabla 5). Por lo tanto, el conjunto de 3 biomarcadores descritos en el presente documento claramente superó el valor pronóstico de NT-pro-BNP. No obstante, se prefiere que el BNP también se pueda analizar en el presente procedimiento, ya sea para aumentar la precisión o confirmar un pronóstico determinado. Este nivel ensayado de BNP puede ser comparado con un nivel basal de BNP o un nivel de referencia como se discute en el presente documento.

5

Las secuencias de nucleótidos y proteínas para pro-BNP se proporcionan en las SEQ ID NO 10-11.

También se desvela en el presente documento un procedimiento para establecer un pronóstico para un paciente con infarto de miocardio, comprendiendo el procedimiento:

- 10
- ensayar, después del infarto, una muestra de fluido corporal del paciente para niveles de factor de crecimiento endotelial vascular B (VEGFB) y niveles plasmáticos bajos de trombospondina-1 (THBS1) y factor de crecimiento placentario (PGF);
  - comparar los niveles de VEGFB, THBS1 y PGF con los niveles correspondientes de VEGFB, THBS1 y PGF a partir de una muestra de referencia, teniendo dicha muestra de referencia un resultado clínico conocido; y
  - determinar el pronóstico para dicho paciente sobre la base de dicha comparación.

- 15
- Los niveles de Trombospondina-1 (THBS1) y/o Factor de Crecimiento Placentario (PGF) son preferentemente niveles en plasma bajos.

También se proporciona un procedimiento para establecer un pronóstico para un paciente con infarto de miocardio, comprendiendo el procedimiento:

- 20
- ensayar, después del infarto, una muestra de fluido corporal del paciente para niveles de factor de crecimiento endotelial vascular B (VEGFB) y niveles plasmáticos bajos de trombospondina-1 (THBS1) y factor de crecimiento placentario (PGF);
  - comparar los niveles de VEGFB, THBS1 y PGF con un control; y
  - determinar el pronóstico para dicho paciente sobre la base de dicha comparación.

- 25
- Esto puede conseguirse comparando los niveles de VEGFB, THBS1 y/o PGF con el nivel correspondiente de VEGFB, THBS1 y/o PGF en una muestra de referencia; en el que al menos uno de los siguientes indica una mayor probabilidad de un pronóstico desfavorable para dicho paciente:

- 30
- niveles más bajos de VEGFB en la muestra del paciente ensayado en comparación con el nivel de control de VEGFB;
  - niveles más altos de THBS1 en la muestra del paciente ensayado en comparación con el nivel de control de THBS1; y
  - niveles más altos de PGF en la muestra de paciente ensayada en comparación con el nivel de control de PGF.

También se proporciona un procedimiento para establecer un pronóstico para un paciente con infarto de miocardio, comprendiendo el procedimiento:

- 35
- analizar, después del infarto, una muestra de fluido corporal del paciente para niveles de factor de crecimiento endotelial vascular B (VEGFB) y niveles plasmáticos bajos de trombospondina-1 (THBS1) y factor de crecimiento placentario (PGF); y
  - consultar un programa estadístico previamente construido (también llamado "clasificador") con los niveles de VEGFB, THBS1 y PGF;

40

en el que

- un alto nivel de VEGFB y bajos niveles de THBS1 y PGF se asocian con una mayor probabilidad de desarrollar dicha afección cardíaca; y
- el clasificador determinará si el paciente tiene una mayor probabilidad de desarrollar dicha afección cardíaca.

- 45
- Los niveles de Trombospondina-1 (THBS1) y/o Factor de Crecimiento Placentario (PGF) son preferentemente niveles en plasma bajos.

El pronóstico desfavorable es preferentemente que el paciente tenga una probabilidad aumentada de sufrir dicha afección cardíaca.

También se desvela en el presente documento un procedimiento para determinar la probabilidad de que un paciente con infarto de miocardio desarrolle una afección cardíaca, que comprende las etapas anteriores.

- 50
- También se desvela en el presente documento un procedimiento para identificar pacientes con infarto de miocardio que tienen un riesgo reducido de desarrollar una afección cardíaca, que comprende:

- ensayar, después del infarto, una muestra de fluido corporal del paciente para niveles de factor de crecimiento

- endotelial vascular B (VEGFB), trombospondina-1 (THBS1) y factor de crecimiento placentario (PGF);
- comparar los niveles de VEGFB, THBS1 y PGF con los niveles correspondientes de VEGFB, THBS1 y/o PGF a partir de una muestra de referencia, teniendo dicha muestra de referencia un resultado clínico conocido; y
- determinar si el paciente tiene un riesgo reducido sobre la base de dicha comparación.

5 De acuerdo con un aspecto, la presente invención proporciona un procedimiento para identificar pacientes con infarto de miocardio que tienen un riesgo reducido de desarrollar insuficiencia cardíaca, que comprende:

- ensayar, después del infarto, una muestra de fluido corporal del paciente para niveles de factor de crecimiento endotelial vascular B (VEGFB), trombospondina 1 (THBS1) y factor de crecimiento placentario (PGF);
- comparar los niveles de VEGFB, THBS1 y PGF con los niveles correspondientes de VEGFB, THBS1 y/o PGF a partir de un control; y
- determinar si el paciente indicó un mayor riesgo sobre la base de dicha comparación.

Este procedimiento puede modificarse para incluir la comparación de los niveles de VEGFB, THBS1 y/o PGF con el nivel correspondiente de VEGFB, THBS1 y PGF en una referencia; en el que al menos uno de los siguientes indica una menor probabilidad de que dicho paciente sufra de dicha afección cardíaca:

- 15 - niveles más altos de VEGFB en la muestra del paciente analizada en comparación con el nivel de control de VEGFB;
- niveles más bajos de THBS1 en la muestra de paciente analizada en comparación con el nivel de control de THBS1; y/o
- niveles más bajos de PGF en la muestra del paciente analizada en comparación con el nivel de control de PGF.

20 Este procedimiento puede mejorarse para incluir:

- consultar un programa estadístico previamente construido (también llamado "clasificador") con los niveles de VEGFB, THBS1 y/o PGF;

en el que

- 25 - un alto nivel de VEGFB y bajos niveles de THBS1 y PGF se asocian a una menor probabilidad de desarrollar dicha afección cardíaca; y
- el clasificador determinará si el paciente tiene una menor probabilidad de desarrollar dicha afección cardíaca.

Se desvela en el presente documento un procedimiento de selección de pacientes con infarto de miocardio para evaluar el riesgo que cada paciente pueda tener de desarrollar una afección cardíaca. Esto puede ser un riesgo mayor o menor.

30 Los procedimientos de la invención y como se desvelan en el presente documento correlacionan la medición tres biomarcadores, con un mejor resultado clínico después de IM. Más preferentemente, el primer biomarcador es VEGFB y si su nivel es alto en el día 1 post-IM, entonces este paciente tiene un resultado clínico más favorable después de IM. Más preferentemente, el segundo biomarcador es THBS1 y si su nivel es bajo en el día 1 post-IM, entonces este paciente tiene un resultado clínico más favorable después de IM. Más preferentemente, el tercer biomarcador es PGF y si su nivel es bajo en el día 1 post-IM, entonces este paciente tiene un resultado clínico más favorable después de IM.

Se apreciará que los presentes procedimientos son útiles para establecer un pronóstico en pacientes con IM, correlacionando una evaluación combinada de múltiples biomarcadores, los cuales, dependiendo de sus niveles, pueden indicar un mejor resultado clínico después de IM.

40 La invención también puede usarse en el establecimiento de una medicina personalizada, por ejemplo en un procedimiento para proporcionar o mejorar la estrategia terapéutica de un paciente después de IM, sobre la base de la identificación de aquellos pacientes en riesgo de desarrollar una afección cardíaca. Esto puede ser a través del análisis de los niveles de ARNm de células sanguíneas o niveles de proteína plasmática de VEGFB, THBS1 y/o PGF.

45 Los kits de diagnóstico para uso en la presente invención son fácilmente disponibles para THBS1 y PGF, tales como los disponibles de R&D Systems, Inc. Sin embargo, para VEGFB, fue necesario construir nuestro propio kit de diagnóstico, como se discute a continuación. De hecho, el único kit comercialmente disponible de VEGFB (de USCNLIFE, VEGFB E0144h) no era lo suficientemente sensible para detectar niveles bajos de VEGFB en plasma. Usando quimioluminiscencia mejorada como el procedimiento de detección y una etapa de amplificación con biotina-estreptavidina, el límite de detección de nuestro kit fue de 10 pg/ml mientras que el USCNLIFE kit, que utiliza una detección colorimétrica clásica, se encontró en alrededor de 100 pg/ml. Por lo tanto, la invención también proporciona un procedimiento para analizar niveles de VEGFB en una muestra, que comprende:

- (a) poner en contacto la muestra con al menos un reactivo de captura inmovilizado sobre un soporte para formar un complejo reactivo de captura inmovilizado-muestra;

- (b) separar la muestra del al menos un reactivo de captura inmovilizado;
- (c) poner en contacto el complejo de reactivo de captura inmovilizado con un anticuerpo secundario específico para VEGFB y, opcionalmente, poner en contacto el anticuerpo secundario con un anticuerpo terciario específico para el anticuerpo secundario;
- 5 (d) poner en contacto el anticuerpo secundario o terciario con una molécula de unión, tal como estreptavidina, conjugada con medios de detección; y
- (e) medir el nivel del anticuerpo secundario o terciario unido a los reactivos de captura usando los medios de detección.

10 El reactivo de captura puede ser un anticuerpo, más preferentemente uno que reconoce el mismo epítipo como clon 58013 monoclonal de ratón de anticuerpo contra VEGFB humano, estando dicho anticuerpo monoclonal unido preferentemente específicamente a VEGFB 167 y/o VEGFB 186.

15 El anticuerpo secundario puede ser un anticuerpo que reconoce el mismo epítipo que el anticuerpo de cabra policlonal que se une específicamente a VEGFB 167 y/o VEGFB 186. El anticuerpo terciario puede ser un anticuerpo conjugado con biotina específico para el anticuerpo secundario, por ejemplo un Ac de burro Anti-cabra. Preferentemente, los medios de detección comprenden una actividad de fosfatasa alcalina.

Se desvela en el presente documento un procedimiento para analizar niveles de VEGFB en una muestra, que comprende:

- 20 (a) poner en contacto y, opcionalmente, incubar la muestra con un reactivo de captura inmovilizado en un soporte sólido, en el que el reactivo de captura es un anticuerpo que reconoce el mismo epítipo que el anticuerpo monoclonal de ratón 58013 contra VEGFB humano, uniendo dicho anticuerpo monoclonal específicamente a VEGFB 167 y VEGFB 186 para formar un complejo reactivo de captura inmovilizado-VEGFB;
- (b) separar la muestra de los reactivos de captura inmovilizados;
- 25 (c) poner en contacto el complejo de reactivo de captura inmovilizado-VEGFB con un anticuerpo secundario, en el que el anticuerpo secundario es un anticuerpo que reconoce el mismo epítipo como anticuerpo policlonal de cabra que se une específicamente a VEGFB 167 y VEGFB 186;
- (d) poner en contacto el anticuerpo secundario con un anticuerpo terciario, en el que el anticuerpo terciario es un anticuerpo de burro anti-cabra conjugado con biotina específico para el anticuerpo secundario;
- (e) poner en contacto el anticuerpo terciario con estreptavidina conjugada con fosfatasa alcalina;
- 30 (f) medir el nivel de VEGFB 186 o VEGFB 167 unido a los reactivos de captura usando un medio de detección para el anticuerpo detectable.

Además, los presentes inventores desvelan un kit ELISA para medir los niveles de VEGFB 186 y VEGFB 167 en fluidos biológicos y su uso como una herramienta de diagnóstico para identificar a los pacientes en riesgo de desarrollar IC después de IM.

El kit comprende preferentemente;

- 35 (a) al menos un reactivo de captura inmovilizado sobre un soporte;
- (b) un anticuerpo secundario específico para VEGFB 186 y/o VEGFB;
- (c) opcionalmente, un anticuerpo terciario específico para el anticuerpo secundario;
- (d) una molécula de unión, tal como estreptavidina, conjugada con medios de detección; y
- 40 (e) medios para medir el nivel del anticuerpo secundario o terciario unido a los reactivos de captura usando los medios de detección.

45 Como anteriormente, el reactivo de captura es un anticuerpo, más preferentemente uno que reconoce el mismo epítipo que el anticuerpo clon monoclonal de ratón 58013 frente a VEGFB humano, estando dicho anticuerpo monoclonal unido preferentemente específicamente a VEGFB 167 y/o VEGFB 186. El anticuerpo secundario puede ser un anticuerpo que reconoce el mismo epítipo que anticuerpo policlonal de cabra que se une específicamente a VEGFB 167 y/o VEGFB 186. El anticuerpo terciario puede ser un anticuerpo conjugado con biotina específico para el anticuerpo secundario, por ejemplo, un Ac de burro anti-cabra. Preferentemente, los medios de detección comprenden una actividad de fosfatasa alcalina.

50 Dicha muestra biológica puede, preferentemente, aislarse de un sujeto humano y puede ser plasma o suero. También se prefiere que los reactivos de captura inmovilizados se recubran sobre una placa de microtitulación. Preferentemente, la detección se amplifica mediante un reactivo quimioluminiscente. El VEGFB 167 humano purificado puede proporcionarse como un patrón de antígeno.

Los presentes inventores compararon su kit VEGFB con el único kit VEGFB comercialmente disponible que encontraron (USCNLIFE VEGFB E0144h). El kit de los presentes inventores es más sensible y por lo tanto permite medir VEGFB en más pacientes que el kit USCNLIFE.

55 De acuerdo con un aspecto adicional, la presente invención proporciona un procedimiento para identificar pacientes con infarto de miocardio que tienen un riesgo disminuido de desarrollar una afección cardíaca, que comprende:

- analizar, después del infarto, una muestra de fluido corporal del paciente para niveles de factor de crecimiento endotelial vascular B (VEGFB), trombospondina 1 (THBS1) y/o factor de crecimiento placentario (PGF); y
- consultar un programa estadístico previamente construido (también llamado "clasificador") con los niveles de VEGFB, THBS1 y PGF;

5 en el que

- un alto nivel de VEGFB y bajos niveles de THBS1 y PGF se asocian a una menor probabilidad de desarrollar insuficiencia cardíaca; y
- el clasificador determinará si el paciente tiene una menor probabilidad de desarrollar dicha afección cardíaca.

10 Los presentes inventores tienen una gran base de datos de pacientes con IM agudo con más de 20 parámetros clínicos y 4 meses y 1 año de seguimiento. Entre otros parámetros clínicos, la fracción de eyección (EF) del corazón, que representa la capacidad del corazón para bombear sangre a las arterias periféricas, se midió por ecocardiografía el día del infarto, 4 meses después y 1 año después. Se asume que los pacientes con EF a los 4 meses  $\leq 40\%$  sufren de remodelación mientras que los pacientes con EF a los 4 meses  $> 40\%$  se están recuperando normalmente. Varios protocolos realizados en el laboratorio de los presentes inventores utilizaron la base de datos para identificar los biomarcadores tempranos de la aparición de la IC después del IM a través de diferentes enfoques. Los biomarcadores identificados por dos enfoques fundamentalmente diferentes se combinaron y la combinación de los biomarcadores más predictivos se definió como el "conjunto pronóstico".

15 El primer enfoque involucró la tecnología de micromatrices de ADN. Se analizaron las biofirmas o perfiles de expresión génica de células sanguíneas circulantes a partir de muestras de sangre extraídas el día del IM. Esta tecnología permitió la identificación de genes regulados diferencialmente entre dos grupos de pacientes con fenotipos extremos, es decir, un grupo de pacientes con un resultado clínico favorable después del IM (grupo EF alto, EF  $> 40\%$ ) y un grupo con un resultado desfavorable después del IM Bajo EF, EF =  $40\%$ ). Los presentes inventores caracterizaron las biofirmas de las células sanguíneas de 32 pacientes, 16 de cada EF alta y baja EF grupo. Utilizando el algoritmo SAM (Análisis Estadístico de Micromatrices) y un doble cambio de 1,3, se encontraron 25 525 genes expresados diferencialmente entre los 2 grupos de pacientes (Tasa de Descubrimiento Falso de 24,5%). Entre estos genes, 9 tenían un valor pronóstico significativo para la IC.

20 El segundo enfoque se basó en una caracterización bioinformática de una red de interacción proteína-proteína de angiogénesis en el IM humano. De hecho, la angiogénesis es uno de los procesos de curación beneficiosos que tienen lugar en el corazón después de IM y un defecto en la angiogénesis puede conducir a la IC. La red fue construida con interacciones proteína-proteína anotadas de la Human Protein Reference Database. Esta red global consistía en 556 nodos (es decir, proteínas) y 686 bordes (es decir, interacciones). Después de la posterior red sobre la base de la expresión génica y análisis, 38 genes de la red de genes mostraron un significativo valor pronóstico. Curiosamente, la combinación de los modelos de clasificación basados en la expresión génica con los modelos de clasificación sobre la base de la red produjo un número reducido de biomarcadores candidatos con un valor pronóstico mucho mejor que cada enfoque considerado por separado. El área bajo la curva (AUC), que representa la capacidad pronóstica de los biomarcadores, estaba entre 0,56 y 0,72 para el modelo de clasificación sobre la base de la expresión génica, y entre 0,56 y 0,73 para el modelo de clasificación sobre la base de la red. Cuando se combinaron los dos modelos, se implementó un conjunto de 3 biomarcadores con un AUC de 0,82 (es decir, con un fuerte valor pronóstico para la aparición de IC): este conjunto se denomina en el presente documento el "conjunto de pronóstico". Estos 3 biomarcadores fueron: trombospondina-1 (THBS1), factor de crecimiento placentario (PGF o PIGF) y factor de crecimiento endotelial vascular B (VEGFB). Aunque THBS1 tiene propiedades anti-angiogénicas, VEGFB y PGF deben su capacidad pro-angiogénica para estimular el crecimiento y la multiplicación de las células endoteliales vasculares.

