

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 614 978**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **22.12.2011 PCT/EP2011/073757**

87 Fecha y número de publicación internacional: **28.06.2012 WO12085188**

96 Fecha de presentación y número de la solicitud europea: **22.12.2011 E 11807918 (5)**

97 Fecha y número de publicación de la concesión europea: **09.11.2016 EP 2655658**

54 Título: **Procedimiento de diagnóstico de trastornos hematológicos**

30 Prioridad:

22.12.2010 EP 10306483

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

02.06.2017

73 Titular/es:

**UNIVERSITE FRANCOIS RABELAIS DE TOURS
(100.0%)**

**3, rue Tanneurs, BP 4103
37041 Tours Cedex 1, FR**

72 Inventor/es:

**HERAULT, OLIVIER y
VIGNON, CHRISTINE**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 614 978 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

Procedimiento de diagnóstico de trastornos hematológicos

La presente invención se refiere a un procedimiento para diagnosticar trastornos hematológicos en un paciente.

La gran cantidad de células hematopoyéticas y los muchos estadios a través de los que pasan complican adicionalmente la clasificación de las neoplasias que se originan de este tipo de células. A pesar de los esfuerzos para establecer una clasificación basada en entidades "reales", algunas de las categorías son ambiguas y en muchos casos contienen grupos muy heterogéneos con respecto a la respuesta a la terapia en el curso clínico. Esta heterogeneidad es la responsable, por una parte, de la incesante búsqueda de marcadores capaces de diferenciar algunos comportamientos de otros, y por otra parte, de que la clasificación discutible de este tipo de neoplasias se someta a continuas revisiones.

Un sistema de clasificación ideal debería ser preciso, fácil de utilizar y debería tener especialmente significación biológica y clínica (Chan WC y col., Croat Med J. 2005; 46:349-59). Los sistemas de diagnóstico actuales y la clasificación de neoplasias hematológicas se basan en el reconocimiento de las características histológicas y morfológicas, inmunofenotípicas y citogenéticas y el estudio de un marcador molecular con valor pronóstico. Sin embargo, en algunas de las categorías diagnósticas definidas de esta manera, se observaba lo siguiente:

– Un marcador de respuesta heterogénea a la terapia: con la misma enfermedad, los pacientes que consiguen la remisión completa, o remisión parcial, o no responden, o recaen tras una cierta terapia. La capacidad para predecir es especialmente importante en este tipo de neoplasias ya que el trasplante de células madre es una respuesta alternativa eficaz pero tóxica. La capacidad para determinar que pacientes responderían a una terapia convencional antes de administrarla puede ser beneficiosa para ser capaz de aplicar el tratamiento más eficaz para cada paciente.

– Un comportamiento clínico variable: en esa categoría para algunos pacientes la enfermedad se mantiene estable durante largos periodos de tiempo y no va a ser necesaria terapia mientras que para otros la enfermedad va a progresar rápidamente necesitando terapia agresiva.

Estas variaciones apuntan a la existencia de heterogeneidad molecular de las categorías diagnósticas, diferencias que los procedimientos convencionales no son capaces de determinar y por lo tanto, la búsqueda de nuevas formas de análisis que proporcionen una mayor resolución en la caracterización de este tipo de neoplasias.

En esta línea, el uso de matrices de expresión ha demostrado ser eficaz no solo para descifrar la diversidad biológica y clínica que se encuentra en muchos tumores, sino para el entendimiento de los procesos biológicos y patológicos que afecta muchos síntomas y, en particular, el sistema hematopoyético. Las matrices de expresión son matrices ordenadas de secuencias asociadas con un soporte sólido, complementarias al ARNm o su ADNc o ARNc correspondiente, que permite el análisis de la expresión diferencial de cientos o miles de genes simultáneamente. Uno de los soportes a los que se unen habitualmente son fragmentos rectangulares de cristal parecidos a portaobjetos, un formato al que se alude habitualmente por los términos micromatriz, biochip o, simplemente chip. Su utilización se está volviendo cada vez más frecuente para el diagnóstico de distintas enfermedades o para la evolución de la evaluación de la susceptibilidad de padecerlas.

En 1999, el grupo Golub publicó uno de los primeros artículos que hacían referencia al papel de las matrices en la clasificación de neoplasias hematológicas (Golub TR y col., Science. 1999; 286:531-7). Se utilizó una matriz con 6817 genes representados para el estudio de los perfiles de expresión en leucemia mieloide aguda (AML) y leucemia linfocítica aguda (ALL). Se seleccionó un grupo de 50 genes con la capacidad para predecir el tipo de leucemia (predictor de clase) y se utilizaron para clasificar un grupo de muestras desconocidas en las categorías correctas. El estudio de la expresión de estos 50 genes es suficiente para la clasificación de una muestra de AML o ALL. A pesar del hecho de que la distinción entre AML y ALL está bien establecida con los procedimientos diagnósticos actuales, el estudio revelaba la existencia de patrones de expresión específicos asociados con cada tipo de leucemia aguda y probaban su uso.

La solicitud US 2007/154931 desvela genes asociados con la progresión y respuesta en la leucemia mieloide crónica y los usos de los mismos.

La solicitud de patente europea EP1947194 ha propuesto el uso de oligonucleótidos específicos para diagnosticar enfermedades leucémicas linfocíticas crónicas (CLL).

El artículo de Mills y col. (Blood. 2009; vol. 114; 5:1063-1072) desvela que los clasificadores basados en micromatrices y los modelos de pronóstico identifican subgrupos con resultados clínicos distintos y alto riesgo de transformación en AML de síndromes mielodisplásicos.

La solicitud internacional WO 2005/080601 desvela procedimientos de análisis genéticos para la clasificación, diagnóstico y pronóstico de leucemia mieloide aguda (AML). Esta solicitud proporciona un procedimiento para diferenciar subtipos de AM, pero no proporciona procedimientos para la discriminación entre estados preleucémicos y leucémicos.

La solicitud internacional WO 2006/125195 desvela un grupo de 24 genes cuya expresión permite clasificar una muestra como síndrome mielodisplásico (MDS), AML, o no enfermo. Sin embargo, este documento no proporciona un procedimiento para clasificar diferentes grados de MDS.

5 De esta manera existe la necesidad de proporcionar un nuevo procedimiento para diagnosticar, clasificar el estado maligno y premaligno de cáncer.

El estrés oxidativo se define en general como un equilibrio entre la generación de especies de oxígeno reactivas (ROS) y los sistemas de defensa antioxidante disfuncionales. Se sabe hace mucho que está implicado en la patofisiología del cáncer.

10 El alto nivel de ROS que se produce bien endógeno o exógenamente puede atacar lípidos, proteínas y ácidos nucleicos simultáneamente en las células vivas. Esta da lugar al desarrollo en las células de distintos mecanismos de defensa antioxidante tanto para evitar la formación excesiva de ROS como para limitar sus efectos perjudiciales. El equilibrio redox apropiado se mantiene por medio de la acción combinada de enzimas antioxidantes.

15 Los síndromes mielodisplásicos y la leucemia aguda se caracteriza por una hematopoyesis patológica. Las ROS tiene ciertamente un importante papel en la hematopoyesis humana. Por ejemplo, está bien establecido en modelos murinos que la activación de MAPK p38 inducida por ROS es crucial en la hematopoyesis y se necesita el aumento de niveles de ROS para desencadenar que las células madre hematopoyéticas (HSC) salgan de la quiescencia y se dirijan a la maduración y diferenciación. Además, un nivel alto de ROS induce una perturbación en la actividad de auto-renovación de HSC. Ahora está bien establecido que la progresión de células normales a transformación neoplásica es el resultado de la acumulación de mutaciones en genes que controlan la proliferación celular, supervivencia, y diferenciación. Aproximadamente, el 30 % de los casos de síndromes mielodisplásicos progresan a leucemia aguda. Las evidencias indirectas sugieren un papel de daño oxidante del ADN en la patogénesis de la mielodisplasia. Además, la cuantificación por citometría de flujo de la ROS en las células de pacientes mielodisplásicos y leucémicos revela un aumento en el nivel de ROS en todos los casos.

25 Así, el estudio de la respuesta antioxidante en el estado premaligno y maligno parece ser un buen comienzo para proporcionar un nuevo procedimiento útil.

30 La hipótesis de célula madre cancerosa (CSC) sugiere que un subgrupo de células de un tumor tiene la capacidad de recapitular la generación de un tumor en crecimiento continuo (Clarke MF y col., Cancer Research. 2006; 66:9339-44). Las CSC se describen mejor en seres humanos en los que se espera aislar las llamadas células madre de leucemia raras (L-HSC) y demuestran transmitir la enfermedad cuando se introducen en ratones inmunocomprometidos (Lapidot T y col., Nature. 1994 645-8). Las células que no comparten este fenotipo a menudo representan la mayoría del clon leucémico, pero fallan al trasplantarse. El primer trabajo en L-HSC se ha extendido actualmente a una lista de tumores que se está expandiendo rápidamente (Bomken S, Br. J. Cancer. 2010; 103:439-45). Debido a que parecen ser resistentes a los fármacos que se utilizan normalmente para tratar la leucemia en seres humanos, las L-HCS pueden ser responsables de recaídas en algunos pacientes (Ishikawa F, Nat. Biotechnol. 2007; 25:1315-21). Los genes que son funcionalmente significativos para la expansión de L-HSC pueden por lo tanto representar en último término dianas terapéuticas. Esto da lugar a la posibilidad de que otros sistemas de búsqueda de ROS tengan importancia reguladora en otras células madre cancerosas. Además, un bajo nivel de ROS en CSC de mama se ha comunicado recientemente, en las que se asociaba con un aumento de expresión de los genes de biosíntesis de glutatión (Diehn M, Nature. 2009; 458:780-3).

40 Por lo tanto, un objetivo de la invención es proporcionar un nuevo procedimiento para diagnosticar el cáncer.

Otro objetivo de la invención es proporcionar un rápido procedimiento eficaz para clasificar muestras patológicas, que no se pueden clasificar por otros procedimientos.

Otro objetivo de la invención es proporcionar un kit para la implementación de los procedimientos anteriores.

45 Otro objetivo de la invención es proporcionar una composición que permita la implementación del procedimiento anterior.

La invención se ilustra por un procedimiento para el diagnóstico, y/o la clasificación, preferentemente *in vitro*, de un trastorno hematológico, en particular un trastorno hematológico mieloide o linfoide, preferentemente un trastorno hematológico mieloide.

Dicho procedimiento comprende las etapas de:

50 a) medir, a partir de las células contenidas en una muestra biológica de un sujeto, preferentemente a partir de células sanguíneas o células de médula ósea que contiene la muestra, el nivel de expresión de al menos los genes de un subgrupo de 6 genes que pertenecen a un grupo de genes que se escoge de entre un grupo de 24 genes, dicho grupo de 24 genes comprende o están constituido por las secuencias de ácido nucleico de SEQ ID NO: 1 a 24,
55 en el que dicho sujeto es sospechoso de padecer un trastorno hematológico, en particular un trastorno

hematológico mieloide y/o linfóide, preferentemente un trastorno hematológico mieloide, perteneciendo dichos 6 genes a dicho subgrupo que comprende o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1 a 6,

- 5 b) comparar el nivel de expresión de cada uno de los genes medidos en la etapa a), con el nivel de expresión de los mismos genes respectivos de células contenidas en una muestra de control preferentemente a partir de células sanguíneas o células de médula ósea, siendo dicha muestra de control de la misma naturaleza que dicha muestra biológica, para establecer una relación de nivel de expresión genética de cada gen de dicho subgrupo, y
- 10 c) determinar el estado de dicha muestra biológica de manera que si la relación establecida en la etapa b) para cada uno de los genes de cualquier combinación de al menos 3 genes de cada subgrupo es de ≥ 2 o $\leq 0,5$, dicha muestra biológica es representativa de células de un trastorno hematológico.

La invención se refiere a un procedimiento *in vitro* para el diagnóstico, de una leucemia mieloide aguda o un trastorno mielodisplásico, comprendiendo dicho procedimiento las etapas de:

- 15 a) medir, a partir de las células contenidas en una muestra biológica de un sujeto, preferentemente a partir de células sanguíneas o células de médula ósea que contiene la muestra, el nivel de expresión de al menos los genes de un subgrupo de 6 genes que pertenecen a un grupo de genes escogidos de entre un grupo de 24 genes,
- 20 comprendiendo o estando constituido dicho grupo de 24 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 a 24,
- comprendiendo o estando constituidos dichos 6 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 a 6,
- 25 b) comparar el nivel de expresión de cada uno de los genes medidos en la etapa a), con el nivel de expresión de los mismos genes respectivos de células contenidas en una muestra de control preferentemente a partir de células sanguíneas o células de médula ósea, siendo dicha muestra de control de la misma naturaleza que dicha muestra biológica, para establecer una relación de nivel de expresión genética de cada gen de dicho subgrupo, y
- c) determinar el estado de dicha muestra biológica de manera que si la relación establecida en la etapa b) para al menos 3 genes de dichos 6 genes es ≥ 2 o $\leq 0,5$, dicha muestra biológica es representativa de células de una leucemia mieloide o un trastorno mielodisplásico.

En otras palabras, la etapa b) de acuerdo con la invención consiste en:

- 30 comparar el nivel de expresión de cada uno de los genes medidos en la etapa a), con el nivel de expresión de los mismos genes respectivos de células contenidas en una muestra de control, preferentemente de una muestra de control que contiene células sanguíneas o células de médula ósea, siendo dicha muestra de control de la misma naturaleza que dicha muestra biológica, para establecer una relación de nivel de expresión genética R_i entre el nivel de expresión entre cada uno de los genes i medidos en la etapa a) y el nivel de expresión de los mismos genes i respectivos de las células contenidas en una muestra de control, para cada uno de los genes de dicho subgrupo.

La etapa c) de acuerdo con la invención consiste en determinar el estado de dicha muestra biológica de manera que si la relación R_i de cada uno de los genes de cualquier combinación de 3 genes de dicho subgrupo es

- ≥ 2 ,
- $\leq 0,5$,

- 40 dicha muestra biológica es representativa de células de un trastorno hematológico.

La invención se basa en la observación inesperada hecha por los inventores de que al menos 6 genes específicos, es decir, los genes que comprenden o están constituidos por las secuencias de ácido nucleicos SEQ ID NO 1-6, que pertenecen a un grupo de 24 genes específicos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO 1-24 son suficientes para determinar el estado de un trastorno hematológico.

- 45 La invención se lleva a cabo preferentemente con muestras de pacientes que no se han tratado para un trastorno hematológico.

Estos genes son específicamente genes que codifican enzimas implicadas en la detoxificación de las células, en las cuales se acumulan ROS.

El proceso natural implicado en la eliminación de ROS se representa en la Figura 1.

- 50 De acuerdo con la invención el grupo (C) de 24 genes comprende un grupo (B) de genes, dicho grupo comprende un subgrupo (A) de 6 genes específicos como se ha definido anteriormente. Las imbricaciones de grupo/grupo/subgrupo de acuerdo con la invención se representan en la Figura 2.

El procedimiento de acuerdo con la invención se lleva a cabo por lo tanto de la siguiente manera:

- a partir de una muestra de un paciente, se extrae las moléculas de ácido nucleico contenidas en dicha muestra, preferentemente moléculas de ARN, de acuerdo con los procedimientos de extracción conocidos en la técnica,
- la cantidad de moléculas de ácido nucleico específicas, correspondientes a las moléculas de ácido nucleico que comprenden o están constituidas por al menos SEQ ID NO: 1-6, se cuantifican, por técnicas bien conocidas como se ilustra posteriormente en el presente documento,
- la cantidad cuantificada en la etapa anterior se compara con la cantidad de las mismas moléculas de ácido nucleico de una muestra de control, y se establece una relación.

En consecuencia, la muestra de control, que se utiliza como referencia, es una muestra de un individuo sano, siendo dicha muestra sana de la misma naturaleza que la muestra del paciente. Ventajosamente, la muestra de control corresponde a un agrupamiento de numerosas muestras de diferentes individuos sanos, es decir, la muestra de control representa la media de numerosos individuos sanos.

Como se ha mencionado anteriormente, la muestra del paciente y la muestra de control son del mismo origen. Esto significa que si la muestra del paciente se origina de la sangre del paciente, la muestra de control se origina de uno o varios individuos sanos. Sangre, en la invención, significa sangre completa, plasma, suero, células mononucleares de sangre periférica (PBMC)...

De la misma manera, si la muestra del paciente se origina de la médula ósea, la muestra de control se origina de la médula ósea de uno o varios individuos sanos.

En la invención, la muestra biológica el paciente, de quien se utiliza la muestra biológica, es sospechosa de padecer un trastorno hematológico, en particular un trastorno hematológico mielóide y/o linfóide, preferentemente un trastorno hematológico mielóide.

Esto significa que el estado patológico del paciente es:

- o bien no determinado (el patólogo no sabe si dicho paciente padece un trastorno hematológico),
- o determinado (el patólogo sabe que el paciente padece un trastorno hematológico).

Esto también significa que el estado patológico del paciente:

- se determinó primero (el patólogo identifica si el paciente padece un trastorno hematológico o no),
- y si el paciente padece un trastorno hematológico, se clasifica dicho trastorno hematológico.

Si el estado patológico es indeterminado, entonces el patólogo mide los niveles de expresión de los genes que consisten en SEQ ID NO: 1-6 en una muestra biológica de dicho paciente y compara dichos niveles de expresión con los niveles de expresión de los mismos genes (es decir, los genes de SEQ ID NO: 1-6) en una muestra de control de la misma naturaleza, como se ha definido anteriormente, (por ejemplo, un agrupamiento de células de donantes sanos que se utilizan como una muestra de control o referencia) con el fin de establecer las relaciones R_i , y para determinar si el paciente padece un trastorno hematológico o no.

Si el estado patológico es determinado (por ejemplo por estudios citológicos), el patólogo mide los niveles de expresión de los genes que consisten en SEQ ID NO: 1-6 en una muestra biológica de dicho paciente, y compara dichos niveles de expresión con los niveles de expresión de los mismos genes (es decir, los genes de SEQ ID NO: 1-6) en una muestra de control de la misma naturaleza, como se ha definido anteriormente, (por ejemplo, un agrupamiento de células de donantes sanos que se utilizan como muestra de control o de referencia), con el fin de establecer las relaciones R_i , y se puede clasificar el trastorno hematológico de acuerdo con la invención.

El procedimiento de acuerdo con la invención proporciona un procedimiento eficaz, rápido y fácil de utilizar para evaluar el estado de una muestra que se identifica como, o se sospecha que es, una muestra correspondiente a un trastorno hematológico, en particular un trastorno mielóide.

Si la relación entre el nivel de expresión de al menos 3 genes de los 6 genes anteriores que pertenecen al subgrupo definido anteriormente, de una muestra de un paciente y el nivel de expresión de los mismos al menos 3 genes de una muestra de control es ≥ 2 o $\leq 0,5$, luego la muestra del paciente se consideraría como que presenta características de una muestra que se corresponde con un trastorno hematológico.

Anteriormente, y de aquí en adelante, la relación se define de la siguiente manera:

$$R_i = [\text{Cantidad (nivel de expresión) de un gen } i \text{ de SEQ ID NO: } i \text{ en la muestra del paciente}] / [\text{cantidad (nivel de expresión) de un gen } i \text{ de SEQ ID NO: } i \text{ en la muestra de control}]$$

variando i de 1 a 24.

Por lo tanto R_3 representa la relación como se define anteriormente, respecto al gen de SEQ ID NO: 3, y por lo tanto, R_i representa la relación como se ha definido anteriormente, respecto al gen de SEQ ID NO: i , variando i de 1 a 24.

En la invención, “la relación R_i de cada gen de cualquier combinación de al menos 3 genes de los 6 genes de dicho subgrupo es ≥ 2 o $\leq 0,5$ ” significa que la relación R_i de cada gen de cualquier combinación de 3, o 4 o 5 o 6 genes es ≥ 2 o $\leq 0,5$.

Las 20 combinaciones de 3 genes escogidos de entre los 6 genes se enumeran a continuación:

- 5 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 4,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 5,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 6,
 SEQ ID NO: 1 + SEQ ID NO: 3 + SEQ ID NO: 4,
 10 SEQ ID NO: 1 + SEQ ID NO: 3 + SEQ ID NO: 5,
 SEQ ID NO: 1 + SEQ ID NO: 3 + SEQ ID NO: 6,
 SEQ ID NO: 1 + SEQ ID NO: 4 + SEQ ID NO: 5,
 SEQ ID NO: 1 + SEQ ID NO: 4 + SEQ ID NO: 6,
 15 SEQ ID NO: 1 + SEQ ID NO: 5 + SEQ ID NO: 6,
 SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 4,
 SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 5,
 SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 6,
 SEQ ID NO: 2 + SEQ ID NO: 4 + SEQ ID NO: 5,
 SEQ ID NO: 2 + SEQ ID NO: 4 + SEQ ID NO: 6,
 20 SEQ ID NO: 2 + SEQ ID NO: 5 + SEQ ID NO: 6,
 SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 5,
 SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 6,
 SEQ ID NO: 3 + SEQ ID NO: 5 + SEQ ID NO: 6 y
 SEQ ID NO: 4 + SEQ ID NO: 5 + SEQ ID NO: 6.

25 Las 15 combinaciones de 4 genes de entre los 6 genes son las siguientes:

- SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 4,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 5,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 6,
 SEQ ID NO: 1 + SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 5,
 30 SEQ ID NO: 1 + SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 6,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 4 + SEQ ID NO: 5,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 4 + SEQ ID NO: 6,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 5 + SEQ ID NO: 6,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 5,
 35 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 6,
 SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 5,
 SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 6,
 SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 5 + SEQ ID NO: 6,
 SEQ ID NO: 2 + SEQ ID NO: 4 + SEQ ID NO: 5 + SEQ ID NO: 6 y
 40 SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 5 + SEQ ID NO: 6.

Las 6 combinaciones de 5 genes de entre los 6 genes son las siguientes:

- SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 5,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 6,
 45 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 5 + SEQ ID NO: 6,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 4 + SEQ ID NO: 5 + SEQ ID NO: 6,
 SEQ ID NO: 1 + SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 5 + SEQ ID NO: 6 y
 SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 5 + SEQ ID NO: 6.

Finalmente, la combinación de los seis genes es SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 4 + SEQ ID NO: 5 + SEQ ID NO: 6.

50 De acuerdo con la invención, una muestra de un paciente en el que la relación R_i de cada gen perteneciente a cualquier combinación de al menos 3 genes del subgrupo de 6 genes es ≥ 2 o $\leq 0,5$ se considerará una muestra correspondiente a un trastorno hematológico.

Por ejemplo, si se estudia la combinación SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3, si las respectivas relaciones R_1 , R_2 , R_3 como se ha descrito anteriormente son las siguientes:

- 55 $R_1 \geq 2$ y $R_2 \geq 2$ y $R_3 \geq 2$, o
 $R_1 \leq 0,5$ y $R_2 \geq 2$ y $R_3 \geq 2$, o
 $R_1 \geq 2$ y $R_2 \leq 0,5$ y $R_3 \geq 2$, o
 $R_1 \geq 2$ y $R_2 \geq 2$ y $R_3 \leq 0,5$, o

$R1 \leq 0,5$ y $R2 \leq 0,5$ y $R3 \geq 2$, o
 $R1 \leq 0,5$ y $R2 \geq 2$ y $R3 \leq 0,5$, o
 $R1 \geq 2$ y $R2 \leq 0,5$ y $R3 \leq 0,5$, o
 $R1 \leq 0,5$ y $R2 \leq 0,5$ y $R3 \leq 0,5$,

5 entonces la muestra del paciente en el que se calculan las relaciones se considerará una muestra correspondiente con un trastorno hematológico.

El ejemplo anterior se aplica mutatis mutandis a las combinaciones de al menos 3 genes mencionadas anteriormente.

10 En consecuencia, si se estudia la combinación SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 4, existen 4 combinaciones de 3 genes:

SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 3,
 SEQ ID NO: 1 + SEQ ID NO: 2 + SEQ ID NO: 4,
 SEQ ID NO: 1 + SEQ ID NO: 3 + SEQ ID NO: 4, y
 SEQ ID NO: 2 + SEQ ID NO: 3 + SEQ ID NO: 4.

15 Por lo tanto, si las respectivas relaciones R1, R2, R3 y R4 son las siguientes:

- para la combinación 1:

20 $R1 \geq 2$ y $R2 \geq 2$ y $R3 \geq 2$, o
 $R1 \leq 0,5$ y $R2 \geq 2$ y $R3 \geq 2$, o
 $R1 \geq 2$ y $R2 \leq 0,5$ y $R3 \geq 2$, o
 $R1 \geq 2$ y $R2 \geq 2$ y $R3 \leq 0,5$, o
 $R1 \leq 0,5$ y $R2 \leq 0,5$ y $R3 \geq 2$, o
 $R1 \geq 0,5$ y $R2 \geq 2$ y $R3 \leq 0,5$, o
 $R1 \geq 2$ y $R2 \leq 0,5$ y $R3 \leq 0,5$, o
 $R1 \leq 0,5$ y $R2 \leq 0,5$ y $R3 \leq 0,5$,

25 - para la combinación 2

30 $R1 \geq 2$ y $R3 \geq 2$ y $R4 \geq 2$, o
 $R1 \leq 0,5$ y $R3 \geq 2$ y $R4 \geq 2$, o
 $R1 \geq 2$ y $R3 \leq 0,5$ y $R4 \geq 2$, o
 $R1 \geq 2$ y $R3 \geq 2$ y $R4 \leq 0,5$, o
 $R1 \leq 0,5$ y $R3 \leq 0,5$ y $R4 \geq 2$, o
 $R1 \leq 0,5$ y $R3 \geq 2$ y $R4 \geq 0,5$, o
 $R1 \geq 2$ y $R3 \leq 0,5$ y $R4 \leq 0,5$, o
 $R1 \leq 0,5$ y $R3 \leq 0,5$ y $R4 \leq 0,5$,

- para la combinación 3,

35 $R1 \geq 2$ y $R2 \geq 2$ y $R4 \geq 2$, o
 $R1 \leq 0,5$ y $R2 \geq 2$ y $R4 \geq 2$, o
 $R1 \geq 2$ y $R2 \leq 0,5$ y $R4 \geq 2$, o
 $R1 \geq 2$ y $R2 \geq 2$ y $R4 \leq 0,5$, o
 $R1 \leq 0,5$ y $R2 \leq 0,5$ y $R4 \geq 2$, o
 $R1 \leq 0,5$ y $R2 \geq 2$ y $R4 \leq 0,5$, o
 $R1 \geq 2$ y $R2 \leq 0,5$ y $R4 \leq 0,5$, o
 $R1 \leq 0,5$ y $R2 \leq 0,5$ y $R4 \leq 0,5$,

- para la combinación 4

45 $R2 \geq 2$ y $R3 \geq 2$ y $R4 \geq 2$, o
 $R2 \leq 0,5$ y $R3 \geq 2$ y $R4 \geq 2$, o
 $R2 \geq 2$ y $R3 \leq 0,5$ y $R4 \geq 2$, o
 $R2 \geq 2$ y $R3 \geq 2$ y $R4 \leq 0,5$, o
 $R2 \leq 0,5$ y $R3 \leq 0,5$ y $R4 \geq 2$, o
 $R2 \leq 0,5$ y $R3 \geq 2$ y $R4 \leq 0,5$, o
 $R2 \geq 2$ y $R3 \leq 0,5$ y $R4 \leq 0,5$, o
 $R2 \leq 0,5$ y $R3 \leq 0,5$ y $R4 \leq 0,5$,

50 entonces la muestra del paciente en el que se calcularon las relaciones se considerará como una muestra correspondiente a un trastorno hematológico.

En la invención, los genes para los que se mide el nivel de expresión se representan por las moléculas de ARN que se obtienen por la transcripción de dichos genes. El proceso de transcripción se conoce bien en la técnica.

5 Por lo tanto, la invención se refiere a un procedimiento como se ha definido anteriormente, en el que el nivel de expresión de los genes anteriores se mide por determinación de la cantidad de moléculas de ARN que son los productos de la transcripción de dichos genes, es decir, que son los productos de la expresión de dichos genes.

10 Algunos genes de la invención son capaces de expresar muchas variantes, es decir, muchas moléculas de ARN que se diferencian en sus secuencias. Estas variantes generalmente se diferencian en sus secuencias tras un corte y empalme alternativo, teniendo dicho corte y empalme alternativo como consecuencia la adición, eliminación y/o la modificación de una o más partes de la secuencia de ácido nucleico contenida en el gen en la molécula de ARN resultante. El experto conoce los mecanismos del corte y empalme alternativo.

Por lo tanto, algunos genes de acuerdo con la invención pueden expresar más de una molécula de ARN y proporcionar variantes.

15 El gen PRDX es capaz de expresar 3 variantes diferentes: la primera variante comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 5, una segunda variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 73 y una tercera variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 74.

El gen SOD2 es capaz de expresar 3 variantes diferentes: la primera variante comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 7, una segunda variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 75 y una tercera variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 76.

20 El gen GSR es capaz de expresar 4 variante diferentes: la primera variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 8, una segunda variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 77, una tercera variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 78 y una cuarta variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 79.

25 El gen GLRX es capaz de expresar 2 variantes diferentes: la primera variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 9 y una segunda variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 80.

30 El gen PDRX5 es capaz de expresar 3 variantes diferentes: la primera variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 11, una segunda variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 81 y una tercera variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 82.

El gen GPX4 es capaz de expresar 2 variantes diferentes: la primera variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 14, y una segunda variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 83.

35 El gen PDRX5 es capaz de expresar 3 variantes diferentes: la primera variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 16, una segunda variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 84 y una tercera variante que comprende o consiste en la secuencia de ácido nucleico de SEQ ID NO: 85.

40 Por lo tanto, de acuerdo con la invención, la medida del nivel de expresión del gen que comprende o está constituido por variante que comprende o consiste por SEQ ID NO: 5, se puede evaluar por la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 73 o SEQ ID NO: 74.

De la misma manera, la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 7, se puede evaluar por la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 75 o SEQ ID NO: 76.

45 Además, la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 8, se puede evaluar por la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 77, SEQ ID NO: 78 o SEQ ID NO: 79.

Además, la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 9 se puede evaluar por la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 80.

50 Además, la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 11 se puede evaluar por la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 81 o SEQ ID NO: 82.

Además, la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 14, se puede evaluar por la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 83.

Además, la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 16, se puede evaluar por la medida del nivel de expresión del gen que comprende o está constituido por SEQ ID NO: 84 o SEQ ID NO: 85.

5 Por lo tanto como se ha desvelado anteriormente y a continuación en la invención, los genes que comprenden o están constituidos por las siguientes secuencias: SEQ ID NO: 5, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 11, SEQ ID NO: 14 y SEQ ID NO: 16 se pueden sustituir por sus respectivas variantes, como se ha definido anteriormente.

Los genes que se utilizan de acuerdo con la invención se representan en la siguiente tabla 1:

Tabla 1

Nombre del gen	SEQ ID
GPX3	SEQ ID NO: 1
GPX1(1)	SEQ ID NO: 2
GLRX2(2)	SEQ ID NO: 3
CAT	SEQ ID NO: 4
PRDX(1-2-3)	SEQ ID NO: 5
	SEQ ID NO: 73
	SEQ ID NO: 74
PRDX5(2)	SEQ ID NO: 6
SOD2(1-2-3)	SEQ ID NO: 7
	SEQ ID NO: 75
	SEQ ID NO: 76
GSR (1-2-3-4)	SEQ ID NO: 8
	SEQ ID NO: 77
	SEQ ID NO: 78
	SEQ ID NO: 79
GLRX(1-2)	SEQ ID NO: 9
	SEQ ID NO: 80
PRDX2(1)	SEQ ID NO: 10
PRDX5(1-3)	SEQ ID NO: 11
	SEQ ID NO: 81
	SEQ ID NO: 82
SOD1	SEQ ID NO: 12
TXN	SEQ ID NO: 13
PRDX3(1-2)	SEQ ID NO: 14
	SEQ ID NO: 83
GPX7	SEQ ID NO: 15
GPX4(1-2-3)	SEQ ID NO: 16
	SEQ ID NO: 84
	SEQ ID NO: 85
TXN2	SEQ ID NO: 17
PRDX4	SEQ ID NO: 18

(continuación)

Nombre del gen	SEQ ID
GPX1 (2)	SEQ ID NO: 19
GLRX3	SEQ ID NO: 20
PRDX2(3)	SEQ ID NO: 21
PRDX6	SEQ ID NO: 22
GLRX5	SEQ ID NO: 23
GLRX2(1)	SEQ ID NO: 24

La Tabla 1 representa SEQ ID de los genes, y las variantes cuando existen, que se utilizan en la invención.

5 De acuerdo con la invención, los trastornos hematológicos corresponden con trastornos que afectan primariamente a la sangre. En particular, los trastornos hematológicos de acuerdo con la invención engloban todas las citopenias (anemia: disminución del recuento de glóbulos rojos o hemoglobina, trombocitopenias: disminución del recuento de plaquetas, y leucopenias: disminución del recuento de leucocitos) cualquiera que sea el mecanismo tales como hemoglobinopatías y síndrome mielodisplásico, trastornos mieloproliferativos (aumento del número de células mieloides o mielofibrosis, que incluye la leucemia mieloide crónica, policitemia vera, trombocitemia esencial, mielofibrosis idiopática), trastornos linfoproliferativos (aumento del número de células linfoides, que incluye leucemia linfocítica crónica, linfomas, mielomas, plasmacitoma), leucemias agudas y coagulopatías (trastornos hemorrágicos y de coagulación).

En una realización ventajosa, la invención se refiere a un procedimiento como se ha definido anteriormente, en el que si la relación establecida en la etapa b) es

- 15 - $\leq 0,3$, para el gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1, y
 - $\geq 3,0$, para los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 y 3,

entonces dicha muestra biológica es representativa de una leucemia mieloide aguda.