35 Luego, se verificó que las diferencias en la expresión de genes observados en el nivel de ARNm por micromatrices fueran eficaces a nivel de proteína. Para este propósito, los niveles plasmáticos de los 3 biomarcadores se midieron mediante ensayo inmunoenzimático (ELISA). Estos experimentos atestiguaron que los niveles de proteína de THBS1 y PGF eran significativamente distinguibles entre pacientes con EF alta y baja.

40 Dado que los kits de VEGFB ELISA comercialmente disponibles no se encontraron lo suficientemente sensibles para detectar VEGFB en las presentes muestras de plasma, los presentes inventores diseñaron su propio kit que permite la cuantificación de VEGFB en fluidos biológicos tales como el plasma humano.

45 Por "resultado favorable", se entenderá que esto significa un menor riesgo de que el paciente desarrolle una afección cardíaca, tal como insuficiencia cardíaca y/o sufra de remodelación ventricular izquierda. Se cree que Remodelación ventricular derecha es menos relevante, por lo que no se prefiere.

50 Por "resultado desfavorable", se entenderá que esto significa un mayor riesgo de que el paciente desarrolle una afección cardíaca, tal como insuficiencia cardíaca y/o sufra de remodelación ventricular izquierda.

55 También se comprenderá que mientras se asocian los riesgos de insuficiencia cardíaca y remodelación ventricular izquierda, estas son afecciones separadas y, por lo tanto, un paciente podría sufrir uno, pero no el otro. Por lo tanto, el mayor riesgo de insuficiencia cardíaca o remodelación ventricular es desfavorable.

Se apreciará que las secuencias de ARN dadas en el listado de secuencias comprenden Timina (T) ya que es cómo se representa en la página web de NCBI. En cada caso, está claro que se contempla el reemplazamiento de T por Uracilo (U).

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

- 5 La Figura 1 es una ilustración de los datos de micromatrices interpretados con el algoritmo SAM.
- La Figura 2 muestra la red de interacción proteína - proteína de la angiogénesis en el IM humano.
- La Figura 3 ilustra la estrategia utilizada para el análisis combinado de los modelos de clasificación basados en la expresión génica con los modelos de clasificación basados en la red.
- 10 La Figura 4 muestra un mapa térmico que ilustra las diferencias en la expresión (micromatrices) de los biomarcadores entre pacientes con fracción de eyección alta (H) y baja (L).
- La Figura 5 muestra gráficas cuantil-cuantil que ilustran la relación entre la fracción de eyección y la expresión de los biomarcadores evaluados por micromatrices y ELISA.
- La Figura 6 muestra gráficos de dispersión que ilustran la relación entre la fracción de eyección y la expresión de los biomarcadores evaluados por micromatrices y ELISA.
- 15 La Figura 7 representa la evolución de los niveles plasmáticos de VEGFB entre el día del infarto (día 0) y el día después (día 1). Mientras que el VEGFB plasmático disminuye entre el día 0 y el día 1 en los pacientes con EF baja (-10 %), los pacientes con EF alta tienen niveles crecientes de VEGFB (+ 15,4 %).
- La Figura 8A muestra los valores de expresión de VEGFB usando PCR cuantitativa y micromatrices para pacientes con EF alto y EF bajo
- 20 La Figura 8B muestra la correlación significativa observada entre la expresión de VEGFB y la fracción de eyección.
- La Figura 9 muestra los niveles de VEGFB entre los grupos EF alto y bajo en el día 0, el día 1 y el día 2 después del IM.
- 25 La Tabla 1 es un resumen de los resultados predictivos de los modelos de clasificación basados en los niveles de ARNm de VEGFB, THBS1 y PGF.
- La Tabla 2 es un resumen de las conclusiones alcanzadas en los experimentos subyacentes a la presente invención.
- La Tabla 3 es un resumen de las estadísticas realizadas para comparar los niveles de VEGFB, THBS1 y PGF entre pacientes con EF alta y pacientes con EF baja.
- 30 La Tabla 4 es un resumen de las estadísticas de la comparación entre los niveles de ARNm de los 3 biomarcadores en los dos grupos.
- La Tabla 5 es un resumen del rendimiento predictivo de NT-pro-BNP.
- La Tabla 6 es una lista de 28 genes angiogénicos expresados diferencialmente entre grupos EF altos y bajos.
- La Tabla 7 muestra rendimientos de predicción usando dos modelos de aprendizaje automático.

### **Descripción detallada**

- La insuficiencia cardíaca (IC) es la principal complicación del infarto de miocardio (IM). Los datos recientes mostraron que el 63 % de los pacientes desarrollan IC en los 6 años posteriores al IM. La angiogénesis es un fenómeno clave involucrado en la reparación del miocardio después de IM. La angiogénesis está estrechamente regulada por un equilibrio que se rige por un gran número de factores angiogénicos, siendo algunos de ellos pro- y otros anti-angiogénicos. Una desregulación de este equilibrio puede conducir a una angiogénesis inapropiada y puede sentar las bases para el desarrollo de IC después de un episodio de IM.
- 40 Los informes recientes indican claramente que la identificación temprana de los pacientes propensos a sufrir de IC después de IM puede mejorar significativamente la adaptación de las estrategias terapéuticas al paciente individual ("medicina personalizada"). Al igual que muchos otros trastornos cardiovasculares, la IC es una enfermedad multifactorial. Por lo tanto, el uso de biomarcadores únicos para predecir la aparición de IC después de IM tiene un valor limitado. La multiplicación de biomarcadores pronósticos puede ser de interés para predecir finamente la IC.
- 45

En primer lugar, los presentes inventores hipotizaron que un conjunto de biomarcadores pronóstico podría identificarse a través del análisis de las biofirmas de las células sanguíneas circulantes. En segundo lugar,

hipotetizaron que un enfoque sobre la base de la red de interacción proteína-proteína también puede tener el potencial de destacar biomarcadores pronósticos de la IC. Y en tercer lugar, los presentes inventores probaron si la combinación de los datos obtenidos por estos dos enfoques independientes permitiría alcanzar un mayor nivel de pronóstico que cada enfoque tomado por separado.

5 Para este propósito, se seleccionaron dos grupos de pacientes (n = 16 por grupo) con IM, un grupo de pacientes con resultado favorable después del IM (Fracción de Eyección (EF) > 40 %) y un grupo de pacientes con resultado desfavorable (EF ≤ 40 %). Esta estrategia para seleccionar "fenotipos extremos" se eligió para aumentar las posibilidades de encontrar diferencialmente expresado genes entre los dos grupos sin la necesidad de un gran tamaño de la muestra. El ARN se extrajo de células enteras aisladas de la sangre periférica de estos pacientes. Las biofirmas se determinaron por perfiles de micromatrices. Después de varios procedimientos de normalización, filtrado y estadística, se encontró que un conjunto de 525 genes se expresaban diferencialmente entre los dos grupos de pacientes (cambio de 1,3 veces, tasa de descubrimiento falso 24,5 %). Entre ellos, un grupo de 47 genes con valor pronóstico moderado se identificó por modelos de clasificación con un máximo AUC de 0,72. El filtrado adicional de estos genes llevó a un grupo de 9 genes con un valor pronóstico equivalente (AUC 0,68).

15 En un intento de aumentar la fuerza de la predicción proporcionada por los genes recuperados de experimentos de micromatriz, y teniendo en cuenta la importancia de la angiogénesis en la reparación del corazón después de IM, se dibujó una red de interacción proteína-proteína de angiogénesis en el IM humano. Esta red se ensambló mediante la extracción de genes conocidos por estar implicados en este proceso y las correspondientes interacciones (curadas) proteína-proteína de bases de datos públicas. Los análisis de agrupación de esta red informaron de un módulo significativamente asociado al crecimiento de las células y la regulación del crecimiento. Dentro de este grupo, se encontraron 38 genes significativamente expresados diferencialmente entre las clases de EF. Diferentes clasificadores independientes construidos con estos 38 genes informaron un valor pronóstico moderado (AUC máx. = 0,73), equivalente al obtenido a partir de micromatrices. Curiosamente, el filtrado adicional de estos genes (selección de características sobre la base de correlación) produjo un conjunto de 3 genes con un valor pronóstico más fuerte (AUC = 0,82 utilizando un alumno sobre la base de la instancia, Tabla 1). Estos 3 genes, VEGFB, THBS1 y PGF, se ven en el presente documento como un nuevo "conjunto pronóstico" de biomarcadores de IC. La expresión diferencial de estos 3 genes se validó por PCR cuantitativa. Además, se midieron los niveles plasmáticos de los 3 biomarcadores.

30 En un descubrimiento, se demostró que los pacientes capaces de montar una respuesta significativa a la IM, caracterizado por altos niveles de ARNm o proteína plasmática del factor de crecimiento endotelial vascular B (VEGFB) y bajos niveles plasmáticos de Trombospondina-1 (THBS1) y factor de crecimiento placentario PGF o PIGF), tienen una baja susceptibilidad a desarrollar HF y/o someterse a Remodelación Ventricular. Por lo tanto la medición de los niveles plasmáticos de estos tres biomarcadores puede servir como una herramienta útil para predecir la aparición de HF y/o Remodelado Ventricular después del IM.

35 **Tabla 1. Rendimiento predictivo de los modelos de clasificación basados en los niveles de ARNm de los 3 biomarcadores.**

Tipo de entrada	Modelo de clasificación	Precisión típica* (%)	AUC*
Solo biomarcadores sobre la base de SAM	K*	65	0,63
Biomarcadores sobre la base de la red	K*	84	0,82
Solo biomarcadores sobre la base de SAM	SVM	68	0,68
Biomarcadores sobre la base de la red	SVM	75	0,75
* Sobre la base de la validación cruzada de "dejar uno fuera"			

Biomarcadores potenciales utilizados: derivados de SAM (9 genes) y basados en la red (3 genes). AUC: área bajo la curva (ROC); SVM: Máquina de Vector de Soporte.

40 Por lo tanto, los presentes inventores proponen una nueva estrategia para identificar a los pacientes en riesgo de desarrollar IC después de IM, sobre la base de la medición de un grupo de 3 biomarcadores, VEGFB, THBS1 y PGF. Estas mediciones pueden hacerse a partir de ARN extraído de células sanguíneas o de niveles plasmáticos de las proteínas correspondientes.

La Tabla 2 resume los descubrimientos particulares detrás de la presente invención.

45

**Tabla 2. Asociación entre el ARNm y los niveles plasmáticos de los 3 biomarcadores y la EF.**

	EF alto (resultado favorable)		EF bajo (resultado desfavorable)	
	ARNm	Plasma	ARNm	Plasma
<b>VEGFB</b>	Alto	Alto	Bajo	Bajo
<b>THBS1</b>	Bajo	Bajo	Alto	Alto
<b>PGF</b>	Bajo	Bajo	Alto	Alto

Los niveles de ARNm se miden mediante micromatrices en muestras de sangre cosechadas el día de MI. Los niveles plasmáticos se miden por ELISA un día después de IM.

Todas las referencias citadas en el presente documento se incorporan en su totalidad en el grado en que no estén en conflicto con la presente invención.

- 5 La invención se describirá a continuación con detalle adicional en los ejemplos no limitantes que acompañan.

### Ejemplos

#### Ejemplo 1

#### PACIENTES Y PROCEDIMIENTOS

##### Pacientes

- 10 Los pacientes con IM agudo fueron tratados con intervención coronaria percutánea primaria. El IM agudo se definió por la presencia de dolor torácico < 12 horas con elevación significativa del segmento ST y enzimas cardíacas positivas. Las muestras de sangre se obtuvieron en el momento de la reperfusión mecánica (para micromatrices y análisis cuantitativos de PCR) y el día después del IM (para la determinación de los niveles plasmáticos). Todos los pacientes firmaron un consentimiento informado.

##### Micromatrices

Para aumentar las posibilidades de detectar biomarcadores relevantes en el contexto del remodelado ventricular, se seleccionaron dos grupos de pacientes con fenotipos "extremos" después del IM, a saber, pacientes que evolucionaron favorablemente después del infarto (EF  $\geq$  45 %, promedio 61 %) y pacientes que evolucionaron desfavorablemente (EF  $\leq$  40 %, promedio del 33 %). Cada grupo contenía 16 pacientes.

- 20 El ARN total se extrajo de células de sangre entera mediante la tecnología PAXgene™. La sangre extraída en tubos de ARN de sangre PAXgene™ (PreAnalytix®, BD Europe, Erembodegem, Bélgica) se almacenó a -20 °C hasta la extracción de ARN. La extracción se realizó con el kit PAXgene™ Blood RNA (Qiagen, Courtaboeuf, Francia) de acuerdo con las instrucciones del fabricante. La cantidad de ARN se midió usando el espectrofotómetro ND-1000 (NanoDrop® Technologies, Wilmington, EE.UU.). La calidad del ARN se evaluó utilizando el aparato Bioanalyzer® 2100 (Agilent Technologies, Massy, Francia) con los Nano chips ARN 6000. Solo el ARN de alta calidad (DO<sub>260</sub>/DO<sub>280</sub> > 1,9 y DO<sub>260</sub>/DO<sub>230</sub> > 1,7) y el ARN sin degradar se consideraron para su análisis adicional.

- 30 Un ARN de referencia común (Universal Human Reference RNA, Stratagene Europe, Amsterdam, Países Bajos), se utilizó una mezcla de ARN de 11 líneas celulares junto con ARN del paciente en todas las etapas siguientes para proporcionar un patrón de referencia interno para comparaciones de los niveles de expresión génica a través de matrices.

- Los ARN mensajeros se amplificaron usando el kit Amino Alil MessageAmp™ (Ambion®, Cambridgeshire, Reino Unido) de acuerdo con el protocolo del fabricante, comenzando con un  $\mu$ g de ARN total. Cinco  $\mu$ g de cada amino alil ARN se marcaron con Cy3 o Cy5 (Amersham, Buckinghamshire, Reino Unido). Se midió el acoplamiento de amino alil ARN usando el espectrofotómetro ND-1000 NanoDrop®. El rendimiento de acoplamiento del colorante > 5 % fue un requisito previo para un análisis posterior. Se combinaron 750 ng de cada amino alil ARN marcado Cy3 o Cy5 (ARN de referencia o ARN donante) y se hibridaron en micromatrices de oligonucleótidos pangenómicos que contenían 25.000 genes (Genomic Platform, Illkirch, Francia). Se hibridaron cuatro micromatrices por paciente y se realizó un intercambio de colorantes (2 micromatrices paciente Cy3/referencia Cy5 y 2 micromatrices paciente Cy5/referencia Cy3). Las etapas de hibridación se realizaron utilizando el sistema Agilent Technologies. Brevemente, el ARN se fragmentó con un tampón de fragmentación antes de mezclar con un tampón de hibridación. Los micromatrices se bloquearon con etanolamina 50 mM en tampón de borato 50 mM pH = 9,0. Las cámaras de hibridación de Agilent y el horno de rotación se usaron para la hibridación a 60°C durante 17h a 4 rpm. Los micromatrices se lavaron durante 10 min en SSC 6X, Triton X-102 al 0,005 %, durante 5 min en SSC 0,1X, Triton X-102 al 0,005 % y luego se secaron por centrifugación antes de escanear usando un escáner de micromatrices Axon

4000B y GenePix Pro 6® (Molecular Devices, Berks, Reino Unido). Se permitió el ajuste de la ganancia del fotomultiplicador y 0,1 % de puntos saturados durante el escaneado.

5 La determinación de los puntos y la cuantificación de los datos brutos de los cuatro micromatrices para cada paciente se realizaron en un análisis por lotes usando el software gratuito MAIA® (Institut Curie, Francia). Este software asigna cada punto con nueve parámetros de calidad que permiten la determinación de "puntos de buena  
10 calidad" entre los cuatro micromatrices. Solo se mantienen buenos puntos para un análisis posterior. Se realizó un paso de normalización no lineal Lowess con el software Acuity® (Molecular Devices) para compensar la distribución desigual de Cy3-Cy5. En los pasos siguientes se utilizó la relación logarítmica normalizada de Cy3/Cy5. A continuación, se realizó una etapa de filtrado para retirar los genes que no estaban presentes en al menos tres micromatrices de cuatro. La calidad y reproducibilidad de cada uno de los cuatro micromatrices por paciente se evaluaron mediante ANOVA, coeficientes de correlación y Mapas de Autoorganización elaborados con el software Acuity®. Los datos se almacenan en el administrador de datos de micromatriz sobre la base de la Web MEDIANTE.

15 Antes del análisis estadístico, los genes no presentes en al menos el 50 % de los pacientes fueron filtrados. El análisis supervisado se realizó utilizando dos enfoques complementarios. El primer enfoque involucró el software de Análisis de Significación de Micromatrices (SAM) que correlaciona la expresión génica con una variable externa como el valor EF. Se utilizó una prueba no apareada de dos clases. La imputación de valores faltantes de genes se realizó a través de un algoritmo K-Vecino más cercano utilizando 10 vecinos.

#### Red de interacción proteína-proteína

20 Un conjunto de genes centrales, conocidos por estar asociados a la angiogénesis en el infarto de miocardio, se recuperó de la base de datos Entrez-Gene, Consulta: "humano Y corazón Y angiogénesis Y miocardio E infarto". Las interacciones proteína-proteína anotadas asociadas con estos genes centrales se recuperaron de la base de datos de referencia de proteínas humanas (HPRD).

Se implementó un análisis de agrupación en red para identificar posibles módulos de red funcionales. Los clústeres se identificaron mediante el algoritmo de agrupación de red MCODE (Cytoscape plug-in).

#### Ensayos bioquímicos para THBS1, PGF y NT-pro-BNP

25 Los niveles plasmáticos de THBS1 y PGF se midieron en muestras de 46 pacientes mediante ELISA utilizando los kits Quantikine DTSP10 y DPG00, respectivamente (R&D Systems, Oxon, RU). Los límites de detección de los ensayos fueron 0,35 ng/ml para THBS1 y 7 p/ml para PGF. Se midió el nivel plasmático de pro-BNP (N-terminal-pro-BNP, NT-pro-BNP) utilizando el dispositivo inmunológico Elecsys 2010 (Roche Diagnostics, Meylan, Francia). El límite de detección de los ensayos fue de 20 pg/ml.  
30

#### Montaje del kit de diagnóstico VEGFB

35 Se desarrolló un ELISA sándwich para detectar VEGFB 167 y VEGF-B 186. Se recubren placas de microtitulación (Lumitrac 600, Greiner, Bélgica) con 100 µl de anticuerpo monoclonal anti-VEGF-B de ratón (2 µg/ml en PBS, MAB751, R&D systems, RU) durante la noche a 4 °C. Después de tres lavados, las placas se bloquearon durante 1 hora con 300 µl de BSA-PBS al 5 % a 500 rpm y temperatura ambiente. Se produce una curva patrón de 2000 pg/ml a 15,6 pg/ml con VEGFB 167 humano (751-VE, R&D Systems) en BSA-PBS al 1 %. Después del bloqueo, las placas se lavan tres veces y se incuban durante 2 horas con 100 µl de plasma, blanco o patrones a 500 rpm y temperatura ambiente. Después de tres lavados, se añaden 100 µl del anticuerpo policlonal de VEGF-B de cabra (400 ng/ml en BSA - PBS al 1 %, AF751, R & D Systems) a cada pocillo y las placas se incuban durante 1 hora a 500 rpm y temperatura ambiente. Después de tres lavados, se añaden 100 µl de anticuerpo anti-cabra de burro conjugado con biotina (127500 en BSA-PBS al 1 %, 705-065-147, Jackson, EE.UU.) a cada pocillo y las placas se incuban durante 1 hora a 500 rpm y temperatura ambiente. Después de tres lavados, se añaden 100 µl de estreptavidina conjugada con fosfatasa alcalina (2 µg / ml en BSA-PBS al 1 %, 016-050-084, Jackson) a cada pocillo y las placas se incuban durante 1 hora a 500 rpm y temperatura ambiente. Las placas se lavan cuatro veces con solución salina tamponada con Tris añadida con Tween 20 (pH 7,5) antes de la incubación con 100 µl por pocillo de Lumiphos 530TM (Lumigen, EE.UU.) durante 30 minutos a 500 rpm, a temperatura ambiente y protegidas de la luz. La quimioluminiscencia se detecta usando un Polarstar Optima (BMG Labtech, París, Francia).  
45

#### Análisis estadístico

50 Para cada conjunto de entrada, se evaluaron diferentes clasificadores estadísticos convencionales y de aprendizaje de máquinas, por ejemplo. Análisis de Predicción para Micromatrices (PAM), Máquina de soporte de vector (SVM) y las técnicas K\*. El K\* es un modelo sobre la base de casos que clasifica una nueva muestra sobre la base de la información de clase proporcionada por sus vecinos más relevantes (o más cercanos) en un conjunto de datos de formación. K\* aplica una medida de distancia sobre la base de entropía para estimar el conjunto de vecinos. Se implementaron modelos con combinación global = 20, y curvas de entropía media de columna para estimar valores faltantes.  
55

El filtrado adicional en los datos de genes sobre la base de la red se implementó con el algoritmo de selección de

características sobre la base de correlación (CFS) utilizando la estrategia de "mejor primera búsqueda" (BF) (Figura 3). El CFS es un procedimiento de selección de características de filtro que encuentra subconjuntos de características (es decir, genes) que maximizan la correlación gen-clase mientras minimizan la correlación gen-gen. Los procedimientos de selección de la característica de filtro se implementan independientemente de cualquier modelo de clasificación. La estrategia de BF se basó en una escalada codiciosa de la colina aumentada con un retroceso de subconjunto.

Los resultados de la evaluación de la clasificación se calcularon utilizando la estrategia de validación cruzada "leave-one-out" (LOO), así como la validación cruzada de 10 veces. Se utilizaron las áreas estimadas bajo las curvas (AUC) de la ROC (Curva Característica Operativa del Receptor) validada de forma cruzada para resumir el rendimiento de clasificación estimado de los clasificadores.

Las diferencias estadísticas entre los grupos EF (sobre la base de cada uno de los biomarcadores) se realizaron mediante la prueba t de Student y se corroboraron con pruebas no paramétricas. Las correlaciones entre estos biomarcadores y los valores de EF se estimaron con coeficientes de Pearson convencionales (Tabla 2).

Software

La implementación de los modelos de aprendizaje automático y la evaluación estadística se realizaron con la plataforma de extracción de datos Weka (v. 3.4). El agrupamiento jerárquico se implementó utilizando el procedimiento de grupo de pares no ponderado con promedios aritméticos y coeficientes de correlación. La visualización de clústeres se realizó con GEPAS. Se generaron ensayos de significación estadística, valores de correlación de Pearson y gráficas con el paquete Statistica (v. 6.0).

**RESULTADOS**

1. El gráfico del Análisis estadístico de Micromatrices (SAM) muestra que 525 genes se expresan diferencialmente entre grupos de baja y alta EF (Figura 1). Se seleccionó un umbral para el cambio de 1,3 veces y se obtuvo un FDR del 24,5 %. Los puntos rojos representan genes regulados positivamente en el grupo EF bajo, los genes verdes regulados positivamente en el grupo EF alto y los puntos negros representan genes cuyo plegamiento de veces es < 1,3 entre los dos grupos.

2. Red de interacción proteína-proteína de angiogénesis en el IM humano (Figura 2). La red resultante consistió en 556 nodos (proteínas) y 686 aristas (interacciones). El análisis de agrupación de redes consistentemente destacó la existencia de un clúster de red (53 proteínas) con una representación altamente significativa de los procesos biológicos (Ontología Génica) relacionados con el crecimiento celular y la regulación del crecimiento.

3. La Figura 3 muestra que el análisis combinado de los modelos de clasificación sobre la base de la expresión génica con los modelos de clasificación basados en la red permitió la identificación de 3 genes con mayor valor pronóstico (AUC máx. = 0,82) que los modelos de clasificación basados en la expresión génica sola (0,56<AUC<0,72) o modelos de clasificación basados en la red solamente (0,56<AUC<0,73). Los resultados pronósticos más elevados (sobre la base de los 3 genes) obtenidos hasta la fecha se han obtenido con el modelo de aprendizaje sobre la base de el caso K\*.

4. Un mapa térmico ilustra las diferencias en la expresión (obtenida por micromatrices) de los biomarcadores entre pacientes con fracción de eyección alta (H) y baja (L) (Figura 4). Los colores (rojo, rosa, azul claro, azul oscuro) muestran el intervalo de valores de expresión (alto, moderado, bajo, más bajo). El color blanco indica valores indetectables. VEGFB se expresa más claramente en el grupo de pacientes con EF alto mientras que THBS1 y PGF es más expresados en el grupo EF bajo.

5. Los Gráficos cuantil-cuantil (Figura 5) y parcelas de dispersión (Figura 6) ilustran las dependencias estadísticas entre la fracción de eyección y la expresión de los biomarcadores evaluados por micromatrices y ELISA. La relación lineal mostrada sugiere que estas variables siguen distribuciones de datos similares. Los resultados de la prueba t de Student y la correlación de coeficientes lineales correspondientes a estas gráficas se resumen en la Tabla 3. VEGFB es positivamente correlacionado con la EF, mientras que THBS1 y PGF están correlacionados negativamente con la EF. Se obtienen resultados concordantes con micromatrices y ELISA para THBS1 y PGF. Para VEGFB el nivel de ARNm se encontró significativamente mayor en pacientes con EF bajo que en pacientes con EF bajo la técnica de micromatrices.

**Tabla 3**

Gen	Micromatrices	Proteínas plasmáticas
THBS1	t = 2,4, p = 0,02 r = -0,3, p = 0,2	t = 2,1, p = 0,04 r = -0,2, p = 0,1
PGF	t = 2,8, p = 0,01 r = -0,2, p = 0,2	t = 2,2, p = 0,04 r = -0,2, p = 0,2
VEGFB	t = 3,1, p = 0,004 r = 0,3, p = 0,05	t = -1,3, p = 0,20 r = 0,2, p = 0,3

(continuación)

Gen	Micromatrices	Proteínas plasmáticas
t: Estadística t r: coeficiente de correlación lineal Las estadísticas realizadas sobre los niveles de ARNm medidos por micromatrices en muestras de sangre recolectadas el día de IM y los niveles plasmáticos medidos por ELISA un día después de IM.		

6. Estadística descriptiva de los niveles de ARNm de los 3 biomarcadores en los dos grupos de pacientes con IC que muestran el mayor nivel de ARNm de VEGFB y los niveles más bajos de ARNm de THBS1 y PGF en pacientes con EF alto en comparación con pacientes con EF baja (Tabla 4). También se proporcionan en esta tabla los umbrales teóricos para los niveles de ARNm asociados con una EF alta (> 40 %) o un EF bajo (≤ 40 %) según se determina a partir de las figuras 5 y 6. Por ejemplo, en la población de pacientes con IC utilizada en este estudio, un paciente que tenía un nivel de ARNm de VEGFB superior a -1,4, un nivel de ARNm de THBS1 inferior a 0 y un nivel de ARNm de PGF inferior a -0,1, era más propenso a tener una EF alta. Por el contrario, un paciente que tiene un nivel de ARNm de VEGFB inferior a -1,4, un nivel de ARNm de THBS1 superior a 0 y un nivel de ARNm de PGF superior a -0,1 era más propenso a tener una EF baja. Es importante mencionar que la combinación de los 3 biomarcadores en lugar de cada biomarcador solo o una combinación de 2 biomarcadores se asocia más exactamente a la EF.

**Tabla 4. Estadísticas descriptivas de los niveles de ARNm de los 3 biomarcadores en los dos grupos.**

	Alta EF (Resultado favorable)		Baja EF (Resultado desfavorable)		Estad, T
	Media ± DE (min / máx.)	Límite	Media ± DE (min / máx.)	Límite	
<b>VEGFB</b>	-1,05 ± 0,42 (-1,78 / -0,20)	> -1,4	-1,46 ± 0,32 (-2,07 / -1,09)	< -1,4	t=3,1 p=0,004
<b>THBS1</b>	-0,13 ± 0,28 (-0,57 / 0,33)	< 0	0,29 ± 0,58 (-0,66 / 1,97)	> 0	t=2,4 p=0,02
<b>PGF</b>	-0,19 ± 0,10 (-0,34 / -0,03)	< -0,1	-0,07 ± 0,15 (-0,34 / 0,15)	> -0,1	t=2,8 p=0,01

t: Estadística t  
 p: valor de probabilidad  
 DE : Desviación estándar

7. Pruebas adicionales del clasificador utilizando los niveles plasmáticos de VEGFB, THBS1 y PGF medidos 1 día después de IM en los mismos 32 pacientes que se han utilizado para construir el clasificador produjeron un AUC de 0,75. Esto sugiere que el rendimiento de predicción del clasificador es bueno incluso cuando se utiliza un conjunto de datos diferente que codifica otros tipos de mediciones (por ejemplo, los niveles plasmáticos).

8. Dado que el único biomarcador establecido de HF es pro-BNP, se comparó el rendimiento pronóstico del conjunto de 3 biomarcadores en el presente documento descrito con el rendimiento pronóstico de NT-pro-BNP. El rendimiento pronóstico del nivel plasmático de NT-pro-BNP, medido 1 día después del IM, fue moderado (AUC = 0,63, Tabla 5). Por lo tanto, el conjunto de 3 biomarcadores descritos en el presente documento claramente superó el valor pronóstico de NT-pro-BNP.

**Tabla 5. Rendimiento predictivo de NT-pro-BNP**

Tipo de entrada	Modelo de clasificación	Precisión típica* (%)	AUC*
NT-pro-BNP	K*	50	0.52
NT-pro-BNP	SVM	63	0.63

\* Sobre la base de la validación cruzada leave-one-out.  
 AUC: Área bajo la curva (ROC); SVM: Máquina de Vector de Soporte.

**Ejemplo 2****PACIENTES Y PROCEDIMIENTOS**

5 Los pacientes con IM agudo fueron inscritos en un registro nacional de IM y tratados con intervención coronaria percutánea primaria. El IM agudo se definió por la presencia de dolor torácico < 12 horas con elevación ST significativa y aumento de la creatina quinasa y la troponina I a más de 2 veces el límite superior de la normalidad. Las muestras de sangre se obtuvieron en el momento de la reperfusión mecánica (para ARN y aislamiento del plasma), un día o dos días después del IM (para el plasma). El protocolo ha sido aprobado por el comité de ética local y se ha obtenido el consentimiento informado de todos los sujetos.

10 La cohorte de validación de 290 pacientes con IM fue de un estudio prospectivo realizado en los Hospitales Universitarios de Leicester NHS Trust (Reino Unido). La ecocardiografía se realizó al alta y 6 meses después del infarto. El volumen diastólico final de LV (LVEDV) se estimó utilizando la regla de Simpson modificada biplanar a partir de vistas apicales de dos y cuatro cámaras. El grado de remodelación del VI se evaluó a partir del cambio en LVEDV (DEDV) entre la descarga y el seguimiento.

**Micromatrices**

15 Los perfiles transcriptómicos de células de sangre completa se obtuvieron utilizando micromatrices de oligonucleótidos que representan 25.000 genes. Los datos están disponibles en la base de datos Gene Expression Omnibus ([www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)) bajo el número de registro GSE8723. El análisis supervisado se realizó mediante el software de Análisis de Significación de Micromatrices (SAM). La significación estadística de la sobre-representación de términos de Ontología Génica (GO) en conjuntos de genes se estimó con la base de datos DAVID. Los mapas de calor se obtuvieron utilizando el software Gene Set Enrichment Analysis (GSEA).

**Medida de la expresión de VEGFB**

La expresión de VEGFB ARNm en células sanguíneas obtenidos el día de MI se determinó por PCR cuantitativa. Se preparó un ELISA en sándwich casero para medir los niveles plasmáticos de VEGFB.

**Modelos de clasificación de pacientes**

25 La Máquina de Vector de Soporte (SVM) y los modelos de clasificación computacional de  $K^*$  fueron evaluados para probar la significación pronóstica de los niveles de expresión de VEGFB. El algoritmo SMO (optimización mínima secuencial) para el entrenamiento de los clasificadores SVM se implementó con los siguientes parámetros: parámetro de complejidad  $C = 1,0$ ,  $\epsilon = 1,0E-12$ , exponente del núcleo polinomial = 1,0. Se implementaron modelos con combinación global = 20, y curvas de entropía media de columna para estimar valores faltantes. Los resultados de la evaluación de la clasificación se calcularon mediante la estrategia de dejar de lado una validación cruzada (LOO). El área bajo la curva característica de funcionamiento del receptor (AUC) se utilizó para resumir el rendimiento de clasificación estimado de los clasificadores.

**Análisis estadístico**

35 Las comparaciones entre las medias de dos grupos de pacientes se realizaron con prueba de t no apareada de dos colas para los datos gaussianos y la prueba de Mann-Whitney para los datos no gaussianos. Las variables categóricas se compararon mediante la prueba exacta de Fisher. La correlación entre los niveles de biomarcadores y la clase EF se estimó con la prueba de Spearman. La implementación de los modelos de aprendizaje automático se realizó con la plataforma de minería de datos Weka (v. 3.4). El agrupamiento jerárquico se implementó utilizando el procedimiento de grupo de pares no ponderado con promedios aritméticos y coeficientes de correlación. La visualización de clústeres se realizó con GEPAS. Se generaron pruebas de significación estadística con el paquete Estadística (v. 6,0). Se consideró estadísticamente significativo un valor de  $P < 0,05$ .