En otras palabras, una realización de la invención se refiere a un procedimiento que se define anteriormente, en el que si

- 20 - la relación R1 es $\geq 0,3$, para el gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1, y
 - las relaciones R2 y R3 son $\geq 3,0$, para los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 y 3,

entonces dicha muestra biológica es representativa de una leucemia mieloide aguda.

25 En una realización de la invención, cuando la relación R1 es $\leq 0,3$ y las relaciones R2 y R3 son ambas ≥ 3 , la muestra originada a partir del paciente es representativa de una leucemia mieloide aguda (AML). Dichos 3 criterios son acumulativos.

30 La AML es la proliferación clónica de células inmaduras de origen mieloide. Estas aparecen de novo o secundariamente en pacientes con síndrome mielodisplásico (MDS). La clasificación preparada por el grupo Francés-Americano-Británico (FAB) considera ocho variedades (M0-M7) basándose en los criterios morfológicos y en el inmunofenotipo de las células neoplásicas (Bennett JM, y col., 1976).

La Organización Mundial de la Salud (OMS) clasifica la AML incorporando datos morfológicos, inmunofenotípicos, genéticos y clínicos para ser capaces de definir entidades biológicas homogéneas y con relevancia clínica. Por lo tanto, la AML se clasifica en cuatro grandes categorías:

- 35 1.- AML con anomalías genéticas recurrentes,
 2.- AML con displasia multilineal,
 3.- AML relacionada con el tratamiento y
 4.- AML no clasificable.

40 Antes de la invención, el análisis citogenético representaba el factor pronóstico más poderoso. Se utiliza para identificar los subgrupos de AML con diferentes pronósticos: bajo riesgo con respuesta favorable al tratamiento (t (8; 21), t (15; 17) o inv (16)), riesgo intermedio (cariotipo normal o t (9; 11) o algo riesgo (inv (3), del (5q) o del (7q), o más de tres alteraciones). Hay heterogeneidad molecular en el grupo de riesgo. En algunos casos de pacientes con un cariotipo normal, la presencia de mutaciones se ha encontrado en algunos genes.

Ventajosamente, la invención se refiere al procedimiento que se ha definido anteriormente, en el que si la relación establecida en la etapa b) es

- $\leq 0,3$, para el gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1, y
- $\geq 3,0$, para los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 y 3,

y además en el que la relación R10 entre el nivel de expresión del gen que consiste en SEQ ID NO: 10 medida en dicha muestra y en dicha muestra de control es menor de 0,5 ($R10 \leq 0,5$), preferentemente es menor de 0,3, ($R10 \leq 0,3$) entonces dicha muestra biológica es representativa de una leucemia mieloide aguda.

En otra realización, la invención se refiere a un procedimiento que se define anteriormente, en el que la etapa c) es de manera que si la relación establecida en la etapa b) para cada uno de los genes de cada combinación de al menos 3 genes de dicho subgrupo es ≥ 2 o $\leq 0,5$, a condición de que

- la relación entre el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1 medida en dicha muestra biológica y medida en dicha muestra de control, no es $\leq 0,3$, o
- las relaciones entre el nivel de expresión de cada uno de los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 o 3 medidas en dicha muestra biológica y medidas en dicha muestra de control, no es ≥ 3 ,

entonces dicha muestra biológica es representativa de un trastorno mielodisplásico, en particular mielodisplasia escogida de entre anemia refractaria (RA), anemia refractaria con sideroblastos anulares (RARS), citopenia refractaria con displasia multilineaje (RCMD), anemia refractaria con exceso de blastos (RAEB), síndrome de 5q y mielodisplasia inclasificable.

En otras palabras, en otra realización ventajosa, la invención se refiere a un procedimiento que se define anteriormente, en el que si la relación establecida en la etapa b) para cada uno de los genes de cualquier combinación de al menos 3 genes a partir de dicho subgrupo es ≥ 2 o $\leq 0,5$, a condición de que la combinación de al menos 3 genes que se corresponden con los genes que comprenden o están constituidos por la secuencia de ácido nucleico de SEQ ID NO: 1, 2 y 3, en la que

- la relación R1 es $\leq 0,3$, para el gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1, y
- las relaciones R1 son $\geq 3,0$, para los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 y 3,

esta combinación se excluye

entonces dicha muestra biológica es representativa de un trastorno mielodisplásico, en particular una mielodisplasia que se escoge de entre anemia refractaria con sideroblastos anulares (RARS), citopenia refractaria con displasia multilineaje (RCMD), anemia refractaria con exceso de blastos (RAEB) o síndrome 5q y mielodisplasia inclasificable.

En otra realización, la invención se refiere a un procedimiento que se ha definido anteriormente, en el que la etapa c) es de manera que si la relación establecida en la etapa b) para al menos 3 genes de dichos 6 genes es ≥ 2 o $\leq 0,5$, a condición de que

- la relación entre el nivel de expresión de cada uno de los genes que comprenden o están constituidos por la secuencia de ácido nucleico de SEQ ID NO: 1 medida en dicha muestra biológica y medida en dicha muestra de control, no es $\leq 0,3$, o
- las relaciones entre el nivel de expresión de cada uno de los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 o 3 medidas en dicha muestra biológica y medida en dicha muestra de control, no es ≥ 3 ,

entonces dicha muestra biológica es representativa de un trastorno mielodisplásico, en particular mielodisplasia escogida de entre anemia refractaria, citopenia refractaria, o anemia refractaria con exceso de blastos, síndrome 5q.

De acuerdo con la invención, si la relación de al menos 3 genes que pertenecen al subgrupo constituido por los genes que comprenden o están constituidos por la secuencia de ácido nucleico de SEQ ID NO: 1-6 es ≥ 2 o $\leq 0,5$, excluyendo la combinación particular de SEQ ID NO: 1, SEQ ID NO: 2 y SEQ ID NO: 2 en la que la relación de SEQ ID NO: 1 es 0,3 y las relaciones de SEQ ID NO: 2 y SEQ ID NO: 3 son ≥ 3 , entonces la muestra biológica es representativa de un trastorno mielodisplásico, en particular mielodisplasia escogida de entre anemia refractaria con sideroblastos anulares (RARS), citopenia refractaria con displasia multilineaje (RCMD), anemia refractaria con exceso de blastos (RAEB) o síndrome 5q y mielodisplasia inclasificable.

De acuerdo con la invención, los trastornos mielodisplásicos se definen como preleucemia, y se corresponden con una colección diversa de afecciones hematológicas unidas por la producción ineficaz (o displasia) de células

sanguíneas mieloides y el riesgo de transformación en leucemia mielógena aguda (AML).

La clasificación Francesa-Americana-Británica (FAB) ha subdividido los trastornos mielodisplásicos de la siguiente manera:

- 5 **Anemia refractaria** (RA), caracterizada por menos de un 5 % de células sanguíneas primitivas (mieloblastos) en la médula ósea y con anomalías patológicas vistas primariamente en precursores de glóbulos rojos,
- Anemia refractaria con sideroblastos anulares** (RARS), también caracterizada por menos de un 5 % de mieloblastos en la médula ósea, pero que se distingue por la presencia de un 15 % o más de precursores de glóbulos rojos en la médula que son células anormales rellenas de hierro llamadas “sideroblastos anulares”,
- 10 **Anemia refractaria con exceso de blastos** (RAEB), caracterizada por un 5-20 % de mieloblastos en la médula, **Anemia refractaria con exceso de blastos en transformación** (RAEB-T), caracterizada por un 21-30 % de mieloblastos en la médula (> 30 % de blastos se define como leucemia mielóide aguda), y **Leucemia mielomonocítica crónica** (CMML), no se tiene que confundir con la leucemia mielógena crónica o CML, caracterizada por menos de un 20 % de mieloblastos en la médula ósea y más de 1000 x 10⁹/μl de monocitos (un tipo de glóbulos blancos sanguíneos) circulantes de la sangre periférica.
- 15 Más recientemente, la Organización Mundial de la Salud (OMS) ha clasificado los síndromes displásicos de la siguiente manera:

Tabla 2

Sistema antiguo	Sistema nuevo
Anemia refractaria (RA)	<u>Citopenia</u> refractaria con displasia unilínea (anemia refractaria, <u>neutropenia refractaria</u> , y trombocitopenia refractaria)
Anemia refractaria con sideroblastos anulares (RARS)	Anemia refractaria con sideroblastos anulares (RARS) Anemia refractaria con sideroblastos anulares – trombocitosis (RARS-t) (entidad provisional) que es en esencia un <u>trastorno mielodisplásico/mieloproliferativo</u> y habitualmente tiene una mutación JAK2 (janus cinasa) – Nueva clasificación de la OMS 2008
Citopenia refractaria con displasia multilineal (RCMD)	<u>Citopenia refractaria con displasia multilineal</u> (RCMD) incluye el subgrupo de citopenia refractaria con displasia multilineal y sideroblastos anulares (RCMD-RS). RCMD incluye pacientes con cambios patológicos no restringidos a glóbulos rojos (es decir, displasia de precursores importantes de glóbulos blancos y de precursores de plaquetas (megacariocitos).
Anemia refractaria con exceso de blastos (RAEB)	Anemia refractaria con exceso de blastos I y II. RAEB se dividió en RAEB-I (un 5-9 % de blastos) y RAEB-II (un 10-19 % de blastos), que tiene un pronóstico peor que RAEB-I. Se pueden ver <u>bastones de Auer</u> en RAEB-II que pueden ser difíciles de distinguir de los de leucemia mielóide aguda.
Anemia refractaria con exceso de blastos en transformación (RAEB-T)	La categoría RAEB-T se eliminó; dichos pacientes se consideran ahora como que tienen leucemia aguda. El síndrome 5q que se ve típicamente en mujeres ancianas con recuentos de plaquetas normales o altos y eliminaciones aisladas del brazo largo del cromosoma 5 en células de médula ósea se añadió a la clasificación.
Leucemia mielomonocítica crónica (CMML)	La CMML se retiró de los síndromes mielodisplásicos y se puso en una nueva categoría de síndromes solapados <u>mielodisplásicos-mieloproliferativos</u> .
	Síndrome 5q
	Mielodisplasia inclasificable (visto en los casos de displasia megacariocítica con fibrosis y otras)
	Citopenia refractaria de la infancia (displasia en la infancia) – Nueva clasificación de la OMS 2008

- 20 El síndrome de eliminación del cromosoma 5q (monosomía del cromosoma 5q, síndrome 5q) es un trastorno raro causado por la pérdida de parte del brazo largo (brazo q) del cromosoma 5 humano.

El síndrome 5q se caracteriza por anemia macrocítica y a menudo trombocitosis, eritroblastopenia, hiperplasia megacariocítica con hipobululación nuclear y una eliminación intersticial aislada del cromosoma 5q. El síndrome 5q se encuentra predominantemente en mujeres de edad avanzada.

En otra realización más, la invención se refiere a un procedimiento definido anteriormente, en el que dicho grupo comprende 10 genes, comprendiendo o estando constituidos dichos 10 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 a 10, y la etapa c) es de tal manera que

- 5 - si la relación establecida en la etapa b) para cada uno de los genes en cualquier combinación de 3 genes de dicho subgrupo es ≥ 2 o $\leq 0,5$,

a condición de que

- 10 - la relación entre el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1 medida en dicha muestra biológica y medida en dicha muestra de control, no es $\leq 0,3$,
o
- las relaciones entre el nivel de expresión de cada uno de los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 o 3 medidas en dicha muestra biológica y medida en dicha muestra de control, no es ≥ 3

y además

- 15 - si la relación establecida en la etapa b) para al menos un gen de dicho grupo que no pertenece a dicho subgrupo es ≥ 2 , y la relación establecida en la etapa b) de al menos otro gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,5$,

entonces dicha muestra biológica es representativa de una anemia refractaria con exceso de blastos o de un síndrome 5q.

20 En otras palabras, en otra realización ventajosa más, la invención se refiere a un procedimiento definido anteriormente, en el que dicho grupo comprende 10 genes, comprendiendo o estando constituidos dichos 10 genes secuencias de ácido nucleico de SEQ ID NO: 1 a 10, y la etapa c) es de tal manera que

- si la relación establecida en la etapa b) para cada uno de los genes de cualquier combinación de 3 genes de dicho subgrupo es ≥ 2 o $\leq 0,5$,

25 a condición de que si la combinación de al menos 3 genes que corresponden a los genes que comprenden o están constituidos por la secuencia de ácido nucleico de SEQ ID NO: 1, 2 y 3, en la que

- 30 - la relación R1 es $\leq 0,3$ para el gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1, y
- las relaciones R1 son $\geq 3,0$, para los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 y 3,

esta combinación se excluye,
y además

- 35 - si la relación establecida en la etapa b) para al menos un gen de dicho grupo que no pertenece a dicho subgrupo es ≥ 2 , y la relación establecida en la etapa b) de al menos otro gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,5$,

entonces dicha muestra biológica es representativa de una anemia refractaria con exceso de blastos o de un síndrome 5q.

En esta realización ventajosa de la invención, los genes pertenecientes al grupo (es decir, los genes que comprenden o están constituidos por SEQ ID NO: 1 a 10) son útiles para discriminar los trastornos mielodisplásicos.

40 Más precisamente, los genes que pertenecen al grupo pero no pertenecen al subgrupo (es decir, los genes que comprenden o están constituidos por SEQ ID NO: 7 a 10 – D en la Figura 2) son útiles para discriminar los trastornos mielodisplásicos.

45 Entonces, si midiendo el nivel de expresión de los genes de SEQ ID NO: 1-6, la muestra biológica se considera que es representativa de un trastorno mielodisplásico, es posible de acuerdo con la invención separar la anemia refractaria con exceso de blastos y un síndrome 5q de otras patologías, cuando la relación de la expresión de al menos un gen del grupo que consiste en SEQ ID NO: 7 a 10 es ≥ 2 y cuando la relación de la expresión de al menos un gen del grupo que consiste en SEQ ID NO: 7 a 10 es $\leq 0,5$.

50 Otra realización ventajosa de la invención se refiere a un procedimiento definido anteriormente, en el que dicho grupo comprende 10 genes, que comprenden o están constituidos dichos 10 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 a 10, y la etapa c) es de tal manera que

- si la relación establecida en la etapa g) para cada los genes de cualquier combinación de 3 genes de dicho

subgrupo es ≥ 2 o $\leq 0,5$,

a condición de que

- 5
- la relación entre el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1 medida en dicha muestra biológica y medida en dicha muestra de control, no es $\leq 0,3$, o
 - las relaciones entre el nivel de expresión de cada uno de los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 o 3 medidos en dicha muestra biológica y medida en dicha muestra de control, no es ≥ 3 ,

y

- 10
- si la relación establecida en la etapa b) para al menos un gen de dicho grupo que no pertenece a dicho subgrupo es ≥ 2 y la relación establecida en la etapa b) de al menos otro gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,5$,

y además

- 15
- si la relación establecida en la etapa b) para al menos 4 genes del grupo de 24 genes que no pertenecen a dicho grupo es ≥ 3 , entonces la muestra biológica es representativa de una anemia refractaria con exceso de blastos.

En otras palabras, otra realización ventajosa de la invención se refiere a un procedimiento definido anteriormente, en el que

dicho grupo comprende 10 genes, comprendiendo o estando constituidos dichos 10 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 a 10, y la etapa c) es de tal manera que

- 20
- si la relación establecida en la etapa b) para cada uno de los genes de cualquier combinación de 3 genes de dicho subgrupo es ≥ 2 o $\leq 0,5$,

a condición de que la combinación de al menos 3 genes se corresponda con los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1, 2 y 3, en la que

- 25
- la relación R_i es $\leq 0,3$ para el gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1, y
 - las relaciones R_i son $\geq 3,0$, para los genes que comprenden o están constituidas por las secuencias de ácido nucleico de SEQ ID NO: 2 y 3,

esta combinación se excluye,

y

- 30
- si la relación establecida en la etapa b) para al menos un gen de dicho grupo que no pertenece a dicho subgrupo es ≥ 2 , y la relación establecida en la etapa b) de al menos otro gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,5$,

y además

- 35
- si la relación establecida en la etapa b) para al menos 4 genes del grupo de 24 genes que no pertenecen a dicho grupo es ≥ 3 ,

entonces dicha muestra biológica es representativa de una anemia refractaria con exceso de blastos.

En esta ventajosa realización de la invención, los genes que pertenecen al grupo (es decir, genes que comprenden o están constituidos por SEQ ID NO: 1 a 24) son útiles para discriminar entre anemia refractaria con exceso de blastos y síndrome 5q.

- 40
- Con más precisión, los genes que pertenecen al grupo pero no pertenecen al grupo (es decir los genes que comprenden o están constituidos por SEQ ID NO: 11 a 24 – E en la Figura 2) son útiles para la discriminación entre anemia refractaria con exceso de blastos y un síndrome 5q.

- 45
- La invención también se refiere a un procedimiento para el diagnóstico, y/o clasificación, preferentemente *in vitro*, de un trastorno hematológico, en particular un trastorno hematológico mieloide y/o linfoide, preferentemente un trastorno hematológico mieloide, comprendiendo dicho procedimiento las etapas de:

- 50
- a) medir, a partir de las células contenidas en una muestra biológica de un sujeto, preferentemente de células sanguíneas o células de médula ósea que contiene la muestra, el nivel de expresión de al menos los genes de un subgrupo de 6 genes que pertenecen a un grupo de genes escogidos de entre un grupo de 24 genes, comprendiendo o estando constituido dicho grupo de 24 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 a 24, perteneciendo dichos 6 genes a dicho subgrupo que comprende o está constituido por las secuencias de ácido nucleico de SEQ ID NO: 1 a 6.

b) comparar el nivel de expresión de cada uno de los genes medidos en la etapa a), con el nivel de expresión de los mismos genes respectivos de células contenidas en una muestra de control, preferentemente de una muestra de control que contiene células sanguíneas o células de médula ósea, siendo dicha muestra de control de la misma naturaleza que dicha muestra biológica, para establecer la relación de nivel de expresión R_i entre el nivel de expresión de cada uno de los genes medidos en la etapa a) y el nivel de expresión de los mismos genes respectivos de células contenidas en una muestra de control para cada uno de los genes de dicho subgrupo, y c) determinar el estado de dicha muestra biológica de manera que si la relación R_i para cada uno de los genes de cualquier combinación de 3 genes de entre dicho subgrupo es

- $0 \geq 2$,
- $0 \leq 0,5$,

dicha muestra biológica es representativa de células de un trastorno hematológico.

Una realización ventajosa de la invención se refiere a un procedimiento que se ha definido anteriormente, en el que si

- la relación R_i es $\leq 0,3$ para el gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1, y
- las relaciones R_i son $\geq 3,0$, para los genes que comprenden o están constituidos por secuencias de ácido nucleico de SEQ ID NO: 2 y 3,

entonces dicha muestra biológica es representativa de una leucemia mieloide aguda.

En esta realización ventajosa del procedimiento de acuerdo con la invención, en la etapa c) cuando la relación R_i del nivel de expresión de los genes GPX3 (SEQ ID NO: 1) $\leq 0,3$, la relación del nivel de expresión de los genes GPX1(1) (SEQ ID NO: 2) ≥ 3 y la relación del nivel de expresión de los genes GLRX2(2) (SEQ ID NO: 3) es $\geq 3,0$, entonces la muestra biológica es representativa de una leucemia mieloide aguda (AML).

En otra realización ventajosa, la invención se refiere a un procedimiento que se define anteriormente, en el que la etapa c) es de tal manera que

si la relación R_i establecida en la etapa b) para cada uno de los genes de cualquier combinación de al menos 3 genes de dicho subgrupo es ≥ 2 o $\leq 0,5$,

a condición de que si la combinación de los 3 genes se corresponde con los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1, 2 y 3, en las que

- la relación R_i es $\leq 0,3$ para el gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1, y
- las relaciones R_i son ≥ 3 para los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 y 3,

dicha combinación se excluye

entonces dicha muestra biológica es representativa de un trastorno mielodisplásico, en particular mielodisplasia que se escoge de entre anemia refractaria con sideroblastos anulares (RARS), citopenia refractaria con displasia multilineaje (RCMD), anemia refractaria con exceso de blastos (RAEB) o síndrome 5q y mielodisplasia inclasificable.

En otras palabras, cuando la relación R_i establecida en la etapa c) del nivel de expresión de cada gen o cualquier combinación de al menos 3 genes escogidos de entre los genes que comprenden o están constituidos por SEQ ID NO: 1-6 es ≥ 2 o $\leq 0,5$, y dicha combinación no se corresponde con la combinación que define una muestra biológica como representativa de una AML, entonces dicha muestra es representativa de un trastorno mielodisplásicos o síndrome 5q.

Otra realización ventajosa de la invención se refiere a un procedimiento definido anteriormente, en el que dicho grupo comprende 10 genes, comprendiendo o estando constituidos dichos 10 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 to 10,

y además, en la etapa c)

si la relación R_i de los genes de dicho grupo que no pertenecen a dicho subgrupo está comprendida entre 0,3 a 2, siendo excluidos los extremos del intervalo

entonces dicha muestra biológica es representativa de una anemia refractaria con sideroblastos anulares o una citopenia refractaria con displasia multilineaje.

En la invención "los genes de dicho grupo que no pertenecen a dicho subgrupo" se corresponde con los genes que pertenecen al grupo D que se define en la Figura 2.

Como el subgrupo consiste en los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1-6 y el grupo consiste en los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1-10, en consecuencia, los genes de dicho grupo que no pertenecen a dicho subgrupo corresponden a los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO:

7-10.

5 En la realización anterior, si la relación R_i de cada gen representado por las secuencias de ácido nucleico de SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9 y SEQ ID NO: 10, es decir todas las relaciones de dichos genes, está comprendida entre 0,3 y 2, estando excluidos 0,3 y 2 del intervalo, dicha muestra biológica es representativa de una anemia refractaria con sideroblastos anulares o una citopenia refractaria con displasia multilineaje.

Es posible decir que si $0,3 < R_7 < 2$, y $0,3 < R_8 < 2$, y $0,3 < R_9 < 2$, y $0,3 < R_{10} < 2$, R_7 , R_8 , R_9 y R_{10} representan la relación respectiva para los genes representados por SEQ ID NO: 7, 8, 9 y 10, además de la evaluación de la relación de los genes de SEQ ID NO: 1-6, entonces dicha muestra biológica es representativa de una anemia refractaria con sideroblastos anulares o una citopenia refractaria con displasia multilineaje.

10 Las expresiones “comprendido entre 0,3 a 2, estando excluidos los extremos del intervalo” se refiere al intervalo representado por el símbolo matemático $]0,3; 2[$. Este intervalo incluye todos los valores comprendidos entre 0,3 y 2, pero excluyendo los valores específicos 0,3 y 2.

Si $R_i = 0,3$, o $R_i = 2$, R_i no pertenece al intervalo $]0,3; 2[$.

15 La evaluación de la relación del nivel de expresión de los genes representados por SEQ ID NO: 7-10, en un intervalo específica, permite discriminar algunos síndromes mielodisplásicos.

20 Para resumir, si la relación de cada gen de una combinación de 3 genes escogidos de entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5, y 6 es ≥ 2 o $\leq 0,5$, excluyendo la combinación que define leucemia, y la relación de cada gen representado por SEQ ID NO: 7, 8, 9 y 10 está comprendido en el intervalo $]0,3; 2[$, entonces la muestra biológica es representativa de una anemia refractaria con sideroblastos anulares o una citopenia refractaria con displasia multilineaje.

En otra realización ventajosa, la invención se refiere a un procedimiento de acuerdo con la definición que se ha mencionado anteriormente, en el que además, en la etapa c)

- si la relación R_i de al menos un gen que pertenece a dicho grupo de genes que no pertenecen a dicho grupo es $\leq 0,3$, entonces dicha muestra biológica es representativa de anemia refractaria con sideroblastos anulares, y
- 25 - si la relación R_i de al menos un gen que pertenece a dicho grupo de genes que no pertenecen a dicho grupo es $\geq 3,0$,

entonces dicha muestra biológica es representativa de citopenia refractaria con displasia multilineaje.

En la invención “un gen que pertenece a dicho grupo de genes que no pertenecen a dicho grupo” se corresponde con los genes que pertenecen al grupo E como se define en la Figura 2.

30 Como el subgrupo consiste en los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1-6, el grupo consiste en los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1-10. y el grupo consiste en los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1-24, en consecuencia, los genes de dicho grupo que no pertenecen a dicho grupo corresponden a los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 11-24.

35 En la realización anterior, los genes representados por SEQ ID NO: 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24 proporcionan información suplementaria con respecto a la naturaleza de la muestra biológica ensayada.

Para resumir,

1)

40 a) si la relación para cada gen de cualquier combinación de 3 genes escogidos de entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5 y 6, es ≥ 2 o $\leq 0,5$, excluyendo la combinación que define la leucemia, y

b) la relación de cada gen representado por SEQ ID NO: 7, 8, 9 y 10 está comprendida en el intervalo $]0,3; 2[$, y

45 c) la relación de al menos 1 gen, es decir, los genes 1, o 2, o 3, o 4, o 5, o 6, o 7, o 8, o 9, o 10, o 11, o 12, o 13, o 14, representados por SEQ ID NO: 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24 es $\geq 3,0$, entonces dicha muestra biológica es representativa de una citopenia refractaria con displasia multilineaje, y

2)

50 a) si la relación para cada gen de cualquier combinación de 3 genes escogidos de entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5, y 6, es ≥ 2 o $\leq 0,5$, excluyendo la combinación que define leucemia, y

b) la relación de cada gen representado por SEQ ID NO: 7, 8, 9 y 10 está comprendida en el intervalo]0,3; 2[, y

c) la relación de al menos 1 gen, es decir los genes 1, o 2, o 3, o 4, o 5, o 6, o 7, o 8, o 9, o 10, o 11, o 12, o 13, o 14, representados por SEQ ID NO: 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24 es $\leq 0,3$, entonces dicha muestra biológica es representativa de una anemia refractaria con sideroblastos anulares.

En otra realización, la invención se refiere a un procedimiento definido anteriormente, en el que dicho grupo comprende 10 genes, comprendiendo o estando constituidos dichos 10 genes por secuencias de ácidos nucleicos SEQ ID NO: 1 a 10,

y además

si

- la relación R_i de al menos un gen de dicho grupo no pertenece a dicho subgrupo es ≥ 2 , o
- la relación R_i de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,3$, o
- la relación R_i de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es ≥ 2 y la relación R_i de otro gen que no pertenece a dicho subgrupo es $\leq 0,3$,

entonces dicha muestra biológica es representativa de escogerse entre citopenia refractaria con displasia multilineaje, anemia refractaria con exceso de blastos, síndrome 5q y mielodisplasia sin clasificar.

En esta realización ventajosa, se tiene en cuenta el caso en el que al menos un gen representado por SEQ ID NO: 7, 8, 9 o 10 no pertenecen al intervalo]0,3; 2[como se ha definido anteriormente.

En un primer caso, al menos un gen tiene una relación ≥ 2 , independientemente de la relación de los otros genes. En un segundo caso al menos un gen tiene una relación $\leq 0,3$, independientemente de la relación de los otros genes. En un tercer caso, ambos al menos un gen tiene una relación ≥ 2 , y otro gen tiene una relación $\leq 0,3$, independientemente de la relación de los otros genes.

En otra realización ventajosa, la invención se refiere a un procedimiento definido anteriormente, en el que si la relación R_i de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es ≥ 2 y la relación de al menos otro gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,5$, entonces dicha muestra biológica es representativa de anemia refractaria con exceso de blastos o síndrome 5q.

Esta realización se refiere al caso en el que al menos un gen tiene una relación ≥ 2 , independientemente de la relación de los otros genes, y en particular el caso en el que al menos un gen tiene una relación ≥ 2 y otro gen tiene una relación $\leq 0,5$, y por lo tanto posiblemente $\leq 0,3$.

Esta situación particular permite detectar, o identificar, síndromes mielodisplásicos que son representativos de anemia refractaria con exceso de blastos o de síndrome 5q.

En otra realización ventajosa, la invención se refiere a un procedimiento definido anteriormente, en el que si la relación R_o de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es ≥ 2 y la relación R_i de al menos otro gen de dicho grupo que no pertenece a dicho subgrupo es $< 0,5$ entonces dicha muestra biológica es representativa de anemia refractaria con exceso de blastos o síndrome 5q, y además

- si las relaciones R_i de al menos cuatro genes que pertenecen a dicho grupo de genes que no pertenecen a dicho grupo son $> 3,0$, entonces dicha muestra biológica es representativa de una anemia refractaria o con exceso de blastos, y
- si las relaciones R_i de al menos tres genes que pertenecen a dicho grupo de genes que no pertenecen a dicho grupo son $> 3,0$, entonces dicha muestra biológica es representativa de un síndrome 5q.

Para resumir,

1)

a) si la relación para cada uno de los genes de cualquier combinación de 3 genes escogidos de entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5 y 6, es > 2 o $< 0,5$ excluyendo la combinación que define leucemia, y

b) la relación de al menos un gen representado por SEQ ID NO: 7, 8, 9 y 10 es > 2 y al menos otro gen representado por SEQ ID NO: 7, 8, 9 y 10 es $< 0,5$, y

c) las relaciones de al menos 4 genes, es decir, 4, o 5, o 6, o 7, o 8, o 9, o 10, o 11, o 12, o 13, o 14 genes, representados por SEQ ID NO: 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24 son $> 3,0$, entonces la muestra biológica es representativa de anemia refractaria con exceso de blastos, y

2)

a) si la relación para cada gen de cualquier combinación de 3 genes escogidos de entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5 y 6, es > 2 o $< 0,5$, excluyendo la combinación que define leucemia, y

b) la relación de al menos un gen representado por SEQ ID NO: 7, 8, 9 y 10 es > 2.0 , y al menos otro gen representado por SEQ ID NO: 7, 8, 9 y 10 es $< 0,5$, y

c) la relación de al menos 3 genes, es decir, 0, o 1, o 2, o 3 genes, representados por SEQ ID NO: 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24 son $> 3,0$, entonces dicha muestra biológica es representativa de un síndrome 5q.

5 En otra realización ventajosa, la invención se refiere a un procedimiento de acuerdo con la definición anterior, en el que si

- la relación R_i de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es > 2 , y
- la relación R_i de ningún gen de dicho grupo que no pertenece a dicho subgrupo es $< 0,5$,

10 entonces dicha muestra biológica es representativa de una citopenia refractaria con displasia multilineaje.

En esta realización específica si la relación de al menos un gen representado por SEQ ID NO: 7, 8, 9 y 10 es > 2 y la relación de los otros genes está incluida en el intervalo $[0,5; +\infty)$, entonces dicha muestra biológica es representativa de una citopenia refractaria con displasia multilineaje.

Para resumir:

- 15 a) si la relación de cada gen de cualquier combinación de 3 genes escogidos de entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5 y 6, es > 2 o $< 0,5$, excluyendo la combinación que define leucemia, y
- b) la relación de al menos un gen representado por SEQ ID NO: 7, 8, 9 y 10 es > 2 y la relación de otro gen representado por SEQ ID NO: 7, 8, 9 y 10 es $< 0,5$, entonces dicha muestra biológica es representativa de una citopenia refractaria con displasia multilineaje.

20 En otra realización ventajosa, la invención se refiere a un procedimiento de acuerdo con la definición anterior, en el que además si

- la relación R_i de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es $< 0,5$ y
- la relación R_i de ningún gen de dicho grupo que no pertenece a dicho subgrupo es > 2 ,

25 entonces dicha muestra biológica es representativa de anemia refractaria con sideroblastos anulares, citopenia refractaria con displasia multilineaje, anemia refractaria con exceso de blastos, síndrome 5q, o mielodisplasia sin clasificar.

En esta realización específica, si la relación de al menos un gen representado por SEQ ID NO: 7, 8, 9 o 10 es $< 0,5$ y la relación de los otros genes está incluida en el intervalo $]-\infty; 2]$, entonces dicha muestra biológica es representativa de anemia refractaria con sideroblastos anulares, citopenia refractaria con displasia multilineaje, anemia refractaria con exceso de blastos, síndrome 5q o mielodisplasia sin clasificar.

30 En esta realización específica, si la relación de al menos un gen representado por SEQ ID NO: 7, 8, 9 o 10 es $< 0,5$ y la relación de los otros genes está incluida en el intervalo $]-\infty; 2]$, entonces dicha muestra biológica es representativa de anemia refractaria con sideroblastos anulares, citopenia refractaria con displasia multilineaje, anemia refractaria con exceso de blastos, síndrome 5q o mielodisplasia sin clasificar.

Para resumir:

- 35 a) si la relación para cada gen de cualquier combinación de 3 genes escogidos de entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5 y 6, es ≥ 2 o $\leq 0,5$, excluyendo la combinación que define leucemia, y
- b) la relación de al menos un gen representado por SEQ ID NO: 7, 8, 9 y 10 es $\leq 0,5$, y la relación de otro gen representado por SEQ ID NO: 7, 8, 9 y 10 es ≥ 2 , entonces dicha muestra biológica es representativa de anemia refractaria con sideroblastos anulares, citopenia refractaria con displasia multilineaje, anemia refractaria con exceso de blastos, síndrome 5q o mielodisplasia sin clasificar.

En otra realización ventajosa, la invención se refiere a un procedimiento como se ha definido anteriormente, en el que la etapa c) si

- 40 - la relación R_i de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,5$ y
- la relación R_i de ningún gen de dicho grupo que no pertenece a dicho subgrupo es ≥ 2 ,

y además si

- 45 - las relaciones R_i de todos los genes que pertenecen a dicho grupo de genes que no pertenecen a dicho grupo están comprendidos entre 0,5 a 2, excluyendo los extremos del intervalo,

entonces dicha muestra biológica es representativa de mielodisplasia sin clasificar.

En esta ventajosa realización los genes representados por SEQ ID NO: 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24 son útiles para discriminar la mielodisplasia sin clasificar de la anemia refractaria con sideroblastos anulares, citopenia refractaria con displasia multilineaje y síndrome 5q.