**Resultados****Selección y características del paciente**

45 Se reclutaron pacientes que presentaban IM con elevación de ST aguda, tratados con reperfusión mecánica. Para los análisis transcriptómicos, dos grupos de 16 pacientes con IM agudo fueron seleccionados sobre la base de su EF 1 mes después de MI. Un grupo de pacientes tenía una función sistólica del VI mantenida con alta FE después de IM ( $> 40\%$ , mediana  $63\%$ , intervalo 45-73), y el otro grupo con función FE baja ( $\leq 40\%$ , mediana  $35\%$  20-40).

**Análisis transcriptómico de las células sanguíneas**

50 Los perfiles de expresión génica de células sanguíneas totales aisladas en el momento de la reperfusión se obtuvieron utilizando micromatrices de 25.000 genes. Entre éstos, 525 genes se encontraron diferencialmente expresado por SAM entre EF alta y pacientes con baja EF de 1,3 veces el cambio de umbral y una tasa de falso descubrimiento de 24,5 %. 226 genes fueron regulados en el grupo de alto EF y 299 fueron regulados en el grupo de baja EF. De los 525 genes, GSEA recuperó los 50 genes más significativamente asociados a uno u otro grupo de

pacientes.

**Genes angiogénicos asociados al resultado clínico después del IM**

5 Siguiendo la presente hipótesis de trabajo que la angiogénesis puede desempeñar un papel significativo en la reparación cardíaca después de IM, el objeto fue identificar, a partir de los 525 genes expresados diferencialmente en pacientes con alta y baja EF, los genes relacionados con la angiogénesis. Para este propósito, se recuperaron de la base de datos Entrez Gene una lista de 494 genes que se sabe que están relacionados con la angiogénesis en los seres humanos con la siguiente consulta: "angiogénesis" Y "homo sapiens". De los 525 genes expresados diferencialmente, 28 se encontraron en este conjunto de 494 genes angiogénicos: hasta 20 regulados positivamente en el grupo de baja EF y hasta 8 regulados positivamente en el grupo de alta EF (*Tabla 6*).

10 **Tabla 6.** Lista de 28 genes angiogénicos expresados diferencialmente entre grupos EF alta y baja.

<b>Sobre-expresados en grupo EF baja</b>						
Pro angiogénico (n=15)	Cambio en veces	q-valor (%)		Anti angiogénico (n=6)	Cambio en veces	q-valor (%)
BMX	1,90	11,36		CLU	1,50	4,97
PBEF1	1,73	20,21		THBS1	1,44	5,88
FOS	1,66	4,09		ITGB1	1,34	25,29
PFKFB3	1,65	0,00		MAPK14	1,31	24,72
CD55	1,65	5,53		STAT1	1,31	25,56
HIF1A	1,63	0,00		MME	1,31	20,21
IL8	1,59	7,51				
PTGS2	1,55	11,36				
TGFBR1	1,50	9,87				
THBS1	1,44	5,88				
SLC2A3	1,40	0,00				
ERO1L	1,32	28,34				
PLAUR	1,31	5,88				
ADM	1,31	25,06				
B2M	1,30	28,12				
<b>Bajo-expresados en grupo EF baja</b>						
Pro angiogénico (n=4)	Cambio en veces	q-valor (%)		Anti angiogénico (n=4)	Cambio en veces	q-valor (%)
VEGFB	0,74	4,97		SOD1	0,70	11,36
RHOC	0,74	8,70		MAGED1	0,75	9,31
CX3CR1	0,76	17,27		ANXA2	0,76	17,27
ATP5B	0,80	21,32		BAI1	0,76	17,27

**q valor:** la tasa más baja de descubrimiento falso a la que el gen se denomina significativo (como 'valor p' adaptado al análisis de un gran número de genes). **Nota:** el número total de genes en esta tabla es 28 puesto que la trombospondina 1 (THBS1) es tanto pro como antiangiogénica,

Un mapa de calor elaborado con estos 28 genes mostró que el resultado clínico después del IM se asocia con una biofirma distinta vinculada a la angiogénesis.

15 En un intento de evaluar si esta biofirma se asocia a la estimulación o la represión de la angiogénesis, se consultó la base de datos Entrez Gene para las propiedades pro o anti angiogénicas de los 28 genes angiogénicos encontrados expresados diferencialmente entre los 2 grupos de pacientes con IM. Como se muestra en la *Tabla 6*, el equilibrio entre los factores conocidos pro y anti angiogénicos tiende a inclinarse hacia el lado pro angiogénico para los pacientes con EF baja, aunque solo se puede especular que esto está asociado con la estimulación de la angiogénesis.

20 Después los presentes inventores estrecharon sus investigaciones sobre VEGFB porque: (1) entre los 4 genes pro angiogénicos sobre expresados en el grupo EF alta (y, por tanto, potencialmente implicados en la remodelación favorable del corazón), solo VEGFB fue recuperado por la base datos Entrez Gene utilizando la consulta: "angiogénesis Y homo sapiens Y corazón"; y (2) la diferencia entre la expresión de VEGFB en pacientes EF altos y bajos fue la más significativa entre los genes pro-angiogénicos (*Tabla 6*).

**La expresión de VEGFB se correlaciona con el resultado después de IM**

La PCR cuantitativa se utilizó para confirmar los datos de micromatrices sobre VEGFB. Los valores de expresión entre los 2 grupos de 16 pacientes con IM se compararon entre los micromatrices y la PCR cuantitativa. Ambas técnicas informaron niveles más altos de expresión de ARNm de VEGFB en pacientes con EF alta en comparación con pacientes con EF baja: 1,3 veces (t = 3,35; P = 0,004) para micromatrices y 1,7 veces (t = 3,35; P = 0,003) para PCR cuantitativa (Figura 8A). La Figura 8B muestra la correlación significativa observada entre la expresión de VEGFB y la EF (r = 0,39; P = 0,03). Por lo tanto, la expresión de VEGFB en las células sanguíneas parece estar correlacionada con el resultado después del IM.

**Los niveles plasmáticos de VEGFB se asocian al resultado clínico después del IM**

Después los presentes inventores midieron la VEGFB en el plasma sanguíneo de 140 pacientes con IM, separados en 2 grupos, es decir, aquellos con función VE preservada (mediana del 57 %, intervalo 45-89) y con deterioro (mediana 37 %, intervalo 17-44). El muestreo de sangre se realizó el día del IM (n = 77), 1 día después del IM (n = 65) o 2 días después del IM (n = 12). Los niveles plasmáticos de VEGFB fueron similares entre los grupos EF alto y bajo en el día 0 y el día 1 después del MI. Sin embargo, en el día 2, los niveles de VEGFB aumentaron en los pacientes con alta EF (2 veces en comparación con el día 0), mientras que disminuyeron en los pacientes con EF bajo (2,5 veces en comparación con el día 0) (Figura 9). Estos datos están de acuerdo con la regulación ascendente de ARNm de VEGFB en pacientes con alta EF (Figura 8). Estos resultados sugieren que los pacientes capaces de aumentar su producción de VEGFB después del IM son más propensos a tener un resultado favorable.

**Validación independiente**

Se utilizó una cohorte independiente de 290 pacientes con IM para estudiar más a fondo la asociación entre los niveles plasmáticos de VEGFB y el remodelado del VI después del IM. Se han publicado las características clínicas de esta población de pacientes. El VEGFB plasmático medio fue 64 % mayor (U Estadística = 8128, P <0,001) en pacientes en los que ΔEDV (n = 138) mostró una caída durante este período (n = 138), en comparación con los pacientes en los que ΔEDV aumento (n= 152). Estos datos confirman nuestra observación de que el VEGFB está asociado con el remodelado del VI después del MI.

**Rendimiento pronóstico del VEGFB**

Los presentes resultados sugieren que VEGFB podría representar un biomarcador potencial de la remodelación después del IM. Dos modelos de aprendizaje de la máquina construidos con varios conjuntos de datos obtenidos, ya sea por micromatrices, PCR cuantitativa o determinación de plasma, se utilizaron para probar la eficiencia pronóstica de VEGFB. Los resultados se muestran en la Tabla 7. El mejor rendimiento se logró cuando el modelo de aprendizaje sobre la base de el caso K\* se construyó con los niveles de expresión de VEGFB medidos en células sanguíneas por micromatrices de los 32 pacientes de la cohorte de prueba. Este modelo alcanzó una especificidad del 75 % (12 de 16 pacientes con EF bajo correctamente clasificado), una sensibilidad del 50 % (8 de 16 pacientes EF altos clasificados correctamente) y una precisión general del 62 % (20 de 32 pacientes correctamente clasificados). La AUC fue de 0,75. Cuando se construyó con los niveles plasmáticos de VEGFB de la cohorte de validación (290 pacientes), la significación pronóstica máxima proporcionó un AUC de 0,52.

**Tabla 7.** Rendimientos de predicción de VEGFB.

		n	Modelo de clasificación	AUC	Especificidad (%)	Sensibilidad (%)	Precisión (%)
VEGFB (micromatrices)	Día 0	32	SVM	0,56	56	56	56
			K*	0,75	75	50	62
VEGFB (PCR)	Día 0	32	SVM	0,68	94	44	69
			K*	0,68	56	56	56
	Día 0	77	SVM	0,5	0	1	66
			K*	0,51	0	92	60
VEGFB (plasma)	Día 1	65	SVM	0,5	0	1	66
			K*	0,47	0	87	57
	Día 2	12	SVM	0,5	0	1	66
			K*	0,07	0	1	66

AUC: Área bajo la curva Característica de funcionamiento de receptor (ROC).

La especificidad indica el porcentaje de pacientes correctamente clasificados como EF baja; la sensibilidad indica el porcentaje de pacientes correctamente clasificados con alta EF; precisión indica el porcentaje de pacientes correctamente clasificados como EF alta y baja.

LISTADO DE SECUENCIAS

<110> Centre de Recherche Public de la Santé

<120> BIOMARCADORES

<130> 08171-2.WO

<160> 13

<170> PatentIn versión 3.3

<210> 1

<211> 1172

<212> ADN

<213> *Homo sapiens*

<400> 1

```

gcgatgcggg cgccccggc gggcggcccc ggcgggcacc atgagccctc tgctccgccc      60
cctgctgctc gccgcactcc tgcagctggc ccccgcccag gccctgtct cccagcctga      120
tgcccctggc caccagagga aagtgggtgc atggatagat gtgtatactc gcgctacctg      180
ccagccccgg gaggtgggtg tgccctgac tgtggagctc atgggcacog tggccaaaca      240
gctgggtgcc agctgcgtga ctgtgcagcg ctgtgggtggc tgctgccctg acgatggcct      300
ggagtgtgtg cccactgggc agcaaccaagt cgggatgcag atcctcatga tccggtaacc      360
gagcagtcag ctgggggaga tgtccctgga agaacacagc cagtgtgaat gcagacctaa      420
aaaaaaggac agtgctgtga agccagacag ggctgccact ccccaccacc gtccccagcc      480
ccgttctgtt ccgggctggg actctgcccc cggagcacc tccccagctg acatcaccca      540
tcccactcca gccccaggcc cctctgcca cgctgcacc agcaccacca ggcgcctgac      600
ccccggacct gccgctgccg ctgccgacgc cgcagcttc tccgttgcca agggcggggc      660
ttagagctca acccagacac ctgcaggtgc cggaaactgc gaaggtgaca catggctttt      720
cagactcagc agggtgactt gcctcagagg ctatatccca gtgggggaac aaagaggagc      780
ctggtaaaaa acagccaagc cccaagacc tcagcccagg cagaagctgc tctaggacct      840
gggcctctca gagggtcttt ctgccatccc ttgtctcctt gaggccatca tcaaacagga      900
cagagttgga agaggagact gggaggcagc aagaggggtc acataccagc tcaggggaga      960
atggagtact gtctcagttt ctaaccactc tgtgcaagta agcatcttac aactggctct     1020
tcctccctc actaagaaga cccaaacctc tgcataatgg gatttgggct ttggtacaag     1080
aactgtgacc cccaaccctg ataaaagaga tggaaggaaa aaaaaaaaaa aaaaaaaaaa     1140
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa                                     1172

```

<210> 2

<211> 624

<212> ADN

<213> *Homo sapiens*

<400> 2

ES 2 640 900 T3

```

atgagccctc tgcctcgcgc cctgctgctc gccgcactcc tgcagctggc ccccgcccag      60
gccoctgtct cccagcctga tgccctggc caccagagga aagtgggtgc atggatagat      120
gtgtatactc gcgctacctg ccagccccgg gaggtgggtgg tgccttgac tgtggagctc      180
atgggcaccg tggccaaaca gctggtgccc agctgcgtga ctgtgcagcg ctgtgggtggc      240
tgetgccctg acgatggcct ggagtgtgtg cccactgggc agcaccaagt cgggatgcag      300
atcctcatga tcoggtaccc gagcagtcag ctgggggaga tgtccctgga agaacacagc      360
cagtgtgaat gcagacctaa aaaaaggac agtgctgtga agccagacag ggctgccact      420
ccccaccacc gtcccagcc ccgttctgtt ccgggctggg actctgccc cggagcaccc      480
tcccagctg acatcaccca tcccactcca gcccaggcc cctctgcca cgctgcaccc      540
agcaccacca gcgccctgac ccccgacct gccgctgccg ctgcgcagc cgcagcttcc      600
tccgttgcca agggcggggc ttag                                             624

```

<210> 3  
 <211> 207  
 <212> PRT  
 <213> *Homo sapiens*

<400> 3

```

Met Ser Pro Leu Leu Arg Arg Leu Leu Leu Ala Ala Leu Leu Gln Leu
 1          5          10          15

Ala Pro Ala Gln Ala Pro Val Ser Gln Pro Asp Ala Pro Gly His Gln
 20          25          30

Arg Lys Val Val Ser Trp Ile Asp Val Tyr Thr Arg Ala Thr Cys Gln
 35          40          45

Pro Arg Glu Val Val Val Pro Leu Thr Val Glu Leu Met Gly Thr Val
 50          55          60

Ala Lys Gln Leu Val Pro Ser Cys Val Thr Val Gln Arg Cys Gly Gly
 65          70          75          80

Cys Cys Pro Asp Asp Gly Leu Glu Cys Val Pro Thr Gly Gln His Gln
 85          90          95

Val Arg Met Gln Ile Leu Met Ile Arg Tyr Pro Ser Ser Gln Leu Gly
 100         105         110

Glu Met Ser Leu Glu Glu His Ser Gln Cys Glu Cys Arg Pro Lys Lys
 115        120        125

```

ES 2 640 900 T3

Lys Asp Ser Ala Val Lys Pro Asp Arg Ala Ala Thr Pro His His Arg  
 130 135 140

Pro Gln Pro Arg Ser Val Pro Gly Trp Asp Ser Ala Pro Gly Ala Pro  
 145 150 155 160

Ser Pro Ala Asp Ile Thr His Pro Thr Pro Ala Pro Gly Pro Ser Ala  
 165 170 175

His Ala Ala Pro Ser Thr Thr Ser Ala Leu Thr Pro Gly Pro Ala Ala  
 180 185 190

Ala Ala Ala Asp Ala Ala Ala Ser Ser Val Ala Lys Gly Gly Ala  
 195 200 205

<210> 4  
 <211> 5820  
 <212> ADN  
 <213> *Homo sapiens*

5

<400> 4

agccgctgcg cccgagctgg cctgcgagtt cagggctcct gtcgctctcc aggagcaacc 60  
 tctactccgg acgcacaggc attccccgcg cccctccagc cctcgccgcc ctcgccaccg 120  
 ctcccggccg ccgcgctccg gtacacacag gatccctgct gggcaccaac agctocacca 180  
 tggggctggc ctggggacta ggcgtcctgt tctgatgca tgtgtgtggc accaacggca 240  
 ttccagagtc tggcggagac aacagcgtgt ttgacatctt tgaactcacc ggggccgccc 300  
 gcaaggggtc tgggcgccga ctggtgaagg gccccgacc ttccagccca gtttccgca 360  
 tcgaggatgc caacctgatc ccccctgtgc ctgatgaaa gttccaagac ctggtggatg 420  
 ctgtgcgggc agaaaagggt ttctctcttc tggcatcctt gaggcagatg aagaagacc 480  
 ggggacgct gctggcctg gagcggaaag accactctgg ccaggtcttc agcgtggtgt 540  
 ccaatggcaa ggcgggcaac ctggacctca gcctgaccgt ccaaggaaag cagcacgtgg 600  
 tgtctgtgga agaagctctc ctggcaaccg gccagtggaa gagcatcacc ctgtttgtgc 660  
 aggaagacag ggcccagctg tacatcgact gtgaaaagat ggagaatgct gagttggacg 720  
 tccccatcca aagcgtcttc accagagacc tggccagcat cgccagactc cgcacgcaa 780  
 agggggcgct caatgacaat ttccaggggg tgctgcagaa tgtgaggttt gtctttggaa 840  
 ccacaccaga agacatctc aggaacaaag gctgctccag ctctaccagt gtctctctca 900  
 cccttgaaa caacgtggtg aatggttoca gcctgccat ccgcactaac tacattggcc 960  
 acaagacaaa ggacttgcaa gccatctgcg gcatctctg tgatgagctg tccagcatgg 1020  
 tcttggact caggggcctg cgcaccattg tgaccagct gcaggacagc atccgcaaag 1080

10

ES 2 640 900 T3

tgactgaaga gaacaaagag ttggccaatg agctgaggcg gcctccccta tgctatcaca 1140  
 acggagttca gtacagaaat aacgaggaat ggactgttga tagctgcact gagtgtcact 1200  
 gtcagaactc agttaccatc tgcaaaaagg tgtcctgccc catcatgccc tgctccaatg 1260  
 ccacagttcc tgatggagaa tgctgtcctc gctggtggcc cagcgactct gcggacgatg 1320  
 gctggtctcc atgggtccgag tggacctcct gttctacgag ctgtggcaat ggaattcagc 1380  
 agcgcggcgg ctccctgcgat agcctcaaca accgatgtga gggctcctcg gtccagacac 1440  
 ggacctgcca cattcaggag tgtgacaaga gatttaaaca ggatggtggc tggagccact 1500  
 ggtccccgtg gtcattctgt tctgtgacat gtggtgatgg tgtgatcaca aggatccggc 1560  
 tctgcaactc tcccagcccc cagatgaacg ggaaaccctg tgaaggcgaa gcgcgggaga 1620  
 ccaaagcctg caagaaagac gcctgcccga tcaatggagg ctggggctcct tggtcacat 1680  
 gggacatctg ttctgtcacc tgtggaggag ggttacagaa acgtagtctg ctctgcaaca 1740  
 accccacacc ccagtttggg ggcaaggact gcgttgggtg tghtaacagaa aaccagatct 1800  
 gcaacaagca ggactgtcca attgatggat gcctgtccaa tccctgcttt gccggcgtga 1860  
 agtgtactag ctaccctgat ggcagctgga aatgtggtgc ttgtccccct ggttacagtg 1920  
 gaaatggcat ccagtgcaca gatgttgatg agtgcaaaga agtgccctgat gcctgcttca 1980  
 accacaatgg agagcaccgg tgtgagaaca cggacccccg ctacaactgc ctgccctgcc 2040  
 ccccacgctt caccggctca cagcccttcg gccagggtgt cgaacatgcc acggccaaca 2100  
 aacaggtgtg caagccccgt aaccctgca cggatgggac ccacgactgc aacaagaacg 2160  
 ccaagtgcaa ctacctgggc cactatagcg accccatgta ccgctgcgag tgcaagcctg 2220  
 gctacgctgg caatggcatc atctgcgggg aggacacaga cctggatggc tggcccaatg 2280  
 agaacctggt gtgcggtggc aatgcgactt accactgcaa aaaggataat tgcccccaacc 2340  
 ttcccaactc agggcaggaa gactatgaca aggatggaat tggatgatgcc tgtgatgatg 2400  
 acgatgacaa tgataaaatt ccagatgaca gggacaactg tccattccat tacaaccag 2460  
 ctcagtatga ctatgacaga gatgatgtgg gagaccgctg tgacaactgt ccctacaacc 2520  
 acaaccaga tcaggcagac acagacaaca atggggaagg agacgcctgt gctgcagaca 2580  
 ttgatggaga oggtatcctc aatgaacggg acaactgcca gtaogtctac aatgtggacc 2640  
 agagagacac tgatatggat ggggttggag atcagtgatga caattgcccc ttggaacaca 2700  
 atccggatca gctggactct gactcagacc gcattggaga tacctgtgac aacaatcagg 2760  
 atattgatga agatggccac cagaacaatc tggacaactg tccctatgtg cccaatgcca 2820  
 accaggctga ccatgacaaa gatggcaagg gagatgcctg tgaccacgat gatgacaacg 2880  
 atggcattcc tgatgacaag gacaactgca gactcgtgcc caatcccagc cagaaggact 2940  
 ctgacggcga tggtcgaggt gatgcctgca aagatgattt tgaccatgac agtgtgccag 3000

ES 2 640 900 T3

acatcgatga catctgtcct gagaatggtg acatcagtga gaccgatttc cgccgattcc 3060  
 agatgattcc tctggacccc aaagggacat cccaaaatga ccoctaactgg gttgtacgcc 3120  
 atcagggtaa agaactcgtc cagactgtca actgtgatcc tggactcgtc gtaggttatg 3180  
 atgagtttaa tgotgtggac ttcagtggca ccttcttcat caacaccgaa agggacgatg 3240  
 actatgctgg atttgtcttt ggctaaccagt ccagcagccg cttttatggt gtgatgtgga 3300  
 agcaagtcac ccagtcctac tgggacacca accccacgag ggctcagga tactcgggcc 3360  
 tttctgtgaa agttgtaaac tccaccacag ggcctggcga gcacctgagg aacgcctgt 3420  
 ggcacacagg aaacaccctt gccaggtgc gcacctgtg gcatgacct cgtcacatag 3480  
 gctggaaga tttcacgcc tacagatggc gtctcagcca caggccaaag acgggtttca 3540  
 ttagagtggg gatgtatgaa ggaagaaaa tcatggotga ctcaggacc atctatgata 3600  
 aaacctatgc tgggtgtaga ctagggttgt ttgtcttctc tcaagaaatg gtgttcttct 3660  
 ctgacctgaa atacgaatgt agagatccct aatcatcaa ttgttgattg aaagactgat 3720  
 cataaaccaa tgcctgtatt gcacctctg gaactatggg cttgagaaaa ccccaggat 3780  
 cacttctctt tggcttctt cttttctgtg cttgcatcag tgtggactcc tagaacgtgc 3840  
 gacctgctc aagaaaatgc agttttcaa aacagactca gcattcagcc tccaatgaat 3900  
 aagacatctt ccaagcatat aaacaattgc ttgggttcc ttttgaaaa gcctctactt 3960  
 gcttcagttg ggaaggtgcc cattcoactc tgcctttgtc acagagcagg gtgctattgt 4020  
 gaggccatct ctgagcagtg gactcaaaag cattttoagg catgtcagag aagggaggac 4080  
 tcaactagaat tagcaaaaa aaccaacctg acatcctcct tcaggaacac ggggagcaga 4140  
 ggccaaagca ctaaggggag ggcgcatacc cgagacgatt gtatgaagaa aatatggagg 4200  
 aactgttaca tgttcggtac taagtcattt tcaggggatt gaaagactat tgcctgattt 4260  
 catgatgctg actggcgta gctgattaac ccatgtaaat aggcaactaa atagaagcag 4320  
 gaaagggaga caaagactgg cttctggact tctcctcga tccccacct tactcatcac 4380  
 ctgcagtggc cagaattagg gaatcagaat caaacagtg taaggcagtg ctggctgcca 4440  
 ttgctggctc acattgaaat tgggtgcttc attctagatg tagcttgtgc agatgtagca 4500  
 ggaaaatagg aaaacctacc atctcagtga gcaccagctg cctcccaaag gagggcagc 4560  
 cgtgcttata tttttatggt tacaatggca caaaattatt atcaacctaa ctaaacatt 4620  
 cctttctctt ttttctcga attatcatgg agttttotaa ttctctctt tggaaatgag 4680  
 atttttttta aatgctttac gatgtaaaat atttatttt tacttattct ggaagactctg 4740  
 gctgaaggat tattcatgga acaggaagaa gcgtaaagac tatccatgct atctttgttg 4800  
 agagtcttcg tgactgtaag attgtaaaata cagattattt attaactctg ttctgctcgg 4860

ES 2 640 900 T3

aaatttaggc ttcatacggg aagtgtttga gagcaagtag ttgacattta tcagcaaatc 4920  
 tcttgcaaga acagcacaag gaaaatcagt ctaataagct gctctgcccc ttgtgctcag 4980  
 agtggatggt atgggattct ttttttctct gttttatctt ttcaagtggg attagttggt 5040  
 tatccatttg caaatgtttt aaattgcaaa gaaagccatg aggtcttcaa tactgtttta 5100  
 ccccatccct tgtgcatatt tocagggaga aggaaagcat atacactttt ttctttcatt 5160  
 tttccaaaag agaaaaaat gacaaaaggt gaaacttaca tacaaatatt acctcatttg 5220  
 ttgtgtgact gagtaagaa tttttggatc aagcggaaag agtttaagtg tctaacaac 5280  
 ttaaagctac tgtagtacct aaaaagtcag tgtgtacat agcataaaaa ctctgcagag 5340  
 aagtattccc aataaggaaa tagcattgaa atgttaaata caatttctga aagttatggt 5400  
 ttttttctat catctggtat accattgctt ttttttata aattattttc tcattgocat 5460  
 tggaatagat atctcagatt gtgtagatat gctatttaaa taatttatca ggaaatactg 5520  
 cctgtagagt tagtatttct atttttatat aatgtttgca cactgaattg aagaattggt 5580  
 ggttttttct tttttttggt ttgttttttt tttttttttt ttttgctttt gacctccat 5640  
 ttttactatt tgccaatacc tttttctagg aatgtgcttt tttttgtaca catttttctc 5700  
 cattttacat tctaaagcag tgtaagttgt atattactgt ttcttatgta caaggaacaa 5760  
 caataaatca tatggaaatt tatatttata aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5820

<210> 5  
 <211> 3513  
 <212> ADN  
 <213> *Homo sapiens*

5

<400> 5

atggggctgg cctggggact aggcgtcctg ttctctgatgc atgtgtgtgg caccaaccgc 60  
 attccagagt ctggcggaga caacagcgtg tttgacatct ttgaactcac cggggcogcc 120  
 cgcaaggggt ctgggcgccg actggtgaag ggccccgacc cttccagccc agctttccgc 180  
 atcgaggatg ccaacctgat cccccctgtg cctgatgaca agttccaaga cctggtggat 240  
 gctgtgcggg cagaaaaggg tttcctcctt ctggcatccc tgaggcagat gaagaagacc 300  
 cggggcacgc tgctggocct ggagcggaaa gaccactctg gccaggtctt cagcgtggtg 360  
 tccaatggca aggcgggcac cctggacctc agcctgaccg tccaaggaaa gcagcacgtg 420  
 gtgtctgtgg aagaagctct cctggcaacc ggccagtggg agagcatcac cctgtttgtg 480  
 caggaagaca gggcccagct gtacatcgac tgtgaaaaga tggagaatgc tgagttggac 540  
 gtcccatcc aaagcgtctt caccagagac ctggccagca tcgccagact cgcacatcga 600  
 aaggggggcg tcaatgacaa tttccagggg gtgctgcaga atgtgaggtt tgtctttgga 660  
 accacaccag aagacatcct caggaacaaa ggctgctcca gctctaccag tgtcctctc 720

10

ES 2 640 900 T3

accottgaca acaacgtggt gaatggttcc agccctgcc tccgcactaa ctacattggc 780  
 cacaagacaa aggacttgca agccatctgc ggcattctct gtgatgagct gtccagcatg 840  
 gtccctggaac tcaggggctt ggcaccatt gtgaccacgc tgcaggacag catccgcaaa 900  
 gtgactgaag agaacaaaga gttggccaat gagctgaggg gccctcccct atgctatcac 960  
 aacggagttc agtacagaaa taacgaggaa tggactggtg atagctgcac tgagtgtcac 1020  
 tgtcagaact cagttacat ctgcaaaaag gtgtcctgcc ccatcatgcc ctgctccaat 1080  
 gccacagttc ctgatggaga atgtgtctct cgtgttggc ccagcgactc tgcggacgat 1140  
 ggctgggtct catgttccga gtggacctcc tgttctacga gctgtggcaa tggaaattcag 1200  
 cagcgcggcc gctcctgcga tagcctcaac aaccgatgtg agggctcctc ggtccagaca 1260  
 cggacctgcc acattcagga gtgtgacaag agatttaaac aggatggtgg ctggagccac 1320  
 tggccccgtt ggtcatcttg ttctgtgaca tgtgtgatg gtgtgatcac aaggatccgg 1380  
 ctctgcaact ctcccagccc ccagatgaac gggaaaccct gtgaaggcga agcgcggggag 1440  
 accaaagcct gcaagaaaga cgcctgcccc atcaatggag gctgggggtcc ttggtcacca 1500  
 tgggacatct gttctgtcac ctgtggagga ggggtacaga aacgtagtgc tctctgcaac 1560  
 aacccacac cccagtttgg aggcaaggac tgcgttgggt atgtaacaga aaaccagatc 1620  
 tgcaacaagc aggactgtcc aattgatgga tgcctgtcca atccctgctt tgccggcgtg 1680  
 aagtgtacta gctaccctga tggcagctgg aaatgtgggt cttgtcccc ttggttacagt 1740  
 ggaaatggca tccagtgcac agatgttgat gactgcaaag aagtgcctga tgcctgcttc 1800  
 aaccacaatg gagagcaccg gtgtgagaac acggaccccc gctacaactg cctgccctgc 1860  
 cccccacgtt tcaccggctc acagccctc ggcagggtg tcgaacatgc cacggccaac 1920  
 aacaggtgt gcaagccccg taaccctgc acggatggga cccacgactg caacaagaac 1980  
 gccaaagtga actacctggg cactatagc gaccccatgt accgctgcga gtgcaagcct 2040  
 ggctacgctg gcaatggcat catctgctgg gaggacacag acctggatgg ctggcccaat 2100  
 gagaacctgg tgtgcgtggc caatgcgact taccactgca aaaaggataa ttgccccaac 2160  
 cttcccaact cagggcagga agactatgac aaggatggaa ttggtgatgc ctgtgatgat 2220  
 gacgatgaca atgataaaat tccagatgac agggacaact gtccattcca ttacaacca 2280  
 gctcagtatg actatgacag agatgatgtg ggagaccgct gtgacaactg tccttacaac 2340  
 cacaacccag atcaggcaga cacagacaac aatggggaag gagacgctg tgctgcagac 2400  
 attgatggag acggtatcct caatgaacgg gacaactgcc agtacgtcta caatgtggac 2460  
 cagagagaca ctgatatgga tggggttggg gatcagtggt acaattgccc ctggaacac 2520  
 aatccggatc agctggactc tgactcagac cgcattggag atacctgtga caacaatcag 2580  
 gatattgatg aagatggcca ccagaacaat ctggacaact gtcctatgt gcccaatgcc 2640

ES 2 640 900 T3

aaccaggctg accatgacaa agatggcaag ggagatgcct gtgaccacga tgatgacaac 2700  
 gatggcattc ctgatgacaa ggacaactgc agactcgtgc ccaatcccga ccagaaggac 2760  
 tctgacggcg atggctgagg tgatgcctgc aaagatgatt ttgaccatga cagtgtgcca 2820  
 gacatcgatg acatctgtcc tgagaatggt gacatcagtg agaccgattt cggccgatcc 2880  
 cagatgattc ctctggaccc caaagggaca tccc aaaatg accctaactg ggttgtacgc 2940  
 catcagggtg aagaactcgt ccagactgtc aactgtgacg ctggactcgc tgtaggttat 3000  
 gatgagttta atgctgtgga cttcagtggc accttcttca tcaacaccga aagggacgat 3060  
 gactatgctg gatttgtctt tggctaccag tccagcagcc gcttttatgt tgtgatgtgg 3120  
 aagcaagtca cccagtcta ctgggacacc aaccccacga gggctcaggg atactcgggc 3180  
 ctttctgtga aagttgtaa ctccaccaca gggcctggcg agcacctgcg gaacgccctg 3240  
 tggcacacag gaaacacccc tggccagggtg cgcaccctgt ggcatgaccc tcgtcacata 3300  
 ggctggaaag atttcaccgc ctacagatgg cgtctcagcc acaggccaaa gacgggttcc 3360  
 attagagtgg tgatgtatga agggaagaaa atcatggctg actcaggacc catctatgat 3420  
 aaaacctatg ctggtggtag actaggggtg tttgtcttct ctcaagaaat ggtgttcttc 3480  
 tctgacctga aatacgaatg tagagatccc taa 3513

<210> 6  
 <211> 1170  
 <212> PRT  
 <213> *Homo sapiens*  
 <400> 6

5

Met Gly Leu Ala Trp Gly Leu Gly Val Leu Phe Leu Met His Val Cys  
 1 5 10 15  
 Gly Thr Asn Arg Ile Pro Glu Ser Gly Gly Asp Asn Ser Val Phe Asp  
 20 25 30  
 Ile Phe Glu Leu Thr Gly Ala Ala Arg Lys Gly Ser Gly Arg Arg Leu  
 35 40 45  
 Val Lys Gly Pro Asp Pro Ser Ser Pro Ala Phe Arg Ile Glu Asp Ala  
 50 55 60  
 Asn Leu Ile Pro Pro Val Pro Asp Asp Lys Phe Gln Asp Leu Val Asp  
 65 70 75 80  
 Ala Val Arg Ala Glu Lys Gly Phe Leu Leu Leu Ala Ser Leu Arg Gln  
 85 90 95

10

ES 2 640 900 T3

Met Lys Lys Thr Arg Gly Thr Leu Leu Ala Leu Glu Arg Lys Asp His  
100 105 110

Ser Gly Gln Val Phe Ser Val Val Ser Asn Gly Lys Ala Gly Thr Leu  
115 120 125

Asp Leu Ser Leu Thr Val Gln Gly Lys Gln His Val Val Ser Val Glu  
130 135 140

Glu Ala Leu Leu Ala Thr Gly Gln Trp Lys Ser Ile Thr Leu Phe Val  
145 150 155 160