50 En otra realización ventajosa, la invención se refiere a un procedimiento como se ha definido anteriormente, en el que,

en la etapa c) si

- la relación R_i de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,3$, y
- la relación R_i de ningún gen de dicho grupo que no pertenece a dicho subgrupo es ≥ 2 ,

y además

5 si

- las relaciones R_i de al menos un gen que pertenece a dicho grupo de genes que no pertenecen a dicho grupo es ≥ 2 , entonces dicha muestra biológica es representativa de síndrome 5q, o citopenia refractaria con displasia multilinaje.

Para resumir:

- 10 a) si la relación de cada gen de cualquier combinación de 3 genes escogidos de entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5 y 6, es ≥ 2 o $\leq 0,5$, excluyendo la combinación que define la leucemia, y
- b) las relaciones R_i de al menos un gen representado por SEQ ID NO: 7, 8, 9 y 10 es ≥ 2 , y
- c) las relaciones R_i de al menos un gen representado por SEQ ID NO: 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24 es ≥ 2 ,

15 entonces dicha muestra biológica es representativa de síndrome 5q o citopenia refractaria con displasia multilinaje.

Otra realización ventajosa de la invención se refiere a un procedimiento como se define anteriormente, en el que si la relación R_i de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,3$ y la relación R_i de ningún gen que no pertenece a dicho subgrupo es ≥ 2 , entonces dicha muestra biológica es representativa de una mielodisplasia sin clasificar, un síndrome 5q, o de una citopenia refractaria con displasia multilinaje.

20

Para resumir:

- 25 a) si la relación de cada gen de cualquier combinación de 3 genes escogidos entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5 y 6, es ≥ 2 o $\leq 0,5$, excluyendo la combinación que define la leucemia, y
- b) la relación de al menos un gen representado por SEQ ID NO: 7, 8, 9 y 10 es $\leq 0,3$, y la relación de ningún gen representado por SEQ ID NO: 7, 8, 9 y 10 es ≥ 2 , dicha muestra biológica es representativa de una mielodisplasia sin clasificar, de un síndrome 5q o de una citopenia refractaria con displasia multilinaje.

Otra realización ventajosa de la invención se refiere a un procedimiento como se ha definido anteriormente, en el que si la relación R_i de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,3$ y la relación R_i de ningún gen que no pertenece a dicho subgrupo es ≥ 2 ,

30

y además

si las relaciones R_i de al menos dos genes de dicho grupo que no pertenecen a dicho subgrupo son $\leq 0,3$ y la relación R_i de ningún gen de dicho grupo que no pertenece a dicho subgrupo es ≥ 2 , entonces dicha muestra biológica es representativa de un síndrome 5q o de citopenia refractaria con displasia multilinaje.

Para resumir:

- 35 a) si la relación de cada uno de los genes de cualquier combinación de 3 genes escogidos de entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5 y 6, es ≥ 2 o $\leq 0,5$, excluyendo la combinación que define la leucemia, y
- b) la relación de al menos un gen representado por SEQ ID NO: 7, 8, 9 y 10 es $\leq 0,3$, la relación de al menos otro gen representado por SEQ ID NO: 7, 8, 9 y 10 es ≥ 2 , dicha muestra biológica es representativa de un síndrome 5q o de una citopenia refractaria con displasia multilinaje.

40

Otra realización ventajosa de la invención se refiere a un procedimiento como se ha definido anteriormente, en el que

si la relación R_i de al menos un gen de dicho grupo que no pertenece a dicho subgrupo es $\leq 0,3$ y la relación R_i de ningún gen que no pertenece a dicho subgrupo es ≥ 2 ,

45

y

si las relaciones R_i de al menos dos genes de dicho grupo que no pertenece a dicho subgrupo son $\leq 0,3$ y la relación R_i de ningún gen de dicho grupo que pertenece a dicho subgrupo es ≥ 2 ,

entonces dicha muestra biológica es representativa de un síndrome 5q, o de citopenia refractaria con displasia multilinaje.

50

y

- si además las relaciones R_i de al menos dos genes que pertenecen a dicho grupo de genes que no pertenecen a dicho grupo son ≥ 2 , entonces dicha muestra biológica es representativa de un síndrome 5q.

Para resumir:

- 5 a) si la relación de cada uno de los genes de cualquier combinación de 3 genes escogidos de entre los 6 genes representados por SEQ ID NO: 1, 2, 3, 4, 5 y 6 es ≥ 2 o $\leq 0,5$, excluyendo la combinación que define leucemia, y
 b) la relación de al menos un gen representado por SEQ ID NO: 7, 8, 9 y 10 es $\leq 0,3$, la relación de al menos otro gen representado por SEQ ID NO: 7, 8, 9 y 10 es ≥ 2 , y
 c) las relaciones de al menos dos genes representados por SEQ ID NO: 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23 y 24 son ≥ 2 ,

dicha muestra biológica es representativa de un síndrome 5q.

10 La invención también se refiere a un procedimiento para el diagnóstico y/o la clasificación, preferentemente *in vitro*, de un trastorno hematológico, en particular un trastorno hematológico mielóide y/o linfóide, preferentemente un trastorno hematológico mielóide, dicho procedimiento comprende las etapas de:

- 15 a) medir a partir de las células contenidas en una muestra biológica de un sujeto, preferentemente a partir de células sanguíneas o células de médula ósea que contiene la muestra, el nivel de expresión de 24 genes, comprendiendo o estando constituidos dichos 24 genes por las secuencias de ácido nucleico de SEQ ID NO:1 a 24,
 b) comparar el nivel de expresión de cada uno de los genes medidos en la etapa a), con el nivel de expresión de los mismos genes respectivos de las células contenidas en una muestra de control preferentemente de células sanguíneas o células de médula ósea contenidas en la muestra, siendo dicha muestra de control de la misma naturaleza que dicha muestra biológica, para establecer una relación del nivel de expresión genética para cada uno de los genes de dicho subgrupo, y
 20 c) determinar el estado de dicha muestra biológica tal como si la relación establecida en la etapa b).

Más ventajosamente, la invención se refiere a un procedimiento definido anteriormente, en el que el nivel de expresión de los genes se mide por un procedimiento que permite la determinación de la cantidad del ARNm o del ADNc correspondiente a dichos genes. Preferentemente dicho procedimiento es un procedimiento cuantitativo.

25 En una realización, la invención se refiere a un procedimiento definido anteriormente, en el que el nivel de expresión de los genes se mide por un procedimiento cuantitativo, en particular RT-qPCR.

30 Los niveles de ARNm se pueden medir cuantitativamente por transferencia de Northern que da el tamaño e información de secuencia acerca de las moléculas de ARNm. Una muestra de ARN se separa en un gel de agarosa y se hibrida con una sonda de ARN radiomarcada que es complementaria a la secuencia diana. El ARN radiomarcado se detecta entonces con una autorradiografía. La transferencia de Northern se utiliza ampliamente como información del tamaño de ARNm que permite la discriminación de transcripciones de corte y empalme alternadamente.

35 Otra estrategia para medir la abundancia de ARNm es la reacción en cadena de la polimerasa cuantitativa (RT-PCR seguido por qPCR). La RT-PCR genera primero una matriz de ADN a partir del ARNm por transcripción inversa, que se llama ADNc. Esta matriz de ADNc se utiliza entonces por qPCR donde el cambio de fluorescencia de una sonda cambia según progresa el procedimiento de amplificación de ADN. Con una curva de referencia construida cuidadosamente qPCR se puede producir una medición absoluta tal como el número de copias de ARNm, normalmente en unidades de copias por nanolitro de tejido homogeneizado o copias por célula. La qPCR es muy sensible (es posible la detección de una única molécula de ARNm), pero puede ser cara debido a la necesidad de sondas fluorescentes.

40 Las transferencias de Northern y la RT-qPCR son buenas para detectar si se expresa un único gen o unos pocos genes.

45 Otros procedimientos conocidos por un experto en la técnica incluyen micromatrices de ADN o tecnologías tipo Análisis en Serie de Expresión Genética (SAGE).

El SAGE puede proporcionar una medición relativa de la concentración celular de diferentes ARN mensajeros. La gran ventaja de los procedimientos basados en marcadores es la "arquitectura abierta", permitiendo la medición exacta de cualquier transcripción presente en las células, la secuencia de dichas transcripciones puede ser conocida o desconocida.

50 El procedimiento preferido utilizado de acuerdo con la invención es RT-qPCR.

55 En otra realización ventajosa más, la invención se refiere a un procedimiento definido anteriormente, en el que la medición del nivel de expresión de los genes se lleva a cabo utilizando los al menos 6 pares de oligonucleótidos que pertenecen a un grupo de 24 pares de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 25 a 72, comprendiendo o estando constituidos dichos al menos 6 pares de oligonucleótidos por las secuencias de ácido nucleico de SEQ ID NO: 25 a 36.

En una realización, la invención se refiere a un procedimiento definido anteriormente, en el que la medición del nivel de expresión de al menos dichos 6 genes de la etapa a) se lleva a cabo utilizando al menos las moléculas de nucleótido de SEQ ID NO: 25-36.

En la invención:

- 5 - el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 25 y SEQ ID NO: 26, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 27 y SEQ ID NO: 28, se utilizan para medir el nivel de expresión del gen que comprende o está constituido
- 10 por la secuencia de ácido nucleico de SEQ ID NO: 2,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 29 y SEQ ID NO: 30, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 3,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 31 y SEQ ID NO: 32, se utilizan para medir el nivel de expresión del gen que comprende o está constituido
- 15 por la secuencia de ácido nucleico de SEQ ID NO: 4,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 33 y SEQ ID NO: 34, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 5,
- 20 - el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 35 y SEQ ID NO: 36, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 6.

En una realización ventajosa, la invención se refiere al procedimiento que se ha definido anteriormente,

- 25 utilizando los al menos 10 pares de oligonucleótidos que pertenecen a un grupo de 24 pares de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 25 a 72, comprendiendo o estando constituidos dichos al menos 6 pares de oligonucleótidos por las secuencias de ácido nucleico de SEQ ID NO: 25 a 44.

En la invención:

- 30 - el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 36 y SEQ ID NO: 37, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 7,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 38 y SEQ ID NO: 39, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 8,
- 35 - el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 41 y SEQ ID NO: 42, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 9,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 43 y SEQ ID NO: 44, se utilizan para medir el nivel de expresión del gen que comprende o está constituido
- 40 por la secuencia de ácido nucleico de SEQ ID NO: 10.

En una realización ventajosa, la invención se refiere al procedimiento que se ha definido anteriormente, utilizando 24 pares de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 25 a 72.

En la invención:

- 45 - el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 45 y SEQ ID NO: 46, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 11,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 47 y SEQ ID NO: 48, se utilizan para medir el nivel de expresión del gen que comprende o está constituido
- 50 por la secuencia de ácido nucleico de SEQ ID NO: 12,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 49 y SEQ ID NO: 50, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 13,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 51 y SEQ ID NO: 52, se utilizan para medir el nivel de expresión del gen que comprende o está constituido
- 55 por la secuencia de ácido nucleico de SEQ ID NO: 14,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 53 y SEQ ID NO: 54, se utilizan para medir el nivel de expresión del gen que comprende o está constituido

- por la secuencia de ácido nucleico de SEQ ID NO: 15,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 55 y SEQ ID NO: 56, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 16,
- 5 - el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 57 y SEQ ID NO: 58, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 17,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 59 y SEQ ID NO: 60, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 18,
- 10 - el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 61 y SEQ ID NO: 62, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 19,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 63 y SEQ ID NO: 64, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 20,
- 15 - el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 65 y SEQ ID NO: 66, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 21,
- 20 - el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 67 y SEQ ID NO: 68, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 22,
- el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 69 y SEQ ID NO: 70, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 23,
- 25 - el par de oligonucleótidos que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 71 y SEQ ID NO: 72, se utilizan para medir el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 24,

La Tabla 3 siguiente representa los genes/variantes/oligonucleótidos que se utilizan de acuerdo con la invención.

30

Tabla 3

Nombre del gen	SEQ ID	SEQ ID variante	SEQ ID Oligo directo	SEQ ID Oligo inverso
GPX3	SEQ ID NO: 1	--	SEQ ID NO: 25	SEQ ID NO: 26
GPX1 (1)	SEQ ID NO: 2	--	SEQ ID NO: 27	SEQ ID NO: 28
GLRX2 (2)	SEQ ID NO:3	--	SEQ ID NO: 29	SEQ ID NO: 30
CAT	SEQ ID NO: 4	--	SEQ ID NO: 31	SEQ ID NO: 32
PRDX (1-2-3)	SEQ ID NO:5	SEQ ID NO: 73 o SEQ ID NO: 74	SEQ ID NO: 33	SEQ ID NO: 34
PRDX5 (2)	SEQ ID NO:6	--	SEQ ID NO: 35	SEQ ID NO: 36
SOD2 (1-2-3)	SEQ ID NO:7	SEQ ID NO: 75 o SEQ ID NO: 76	SEQ ID NO: 37	SEQ ID NO: 38
GSR (1-2-3-4)	SEQ ID NO:8	SEQ ID NO: 77 o SEQ ID NO: 78 o SEQ ID NO: 79	SEQ ID NO: 39	SEQ ID NO: 40
GLRX (1-2)	SEQ ID NO:9	SEQ ID NO: 80	SEQ ID NO: 41	SEQ ID NO: 42
PRDX2 (1)	SEQ ID NO: 10	--	SEQ ID NO: 43	SEQ ID NO: 44
PRDX5 (1-3)	SEQ ID NO: 11	SEQ ID NO: 81 o SEQ ID NO: 82	SEQ ID NO: 45	SEQ ID NO: 46

(continuación)

Nombre del gen	SEQ ID	SEQ ID variante	SEQ ID Oligo directo	SEQ ID Oligo inverso
SOD1	SEQ ID NO: 12	--	SEQ ID NO: 47	SEQ ID NO: 48
TXN	SEQ ID NO: 13	--	SEQ ID NO: 49	SEQ ID NO: 50
PRDX3 (1-2)	SEQ ID NO: 14	SEQ ID NO: 83	SEQ ID NO: 51	SEQ ID NO: 52
GPX7	SEQ ID NO: 15	--	SEQ ID NO: 53	SEQ ID NO: 54
GPX4 (1-2-3)	SEQ ID NO: 16	SEQ ID NO: 84 o SEQ ID NO: 85	SEQ ID NO: 55	SEQ ID NO: 56
TXN2	SEQ ID NO: 17	--	SEQ ID NO: 57	SEQ ID NO: 58
PRDX4	SEQ ID NO: 18	--	SEQ ID NO: 59	SEQ ID NO: 60
GPX1 (2)	SEQ ID NO: 19	--	SEQ ID NO: 61	SEQ ID NO: 62
GLRX3	SEQ ID NO: 20	--	SEQ ID NO: 63	SEQ ID NO: 64
PRDX2 (3)	SEQ ID NO: 21	--	SEQ ID NO: 65	SEQ ID NO: 66
PRDX6	SEQ ID NO: 22	--	SEQ ID NO: 67	SEQ ID NO: 68
GLRX5	SEQ ID NO: 23	--	SEQ ID NO: 69	SEQ ID NO: 70
GLRX2 (1)	SEQ ID NO: 24	--	SEQ ID NO: 71	SEQ ID NO: 72
GPX3	SEQ ID NO: 1	--	SEQ ID NO: 25	SEQ ID NO: 26
GPX1 (1)	SEQ ID NO: 2	--	SEQ ID NO: 27	SEQ ID NO: 28
GLRX2 (2)	SEQ ID NO: 3	--	SEQ ID NO: 29	SEQ ID NO: 30
CAT	SEQ ID NO:4	--	SEQ ID NO: 31	SEQ ID NO:32
PRDX (1-2-3)	SEQ ID NO:5	SEQ ID NO: 73 o SEQ ID NO: 74	SEQ ID NO: 33	SEQ ID NO: 34
PRDX5 (2)	SEQ ID NO:6	--	SEQ ID NO: 35	SEQ ID NO: 36
SOD2 (1-2-3)	SEQ ID NO:7	SEQ ID NO: 75 o SEQ ID NO: 76	SEQ ID NO: 37	SEQ ID NO: 38

Tabla 3

Los oligonucleótidos definidos anteriormente se utilizan preferentemente para llevar a cabo la reacción qPCR.

- 5 La qPCR es bien conocida en la técnica, y se puede llevar a cabo utilizando, en asociación con oligonucleótidos que permiten una amplificación específica del gen diana, con colorantes o con una sonda indicadora.

Ambas técnicas se resumen brevemente a continuación.

- PCR en tiempo real con colorantes unidos al ADN de doble cadena como indicadores:

5 Se une un colorante de unión a ADN a todo el ADN de doble cadena (ds) en una PCR, produciendo una fluorescencia del colorante. Un aumento en el producto del ADN durante la PCR da lugar por lo tanto a un aumento de la intensidad de la fluorescencia y se mide en cada ciclo, permitiendo de esta manera que se cuantifiquen las concentraciones de ADN.

Sin embargo, los colorantes de ADNds tales como el Verde SYBR se unirá a todos los productos ADNds de la PCR, incluyendo los productos de la PCR no específicos (tales como el dímero cebador). Esto puede interferir potencialmente o evitar la cuantificación exacta de la secuencia diana que se pretende.

10 La reacción se prepara como habitualmente, con la adición del colorante de ADNds fluorescente.

La reacción se ejecuta en un instrumento de PCR en tiempo real y después de cada ciclo, se miden los niveles de fluorescencia con un detector; el colorante solo es fluorescente cuando se une al ADNds (es decir, el producto de la PCR). En referencia a una dilución convencional, se puede determinar la concentración en la PCR.

15 Como en otros procedimientos de PCR en tiempo real, los valores que se obtienen no tienen unidades absolutas asociados con ellos (es decir, copias de ARNm/célula). Como se ha descrito anteriormente, la comparación de una medición de una muestra de ADN/ARN respecto a una dilución convencional solamente dará una fracción o relación de la muestra con respecto a la referencia, permitiendo solamente comparaciones relativas entre diferentes tejidos o condiciones experimentales. Para asegurar la exactitud de la cuantificación, habitualmente es necesario normalizar la expresión de un gen diana para un gen que se expresa establemente (véase posteriormente). Esto puede corregir
20 posibles diferencias en la cantidad o calidad de ARN a través de las muestras experimentales.

- Procedimiento de sonda indicadora fluorescente

25 Las sondas indicadoras fluorescentes detectan solamente el ADN que contiene la secuencia de la sonda; por lo tanto, el uso de la sonda indicadora aumenta significativamente la especificidad, y hace posible la cuantificación incluso en presencia de amplificaciones de ADN no específicas. Las sondas fluorescentes se pueden utilizar en ensayos múltiples para la detección de varios genes en la misma reacción basándose en sondas específicas con diferentes marcadores coloreados, a condición de que todos los genes dirigidos se amplifiquen con eficacia similar. La especificidad de las sondas indicadoras fluorescentes también evita la interferencia de mediciones producidas por los dímeros cebadores, que son productos colaterales potenciales no deseados en la PCR. Sin embargo, las sondas indicadoras fluorescentes no evitan el efecto inhibitorio de los dímeros cebadores, que pueden deprimir la
30 acumulación de productos deseados en la reacción.

35 El procedimiento se basa en una sonda basada en ADN con un indicador fluorescente en un extremo y un desactivador de fluorescencia en el extremo opuesto de la sonda. La estrecha proximidad del indicador al desactivador evita la detección de su fluorescencia; la descomposición de la sonda por la actividad exonucleasa de 5' a 3' de la polimerasa Taq rompe la proximidad indicador-desactivador y así permite la emisión no desactivada de fluorescencia, que se puede detectar tras la excitación con un láser. Un aumento en el producto dirigido por la sonda indicadora en cada ciclo de PCR por lo tanto, produce un aumento proporcional de la fluorescencia debido a la descomposición de la sonda y liberación del indicador.

La PCR se prepara como habitualmente, y se añade la sonda indicadora.

Durante el estadio de hibridación de la PCR tanto la sonda como los cebadores se hibridan al ADN diana.

40 La polimerización de una nueva cadena de ADN se inicia a partir de los cebadores, y una vez que la polimerasa alcanza la sonda, su exonucleasa 5'-3' degrada la sonda, separando físicamente el indicador fluorescente del desactivador, dando como resultado un aumento de la fluorescencia.

La fluorescencia se detecta y se mide en el termociclador de PCR en tiempo real y se utiliza su aumento geométrico correspondiente con el aumento exponencial del producto para determinar el umbral de ciclo (CT) en cada reacción.

45 En una realización particular, se consigue la medición del nivel de expresión de los genes que se han definido anteriormente, además de los oligonucleótidos específicos definidos anteriormente, utilizando una sonda disponible en el mercado. Cada gen para el que se espera un nivel de expresión se asocia con una sonda específica, una sonda que reconoce un gen no es capaz de reconocer otro gen. Además una sonda específica de un gen puede detectar también, cuando existen, variantes de dichos genes.

50 Las sondas ventajosas que se utilizan en la invención se enumeran en la Tabla 4.

La asociación entre gen/variante/oligonucleótidos y sondas se representan en la siguiente Tabla 4.

Tabla 4

Nombre del gen	SEQ ID	SEQ ID variante	SEQ ID Oligo directo	SEQ ID Oligo inverso	SEQ de sondas
GPX3	SEQ ID NO: 1	--	SEQ ID NO: 25	SEQ ID NO: 26	CCAGCC GC
GPX1 (1)	SEQ ID NO: 2	--	SEQ ID NO: 27	SEQ ID NO: 28	GGTGGT GG
GLRX2 (2)	SEQ ID NO:3	--	SEQ ID NO: 29	SEQ ID NO: 30	GGCGGC GG
CAT	SEQ ID NO:4	--	SEQ ID NO: 31	SEQ ID NO: 32	TGCTGG AG
PRDX (1-2-3)	SEQ ID NO: 5	SEQ ID NO: 73 o SEQ ID NO: 74	SEQ ID NO: 33	SEQ ID NO: 34	CTGGCT GG
PRDX5 (2)	SEQ ID NO: 6	--	SEQ ID NO: 35	SEQ ID NO: 36	GGAAGG AG
SOD2 (1-2-3)	SEQ ID NO: 7	SEQ ID NO: 75 o SEQ ID NO: 76	SEQ ID NO: 37	SEQ ID NO: 38	CTGCTG GG
GSR (1-2-3-4)	SEQ ID NO: 8	SEQ ID NO: 77 o SEQ ID NO: 78 o SEQ ID NO: 79	SEQ ID NO: 39	SEQ ID NO: 40	GCTGGA AG
GLRX (1-2)	SEQ ID NO: 9	SEQ ID NO: 80	SEQ ID NO: 41	SEQ ID NO: 42	GGTGGC TG
PRDX2 (1)	SEQ ID NO: 10	--	SEQ ID NO: 43	SEQ ID NO: 44	TGGGGA AG
PRDX5 (1-3)	SEQ ID NO: 11	SEQ ID NO: 81 o SEQ ID NO: 82	SEQ ID NO: 45	SEQ ID NO: 46	GGAAGG AG
SOD1	SEQ ID NO: 12	--	SEQ ID NO: 47	SEQ ID NO: 48	TGGGGA AG
TXN	SEQ ID NO: 13	--	SEQ ID NO: 49	SEQ ID NO: 50	CAGCAG CC
PRDX3 (1-2)	SEQ ID NO: 14	SEQ ID NO: 83	SEQ ID NO: 51	SEQ ID NO: 52	CTGCTTC C
GPX7	SEQ ID NO: 15	--	SEQ ID NO: 53	SEQ ID NO: 54	GGAAGG AG
GPX4 (1-2-3)	SEQ ID NO: 16	SEQ ID NO: 84 o SEQ ID NO: 85	SEQ ID NO: 55	SEQ ID NO: 56	CTGCCA CA
TXN2	SEQ ID NO: 17	--	SEQ ID NO: 57	SEQ ID NO: 58	GGCCCC AG
PRDX4	SEQ ID NO: 18	--	SEQ ID NO: 59	SEQ ID NO: 60	ACTGGG AA
GPX1 (2)	SEQ ID NO: 19	--	SEQ ID NO: 61	SEQ ID NO: 62	CTCCTC CT
GLRX3	SEQ ID NO: 20	--	SEQ ID NO: 63	SEQ ID NO: 64	TGGTGG AA
PRDX2 (3)	SEQ ID NO: 21	--	SEQ ID NO: 65	SEQ ID NO: 66	GGAGGC TG

(continuación)

Nombre del gen	SEQ ID	SEQ ID variante	SEQ ID Oligo directo	SEQ ID Oligo inverso	SEQ de sondas
PRDX6	SEQ ID NO: 22	--	SEQ ID NO: 67	SEQ ID NO: 68	CCTGGA GC
GLRX5	SEQ ID NO: 23	--	SEQ ID NO: 69	SEQ ID NO: 70	TGCTGG AG
GLRX2 (1)	SEQ ID NO: 24	--	SEQ ID NO: 71	SEQ ID NO: 72	GGATGG AG

Tabla 4

5 La Tabla 4 se puede leer de la siguiente manera: El nivel de expresión de GPX3 (SEQ ID NO: 1) se puede medir cuantitativamente utilizando los oligonucleótidos SEQ ID NO: 24 y 25, y utilizando la sonda, acoplada a un colorante fluorescente y un desactivador, que tiene la siguiente secuencia CCAGCCGC.

También, en otro ejemplo: El nivel de expresión de PDRX (SEQ ID NO: 5) o una de sus variantes (SEQ ID NO: 73 o 74) se pueden medir cuantitativamente utilizando los oligonucleótidos SEQ ID NO: 33 y 34, y utilizando como sonda, acoplada con un colorante fluorescente y un desactivador, que tiene la siguiente secuencia CTGGCTGG.

10 Las definiciones anteriores se aplican mutatis mutandis a los otros genes.

El desactivador y colorante mencionados anteriormente se pueden escoger por el experto, dependiendo del ensayo.

15 En otra realización ventajosa, la invención se refiere al procedimiento que se ha definido anteriormente, utilizando al menos uno de los siguientes oligonucleótidos permitiendo la medición del nivel de expresión de cada uno de al menos los 6 genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1-6, preferentemente utilizando al menos uno de los oligonucleótidos que permiten la medición del nivel de expresión de cada uno de al menos 10 genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1-10, en particular utilizando al menos uno de los oligonucleótidos que permiten la medición del nivel de expresión de cada uno de los 24 genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1-24.

20 En esta realización, el procedimiento se adapta para el ensayo de transferencia de Northern.

25 La invención se ilustra por una composición que comprende oligonucleótidos que permiten la medición de la expresión de al menos los genes de un subgrupo de 6 genes que pertenecen a un grupo de genes escogidos de entre un grupo de 24 genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1 a 24, perteneciendo dichos 6 genes a dicho subgrupo que comprende o está constituido por las secuencias de ácido nucleico de SEQ ID NO: 1 a 6.

En una realización ventajosa, la invención se ilustra por una composición como se ha definido anteriormente, que comprende oligonucleótidos que permiten la medición de la expresión de al menos los genes de un grupo de 10 genes escogidos de entre un grupo de 24 genes, comprendiendo o estando constituido dicho grupo de 24 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 a 24,

30 dichos 10 genes que pertenecen a dicho grupo comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1 a 10.

En otra realización ventajosa, la invención se ilustra por una composición como se ha definido anteriormente que comprende los oligonucleótidos que permiten medir la expresión de los 24 de dicho grupo de 24 genes.

35 En otra realización ventajosa, la invención se ilustra por una composición como se ha definido anteriormente, comprendiendo dicha composición al menos 12 oligonucleótidos escogidos de entre una biblioteca de 48 oligonucleótidos que comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO: 25-72, permitiendo la medición de la expresión de al menos los genes de un subgrupo de 6 genes que pertenecen a un grupo de genes escogidos de entre un grupo de 24 genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1 a 24,

40 comprendiendo o consistiendo dichos al menos 12 oligonucleótidos las secuencias de ácido nucleico de SEQ ID NO: 25-36.

La invención se refiere a una composición que comprende al menos 12 oligonucleótidos escogidos de entre una biblioteca de 48 oligonucleótidos que comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO:

25-72, comprendiendo dicha composición al menos los 12 oligonucleótidos que comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO: 25-36.

5 En una realización ventajosa, la invención se refiere a una composición como se ha descrito anteriormente, comprendiendo dicha composición al menos 20 oligonucleótidos escogidos de entre una biblioteca de 48 oligonucleótidos que comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO: 25-72, permitiendo la medición de la expresión de al menos los genes de un grupo de 10 genes escogidos de entre un grupo de 24 genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1 a 24, comprendiendo o consistiendo dichos al menos 20 oligonucleótidos en las secuencias de ácido nucleico de SEQ ID NO: 25-44.

10 En otra realización ventajosa, la invención se refiere a una composición como se ha definido anteriormente, comprendiendo dicha composición 48 oligonucleótidos que comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO: 25-72, permitiendo la medición de la expresión de un grupo de 24 genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 1 a 24.

15 La composición anterior puede comprender adicionalmente las sondas que se definen anteriormente en la Tabla 4.

En otro aspecto, la invención se ilustra por una composición como se ha mencionado anteriormente, para su uso en el diagnóstico, y/o la clasificación, preferentemente *in vitro*, de un trastorno hematológico, en particular un trastorno hematológico mieloide y/o linfóide, preferentemente un trastorno hematológico mieloide.

20 La invención se refiere a una composición como se ha mencionado anteriormente, para su uso en el diagnóstico, y/o la clasificación, preferentemente *in vitro*, de una leucemia mieloide aguda o un trastorno mielodisplásico.

Por lo tanto, la invención se refiere a la composición anterior per se, y se refiere a dicha composición para su uso como se ha mencionado anteriormente.

25 La invención se refiere a un kit que comprende al menos 12 oligonucleótidos escogidos de entre un grupo de 48 oligonucleótidos que comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO: 25-72, comprendiendo o consistiendo dichos al menos 12 oligonucleótidos en las secuencias de ácido nucleico de SEQ ID NO: 25-36.

En una realización ventajosa, la invención se refiere al kit que se ha definido anteriormente, que comprende al menos 20 oligonucleótidos, comprendiendo o consistiendo dichos al menos 20 oligonucleótidos las siguientes secuencias de ácido nucleico de SEQ ID NO: 25-44.

30 En una realización ventajosa, la invención se refiere a un kit como se ha definido anteriormente que comprende 48 oligonucleótidos que comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO: 25-72.

35 La invención también se refiere a, en una realización ventajosa, un kit como se ha definido anteriormente, que comprende además al menos 6 sondas específicas que interactúan respectivamente con las moléculas de ácido nucleico que comprenden o consisten en SEQ ID NO: 1-6, preferentemente comprenden adicionalmente al menos 10 sondas específicas que interactúan respectivamente con moléculas de ácido nucleico que comprenden o consisten en SEQ ID NO: 1-10, en particular comprenden adicionalmente al menos 24 sondas específicas que interactúan respectivamente con las moléculas de ácido nucleico que comprenden o consisten en SEQ ID NO: 1-24.

40 En otra realización ventajosa, la invención se refiere a un kit como se ha definido anteriormente, que comprende adicionalmente moléculas de ácido nucleico que corresponden con los genes de SEQ ID NO: 1-24, en una cantidad representativa de al menos una patología escogida de entre: leucemia mieloide aguda (AML), anemia refractaria con sideroblastos anulares (RARS), citopenia refractaria con displasia multilineal (RCMD), anemia refractaria con exceso de blastos (RAEB) o síndrome 5q y mielodisplasia inclasificable.

En otra realización ventajosa, la invención también se refiere al kit que se ha definido anteriormente, que comprende además moléculas de ácido nucleico de una muestra de control como se ha definido anteriormente.

45 La invención también se refiere a una muestra de control positivo que comprende o está constituida por al menos las moléculas de ácido nucleico que se corresponden con los genes representados en SEQ ID NO: 1-6, que se escogen de entre el grupo de 24 genes representados por SEQ ID NO: 1-24, estando presentes dichas moléculas de ácido nucleico en dicha muestra en una cantidad que se representa en las 6 primeras líneas de la Tabla 5 o la Tabla 6, en comparación con una muestra sana en la que cada molécula de ácido nucleico respectiva están presentes en una cantidad de 1.

50

Tabla 5

Tabla 5	RARS	RCMD	RAEB	AML
SEQ ID NO: 1	1,32-2,62	0,01 - 1,67	0,43 - 2,88	0,03-0,30
SEQ ID NO: 2	2,32-2,98	0,12-9,95	4,48 - 18,89	2,98-6,98
SEQ ID NO: 3	3,74-9,79	0,08-10,50	0,14-10,87	3,12-15,80
SEQ ID NO: 4	1,28-2,20	0,02-3,94	0,99-8,75	0,32-10,20
SEQ ID NO: 5	1,25-3,68	0,18-5,77	4,59-10,04	2,00-13,34
SEQ ID NO: 6	1,84-6,09	0,00-10,24	0,05-5,53	1,72-5,56
SEQ ID NO: 7	0,65-0,95	0,54-4,37	0,85-2,13	0,08-0,79
SEQ ID NO: 8	0,67-1,09	0,00 - 1,18	0,44 - 1,15	0,20-0,58
SEQ ID NO: 9	0,90-1,25	0,43 - 2,05	0,07-1,05	0,08-0,79
SEQ ID NO: 10	0,48 - 1,87	0,02-2,13	0,49 - 4,41	0,04-0,25
SEQ ID NO: 11	0,81 - 1,69	0,00-2,01	1,19-1,47	0,65-1,76
SEQ ID NO: 12	0,64-2,12	0,23-3,74	1,48 - 2,74	0,45 - 2,94
SEQ ID NO: 13	0,96-2,43	0,63-4,99	1,14-4,22	0,42-5,34
SEQ ID NO: 14	1,20-2,02	0,16-3,16	0,63-1,47	0,70-2,15
SEQ ID NO: 15	0,63-1,44	0,09-1,18	0,00-1,37	0,55-3,71
SEQ ID NO: 16	0,94-2,53	0,10-3,29	2,58-6,22	1,35-4,25
SEQ ID NO: 17	1,07-2,22	0,15-2,58	2,30-2,85	0,90-3,89
SEQ ID NO: 18	1,48 - 2,79	0,37-2,86	3,31 - 4,99	1,10-6,67
SEQ ID NO: 19	1,04-5,32	0,49 - 3,87	1,96-8,77	0,64-5,18
SEQ ID NO: 20	1,49 - 5,23	0,09-2,58	1,81 - 11,93	0,84-7,09
SEQ ID NO: 21	0,93-8,83	0,01 - 2,87	0,15-3,76	0,08-0,63
SEQ ID NO: 22	0,81 - 3,63	0,14-36,64	0,25-12,98	0,98-5,93
SEQ ID NO: 23	1,47 - 3,45	0,19-7,04	2,80-53,30	0,26-3,65
SEQ ID NO: 24	0,78-2,25	0,35-1,60	0,40-0,87	0,18-0,83

La Tabla 5 representa para cada gen de SEQ ID NO: 1-24, el intervalo específico correspondiente con la patología mencionada.

5

Tabla 6

Tabla 6	RARS	RCMD	RAEB	AML
SEQ ID NO: 1	2,14 ± 0,71	0,76 ± 0,56	1,60 ± 1,23	0,16 ± 0,13
SEQ ID NO: 2	2,64 ± 0,33	3,65 ± 3,95	18,89 ± 17,93	4,95 ± 1,70
SEQ ID NO: 3	5,97 ± 3,33	4,04 ± 3,49	5,47 ± 5,36	9,73 ± 5,24
SEQ ID NO: 4	1,63 ± 0,50	1,46 ± 1,64	4,31 ± 4,00	4,01 ± 3,73
SEQ ID NO: 5	2,52 ± 1,22	1,89 ± 2,14	7,40 ± 2,73	5,03 ± 4,74
SEQ ID NO: 6	4,55 ± 2,36	3,64 ± 4,01	5,30 ± 5,13	4,19 ± 1,67
SEQ ID NO: 7	0,84 ± 0,17	1,80 ± 1,61	1,32 ± 0,70	0,33 ± 0,27
SEQ ID NO: 8	0,85 ± 0,22	0,76 ± 0,42	0,78 ± 0,35	0,40 ± 0,18
SEQ ID NO: 9	1,07 ± 0,18	1,22 ± 0,67	0,54 ± 0,49	0,45 ± 0,26

(continuación)

Tabla 6	RARS	RCMD	RAEB	AML
SEQ ID NO: 10	1,06 ± 0,72	0,74 ± 0,73	2,40 ± 1,96	0,11 ± 0,08
SEQ ID NO: 11	1,33 ± 0,46	0,93 ± 0,83	1,47 ± 0,29	1,11 ± 0,44
SEQ ID NO: 12	1,31 ± 0,75	1,12 ± 1,30	1,92 ± 0,71	1,27 ± 0,97
SEQ ID NO: 13	1,51 ± 0,80	2,06 ± 1,70	2,32 ± 1,67	2,52 ± 1,87
SEQ ID NO: 14	1,68 ± 0,43	1,64 ± 1,21	1,17 ± 0,46	1,70 ± 0,59
SEQ ID NO: 15	0,99 ± 0,41	0,79 ± 0,39	0,70 ± 0,68	2,57 ± 1,32
SEQ ID NO: 16	1,61 ± 0,82	1,58 ± 1,28	6,22 ± 4,17	2,62 ± 1,37
SEQ ID NO: 17	1,69 ± 0,58	1,15 ± 0,96	2,79 ± 0,47	2,40 ± 1,33
SEQ ID NO: 18	2,17 ± 0,65	1,19 ± 0,88	4,36 ± 0,91	2,59 ± 2,33
SEQ ID NO: 19	2,59 ± 2,37	1,67 ± 1,25	4,89 ± 3,50	2,12 ± 1,80
SEQ ID NO: 20	2,75 ± 2,15	1,27 ± 1,12	5,31 ± 5,74	3,96 ± 2,61
SEQ ID NO: 21	3,80 ± 4,37	1,38 ± 1,18	2,41 ± 1,97	0,36 ± 0,23
SEQ ID NO: 22	1,83 ± 1,56	7,33 ± 14,42	5,69 ± 6,56	2,90 ± 2,21
SEQ ID NO: 23	2,76 ± 1,12	2,60 ± 2,88	21,67 ± 27,56	1,49 ± 1,34
SEQ ID NO: 24	1,56 ± 0,74	0,95 ± 0,46	0,87 ± 0,62	0,41 ± 0,25

La Tabla 6 representa, para cada gen de SEQ ID NO: 1-24, la media ± la desviación estándar correspondiente con la patología mencionada.

- 5 La invención también se refiere a un procedimiento para determinar la eficacia de un tratamiento de un trastorno hematológico, siendo el tratamiento adecuado para administrarse a un paciente, comprendiendo dicho procedimiento
- a) una etapa de poner en contacto, preferentemente *in vitro*, una muestra biológica de un sujeto que padece un trastorno hematológico, preferentemente un trastorno hematológico mieloide y/o linfoide, más preferentemente un trastorno hematológico mieloide, con un fármaco adecuado para utilizarse en el tratamiento, o adecuado para tratar dicho trastorno hematológico,
- 10 b) una etapa de medición en la muestra biológica que se puso en contacto con un fármaco en la etapa a), el nivel de expresión de al menos los genes de un subgrupo de 6 genes que pertenecen a un grupo de genes escogidos de entre un grupo de 24 genes, comprendiendo o estando constituido dicho grupo de 24 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 a 24, , perteneciendo dichos 6 genes a dicho subgrupo que comprende o está constituido por las secuencias de ácido nucleico de SEQ ID NO: 1 a 6.
- 15 y
- c) una etapa de comparación del nivel de expresión de dichos al menos los genes de un subgrupo de 6 genes que pertenecen a un grupo de genes escogidos de entre un grupo de 24 genes obtenidos en la etapa b) con el nivel de expresión de dichos al menos los genes de un subgrupo de 6 genes que pertenecen a un grupo de genes escogidos de entre un grupo de 24 genes obtenidos de la medición en dicha muestra biológica que no se ha puesto en contacto con dicho fármaco adecuado para utilizarse para tratar, o adecuado para tratar, dicho trastorno hematológico.
- 20
- 25 El procedimiento anterior es fácil de llevar a cabo, y permite evaluar la susceptibilidad de la muestra de AML a un fármaco. Esto es muy importante para reducir el coste de los tratamientos, y que pueden ser ineficaces en un paciente, debido a que el tumor sea resistente al fármaco.
- El procedimiento anterior se utiliza ventajosamente para explorar, *in vitro*, fármacos que tengan eficacia en la progresión de AML, y que se podrían utilizar *in vivo* para el tratamiento del paciente.
- 30 Esto es más ventajosamente importante para la exploración de fármacos, o compuestos que son capaces de modular la modificación epigenética, en particular de agentes desmetilantes tales como la azacitidina (5-azacitidina) o decitabina (5-azadesoxicitidina).

La azacitidina y decitabina son agentes quimioterápicos potentes que se utilizan para tratar AML y MDS de alto grado, pero solo aproximadamente el 40 % de AML y MDS de alto grado son sensibles a sus efectos. Por lo tanto,

con el fin de reducir los costes, y los efectos secundarios de un tratamiento ineficaz, es ventajoso verificar *in vitro*, si la AML y MDS de alto grado que se van a tratar son sensibles a estos compuestos.

El ejemplo 9 muestra que el tratamiento de una muestra de AML con azacitidina puede modular la relación de expresión de los genes de SEQ ID NO: 1-24, (y consecuentemente los genes de SEQ ID NO: 1-6) que demuestra que la azacitidina, en este particular del que deriva la muestra de AML, sería eficaz si se utilizara *in vivo*.

Ventajosamente, la invención se refiere a un procedimiento como se ha definido anteriormente, en el que dicho grupo consiste en 10 genes que consisten en las secuencias de ácido nucleico de SEQ ID NO: 1-10.

Ventajosamente, la invención se refiere a un procedimiento como se ha definido anteriormente, en el que se mide el nivel de expresión de los 24 genes que consisten en SEQ ID NO: 1-24.

10 Leyenda de las figuras

La **Figura 1** representa el mecanismo natural utilizado en la célula para eliminar las especies de óxido reactivo (ROS, O₂^{*}). O₂^{*} se convierte en moléculas de H₂O₂, convirtiéndose las propias H₂O₂ en H₂O y O₂ como proceso de detoxificación, o en •OH y OH⁻ que ejercen efectos biológicos en las células.

La **Figura 2** representa las imbricaciones entre el subgrupo (A), el grupo (B) y el grupo (C) de los genes de acuerdo con la invención. D representan el conjunto correspondiente a los genes que pertenecen al grupo pero no pertenecen al subgrupo. E representa el conjunto correspondiente a los genes que pertenecen al grupo pero no pertenecen al grupo.

Las **Figuras 3A-E** representan la representación esquemática del nivel de expresión de cada uno de los genes representados por SEQ ID NO: 1-24, para cada patología RARS, RCMD, RAEB y AML.

La **Figura 3A** representa un histograma que muestra el nivel de expresión de los genes indicados en el eje x, que se obtiene por qRT-PCR, en comparación con el gen constitutivo GAPDH, utilizando 5 muestras de médula ósea sana (muestra de control).

El eje Y representa la relación

$$\Delta CT_{gen} (\Delta CT_{gen} = CT_{gen} - CT_{GAPDH}).$$

La **Figura 3B** es una representación gráfica de la media (línea negra) y error estándar de la media (área gris) de la variación de expresión de cada uno de los genes indicados (eje x) en una muestra de RARS en comparación con muestras de control (muestras de médula ósea sana). El eje Y representa la variación de la cantidad en comparación con muestras de control expresada en log₁₀.

La **Figura 3C** es una representación gráfica de la media (línea negra) y el error estándar de la media (área gris) de la variación de expresión de cada uno de los genes indicados (eje x) en muestras de RMCD en comparación con muestras de control (muestras de médula ósea sana). El eje Y representa la variación de la cantidad comparada con las muestras de control, expresada en log₁₀.

La **Figura 3D** es una representación gráfica de la media (línea negra) y el error estándar de la media (área gris) de la variación de expresión de cada uno de los genes indicados (eje x) en muestras de RAEB en comparación con muestras de control (muestras de médula ósea sana). El eje Y representa la variación de la cantidad comparada con las muestras de control, expresada en log₁₀.

La **Figura 3E** es una representación gráfica de la media (línea negra) y el error estándar de la media (área gris) de la variación de expresión de cada uno de los genes indicados (eje x) en muestras de AML en comparación con muestras de control (muestras de médula ósea sana). El eje Y representa la variación de la cantidad comparada con las muestras de control, expresada en log₁₀.

Las **Figuras 4A-D** representan la representación esquemática del nivel de expresión de cada uno de los genes representados por SEQ ID NO: 1-24, para cada una de las patologías: RARS (Figura 4A), RMCD (Figura 4B), RAEB (Figura 4C), y AML (Figura 4D). La representación del nivel de cada uno de los genes representados por SEQ ID NO: 1-24 de una muestra de síndrome mielodisplásico inclasificable se representa en cada una de las Figuras 4A-D por una línea discontinua.

Las **Figuras 5A y B** representan la representación esquemática del nivel de expresión de cada uno de los genes representados por SEQ ID NO: 1-24 de dos muestras independientes de síndromes mielodisplásico inclasificables (Figura 5B), siendo dicha expresión similar al nivel de expresión observado en las muestras de RCMD (Figura 5A).

La **Figura 6** representa la representación esquemática del nivel de expresión de cada uno de los genes representados por SEQ ID NO: 1-24 de dos muestras independientes de leucemia mielomonocítica crónica (Figura 6B) siendo dicha expresión diferente del nivel de expresión observado en las muestras de RAEB (Figura 6A)

La **Figura 7** representa la representación esquemática del nivel de expresión de cada uno de los genes representados por SEQ ID NO: 1-24 de un paciente en el momento del diagnóstico de RCMD (panel B) y después de 12 meses (panel B, línea discontinua). La muestra en el momento del diagnóstico es similar a la muestra de RMCD (panel A), y la muestra tras 12 meses ha adquirido las características de RAEB (véase PDRX2(1) y GLRX5 en el panel B y panel C).

La **Figura 8** representa la representación esquemática del nivel de expresión de cada uno de los genes

representados por SEQ ID NO: 1-24 de AML, comparando la misma AML tras el tratamiento con azacitidina (línea discontinua).

Ejemplos

Comentario preliminar

5 Todas las muestras utilizadas en los siguientes ejemplos se han ensayado primero de acuerdo con la invención, en un ensayo con ocultación, y se compararon con las muestras de control.

Todas las muestras de SMD y leucémicas satisfacen las provisiones del procedimiento de acuerdo con la invención, es decir, en cada una de las muestras al menos 3 genes de los genes de SEQ ID NO: 1-6 se expresan de manera que su relación comparada con la expresión de los mismos genes correspondientes de las muestras de control es o menor de 0,5 o mayor de 2.

Ejemplo 1

En los siguientes ejemplos, las muestras del paciente y el análisis de las expresiones genéticas se analizan de la siguiente manera.

Materiales y Procedimientos

15 Recolección de muestras

Se obtuvieron muestras normales de médula ósea (BM) (que se utilizan como referencia) de pacientes sometidos a cirugía ortopédica. Las muestras de médula ósea de pacientes de MDS y AML se obtuvieron en el momento del diagnóstico y durante el seguimiento. Las células de AML y MDS se clasificaron de acuerdo con los hallazgos morfológicos, citoquímicos y citogenéticos. Los pacientes se informaron y consintieron de seguir un procedimiento aprobado por el comité ético. Las muestras de BM se aspiraron en jeringas heparinizadas y se transfirieron a tubos con EDTA.

Lisis de eritrocitos

Se llevó a cabo la lisis en 47 ml de tampón de lisis que contenía EDTA (0,12 mM), bicarbonato potásico (KHCO₃) (10 mM), y cloruro amónico (NH₄Cl) (150 mM) para 3 ml de BM. Después de la incubación a temperatura ambiente durante 15 min, se centrifugaron a 700 g durante 10 min, y se lavaron dos veces, en 20 ml de solución salina tampón fosfato (PBS) (Invitrogen). El aglomerado se re-suspendió en Trizol® (Invitrogen) (1 ml/ 8.10⁶ células), se mezcló vigorosamente durante 15 minutos. El lisado se almacenó a -80 °C hasta la etapa de extracción de ARN.

Extracción de ARN

30 La extracción del ARN total se llevó a cabo según el procedimiento de cloroformo / isopropanol / etanol. La fase de separación se obtuvo añadiendo cloroformo (0,2 ml por 1 ml de Trizol® obtenido en Invitrogen) al lisado. Tras mezclar durante 50 s, la solución se separó en tres fases por centrifugación a 12.000 g durante 15 min a 4 °C. Se precipitó el ARN a partir de la fase acuosa, añadiendo isopropanol (0,5 ml por 1 ml de Trizol®). Después de la incubación a temperatura ambiente durante 10 min y centrifugación a 12.000 g durante 10 min a 4 °C, se lavó el ARN en un 75 % de etanol (1 ml por 1 ml de Trizol®) y se centrifugó a 7500 g durante 5 min a 4 °C. Esta etapa se llevó a cabo dos veces. Tras retirar el sobrenadante de etanol, el aglomerado de ARN se secó al aire durante 20 min. después, se disolvió el ARN en 50 µl de agua tratada con UltraPure™ DEPC (Invitrogen) y se almacenó a -80 °C.

Cuantificación y cualificación de ARN

40 El ARN celular total se cuantificó utilizando un espectrofotómetro Nano-Drop 1000 (Nano-Drop Technologies) y se analizó la pureza de ARN utilizando un Bioanalizador Agilent 2100 (Agilent Technologies).

Transcripción inversa y análisis de PCR cuantitativa en tiempo real (qRT-PCR)

Se transcribieron tres microgramos de ARN total de cada muestra se transcribieron inversamente utilizando el kit de síntesis de ADNc SuperScript® VILO™ (Invitrogen) de acuerdo con el protocolo del suministrador. La cuantificación relativa de la expresión genética se hizo por PCR en tiempo real en la plataforma cicladora basada en placas de 480 micropocillos LightCycler® (Roche Applied Science) utilizando ensayos Universal ProbeLibrary diseñados con el software ProbeFinder (Roche Applied Science, www.roche-applied-science.com/sis/rtPCR/upl/ezhome.html). Los cebadores se adquirieron en Invitrogen y las sondas Universal ProbeLibrary en Roche Applied Science. Las secuencias de nucleótidos de los cebadores y sondas de cada diana se muestran en la Tabla A. Todas las dianas se analizaron concomitantemente. Las reacciones de qRT-PCR se llevaron a cabo en un volumen total de 10 µl de 20 ng de ADNc utilizando LightCycler® 480 Probes Master (Roche Applied Science). El LightCycler® 480 se programó para una desnaturalización inicial (95 °C, 10 min) seguido por 45 ciclos de 10 s a 95 °C, 30 s a 60 °C, 1 s a 72 °C y una etapa de enfriamiento final a 40 °C durante 10 s. Todas las reacciones se ejecutaron por triplicado, y los

valores medios se utilizaron para la cuantificación. Los resultados se analizaron por el procedimiento de cuantificación relativa ($\Delta\Delta CT = \Delta CT_{\text{paciente}} - \Delta CT_{\text{referencia}}$) utilizando los valores de umbral de ciclo (CT) determinados con el software LightCycler® 480 (publicación 1.5.0) de Roche Applied Science. Se utilizó el gen de gliceraldehído-3-fosfato deshidrogenasa humana (*GAPDH*) como control endógeno para normalizar la expresión de la diana ($\Delta CT = CT_{\text{diana}} - CT$ del control endógeno). El cambio de la expresión relativa de ARNm entre un paciente y la referencia se determinó para cada diana utilizando el procedimiento $2^{\text{exp}(-\Delta\Delta CT)}$ (Livak KJ y Schmittgen TD. Analysis of relative gene expression data using real time quantitative PCR and the $2^{\text{exp}(-\Delta\Delta CT)}$ method. Methods 2001; 25:402-408).

Ejemplo 2: Uso de 6 genes para el diagnóstico de mielodisplasia y leucemia

Con el fin de ensayar la validez del procedimiento, se utilizaron las muestras del paciente identificadas como representativas de las muestras leucémicas, y se evaluó el nivel de expresión de cada uno de los genes representados por SEQ ID NO: 1 a SEQ ID NO: 6 por RT-qPCR como se desvela en el Ejemplo 1, utilizando oligonucleótidos y sondas de la Tabla 4.

Los resultados se presentan en la Tabla 7 siguiente:

Tabla 7

	AML nº 1	AML nº 2	AML nº 3	AML nº 4	AML nº 5
SEQ ID NO: 1	0,03	0,29	0,30	0,08	0,08
SEQ ID NO: 2	3,54	6,16	6,98	2,98	5,10
SEQ ID NO: 3	14,24	7,07	8,41	15,80	3,12
SEQ ID NO: 4	10,20	4,06	3,32	2,17	0,32
SEQ ID NO: 5	13,34	2,96	2,00	4,46	2,36
SEQ ID NO: 6	5,56	3,22	5,12	5,34	1,72

La Tabla 7 representa la relación del nivel de expresión de cada uno de los genes indicados (SEQ ID NO: 1-6) para cada muestra de AML nº 1-nº 5.

Los criterios definidos en la invención se satisfacían.

Una muestra se considera como representativa de un trastorno hematológico cuando la relación del nivel de expresión de cada uno de los genes de cualquier combinación de 3 genes de entre los genes representados por SEQ ID NO: 1-6 es ≥ 2 o $\leq 0,5$.

Cualquiera de las combinaciones que se tomaron en consideración, satisfacían los criterios anteriores.

El otro criterio es que: si la relación de SEQ ID NO: 1 es $\leq 0,3$ y la relación de tanto SEQ ID NO: 2 como 3 son $\geq 3,0$, la muestra es representativa de una AML.

De nuevo para las 5 muestras de AML anteriores que se ensayaron, se satisfacían los criterios.

Todas las muestras de AML ensayadas satisfacían los criterios con respecto a AML. Por lo tanto, el procedimiento de acuerdo con la invención permite la discriminación entre AML y trastornos mielodisplásicos.

Ejemplo 3: Validación del procedimiento con líneas celulares leucémicas

Con el fin de validar el procedimiento, se evaluó el nivel de expresión de los genes representados por SEQ ID NO: 1-6 en 11 líneas celulares correspondientes a casi todos los subtipos de AML definidos de acuerdo con la clasificación FAB.

Se ensayaron las siguientes líneas celulares:

KG1a y KG1 (FAB M0/M1), HL60 (FAB M2), KASUMI-1 (FAB M2), ML-2 (FAB M4), MV4-11 (FAB M5), THP-1 (FAB M5), U937 (FAB M5), K562 (FAB M6), TF-1 (FAB M6) y UT7 (FAB M7).

Los niveles de expresión se indican en la Tabla 8 siguiente.

Tabla 8

SEQ ID NO:	KG1 a	KG 1	HL60	Kasumi-1	ML-2	MV4 -11	TH P-1	U93 7	K56 2	TF-1	UT 7
1	0,05	0,06	0,05	0,04	0,06	0,03	0,15	0,07	0,10	0,03	0,30
2	3,42	5,26	4,74	6,23	5,61	17,95	45,70	12,20	4,58	7,92	48,90
3	37,71	54,38	33,49	14,34	34,17	17,19	77,40	45,84	69,24	18,38	53,66
4	0,53	0,91	2,73	1,25	2,13	5,19	8,80	1,91	2,08	1,54	2,44
5	197,83	39,98	14,21	9,46	6,10	8,08	29,76	17,93	48,00	31,97	33,70
6	30,91	32,92	13,85	12,11	9,14	5,84	20,37	10,29	24,11	13,99	30,32

5 La Tabla 8 representa las relaciones R_i entre los niveles de expresión del gen indicado en la línea celular correspondiente, en comparación con los niveles de expresión del gen correspondiente en las muestras de control (muestras de médula ósea sana).

Los criterios definidos en la invención se satisfacían:

Una muestra se consideraba como representativa de un trastorno hematológico cuando la relación del nivel de expresión de cada uno de los genes de cualquier combinación de 3 genes entre los genes representados por SEQ ID NO: 1-6 es ≥ 2 o $\leq 0,5$.

10 Cualquiera de las combinaciones de 3 genes que se toman en consideración, satisfacía los criterios anteriores.

El otro criterio era que: si la relación de SEQ ID NO: 1 es $\leq 0,3$ y la relación de tanto SEQ ID NO: 2 como 3 son $\geq 3,0$, la muestra es representativa de AML.

15 De nuevo, para las 11 líneas celulares ensayadas anteriores (Tabla 3), se satisfacían los criterios, y el procedimiento confirmaba que estos resultados obtenidos con estas líneas celulares se corresponden con los obtenidos con células leucémicas primarias.

Ejemplo 4: Clasificación de la muestra mielodisplásica

Como se ha definido anteriormente, se ha propuesto que la medición del nivel de expresión representados por SEQ ID NO: 1-24 permite identificar el subtipo de síndrome mielodisplásico.

20 Se ensayó un panel de muestras de paciente, clasificadas por otras técnicas para validar el procedimiento de acuerdo con la invención.

Se utilizaron 3 RARS, 6 RCMD y 3 RAEB y se evaluó el nivel de expresión de los genes representados por SEQ ID NO: 1-24.

Los resultados se presentan a continuación:

Tabla 9

	RARS nº 1	RARS nº 2	RARS nº 3
SEQ ID NO: 1	2,48	1,32	2,62
SEQ ID NO: 2	2,63	2,32	2,98
SEQ ID NO: 3	9,79	3,74	4,38
SEQ ID NO: 4	2,20	1,41	1,28
SEQ ID NO: 5	3,68	2,64	1,25
SEQ ID NO: 6	5,72	6,09	1,84
SEQ ID NO: 7	0,65	0,93	0,95
SEQ ID NO: 8	1,09	0,79	0,67
SEQ ID NO: 9	1,25	1,06	0,90
SEQ ID NO: 10	1,87	0,48	0,82

ES 2 614 978 T3

(continuación)

	RARS nº 1	RARS nº 2	RARS nº 3
SEQ ID NO: 11	1,69	1,50	0,81
SEQ ID NO: 12	2,12	0,64	1,17
SEQ ID NO: 13	2,43	1,16	0,96
SEQ ID NO: 14	2,02	1,82	1,20
SEQ ID NO: 15	1,44	0,89	0,63
SEQ ID NO: 16	2,53	1,35	0,94
SEQ ID NO: 17	2,22	1,07	1,79
SEQ ID NO: 18	2,79	2,25	1,48
SEQ ID NO: 19	1,43	5,32	1,04
SEQ ID NO: 20	5,23	1,49	1,53
SEQ ID NO: 21	1,64	8,83	0,93
SEQ ID NO: 22	3,63	1,06	0,81
SEQ ID NO: 23	3,35	3,45	1,47
SEQ ID NO: 24	1,64	0,78	2,25

Tabla 9: Muestras de RARS

Tabla 10

	RCMD nº 1	RCMD nº 2	RCMD nº 3	RCMD nº 4	RCMD nº 4	RCMD nº 5
SEQ ID NO: 1	0,55	1,05	1,67	0,66	0,61	0,01
SEQ ID NO: 2	0,52	9,95	5,75	0,54	0,12	5,05
SEQ ID NO: 3	3,44	4,57	10,50	2,71	0,08	2,96
SEQ ID NO: 4	2,07	2,53	3,94	0,06	0,02	0,13
SEQ ID NO: 5	2,43	5,77	2,23	0,48	0,18	0,28
SEQ ID NO: 6	3,96	6,09	10,24	1,46	0,10	0,00
SEQ ID NO: 7	0,58	3,22	4,37	1,27	0,54	0,82
SEQ ID NO: 8	0,91	1,05	1,18	0,84	0,57	0,00
SEQ ID NO: 9	1,48	2,05	1,86	0,88	0,62	0,43
SEQ ID NO: 10	0,68	2,13	0,75	0,31	0,53	0,02
SEQ ID NO: 11	0,71	2,01	1,89	0,69	0,30	0,00
SEQ ID NO: 12	0,72	3,74	0,78	0,55	0,71	0,23
SEQ ID NO: 13	1,98	4,99	2,99	1,09	0,63	0,66
SEQ ID NO: 14	1,81	3,16	2,89	0,69	1,10	0,16
SEQ ID NO: 15	0,87	1,18	0,95	1,01	0,64	0,09
SEQ ID NO: 16	1,43	3,29	2,80	0,35	0,10	1,48
SEQ ID NO: 17	1,24	2,58	1,95	0,70	0,15	0,30
SEQ ID NO: 18	1,08	2,86	1,23	0,37	0,95	0,64
SEQ ID NO: 19	1,58	3,87	2,28	0,49	0,71	1,10

ES 2 614 978 T3

(continuación)

	RCMD nº 1	RCMD nº 2	RCMD nº 3	RCMD nº 4	RCMD nº 4	RCMD nº 5
SEQ ID NO: 20	2,58	2,46	1,71	0,47	0,09	0,31
SEQ ID NO: 21	1,96	2,40	2,87	0,56	0,47	0,01
SEQ ID NO: 22	0,59	3,60	2,49	0,51	0,14	36,64
SEQ ID NO: 23	2,46	7,04	5,12	0,40	0,40	0,19
SEQ ID NO: 24	0,81	1,60	1,37	0,67	0,93	0,35

Tabla 10: Muestras de RCMD

Tabla 11

	RAEB nº 1	RAEB nº 2	RAEB nº 3
SEQ ID NO: 1	1,49	0,43	2,88
SEQ ID NO: 2	38,96	13,21	4,48
SEQ ID NO: 3	5,40	10,87	0,14
SEQ ID NO: 4	3,18	0,99	8,75
SEQ ID NO: 5	4,59	7,56	10,04
SEQ ID NO: 6	10,31	5,53	0,05
SEQ ID NO: 7	0,85	2,13	0,99
SEQ ID NO: 8	0,76	1,15	0,44
SEQ ID NO: 9	0,50	1,05	0,07
SEQ ID NO: 10	2,30	0,49	4,41
SEQ ID NO: 11	1,76	1,19	1,46
SEQ ID NO: 12	1,48	2,74	1,54
SEQ ID NO: 13	1,59	1,14	4,22
SEQ ID NO: 14	0,63	1,47	1,40
SEQ ID NO: 15	0,72	1,37	0,00
SEQ ID NO: 16	10,76	5,31	2,58
SEQ ID NO: 17	3,23	2,85	2,30
SEQ ID NO: 18	3,31	4,99	4,78
SEQ ID NO: 19	1,96	3,95	8,77
SEQ ID NO: 20	2,18	11,93	1,81
SEQ ID NO: 21	3,32	3,76	0,15
SEQ ID NO: 22	3,83	12,98	0,25
SEQ ID NO: 23	8,91	2,80	53,30
SEQ ID NO: 24	1,58	0,63	0,40

Tabla 11: Muestras de RAEB

Todas las muestras ensayadas satisfacían los criterios definidos en los procedimientos de acuerdo con la invención.

Ejemplo 5: Representación esquemática de los perfiles específicos de RARS, RMCD, RAEB y AML

Los resultados correspondientes a la clasificación anterior se pueden ilustrar como un “antioxidograma”, que se corresponde a áreas representativas de una muestra determinada.

5 La expresión de cada gen de SEQ ID NO: 1-24 se midió por qRT-PCR como se ha definido anteriormente, y se comparó con el nivel de expresión de un gen GAPDH constitutivo.

La Figura 3A representa la variación de expresión de cada uno de los genes de SEQ ID NO: 1-24 en comparación con GAPDH, clasificados de acuerdo con su nivel de expresión. El número de gen se encuentra fácilmente utilizando la Tabla 1 anterior.

10 El antioxidograma para la muestra RARS (Figura 3B), muestra de RCMD (Figura 3C), muestras de RAEB (Figura 3D) y muestras de AML (3E) se pueden utilizar para clasificar una nueva muestra.

Además, midiendo la relación de cada uno de los genes representados por SEQ ID NO: 1-24, se puede trazar una curva. Esta curva se puede comparar entonces con los antioxidogramas, y de esta manera, es fácil determinar a qué tipo de patología pertenece la muestra estudiada.

Ejemplo 6: clasificación de síndromes mielodisplásicos inclasificables

15 El antioxidograma anterior se puede utilizar para clasificar síndromes mielodisplásicos para los que fallan otras técnicas de clasificación.

El nivel de expresión de cada uno de los genes de SEQ ID NO: 1-24 se midió de acuerdo con la invención, y se estableció una curva.

Esta curva (línea discontinua) se comparó con el antioxidograma “específico” de RARS, RCDM, RAEB y AML.

20 La Figura 4 muestra que la mayoría de las relaciones R_i de expresión están cerca a las que son características de una muestra de RCDM.

La Figura 5 muestra que dos muestras independientes de síndrome mielodisplásico inclasificable presente un nivel de expresión similar de los genes de SEQ ID NO: 1-24, y de esta manera están cerca de las que son características de una muestra de RCDM.

25 Ejemplo 7: Distinción de MDS utilizando oxidogramas

La leucemia mielomonocítica crónica (CMML) es una forma de leucemia caracterizada por monocitosis. La categorización de esta enfermedad ha sido controvertida.

Los pacientes con CMML pueden presentar varias características clínicas, imitando síndromes mielodisplásicos o neoplasias mieloproliferativas, dependiendo de la presentación específica en el paciente.

30 Debido a esta controversia se clasificó por la Organización Mundial de la Salud en una categoría “mielodisplásica/mieloproliferativa” de afecciones médicas a principios de los 2000.

El oxidograma de acuerdo con la invención puede ayudar a determinar el estado de una muestra de CMML, es está cerca de RAEB por análisis histológico.

35 Como se muestra en la Figura 6, el nivel de expresión de los genes de SEQ ID NO: 1-24 de muestras de dos pacientes con CMML es diferente del nivel de expresión de los genes SEQ ID NO: 1-24 de RAEB.

Estos datos demuestran que CMML es distinta de MDS.

Ejemplo 8: Utilización del oxidograma para el seguimiento de MDS evolutivo

Como se ha mencionado anteriormente, un MDS evoluciona progresivamente a AML. De esta manera, es importante saber si esta progresión es rápida o lenta.

40 El oxidograma se puede utilizar para determinar si la expresión molecular de los genes de SEQ ID NO: 1-24 ha evolucionado desde el diagnóstico a una determinada fecha.

Un ejemplo se muestra en la Figura 7. Se puede diagnosticar un paciente en un $t = 0$ como que tiene RCMD.

A los doce meses del diagnóstico, el nivel de expresión de los genes de SEQ ID NO: 1-24 también se midieron.

45 La Figura 7 muestra una diferencia entre la expresión en el momento del diagnóstico y tras doce meses. Después de doce meses, el oxidograma muestra que el paciente ha sido capaz de evolucionar de RCMD hacia RAEB, aunque en el análisis histológico no hay diferencias.

Ejemplo 9: Medición *in vitro* de efecto antitumoral de un agente desmetilante

Se buscan activamente nuevos agentes terapéuticos con el fin de tratar la AML. Estos compuestos son caros debido a las búsquedas extensas, y desafortunadamente no son eficaces universalmente en todas las muestras de AML.

5 Una media del 30 % de los subtipos de AML es sensible a los nuevos fármacos terapéuticos, y se podrían utilizar eficazmente en pacientes, con el fin de disminuir la progresión leucémica.

Sin embargo, no se puede obtener una respuesta positiva o negativa *in vivo* a un fármaco antes de 6 a 9 meses del inicio del tratamiento.