Gln Glu Asp Arg Ala Gln Leu Tyr Ile Asp Cys Glu Lys Met Glu Asn  
165 170 175

Ala Glu Leu Asp Val Pro Ile Gln Ser Val Phe Thr Arg Asp Leu Ala  
180 185 190

Ser Ile Ala Arg Leu Arg Ile Ala Lys Gly Gly Val Asn Asp Asn Phe  
195 200 205

Gln Gly Val Leu Gln Asn Val Arg Phe Val Phe Gly Thr Thr Pro Glu  
210 215 220

Asp Ile Leu Arg Asn Lys Gly Cys Ser Ser Ser Thr Ser Val Leu Leu  
225 230 235 240

Thr Leu Asp Asn Asn Val Val Asn Gly Ser Ser Pro Ala Ile Arg Thr  
245 250 255

Asn Tyr Ile Gly His Lys Thr Lys Asp Leu Gln Ala Ile Cys Gly Ile  
260 265 270

Ser Cys Asp Glu Leu Ser Ser Met Val Leu Glu Leu Arg Gly Leu Arg  
275 280 285

Thr Ile Val Thr Thr Leu Gln Asp Ser Ile Arg Lys Val Thr Glu Glu  
290 295 300

Asn Lys Glu Leu Ala Asn Glu Leu Arg Arg Pro Pro Leu Cys Tyr His  
305 310 315 320

Asn Gly Val Gln Tyr Arg Asn Asn Glu Glu Trp Thr Val Asp Ser Cys  
325 330 335

Thr Glu Cys His Cys Gln Asn Ser Val Thr Ile Cys Lys Lys Val Ser  
340 345 350

ES 2 640 900 T3

Cys Pro Ile Met Pro Cys Ser Asn Ala Thr Val Pro Asp Gly Glu Cys  
 355 360 365  
 Cys Pro Arg Cys Trp Pro Ser Asp Ser Ala Asp Asp Gly Trp Ser Pro  
 370 375 380  
 Trp Ser Glu Trp Thr Ser Cys Ser Thr Ser Cys Gly Asn Gly Ile Gln  
 385 390 395 400  
 Gln Arg Gly Arg Ser Cys Asp Ser Leu Asn Asn Arg Cys Glu Gly Ser  
 405 410 415  
 Ser Val Gln Thr Arg Thr Cys His Ile Gln Glu Cys Asp Lys Arg Phe  
 420 425 430  
 Lys Gln Asp Gly Gly Trp Ser His Trp Ser Pro Trp Ser Ser Cys Ser  
 435 440 445  
 Val Thr Cys Gly Asp Gly Val Ile Thr Arg Ile Arg Leu Cys Asn Ser  
 450 455 460  
 Pro Ser Pro Gln Met Asn Gly Lys Pro Cys Glu Gly Glu Ala Arg Glu  
 465 470 475 480  
 Thr Lys Ala Cys Lys Lys Asp Ala Cys Pro Ile Asn Gly Gly Trp Gly  
 485 490 495  
 Pro Trp Ser Pro Trp Asp Ile Cys Ser Val Thr Cys Gly Gly Gly Val  
 500 505 510  
 Gln Lys Arg Ser Arg Leu Cys Asn Asn Pro Thr Pro Gln Phe Gly Gly  
 515 520 525  
 Lys Asp Cys Val Gly Asp Val Thr Glu Asn Gln Ile Cys Asn Lys Gln  
 530 535 540  
 Asp Cys Pro Ile Asp Gly Cys Leu Ser Asn Pro Cys Phe Ala Gly Val  
 545 550 555 560  
 Lys Cys Thr Ser Tyr Pro Asp Gly Ser Trp Lys Cys Gly Ala Cys Pro  
 565 570 575  
 Pro Gly Tyr Ser Gly Asn Gly Ile Gln Cys Thr Asp Val Asp Glu Cys  
 580 585 590  
 Lys Glu Val Pro Asp Ala Cys Phe Asn His Asn Gly Glu His Arg Cys  
 595 600 605

ES 2 640 900 T3

Glu Asn Thr Asp Pro Gly Tyr Asn Cys Leu Pro Cys Pro Pro Arg Phe  
 610 615 620

Thr Gly Ser Gln Pro Phe Gly Gln Gly Val Glu His Ala Thr Ala Asn  
 625 630 635 640

Lys Gln Val Cys Lys Pro Arg Asn Pro Cys Thr Asp Gly Thr His Asp  
 645 650 655

Cys Asn Lys Asn Ala Lys Cys Asn Tyr Leu Gly His Tyr Ser Asp Pro  
 660 665 670

Met Tyr Arg Cys Glu Cys Lys Pro Gly Tyr Ala Gly Asn Gly Ile Ile  
 675 680 685

Cys Gly Glu Asp Thr Asp Leu Asp Gly Trp Pro Asn Glu Asn Leu Val  
 690 695 700

Cys Val Ala Asn Ala Thr Tyr His Cys Lys Lys Asp Asn Cys Pro Asn  
 705 710 715 720

Leu Pro Asn Ser Gly Gln Glu Asp Tyr Asp Lys Asp Gly Ile Gly Asp  
 725 730 735

Ala Cys Asp Asp Asp Asp Asp Asn Asp Lys Ile Pro Asp Asp Arg Asp  
 740 745 750

Asn Cys Pro Phe His Tyr Asn Pro Ala Gln Tyr Asp Tyr Asp Arg Asp  
 755 760 765

Asp Val Gly Asp Arg Cys Asp Asn Cys Pro Tyr Asn His Asn Pro Asp  
 770 775 780

Gln Ala Asp Thr Asp Asn Asn Gly Glu Gly Asp Ala Cys Ala Ala Asp  
 785 790 795 800

Ile Asp Gly Asp Gly Ile Leu Asn Glu Arg Asp Asn Cys Gln Tyr Val  
 805 810 815

Tyr Asn Val Asp Gln Arg Asp Thr Asp Met Asp Gly Val Gly Asp Gln  
 820 825 830

Cys Asp Asn Cys Pro Leu Glu His Asn Pro Asp Gln Leu Asp Ser Asp  
 835 840 845

Ser Asp Arg Ile Gly Asp Thr Cys Asp Asn Asn Gln Asp Ile Asp Glu

ES 2 640 900 T3

850						855										860
Asp	Gly	His	Gln	Asn	Asn	Leu	Asp	Asn	Cys	Pro	Tyr	Val	Pro	Asn	Ala	
865					870					875					880	
	Asn	Gln	Ala	Asp	His	Asp	Lys	Asp	Gly	Lys	Gly	Asp	Ala	Cys	Asp	His
				885					890						895	
	Asp	Asp	Asp	Asn	Asp	Gly	Ile	Pro	Asp	Asp	Lys	Asp	Asn	Cys	Arg	Leu
				900					905					910		
	Val	Pro	Asn	Pro	Asp	Gln	Lys	Asp	Ser	Asp	Gly	Asp	Gly	Arg	Gly	Asp
			915					920					925			
	Ala	Cys	Lys	Asp	Asp	Phe	Asp	His	Asp	Ser	Val	Pro	Asp	Ile	Asp	Asp
		930					935					940				
	Ile	Cys	Pro	Glu	Asn	Val	Asp	Ile	Ser	Glu	Thr	Asp	Phe	Arg	Arg	Phe
	945					950					955					960
	Gln	Met	Ile	Pro	Leu	Asp	Pro	Lys	Gly	Thr	Ser	Gln	Asn	Asp	Pro	Asn
					965					970						975
	Trp	Val	Val	Arg	His	Gln	Gly	Lys	Glu	Leu	Val	Gln	Thr	Val	Asn	Cys
				980					985					990		
	Asp	Pro	Gly	Leu	Ala	Val	Gly	Tyr	Asp	Glu	Phe	Asn	Ala	Val	Asp	Ph
			995					1000					1005			
	Ser	Gly	Thr	Phe	Phe	Ile	Asn	Thr	Glu	Arg	Asp	Asp	Asp	Tyr	Ala	
		1010					1015					1020				
	Gly	Phe	Val	Phe	Gly	Tyr	Gln	Ser	Ser	Ser	Arg	Phe	Tyr	Val	Val	
		1025					1030					1035				
	Met	Trp	Lys	Gln	Val	Thr	Gln	Ser	Tyr	Trp	Asp	Thr	Asn	Pro	Thr	
		1040					1045					1050				
	Arg	Ala	Gln	Gly	Tyr	Ser	Gly	Leu	Ser	Val	Lys	Val	Val	Asn	Ser	
		1055					1060					1065				
	Thr	Thr	Gly	Pro	Gly	Glu	His	Leu	Arg	Asn	Ala	Leu	Trp	His	Thr	
		1070					1075					1080				
	Gly	Asn	Thr	Pro	Gly	Gln	Val	Arg	Thr	Leu	Trp	His	Asp	Pro	Arg	
		1085					1090					1095				

ES 2 640 900 T3

His Ile Gly Trp Lys Asp Phe Thr Ala Tyr Arg Trp Arg Leu Ser  
1100 1105 1110

His Arg Pro Lys Thr Gly Phe Ile Arg Val Val Met Tyr Glu Gly  
1115 1120 1125

Lys Lys Ile Met Ala Asp Ser Gly Pro Ile Tyr Asp Lys Thr Tyr  
1130 1135 1140

Ala Gly Gly Arg Leu Gly Leu Phe Val Phe Ser Gln Glu Met Val  
1145 1150 1155

Phe Phe Ser Asp Leu Lys Tyr Glu Cys Arg Asp Pro  
1160 1165 1170

<210> 7  
<211> 1758  
<212> ADN  
<213> *Homo sapiens*  
  
<400> 7

5

ES 2 640 900 T3

ctgctgtctg cggaggaaac tgcacgacg gacggccgcc cagctacggg aggacctgga 60  
 gtggcaactgg gcgcccgcag gaccatcccc gggacccgcc tgcccctcgg cgcgccgccc 120  
 cgcggggccg ctccccgtcg gggtcccccag ccacagcctt acctacgggc tcctgactcc 180  
 gcaaggcttc cagaagatgc togaaccacc ggcgggggoc tcggggcagc agtgagggag 240  
 gcgtccagcc cccactcag ctcttctcct cctgtgccag gggctccccg ggggatgagc 300  
 atggtggttt tcctcggag cccctggct cgggacgtct gagaagatgc cggatcatgag 360  
 gctgttcctt tgcttctgc agctcctggc cgggctggcg ctgcctgctg tgcccccca 420  
 gcagtgggccc ttgtctgctg ggaacggctc gtcagaggtg gaagtgttac ccttccagga 480  
 agtgtggggc cgcagctact gccgggctc ggagaggctg gtggacgtcg tgtccgagta 540  
 cccagcggag gtggagcaca tggtcagccc atcctgtgtc tccctgctgc gctgcaccgg 600  
 ctgctgcggc gatgagaatc tgcactgtgt gccggtggag acggccaatg tcaccatgca 660  
 gctcctaaag atccgttctg gggaccggcc ctccctacgtg gagctgacgt tctctcagca 720  
 cgttcgctgc gaatgccggc ctctgcggga gaagatgaag ccggaagga ggagacccaa 780  
 gggcaggggg aagaggagga gagagaagca gagaccaca gactgccacc tgtgcggcga 840  
 tgctgttccc cggaggtaac ccacccttg gaggagagag accccgcacc cggctcgtgt 900  
 atttattacc gtcacactct tcagtactc ctgctggtac ctgccctcta tttattagcc 960  
 aactgtttcc ctgctgaatg cctogctccc ttcaagacga ggggcagggg aggacaggac 1020  
 cctcaggaat tcagtgcctt caacaacgtg agagaaagag agaagccagc cacagacccc 1080  
 tgggagcttc cgctttgaaa gaagcaagac acgtggcctc gtgaggggca agctaggccc 1140  
 cagagggcct ggaggtctcc aggggcctgc agaaggaaag aagggggccc tgctacctgt 1200  
 tcttgggect caggctctgc acagacaagc agcccttget ttcggagctc ctgtccaaag 1260  
 tagggatgcg gatcctgctg gggccgccac gccctggctg gtgggaaggc cggcagcggg 1320  
 cggaggggat ccagccactt cccctcttc ttctgaagat cagaacattc agctctggag 1380  
 aacagtgggt gcctggggc ttttgccact ccttgtccc cgtgatctcc cctcacactt 1440  
 tgccatttgc ttgtactggg acattgttct ttccggccaa ggtgccacca ccctgcccc 1500  
 cctaagagac acatacagag tgggccccgg gctggagaaa gagctgcctg gatgagaaac 1560  
 agctcagcca gtggggatga ggtcaccagg ggaggagcct gtgcgtcca gctgaaggca 1620  
 gtggcagggg agcaggttcc ccaagggccc tggcaccccc acaagctgtc cctgcagggc 1680  
 catctgactg ccaagccaga ttctcttgaa taaagtattc tagtgtggaa aaaaaaaaaa 1740  
 aaaaaaaaaa aaaaaaaaaa 1758

<210> 8  
 <211> 513  
 <212> ADN  
 <213> *Homo sapiens*

ES 2 640 900 T3

<400> 8

```

atgccggtca tgaggctggt cccttgcttc ctgcagctcc tggccgggct ggcgctgcct      60
gctgtgcccc cccagcagtg ggccttgtct gctgggaacg gctcgtcaga ggtggaagtg      120
gtacccttcc aggaagtgtg gggccgcagc tactgccggg cgctggagag gctggtggac      180
gtcgtgtccg agtaccocag cgaggtggag cacatgttca gcccatcctg tgtctccctg      240
ctgcgctgca cgggctgctg cggcgatgag aatctgcact gtgtgccggt ggagacggcc      300
aatgtcacca tgcagctcct aaagatccgt tctggggacc ggccctocta cgtggagctg      360
acgttctctc agcacgttog ctgcgaatgc cggcctctgc gggagaagat gaagccgaa      420
aggaggagac ccaagggcag ggggaagagg aggagagaga agcagagacc cacagactgc      480
cacctgtgcg gcgatgctgt tccccggagg taa                                     513

```

5

<210> 9  
 <211> 170  
 <212> PRT  
 <213> *Homo sapiens*

10

<400> 9

```

Met Pro Val Met Arg Leu Phe Pro Cys Phe Leu Gln Leu Leu Ala Gly
 1           5           10           15

Leu Ala Leu Pro Ala Val Pro Pro Gln Gln Trp Ala Leu Ser Ala Gly
          20           25           30

```

ES 2 640 900 T3

Asn Gly Ser Ser Glu Val Glu Val Val Pro Phe Gln Glu Val Trp Gly  
 35 40 45

Arg Ser Tyr Cys Arg Ala Leu Glu Arg Leu Val Asp Val Val Ser Glu  
 50 55 60

Tyr Pro Ser Glu Val Glu His Met Phe Ser Pro Ser Cys Val Ser Leu  
 65 70 75 80

Leu Arg Cys Thr Gly Cys Cys Gly Asp Glu Asn Leu His Cys Val Pro  
 85 90 95

Val Glu Thr Ala Asn Val Thr Met Gln Leu Leu Lys Ile Arg Ser Gly  
 100 105 110

Asp Arg Pro Ser Tyr Val Glu Leu Thr Phe Ser Gln His Val Arg Cys  
 115 120 125

Glu Cys Arg Pro Leu Arg Glu Lys Met Lys Pro Glu Arg Arg Arg Pro  
 130 135 140

Lys Gly Arg Gly Lys Arg Arg Arg Glu Lys Gln Arg Pro Thr Asp Cys  
 145 150 155 160

His Leu Cys Gly Asp Ala Val Pro Arg Arg  
 165 170

<210> 10  
 <211> 708  
 <212> ADN  
 <213> *Homo sapiens*

<400> 10

ccccgcaggc tgagggcagg tgggaagcaa acccggacgc atcgcagcag cagcagcagc 60

agcagaagca gcagcagcag cctccgcagt ccctccagag acatggatcc ccagacagca 120

ccttcccggg cgctcctgct cctgctcttc ttgcatctgg ctttctggg aggtcgttcc 180

caccgcctgg gcagccccgg ttcagcctcg gacttggaaa cgtccgggtt acaggagcag 240

cgcaaccatt tgcaaggcaa actgtcggag ctgcaggtgg agcagacatc cctggagccc 300

ctccaggaga gccccgtcc cacaggtgtc tggaagtecc gggaggtagc caccgagggc 360

atccgtgggc accgcaaaat ggtcctctac accctgctgg caccacgaag cccaagatg 420

gtgcaagggt ctggctgctt tgggaggaag atggaccgga tcagctcctc cagtggcctg 480

ggctgcaaag tgctgagggc gcattaagag gaagtccctg ctgcagacac ctgcttctga 540

ttccacaagg ggctttttcc tcaaccctgt ggccgccttt gaagtgactc atttttttaa 600

5

10

ES 2 640 900 T3

tgtatttatg tatttatttg attgttttat ataagatggt ttcttacott tgagcacaaa 660  
 atttccacgg tgaataaag tcaacattat aagctttaaa aaaaaaaaa 708