El oxidograma de acuerdo con la invención se puede utilizar para evaluar, *in vitro*, si el fármaco será eficaz en la muestra leucémica, estudiando la modulación del nivel de expresión de los genes de SEQ ID NO: 1-24.

10 La Figura 8 muestra un ejemplo de una muestra de AML tratada con el agente desmetilante azacitidina. Esta figura demuestra que, en esta muestra específica, el tratamiento modifica el nivel de expresión de los genes, y podrían ser eficaz *in vivo* para el paciente.

LISTADO DE SECUENCIAS

15 <110> UNIVERSITE FRANCOIS RABELAIS
HERAULT, Olivier
VIGNON, Christine

20 <120> PROCEDIMIENTO DE DIAGNÓSTICO DE TRASTORNOS HEMATOLÓGICOS

<130> WOB 10 TOU OXYD

<150> EP 10306483.8

<151> 22-12-2010

25 <160> 85

<170> PatentIn versión 3.5

30 <210> 1
<211> 1779
<212> ADN
<213> *Homo sapiens*

35 <400> 1

ES 2 614 978 T3

gtcgccccgg gacggggagg tggggagctg agggcaagtc gcgccccccc ctgaaatccc 60
agccgcctag cgattggctg caagggctct ggcttggccg cggattggtc acacccgagg 120
gcttgaaagg tggctgggag cgccggacac ctcagacgga cggtgccag ggatcaggca 180
gcggtcagc cgaccctgag tgtgccccca ccccgccatg gcccggtgc tgcaggcgtc 240
ctgcctgctt tccctgctcc tggccggctt cgtctcgcag agccggggac aagagaagtc 300
gaagatggac tgccatggtg gcataagtgg caccatttac gagtacggag ccctcacat 360
tgatggggag gagtacatcc ccttcaagca gtatgctggc aaatacgtcc tctttgtcaa 420
cgtggccagc tactgaggcc tgacgggcca gtacattgaa ctgaatgcac tacaggaaga 480
gcttgacca ttcggtctgg tcattctggg ctttccctgc aaccaatttg gaaaacagga 540
accaggagag aactcagaga tccttcctac cctcaagtat gtccgaccag gtggaggctt 600
tgtccctaat tccagctct ttgagaaagg ggatgtcaat ggagagaaag agcagaaatt 660
ctacactttc ctaaagaact cctgtcctcc cacctcggag ctctgggta catctgaccg 720
cctcttctgg gaacccatga aggttcacga catccgctgg aactttgaga agttcctggt 780
ggggccagat ggtataccca tcatgcgctg gcaccaccgg accacggtca gcaacgtcaa 840
gatggacatc ctgtcctaca tgaggcggca ggcagccctg ggggtcaaga ggaagtaact 900
gaaggccgtc tcatcccatg tccaccatgt aggggaggga ctttgttcag gaagaaatcc 960
gtgtctccaa ccacactatc taccatcac agaccocctt cctatcactc aaggccccag 1020
cctggcacia atggatgcat acagttctgt gtactgccag gcatgtgggt gtgggtgcat 1080
gtgggtgttt acacacatgc ctacaggtat gcgtgattgt gtgtgtgtgc atgggtgtac 1140
agccacgtgt ctacctatgt gtctttctgg gaatgtgtac catctgtgtg cctgcagctg 1200
tgtagtgtg gacagtgaca accctttctc tccagttctc cactccaatg ataatagttc 1260
acttacacct aaacccaaag gaaaaaccag ctctaggtcc aattgttctg ctctaactga 1320
tacctcaacc ttggggccag catctcccac tgcctccaaa tattagtaac tatgactgac 1380
gtccccagaa gtttctgggt ctaccacact ccccaacccc ccaactcctac ttcctgaagg 1440
gccctcccaa ggctacatcc ccaccccaca gttctccctg agagagatca acctccctga 1500
gatcaaccaa ggcagatgtg acagcaaggg ccacggaccc catggcaggg gtggcgtctt 1560
catgagggag gggcccaaag cccttgtggg cggacctccc ctgagcctgt ctgaggggcc 1620
agcccttagt gcattcagc taaggcccct gggcagggat gccaccctg ctcttccgga 1680
ggacgtgcc tcaccctca ctggtccact ggcttgagac tcaccctgc tgcccagtaa 1740
aagcctttct gcagcagctg aaaaaaaaa aaaaaaaaa 1779

<210> 2
<211> 921
<212> ADN
<213> *Homo sapiens*

ES 2 614 978 T3

<400> 2

```

cagttaaaag gaggcgcctg ctggcctccc cttacagtgc ttgttcgggg cgctccgctg      60
gcttcottgga caattgcgcc atgtgtgctg ctcggttagc ggoggcggcg gcggcggccc      120
agtgggtgta tgccttctcg gcgcgcccgc tggccggcgg ggagcctgtg agcctgggct      180
ccctgcgggg caaggtacta cttatogaga atgtggcgtc cctctgaggc accacgggtcc      240
gggactacac ccagatgaac gagctgcagc gggcctcgg accccggggc ctggtggtgc      300
tcggottccc gtgcaaccag tttgggcatac aggagaaagc caagaacgaa gagattctga      360
attccctcaa gtaogtccgg cctggtggtg gggtcgagcc caacttcatg ctcttogaga      420
agtgcgaggt gaacgggtgc ggggcgcacc ctctcttcgc cttcctgagg gaggcctgc      480
cagctcccag cgacgacgcc accgcgctta tgaccgacc caagctcacc acctggtctc      540
cgggtgtgctg caacgatggt gcctggaact ttgagaagtt cctggtgggc cctgacggtg      600
tgcccctacg caggtacagc cgccgcttcc agaccattga catcgagcct gacatcgaag      660
ccctgctgtc tcaagggccc agctgtgcct agggcgcgcc tcctaccccg gctgcttggc      720
agttgcagtg ctgctgtctc gggggggttt tcatctatga ggggtgttcc tctaaacctt      780
cgagggagga acaoctgatc ttacagaaaa taccacctcg agatgggtgc tggctcctgtt      840
gatcccagtc tctgccagac caaggcgagt ttccccacta ataaagtgcc ggggtgtcagc      900
agaaaaaaaa aaaaaaaaaa a                                             921

```

- 5 <210> 3
- <211> 735
- <212> ADN
- <213> *Homo sapiens*

10 <400> 3

```

gcggtgtccg gcagtagagc tcgctgcaga tccgggctct gaccatgatt tggcgcgcgcg      60
cggcgcctggc ggggacgcgg ctggtttggg gcaggagcgg ctcggcaggc tggcttgaca      120
gggcggcggg agctgcggga gctgcggcag ctgcggcctc tgggatggag agcaatacat      180
catcatcttt ggagaattta gcgacggcgc ctgtgaacca gatccaagaa acaatttctg      240
ataattgtgt ggtgattttc tcaaaaacat cctgttctta ctgtacaatg gcaaaaaagc      300
ttttccatga catgaatggt aactataaag tgggtggaact ggacctgctt gaatatggaa      360
accagttcca agatgctctt tacaaaatga ctggtgaaag aactgttcca agaatatttg      420
tcaatggtac ttttatttgg ggtgcaactg aactcatag gcttcacaaa gaaggaaaat      480
tgctcccact agttcatcag tgttatttaa aaaaaagtaa gaggaagaa tttcagtgat      540
gtttatacta ataagtttgc tagtacagtg tcagttatth aaagtggtaa tgcccgataa      600
tgtcttttaa atgtttgagg atgttttaa tacatgcatt gtcttcacga agaagatgta      660
aaaataatga acaataaatt gcggtggaac cctcaaaaaa aaaaaaaaaa aaaaaaaaaa      720
aaaaaaaaaa aaaaaa                                             735

```

ES 2 614 978 T3

<210> 4
 <211> 2300
 <212> ADN
 <213> *Homo sapiens*

5

<400> 4

```

actcggggca acaggcagat ttgctgctg aggggtggaga cccacgagcc gaggcctcct      60
gcagtgttct gcacagcaaa ccgcacgcta tggttgacag ccgggatccc gccagcgacc      120
agatgcagca ctggaaggag cagcgggccg cgcagaaagc tgatgtcctg accactggag      180
ctggtaacct agtaggagac aaacttaatg ttattacagt agggccccgt gggccccttc      240
ttgttcagga tgtggttttc actgatgaaa tggctcattt tgaccgagag agaattcctg      300
agagagttgt gcatgctaaa ggagcagggg cttttggcta ctttgaggtc acacatgaca      360
ttaccaaata ctccaaggca aaggtatttg agcatattgg aaagaagact cccatcgcag      420
ttcggttctc cactgttgct ggagaatcgg gttcagctga cacagttcgg gaccctcgtg      480
ggtttgagct gaaattttac acagaagatg gtaactggga tctcgttggg aataacaccc      540
ccattttctt catcagggat cccatattgt ttccatcttt tatccacagc caaaagagaa      600
atcctcagac acatctgaag gatccggaca tggctctggga cttctggagc ctacgtcctg      660
agtctctgca tcaggtttct ttcttgttca gtgatcgggg gattccagat ggacatcgcc      720
acatgaatgg atatggatca catactttca agctggttaa tgcaaatggg gaggcagttt      780
attgcaaatt ccattataag actgaccagg gcatcaaaaa cttttctggt gaagatgcgg      840
    
```

ES 2 614 978 T3

cgagactttc ccaggaagat cctgactatg gcatccggga tctttttaac gccattgcc 900
 caggaaagta cccctcctgg actttttaca tccaggcat gacatttaac caggcagaaa 960
 cttttccatt taatccattc gatctcacca aggtttggcc tcacaaggac taccctctca 1020
 tcccagttgg taaactggtc ttaaaccgga atccagttaa ttactttgct gaggttgaac 1080
 agatagcctt cgacccaagc aacatgccac ctggcattga ggccagtcct gacaaaatgc 1140
 ttcagggccg cctttttgcc taccctgaca ctaccgccca tcgcctggga cccaattatc 1200
 ttcataatcc tgtgaactgt ccctaccgtg ctgcagtggtc caactaccag cgtgacggcc 1260
 cgatgtgcat gcaggacaat caggggtggtg ctccaaatta ctacccaac agctttggtg 1320
 ctccggaaca acagccttct gccctggagc acagcatcca atattctgga gaagtgcgga 1380
 gattcaacac tgccaatgat gataacgtta ctcaggtgctg ggcattctat gtgaacgtgc 1440
 tgaatgagga acagaggaaa cgtctgtgtg agaacattgc cggccacctg aaggatgcac 1500
 aaattttcat ccagaagaaa gcggtcaaga acttcaactga ggtccaccct gactacggga 1560
 gccacatcca ggctcttctg gacaagtaca atgctgagaa gcctaagaat gcgattcaca 1620
 cctttgtgca gtccggatct cacttggcgg caagggagaa ggcaaatctg tgaggccggg 1680
 gccctgcacc tgtgcagcga agotttagcgt tcatccgtgt aaccgcctca tcaactggatg 1740
 aagattctcc tgtgctagat gtgcaaatgc aagctagtgg cttcaaaata gagaatccca 1800
 ctttctatag cagattgtgt aacaatttta atgctatttc cccaggggaa aatgaaggtt 1860
 aggatttaac agtcatttaa aaaaaaatt tgttttgacg gatgattgga ttattcattt 1920
 aaaatgatta gaaggcaagt ttctagctag aaatatgatt ttatttgaca aaatttgttg 1980
 aaattatgta tgtttacata tcacctcatg gcctattata ttaaaatag gctataaata 2040
 tataaaaaga aaagataaag atgatctact cagaaatttt tatttttcta aggttctcat 2100
 aggaaaagta catttaatac agcagtgctc tcagaagata acttgagcac cgtcatggct 2160
 taatgtttat tctgataat aattgatcaa attcattttt tcaactggag ttacattaat 2220
 gttaattcag cactgatttc acaacagatc aatttgtaat tgettacatt ttacaataa 2280
 ataactctga cgtaagaaca 2300

<210> 5
 <211> 1262
 <212> ADN
 <213> *Homo sapiens*

5

<400> 5

ES 2 614 978 T3

```

actctcgcga gatccctact ggctataaag gcagcgcgcc ggagagctct tgcgcgtctt      60
gttctctgct ggtgtcgggtg gttagtttct gcgacttggtg ttgggactgg tgagtgtggg      120
cagtgcggcc cctgcggagt gaggcgcggc gcgccttct tgcctgttgc ctcttctcc      180

tctgtccgg gccccgccg cgctcgggtg ggggtgctgt gatgcgtgag gcagccgggg      240
gaggcccgga gtccgagact gcttgagcgc tgcgcacacc cctctcgtgg gccccccag      300
taggtgcggg aacctggtt aaccccaagc tgataggaag atgtcttcag gaaatgctaa      360
aattgggcac cctgccccca acttcaaagc cacagctggt atgccagatg gtcagtttaa      420
agatatcagc ctgtctgact acaaaggaaa atatgttggtg ttcttctttt acctcttga      480
cttcaccttt gtgtgcccc aaggagatcat tgctttcagt gatagggcag aagaatttaa      540
gaaactcaac tgccaagtga ttggtgcttc tgtggattct cacttctgtc atctagcatg      600
ggtcaataca cctaagaaac aaggaggact gggacccatg aacattcctt tggtatcaga      660
cccgaagcgc accattgctc aggattatgg ggtcttaaag gctgatgaag gcatctcgtt      720
caggggcctt tttatcattg atgataaggg tattcttcgg cagatcactg taaatgacct      780
ccctgttggc cgctctgtgg atgagacttt gagactagtt caggccttcc agttcactga      840
caaacatggg gaagtgtgcc cagctggctg gaaacctggc agtgatacca tcaagcctga      900
tgtccaaaag agcaaagaat atttctcaa gcagaagtga gcgctgggct gttttagtgc      960
caggctcggg tgggcagcca tgagaacaaa acctcttctg tatttttttt ttccattagt      1020
aaaacacaag acttcagatt cagccgaatt gtggtgtcct acaaggcagg cctttcctac      1080
agggggtgga gagaccagcc tttcttctt tggtaggaat gccctgagtt ggcgttgtgg      1140
gcaggctact ggtttgtatg atgtattagt agagcaaccc attaatcttt tgtagtttgt      1200
attaaacttg aactgagacc ttgatgagtc ttaaaaaaa aaaaaaaaa aaaaaaaaa      1260
aa                                                                                   1262

```

5 <210> 6
 <211> 827
 <212> ADN
 <213> *Homo sapiens*
 <400> 6

ES 2 614 978 T3

```

gcagtggagg cggcccaggc ccgccttccg caggggtgtcg ccgctgtgcc gctagcgggtg      60
ccccgcctgc tgcggtggca ccagccagga ggcggagtgg aagtggccgt ggggcgggta      120
tgggactagc tggcgtgtgc gccctgagac gctcagcggg ctatatactc gtcggtgggg      180
ccggcgggtca gtctgcccga gcggcagcaa gacgggtgag tgaaggagag tgggcgtctg      240
gcggggtccg cagtttcagc agagccgctg cagccatggc cccaatcaag gtgggagatg      300
ccatcccagc agtggagggtg tttgaagggg agccagggaa caaggtgaac ctggcagagc      360
tgttcaaggg caagaagggt gtgctgtttg gagttcctgg ggccttcacc cctggatggt      420
ccaaggttcg gctcctggct gatcccactg gggcctttgg gaaggagaca gacttattac      480
tagatgattc gctggtgtcc atctttggga atcgacgtct caagaggttc tccatggtgg      540

tacaggatgg catagtgaag gccctgaatg tggaaaccaga tggcacaggc ctcacctgca      600
gcctggcacc caatatcatc tcacagctct gaggcctctg gccagattac ttctccacc      660
cctccctatc tcacctgcc agccctgtgc tggggccctg caattggaat gttggccaga      720
tttctgcaat aaacacttgt ggtttgcggc caaaaaaaaa aaaaaaaaaa aaaaaaaaaa      780
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa      827

```

<210> 7
 <211> 1593
 <212> ADN
 <213> *Homo sapiens*

 <400> 7

5

ES 2 614 978 T3

gcggtgccct tgcggcgcag ctggggtcgc ggcctgctc cccgcgcttt cttaaagccc 60
 gcggggggcg caggagcggc actcgtggct gtggtggctt cggcagcggc ttcagcagat 120
 cggcggcatc agcggtagca ccagcactag cagcatggtg agccgggcag tgtgcggcac 180
 cagcaggcag ctggctccgg ttttggggta tctgggctcc aggcagaagc acagcctccc 240
 cgacctgccc tacgactacg gggccctgga acctcacatc aacgcgcaga tcatgcagct 300
 gcaccacagc aagcaccacg cggcctacgt gaacaacctg aacgtcaccg aggagaagta 360
 ccaggaggcg ttggccaagg gagatgttac agcccagata gctcttcagc ctgcactgaa 420
 gttcaatggt ggtggtcata tcaatcatag cttttctggt acaaacctca gccctaacgg 480
 tgggtggagaa cccaaagggg agttgctgga agccatcaaa cgtgactttg gttcctttga 540
 caagtttaag gagaagctga cggctgcatc tgttgggtgc caaggctcag gttgggggtg 600
 gcttggtttc aataaggaac ggggacactt acaaattgct gcttgtcaa atcaggatcc 660
 actgcaagga acaacaggcc ttattccact gctggggatt gatgtgtggg agcacgctta 720
 ctacctcag tataaaaatg tcaggcctga ttatctaaaa gctatttggg atgtaatcaa 780
 ctgggagaat gtaactgaaa gatacatggc ttgcaaaaag taaaccacga tcgttatgct 840
 gagtatgta agctctttat gactgttttt gtagtggtat agagtactgc agaatacagt 900
 aagctgctct attgtagcat ttcttgatgt tgcttagtca cttatttcat aaacaactta 960
 atgttctgaa taatttctta ctaaacattt tgttattggg caagtgattg aaaatagtaa 1020
 atgctttgtg tgattgaatc tgattggaca ttttcttcag agagctaaat tacaattgct 1080
 atttataaaa ccatcaaaaa tattccatcc atatactttg gggacttgta gggatgcctt 1140
 tctagtctta ttctattgca gttatagaaa atctagtott ttgccccagt tacttaaaaa 1200
 taaaatatta acactttccc aagggaaaca ctcggcttcc tatagaaaat tgcacttttt 1260
 gtcgagtaat cctctgcagt gatactctg gtagatgtca cccagtgggt tttgtaggt 1320
 caaatgttcc tgtatagttt ttgcaaatag agctgtatac tgtttaaatg tagcaggtga 1380

 actgaactgg ggtttgctca cctgcacagt aaaggcaaac ttcaacagca aaactgcaaa 1440
 aaggtggttt ttgcagtagg agaaaggagg atgtttatatt gcagggcgcc aagcaaggag 1500
 aattgggcag ctcatgcttg agacceaatc tccatgatga cctacaagct agagtattta 1560
 aaggcagtgg taaatttcag gaaagcagaa gtt 1593

5 <210> 8
 <211> 3174
 <212> ADN
 <213> *Homo sapiens*
 <400> 8

ES 2 614 978 T3

tcttcctggg tcttgcctag cggcggggcg atgcttagtc accgtgaggc tgcgcttgcc	60
cggggcccgc gccccctac ccoggggacc gccccgggc cgccccccc acttggcgcg	120
ccacttccgc gtgcatggcc ctgctgcccc gagccctgag cgccggcgcg ggaccgagct	180
ggcggcgggc ggcgcgcgcc ttccgagget tectgctgct tctgcccag cccgcggccc	240
tcacgcgcgc cctctcccgt gccatggcct gcaggcagga gccgcagccg cagggcccgc	300
cgcccgctgc tggcgcogtg gcctcctatg actacctggt gatcgggggc ggctcgggcg	360
ggctggccag cgcgcgcagg gcggccgagc tgggtgccag ggccgccgtg gtggagagcc	420
acaagctggg tggcacttgc gtgaatggtg gatgtgtacc caaaaaggta atgtggaaca	480
cagctgtcca ctctgaattc atgcatgac atgctgatta tggctttcca agttgtgagg	540
gtaaattcaa ttggcgtggt attaaggaaa agcgggatgc ctatgtgagc cgctgaatg	600
ccatctatca aaacaatctc accaagtccc atatagaaat catccgtggc catgcagcct	660
tcacgagtga tccaagccc acaatagagg tcagtgggaa aaagtacacc gccccacaca	720
tctgatcgc cacagtggt atgccctcca ccctcatga gagccagatc cccggtgcca	780
gcttaggaat aaccagcgat ggattttttc agctggaaga attgccggc cgcagcgtca	840
ttgttggtgc aggttacatt gctgtggaga tggcagggat cctgtcagcc ctgggttcta	900
agacatcact gatgatacgg catgataagg tacttagaag ttttgattca atgatcagca	960
ccaactgcac ggaggagctg gagaacgctg gcgtggaggt gctgaagttc tcccaggtca	1020
aggaggttaa aaagactttg tcgggcttgg aagtcagcat ggttactgca gttcccggta	1080
ggctaccagt catgaccatg attccagatg ttgactgcct gctctgggcc attgggcggg	1140
tcccgaatac caaggacotg agtttaaaca aactggggat tcaaaccgat gacaagggtc	1200
atatcatcgt agacgaattc cagaatacca acgtcaaagg catctatgca gttggggatg	1260
tatgtggaaa agctcttctt actccagttg caatagctgc tggccgaaaa cttgcccatc	1320
gactttttga atataaggaa gattccaaat tagattataa caacatccca actgtggtct	1380
tcagccaccc ccctattggg acagtgggac tcacggaaga tgaagccatt cataaatatg	1440

ES 2 614 978 T3

gaatagaaaa tgtgaagacc tattcaacga gctttacccc gatgtatcac gcagttacca 1500
 aaaggaaaac aaaatgtgtg atgaaaatgg tctgtgctaa caaggaagaa aaggtgggtg 1560
 ggatccatat gcagggactt ggggtgtgatg aaatgctgca gggttttgct gttgcagtga 1620
 agatgggagc aacgaaggca gactttgaca acacagtgc cattcacccct acctcttcag 1680
 aagagctggt cacacttctg tgagaaccag gagacacgtg tggcgggcag tgggacccat 1740
 agatcttctg aaatgaaaca aataatcaca ttgacttact gtttgagttt tatgtatttc 1800
 tttattttta tcaggatctt ctgatagtgg aaatttttag tacataatag aacttattta 1860
 tggagttaga aattttagt gttatccagg attgattttc atttgatcac atctcacagt 1920
 aattaatatt ttcaagttt ttttttatta acagctctgt gctagttttt ttttctggt 1980
 ttagcctcat cccaaatata aagctttgtg aagtacaatt aacttaatgt acttgaatga 2040
 atagaacttg ctactttttt tttttttttt tttgagacag agttttgctc tcattgcca 2100
 ggctggagtg cgggtgtgct atttcagctc accacaacct ctgcctcctg ggttcaagtg 2160
 attctcctgc cttagcctcc cgaatagctg gaattacagg cacgcaccac catgcctgac 2220
 taattttgta ttttagtag acatggggtt tctccatggt ggtcaggctg gtctcaaact 2280
 cccaccttca ggtgatccgc ccacctcggc ctctgaggt gctgagatta caggcgtgag 2340
 ccactgtgcc agcttgctaa ttttcacaga agttgatggc aattcttcac atgtaaacag 2400
 tgccagtga cagaacctt atatatattt tgaagccagt actgtgctct gcatataaca 2460
 aagctgcttc aaggatgaga ctttttcta aaagcatgta atgtgagaag ccggcctgcc 2520
 ttattttctt tttcttttt taatgattaa aaatagtttg tggcaaggca cgggtggctca 2580
 ggctgtaat tctagcactt tgggagggc aggcaggagg attacttgag cctacaagtt 2640
 tgaggccagc atgcacagca tagcaagact gcctctctac agagagtaaa aaaaattacc 2700
 cgagtgtggt gatgtgcatc tgtaatctca gctacttggg aggctgaggt gagaggatca 2760
 cttgagcttg ggtgaggtga ggctgcagtg agtcctgatc atgtgctgc actcaatctt 2820
 ggacaacaga gcaagacct gtctcaaaaa aaaaaaaaaa aaatatatat atatatatat 2880
 attattttta tgaggatgaag tgcatcaaac ttgggaaaga tttgaggagg ctgggaacct 2940
 cctggaaaac cactccttga agaaagatat gagagacatt tagaagtgat tcctgctttc 3000
 agaaggaggt ggattcaaat acatcaaaag tccttctctc tgctaagtgt ttatagttca 3060
 atgaataatt tcaatatttg tatgtgttct tgtcatttta ttttttctg aaaaacttcc 3120
 aaaaatttga aaataaaatt acagcctttt cttcttataa aaaaaaaaaa aaaa 3174

<210> 9
 <211> 1661
 <212> ADN
 <213> *Homo sapiens*

5

<400> 9

ES 2 614 978 T3

attgcattcc tgggcattgc taactagtga agtataccag atggaaatgt cttcgaagct 60
gtccctttaa aactcgagca agctaccagg caaactccgc ctccagggag gttccttatt 120
aataggagc caactggctg ggtcggggct caatacccca agcaatacct gcaactgagg 180
attcttcccg gggagaccgc agcccatcgg catggctcaa gagtttgtga actgcaaaat 240
ccagcctggg aaggtggttg tgttcatcaa gcccaoctgc ccgtactgca ggagggccca 300
agagatcctc agtcaattgc ccatcaaca agggcttctg gaatttgtcg atatcacagc 360
caccaaccac actaacgaga ttcaagatta tttgcaacag ctcacgggag caagaacggt 420
gcctcgagtc tttattggta aagattgtat aggcggatgc agtgatctag tctctttgca 480
acagagtggg gaactgctga cgcggctaaa gcagattgga gctctgcagt aaccacagaa 540
caggcccat gctgacgtcc ctctcaaga gctggatggc attgcaaatg atgacagcac 600
ttctggtgga tgaatttggg ggcacaaaca gctttttcc tcttttggct cagtatttaa 660
aagtggacca acttgctctt aatcacaggg ccaagaaggt tgacgggcca tcttggtttt 720
cttctggatg tgctctttgg ttttcagaag actgtgacaa gttctggccc aggattcgct 780
cactgaccct caattgtcct ctttggcatg cgtttcttac tgttctccat gtgtcggcat 840
gtctctacct ctaagccagt gttttcaac tatgtttatc cagactcctt ctccacaatg 900
atgaatccac agttggttat ctgctactgc ccattagcta aatcatttt gctgcttgac 960
tttatggagt ttgtattatg aatcagtggt gtattttgaa tgtgttcttt ctaactacat 1020
gcatctctcc actcaactcc accccatccc atcccacctt gaaaatcact gctctgaacc 1080
agtgttctcc accttgctct ccacagatct cataggaaat gttcaacaat tctgtgaaag 1140
gtcacaggac ccaattggag aatcatatg aaaagcatag ttggtcttgg tgtcatatgg 1200
atcagaggca caagtgcaga ggctgtggtc atgcggaaca ctctgttatt taagatggct 1260
atccagataa tcctgaacac tgtgtattta ttttatttag actaccagca aagattaag 1320
catgaaatgt aaaacatctg ataaaactta cagcccccta caccaagagt gtatctgtga 1380
aagagctcct acactttgaa aacttaagaa tcccttatca tgaagtttgc ctgttctaga 1440
attgtaagat tgttaatttc cttcaatctc tagtgacaac acttaatttc ttttctaata 1500
aaaaaacct atagatgatt cagtgatttt tgtccaattc atttgcattg tctcaagaca 1560
ttaaggaatg ttatgcgaaa tacactaact taaaactgtg tttatatttg gccctgccat 1620
tataaataaa gacacgtgct gctgtcaaaa aaaaaaaaaa a 1661

<210> 10
<211> 1039
<212> ADN
<213> *Homo sapiens*

5

<400> 10

ES 2 614 978 T3

gctcgtccgc tccctcccc gcgccgtgca cgtcttggtt cgggccgggc ataaaaggct 60
 tcgcgggcca gggctcactt ggcgctgaga acgcggtcc acgctgtga tcgtccgtgc 120
 gtctagecct tgcccacgca gtttcagtc atggcctccg gtaacgcgcg catcgaaag 180
 ccagcccctg acttcaaggc cacagcggtg gttgatggcg ccttcaaaga ggtgaagctg 240
 tcggactaca aaggaagta cgtggctctc tttttctacc ctctggaett cacttttgtg 300
 tgccccacog agatcatcgc gttcagcaac cgtgcagagg acttccgcaa gctgggctgt 360
 gaagtgtctg gcgtctcggg ggactctcag ttcaccacc tggcttgat caacaccccc 420
 cggaaagagg gaggcttggg cccctgaac atccccctgc ttgctgacgt gaccagacgc 480
 ttgtctgagg attacggcgt gctgaaaaca gatgagggca ttgcctacag gggcctcttt 540
 atcatcgatg gcaaggggtg ccttcgccag atcactgta atgatttgc tgtgggacgc 600
 tccgtggatg aggctctgcg gctggctccag gcctccagt acacagacga gcatggggaa 660
 gtttgtcccg ctggctggaa gcctggcagt gacacgatta agcccaacgt ggatgacagc 720
 aaggaatatt tctccaaaca caattaggct ggctaaccga tagtgagctt gtgcccctgc 780
 ctaggtgcct gtgctgggtg tccacctgtg cccccacctg ggtgccctat gctgacctag 840
 gaaaggccag acctgccct ccaaactcca cagtatggga ccctggaggg ctaggccaaag 900
 gccttctcat gcctccacct agaagctgaa tagtgacgcc ctcccccaag cccaccacgc 960
 cgcacacagg cctagaggta accaataaag tattagggaa aggtgtgaaa aaaaaaaaaa 1020
 aaaaaaaaaa aaaaaaaaaa 1039

<210> 11
 <211> 2799
 <212> ADN
 <213> *Homo sapiens*

5

<400> 11

acgtgtgagg gaggaagca ggaagtgact gcgggagtg agccggcgag agagtggcag 60
 cgggggctga tggaagtgca gtgggggctg gagagggcac cctactgtat ccagcatgct 120
 ccaaggccac agctctgtgt tccaggcctt gctggggacc ttcttcacct gggggatgac 180
 agcagctggg gcagctctcg tgttcgtatt ctctagtga cagagcgga tottagatgg 240
 aagtcttggc tttgctgcag gggcatggt ggcagcttcc tattggtctc ttctggcccc 300
 agcagttgag atggccacgt cctctggggg cttcgggtgcc tttgccttct tcctgtggc 360
 tgttggettcc acccttgag cggttttgt ctacttggt gacctcctga tgectcactt 420
 ggggtgcagca gaagaccccc agacgacct ggcactgaac ttcggtcta cgttgatgaa 480
 gaagaagtct gatcctgagg gtcccgcgt gctcttcct gagagtgaac tttccatccg 540

10

ES 2 614 978 T3

gataggtaga gctgggcttc tttcagacaa gagtgagaat ggtgaggcat atcagagaaa 600
gaaggcggca gccactggcc ttccagaggg tcctgctgtc cctgtgcctt ctcgagggaa 660
tctggcacag cccggcggca gcagctggag gaggatcgca ctgctcatct tggccatcac 720
tatacacaac gttccagagg gtctcgctgt tggagttgga tttggggcta tagaaaagac 780
ggcatctgct acctttgaga gtgccaggaa tttggccatt ggaatcggga tccagaattt 840
ccccgagggc ctggctgtca gccttccctt gcgaggggca ggcttctcca cctggagagc 900
tttctggtat gggcagctga ggggcatggt ggagcccctg gccggggtct ttggtgcctt 960
tgccgtggtg ctggctgagc ccatcctgcc ctacgctctg gcctttgctg ccggtgccat 1020
ggtctacgtg gtcattggac acatcatccc cgaagcccag atcagtggta atgggaaact 1080
ggcatcctgg gcctccatcc tgggatttgt agtgatgatg tcaactggacg ttggcctggg 1140
ctagggctga gacgcttcgg accccgggaa aggccatacg aagaacagc agtggttggc 1200
ttctatggga caacaagctt ctttcttcac attaaaactt ttttcttcc tctcttctc 1260
atctcattat cctgattgac tctgattata atagaacat ttttactttg ctttgaggga 1320
gatttttgat ttaatgggga attttaaggt gtcattggaaa tacagattct ttgttttggc 1380
cactgaatgg actctctctt cagtgggatt atcaaggaac ttcagatcag ggaaatctcc 1440
acttcgggac cttctatctg cctcccaact cctcaaggtc acctatagaa gcgagctacc 1500
aaaagacgtc tcctaagcat tttggtggcc tagtgactca gggcagagtg gccagcacac 1560
ctctcatccg cccctcctgc tccatcactg ctgagcctct ccccatctag aatgttggaa 1620
ctggagcatc ataaagatag caagctacct tccaaggccg agccagcca gagaggagca 1680
tgtcttctt tacctcccc taaggagata ctacatggga gggggacaca gaaaaaggga 1740
aggaaattgg ctagtctggc tttttttttt ttttttttta aaggcaaaga ttgacattat 1800
tgaaggaaag gggatgagga caactgtgaa ctcacagtga gccctgtgga aagaagagac 1860
agacagagtg tgggtttggt cggaggcctc tgctgtcaat ggattccagg agcaaggcca 1920
tttgtcgcgc tttccaaatt tcttaggcat ttattttgat aagtttatag ccatcatggt 1980
tctaagagac ttggagacac cagcaaactg ctagaactca aactcttcaa ttactcaaag 2040
aaggagccat ttcagttaac tcaagtgaat gaaagagttt tggaatctgc tgtgggtcct 2100
tcctgttga ccatttggtg acttataatc tgacaaaaac tcttgagctg caacaggcct 2160
tgccagaggg ctcaggatgg gaaaggaaga aggggatagg aaaagaagag gtaattttac 2220
atctccctt taaagtaaat tttagccaac tcatcattct gaaatgtccc tataaagaat 2280
gagtcgaact agaccagaag ccagcctact ccttcttaca tagcttctcc aacaggggta 2340
gcaatgacct gtccacttca aacacagata aggcctgcc tctctattgg ttaaaggcac 2400

ES 2 614 978 T3

acgtgagact ttcagtgggc tctgctgaga aggaaggcag cccaggagtc aggtatgcag 2460
gcattgcatt gtcagtgtct gctctcagag tttacacatt caattgcttc caagggtgaa 2520
tctcctgctc tgtgaatgct atcagacccc aaaggccaac cttgggctgg gtctatgtac 2580
gttcttccga agcactgatg atcaaaattg aagacacatt cagaggtttg attggttgag 2640
attaactggg gtgggtggtg gtgtatgtat gttttatfff tatgtctttg tatgtagtcc 2700
tacataatgc aaattgtgct ttctgatgga caagacctca taactgtgat taatatcaat 2760
aaaaagggga tgttgtggat gaaaaaaaa aaaaaaaaa 2799

5 <210> 12
<211> 981
<212> ADN
<213> *Homo sapiens*
<400> 12

gtttggggcc agagtgggcg aggcgaggag gtctggccta taaagtagtc ggggagacgg 60
ggtgctgggt tgcgtcgtag tctcctgcag cgtctggggg ttccgttgca gtcctcggaa 120
ccaggacctc ggcgtggcct agcgagttat ggcgacgaag gccgtgtgcg tgctgaaggg 180
cgacggccca gtgcagggca tcatcaattt cgagcagaag gaaagtaatg gaccagtgaa 240
ggtgtgggga agcattaaag gactgactga aggcctgcat ggattccatg ttcattgagtt 300
tgagagataat acagcaggct gtaccagtgc aggtcctcac tttaatcctc tatccagaaa 360
acacgggtggg ccaaaggatg aagagaggca tgttgagac ttgggcaatg tgactgctga 420
caaagatggt gtggccgatg tgtctattga agattctgtg atctcactct caggagacca 480
ttgcatcatt ggccgcacac tgggtgtcca tgaaaaagca gatgacttgg gcaaagggtg 540
aatgaagaa agtaciaaaga caggaaacgc tggaaagtcgt ttggcttgtg gtgtaattgg 600
gatcgcccaa taaacattcc cttggatgta gtctgaggcc ccttaactca tctgttatcc 660
tgctagctgt agaaatgtat cctgataaac attaaacact gtaatcttaa aagtgttaatt 720
gtgtgacttt ttcagagttg ctttaaagta cctgtagtga gaaactgatt tatgatcact 780
tggaagattt gtatagtttt ataaaactca gttaaaatgt ctgtttcaat gacctgtatt 840
ttgccagact taaatcacag atgggtatta aacttgtcag aatttctttg tcattcaagc 900
ctgtgaataa aaacctgta tggcacttat tatgaggcta ttaaaagaat ccaaattcaa 960
actaaaaaaaa aaaaaaaaa a 981

10
15 <210> 13
<211> 508
<212> ADN
<213> *Homo sapiens*
<400> 13

ES 2 614 978 T3

```

tttggtgctt tggatccatt tccatcggtc ottacagccg ctgctcagac tccagcagcc      60

aagatgggta agcagatcga gagcaagact gcttttcagg aagccttgga cgctgcagggt      120
gataaacttg tagtagttga cttctcagcc acgtgggtg ggccttgcaa aatgatcaag      180
cctttctttc attccctctc tgaaaagtat tccaacgtga tattccttga agtagatgtg      240
gatgactgtc aggatgttgc ttcagagtgt gaagtcaaat gcatgccaac attccagttt      300
tttaagaagg gacaaaaggt ggggtgaattt totggagcca ataaggaaaa gcttgaagcc      360
accattaatg aattagtcta atcatgtttt ctgaaaatat aaccagccat tggctattta      420
aaacttgtaa tttttttaat ttacaaaaat ataaaatatg aagacataaa cccagttgcc      480
atctgogtga caataaaaca ttaatgct                                          508

```

5 <210> 14
 <211> 1591
 <212> ADN
 <213> *Homo sapiens*
 <400> 14

ES 2 614 978 T3

ccctgogtct ctgcccgcc cgtggcgccc gagtgcactg aagatggcgg ctgctgtagg 60
 acggttgctc cgagcgtcgg ttgcccgaca tgtgagtgcc attccttggg gcatttctgc 120
 cactgcagcc ctgagcctg ctgcatgtgg aagaacgagc ttgacaaatt tattgtgttc 180
 tggttccagt caagcaaaat tattcagcac cagttcctca tgccatgcac ctgctgtcac 240
 ccagcatgca ccctatttta aggttacagc cgttgtcaat ggagagttca aagacctaa 300
 ccttgatgac ttaaggggga aatatttggg gcttttcttc tatecttggg atttcacctt 360
 tgtgtgtcct acagaaattg ttgcttttag tgacaaagct aacgaatttc acgacgtgaa 420
 ctgtgaagtt gtcgcagtct cagtggattc ccactttagc catcttgcct ggataaatac 480
 accaaggaag aatggtgggt tgggccacat gaacatcgca ctcttgtcag acttaactaa 540
 gcagatttcc cgagactacg gtgtgctgtt agaaggttct ggtcttgac taagaggtct 600
 cttcataatt gacccaatg gagtcatcaa gcatttgagc gtcaacgatc tcccagtggg 660
 ccgaagcgtg gaagaaacc tccgcttggg gaaggcgttc cagtatgtag aaacacatgg 720
 agaagtctgc ccagogaact ggacaccgga ttctcctaog atcaagcaa gtccagctgc 780
 ttccaaagag tactttcaga aggtaaatca gtagatcacc catgtgtatc tgcaccttct 840
 caactgagag aagaaccaca gttgaaacct gcttttatca ttttcaagat ggttatttgt 900
 agaaggcaag gaaccaatta tgcttgtatt cataagtatt actctaaatg ttttgttttt 960
 gtaattctgg ctaagacctt ttaaaccatgg ttagttgcta gtacaaggaa tcctttattg 1020
 gtaacatctt ggtggctggc tagctagttt ctacagaaca taatttgcct ctatagaagg 1080
 ctattcttag atcatgtctc aatggaaaca ctcttcttcc ttagccttac ttgaatcttg 1140
 cctataataa agtagagcaa cacacattga aagcttctga tcaacgggcc tgaattttc 1200

 atcttgaatg tctttgtatt aaactgaatt ttcttttaag ctaacaaaga tcataatfff 1260
 caatgattag ccgtgtaact cctgcaatga atgtttatgt gattgaagca aatgtgaatc 1320
 gtattatfff aaaaagtggc agagtgactt aactgatcat gcatgatccc tcatccctga 1380
 aattgagttt atgtagtcat tttacttatt ttattcatta gctaactttg tctatgtata 1440
 tttctagata ttgattagtg taatcgatta taaaggatat ttatcaaate cagggattgc 1500
 attttgaaat tataattatt ttctttgctg aagtattcat tgtaaaacat acaaaataaa 1560
 catattttaa aacatttgca ttttaccacc a 1591

5

- <210> 15
- <211> 1246
- <212> ADN
- <213> *Homo sapiens*
- <400> 15

ES 2 614 978 T3

```

gtctttgccc tcgcgacgcc gccacctccg gaacaagcca tggtagcgcc gacggtaggca      60
gcggcgtggc tgctcctgtg ggctgcggcc tgcgcgcagc aggagcagga cttctacgac      120
ttcaaggcgg tcaacatccg gggcaaactg gtgtcgctgg agaagtaccg cggatcgggtg      180
tccttggtgg tgaatgtggc cagcgagtgc ggcttcacag accagcacta ccgagccctg      240
cagcagctgc agcgagacct gggccccac cactttaacg tgctgcctt cccctgcaac      300
cagtttgccc aacaggagcc tgacagcaac aaggagattg agagcttgc ccgccgacc      360
tacagtgtct cattccccat gtttagcaag attgcagtca ccggtactgg tgcccatcct      420
gccttcaagt acctggccca gacttctggg aaggagccca cctggaactt ctggaagtac      480
ctagtagccc cagatggaaa ggtggtaggg gcttgggacc caactgtgtc agtggaggag      540
gtcagacccc agatcacagc gctcgtgagg aagctcatcc tactgaagcg agaagactta      600
taaccaccgc gtctcctcct ccaccacctc atcccgccca cctgtgtggg gctgaccaat      660
gcaaactcaa atggtgcttc aaaggagag acccactgac tctcctcct ttactcttat      720
gccattggtc ccatcattct tgtgggggaa aaattctagt attttgatta tttgaaatctt      780
acagcaacia ataggaactc ctggccaatg agagctcttg accagtgaat caccagccga      840
tacgaacgtc ttgccaacia aatgtgtgg caaatagaag tatatcaagc aataatctcc      900
caccgaaggc ttctgtaaac tgggaccaat gattacctca tagggctgtt gtgaggatta      960
ggatgaaata cctgtgaaag tgccataggc gtgccagcca aataggaggc attcaatgaa     1020
cattttttgc atataaacca aaaaataact tgttatcaat aaaaacttgc atccaacatg     1080
aatttccagc cgatgataat ccaggccaaa ggttttagttg ttgttatttc ctctgtatta     1140
tttcttcat tacaaaagaa atgcaagttc attgtaacia tccaaacaat acctcacgat     1200
ataaaataaa aatgaaagta tcctcctcaa aaaaaaaaaa aaaaaa                       1246

```

- <210> 16
- <211> 942
- <212> ADN
- <213> *Homo sapiens*
- <400> 16

ES 2 614 978 T3

gagcgcctctg gagggcgtgg ccgtgggaaa ggagggcgcgg aaagccgacg cgcgtccatt 60
 ggtcggcttg acgaggggag gagccgctgg ctcccagccc cgccgcgatg agcctcggcc 120
 gcctttgccg cctactgaag cgggcgctgc tctgtggggc tctggccgcg cctggcctgg 180
 ccgggaccat gtgcgcgtcc cgggacgact ggcgctgtgc gcgctccatg cacgagtttt 240
 cggccaagga catcgacggg cacatggtta acctggacaa gtaccggggc ttcgtgtgca 300
 tcgtcaccaa cgtggcctcc cagtgaggca agaccgaagt aaactacact cagctcgtcg 360
 acctgcacgc ccgatacget gagtgtggtt tgcggatcct ggccctcccg tgtaaccagt 420
 tcgggaagca ggagccaggg agtaacgaag agatcaaaga gttcgccgcg ggctacaacg 480
 tcaaattcga tatgttcagc aagatctcgc tgaacgggga cgacgccac ccgctgtgga 540
 agtggatgaa gatccaaccc aagggcaagc gcatcctggg aatgccatc aagtggact 600
 tcaccaagtt cctcatcgac aagaacggct gcgtggtgaa gcgctacgga cccatggagg 660
 agcccctggt gatagagaag gacctgcccc actatttcta gctccacaag tgtgtggccc 720
 cgcccagacc cctgcccacg cccttgagc cttccaccgg cactcatgac ggccctgcctg 780
 caaacctgct ggtggggcag acccgaaaat ccagcgtgca ccccgccgga ggaagggtccc 840
 atggcctgct gggcttggct cggcgcccc acccctggct accttgtggg aataaacaga 900
 caaattagcc tgctggaaaa aaaaaaaaaa aaaaaaaaaa aa 942

<210> 17
 <211> 1342
 <212> ADN
 <213> *Homo sapiens*

5

<400> 17

gaggaagtga cgacagggct gcccttgaca ggcagggagg gctaggctgt gcatccctcc 60
 gctcgcattg cagggagatg gtcacgcgac ttcttctgag gaggttcctg gcctctgtca 120
 tctccaggaa gccctctcag ggtcagtggc caccctcac ttccagagcc ctgcagaccc 180
 cacaatgcag tcctggtggc ctgactgtaa cacccaaccc agcccggaca atatacacca 240
 cgaggatctc cttgacaacc tttaatatcc aggatggacc tgactttcaa gaccgagtgg 300
 tcaacagtga gacaccagtg gttgtggatt tccacgcaca gtgggtgtgga ccttgcaaga 360
 tcctggggcc gaggttagag aagatggtgg ccaagcagca cgggaagggtg gtgatggcca 420
 aggtggatat tgatgaccac acagacctcg ccattgagta tgaggtgtca gcggtgcccc 480

10

ES 2 614 978 T3

```

ctgtgctggc catgaagaat ggggacgtgg tggacaagtt tgtgggcatc aaggatgagg      540
atcagttgga ggccttcctg aagaagctga ttggctgaca agcagggatg agtcctggtt      600
cccttgcccg cgtgggaccc caatagaact cagcccttcc atgccagccc ttctgtctgc      660
ctccctcctg tctggctcct ggggcccattg cttagagccc aggctccagc cctgagtgtc      720
tccgagctgg cggactgccc aggggccatc agaggatggt ggtgctgctg ctgatccggg      780
gaccgctgtc ttccctcca tacgcctttc atccctcctt ctagggccta tggcagttct      840
cccaggatgt gtggcgagag cctgggcccag cccacagcgt tcctagtcag gcagccacac      900
cttggctcctc atcttggctc cttccaatct gaaacctcgt gcctggctcg tctgccacct      960
acatttctct ttccagctgc tgttttgtaa aaagaaaag aaaaaagaag cccaaactag     1020
tgagagtaat atctaattat ctcatTTTTT gtaggctctgt gataaagaac ttagtcatcc     1080
cttccacctc ctactgtgaa gaacagaccc tgggtcccac actgaaatcc cctctagtca     1140
cccattcca cccccaggg agctgcctcc caggcagggg gtgcagaaaa tgattgatgg     1200
gctggggaac cctggagagc ctcgactccg gaagtctcaa ggtgcctcct cctctcctta     1260
gctggcccgt tggTTTTctg agcagggggc tgaactgtga acaagtcaga caaataaagc     1320
aagggtctgc accatcaaaa aa                                             1342

```

5

<210> 18
 <211> 921
 <212> ADN
 <213> *Homo sapiens*

 <400> 18

ES 2 614 978 T3

gcggcgctcg egccaagga cgtgtttctg cgctcgctg gtcattggagg cgetgcccgt 60
 gctagccgcg acaactccgg accacggccg ccaccgaagg ctgcttctgc tgcgctact 120
 gctgttcctg ctgcccggctg gagctgtgca gggctgggag acagaggaga ggccccggac 180
 tcgcaagag gagtgccact tctacgaggg tggacaagtg taccgggag aggcattccc 240
 ggtatcggtc gccgaccact ccctgcacct aagcaaagcg aagatttcca agccagcgcc 300
 ctactgggaa ggaacagctg tgatcgatgg agaatttaag gagctgaagt taactgatta 360
 tcgtgggaaa tacttggttt tcttcttcta cccacttgat ttcacatttg tgtgtccaac 420
 tgaaattatc gcttttggcg acagacttga agaattcaga tctataaata ctgaagtggg 480
 agcatgctct gttgattcac agtttaccga tttggcctgg attaataccc ctccaagaca 540
 aggaggactt gggccaataa ggattccact tctttcagat ttgaccatc agatctcaaa 600
 ggactatggt gtatacctag aggactcagg ccacactctt agaggtctct tcattattga 660
 tgacaaagga atcctaagac aaattactct gaatgatctt cctgtgggta gatcagtggg 720
 tgagacacta cgtttgggtc aagcattcca gtacactgac aaacacggag aagtctgccc 780
 tgctggctgg aaacctggta gtgaaacaat aatcccagat ccagctggaa agctgaagta 840
 ttcgataaa ctgaattgag aaatacttct tcaagttatg atgcttgaaa gttctcaata 900
 aagttcacgg tttcattacc a 921

<210> 19
 <211> 1200
 5 <212> ADN
 <213> *Homo sapiens*
 <400> 19

ES 2 614 978 T3

cagttaaaag gaggcgctg ctggcctccc cttacagtgc ttgttcgggg cgctccgctg 60
gcttcttga caattgcgcc atgtgtgctg ctcgctagc ggcggcgccg gcggcgccc 120
agtcggtgta tgccttctcg gcgcgcccgc tggccggcgg gagcctgtg agcctgggct 180
ccctgcgggg caaggtacta cttatcgaga atgtggcgtc cctctgagge accacggctc 240
gggactacac ccagatgaac gagctgcagc ggcgcctcgg accccggggc ctggtggtgc 300
toggettccc gtgcaaccag tttgggcata aggtgcgccc ggcggagcgg ggcggggcgg 360
gggcgacagt gcagtagtgg ctggggcgc cgccggtgtg ctggtgggtg ccgtcggctc 420
catgcgcca gagtctggct actctctcgt ttcctttctg ttgctcgtag ctgctgaaat 480
tcctctccgc ccttgggatt gcgcatggag ggcaaatcc cggtgactca tagaaaatct 540
cccttgtttg tggtagaac gtttctctcc tcctcttgac cccgggttct agctgccctt 600
ctctcctgta ggagaacgcc aagaacgaag agattctgaa ttccctcaag tacgtccggc 660
ctggtggtgg gttcagccc aacttcatgc tcttcgagaa gtgcgaggtg aacgggtcgg 720
gggcgcccc tctcttcgcc ttctcgggg aggcctgcc agctccagc gacgaccca 780
ccgcgcttat gaccgacccc aagctcatca cctggtctcc ggtgtgtcgc aacgatgtt 840
cctggaactt tgagaagttc ctggtgggcc ctgacggtgt gccctacgc aggtacagcc 900
gocgctcca gaccattgac atcgagcctg acatcgaagc cctgctgtct caagggcca 960
gctgtgceta gggcgccct cctaccocgg ctgcttgcca gttgcagtgc tgetgtctcg 1020
ggggggttt catctatgag ggtgttctc ctaaacctac gagggaggaa cacctgatct 1080
tacagaaaat accacctcga gatgggtgct ggtcctgtt atcccagtct ctgccagacc 1140
aaggcgagt tcccactaa taaagtccg ggtgtcagca gaaaaaaaa aaaaaaaaa 1200

<210> 20
<211> 1274
<212> ADN
<213> *Homo sapiens*

5

<400> 20

gcttctgtct ggcggcgcca gcatggcggc gggggcggt gaggcagctg tagcggccgt 60
ggaggaggtc ggctcagccc ggcagtttga ggagctgctg cgcctcaaag ccaagtcctc 120

10

ES 2 614 978 T3

```

ccttgtggtc catttctggg caccatgggc tccacagtgt gcacagatga acgaagttat 180
ggcagagtta gctaaagaac tccctcaagt ttcatttgtg aagttggaag ctgaaggtgt 240
tcctgaagta tctgaaaaat atgaaattag ctctgttccc acttttctgt ttttcaagaa 300
ttctcagaaa atcgaccgat tagatgggtgc acatgccccca gagttgacca aaaaagttca 360
gogacatgca tctagtggct ccttcctacc cagcgcta at gaacatctta aagaagatct 420
caaccttcgc ttgaagaaat tgactcatgc tgccccctgc atgctgttta tgaaaggaac 480
tcctcaagaa ccacgctgtg gtttcagcaa gcagatgggtg gaaattcttc acaaacataa 540
tattcagttt agcagttttg atatcttctc agatgaagag gttcgacagg gactcaaagc 600
ctattccagt tggcctacct atcctcagct ctatgtttct ggagagctca taggaggact 660
tgatataatt aaggagctag aagcatctga agaactagat acaatttgc ccaaagctcc 720
caaattagag gaaaggctca aagtgtgac aaataaagct tctgtgatgc tctttatgaa 780
aggaaacaaa caggaagcaa aatgtggatt cagcaaacaa attctggaaa tactaaatag 840
tactggtggt gaatatgaaa cattcgatat attggaggat gaagaagttc ggcaaggatt 900
aaaagcttac tcaaattggc caacatacc tcagctgtat gtgaaagggg agctggtggg 960
aggattggat attgtgaagg aactgaaaga aaatggtgaa ttgctgccta tactgagagg 1020
agaaaattaa taaatcttaa acttggtgcc caactattgt aagaaatatt taattacatt 1080
gggagcagtt catgatttag tcctcagaaa tggactagga atagaaaatt cctgctttct 1140
cagttacatg ttttgtgtat ttcacaatgt cgtgctaaat aaatgtatgt tacatttttt 1200
tcccacccaaa aatagaatgc aataaacatc ttcaaattat taacgaaaaa aaaaaaaaaa 1260
aaaaaaaaaa aaaa 1274

```

5 <210> 21
 <211> 710
 <212> ADN
 <213> *Homo sapiens*
 <400> 21

ES 2 614 978 T3

gctcgtccgc tccctcccc gcgccgtgca cgtcttggtt cgggccgggc ataaaaggct	60
tcggggccca gggctcactt ggcgctgaga acgcggtcc acgcgtgtga tcgtccgtgc	120
gtctagcctt tgcccacgca gctttcagtc atggcctccg gtaacgcgcg catcggaaaag	180
ccagcccctg acttcaaggc cacagcgggtg gttgatggcg ccttcaaaga ggtgaagctg	240
tcggactaca aagggaagta cgtggtcctc tttttctacc ctctggactt cacttttgtg	300
tgccccaccg agatcatcgc gttcagcaac cgtgcagagg acttccgcaa gctgggctgt	360
gaagtgctgg gcgtctcggg ggactctcag ttcaccacc tggcttggta tgagcagggg	420
ccaaagaggg aggttgcagc taagctcaca ccctcaggtc ctagcagtgt ggcttcgtgg	480
ccattgctca acctctggaa cctgcgtttc cccatcgtga aaataatgga aacattgccg	540
cccaagtctt taaggatgat gacagtaatt agcatttgac aactagttgc ctggtatata	600
gagttgcaga tgcaactcag atgcaactct atctactcta tgtacttagt tcccaggagg	660
gaggctgtgc tgccctatth catgaagatg gaaactccag ttcaccgaag	710

<210> 22
 <211> 1715
 <212> ADN
 <213> *Homo sapiens*
 <400> 22

5

ES 2 614 978 T3

gaaccaaccg gttgcttgct gtcccagcgg cgccccctca tcaccgtcgc catgcccgga 60
 ggtctgcttc tcggggacgt ggctcccaac tttgaggcca ataccaccgt cggccgcctc 120
 cgtttccacg actttctggg agactcatgg ggcattctct tctcccaccc tcgggacttt 180
 accccagtgt gcaccacaga gcttggcaga gctgcaaagc tggcaccaga atttgccaaag 240
 aggaatgtta agttgattgc cttttcaata gacagtgttg aggaccatct tgccctggagc 300
 aaggatatca atgottacaa ttgtgaagag cccacagaaa agttaccttt tcccctcctc 360
 gatgatagga atcgggagct tgccatcctg ttgggcatgc tggatccagc agagaaggat 420
 gaaaagggca tgccctgtgac agctcgtgtg gtgtttgttt ttggtcctga taagaagctg 480
 aagctgtcta tcctctaccc agctaccact ggcaggaact ttgatgagat tctcagggtg 540
 gtcatctctc tccagctgac agcagaaaaa agggttgccca cccagttga ttggaaggat 600
 ggggatagtg tgatggctct tccaaccatc cctgaagaag aagccaaaaa acttttcccg 660
 aaaggagtct tcaccaaaga gctcccctct ggcaagaaat acctccgcta cacaccccag 720
 ccttaagtct cttggagaag ctgggtgctgt gagccagagg atgtcagctg ccaattgtgt 780
 tttcctgcag caattccata aacacatcct ggtgtcatca cagccaaggt ttttaggttg 840
 ctataccaat ggcttattaa atgaaaatgg cactaaaagt ttcttgagat tctttatact 900
 ctctgccttc agcaatcaat tccattcata catcagcact ctgctgggtc tgtttgaaat 960
 atgttctgta tttaaaactc aaatcttgtt ggatctctgc agggcttgtg accaatgaag 1020
 tcatatattg tgatggttga caaagcttgc ttcactccat cagagaaatga ctatcaattt 1080
 ttttttaact gtccatcac gtccctcctc gtcaccatt ttgaagagtg gcagaacttg 1140
 aagttcaact tcctctgtaa atatccaagt ataaagccca ggaacttcta gaataacca 1200
 gatgcgcttt aatttttttt aatatgtttt gatcacagaa cttctagaat aaccagatg 1260
 ctctttcata ttcttttaac acatcttgat cacagctggg ggaaaaaag ctttttaatt 1320
 ctataccttc ctagtagata agtgaagagc agggaaagag acctttaaat attttgctat 1380
 aaaaaaattt gtgataagtt tctatcaaaa tggggagatt gcagaaaag cttcccttgg 1440
 ctoccaaagga ggtgtagcag gtgtgagcaa tattagtgcc atgtgccttt cacacagggt 1500
 ttgcatttat cagtctgttt tccgatgatg tgtacatgaa agagtacacc atgtgaagag 1560
 aagagagaat gattgaaaat gttttagtat agaactcttc ttgcagtggg ttgctatttt 1620
 ctagatttta ctttttaggg aacaaaataa aatcctttgt taaaactggg aaaaaaaaaa 1680
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaa 1715

<210> 23
 <211> 1182
 <212> ADN
 <213> *Homo sapiens*

<400> 23

ES 2 614 978 T3

```

aatgagggcc tccagggggc gggtcggact gccgcgggcc ggggagcgcct ctgggtggcc      60
agctgtgggc ccgggcccgtc gtgggctccg gcttgcgtgc ggagatgagc gggtcacctcg      120
gccgagctgc ggcggctctg ctccgctggg ggcgcggcgc gggcggcggg ggcctttggg      180
gtccgggctg gcgggcccgc ggctcgggcg cgggcggcgg cggctcggcg gagcagttgg      240
acgcgctggg gaagaaggac aaggtggtgg tcttcctcaa ggggacgcgg gagcagcccc      300
agtgcggctt cagcaacgcc gtggtgcaga tcttcggcgt gcacggcgtc cgcgattacg      360
cggcctacaa cgtgctggac gacccggagc tccgacaagg cattaagac tattccaact      420
ggcccaccat ccogcaagtg tacctcaatg gcgagtttgt agggggctgt gacattcttc      480
tgcagatgca ccagaatggg gacttgggtg agaactgaa aaagctgggg atccactccg      540
cccttttaga tgaaaagaaa gaccaagact ccaagtgagg gcggccaagt cctcgtgag      600
cagagagggg gccgttcatg tcagagactc actgccagaa aagcottaacc cattttggtt      660
ttcactattg agaaccgcaac tgcttgcact gatcattttg gttcgtgagc agttggtgat      720
tttagttggt ctggtgttcg ggctaagaat attttattgt ggacttaatt acaaccactg      780
cactgtaatg attcaatgct gtattatgat attgctgtaa acaaaattca ttcttatatt      840
gtcacttatt ctttgcctga ttcagaagtt aataggagc tttggaatca ttattcatga      900
cccctctgca aatgtgtcag tctccaaaga gagtatctcc ccccaaattt tgtgtagctt      960
cttttgttat ggaaaatggt gaacaaaaaa agaaactgtg ataactgggg cgttgttttt     1020
taaaataaac tccagcacag ggatgctgtg catgcctgag ttgattccga agtgcataatg     1080
tctgtaagga tttggagtgc ctgcagtgtt ttatgtgtgg gaagtaaggg tgagctctcat     1140
attcttctat taaatttgcc acaagaattg caaaaaaaaa aa                          1182

```

<210> 24
 <211> 1170
 <212> ADN
 <213> *Homo sapiens*

5

<400> 24

ES 2 614 978 T3

aaggctatta ttaccaccac tgagtggctt aaataatcct gtcaacagca atcgcccatt 60
 tccaaagcca tggtgaaaca tctctgtgct aatttctttt gttttgtttc ctaatttttt 120
 ttttttggca ggtggtggga aataatcttt gtcttctttg gagtaaacct tcaacaccgg 180
 attttttctt ttaattatgg atgtaaacc ccaatcccc ataatttaca ttgggtctcg 240
 accaattgcc taattataag aggatattatt taggctctta tttcatccac acaaaaactt 300
 gtgtaacagg tagttgaaa catctgaggc accactttga ttctgttttg gatggctcatg 360
 ttttttctcc tccgtttccc cagcatgtct gccaccatcc tcatgcactg cttccaagtg 420
 cctgggagcc tttatgagcg tccctaaacc taaaagaatc cagagggcggg gctcggatga 480
 accctcgaga taagcaagtg agccgcttct cccctctaaa ggatgtttac acgtgggtgg 540
 cactcgctgg aatccagcgc tcgggcagcc ctgggaggac gcgctcagct gcgaggagga 600
 tggagagcaa tacatcatca tctttggaga atttagcgac ggcgcctgtg aaccagatcc 660
 aagaaacaat ttctgataat tgtgtggtga ttttctcaaa aacatcctgt tcttactgta 720
 caatggcaaa aaagcttttc catgacatga atgttaacta taaagtgggt gaactggacc 780
 tgcttgaata tggaaaccag ttccaagatg ctctttacaa aatgactggt gaaagaactg 840
 ttccaagaat atttgtcaat ggtactttta ttggaggtgc aactgacact cataggcttc 900
 acaaagaagg aaaattgctc ccactagttc atcagtgtta tttaaaaaaa agtaagagga 960
 aagaatttca gtgatgttta tactaataag ttgctagta cagtgtcagt tattttaaagt 1020
 ggtaatgccc gataatgtct tttaaatggt tgaggatggt ttaaatacat gcattgtctt 1080
 cacgaagaag atgtaaaaat aatgaacaat aaattgcggt ggaaacctaa aaaaaaaaaa 1140
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1170

5 <210> 25
 <211> 18
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> oligonucleótido sentido

<400> 25
 gggagctgag ggcaagtc 18

15 <210> 26
 <211> 19
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> oligonucleótido antisentido

<400> 26
 gagaccctg cagccaatc 19

25 <210> 27
 <211> 19
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> oligonucleótido sentido

 <400> 27
 5 caaccagttt gggcatcag 19

 <210> 28
 <211> 19
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 28
 15 gttcacctcg cacttctcg 19

 <210> 29
 <211> 18
 <212> ADN
 20 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido

 <400> 29
 25 gctggtttg agcaggag 18

 <210> 30
 <211> 24
 <212> ADN
 30 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 30
 35 ccaaagatga tgatgtattg ctct 24

 <210> 31
 <211> 18
 <212> ADN
 40 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido

 <400> 31
 50 cgcagttcgg ttctccac 18

 <210> 32
 <211> 18
 <212> ADN
 <213> Secuencia artificial
 55

 <220>
 <223> oligonucleótido antisentido

 <400> 32
 60 gggccccgaa ctgtgtca 18

 <210> 33
 <211> 21
 <212> ADN
 65 <213> Secuencia artificial

<220>
 <223> oligonucleótido sentido

 <400> 33
 5 cactgacaaa catggggaag t 21

 <210> 34
 <211> 21
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 34
 15 ttgctcttt tggacatcag g 21

 <210> 35
 <211> 18
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido
 25
 <400> 35
 caccctgga tgttcaa 18

 <210> 36
 <211> 22
 30 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido
 35
 <400> 36
 ggacaccagc gaatcatcta gt 22

 <210> 37
 <211> 20
 40 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido
 45
 <400> 37
 50 tccactgcaa ggaacaacag 20

 <210> 38
 <211> 19
 <212> ADN
 <213> Secuencia artificial
 55
 <220>
 <223> oligonucleótido antisentido

 <400> 38
 60 taagcgtgct cccacacat 19

 <210> 39
 <211> 21
 <212> ADN
 65 <213> Secuencia artificial

<220>
 <223> oligonucleótido sentido

 <400> 39
 5 tgccagctta ggaataacca g 21

 <210> 40
 <211> 18
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 40
 15 cctgcaccaa caatgacg 18

 <210> 41
 <211> 20
 <212> ADN
 20 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido

 <400> 41
 25 ggcttctgga attgtcgat 20

 <210> 42
 <211> 21
 <212> ADN
 30 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 42
 35 tgcacccgcc tatacaatct t 21

 <210> 43
 <211> 21
 <212> ADN
 40 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido

 <400> 43
 45 gcctccagt acacagacga g 21

 <210> 44
 <211> 21
 <212> ADN
 50 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 44
 55 gttgggctta atcgtgtcac t 21

 <210> 45
 <211> 18
 <212> ADN
 60 <213> Secuencia artificial

 <210> 45
 <211> 18
 <212> ADN
 65 <213> Secuencia artificial

<220>
 <223> oligonucleótido sentido

 <400> 45
 5 tcctggctga tcccactg 18

 <210> 46
 <211> 20
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 46
 15 atgcatcct gtaccacat 20

 <210> 47
 <211> 20
 <212> ADN
 20 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido

 <400> 47
 25 gcatcatcaa tttcgagcag 20

 <210> 48
 <211> 20
 <212> ADN
 30 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 48
 35 caggcctca gtcagtcct 20

 <210> 49
 <211> 18
 <212> ADN
 40 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido

 <400> 49
 45 ttacagccgc tcgtcaga 18

 <210> 50
 <211> 21
 <212> ADN
 50 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 50
 55 ggcttctga aaagcagtct t 21

 <210> 51
 <211> 20
 <212> ADN
 60 <213> Secuencia artificial

 65

<220>
 <223> oligonucleótido sentido

 <400> 51
 5 ctggacaccg gatttccta 20

 <210> 52
 <211> 25
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 52
 15 gggatgatcta ctgatttacc ttctg 25

 <210> 53
 <211> 20
 <212> ADN
 20 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido

 <400> 53
 25 ccatcctgcc ttcaagtacc 20

 <210> 54
 <211> 20
 <212> ADN
 30 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 54
 35 ttccatctgg ggctactagg 20

 <210> 55
 <211> 18
 <212> ADN
 40 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido

 <400> 55
 45 tacggaccga tggaggag 18

 <210> 56
 <211> 22
 <212> ADN
 50 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 56
 60 ccacacactt gtggagctag aa 22

 <210> 57
 <211> 20
 <212> ADN
 65 <213> Secuencia artificial

<220>
 <223> oligonucleótido sentido

 <400> 57
 5 gagacaccag tgggttgga 20

 <210> 58
 <211> 19
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 58
 15 gctggccac catctctc 19