5 <210> 11  
 <211> 405  
 <212> ADN  
 <213> *Homo sapiens*

<400> 11

atggatcccc agacagcacc ttcccgggag ctctgctcc tgetcttctt gcactctggct 60  
 ttctgggag gtogttcca cccgctgggc agccccggtt cagcctcgga cttggaaacg 120  
 tccgggttac aggagcagcg caaccatttg cagggcaaac tgtcggagct gcaggtggag 180  
 cagacatccc tggagcccct ccaggagagc ccccgtcca caggtgtctg gaagtcccg 240  
 gaggtagcca cagagggcat ccgtgggcac cgcaaatgg tcctctacac cctgcgggca 300  
 ccacgaagcc ccaagatggt gcaagggtct ggctgcttg ggaggaagat ggaccggatc 360  
 agctcctcca gtggcctggg ctgcaaagtg ctgaggcggc attaa 405

10 <210> 12  
 <211> 134  
 <212> PRT  
 15 <213> *Homo sapiens*

<400> 12

ES 2 640 900 T3

Met Asp Pro Gln Thr Ala Pro Ser Arg Ala Leu Leu Leu Leu Leu Phe  
 1 5 10 15

Leu His Leu Ala Phe Leu Gly Gly Arg Ser His Pro Leu Gly Ser Pro  
 20 25 30

Gly Ser Ala Ser Asp Leu Glu Thr Ser Gly Leu Gln Glu Gln Arg Asn  
 35 40 45

His Leu Gln Gly Lys Leu Ser Glu Leu Gln Val Glu Gln Thr Ser Leu  
 50 55 60

Glu Pro Leu Gln Glu Ser Pro Arg Pro Thr Gly Val Trp Lys Ser Arg  
 65 70 75 80

Glu Val Ala Thr Glu Gly Ile Arg Gly His Arg Lys Met Val Leu Tyr  
 85 90 95

Thr Leu Arg Ala Pro Arg Ser Pro Lys Met Val Gln Gly Ser Gly Cys  
 100 105 110

Phe Gly Arg Lys Met Asp Arg Ile Ser Ser Ser Ser Gly Leu Gly Cys  
 115 120 125

Lys Val Leu Arg Arg His  
 130

- <210> 13
- <211> 184
- <212> PRT
- <213> *Homo sapiens*
- <400> 13

ES 2 640 900 T3

Phe Asp Ala Ile Ala Glu Ile Gly Asn Gln Leu Tyr Leu Phe Lys Asp  
 1 5 10 15  
 Gly Lys Tyr Trp Arg Phe Ser Glu Gly Arg Gly Ser Arg Pro Gln Gly  
 20 25 30  
 Pro Phe Leu Ile Ala Asp Lys Trp Pro Ala Leu Pro Arg Lys Leu Asp  
 35 40 45  
 Ser Val Phe Glu Glu Arg Leu Ser Lys Lys Leu Phe Phe Phe Ser Gly  
 50 55 60  
 Arg Gln Val Trp Val Tyr Thr Gly Ala Ser Val Leu Gly Pro Arg Arg  
 65 70 75 80  
 Leu Asp Lys Leu Gly Leu Gly Ala Asp Val Ala Gln Val Thr Gly Ala  
 85 90 95  
 Leu Arg Ser Gly Arg Gly Lys Met Leu Leu Phe Ser Gly Arg Arg Leu  
 100 105 110  
 Trp Arg Phe Asp Val Lys Ala Gln Met Val Asp Pro Arg Ser Ala Ser  
 115 120 125  
 Glu Val Asp Arg Met Phe Pro Gly Val Pro Leu Asp Thr His Asp Val  
 130 135 140  
 Phe Gln Tyr Arg Glu Lys Ala Tyr Phe Cys Gln Asp Arg Phe Tyr Trp  
 145 150 155 160  
 Arg Val Ser Ser Arg Ser Glu Leu Asn Gln Val Asp Gln Val Gly Tyr  
 165 170 175  
 Val Thr Tyr Asp Ile Leu Gln Cys  
 180

**REIVINDICACIONES**

1. Un procedimiento de identificación de pacientes con infarto de miocardio que tienen un mayor riesgo de desarrollar insuficiencia cardíaca, que comprende:
  - 5 - analizar, después del infarto, una muestra de fluido corporal del paciente para niveles de factor de crecimiento endotelial vascular B (VEGFB), de trombospondina-1 (THBS1) y de factor de crecimiento placentario (PGF);
  - comparar los niveles de VEGFB, THBS1 y PGF con un control; y
  - determinar si el paciente tiene un mayor riesgo de desarrollar insuficiencia cardíaca, sobre la base de dicha comparación.
- 10 2. Un procedimiento de acuerdo con la reivindicación 1, en el que control es un valor de referencia obtenible de una población de pacientes infartados con un intervalo de resultados clínicos o en el que el control es un nivel correspondiente de VEGFB, THBS1 y PGF de una muestra adicional de dicho paciente
3. Un procedimiento de acuerdo con la reivindicación 2, en el que la muestra adicional es una muestra anterior del paciente y uno de los siguientes niveles indica una mayor probabilidad de que dicho paciente sufra insuficiencia cardíaca:
  - 15 - niveles más bajos de VEGFB en la muestra ensayada del paciente en comparación con el VEGFB de la muestra anterior;
  - niveles más altos de THBS1 en la muestra ensayada del paciente en comparación con el THBS1 de la muestra anterior; y
  - 20 - niveles más altos de PGF en la muestra ensayada del paciente en comparación con el PGF de la muestra anterior.
4. Un procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la muestra de fluido corporal tomada del paciente es una muestra de sangre, una muestra de plasma, o una muestra de suero y/o en el que los niveles de VEGFB, THBS1 y PGF analizados son niveles de ARNm o niveles de proteína plasmática, y/o en el que la determinación es mediante técnicas de comparación de análisis estadístico del vecino más cercano y/o el procedimiento comprende además recopilar datos sobre uno o más pacientes con IM, incluyendo el análisis de los niveles o valores de VEGFB, THBS1 y PGF y el resultado clínico asociado para ese paciente para crear datos característicos para VEGFB, THBS1 y PGF asociados a un resultado clínico particular, y/o en el que se utiliza un clasificador para determinar el pronóstico, incluyendo el clasificador programas tales como PAM, Kstar y SVM.
- 25 5. Un procedimiento de identificación de pacientes con infarto de miocardio (IM) que tienen un mayor riesgo de desarrollar insuficiencia cardíaca, que comprende:
  - 30 - analizar una muestra de fluido corporal de un paciente post-IM para los niveles de ARNm o proteína plasmática de VEGFB, THBS1 y PGF para determinar datos característicos del paciente, asociando estos datos característicos a un resultado clínico particular relacionado con la incidencia de insuficiencia cardíaca en dicho paciente, e introducir estos datos característicos en una base de datos;
  - 35 - repetir el análisis para una pluralidad de pacientes post-IM, para llenar la base de datos para contener información de referencia sobre la relación entre IC y niveles de VEGFB, THBS1 y PGF;
  - determinar los datos característicos de un paciente post-IM con pronóstico desconocido; y
  - procesar los datos característicos para los niveles de VEGFB, THBS1 y/o PGF del paciente con pronóstico desconocido para compararlos con los datos característicos de la base de datos; y
  - 40 - emitir un pronóstico probable para el paciente con pronóstico desconocido en función del resultado de la comparación.
6. Un procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la determinación de un mayor riesgo de insuficiencia cardíaca es relativa a aquellas muestras de referencia infartadas que tienen niveles comparativamente altos de ARNm de VEGFB (relación  $\log > -1,4$ ), niveles relativamente bajos de THBS1 (relación  $\log < 0$ ) y niveles relativamente bajos de PGF (relación  $\log < -0,1$ ), en comparación con la muestra consultada, respectivamente y preferentemente en el que los valores de relación varían en una cantidad seleccionada del grupo que consiste en: al menos el 1 %, al menos el 5 %, al menos el 7 %, al menos el 10 %, al menos el 15 %, al menos el 20 %, al menos el 30 %, al menos el 40 % y hasta el 50 %.
7. Un procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores, que comprende además analizar el nivel de BNP y compararlo con una muestra de referencia de péptido natriurético cerebral (BNP).
8. Un procedimiento para establecer un pronóstico para un paciente con infarto de miocardio de desarrollar insuficiencia cardíaca, comprendiendo el procedimiento:
  - 55 - analizar, después del infarto, una muestra de fluido corporal del paciente para niveles de factor de crecimiento endotelial vascular B (VEGFB) y de trombospondina-1 (THBS1) y de factor de crecimiento placentario (PGF); y
  - comparar los niveles de VEGFB, THBS1 y/o PGF con un control; y
  - determinar el pronóstico para dicho paciente sobre la base de dicha comparación.

9. Un procedimiento para establecer un pronóstico para un paciente con infarto de miocardio de desarrollar insuficiencia cardíaca, comprendiendo el procedimiento:

- 5
- analizar, después del infarto, una muestra de fluido corporal del paciente para niveles de factor de crecimiento endotelial vascular B (VEGFB) y de trombospondina-1 (THBS1) y de factor de crecimiento placentario (PGF); y
  - consultar un programa estadístico previamente construido (también llamado "clasificador") con los niveles de VEGFB, THBS1 y PGF;

en el que

- 10
- un alto nivel de VEGFB y bajos niveles de THBS1 y PGF se asocian a una menor probabilidad de desarrollar insuficiencia cardíaca; y
  - el clasificador dictará si el paciente tiene una menor probabilidad de desarrollar insuficiencia cardíaca.

10. Un procedimiento de identificación de pacientes con infarto de miocardio que tienen un riesgo reducido de insuficiencia cardíaca, que comprende:

- 15
- analizar, después del infarto, una muestra de fluido corporal del paciente para niveles de factor de crecimiento endotelial vascular B (VEGFB) y de trombospondina-1 (THBS1) y de factor de crecimiento placentario (PGF);
  - comparar los niveles de VEGFB, THBS1 y PGF con un control; y
  - determinar si el paciente tiene un riesgo reducido de desarrollar insuficiencia cardíaca, sobre la base de dicha comparación.

20

11. Un sistema que comprende una base de datos, comprendiendo la base de datos datos característicos de pacientes con IM e incluyendo los datos los resultados clínicos que corresponden con los niveles de VEGFB, THBS1 y PGF obtenidos en el análisis post-IM y al menos un dispositivo de comparación para acceder y/u operar la base de datos y para procesar la consulta y adaptado para realizar el procedimiento de cualquiera de las reivindicaciones 1-10.

Figura 1

Significativo: 525  
Número medio de falsos positivos: 128,68  
Tasa de descubrimiento falsa (%): 24,51