 <210> 59
 <211> 20
 <212> ADN
 20 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido

 <400> 59
 25 gcacctaagc aaagcgaaga 20

 <210> 60
 <211> 21
 <212> ADN
 30 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 60
 35 aaattctcca tcgatcacag c 21

 <210> 61
 <211> 21
 <212> ADN
 40 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido

 <400> 61
 45 ccctgttg tggtagaac g 21

 <210> 62
 <211> 21
 <212> ADN
 50 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 62
 55 gagagaaggg cagctagaac c 21

 <210> 63
 <211> 19
 <212> ADN
 60 <213> Secuencia artificial

 65

<220>
 <223> oligonucleótido sentido

 <400> 63
 5 tcctcaagaa ccacgctgt 19

 <210> 64
 <211> 27
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido

 <400> 64
 15 tgagaagata tcaaaactgc taaactg 27

 <210> 65
 <211> 26
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido
 25
 <400> 65
 gcaactcaga tgcaactcta tctact 26

 <210> 66
 <211> 23
 30 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> oligonucleótido antisentido
 35
 <400> 66
 tgaactggag ttccatctt cat 23

 <210> 67
 <211> 24
 40 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> oligonucleótido sentido
 45
 <400> 67
 50 caatagacag tgttgaggac catc 24

 <210> 68
 <211> 20
 <212> ADN
 <213> Secuencia artificial
 55
 <220>
 <223> oligonucleótido antisentido

 <400> 68
 60 ttctgtggg ctctcaca 20

 <210> 69
 <211> 20
 <212> ADN
 65 <213> Secuencia artificial

ES 2 614 978 T3

<220>
 <223> oligonucleótido sentido

5 <400> 69
 gtgataactg gggcggtgtt 20

10 <210> 70
 <211> 18
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> oligonucleótido antisentido

20 <400> 70
 actcaggcat gcacagca 18

25 <210> 71
 <211> 18
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> oligonucleótido sentido

35 <400> 71
 gtggcactcg ctggaatc 18

40 <210> 72
 <211> 22
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> oligonucleótido antisentido

<400> 72
 cgtcgctaaa ttctcaaag at 22

<210> 73
 <211> 1236
 <212> ADN
 <213> *Homo sapiens*

45 <400> 73

actctcgoga gatccctact ggctataaag gcagcgcccc ggagagctct tgcgcgtctt 60
 gttcttgccct ggtgtcgggtg gttagtctct gcgacttgtg ttgggactgg tgagtgtggg 120
 cagtgcggcc cctgcggagt gaggcgcggc gcgcccttct tgccctgttc ctcttctcc 180

ES 2 614 978 T3

tccctgtccgg ggcccgcccg cgctcgggtg ggggtgctgt gatgcgtgag gcagccgggg	240
gaggcccgga gtccgagact gcttgagcgc tgcgcacacc cctctcgtgg gccccccacg	300
tagctgatag gaagatgtct tcaggaaatg ctaaaattgg gcaccctgcc cccaacttca	360
aagccacagc tgttatgcca gatggtcagt ttaaagatat cagcctgtct gactacaaag	420
gaaaatatgt tgtgttcttc ttttaccctc ttgacttcac ctttgtgtgc cccacggaga	480
tcattgcttt cagtgatagg gcagaagaat ttaagaaact caactgccaa gtgattggtg	540
cttctgtgga ttctcacttc tgtcatctag catgggtcaa tacacctaag aaacaaggag	600
gactgggacc catgaacatt cctttggtat cagacccgaa gcgcaccatt gctcaggatt	660
atggggctct aaaggctgat gaaggcatct cgttcagggg cctttttatc attgatgata	720
agggtattct tcggcagatc actgtaaatg acctccctgt tggccgctct gtggatgaga	780
ctttgagact agttcaggcc ttccagttca ctgacaaaca tggggaagtg tgcccagctg	840
gctggaaacc tggcagtgat accatcaagc ctgatgtcca aaagagcaaa gaatatttct	900
ccaagcagaa gtgagcgtg ggctgtttta gtgccaggct gcggtgggca gccatgagaa	960
caaaacctct tctgtatfff tttttccat tagtaaaaca caagacttca gattcagccg	1020
aattgtggtg tcttacaagg caggcctttc ctacaggggg tggagagacc agcctttctt	1080
cctttggtag gaatggcctg agttggcgtt gtgggcaggc tactggtttg tatgatgat	1140
tagtagagca acccattaat cttttgtagt ttgtattaaa cttgaactga gaccttgatg	1200
agtctttaa aaaaaaaaaa aaaaaaaaaa aaaaaa	1236

<210> 74
 <211> 1042
 <212> ADN
 <213> *Homo sapiens*

 <400> 74

5

ES 2 614 978 T3

```

actctcgcga gatccctact ggctataaag gcagcgcccc ggagagctct tgcgcgtctt      60
gttcttgcct ggtgtcgggtg gttagtttct gcgacttgtg ttgggactgc tgataggaag      120
atgtcttcag gaaatgctaa aattgggcac cctgccccca acttcaaagc cacagctggt      180
atgccagatg gtcagtttaa agatatcagc ctgtctgact acaaaggaaa atatgttgtg      240
ttcttctttt accctcttga cttcaccttt gtgtgccccca cggagatcat tgctttcagt      300
gatagggcag aagaatttaa gaaactcaac tgccaagtga ttggtgcttc tgtggattct      360
cacttctgtc atctagcatg ggtcaataca cctaagaaac aaggaggact gggacccatg      420
aacattcctt tggatcaga cccgaagcgc accattgctc aggattatgg ggtcttaaag      480
gctgatgaag gcatctcgtt caggggcctt tttatcattg atgataaggg tattcttcgg      540
cagatcactg taaatgacct ccctgttggc cgctctgtgg atgagacttt gagactagtt      600

caggccttcc agttcactga caaacatggg gaagtgtgcc cagctggctg gaaacctggc      660
agtgatacca tcaagcctga tgtccaaaag agcaaagaat atttctcaa gcagaagtga      720
gcgctgggct gttttagtgc caggctgcgg tgggcagcca tgagaacaaa acctcttctg      780
tatttttttt ttccattagt aaaacacaag acttcagatt cagccgaatt gtggtgtctt      840
acaaggcagg cctttcctac aggggggtgga gagaccagcc tttcttcctt tggtaggaat      900
ggcctgagtt ggcgttgtgg gcaggctact ggtttgtatg atgtattagt agagcaaccc      960
attaatcttt tgtagtttgt attaaacttg aactgagacc ttgatgagtc tttaaaaaaa     1020
aaaaaaaaaa aaaaaaaaaa aa                                             1042

```

<210> 75
 <211> 1035
 <212> ADN
 <213> *Homo sapiens*
 <400> 75

5

ES 2 614 978 T3

gcggtgccct tgcggcgag ctggggtcgc ggcctgctc cccgcgcttt cttaggccc 60
 gcggggcgcg caggagcggc actcgtggct gtggtggctt cggcagcggc ttcagcagat 120
 cggcggcatc agcggtagca ccagcactag cagcatgttg agccgggcag tgtgcggcac 180
 cagcaggcag ctggctccgg ttttggggta tctgggctcc aggcagaagc acagcctccc 240
 cgacctgcc c tacgactacg gcgcctgga acctcacatc aacgcgcaga tcatgcagct 300
 gcaccacagc aagcaccacg cggcctacgt gaacaacctg aacgtcaccg aggagaagta 360
 ccaggaggcg ttggccaagg gagatgttac agcccagata gctcttcagc ctgcactgaa 420
 gttcaatggt ggtggtcata tcaatcatag cttttctggy acaaacctca gccctaacgg 480
 tgggtggagaa cccaagggg agttgctgga agccatcaaa cgtgactttg gttcctttga 540
 caagtttaag gagaagctga cggtgcatc tgttggtgtc caaggctcag gttggggttg 600
 gcttggtttc aataaggaac ggggacactt acaaattgct gcttgtccaa atcaggatcc 660
 actgcaagga acaacaggcc ttattccact gctggggatt gatgtgtggg agcacgctta 720
 ctaccttcag tataaaaatg tcaggcctga ttatctaaaa gctatttga atgtaatcaa 780
 ctgggagaat gtaactgaaa gatacatggc ttgcaaaaag taaaccacga tcgttatgct 840
 gatcataccc taatgatccc agcaagataa tgcctgtct tctaagatgt gcatcaagcc 900
 tggtagatac tgaaaacct ataaggctct ggataatttt tgtttgatta ttcattgaag 960
 aaacatttat tttccaattg tgtgaagttt ttgactgta ataaaagaat ctgtcaacca 1020
 tcaaaaaaaaa aaaaa 1035

<210> 76
 <211> 918
 <212> ADN
 <213> *Homo sapiens*

 <400> 76

5

ES 2 614 978 T3

gcggtgccct tgcggcgcag ctggggtcgc ggcctgctc cccgcgctt cttaggccc 60
 gcgggcggcg caggagcggc actcgtggct gtggtggctt cggcagcggc ttcagcagat 120
 cggcggcatc agcggtagca ccagcactag cagcatggtt agccgggcag tgtgcggcac 180
 cagcaggcag ctggctccgg ttttgggta tctgggctcc aggcagaagc acagcctccc 240
 cgacctgcc tacgactacg gcgccttggg acctcacatc aacgcgcaga tcatgcagct 300
 gcaccacagc aagcaccacg cggcctacgt gaacaacctg aacgtcaccg aggagaagta 360
 ccaggaggcg ttggccaagg gggagttgct ggaagccatc aaacgtgact ttggttcctt 420
 tgacaagttt aaggagaagc tgacggctgc atctgttggg gtccaaggct caggttgggg 480
 ttggttggg ttcaataagg aacggggaca cttacaaatt gctgcttgtc caaatcagga 540
 tccactgcaa ggaacaacag gccttattcc actgctgggg attgatgtgt gggagcacgc 600
 ttactacctt cagtataaaa atgtcaggcc tgattatcta aaagctattt ggaatgtaat 660
 caactgggag aatgtaactg aaagatacat ggcttgcaaa aagtaaacca cgatcgttat 720
 gctgatcata ccctaatagat cccagcaaga taatgtcctg tcttctaaga tgtgcatcaa 780
 gcctgttaca tactgaaaac cctataaggt cctggataat ttttgttga ttattcattg 840
 aagaaacatt tattttccaa ttgtgtgaag tttttgactg ttaataaaag aatctgtcaa 900
 ccatcaaaaa aaaaaaaaa 918

<210> 77
 <211> 3087
 <212> ADN
 <213> *Homo sapiens*

5

<400> 77

tcctcctggg tcttgccctag cggcggggcgc atgcttagtc accgtgaggc tgcgcttgcc 60
 cggggcccgc gccccctac cccggggacc gccccgggc cggccgccc acttggcgcg 120
 ccacttccgc gtgcatggcc ctgctgcccc gagccctgag cggcggcgcg ggaccgagct 180
 ggcggcgggc ggcgcgcgcc ttccgaggct tcctgctgct tctgcccag cccgcggccc 240
 tcacgcgcgc cctctcccgt gccatggcct gcaggcagga gccgcagccg cagggccgcg 300
 cgcccgtgct tggcgcctg gcctcctatg actacctggt gatcgggggc ggctcgggcg 360
 ggctggccag cgcgcgcagg gcggccgagc tgggtgccag ggccgcctg gtggagagcc 420
 acaagctggg tggcacttgc gtgaatggtg gatgtgtacc caaaaaggta atgtggaaca 480
 cagctgtcca ctctgaattc atgcatgac atgctgatta tggctttcca agttgtgagg 540
 gtaaattcaa ttggcgtggt attaaggaaa agcgggatgc ctatgtgagc cgctgaaatg 600

10

ES 2 614 978 T3

ccacatcatca aaacaatctc accaagtccc atatagaaat catccgtggc catgcagcct 660
 tcacgagtga tcccaagccc acaatagagg tcagtgggaa aaagtacacc gccccacaca 720
 tcctgatcgc cacaggtggt atgccctcca cccctcatga gagccagatc cccggtgcca 780
 gcttaggaat aaccagcgat ggattttttc agctggaaga attgcccggc cgcagcgtca 840
 ttgttggtgc aggttacatt gctgtggaga tggcagggat cctgtcagcc ctgggttota 900
 agacatcact gatgatacgg catgataaag tcaaggaggt taaaaagact ttgtcgggct 960
 tggaagtcat catggttact gcagttcccg gtaggtacc agtcatgacc atgattccag 1020
 atgttgactg cctgctctgg gccattgggc gggccccgaa taccaaggac ctgagtttaa 1080
 aaaaactggg gattcaaacc gatgacaagg gtcatatcat cgtagacgaa ttccagaata 1140
 ccaacgtcaa aggcacatctat gcagttgggg atgtatgtgg aaaagctctt cttactccag 1200
 ttgcaatagc tgctggccga aaacttgccc atcgactttt tgaatataag gaagattcca 1260
 aattagatta taacaacatc ccaactgtgg tcttcagcca cccccctatt gggacagtgg 1320
 gactcacgga agatgaagcc attcataaat atggaataga aaatgtgaag acctattcaa 1380
 cgagctttac cccgatgtat cacgcagtta ccaaaaggaa aacaaaatgt gtgatgaaaa 1440
 tggctctgtc taacaaggaa gaaaaggtgg ttgggatcca tatgcagggg cttgggtgtg 1500
 atgaaatgct gcagggtttt gctggtgcag tgaagatggg agcaacgaag gcagactttg 1560
 acaacacagt cgccattcac cctacctctt cagaagagct ggtcacactt cgttgagaac 1620
 caggagacac gtgtggcggg cagtgggacc catagatctt ctgaaatgaa acaaataatc 1680
 acattgactt actgtttgag ttttatgtat tcttttattt taatcaggat cttctgatag 1740
 tggaaatfff tagtacataa tagaacttat ttatggagtt agaaatftgt agtgttatcc 1800
 aggattgatt ttcatttgat cacatctcac agtaattaat attttcaagt ttttttttta 1860
 ttaacagctc tgtgctagtt tttttttct gttttagcct catcccaaat ataaagcttt 1920
 gtgaagtaca attaacttaa tgtacttgaa tgaatagaac ttgctacttt tttttttttt 1980
 ttttttgaga cagagtfttg ctctcattgc ccaggctgga gtgcgggtgg gctatttccag 2040
 ctaccacaaa cctctgcctc ctgggttcaa gtgattctcc tgccttagcc tcccgaatag 2100
 ctggaattac aggcacgcac caccatgcct gactaatftt gtattftttag tagacatggg 2160
 gtttctccat gttggtcagg ctggctcaa actcccacct tcaggtgatc cgcccacctc 2220
 ggcctcctga ggtgctgaga ttacaggcgt gagccactgt gccagcttgc taatfttcac 2280
 agaagttgat ggcaattctt cacatgtaa cagtgccagt gcacagaacc tttatatatt 2340
 ttttgaagcc agtactgtgc tctgcatata acaaagctgc ttcaaggatg agacctfttt 2400
 ctaaaagcat gtaatgtgag aagccggcct gcctatfttt cttftttctt ttttaatgat 2460
 taaaaatagt ttgtggcaag gcacgggtggc tcaggcctgt aattctagca ctttgggagg 2520

ES 2 614 978 T3

ccgaggcagg aggattactt gagcctacaa gtttgaggcc agcatgcaca gcatagcaag 2580
 actgcatctc tacagagagt aaaaaaatt acccgagtgt ggtgatgtgc atctgtaatc 2640
 tcagctactt gggaggctga ggtgagagga tcaacttgagc ttgggtgagg tgaggctgca 2700
 gtgagtctctg atcatgctgc tgcactcaat cttggacaac agagcaagac cctgtctcaa 2760
 aaaaaaaaaa aaaaaatata tatatatata tatattattt ttatgaggtg aagtgcataca 2820
 aacttgggaa agatttgagg aggctgggaa cctcctggaa aaccactcct tgaagaaaga 2880
 tatgagagac atttagaagt gattcctgct ttcagaagga ggtggattca aatacatcaa 2940
 aagtcccttc ctctgctaag tgtttatagt tcaatgaata atttcaatat ttgtatgtgt 3000
 tcttgtcatt ttatTTTTT ctgaaaaact tccaaaaatt tgaaaataaa attacagcct 3060
 tttcttctta taaaaaaaaa aaaaaaa 3087

5 <210> 78
 <211> 3014
 <212> ADN
 <213> *Homo sapiens*
 <400> 78

tcctcctggg tcttgcctag cggcggggcgc atgcttagtc accgtgaggc tgcgcttgcc 60
 cggggcccgc gccccctac ccoggggacc gccccgggc cggccgccc acttggcgcg 120
 ccaactccgc gtgcatggcc ctgctgcccc gagccctgag cggcggcgcg ggaccgagct 180
 ggcggcgggc ggcgcgcgcc ttccgaggct tcctgctgct tctgcccag cccgcggccc 240
 tcacgcgcgc cctctcccgt gccatggcct gcaggcagga gccgcagccg cagggcccgc 300
 cgcccgetgc tggcgcctg gcctcctatg actacctggt gatcgggggc ggctcgggcg 360
 ggctggccag cgcgcgcagg gcggccgagc tgggtgccag ggccgcccgt gtggagagcc 420
 acaagctggg tggcacttgc gtgaatggtg gatgtgtacc caaaaaggtg atgtggaaca 480
 cagctgtcca ctctgaattc atgcatgatc atgctgatta tggctttcca agttgtgagg 540
 gtaaattcaa ttggcgtggt attaaggaaa agcgggatgc ctatgtgagc cgcctgaatg 600
 ccatctatca aaacaatctc accaagtccc atatagaaat catccgtggc catgcagcct 660
 tcacgagtga tcccaagccc acaatagagg tcagtgggaa aaagtacacc gccccacaca 720
 tcctgatcgc cacaggtggt atgcctcca cccctcatga gagccagatc cccggtgcca 780
 gcttaggaat aaccagcgat ggattttttc agctggaaga attgcccggc cgcagcgtca 840
 ttgttggtgc aggttacatt gctgtggaga tggcagggat cctgtcagcc ctgggttcta 900
 agacatcact gatgatacgg catgataag tacttagaag ttttgattca atgatcagca 960
 ccaactgcac ggaggagctg gagaacgctg gcgtggaggt gctgaagttc tcccagggga 1020
 ttcaaaccga tgacaagggt catatcatcg tagacgaatt ccagaatacc aacgtcaaag 1080

10

ES 2 614 978 T3

gcatctatgc agttggggat gtatgtggaa aagctcttct tactccagtt gcaatagctg 1140
 ctggccgaaa acttgcccat cgactttttg aatataagga agattccaaa ttagattata 1200
 acaacatccc aactgtggtc ttcagccacc cccctattgg gacagtggga ctcacggaag 1260
 ataagccatt cataaatatg gaatagaaaa tgtgaagacc tattcaacga gctttacccc 1320
 gatgtatcac gcagttacca aaaggaaaaac aaaatgtgtg atgaaaatgg tctgtgctaa 1380
 caaggaagaa aaggtggttg ggatccatat gcagggactt ggggtgtgatg aaatgctgca 1440
 gggttttgct gttgcagtga agatgggagc aacgaaggca gactttgaca acacagtcgc 1500
 cattcacctt acctcttcag aagagctggt cacacttcgt tgagaaccag gagacacgtg 1560
 tggcgggcag tgggacccat agatcttctg aaatgaaaca aataatcaca ttgacttact 1620
 gtttgagttt tatgtatttc tttattttaa tcaggatctt ctgatagtgg aaatttttag 1680
 tacataatag aacttattta tggagttaga aattttagt gttatccagg attgattttc 1740
 atttgatcac atctcacagt aattaatatt ttcaagtttt ttttttatta acagctctgt 1800
 gctagttttt tttttctggt ttagcctcat cccaaatata aagctttgtg aagtacaatt 1860
 aacttaatgt acttgaatga atagaacttg ctactttttt tttttttttt tttgagacag 1920
 agttttgctc tcattgccc a ggctggagtg cgggtgggtgct atttcagctc accacaacct 1980
 ctgcctcctg ggttcaagtg attctcctgc cttagcctcc cgaatagctg gaattacagg 2040
 cacgcaccac catgcctgac taattttgta ttttttagtag acatgggggt tctccatggt 2100
 ggtcaggctg gtctcaaaact cccaccttca ggtgatccgc ccacctgggc ctctgaggt 2160
 gctgagatta caggcgtgag ccactgtgcc agcttgctaa ttttcacaga agttgatggc 2220
 aattcttcac atgtaaacag tgccagtgca cagaaccttt atatattttt tgaagccagt 2280
 actgtgctct gcatataaca aagctgcttc aaggatgaga cctttttcta aaagcatgta 2340
 atgtgagaag ccggcctgcc ttattttctt ttttctttt taatgattaa aaatagtttg 2400
 tggcaaggca cggtggtca ggcctgtaat tctagcactt tgggaggccg aggcaggagg 2460
 attacttgag cctacaagtt tgaggccagc atgcacagca tagcaagact gcatctctac 2520
 agagagtaaa aaaaattacc cgagtgtggt gatgtgcatc tgtaatctca gctacttggg 2580
 aggctgaggt gagaggatca cttgagcttg ggtgaggtga ggctgcagtg agtcctgatc 2640
 atgctgctgc actcaatctt ggacaacaga gcaagaccct gtctcaaaaa aaaaaaaaaa 2700
 aaatatatat atatatatat attattttta tgaggtgaag tgcatcaaac ttgggaaaga 2760
 tttgaggagg ctgggaacct cctggaaaac cactccttga agaaagatat gagagacatt 2820
 tagaagtgat tctgctttc agaaggaggt ggattcaaat acatcaaaag tcccttcctc 2880
 tgctaagtgt ttatagttca atgaataatt tcaatatttg tatgtgttct tgtcatttta 2940
 ttttttctg aaaaacttcc aaaaatttga aaataaaatt acagcctttt cttcttataa 3000
 aaaaaaaaaa aaaa 3014

ES 2 614 978 T3

<210> 79
 <211> 2928
 <212> ADN
 <213> *Homo sapiens*

5

<400> 79

tcctcctggg tcttgccctag cggcggggcgc atgcttagtc accgtgaggc tgcgcttgcc	60
cggggcccgc gccccctac cccggggacc gccccgggc cgcccgcgcc acttggcgcg	120
ccacttcgcg gtgcatggcc ctgctgcccc gagccctgag cgccggcgcg ggaccgagct	180
ggcggggggc ggcgcgogcc ttccgaggct tctgctgct tctgcccag cccgcggccc	240
tcacgcgcgc cctctcccgt gccatggcct gcaggcagga gcgcagccg caggggccgc	300
cgccogetgc tggcgccgtg gcctcctatg actacctggt gatcgggggc ggctcgggcg	360
ggctggccag cgcgcgcagg gcggccgagc tgggtgccag ggccgcctg gtggagagcc	420
acaagctggg tggcacttgc gtgaatggtg gatgtgtacc caaaaaggta atgtggaaca	480
cagctgtcca ctctgaattc atgcatgac atgctgatta tggctttcca agttgtgagg	540
gtaaattcaa ttggcgtggt attaaggaaa agcgggatgc ctatgtgagc cgctgaatg	600
ccatctatca aaacaatctc accaagtccc atatagaaat catccgtggc catgcagcct	660
tcacgagtga tcccagccc acaatagagg tcagtgggaa aaagtacacc gccccacaca	720
tcctgatcgc cacaggtggt atgcctcca cccctcatga gagccagatc cccggtgcca	780
gcttaggaat aaccagcgat ggattttttc agctggaaga attgcccggc cgcagcgtca	840
ttgttggtgc aggttacatt gctgtggaga tggcagggat cctgtcagcc ctgggttcta	900
agacatcact gatgatacgg catgataagg ggattcaaac cgatgacaag ggtcatatca	960
tcgtagacga attccagaat accaacgtca aaggcatcta tgcagttggg gatgtatgtg	1020
gaaaagctct tcttactcca gttgcaatag ctgctggccg aaaacttgcc catcgacttt	1080
ttgaatataa ggaagattcc aaattagatt ataacaacat cccaactgtg gtcttcagcc	1140
acccccctat tgggacagtg ggactcacgg aagatgaagc cattcataaa tatggaatag	1200
aaaatgtgaa gacctattca acgagcttta ccccgatgta tcacgcagtt accaaaagga	1260
aaacaaaatg tgtgatgaaa atggtctgtg ctaacaagga agaaaagggtg gttgggatoc	1320
atatgcaggg acttggtgtg gatgaaatgc tgcagggttt tgctgttgca gtgaagatgg	1380
gagcaacgaa ggcagacttt gacaacacag tcgccattca ccctacctct tcagaagagc	1440
tggtcacact tcgttgagaa ccaggagaca cgtgtggcgg gcagtgggac ccatagatct	1500
tctgaaatga aacaaataat cacattgact tactgtttga gttttatgta tttctttatt	1560

ES 2 614 978 T3

ttaatcagga tcttctgata gtggaaattt ttagtacata atagaactta tttatggagt 1620
 tagaaatttg tagtgttatc caggattgat tttcatttga tcacatctca cagtaattaa 1680
 ttttttcaag tttttttttt attaacagct ctgtgctagt tttttttttc tgttttagcc 1740
 tcatcccaa tataaagctt tgtgaagtac aattaactta atgtacttga atgaatagaa 1800
 cttgctactt tttttttttt tttttttgag acagagtttt gctctcattg occaggctgg 1860
 agtgcggttg tgctatttca gctcaccaca acctctgcct cctgggttca agtgattctc 1920
 ctgccttagc ctcccgaata gctggaatta caggcacgca ccaccatgcc tgactaattt 1980
 tgtattttta gtagacatgg ggtttctcca tgttggtcag gctggtctca aactcccacc 2040
 ttcaggtgat ccgcccacct cggcctcctg aggtgctgag attacaggcg tgagccactg 2100
 tgccagcttg ctaattttca cagaagttga tggcaattct tcacatgtaa acagtgccag 2160
 tgcacagaac ctttatatat tttttgaagc cagtactgtg ctctgcatat aacaaagctg 2220
 cttcaaggat gagacctttt tctaaaagca tgtaatgtga gaagccggcc tgccttattt 2280
 tcttttttct tttttaatga ttaaaaatag tttgtggcaa ggcacggtgg ctcaggcctg 2340
 taattctagc actttgggag gccgaggcag gaggattact tgagcctaca agtttgaggc 2400
 cagcatgcac agcatagcaa gactgcatct ctacagagag taaaaaaaaat taccggagtg 2460
 tgggtgatgtg catctgtaat ctcagctact tgggaggctg aggtgagagg atcacttgag 2520
 cttgggtgag gtgaggctgc agtgagtcct gatcatgctg ctgcactcaa tcttggacaa 2580
 cagagcaaga ccctgtctca aaaaaaaaaa aaaaaaatat atatatatat atatattatt 2640
 tttatgaggt gaagtgcac aaacttggga aagatttgag gaggctggga acctcctgga 2700
 aaaccactcc ttgaagaaag atatgagaga catttagaag tgattcctgc tttcagaagg 2760
 aggtggattc aaatacatca aaagtccctt cctctgctaa gtgtttatag ttcaatgaat 2820
 aatttcaata tttgtatgtg ttcttgtcat tttatttttt tctgaaaaac ttccaaaaat 2880
 ttgaaaataa aattacagcc ttttcttctt ataaaaaaaa aaaaaaaa 2928

<210> 80
 <211> 1093
 <212> ADN
 <213> *Homo sapiens*

5

<400> 80

ES 2 614 978 T3

attgcattcc tgggcattgc taactagtga agtataccag atggaaatgt cttcgaagct 60
 gtcocctttaa aactcgagca agctaccagc caaactccgc ctccaggag gttccttatt 120
 aaataggagc caactggctg ggtoggggct caatacccca agcaatacct gcaactgagg 180
 attcttcccg gggagaccgc agcccatcgg catggctcaa gagtttgtga actgcaaaat 240
 ccagcctggg aaggtggttg tgttcatcaa gccacactgc ccgtactgca ggagggccca 300

agagatcctc agtcaattgc ccatcaaaca agggcttctg gaatttgtcg atatcacagc 360
 caccaaccac actaacgaga ttcaagatta tttgcaacag ctccaggag caagaacggt 420
 gcctcgagtc tttattggta aagattgtat aggcggatgc agtgatctag tctctttgca 480
 acagagtggg gaactgctga cgcggtctaa gcagattgga gctctgcagt aaccacagat 540
 ctcataggaa atgttcaaca attctgtgaa aggtcacagg acccaattgg agaaatcata 600
 tgaaaagcat agttggtcct ggtgtcatat ggatcagagg cacaagtgca gaggctgtgg 660
 tcatgoggaa cactctgtta ttttaagatgg ctatccagat aatcctgaac actgtgtatt 720
 tattttatth agactaccag caaagattaa agcatgaaat gtaaaacatc tgataaaaact 780
 tacagcccc tacaccaaga gtgtatctgt gaaagagctc ctacactttg aaaacttaag 840
 aatcccttat catgaagttt gcctgttcta gaattgtaag attgttaatt tccttcaatc 900
 tctagtgaca aacttaatt tcttttctaa taaaaaaaaac ctatagatga ttcagtgatt 960
 tttgtccaat tcatttgcac gttctcaaga cattaaggaa tgttatgca aatacactaa 1020
 cttaaaactg tgtttatatt tggccctgcc attataaata aagacacgtg ctgctgtcaa 1080
 aaaaaaaaaaaa aaa 1093

5 <210> 81
 <211> 827
 <212> ADN
 <213> *Homo sapiens*

<400> 81

ES 2 614 978 T3

gcagtggagg cggcccaggc ccgccttccg caggggtgtcg ccgctgtgcc gctagcggtg 60
 ccccgctgc tgcggtggca ccagccagga ggcggagtgg aagtggccgt ggggcgggta 120
 tgggactagc tggcgtgtgc gccctgagac gctcagcggg ctatatactc gtcggtgggg 180
 ccggcgggtca gtctgctggca gcggcagcaa gacgggtgcag tgaaggagag tgggcgtctg 240
 gcgggggtccg cagtttcagc agagccgctg cagccatggc cccaatcaag gtgggagatg 300
 ccatcccagc agtggaggtg tttgaagggg agccagggaa caaggtgaac ctggcagagc 360
 tgttcaaggg caagaagggt gtgctgtttg gagttcctgg ggccttcacc cctggatgtt 420
 ccaaggttcg gtcctggct gatcccactg ggccttttg gaaggagaca gacttattac 480
 tagatgattc gctggtgtcc atctttggga atcgacgtct caagaggttc tccatggtgg 540
 tacaggatgg catagtgaag gccctgaatg tggaaaccaga tggcacaggc ctcacctgca 600
 gcctggcacc caatatcacc tcacagctct gaggccttg gccagattac ttcctccacc 660
 cctccctacc tcacctgcc agccctgtgc tggggccctg caattggaat gttggccaga 720
 tttctgcaat aaacacttgt ggtttgcggc caaaaaaaaa aaaaaaaaaa aaaaaaaaaa 780
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaa 827

<210> 82
 <211> 692
 <212> ADN
 <213> *Homo sapiens*

5

<400> 82

gcagtggagg cggcccaggc ccgccttccg caggggtgtcg ccgctgtgcc gctagcggtg 60
 ccccgctgc tgcggtggca ccagccagga ggcggagtgg aagtggccgt ggggcgggta 120
 tgggactagc tggcgtgtgc gccctgagac gctcagcggg ctatatactc gtcggtgggg 180
 ccggcgggtca gtctgctggca gcggcagcaa gacgggtgcag tgaaggagag tgggcgtctg 240
 gcgggggtccg cagtttcagc agagccgctg cagccatggc cccaatcaag gttcggctcc 300
 tggctgatcc cactggggcc tttgggaagg agacagactt attactagat gattcgtctg 360
 tgtccatctt tgggaatcga cgtctcaaga ggtctccat ggtggtacag gatggcatag 420
 tgaaggccct gaatgtgaa ccagatggca caggcctcac ctgcagcctg gcaccaata 480
 tcactcaca gctctgaggc cctgggcccag attacttctt ccacctctc ctatctcacc 540
 tgcccagccc tgtgctgggg ccctgcaatt ggaatgttg ccagatttct gcaataaaca 600
 cttgtggttt gcggccaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 660
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aa 692

10

<210> 83
 <211> 1537
 <212> ADN
 <213> *Homo sapiens*

15

<400> 83

ES 2 614 978 T3

ccctgctct ctgcccgcc cgtggcgccc gagtgcactg aagatggcgg ctgctgtagg	60
acggttgctc cgagcgtcgg ttgcccgaca tgtgagtgcc attccttggg gcatttctgc	120
cactgcagcc ctcaggcctg ctgcatgtgg aagaacgagc ttgacaaatt tattgtgttc	180
tggttccagt caagcaccct attttaaggg tacagccgtt gtcaatggag agttcaaaga	240
cctaagcctt gatgacttta aggggaaata tttgggtgctt ttcttctatc ctttggattt	300
cacctttgtg tgtcctacag aaattggtgc ttttagtgac aaagctaacg aatttcacga	360
cgtgaactgt gaagtgtcgc cagtctcagt ggattccac tttagccatc ttgcctggat	420
aaatacacca aggaagaatg gtggtttggg ccacatgaac atcgcaactc tgtcagactt	480
aactaagcag atttcccgag actacgggtg gctgtagaa ggttctggtc ttgcactaag	540
aggtctcttc ataattgacc ccaatggagt catcaagcat ttgagcgtca acgatctccc	600
agtgggcca agcgtggaag aaaccctccg cttgggtgaag gcgttccagt atgtagaac	660
acatggagaa gtctgccag cgaactggac accggattct cctacgatca agccaagtcc	720
agctgcttcc aaagagtact ttcagaaggt aaatcagtag atcacccatg tgtatctgca	780
ccttctcaac tgagagaaga accacagttg aaacctgctt ttatcatttt caagatggtt	840
atgttagaa ggcaaggaac caattatgct tgtattcata agtattactc taaatgtttt	900
gtttttgtaa ttctggctaa gaccttttaa acatggtag ttgctagtac aaggaatcct	960
ttattgtaa catcttggtg gctggctagc tagtttctac agaacataat ttgcctctat	1020
agaaggctat tcttagatca tgtctcaatg gaaacactct tctttcttag ccttacttga	1080
atcttgcta taataaagta gagcaacaca cattgaaagc ttctgatcaa cggtcctgaa	1140
atcttcatct tgaatgtctt tgtattaaac tgaattttct ttaagctaa caaagatcat	1200
aattttcaat gattagccgt gtaactcctg caatgaatgt ttatgtgatt gaagcaaatg	1260
tgaatcgtat tttttaaaa agtggcagag tgacttaact gatcatgcat gatccctcat	1320
ccotgaaatt gagtttatgt agtcatttta cttattttat tcattagcta actttgtcta	1380
tgtatatttc tagatattga ttagtgtaat cgattataaa ggatatttat caaatccagg	1440
gattgcattt tgaaattata attattttct ttgctgaagt attcattgta aaacatacaa	1500
aataaacata ttttaaaaca tttgcatttt accacca	1537