Hoja de gráfico SAM

Resistencia de cola (%): 32  
se (%): 24,6

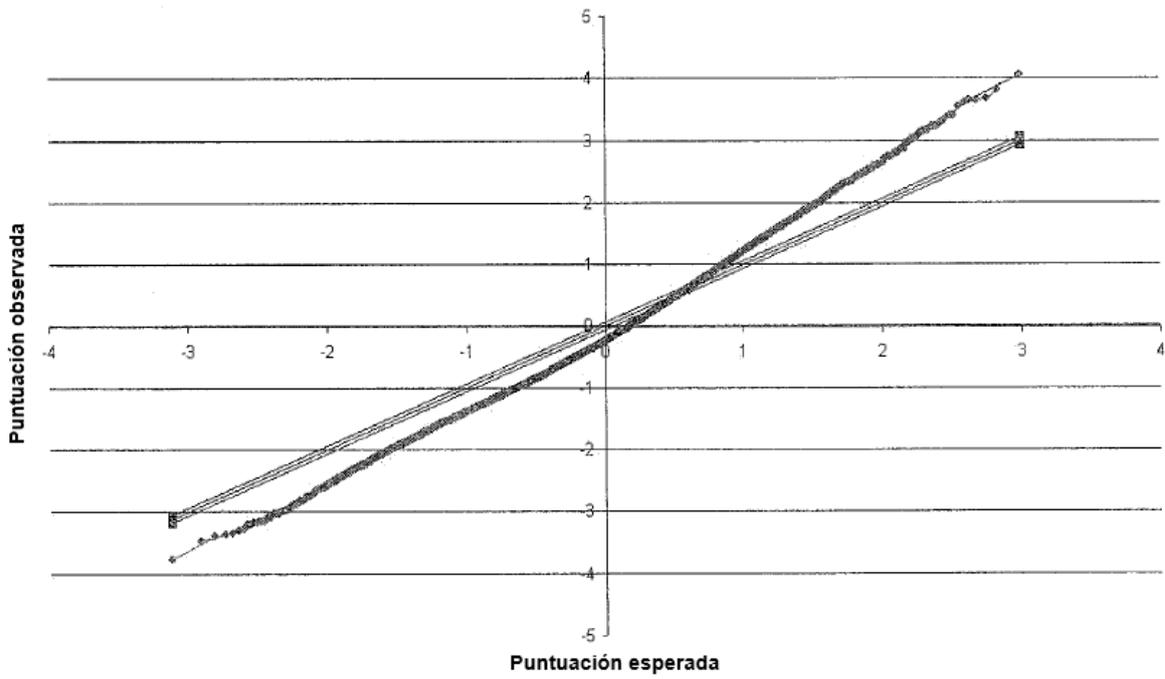


Figura 2

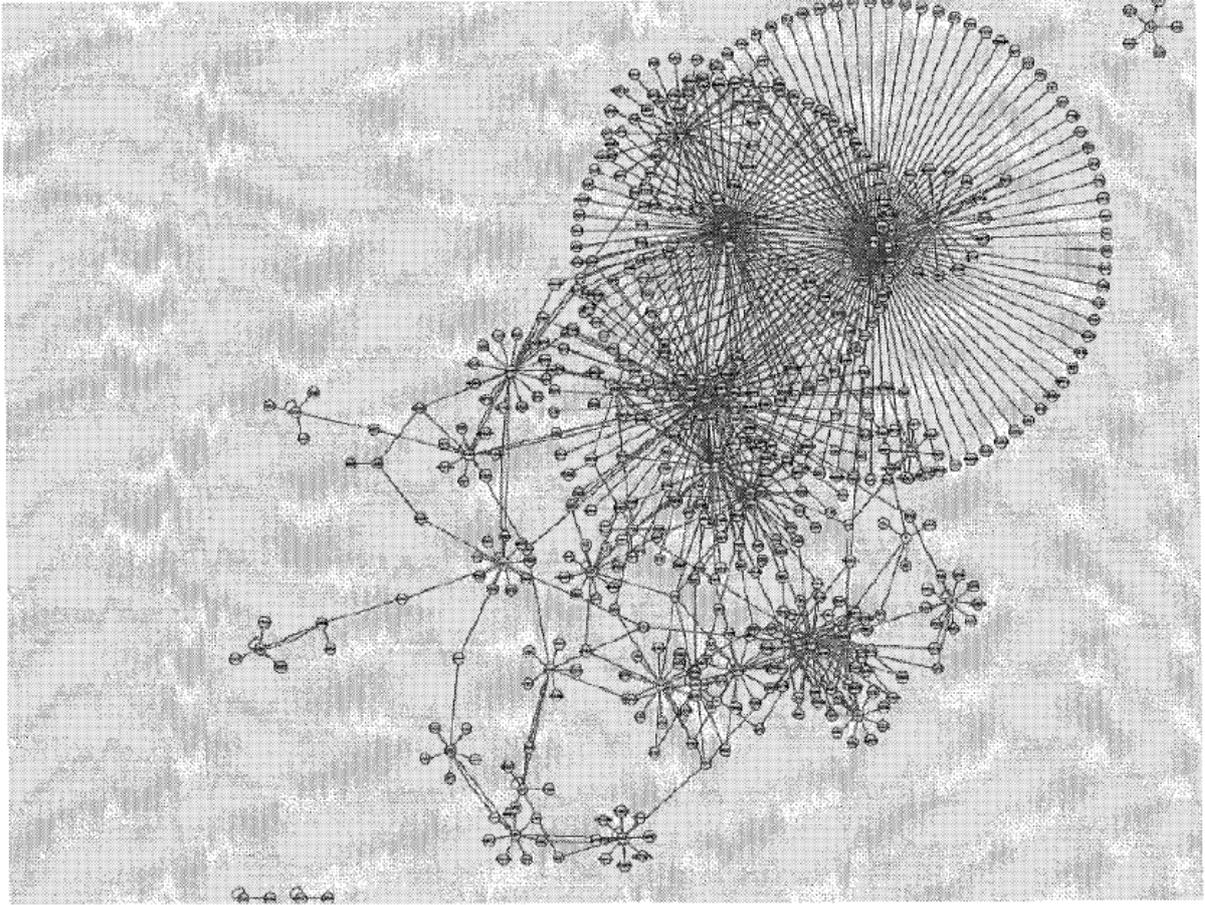


Figura 3

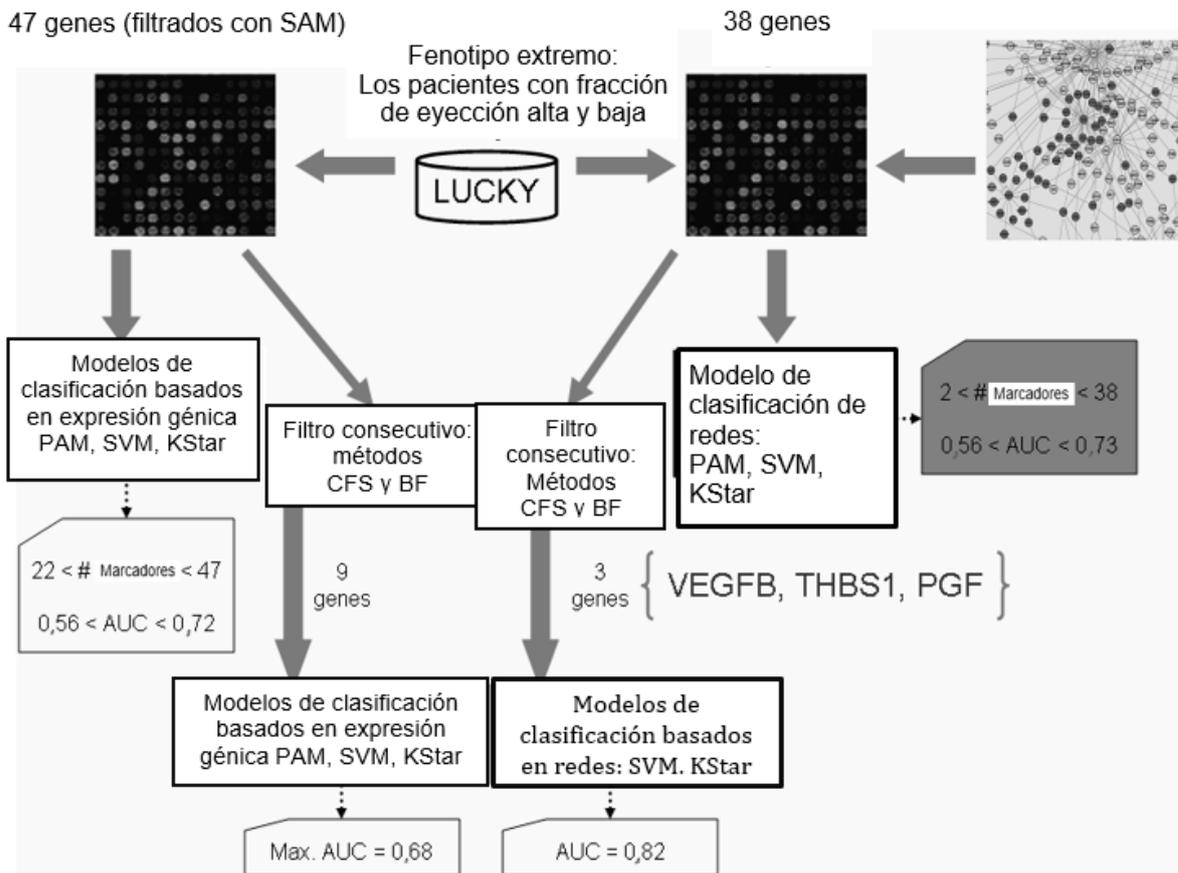


Figura 4

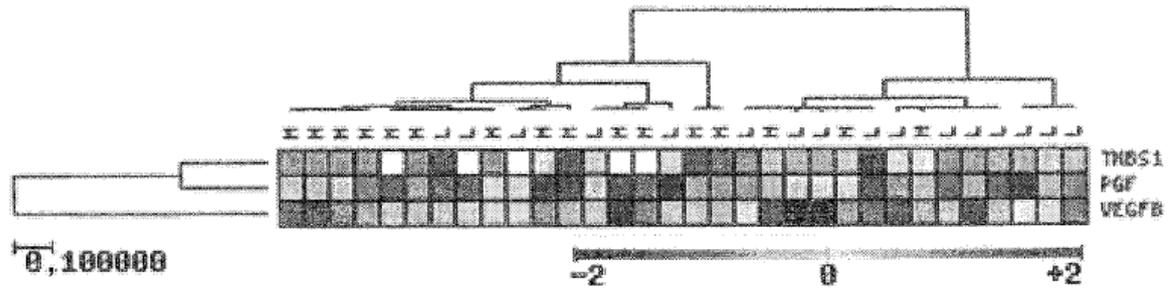


Figura 5

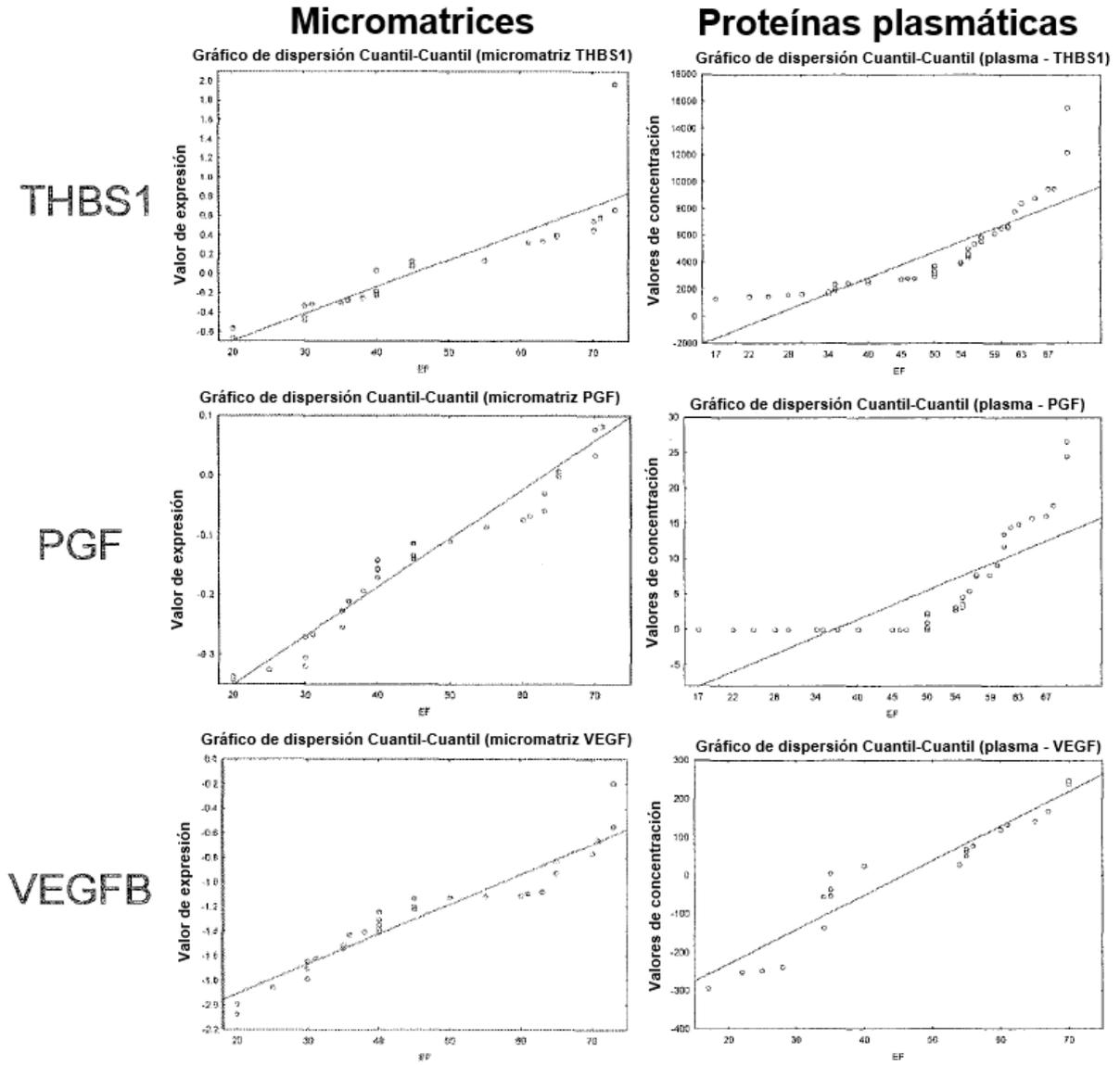


Figura 6

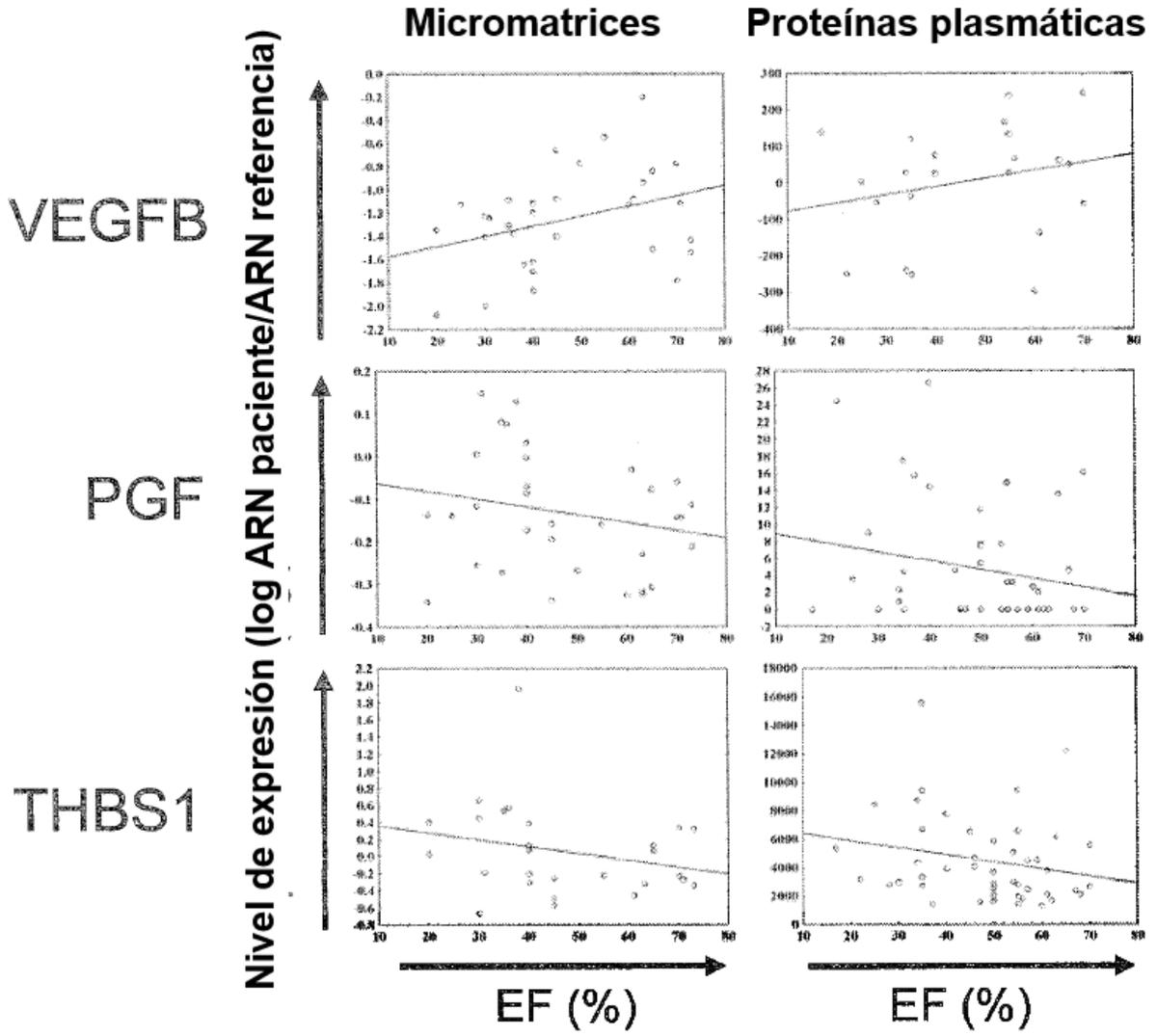


Figura 7

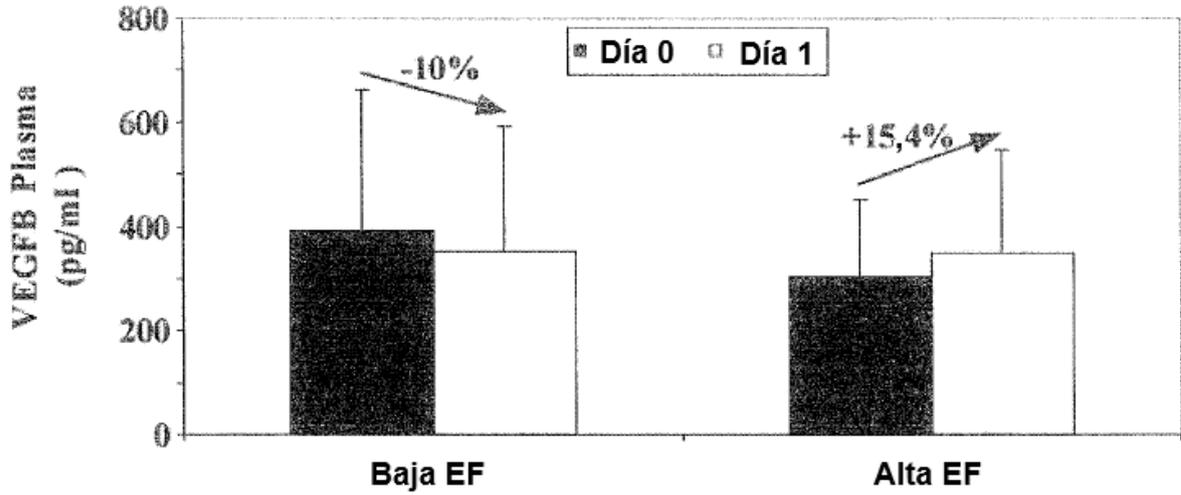
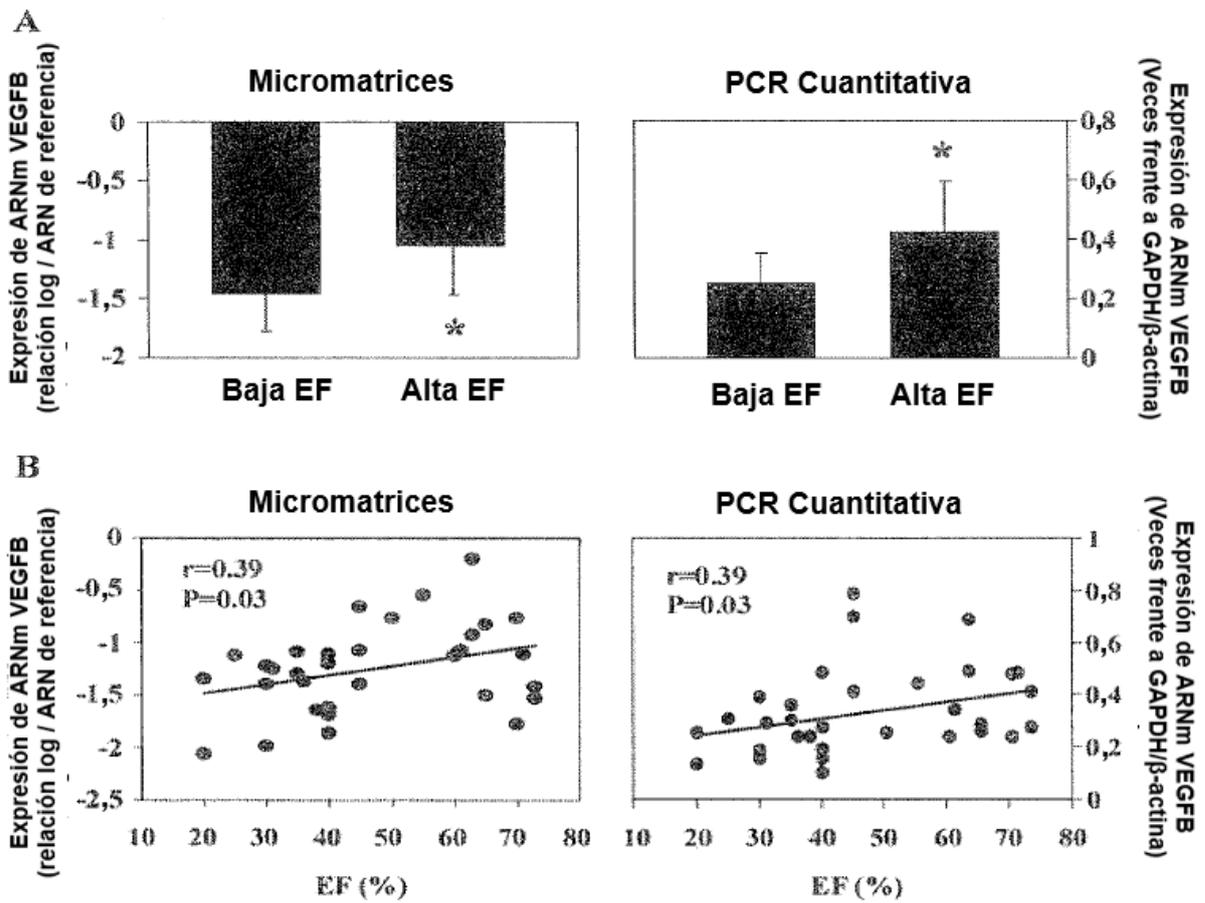


Figura 8



**Figura 9**