<210> 84
 <211> 964
 <212> ADN
 <213> *Homo sapiens*

 <400> 84

5

ES 2 614 978 T3

```

gagcgctctg gagggcgtgg ccgtgggaaa ggagggcgcgg aaagccgacg cgcgtccatt      60
ggtcggctgg acgaggggag gagccgctgg ctcccagccc cgcgcgatg agcctcggcc      120
gcctttgccg cctactgaag ccggcgctgc tctgtggggc tctggccgcg cctggcctgg      180
ccgggaccat gtgcgcgtcc cgggacgact ggcgctgtgc gcgctccatg cacgagtttt      240
ccgccaagga catogacggg cacatggtta acctggacaa gtaccggggc ttcgtgtgca      300
tcgtcaccaa cgtggcctcc cagtgaggca agaccgaagt aaactacact cagctcgtcg      360
acctgcacgc ccgatacgtc gagtgtggtt tgcggatcct ggccttcccg tgtaaccagt      420
tcgggaagca ggagccaggg agtaacgaag agatcaaaga gttcgccgcg ggctacaacg      480
tcaaattcga tatgttcagc aagatctgcg tgaacgggga cgacgccac ccgctgtgga      540
agtggatgaa gatccaaccc aagggcaagg gcatcctggg aaatgccatc aagtggaact      600
tcaccaagtt tggacacegt ctctccacag ttctcatcg acaagaacgg ctgcgtggtg      660
aagcgctacg gacccatgga ggagcccctg gtgatagaga aggacctgcc cactatttc      720
tagctccaca agtgtgtggc cccgcccag cccctgccca cgcccttga gccttcacc      780
ggcactcatg acggcctgcc tgcaaacctg ctggtggggc agaccgaaa atccagcgtg      840
caccocgccc gaggaaggtc ccatggcctg ctgggcttgg ctcggcgcc ccaccctgg      900

ctaccttgtg ggaataaaca gacaaattag cctgctggaa aaaaaaaaaa aaaaaaaaaa      960
aaaa                                             964

```

<210> 85
 <211> 1031
 <212> ADN
 <213> *Homo sapiens*

 <400> 85

5

ES 2 614 978 T3

agtcctgact	acggcctcog	ggccctttgt	ccccgctagc	ggcgcctcggg	gtgggggagc	60
caggaggggc	gggagacggg	cgggtatggg	ccgcgcgggc	gcaggctccc	ccgggcgccc	120
caggcagcgg	tgccagagcc	ggggcagggc	gcggccgcga	gcccctcggc	ggcggaaggc	180
cccagcgtgc	aggcgcagga	gggcgcggcg	ccgcgcggaag	aagccctgtc	cccgcagctt	240
gcgaccggag	atccacgaat	gtcccaagtc	ccaggacccg	tgcgcgtccc	gggacgactg	300
gcgctgtgcg	cgctccatgc	acgagttttc	cgccaaggac	atcgacgggc	acatggttaa	360
cctggacaag	taccggggct	togtgtgcat	cgtcaccaac	gtggcctccc	agtgaggcaa	420
gaccgaagta	aactacactc	agctcgtcga	cctgcacgcc	cgatacgtg	agtgtggttt	480
gcggatcctg	gccttcccgt	gtaaccagtt	cgggaaagcag	gagccaggga	gtaacgaaga	540
gatcaaagag	ttcgccgcgg	gctacaacgt	caaattcgat	atgttcagca	agatctgcgt	600
gaacggggac	gacgccacc	cgctgtggaa	gtggatgaag	atccaacca	agggcaaggg	660
catcctggga	aatgccatca	agtggaactt	caccaagttc	ctcatcgaca	agaacggctg	720
cgtggtgaag	cgctacggac	ccatggagga	gcccctggtg	atagagaagg	acctgcccc	780
ctatcttag	ctccacaagt	gtgtggcccc	gcccagagccc	ctgcccacgc	ccttggagcc	840
ttccaccggc	actcatgacg	gcctgcctgc	aaacctgctg	gtggggcaga	cccgaaaatc	900
cagcgtgcac	cccgcggag	gaaggtccca	tggcctgctg	ggcttggctc	ggcgccccca	960
ccctggcta	ccttgtggga	ataaacagac	aaattagcct	gctggaaaaa	aaaaaaaaaa	1020
aaaaaaaaa	a					1031

REIVINDICACIONES

1. Un procedimiento *in vitro* para el diagnóstico, de una leucemia mieloide aguda o un trastorno mielodisplásico, comprendiendo dicho procedimiento las etapas de:
 - 5 a) medir, a partir de las células contenidas en una muestra biológica de un sujeto, preferentemente de las células sanguíneas o células de médula ósea que contiene la muestra, el nivel de expresión de al menos 6 genes que pertenecen a un grupo de genes escogidos de entre un grupo de 24 genes, comprendiendo o estando constituido dicho grupo de 24 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 a 24, comprendiendo o estando constituidos dichos 6 genes por las secuencias de ácido nucleico de SEQ ID NO: 1 a 6,
 - 10 b) comparar el nivel de expresión de cada uno de los genes medidos en la etapa a), con el nivel de expresión de los mismos genes respectivos de células contenidas en una muestra de control, preferentemente de células sanguíneas o células de médula ósea que contiene la muestra, siendo dicha muestra de control de la misma naturaleza que dicha muestra biológica, para establecer una relación de nivel de expresión para cada uno de los genes de dichos 6 genes, y
 - 15 c) determinar el estado de dicha muestra biológica de manera que si la relación establecida en la etapa b) para al menos 3 genes de entre dichos 6 genes es ≥ 2 o $\leq 0,5$, dicha muestra biológica es representativa de una leucemia mieloide aguda o un trastorno mielodisplásico.
2. El procedimiento de acuerdo con la reivindicación 1, en el que si la relación establecida en la etapa b) es
 - $\leq 0,3$ para el gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1, y
 - 20 - ≥ 3 para los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 y 3,
 entonces, dicha muestra biológica es representativa de una leucemia mieloide aguda.
3. El procedimiento de acuerdo con la reivindicación 1, en el que la etapa c) es de tal manera que si la relación establecida en la etapa b) para al menos 3 genes de entre dichos 6 genes es ≥ 2 o $\leq 0,5$, a condición de que
 - 25 - la relación entre el nivel de expresión del gen que comprende o está constituido por la secuencia de ácido nucleico de SEQ ID NO: 1 medida en dicha muestra biológica y medida en dicha muestra de control, no es $\leq 0,3$, o
 - las relaciones entre el nivel de expresión de cada uno de los genes que comprenden o están constituidos por las secuencias de ácido nucleico de SEQ ID NO: 2 o 3 medido en dicha muestra biológica y medido en dicha muestra de control, no es ≥ 3 ,
- 30 entonces dicha muestra biológica es representativa de un trastorno mielodisplásico, en particular una mielodisplasia escogida de entre anemia refractaria, citopenia refractaria o anemia refractaria con exceso de blastos, síndrome 5q.
4. El procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 3, en el que el nivel de expresión de los genes se mide por un procedimiento cuantitativo, en particular RT-qPCR.
5. El procedimiento de acuerdo con cualquiera de las reivindicaciones 1 a 4, en el que la medición del nivel de expresión de al menos dichos 6 genes en la etapa a) se lleva a cabo utilizando al menos las moléculas de nucleótidos de SEQ ID NO: 25-36.
- 35 6. Una composición que comprende al menos 12 oligonucleótidos escogidos de entre una biblioteca de 48 oligonucleótidos que comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO: 25-72, comprendiendo dicha composición al menos los 12 oligonucleótidos que comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO: 25-36.
- 40 7. La composición de acuerdo con la reivindicación 6, para su uso en el diagnóstico, y/o la clasificación, preferentemente *in vitro*, de una leucemia mieloide aguda o un trastorno mielodisplásico.
8. Un kit que comprende al menos 12 oligonucleótidos escogidos de entre un grupo de 48 oligonucleótidos que comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO: 25-72, dichos al menos 12 oligonucleótidos comprenden o consisten en las secuencias de ácido nucleico de SEQ ID NO: 25-36.
- 45

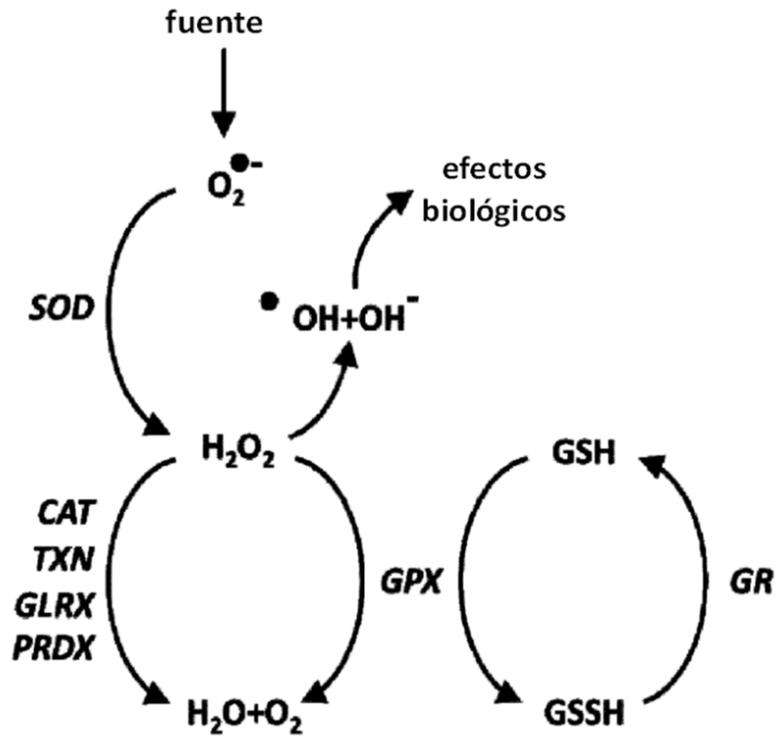


Fig. 1

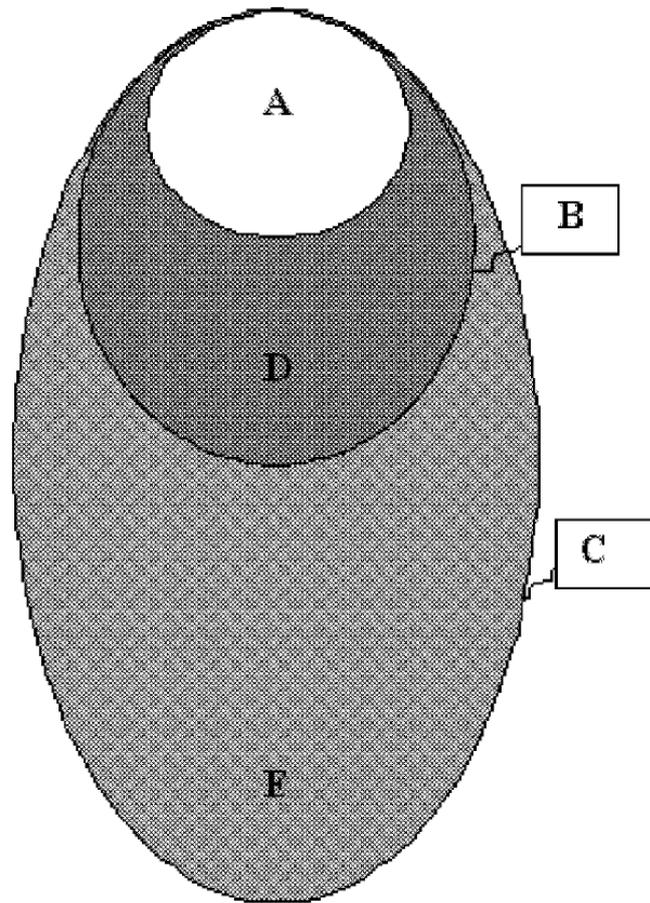


Fig. 2

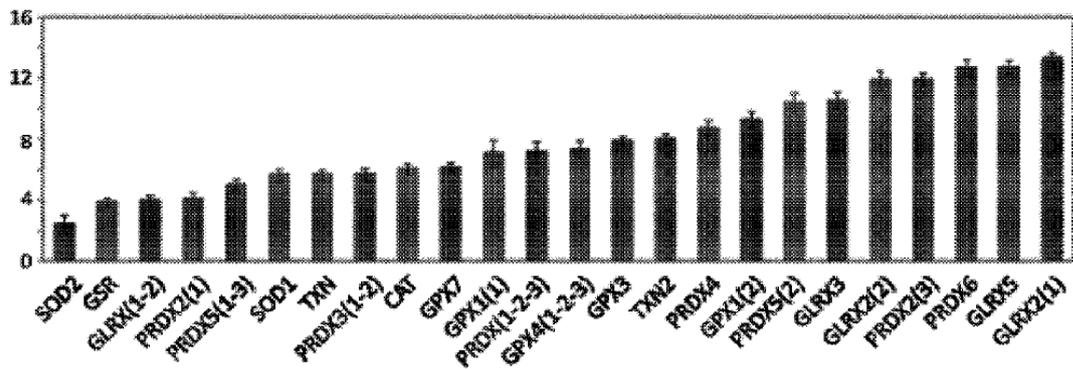


Fig. 3A

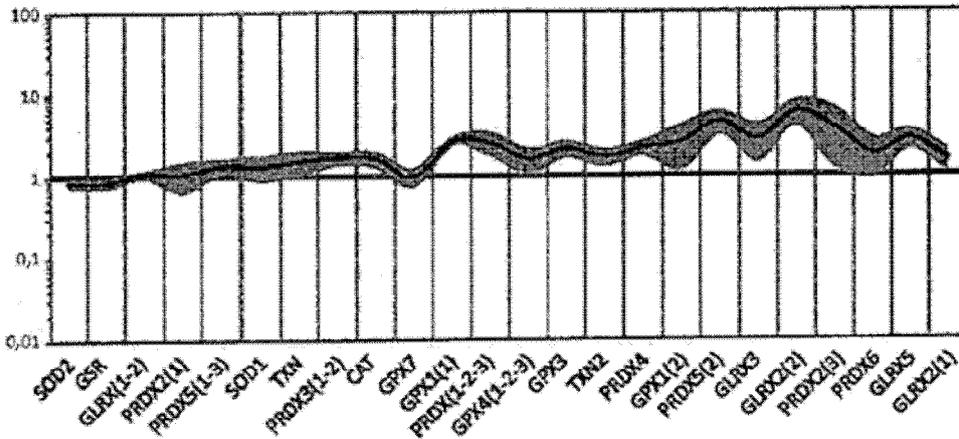


Fig. 3B

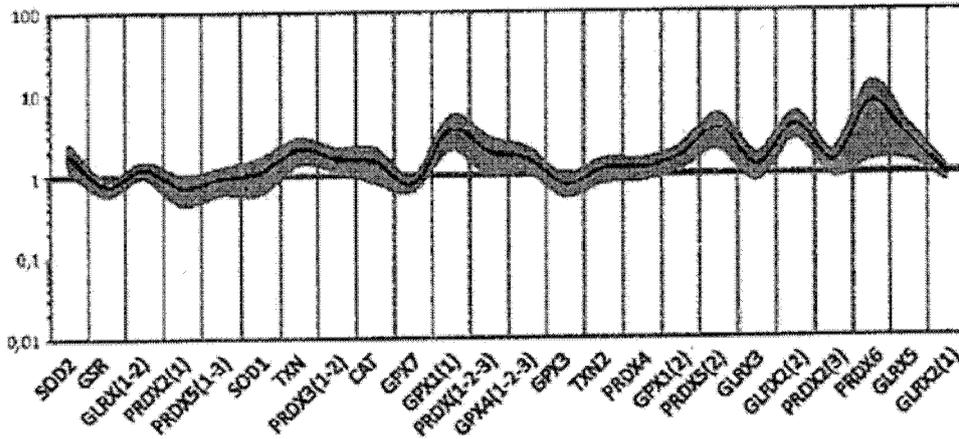


Fig. 3C

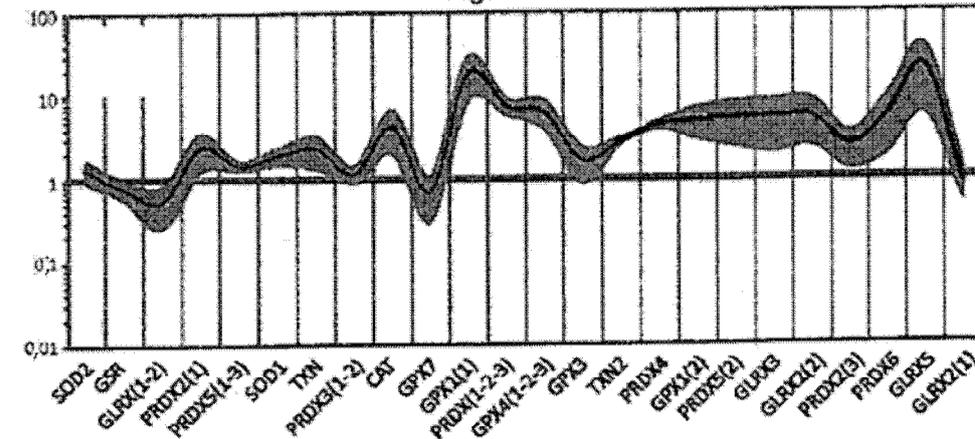


Fig. 3D

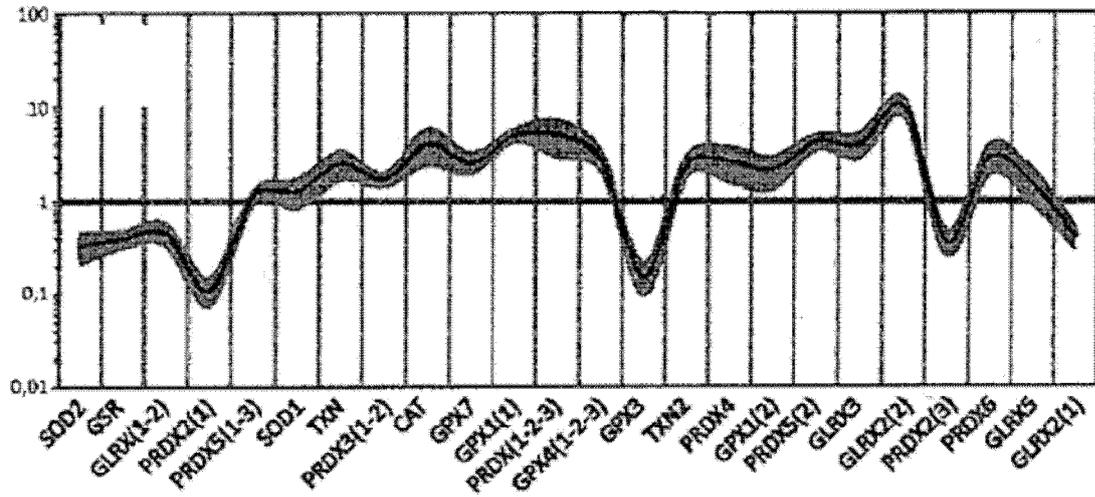


Fig. 3E

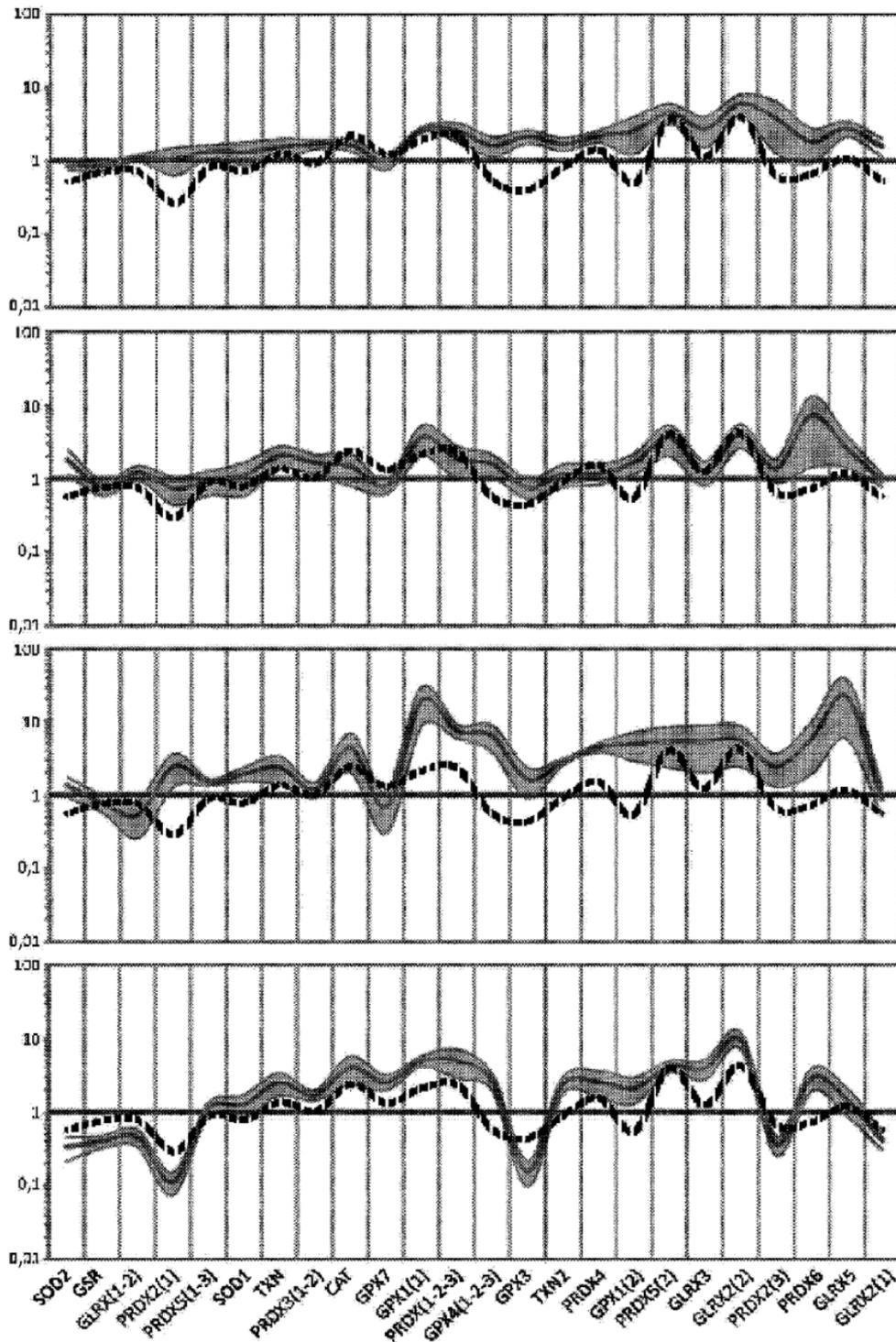


Fig. 4

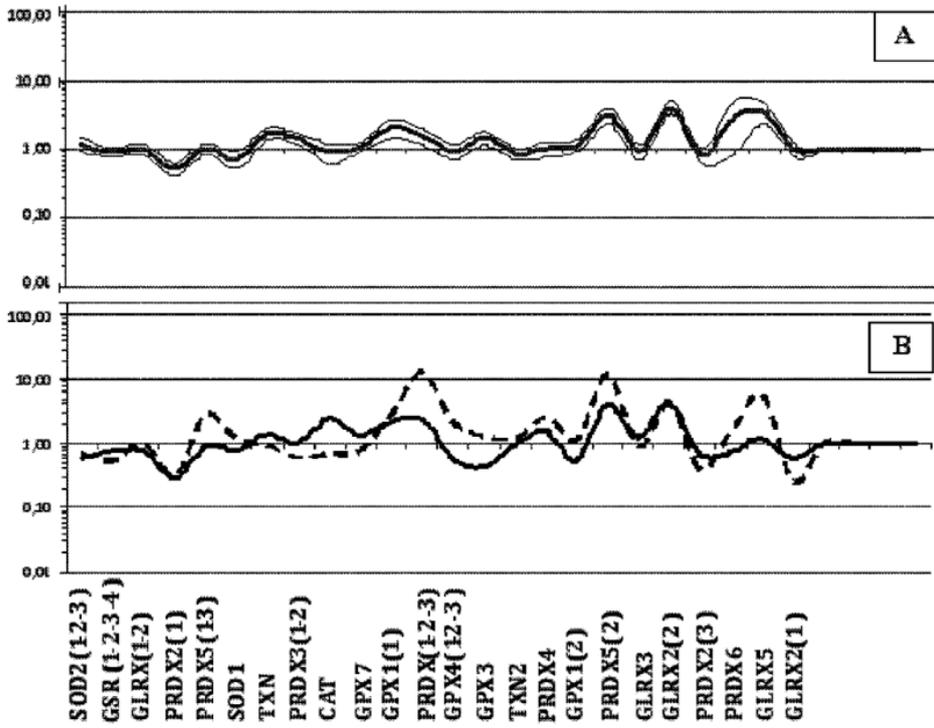


Fig. 5

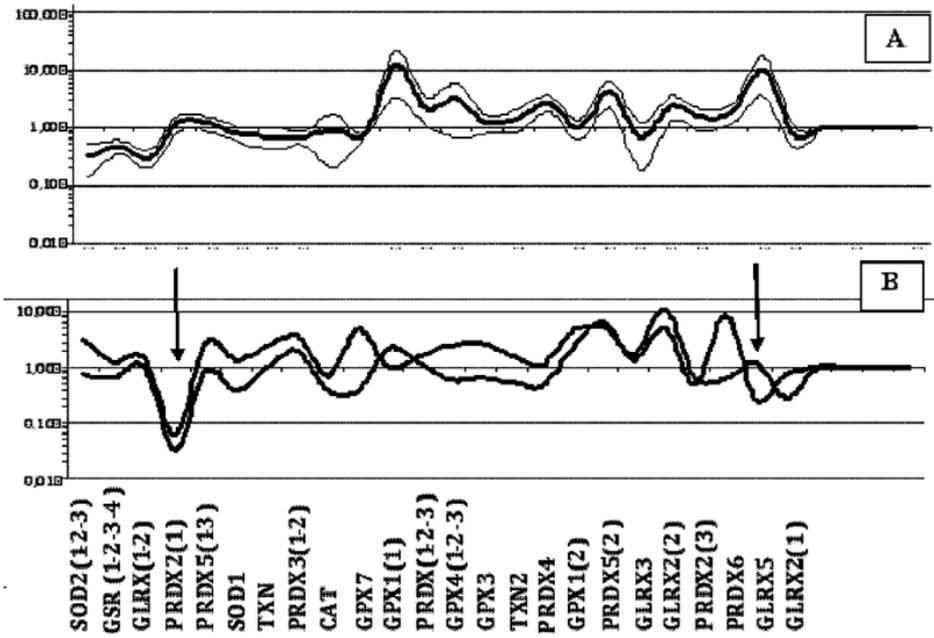


Fig. 6

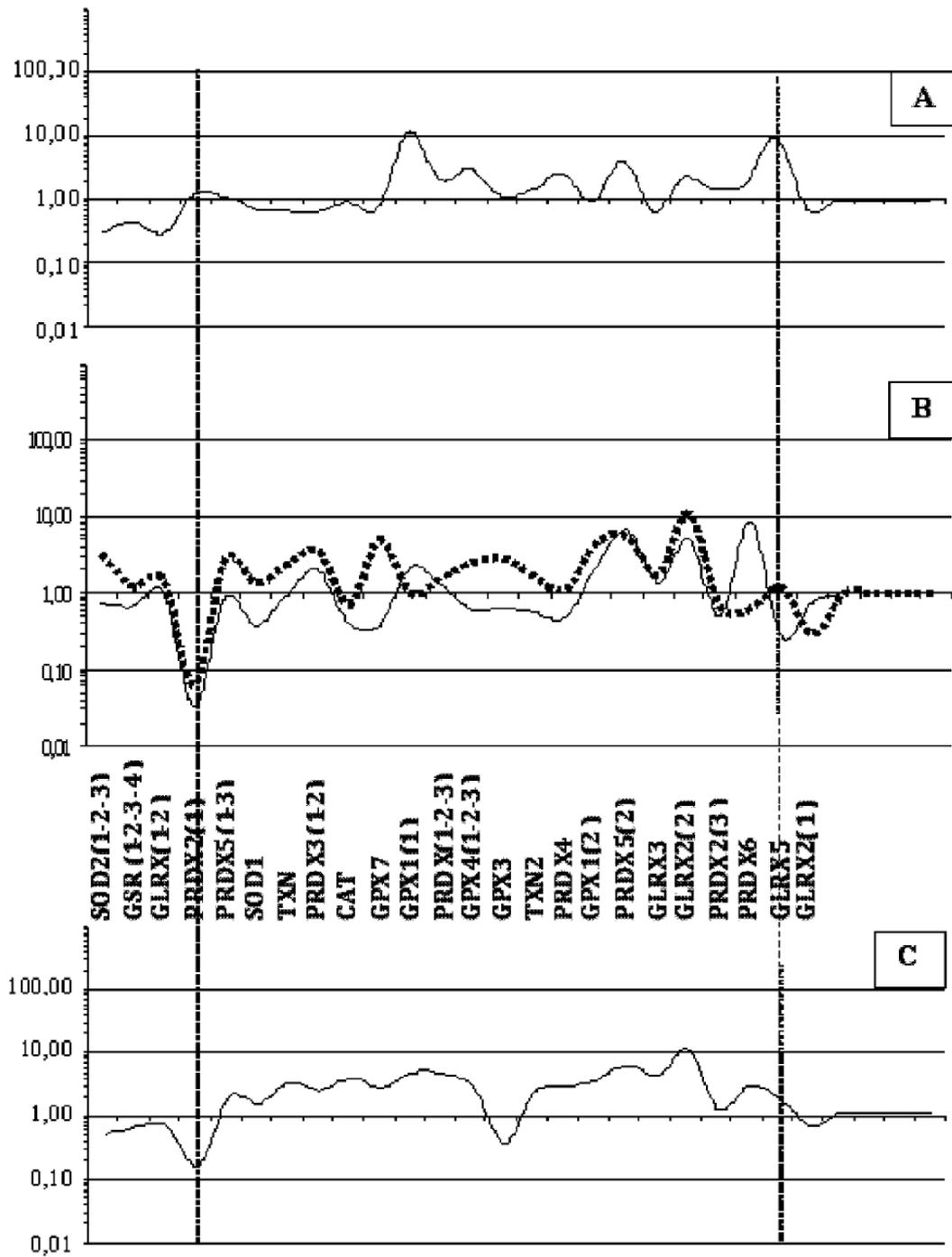


Fig. 7

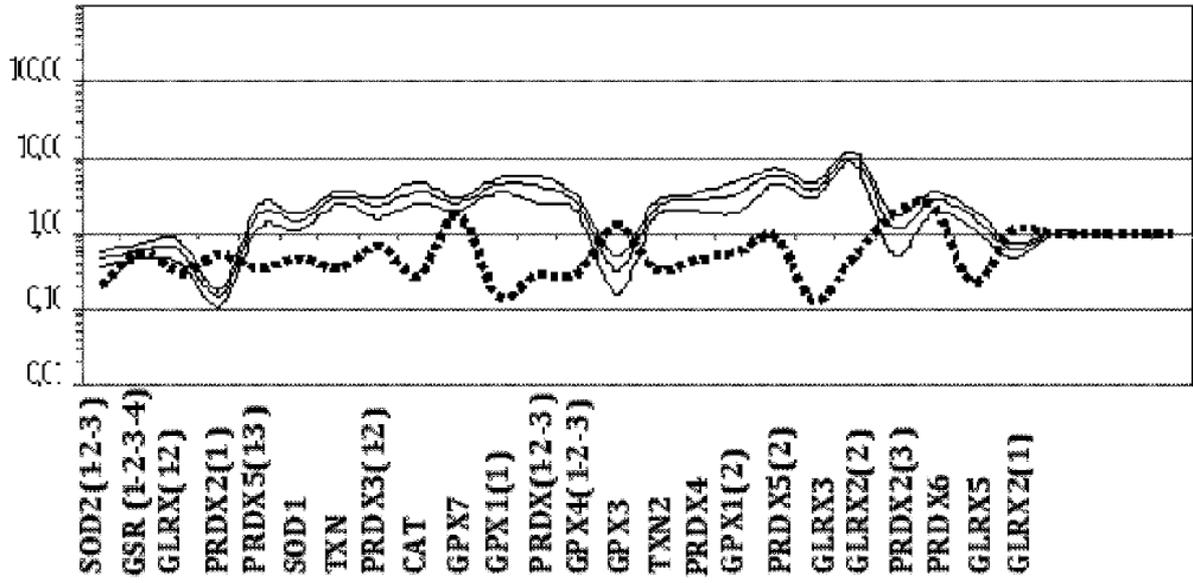


Fig. 8